NORGES HANDELSHØYSKOLE

Bergen, June 17, 2011

The Theory of Long-term Socially Efficient Discount Rates

A report on fundamental issues concerning long-term discounting

Steinar Nilsen

Veileder: Kåre P. Hagen

Master`s Thesis in Economic Analysis

NORGES HANDELSHØYSKOLE

Abstract

This thesis discusses the theory of long-term discount rates for evaluation of long-term public projects. While there are many public projects with a distant time horizon, the threat of global climate change has vastly stimulated economic research and debate on the socially efficient discount rate for such applications. Due to absence of efficient interest rate markets with distant maturities, the work in progress is heavily relying on the sophistication of economic theory. I present the essence of the debate following the Stern review centered around the so-called Ramsey-rule, by which I depart from to investigate the recent development on declining discount rates that incorporates risk and uncertainty of the future. I hope to clarify some of the basis upon which arguments are held in the literature.

Preface

The process leading to this product began with a lot of reading of academic papers, vast amount of information to take in, and too many interesting topics to study. The scope of this thesis was formed continously as I was reading more and more papers, and repetitively discussing with my supervisor. At the time I actually began to write, I thought the topic and my focus were more or less settled, as well as the main conclusions. In retrospect, I am quite sure that the writing itself was the most important learning process. While I was dwelling on many topics, and having a need to be sure of the outcome of the thesis before starting to write, I have now learned that having faith in the process of writing is the most efficent way of organizing findings from the literature. I guess the road is made by walking.

This is a theoretical report. Naturally, I have focused on analyzing theoretical models. I present models from selected literature and try to explain the intuition and the way in which they are different. I think this is necessary for, and really a part of, the discussion itself. Of course, what really makes theoretical analysis interesting is when it matters for practical purposes. In this report, I try to perform a balancing of derivations and discussions, and I hope the reader finds it interesting in the sense that it seems to matter for economic policy.

I would like to express my gratitude for the valuable comments and directional guiding of my supervisor, prof. Kåre P. Hagen, who showed great flexibility of his own time at the final stage of completion. I owe thanks to my brother Vegard Nilsen, who kindly examined some of my calculations and gave helpful comments. I am grateful for all the help of my good friend Kristoffer Thoner, who taught me the basics of LaTeX-programming as willingly as he helped improving my English writing. Finally, I would like to dedicate this thesis to my dear Elisabeth, who patiently encouraged me from the beginning as well as during times of intense work.

While thankful for the help I have received, the responsibility for any remaining errors is mine alone.

Table of Contents

Part 1

Introduction

This part motivates the focus on long-term socially efficient discount rates and clarifies the scope of this thesis. As an introduction to subsequent parts, I present three general rationales for discounting and the common model for use in dynamic analysis is derived. A simple decision criterion (the NPV-rule) is illustrated and linked to the Ramsey model at the end of this part.

1.1 Motivation

The theory of discounting the future in public cost-benefit analysis (CBA) has attracted economists' attention for many decades, gaining its renaissance alongside the popularization of global warming issues. Great effort has been put into deriving socially efficient discount rates from sophisticated theory, as a contrast to merely studying interest rates in the market. While there may be a handful of reasons for failures of this market, including taxation of capital and liquidity constraints, the theory investigated here is mostly motivated by the fact that interest rate markets for distant horizons barely exist. The most important reason for this is future generations' lack of representation in today's market. Even though some traded bonds with maturities up to 50 and a few up to 100 years are not uncommon[1], these markets have hardly been regarded as efficient[2].

Public projects having an economic feature in horizons for which there doesn‘t exist an efficient interest rate market are many, including long-lived infrastructure, pension and health reforms, medical research, education and research in general, biodiversity, nuclear power plants and mitigation of climate change. The applications of long-term socially efficient discount rates are many and so are the different practices of governments. Table 1 is an illustration of some governments‘ recommendations:

Nation	$t(years) \leq 30$	$t \geq 30$	Comments
Australia	-	-	no rules of thumb
France	4%	2%	-
Norway	4%	4%	up to 6% risky projects
USA	7%	-	est. pre-tax return private capital

Nation	$t \leq 30$	$t \in [31, 75]$	$t \in [76, 125]$	$t \in [126, 200]$	$t \in [201, 300]$	$t \geq 301$
GB	3.5%	3%	2.5%	2%	1.5%	1%

Table 1. Governments‘ recommended discount rates[3]

Both France and Great Britain recommend a lower future discount rate to reflect uncertainty about the future. Great Britain uses the steady state Ramsey equation[4] $\delta + \theta g = 0.015 + 1 \cdot 0.02$ to justify 3.5% in the near future. Norway has estimated the risk free real rate to be around 2 %, and rounds up to 4 % for public projects with normal risk. Australia does not recommend specific numbers, but prescribes the far most flexible discount rate adapted to the specific project‘s properties and cir-

[1]There are even a few examples of 1,000-year bonds, (the Canadian Pacific Corporation has issued such bonds in the past) and bonds without maturity (Investopedia). The British government has issued bonds which make coupon payments indefinitely, called Consolidated Annuities (Encyclopedia Britannica).

[2]Conventional economics will regard 30 years as a maximum time horizon for liquid capital asset markets.

[3](Commonwealth of Australia, 2006), (Rapport Lebégue, 2005), (Finansdepartementet, 2005), (Office of Management and Budget, 1992), (HM Treasury, 2003).

[4]See section 1.4 for an explanation of this.

cumstances. Both Australia and the United States argue that the so-called shadow price of capital method should be used[5] on theoretical justification. The United States recommends a pre-tax opportunity cost of capital of 7% for operational use in general.

The timeliness of discounting issues comes from the increasing awareness of anthropogenic climate change, and the arising attention for economic research on climate change. No doubt have these overwhelming threats to our planet fostered a whole new wave of publications on issues directly or indirectly related to discounting. Still, it is hard to see that researchers are moving towards some kind of convergence on prescriptions for which discount rate to use, or even the shape of the term structure. The meta-analysis performed by Weitzman (2001) is one indicator[6]: He received 2,160 numerical answers from PhD-level economists from 48 countries to the "simple" question of which discount rate to use in cost-benefit analysis. The answers were ranging from -3 to 27 % with three-digit number of responses in the range of 1 to 6 %.

There are most certainly large caveats to overcome on the whole range of topics related to the economics of climate change, not at least on the dynamics of the global warming control system itself. Still, since the publication of the Stern Review (Stern and Treasury, 2007), it seems that most of the debate has centered around prescriptive analysis of the discount rate. If the widespread disagreements on this weren‘t clear prior to the publication, the aftermath of the Stern Review effectively illustrated how inconclusive the literature on long-term discounting as a whole stands today:

"They [Questions of discounting] lie at the heart of the Review‘s [The Stern review] radical view of the grave damages from climate change and the need for immediate steps to reduce greenhouse gas emissions sharply"(Nordhaus, 2007)

[5]The shadow price of capital is defined as the present value of consumption produced by one unit of capital (Arrow et al., 1996). It is widely held that the shadow price is difficult to estimate, and it varies among projects in general.

[6]A problem of this simplistic approach, in which respondents were likely to exhibit personal ethical values as well, is that they were not allowed to specify assumptions or state the framework under operation.

"The strong, immediate action on climate change advocated by the authors [of the Stern review] is an implication of their views on intergenerational equity; it isn't driven so much by the new climatic facts the authors have stressed"(Dasgupta, 2007)
"In fact, it is not an exaggeration to say that the biggest uncertainty of all in the economics of climate change is the uncertainty about which interest rate to use for discounting."(Weitzman, 2007)
"In reaction to those conclusions [of the Stern Review], the most widely debated economic issue was the choice of discount rate."(Gollier and Weitzman, 2010)

The importance of the discount rate in determining profitability of public projects is easily illustrated by a numerical example. By the power of compounding, the present value of $1 to be received 100 years from now with a 1 % discount rate is $0.368, while it is $0.0025 with a 6 % discount rate, the former being 147 times higher than the latter. Adding to disagreements on which discount rate to use, the shape of the term structure is also subject to debate, as we will see.

Thus, in a field of economics where applications are many, as well as disagreements within both academia and practice; the discount rate used is also often the most crucial factor in determining the economic value of public projects.

1.2 The Scope and Limitations of This Thesis

Suppose that a decision-maker evaluating a long-term public project finds himself wondering how to compare net benefits at different points of time. He asks: "What is the theoretical foundation for long-term discounting? What are the most important questions concerning ethics in this theory? Why should discount rates vary with the time horizon, and is that an issue of time inconsistency?"

This thesis is an attempt to answer these questions by collecting and discussing insights from selected literature.

Part 2 will very briefly look into the academic debate within the conventional framework of exponential discounting following the Stern Review. In addition to the Stern Review‘s circumspect discussion of discounting issues, there already exists an excellent, and complete at the time, review on discounting and intergenerational equity by Arrow et al. (1996)[7]. I will not replicate this work, or even try to reach up to the level of their extensive treatment. But in acknowledging the need to outline central issues of disagreement before investigating the recent progress of non-exponential discounting, I think the academic debate deserves to be dedicated a small part of this thesis. The presentation of the debate will be structured around the steady state Ramsey equation, which I derive in section 1.4. Part 2 is more of a sentimental analysis of how different philosophical views work through the Ramsey model. The advantage of an economic framework applied to philosophical questions, is in its ability to structure implications of ethical values for policy analysis and vice versa.

Part 3 is the main part of this thesis. I thoroughly follow the reader through the different rationales for using declining discount rates in long-term public CBA. Declining discount rates are not to be confused with hyperbolic *preferences* (sometimes also referred to as hyperbolic discounting). Hyperbolic preferences is a well-studied field within the behavioral theory of individual intertemporal decision-making, flourishing with the work of Strotz (1955) who pointed to the problems of time inconsistency. The theory of hyperbolic preferences deals with discounting of utility. Declining discount rates describe an optimal term structure for discounting public projects. The theory is based on external factors to the agent and deals with discounting of goods. This important distinction will be made clear in part 3, as well as implications for time inconsistency. As we will see, the literature is not yet fully developed on this particular issue.

In part 4, I summarize my reflections and conclude the main insights I draw from the investigated literature. I mention a few of the major caveats in long-term project evaluation that relates to the discount rate but falls short of the scope of this thesis.

Throughout the thesis I will not look at effects of the intrinsic risks of projects

[7]As well as the collection of essays on discounting and intergenerational equity edited by Portney and Weyant (1999).

on discount rates. Rather, my focus is on effects of introducing risk and uncertainty of the macroeconomic conditions in the future. These future macroeconomic conditions are captured either in the consumption growth or the interest rate, and these are dual to each other in an equilibrium framework.

It is commonly held that project risk should be adjusted for by using certainty-equivalent costs and benefits, when these accrue asymmetrically over time[8]. The discount rate is then left to reflect opportunity costs of time, not riskiness of projects. This would be a risk-free interest rate with regard to the project, but risks of future consumption growth, or risks of the future risk-free interest rate will affect this discount rate in general. The macroeconomic risk comes from not knowing what our real economic situation, and hence our marginal utility will be in the future. This kind of risk will be introduced and analyzed in part 3, and is at the heart of the theory of declining discount rates.

The distinction between risk and uncertainty is explained by the degree of knowledge about the stochastic process. If we know the parameters of such a process for sure, we are left with pure risks of nature. If we don‘t know the parameters for sure, we are dealing with parametric uncertainty. Throwing a fair dice is a good example of a pure risk environment, while real developments of macroeconomic conditions in a strict sense always is bound to contain some varying degree of parametric uncertainty, like for instance throwing a dice with unknown probabilities.

As always, real economic values and not nominal values matter for public CBA. Therefore, if not stated, I mean real values.

I must point to the review performed by Pearce et al. (2003), since this work aspires to review the recent progress in social discounting and the theory of declining discount rates. My thesis differs in at least three ways: It is not a review per se, but rather a kind of textbook-approach. Therefore, I also perform an explicit investigation of the rationale behind long-term discounting. Lastly, I focus more extensively on the timing issues that arises from a declining term structure and aspire

[8]Weitzman (2007) looks at a combined framework treating project risks of climate change mitigation directly in the discount rate and obtains a declining pattern of discount rates. See section 4.3.

to conclude on these issues[9]. An additional review by Gollier et al. (2008) investigates declining discount rates and focuses, to a large extent, on transforming the theory into operational prescriptions. I will refrain from leaving the environment of abstract economic theory in part 3, and devote more attention to the timing issues, in addition to investigate the whole range of rationales behind declining discount rates. This thesis is purely theoretical and will not try to implicate numerical answers to the question of a socially efficient discount rate. I have accordingly made many simplifying assumptions, and overlooked the many caveats appearing when operationalizing public CBA for real applications. In this way, I also think my work differs from the reviews mentioned.

The rest of part 1 will shortly discuss some basic features of discounting, and the last section entails a "back of the envelope"-derivation of the Ramsey model. The end of part 1 contains a few remarks on the Ramsey model and public CBA.

Further assumptions are stated when needed.

1.3 About the Discount Rate

Here I will shortly answer the question: Why do economic agents discount the future at all? In the theory, there are three main reasons for discounting[10].

Impatience

On the individual level there is a well-documented tendency to put lower weight on the future simply because we prefer utility today versus tomorrow, ceteris paribus. This is called the pure rate of time preference or the impatience rate, and is generally applied to discounting of utility rather than consumption goods. The tendency to be impatient is observed among animals as well as human beings, having features of animal spirits. This behavior might reveal instincts, biologically explained by survival propensities[11] or could be explained by other external factors driving

[9]In those ways I also believe that my work differs from OXERA (2002).

[10]As stated in the previous section, I don't look at individual project risks.

[11]The observed impatience have been studied among economists, as well as biologists, psychologists and philosophers.

uncertainty about the future, or even uncertainty about own preferences in the future[12]. Hence, on both the individual and collective level we might interpret the impatience rate as a tragedy rate or extinction rate. Such an interpretation implies that the exponential discount rate reflects the probability of a tragedy or extinction per instant of time. Whether this impatience rate applies to social discounting or not is discussed briefly in part 2.

Opportunity Cost of Capital

As long as there is productive technology, there is a cost of not putting the capital into its best alternative use. To forgo the benchmark investment is an opportunity cost and hence any project under evaluation is required to yield at least as high return on capital as the marginal return on capital. This is true for the individual as well as the society in general [13]. This is a discounting rate that applies to discounting of consumption goods rather than utility.

Technological Progress and Decreasing Marginal Utility

If there is technological progress, or on the individual level some other reason that income grows over time, transferring consumption from today and into the future should be discounted because of decreasing marginal utility of consumption. The simple intuition is that in the future we are on a higher consumption level, and by decreasing marginal utility[14], the marginal utility of one unit of consumption is lower compared to that of one unit of consumption invested today. Hence, there is a utility loss of transferring consumption into the richer future.

[12]One extensive review of theory and empirics on the impatience rates is the meta-analysis of Frederick et al. (2002).

[13]The historical marginal return on capital is elicited from the financial markets, or estimated directly from the return on real capital.

[14]In fact, this rationale for discounting falls apart if we assume linear utility.

1.4 The Ramsey Model

Much of the debate on long-term discounting, especially post-Stern, is centered around a formal model describing dynamic optimal planning of consumption. The framework has its origin in the seminal paper *A Mathematical Theory on Saving* by Ramsey (1928). The model has evolved through new mathematical methods and many bright economic scholars[15], but I will nevertheless call it the Ramsey Model. Before we look into the problems of this kind of modelling and the debate that follows, I believe it is useful to explicitly derive its assumptions and state the results, rather than just take the Ramsey equation at face value. The presentation of this model is based on the textbook on economic growth by Barro and Sala-i Martin (2004). For an almost complete derivation of the model, I refer to appendix A, which is recommended to the reader who is unfamiliar with the model.

Simplifying Assumptions

1. The world is assumed to be completely predictable and free of risk!
2. There is one optimizing representative agent, and problems of aggregating agents‘ utility across space and time are ignored.
3. The representative agent has power utility, which gives constant relative risk aversion(θ) and satisfies the general criteria: $u'(c) > 0$, $u''(c) < 0$ and Inada conditions: $\lim_{c\to 0} u'(c) = \infty$, $\lim_{c\to\infty} u'(c) = 0$
4. The representative agent works at equilibrium wage, $W(t)$, and has initial endowment $A(0)$. Total capital income is $r(t)A(t)$
5. There is one representative consumption good: $C(t)$
6. Population is equal to the labor force and is normalized to 1 at time 0 and has relative growth n such that: $L(t) = e^{nt}$
7. Technology is normalized to 1 at date 0 and has relative growth g such that: $G(t) = e^{gt}$

[15]Especially economists Robert M. Solow, Tjalling Koopmans and David Cass.

8. A positive impatience rate is assumed, $\delta > 0$ if not for other reasons to get bounded lifetime utility.

9. $Y = F(K, \hat{L})$, where $\hat{L} = L(t)G(t)$ denotes effective labor force. K denotes total capital stock, and we will introduce the variables $k = \frac{K}{L}$ meaning capital per capita, $\hat{k} = \frac{K}{\hat{L}}$ which denotes capital per effective capita, and $\dot{\hat{k}}$, which denotes the time derivative of $\hat{k}$. We use equivalent notation for $C(t)$, c, $\hat{c}$ and $\dot{\hat{c}}$ which denotes consumption per capita and so on. In practice, this means that over time, these variables are "depreciated" with technology and population growth. Cobb-Douglas production technology is assumed for its convenient properties. That is, we want the production function to satisfy constant returns to scale in both arguments and Inada conditions.

10. All markets clear at all points of time, one unit of labor W(t) is inelastically supplied per capita. Depreciation is assumed to be zero, so the rental rate of capital is $r(t)$, and frictionless markets provide the same rate for borrowing and lending.

The Consumers' Problem

Consumption c is generalized to include all goods, abstract, physical, health and environmental goods and so on. The instantaneous utility function

$$u(c) = \frac{c^{1-\theta} - 1}{1-\theta}, \quad \theta > 1 \tag{1.1}$$

$$\theta = 1 \implies u(c) = ln(c) \tag{1.2}$$

$$\lim_{\theta \to 1} u(c) \overset{\text{l'Hopital}}{=} \lim_{\theta \to 1} \left(\frac{-\theta ln(c)}{-1} \right) \tag{1.3}$$

$$= ln(c) \tag{1.4}$$

The objective is to maximize lifetime utility:

$$U = \int_0^\infty u[c(t)] e^{-(\delta - n)} dt, \tag{1.5}$$

subject to the constraint in dynamics of income:

$$\dot{a} = w(t) - c(t) - a(t)\big(n - r(t)\big) \tag{1.6}$$

,and the transversality condition, where $v(t)$ denotes the present value shadow price of income:

$$\lim_{t\to\infty}\left(v(t)\cdot a(t)\right)=0 \tag{1.7}$$

The transversality condition‘s intuition is best explained if one imagines a thought ”end of time”. At that time, it is not optimal having consumed all resources, and the credit markets don‘t allow any negative values on assets. The formulation of the problem will ensure this. The Hamiltonian[16]:

$$H=u[c(t)]e^{-(\delta-n)t}+v(t)\big(w(t)-c(t)-a(t)\big(n-r(t)\big)\big) \tag{1.8}$$

In appendix A.1, I solve for these first-order condition and substitute the general utility function for power utility:

$$\left.\begin{aligned}\frac{\partial H}{\partial c}&=0\\ -\frac{\partial H}{\partial a}&=\dot{v}\end{aligned}\right\}\qquad \text{FOC} \tag{1.9}$$

Euler Equation

According to the model we are now to end up with the famous Euler equation, governing the optimal path of per capita consumption at *any* time t.

$$\text{Euler equation: }\frac{\dot{c}}{c}=\frac{1}{\theta}(r-\delta) \tag{1.10}$$

Here it is important to note that this equation describes the optimal path of consumption per capita regardless of steady state.

The Firms‘ Problem

Capital K is generalized to include all forms of capital except human capital which is treated directly as labor-augmenting technology. Assuming that firms optimize profits each instant of time,

$$\pi=F[K,L]-rK-wL \tag{1.11}$$

[16]A good source to learn the intuition of this mathematical tool is Dixit (1990).

Optimization gives the first-order conditions:

$$\left.\begin{aligned}\frac{\partial \pi}{\partial K} &= 0\\ f'(\hat{k}) &= r\end{aligned}\right\}\qquad \text{FOC capital} \tag{1.12}$$

$$\left.\begin{aligned}\frac{\partial \pi}{\partial L} &= 0\\ e^{gt}\left(f(\hat{k}) - f'(\hat{k})\hat{k}\right) &= w\end{aligned}\right\}\qquad \text{FOC labor} \tag{1.13}$$

Equilibrium Optimal Dynamics

The dynamics of capital per effective labor:

$$\dot{\hat{k}} = f(\hat{k}) - \hat{c} - \hat{k}(n+g) \tag{1.14}$$

The relative consumption growth per effective labor (Euler equation)

$$\frac{\dot{\hat{c}}}{\hat{c}} = \frac{1}{\theta}\left(f'(\hat{k}) - \delta - \theta g\right) \tag{1.15}$$

This is the Euler equation in per effective capita form. Rearranging, we arrive at the Ramsey equation in per effective capita form:

$$f'(\hat{k}) = \delta + \theta\left(g + \frac{\dot{\hat{c}}}{\hat{c}}\right) \tag{1.16}$$

The Steady State Ramsey Equation

By imposing steady state, $\dot{\hat{c}} = \dot{\hat{k}} = 0$, we arrive at the very important steady state Ramsey equation (s.s. Ramsey equation):

$$f'(\hat{k}) = \delta + \theta g \tag{1.17}$$

Note that this equation in general doesn‘t hold until reaching steady state. This equation is often used in debates about long-term discounting, and I think it is important to have in mind that it relies on the economy having arrived at steady state.

We know that in our model, firms simply adjust investments such that ($f'(\hat{k}) = r$) the marginal product is equal to the interest rate at every instant of time. Along the optimal consumption path, this interest rate will vary given concavity of $f(\hat{k})$, but we could say something about the interest rate in steady state of the economy in this model. As shown, the only source of consumption growth per capita in steady state is the technological process.

The rule for optimal consumption growth has now been translated into a socially efficient discount rate, with infinite time horizon *given* that the economy is in steady state. The right hand side of eq. (1.17) is often called the social rate of time preference (SRTP), because the right hand side is derived on the basis of society‘s dynamic preferences. That is:

$$SRTP = \delta + \theta g \tag{1.18}$$

1.5 The Ramsey Model and CBA

Now that we have described the underlying assumptions and forces governing optimal consumption growth, we are able to say something about the implications for normative analysis in the field of public CBA. Before we proceed we need to specify the decision criterion.

1.5.1 The NPV-Criterion

I am not considering project risks but the NPV-criterion is generalized using certainty equivalent costs and benefits. Define

$$C_t = \text{certainty equivalent costs at time } t \tag{1.19}$$

$$B_t = \text{certainty equivalent benefits at time } t \tag{1.20}$$

$$I_t = \text{investment required at time } t \text{ to undertake a project} \tag{1.21}$$

$$k_t = \text{the socially efficient discount rate at time } t \tag{1.22}$$

Then, we can define

$$\mathrm{NPV}_t = \int_t^{\infty} (B_v - C_v) e^{-k_v v} dv - I_t \geq 0 \tag{1.23}$$

The criterion requires that the net present value of the project is positive. If the constraint on capital availability is not binding, all projects with positive NPV should be undertaken.

Note that costs and benefits are defined in equivalent measures. Usually, benefits in terms of social utility is converted into consumption equivalents, as with costs and investments.

Even though the upper limit is ∞ in the above integral, it is assumed that the integral is bounded[17].

Define π_t = profits at time t on the marginal private project and r = return on the marginal project. Then, by accounting relationship,

$$\int_t^\infty \pi_v e^{-rv} dv - I_t = 0 \tag{1.24}$$

This is highly trivial and simply illustrates that the marginal private project has return equal to the marginal rate of return in the economy.

1.5.2 Conclusions

In practice, the observed marginal return on capital is different from the estimated social rate of time preference, as defined by eq. (1.18).

In order to say something about public CBA when we are not in steady state, define:

From the general Ramsey eq.:

$$f'(\hat{k}) = \delta + \theta\left(g + \frac{\dot{\hat{c}}}{\hat{c}}\right) = \rho^g \tag{1.25}$$

From the s.s. Ramsey eq.:

$$f'(\hat{k}) = \delta + \theta \cdot g = \rho^{ss} \tag{1.26}$$

From frictionless markets:

$$r = f'(\hat{k}) \tag{1.27}$$

Proposition 1.1

Trivially, if the real world economy were a Ramsey economy in steady state, we could simply apply the observed market interest rate for CBA. This means that $\rho^{ss} = r$

[17]We assume that the discount rate is positive and net benefits are increasing sufficiently slowly.

Suppose a new public project is under evaluation. If the project survives the NPV-rule with r, the marginal loss of utility from reducing consumption today is at least offset by the marginal increase in present value of utility in the future, taking into account that marginal utility is lower in the future.

Proposition 1.2

If we were to observe the right hand side of the s.s. Ramsey equation as well as the frictionless market interest rate, and find that $r \neq \rho^{ss}$, then we are not in steady state, and cannot apply ρ^{ss} as a discount rate in general.

In deriving the Ramsey model, we have seen that $r = \rho^{ss}$ is a necessary, though not sufficient condition for steady state. If we are not in steady state, we need the general Ramsey equation to know if we are moving along the optimal path. That is, if eq. (1.25) holds.

Proposition 1.3

If we were to observe all the parameters on the left and right hand side of the general Ramsey equation, and find inequality, we should adjust the path of consumption in order to get equality, and hence, achieve the optimal path.

The optimal change of relative consumption growth is given by the Euler equation. This will be achieved by investing or disinvesting in the economy. Adjusting investments will in response change $\hat{k}$, and in turn $f'(\hat{k})$.

Proposition 1.4

If $r > \rho_g$ the Euler equation states that we should invest by reducing consumption today in order to get equality and again move along the optimal path. The socially efficient discount rate will then be r.

We should always invest in the most profitable project first.

Proposition 1.5

If $r < \rho_g$ the Euler equation states that we should disinvest by increasing consumption today in order to get equality and again move along the optimal path. The *socially efficient discount rate will then be* ρ_g.

If we use the market interest rate as a discount rate, we will not be sure of compensating the decreasing marginal utility effect and the impatience rate effect.

Part 2

Background: The Academic Debate

This part discusses some selected topics of the debate raised by the Stern Review. It will not discuss the Stern Review itself, but rather look at the publications in response and occasionally draw on earlier publications. The topics of the Stern Review debate are not particularly new to academia, but the conclusions relying on these topics and assumptions are very much controversial. The discount rate part of the climate change debate is important for public projects in general as well as for green projects. But if the project under evaluation seeks to limit the probability or consequences of a possible catastrophe, the rate of return hardly matter the most[1]. The main point of this part is the difficulty of applying the s.s. Ramsey equation in a way meeting consensus in academia, even when economic growth is taken to be deterministic. There are, as I see it, three important reasons for this:

1. There is no general consensus that the s.s. Ramsey equation is a good formula for discount rates in real world applications, especially when opportunity cost of capital is higher than the social rate of time preference.

2. There are large uncertainties on estimates of the parameters (δ) pure rate of time preference and (θ) elasticity of intertemporal substitution.

[1]See section 4.3 for a short discussion of this.

3. There are huge disagreements regarding ethical values, and how to assess such values for normative applications.

The first point here was partly made in part 1 of the thesis. It will be somewhat elaborated in section 2.1.2. The second point is demonstrated in Table 2 here, as well as discussed shortly in section 2.1.1: The third point is the subject of section 2.2.

Authors	E[θ]	Range of E[θ]
Stern(1977)	2	[0,10]
Hall (1988)	10	-
Epstein & Zin (1991)	-	[1.25,5]
Pearce & Ulph (1995)	-	[0.7,1.5]

Table 2. Estimates of θ [2]

Frederick et al. (2002) present a long list of different estimates of δ, varying from negative values to more than 100 %. The estimates are as many as the methods used in the empirical literature, and suggest that individuals‘ δ varies with different situations.

2.1 Repercussions of the Stern Review

The importance of the Stern Review in the academic debate on climate change extends to a whole range of topics. With its 2,904 citations (according to Google Scholar), it has certainly served as a benchmark for the climate change debate. I will however limit myself to concentrate on the debate induced by controversial discounting. Alongside introducing insurance arguments for using a low discount rate, Stern and Treasury (2007) uses the framework of the s.s. Ramsey Model to rationalize a discount rate of 1.4%. He argues that there are good reasons to believe that $\delta = 0.001$, $\theta = 1$, and $g = 0.013$ resulting into

[2]The information in the table is obtained from Gollier (unpublished manuscript), "Pricing the future: The economics of discounting and sustainable development", an unpublished work by prof. Christian Gollier at Toulouse School of Economics, France.

$$\rho = \delta + \theta g = 0.001 + 1 \cdot 0.013 = 1.4\% \tag{2.1}$$

Note that δ was set solely to reflect the society‘s extinction rate[3]. I will not go into details on all the different suggestions of these parameter values, but as a comparison simply refer to Weitzman (2007), whose best informal guess was the "trio of two‘s":

$$\rho = 0.02 + 2 \cdot 0.02 = 6\% \tag{2.2}$$

The balancing of a constant discount rate of 1.4 % which gives a *proposed* unacceptably high relative weight on the distant future, and a constant discount rate of 6 % which gives a *proposed* unacceptably low relative weight on the distant future, is an interesting motivation for exploring alternatives to the s.s. Ramsey equation in exploring socially efficient discount rates. Still, as I am trying to stress; arguments for unacceptable relative weights on generations based on mere sentiment, that one often finds in the literature, hardly suffice for bringing the debate forward. It seems unlikely that we can expect consensus on a socially efficient discount rate based on ethical grounds. There may be some extreme parameter values for both δ and θ that can be ruled out merely based on unacceptable ethical implications, but the question of how the discount rate should be set remains unanswered.

Less controversial is the Stern Review‘s choice of the δ parameter, describing impatience as motivated by the extinction rate. Many prominent scholars have advocated a zero discounting rate of utility when considering intergenerational welfare from a normative point of view. The most famous quote in this regard most likely belong to Ramsey himself when commenting on the practice of discounting utility: "a practice which is ethically indefensible and arises from the weakness of the imagination"(Ramsey, 1928). This conclusion is followed up by most others working on this topic. Yet, the literature is most certainly not unanimous. In the aftermath of the Stern Review, Nordhaus (2007) has probably been the foremost critic of this ethical view, suggesting a handful other plausible ethical views. The long tradition of "agent-relativism", also suggests that we indeed care more for our

[3]Such a constant rate implies a Poisson process (memoryless) with probability δdt for extinction per instant time dt.

nearest relations and defends it as a legitimate moral point of view. This part of the debate is commented further in section 2.2[4].

2.1.1 The Instantaneous Utility Function

In deriving the Ramsey model, we simply assumed iso-elastic utility. This is not necessarily a good description of people‘s preferences. Using general utility, the s.s. Ramsey equation becomes

$$r = \delta - \left(c \frac{u''(c)}{u'} \right) \frac{\dot{c}}{c} \tag{2.3}$$

The power utility function was not randomly picked; it is analytically easy to work with. An important disadvantage of this assumption, however, is that in steady state, the level of consumption doesn‘t matter at all for willingness to transfer consumption between different points of time. Using power utility, the last term of eq. (2.3) becomes θg in steady state. In equation (2.3) on the other hand, the level of consumption in general does matter for the interest rate. It has been proposed that as we get richer, it is plausible that we are willing to save a larger share of our wealth for the future, implying decreasing elasticity of marginal utility. If this were constant, and given g, a poor country with a below subsistence consumption level, and a rich country would be willing to sacrifice equally as much in relative terms to improve the well-being of future generations. If this is an unacceptable feature, one could for instance modify the power utility function,

$$u(c) = \frac{(c-x)^{1-\theta}-1}{1-\theta}, \quad \theta > 1 \tag{2.4}$$

where x denotes the subsistence level in the economy. Then we get decreasing elasticity of marginal utility:

$$\left(c \frac{u''(c)}{u'(c)} \right) = \frac{\theta c}{c - x}, c > x \tag{2.5}$$

[4]In part 3, section 3.4 we will look at the implications of combining the two points of view in the models of Chichilnisky (1996) and Li and Löfgren (2000).

It is easy to see that it is decreasing with consumption, and approaches θ in the limit when consumption goes to infinity:

$$\frac{\partial}{\partial c}\frac{\theta c}{c-x} = \frac{\theta}{c-x}\left(1-\frac{c}{c-x}\right) < 0, c > x \tag{2.6}$$

$$\lim_{c\to\infty}\left(\frac{\theta c}{c-x}\right) = \theta \tag{2.7}$$

Usually, the instantaneous utility function describes aversion to inequality in three dimensions: aversion to inequality across states of nature[5], intertemporal inequality and spatial inequality[6]. Decreasing marginal utility is often regarded as reasonable in all three applications, but the magnitude of the relative curvature is not necessarily the same for each application. A person might be willing to bear relatively high amounts of risk, and at the same time accepting relatively low inequality in consumption over time. In the same line, society‘s valuation of spatial equality might differ from intertemporal equality. As Weitzman (2007) and Nordhaus (2007) criticize the Stern Review‘s choice of a low parameter of inequality aversion ($\theta = 1$), Gollier (2006) points out that this is deeply unrealistic with observed aversion to risk in financial markets. Overall, there is little dispute about reasonable values for θ when power utility is assumed, at least compared to the absurd dispersion of the many δ-estimates. Still, it is not clear that θs automatically applies identically for every dimension, a point stressed by Schelling (1995).

2.1.2 Normative versus Positive Theory

It is commonly held that economists have two important tasks in society: explaining observed economic behavior (positive or descriptive theory) and deriving rules for optimal decisions (normative or prescriptive theory). It is a fact that these two disciplines work together, still; there is a long way between how the world looks like and how it should look like. The descriptive and prescriptive approaches ask different questions, and so in general the answers cannot be the same. Observed

[5]The Arrow-Pratt measure of relative risk aversion is identical to elasticity of marginal utility.

[6]Spatial inequality, that is inequality across people, is not considered in our model as we have assumed a single representative consumer. In theory, it is possible to adjust for spatial inequality in the function describing the "representative" consumer.

behavior does not automatically translate into optimal policy rules. There exists, however, a solid tradition that people‘s preferences should count when deriving economic policy rules. When an economist or a decision-maker chooses an ethical value for the society‘s time preference on a normative basis, it does not reflect anything else than this particular person‘s view of intergenerational distribution. This is at the heart of Nordhaus‘ (Nordhaus, 2007) critique of Stern. That might lead us to search for revealed preferences, for instance through the interest rate markets. This is the equilibrium rate describing people‘s actual choices. Depending on the extent to which markets are efficient, we might get some answers. The only problem is that we need to know which question we are asking. Do consumers reveal ethical values on intergenerational distribution on the social level, when they make individual consumption planning in short-term markets? It is widely held that consumers have a positive pure rate of time preference, and on an individual level, this reflects not only the instinctive impatience as mentioned in section 1.3, but also that we care more for ourselves, and more for our children than future grandchildren and so on. Even when individuals have private bequeath motives for their children, they are not necessarily revealing what they think society as a whole should leave for the next generation. Sen (1967) proposes a rationale for this, famously known as the Isolation Paradox. The paradox is that each person is better off entering a contract that ensures everyone to save more for the future. But when they act in isolation, total saving for the future will be below what they collectively desire. The Isolation Paradox seeks to explain a situation of collective action, where a certain threshold for saving must be achieved; conditioning on the threshold being reached, it is desirable to save for the future. It is somewhat similar to the standard equilibrium of a game where everyone benefit from each individual‘s effort; both situations will require coordination. According to the Isolation Paradox, no one will benefit when level of saving is below the threshold. So when individuals act in isolation in the market, they undersave for future generations compared to a coordinated governmental program. This rationale of Sen is much debated, but it is still a good example of the friction between descriptive and prescriptive theory.

In the concluding section of Beckerman and Hepburn (2007) the conflict between the a priori choice of ethical values and the revealed ethical values of markets is doomed to continue endlessly, and the authors propose to undertake methods directly pointed to people‘s preferences, such as ”the use of stated preference surveys, behavioral experiments and methods to reveal the social preferences inherent in our social institutions”. These methods have problems of their own, of course. As controversial and innovative as they may be, they lie in the tradition of cementing economic policy in people‘s actual preferences.

2.2 The Ethics of Discounting

While future progress on socially efficient discount rates will have its source in economic theory, it seems in no way possible to get around the underlying ethics of the economic models, economic literature or in the concept of long-term discounting itself. When considering the many caveats of assessing the present generation‘s preferences for future generations‘ welfare, I take the stand that ethical issues deserve a great deal of attention. Therefore I will devote a section to review the ethics separately.

In the intercepting field of philosophy and economics, discounting issues from an economic point of view is thoroughly explained by Broome (1994). The typical philosopher‘s question is, as he writes: ”How, they ask, can the mere date at which a good occurs make any difference of its value?”(Broome, 1994). Ruling out the practice of discounting utility, the disagreement is more of a misunderstanding in Broome‘s view. One might get the impression that he believes in consensus on the ethics of discounting utility. This is not true of course. Zero-discounting of utility entails utilitarianism, which is much debated as with utilitarianism in economics in general (Sen et al., 1982). Agent-relative ethics in terms of utility discounting is discussed in Beckerman and Hepburn (2007). Agent-relative ethics, as they see it, can be tracked all the way back to the work of David Hume, describing the moral pattern that agents care more for their closer relations. We care more for our close family, than our distant, and we care more for fellowing citizens than citizens of other towns and nations. This is just describing the prevailing moral of course, not making this line of thinking automatically normative. The authors' main point is that agent-relative ethics is a fair alternative to utilitarianism, and is more in the line with the moral people actually seem to exhibit. Compared to the Stern Review's argument that we care for our children and grandchildren to the extent that δ reflects the extinction rate of society, agent-relative ethics is compatible with utility-discounting to the extent that we care relatively more for our children than grandchildren and so on.

In the recent literature, Nordhaus (2007) seems to be the strongest critique of zero utility discounting, while Dasgupta (2005, 2007) clearly criticizes the inconsistency between a low δ and low θ, see section 2.2.1. There are several possible ethical evaluation concepts and Nordhaus (2007) mentions some of them:

- Sustainable development[7], which he describes as leaving at least as much capital as inherited, using a broad definition of societal capital. Sustainable development is widely known as "development that meets the needs of the present without compromising the ability of future generations to meet their own needs" (World Commission on Environment and Development, 1987). Arrow et al. (2004) interpret this formally, stating that intertemporal social welfare must be non-decreasing over time. If V_t is the intertemporal social welfare function, then the criterion is $\frac{dV_t}{dt} \geq 0$. As the authors point out, the criterion will not determine a unique path of consumption, neither will it ensure the efficient path.

- The perspective of Rawls' "veil of ignorance", which implies maximizing the welfare of the poorest generation, is consistent with a very high value of θ. This means that ex-ante, before we are assigned to a generation, we want to level the field, by perfectly smoothing consumption across generations and states of nature, as far as possible. In the discussion of such an ethical view Dasgupta (2005) states that there would either be no saving, or we would get intergenerationally inconsistent behavior. If we save, the future generation

[7]Chichilnisky (1996) shows that the concept of sustainable development is incompatible with utility discounting.

will be better off, and this is ruled out by maximizing the welfare of the poorest. If we care for a finite number of descendants, the plans of dissaving at some point in the future would be reversed, as the generation at that point of time cares for a finite number of their descendants. If we care for every future generations, we would get a sort of a Ramsey model, with or without utility discounting. Saving would again be zero by maximizing the welfare of the poorest generation.

- A precautionary principle, which he interprets as maximizing the minimum consumption along the riskiest path. The precautionary principle states in most applications, such as innovation with possible negative side-effects, that when uncertainty about consequences is sufficiently great, one should refrain from utilizing the innovation. In terms of global warming, this implies taking huge costs to reduce emissions of greenhouse gases and it will be consistent with maximizing consumption along the riskiest path.

- Considering non-anthropocentric values, such as, intrinsic values derived from ecology and religion.

Nordhaus (2007) seems to take the position that none of the competing ethical views actually helps us narrow down an efficient discount rate, but concludes that opportunity cost of capital is the relevant baseline. The premise for this thesis is that we can‘t observe the efficient market rate for long horizons, and we will soon treat the subject of opportunity cost of capital in the long run based on theoretical models.

2.2.1 Ethical Consistency: δ and θ

As Dasgupta (2007) has argued, a low δ combined with a low θ will put disproportionally high weight on intergenerational equity, compared to intragenerational equity. Nordhaus argues that $\theta = 1$ is about the lowest plausible value held by economists. In any way it seems unreasonable to combine low values of both δ and θ, not only because of its unreasonable implications, but most of all because of the paradoxical issues with respect to equality. Zero discounting of utility is a complete egalitarian view of present and future generations. A low elasticity of marginal utility is a very inegalitarian view of distribution within a particular generation. So why should one care more for people in the future than for people in the present world? This is one of the main points of Schelling (1995), who stresses that the discounting model of Ramsey‘s legacy is perfectly reasonable for the individual. On the social level, when considering future generations, he argues that the question is more of a pure ethical character. Just like aid to poor countries is a political issue today. Even though the utilitarian view, by decreasing marginal utility, implies transferring vast amounts of goods to countries on lower consumption levels compared to the western, richer world; this is not what we observe anywhere in the world today. Schelling (1995) implies that political goals should determine redistribution in both present time from rich too poor, and through time from rich generations to poor generations. This could help determining the amount to invest for our descendants, still; the most efficient way to choose between redistributional projects involves determining a socially efficient discount rate.

The duality of the choice of these ethical parameters is pointed out by Dasgupta (2007): By fixing the growth rate g and the welfare-preserving rate ρ, all combinations of δ and θ that satisfy $\delta = \rho - \theta g$ will give the same value of ρ. Even though the ethical parameters are conceptually different, different choices could yield the same result. Both Dasgupta (2007) and Nordhaus (2007) refer to the doctrine of Koopmans, that we should be careful in making a priori choices for the ethical parameters δ and θ, since the model is far too complicated to have any feeling to their implications. Stern‘s and many other scholars‘ choices for parameters in the long run are nothing less than ethical choices; they are not universally embraced or necessarily representative and they seem to be inconsistent with actual observations of behavior.

Part 3

Declining Discount Rates

I have not yet analyzed models treating risk of future consumption growth or interest rates. As we will see, the assumptions about risk and attitudes towards risk are crucial in determining socially efficient discount rates. Firstly, they might behave very differently. Secondly, they might be more suitable for real world applications. As a substitute for using market interest rates, which by assumption fail to incorporate the valuation effects of long-term risks, properly modelling of risk is a powerful tool, at least in saying something about the term structure of the discount rate. Having in mind the academic debate outlined in the previous section, the framework which I will look at is not immune against ethical considerations or the need to estimate the utility parameters properly. But the models will be much more realistic, and they have the potential to bridge the gap between the "opportunity cost view" and the advocates of "exponential absurdity", based on economic principles rather than ethics alone. I start by making clear the distinction between hyperbolic preferences and declining discount rates. Then I derive necessary conditions for applying declining discount rates based on macro-risk. The implications for timing issues is discussed and as we will see, we may be able to reject the usual assumption of time inconsistency.

3.1 Hyperbolic Preferences vs. Declining Discount Rates

The concept of hyperbolic discounting is in general not referring to a mathematical hyperbolic function, but to a discount rate that declines as the future date becomes more distant relative to the evaluation date. This is not to be confused with a discount rate that declines due to circumstances related to specific points of time (Rasmusen, 2008). Hyperbolic discounting has been associated with utility discounting for decades but in recent times, the concept has entered into the studying of discounting goods as well. The term "hyperbolic preferences" is still reserved to utility discounting, δ, which is a form of hyperbolic discounting in the spirit of Strotz (1955), Phelps and Pollak (1968) and Laibson (1997, 1998).

One could obtain declining discount rates that depend on the length of time[1], or simply because something related to specific dates gives a declining pattern (the trivial case). In the recent literature, "declining discount rates" is the common term used for discount rates depending on the length of time[2], even though the term in a strict sense encompasses all the different types of declining discount rates, including the trivial case.

The definitions and the uses of them are obviously confusing, but incredibly important when it comes to possible implications for timing issues. It almost seems as if there is a need to come up with a new term, defined as "hyperbolic discounting of economic goods". It could solve two problems: the word "hyperbolic" is by many scholars associated with utility discounting, and "declining" fails to differentiate between trivial and nontrivial cases of declining discount rates.

I will specify "hyperbolic preferences" for utility discounting. In most cases when I use the term "declining discount rates", I am referring to "declining" in the sense that the discount rate declines as the discounted payoff becomes more distant

[1] As we will see, appropriate assumptions on the stochastic process rationalize such discount rates.

[2] Gollier et al. (2008), Pearce et al. (2003), Weitzman (1998, 2001) and many others use this term in order to describe discount rates declining with the time horizon.

from today's point of view. It will be explicitly specified when it refers to a trivially declining discount rate.

3.1.1 Trivial Cases of Declining Discount Rates

Assume δ is constant, i.e. non-hyperbolic preferences. If we look at the s.s. Ramsey equation (1.17), there are two reasons for getting declining discount rates in the absence of risk. One can either believe that the labor-augmenting technology growth (g) will decline in time, or the utility function could reflect decreasing elasticity of marginal utility (θ) as in equation (2.4)[3]. As long as the product $(\theta \cdot g)$ is declining with time, the Ramsey discount rate will as well. These reasons for getting a declining pattern of discount rates are more of a trivial character. The discount rates will be declining due to predictable or expected events, and they are not related to the time horizon itself.

3.2 Introducing Risk in The S.S. Ramsey Equation

3.2.1 Prudence

Prudence is the idea that the optimal response when the future becomes more uncertain is to increase saving to be prepared. This section derives the sufficient condition for prudence. Let us introduce a two-periodic model where the agent is able transfer income at the sure interest rate. There is no project risk, but the agent faces risk from uncertain income the next period. As usual, $u'(c) > 0$ and $u''(c) < 0$. Define m_0 as the first period's income, s as the savings rate, and $\tilde{m}_t$ as the second period's stochastic income. We will also use m_t, the deterministic income. Define $\tilde{m}_t = m_t(1 + \epsilon_t)$ where ϵ_t is a stochastic, mean zero variable. By accounting relationships, the total consumption the first period is $(m_0 - s)$. The second period consumption is $\tilde{m}_t$ plus the gross return on savings $(\tilde{m}_t + e^{rt}s)$. The

[3] This argument for a declining pattern of discount rates is analyzed by Gollier (2002, 2007).

constrained optimization problem is

$$max_s U(s) = u(m_0 - s) + e^{-\delta t} E[u(\tilde{m}_t + e^{rt} s)] \tag{3.1}$$

A sum of strictly concave functions is strictly concave, so the first-order condition is sufficient:

$$U'(s) = u'(m_0 - s)(-1) + e^{-\delta t} E[u'(\tilde{m}_t + e^{rt} s)] e^{rt} = 0 \tag{3.2}$$

Assume there is no risk. Then the first-order condition looks like this:

$$U'(s) = u'(m_0 - s)(-1) + e^{-\delta t} E[u'(m_t + e^{rt} s)] e^{rt} = 0 \tag{3.3}$$

Now, we introduce the risk. The only change is in the last term in the first-order condition. If $U'(s) > 0$ after introducing risk in m_t, then

$$e^{r-\delta} E[u'(\tilde{m}_t + e^{rt} s)] > e^{r-\delta} [u'(m_t + e^{rt} s)] \tag{3.4}$$

This is true if, and only if,

$$E[u'(c_t)] > u'(E[c_t]) \tag{3.5}$$

By Jensen‘s inequality, this is true if, and only if, marginal utility is strictly convex. The necessary condition for prudence is that $u'''(c) > 0$. This implies that a socially efficient discount rate in a situation with risk in future income is lower than compared to the certainty case. The general class of utility functions satisfies the condition of prudence, except the quadratic utility:

$$u(c) = ac - bc^2 \implies u'''(c) = 0 \tag{3.6}$$

For the rest of the analysis, I will assume convex marginal utility and the only specific utility function I will work with is the general power functions:

$$u(c) = \frac{(c)^{1-\theta} - 1}{1-\theta}, \quad \theta > 1 \implies u'''(c) = \theta(\theta+1)c^{-(\theta+2)} > 0 \tag{3.7}$$

$$u(c) = ln(c) \implies u'''(c) = 2c^{-3} > 0 \tag{3.8}$$

I have now shown that $u'''(c) > 0$ is a necessary condition for an optimizing agent to save more when the future becomes more risky.

3.2.2 An i.i.d. Stochastic Process

Consistent with the s.s. Ramsey equation (1.17), we now extend the framework of section 3.2.1 to analyze an independent and identically distributed growth process of consumption. As a part of this, we need to define the general socially efficient discount rate function. Without loss of generality, we assume two periods, 0 and t as in section 3.2.1. The social welfare function is

$$U = u(c_0) + e^{-\delta t} E[u(c_t)] \tag{3.9}$$

Assume optimizing agents and frictionless markets. Assume $1 + r$ is the marginal rate of transformation between time 0 and t. Equation (3.9) is concave, hence equation (3.10) is the sufficient condition to solve the consumers‘ problem:

$$-u'(c_0) + e^{(r-\delta)t} E[u'(c_t)] = 0 \tag{3.10}$$

Taking logarithms and solving for r gives

$$r_t = \delta - \frac{1}{t} ln\left(\frac{E[u'(c_t)]}{u'(c_0)}\right) \tag{3.11}$$

This is the general socially efficient discount rate function. Compared to the s.s. Ramsey equation (2.3), the last term now incorporates the effect of prudence in addition to the effect of decreasing marginal utility. To illustrate this, consider the case with power utility, $t = 1$ and assume that the increase in logarithm of consumption is normally distributed[4]. Then we get

$$r = \delta + \theta g - \frac{1}{2}\theta(\theta + 1)\sigma^2 \tag{3.12}$$

See appendix B.1 for proof.

This is just the s.s. Ramsey equation with an extra term describing the isolated effect of prudence, which lowers the Ramsey discount rate. In a Ramsey model equilibrium, the lower market interest rate comes from the fact that consumers are investing relatively more in productive capital which has decreasing marginal product. Obviously, the prudence effect increases with the memoryless variance,

[4]In fact, normal distribution is not that restrictive compared to a general i.i.d. process, because when time is large enough, the general i.i.d. process approaches a normal distribution.

and elasticity of marginal utility (θ), which also is interpreted as the coefficient of relative risk aversion. The higher aversion to inequality, the higher the willingness to save for the risky future. Observe that relative prudence can be defined in line with the Arrow-Pratt measure of relative risk aversion:

$$P = -c\left(\frac{u'''(c)}{u''(c)}\right) = (1+\theta) \tag{3.13}$$

The more prudent we are, the more willing we are to save for the future. When marginal utility is not convex ($P = 0$) the prudence term disappears, even though we are risk averse.

What is left to do now is to show that an i.i.d. stochastic process implies a constant discount rate departing from equation (3.11). This is exactly what is done in a working paper by Gollier (2007). I will focus on the result and intuition here while a formal illustration of the result is found in appendix B. Define the consumption growth technology:

$$c_{t+1} = c_t e^x \tag{3.14}$$

where x is independent and identically distributed each short period of time. Combining these properties with equation (3.11), we end up with a constant discount rate, without imposing normal distribution:

$$r = \delta - ln\left(E[e^{-\theta x}]\right) \tag{3.15}$$

This is, of course given power utility and constant pure rate of time preference (δ). When the growth of consumption is i.i.d., the per periodic risk does not increase with the time interval. The representative agent faces the exact same risk in each time period. The variance of log consumption and expected log consumption increase proportionally with time. This means that the effect of prudence and the effect of decreasing marginal utility cancel each other out every instant of time, and hence the term structure of the discount rate is flat[5].

[5]This insight is pointed out in Gollier (2007).

3.2.3 A Persistent Shock Stochastic Process

There are many stochastic processes which entail persistent shocks to consumption, but I only show the case of a non mean reversion process. The implications of a mean reversion process are illustrated in Gollier (2007) and Gollier (unpublished manuscript). I will not illustrate the mean reversion for two reasons:

1. A mean reversion process will give two effects. The expectation of future interest rates would be increasing or decreasing due to date-specific events, depending on the expected movement of the business cycles. This is similar to the yield curve of market interest rates following a mean reversion process. The other effect is that increased risk gives a lower long-term discount rate, compared to an i.i.d. stochastic process. We are mostly interested in this effect in isolation. Even though a mean reversion process has some degree of persistence in it, it fails to illustrate our point clearly [6].

2. The calculations are messy and the effect of persistence is not intuitively easy to show, at least not as efficiently as with the model in this section.

The presentation that follows here is based on own calculations.

We still have the technology of consumption growth as before:

$$c_{t+1} = c_t e^{x_t} \tag{3.16}$$

Gollier (unpublished manuscript) analyzes the following model:

$$x_t = \phi x_{t-1} + (1-\phi)\mu + \epsilon_t \tag{3.17}$$

$$\epsilon_t \sim N(0, \sigma^2) \tag{3.18}$$

$$\phi \in (0,1) \tag{3.19}$$

We will set $\phi = 1$, and the process will then become purely persistent:

$$x_t = x_{t-1} + \epsilon_t \tag{3.20}$$

[6] In fact, the mean reversion properties work to limit the effects of persistence on long-term valuation (Gollier, unpublished manuscript).

$$\epsilon_t \sim N(0, \sigma^2) \tag{3.21}$$

Rewrite eq. (3.16)

$$ln(c_{t+1}) = ln(c_t) + x_t \tag{3.22}$$

By forward induction, we end up with:

$$ln(c_t) - ln(c_0) = tx_{-1} + t\epsilon_0 + (t-1)\epsilon_1 + (t-2)\epsilon_2 + \cdots + \epsilon_{t-1} \tag{3.23}$$

Note that the relative growth rate of consumption g is normally distributed (because ϵs are) and defined as

$$g = ln(c_t) - ln(c_0) \tag{3.24}$$

Inserting into equation (3.11), and using power utility, we obtain:

$$r_t = \delta - \frac{1}{t} ln \left(E[e^{-\theta(lnc_t - lnc_0)}] \right) \tag{3.25}$$

This describes the term structure of the socially efficient discount rate. Because the ϵs are uncorrelated we can write the variance as:

$$Var\left(ln(c_t) - ln(c_0)\right) = \sigma^2 \sum_{\tau=1}^{t} \tau^2 \tag{3.26}$$

By using a formula for log-normal distributions (see appendix B.3), we end up with the (specific) socially efficient discount rate function:

$$r_t = \delta + \theta x_{-1} - \frac{1}{t} 0.5\theta\sigma^2 \sum_{\tau=1}^{t} \tau^2 \tag{3.27}$$

It is easy to see that the socially efficient discount rate is declining, and this is proven in appendix B.3. It has nothing to do with date-specific events. It is the length of the time interval that increases the riskiness. This is because of positive serial correlation of shocks. The annualized risk in log consumption increases with time when shocks are persistent.

See appendix B.3 for a complete derivation.

3.2.4 Parametric Uncertainty

Another approach to rationalize declining discount rates is based on the realistic assumption of parametric uncertainty. It is unlikely that we know the parameters driving the stochastic process of consumption growth. This is often related to extreme events happening too rarely to be counted properly in the learning process. I will not discuss the extent to which we are ignorant about the parameters, but present a model that leads us to average discount factors rather than discount rates, based on subjective probabilities. The model is borrowed from Gollier (unpublished manuscript). Suppose the stochastic process of c_t is a function of an unknown parameter γ. Assume there are n number of possible values for γ, and that we have subjective probabilities q_γ of the value of γ. Then we can write expected utility as:

$$E[u'(c_t)] = \sum_{\gamma=1}^{n} q_\gamma E[u'(c_t)|\gamma] \tag{3.28}$$

We insert this directly into eq. (3.11):

$$r_t = \delta - ln \sum_{\gamma=1}^{n} q_\gamma \frac{E[u'(c_t)|\gamma]}{u'(c_0)} \tag{3.29}$$

Then, for a given γ, the efficient discount rate is

$$r_t(\gamma) = \delta - \frac{1}{t} ln \frac{E[u'(c_t)|\gamma]}{u'(c_0)} \tag{3.30}$$

If eq. (3.29) and (3.30) are put together, we end up with the discount factor function:

$$DF(\gamma, t) = e^{-r_t t} = \sum_{\gamma=1}^{n} q_\gamma e^{-r_t(\gamma)t} \tag{3.31}$$

This function is strictly convex in r, so by Jensen‘s Inequality,

$$DF_t(E_\gamma[r_t]) > E_\gamma[DF_t(r_t)] \tag{3.32}$$

$$\Longrightarrow \; r_t < E_\gamma[r_t(\gamma)], \forall t > 0 \tag{3.33}$$

This states that the socially efficient discount rate in general is lower than the expected interest rate. This result comes from increasing and concave utility combined with parametric uncertainty. But we want to say something more. By the same assumptions it can be shown that r_t as defined by eq. (3.31) has the following properties:

$$r_0 = E_\gamma[r_0(\gamma)] \tag{3.34}$$

$$\frac{dr_t}{dt} < 0 \tag{3.35}$$

$$\lim_{t \to \infty} r_t = min\{r_t(\gamma)\} \tag{3.36}$$

Proof of this is identical to the proof in appendix C.1. The socially efficient discount rate declines with time, and approaches its lowest value in the limit when time goes to infinity. This result is obtained by assuming parametric uncertainty and that leads us to calculate discount rates based on the expected discount factors conditioning on γ, rather than the expected interest rates. As t gets larger, the small $r_t(\gamma)$s become relatively more important in the discount factor.

3.3 The Weitzman-Gollier Approach

In this section, Weitzman‘s argument for declining discount rates is explored. We present the puzzle that arises from the critique, and a possible solution to solve the puzzle.

3.3.1 Weitzman‘s Argument

In a series of papers, Weitzman (1998, 2001) (as well as Weitzman (2007)), advocated the idea that when the socially efficient discount rate is uncertain, it isn‘t discount rates that should be probability-averaged, but rather the discount factors. The implication of this is that discount rates are declining with time, and in the limit when $t \to \infty$, denoted as the far-distant future, the lowest possible rate should be used. Note, as Weitzman (1998) is stressing, that this result relies on the assumption of a non mean reverting stochastic process for the interest rate. Technically, he assumes a constant discount rate to be revealed after the decision, while he points out that the result holds for a mean reverting stochastic process with a coefficient of reversion that goes towards zero in the limit. Based on the argument that discount factors should be averaged, given a pure persistent shock stochastic process he obtains the definition of a socially efficient discount rate $R^W(t)$ [7]:

$$R^W(t) = -\frac{1}{t}\sum_{i=1}^{n} ln\left(p_i e^{-r_i t}\right) \tag{3.37}$$

[7]Note that Weitzman (1998) uses a different notation, but one that is equivalent. The notation here is the exact notation used in Gollier and Weitzman (2010).

$$R_0^W = \sum_j^n p_j r_j = E[r] \tag{3.38}$$

$$\frac{dR^W}{dt} < 0 \tag{3.39}$$

$$\lim_{t \to \infty} R^W = min\{r_j\} \tag{3.40}$$

Proofs of these properties are found in appendix C.3.

Weitzman (2001) gathers numerical opinions on a long-term discount rate from 2,160 PhD-level economists from 48 countries, and he uses the data to estimate a Gamma-distribution of what the socially efficient discount rate for long-term applications should be, based on the survey. He landed on the following estimated interest rate function:

$$R(t) = \frac{\mu}{1+\frac{t\sigma^2}{\mu}} \tag{3.41}$$

This function has the similar properties as eq. (3.37).

$$t = 0 \implies R(0) = \mu \tag{3.42}$$

$$\frac{dR(t)}{dt} = -\left(1 + \frac{t\sigma^2}{\mu}\sigma^2\right) < 0 \tag{3.43}$$

$$\lim_{t \to \infty} R(t) = 0 \tag{3.44}$$

The only difference is that the interest rate goes to zero as t approaches infinity, due to the particular specification of the discount rate function. The key assumption is on averaging discount factors rather than rates, and Weitzman (2001)-result may just be interpreted as a specific form of a declining discount rate function with the probability distribution estimated from 2,160 professionals. Note, and this is fairly important, that even though it is not stated explicitly, the assumption of a pure non mean reverting process is just as important here as it is in Weitzman (1998). Once the interest rate is revealed it is constant. This can be observed from his formulation of the discount factor function (Weitzman, 2001):

$$A(t) \stackrel{\text{def}}{=} \int_0^\infty e^{-xt} f(x) dx \tag{3.45}$$

For each state, x goes into the discounting function as a constant value for all ts.

In Weitzman (2007), the argument is illustrated in yet another perspective, where he uses a similar model and argumentation treating project risk. This model falls out of the strictly interpreted scope of this thesis, but is briefly illustrated in section 4.3.

3.3.2 Gollier‘s Critique

In a critique of Weitzman‘s argument, Gollier (2004) pointed out that the NPV-rule used is inconsistent with the opposite financing strategy based on net forward value, the NFV-rule. The two criteria are mathematically equivalent, but they entail complete different implications for risk bearing. Since preferences are not stated in Weitzman‘s argument, the correct answer to the question is out of the model‘s reach. It is not about the hypothesis of declining discount rates per se, but rather on the fact that a different financing strategy turns the argument upside down, suggesting that the scientific basis of Weitzman (1998) is insufficient in providing a good story to back up the theory. This point is stressed in Gollier (2004) as he writes: "In fact, to tell the truth, I believe that we are both wrong, because our criteria are arbitrary, as they do not rely on actual preferences." Before we move to a presentation of the puzzle solution, let‘s briefly illustrate the different criteria.

Net Forward Value (NFV)-evaluation

Gollier considers a project with a sure gross return Z_t in comparison with an investment yielding the marginal rate of return on capital in the economy. We are uncertain about the $\tilde{x}$, and Gollier considers the evaluation criterion where risk is allocated to the future date of payoff. The NFV-criterion is then:

$$-e^{R(t)t} + Z_t \geq 0 \tag{3.46}$$

, where $R(t)$ is defined by

$$e^{R(t)t} = E[e^{\tilde{x}t}] \tag{3.47}$$

This implies that the "certainty equivalent" discount rate $R(t)$ is increasing with time. As t becomes larger, the higher values of $\tilde{x}$ become relatively more important[8].

[8] Or you could, as Gollier (2004) does, view $E[e^{\tilde{x}t}]$ as an implicit utility function with absolute degree of risk aversion equal to $-t$. It is well known that a decrease in absolute risk aversion raises the certainty equivalent.

Net Present Value (NPV)-evaluation

Weitzman's approach is to evaluate the payoff at present time, in which risk is allocated:

$$e^{-R^W(t)t} = E[e^{-\tilde{x}t}] \tag{3.48}$$

and the NPV-criterion advising to undertake the project if, and only if,

$$Ze^{-R^W(t)t} - 1 \geq 0 \tag{3.49}$$

The same intuition as with the NFV-rule applies with opposite sign. The larger the t, the more relative weight is attached to the smaller values of $\tilde{x}$.

The NFV-rule and the NPV-rule should yield the same policy advices if the model is right. This is not happening here and that is why Gollier criticizes the story of Weitzman and asks for building models that rely on preferences.

3.3.3 The Puzzle Solution

There are several proposals for solving the Weitzman-Gollier Puzzle, including Buchholz and Schumacher (2008) and Hepburn and Groom (2007)[9], but here I will focus on the solution proposed by the very authors from which the puzzle originated, as presented in the conciliatory paper of Gollier and Weitzman (2010). I attempt to explain the main insight here.

The solution relies on the agent´s preferences and trivially the result of Weitzman holds if the agent's utility is $ln(c)$. This is because optimal consumption at time zero is independent of the interest rate[10]. Then a NPV-rule can be applied, as opposed to a NFV-rule. In order to solve the puzzle in general, Gollier and Weitzman (2010) show that in the context of expected utility, the NPV-rule and NFV-rule equate perfectly when using risk-adjusted probabilities.

[9]These two approaches are similar in the sense that the puzzle is resolved in favor of a declining pattern of discount rates. Buchholz and Schumacher (2008) rely on a proper modelling of risk aversion, while Hepburn and Groom (2007) study the evaluation date more closely.

[10]It is a well-known property of logarithmic utility that the substitution effect is exactly offset by the wealth effect.

Consider the three different dates: 0^* is the time when a decision must be made, and right after comes time 0, the date at which the true, permanent interest rate is revealed. At last, time t is the date at which the payoff is received. At time 0^* we have probabilities p_s of the interest rate r_s, where $\sum_s p_s = 1$.

In line with the previous section, define Z as the present value of the required capital to undertake the investment, and ϵ as the future value of a sure payoff to be received at time t.

Before an investment opportunity comes along, the agent is presumed to make state-contingent optimal consumption plans. To show this, consider a standard intertemporal framework similar to the one derived in part 1:

$$V(C) = \sum_{t=0}^{\infty} e^{-\delta t} U(C_t) \tag{3.50}$$

The Inada conditions on $U(C)$ combined with $\delta > 0$, will insure a bounded solution.

For the purpose of this exercise think of a general function:

$$V(C) = \text{intertemporal welfare} \tag{3.51}$$

We are mainly interested in how the optimal rule looks like, and not so much in how the intertemporal welfare function $V(C)$ looks like.

If we assume that the underlying production function is linear, we get constant state-contingent interest rates independent of the level of investment. Consistent with the risk-free Ramsey model presented in part 1, the consumer now makes optimal state-contingent plans[11]:

$$\frac{\partial V(C_s*)}{\partial C_0} = \frac{\partial V(C_s*)}{\partial C_t} e^{r_s t}, \forall t \tag{3.52}$$

, where r_s is the state-dependent interest rate.

Relying on departing from optimum, we can use the Envelope Theorem[12] to derive the following evaluation criterion: It is optimal to undertake the project if, and only if expected gain in intertemporal welfare is higher than expected loss in intertemporal welfare:

$$\epsilon \sum_{s=1}^{n} p_s \frac{\partial V(C_s^*)}{\partial C_t} \geq Z \sum_{s=1}^{n} p_s \frac{\partial V(C_s^*)}{\partial C_0} \tag{3.53}$$

[11]We assume that the transversality condition holds.

[12]If we start out with optimal plans, the initial consumption plans will not change on the margin.

The Weitzman Approach

Weitzman would want to express the criterion using expected marginal intertemporal welfare at time 0. Substituting for $\frac{\partial V(C_s^*)}{\partial C_t}$ from eq. (3.52) into eq. (3.53) we end up with

$$\epsilon \sum_{s=1}^{n} q_s^W e^{-r_s t} \geq Z \tag{3.54}$$

, where

$$q_s^W \stackrel{\text{def}}{=} \frac{p_s \frac{\partial V(C_s^*)}{\partial C_0}}{\sum_{s=1}^{n} p_s \frac{\partial V(C_s^*)}{\partial C_0}} \tag{3.55}$$

This expression goes into the usual Weitzman discount rate function:

$$R^W(t) = -\frac{1}{t} ln \left(\sum_{s=1}^{n} q_s^W e^{-r_s t} \right) \tag{3.56}$$

This should be familiar. It is similar to eq. (3.37) except that we now are using risk-adjusted probabilities. The difference is in how the utility function is used to price the different states by adjusting the probabilities.

The Gollier Approach

Gollier would like to evaluate the payoff at date t, so he would want to substitute $\frac{\partial V(C_s^*)}{\partial C_0}$ from eq. (3.52) into eq. (3.53). Then we end up with:

$$\epsilon \geq Z \sum_{s=1}^{n} q_s^G e^{r_s t} \tag{3.57}$$

,where

$$q_s^G \stackrel{\text{def}}{=} \frac{p_s \frac{\partial V(C_s^*)}{\partial C_t}}{\sum_{s=1}^{n} p_s \frac{\partial C_s^*}{\partial C_t}} \tag{3.58}$$

In line with Gollier (2004), the discount rate function can now be written as:

$$R^G(t) = \frac{1}{t} ln \left(\sum_{s=1}^{n} q_s^G e^{r_s t} \right) \tag{3.59}$$

This is also familiar and R^G is similar to the discount rate defined by eq. (3.47), except that probabilities are adjusted in order to price the different states efficiently.

The Properties

$$R^W(t) = R^G(t), \forall t \tag{3.60}$$

These discount rate functions are quantitatively and qualitatively equal, so let‘s denote this unique function as R^{WG}. It can be shown that R^{WG} has the same properties (1-3) as eq. (3.37) (Weitzman, 1998).

$$R_0^{WG} = \sum_j^n p_j r_j = E[r] \tag{3.61}$$

$$\frac{dR^{WG}}{dt} < 0 \tag{3.62}$$

$$\lim_{t \to \infty} R^{WG} = min\{r_j\} \tag{3.63}$$

We now see that using a framework where preferences are defined and the agent makes optimal decisions, the qualitative result of Weitzman (1998) holds and we get declining discount rates. The new, risk-adjusted probabilities come from a type of no-arbitrage condition in an economy where agents behave optimally. In fact, the term "no-arbitrage" here is slightly misleading. We are not dealing with arbitrage per se, rather we are speaking of arbitrage between present and future selves. If NFV- and NPV-evaluation do not give the same valuation, the agent will be better off simply by changing the financing strategy.

3.4 Approaches Concerning Pure Rate of Time Preference

As part 2 shows, there is some controversy in exponential discounting for its negligence of future generations. A too low or a zero discount rate place too much weight on the future, and it is inefficient in the short and medium run.

The literature on intergenerational equity is simply about the balancing of two conflicting considerations: the present generations‘ welfare versus the future generations‘ welfare. This ethical view could be represented by a declining utility discounting rate, δ_t, as I interpret the combined work of Chichilnisky (1996) and Li and Löfgren (2000). In a paper on sustainable development, Chichilnisky (1996) formulated this conflict formally in two axioms, which together require no dictatorship of the present generation over the future generation, and vice versa. Chichilnisky (1996) finds that none of the conventional optimality criteria satisfies the stated axioms. If we take the discounted utility model, exponential discounting entails dictatorship of the present. Zero discounting on the other hand, will give an unbounded solution, and a likely implication is the dictatorship of the future. By combining these two considerations, a sustainable path is attainable. That is, applying a positive utility discount rate and adding a term with the long run average of the sequence of utilities:

$$W = e^{-\delta t}U(C_t) + \Phi(U(C)) \tag{3.64}$$

Li and Löfgren (2000) show that this approach applied for renewable resources, such as productive capital, implies declining discount rates. This would be a kind of discount rate function reflecting preferences for equal treatment of different generations, and this function would be deterministic in time. Hence, it has similar implications for time inconsistency as hyperbolic preferences[13].

Another approach when defending declining pure rate of time preference is to observe that individual behavior reveals hyperbolic preferences. Frederick et al. (2002), as an example, suggest that there is some evidence of behavior consistent with hyperbolic preferences. In their review of "Recent advances in social discounting", Pearce et al. (2003) use the assumption that individuals seem to discount utility hyperbolically as an argument for applying declining discount rates in public CBA, under the parole that actual preferences should count for policy. However, the literature is ambiguous in its conclusions on hyperbolic preferences being efficient in the sense that it will ensure the optimal path for the consumer. One debate is about the optimality of individuals that discount utility hyperbolically and commitment strategies[14], the other is on the link between positive and normative theory as discussed in section 2.1.2. It is therefore not obvious that the practice of hyper-

[13]See section 3.5.2 for a discussion of these implications.

[14]Much treated in the literature inspired by Strotz (1955) and Phelps and Pollak (1968).

bolic discounting of utility observed among individuals translates into social level hyperbolic preferences.

3.5 Time Inconsistency

3.5.1 Definition of Time Inconsistency

If an agent‘s optimal actions change with time, assuming no new information is made available in a strict sense, then we are dealing with time inconsistent behavior. In such cases, what we plan for will be regrettable due to the passage of time alone. Time consistency is not in contradiction to behavior that changes optimally in response to new information, but time consistency requires optimal state-dependent plans to hold through the passage of time.

3.5.2 Hyperbolic Preferences

This section contains a simple illustration of time inconsistency when pure rate of time preference, δ_t declines hyperbolically. This model, which is inspired by Gollier et al. (2008), is an adapted version of the model that is derived in the next section. Pure rate of time preference is denoted by δ_t and the interest rate is for simplicity assumed to be zero. Total income is m and c_t is consumption in period t. Utility is as usual monotonically increasing and strictly concave. The optimization program in period 0:

$$max_{c_0,c_1,c_2} L_0 = u(c_0) + u(c_1)e^{-\delta_1} + u(c_2)e^{-2\delta_2} + \lambda_0 \left(m - c_0 - c_1 - c_2\right) \tag{3.65}$$

The first-order condition ensuring optimal bundles of consumption period 0 is:

$$\lambda_0 = u'(c_0) = u'(c_1)e^{-\delta_1} = u'(c_2)e^{-2\delta_2} \tag{3.66}$$

We denote the resulting optimal bundle as (c_0^*, c_1^*, c_2^*). In period 1, the optimization program the agent faces is like this:

$$max_{c_1,c_2} L_1 = u(c_1) + u(c_2)^{-\hat{\delta}} + \lambda_1 \left((m - c_0^*) - c_1 - c_2\right) \tag{3.67}$$

The first-order condition ensuring optimal bundles of consumption period 1 is:

$$\lambda_1 = u'(c_1) = u'(c_2)e^{-\hat{\delta}} \tag{3.68}$$

By substituting for $u'(c_1)$ in the the first-order condition for period 0, we can check if the plan is still optimal in period 1[15]:

$$u'(c_2)e^{-(\hat{\delta}-\delta_1)} = u'(c_2)e^{-2\delta_2} \tag{3.69}$$

$$\implies \hat{\delta} = 2\delta_2 - \delta_1 \tag{3.70}$$

As we can see, the plans made in period 0 will prevail if, and only if, eq. (3.70) holds. This condition does not rule out the possibility of a declining δ_t due to date-specific events, but it rules out hyperbolic preferences. If preferences were hyperbolic, then $\delta_2 < \delta_1$ and $\hat{\delta} = \delta_1$, which is in contradiction to the condition of time consistency as defined by eq. (3.70).

Assume that all information about the present and the future is available to an agent with hyperbolic preferences. One problem is that the agent could wish to reverse an investment made in the past or to delay the initiation of a project with positive NPV. When the planned date of initiation approaches, he could wish to delay the project further into the future and so on. Such a project would never be undertaken. This problem is related to the availability of commitment instruments. A good commitment strategy would ensure that all plans made at the present time actually will be undertaken.

The other problem is more philosophical: When making optimal plans for the future, should the present or the future self count? There is at least one example in which commitment is intertemporally optimal. When a project has positive NPV, both in present and in future evaluation, and the agent still wishes to delay the project at each point of time, a commitment to undertake the project would make the agent better off than never undertaking the project.

[15]Marginal utility is an injective function of c_i. Hence, this approach will suffice for proving that (c_1^*, c_2^*) still is an optimal bundle.

3.5.3 Declining Discount Rates

A model from Gollier et al. (2008)

What follows here is a simple illustration from the review on "Declining Discount Rates" by Gollier et al. (2008), which shows the condition for time consistent discount rates. Unfortunately, it is unclear whether this model is pointed at discount rates that decline with the the time horizon or the trivial case of declining discount rates. In this section, I will try to show that Gollier et al. (2008) prove time consistency in the trivial case where discount rates are declining due to expected trends or specific events of the future.

They introduce a three-periodic deterministic model to illustrate their point. Define m as present value of total income, c_i as consumption period i, and $u(c_i)$ as utility of consumption. Utility is, as usual, monotonically increasing in consumption and strictly concave. The utility discount rate is δ_i, and the one-periodic discount rate for economic goods is r_i, for period i (r_2 is then the long rate), $\hat{r}$ is the second period‘s short rate. The optimization program looks like this:

$$max_{c_0,c_1,c_2} L_0 = u(c_0)+e^{-\delta_1}u(c_1)+e^{-2\delta_2}u(c_2)+\lambda_0\left(m-c_0-e^{-r_1}c_1-e^{-r_2}c_2\right) \quad (3.71)$$

By Lagrange, the first-order condition is:

$$\lambda_0 = u'(c_0) = e^{r_1-\delta_1}u'(c_1) = e^{2(r_2-\delta_2)}u'(c_2) \quad (3.72)$$

We assume that the second-order condition is satisfied. The first-order condition is then sufficient and will give an optimal bundle of consumption (c_0^*, c_1^*, c_2^*). Let us move to the next period‘s optimization program. The initial income minus the first period consumption has been transfered to period 2, at the first period rate of return, δ_1

$$max_{c_1,c_2} L_1 = u(c_1) + e^{-\delta_1}u(c_2) + \lambda_1\left(e^{-r_1}(m-c_0^*) - c_1 - e^{-\hat{r}}c_2\right) \quad (3.73)$$

By Lagrange, the first-order condition is:

$$\lambda_1 = u'(c_1) = e^{\hat{r}-\delta_1}u'(c_2) \quad (3.74)$$

Gollier et al. (2008) state that in order to avoid arbitrage, we must define $\hat{r} = 2r_2 - r_1$. This is a crucial statement, so let us stop at this stage for a moment. It is easier to see this if we compare the two strategies: 1) investing in a 2 years project at the long rate, or 2) investing in a 1 year project at the short rate, and then reinvesting the gross return at the next period‘s 1 year short rate. That is,

$$e^{2r_2} = e^{r_1}e^{\hat{r}} \tag{3.75}$$

Taking logarithms and solving for $\hat{r}$ gives

$$\hat{r} = 2r_2 - r_1 \tag{3.76}$$

I will return to this condition later. Now, we check that the bundle (c_1^*, c_2^*) is still optimal at time $t = 1$. Substituting for $u'(c_1)$ in eq. (3.72) and using eq. (3.76), we end up with the following equation[16]:

$$2(r_2 - \delta_1) = 2(r_2 - \delta_2) \tag{3.77}$$

As long as $\delta_1 = \delta_2$ holds (non-hyperbolic preferences), eq. (3.77) will hold in general.

This requires further investigation. If the no-arbitrage condition (eq. (3.76)) holds, and discount rates are declining, then $\hat{r} < r_2 < r_1$[17]. This condition however, implies that the interest rate is declining due to an expected or deterministic trend. That is, the declining pattern is attached to dates, and not the length of the time interval[18]. To see this, look at eq. (3.76), and have in mind that our previous models of declining discount rates entail a flat structure of the expectation of future short rates; it is increasing risk alone that drives the declining pattern. Eq. (3.76) holds in period 0, but we will seen from period 1, expect that the short rate of period 2 ($\hat{r}$) will equal the short rate of period 1 (r_1). If we observe that $\hat{r} = r_1$, we would want to change the plans that were made in period 0. And if we change our plans it

[16]As long as utility is monotonically increasing, marginal utility is a one-to-one (injective) function of c_i. Hence, this approach will suffice for proving that (c_1^*, c_2^*) still is an optimal bundle.

[17]See appendix C.2.

[18]If δ_ts here for some reason were declining due to date-specific events, which is fully compatible with non-hyperbolic preferences, we would still obtain the result of time consistency. This is in the line with the statements of Rasmusen (2008). I have actually shown this in section 3.5.2.

will be due to new information of the future risk (shorter time horizon means lower risk).

It follows from the model that the condition for time consistency operates in a deterministic world. Gollier et al. (2008) point out that the result holds for each state when operating in a risky world, but this is not explicitly analyzed. In our models of declining discount rates, the discount rate satisfies the time consistency condition of Gollier et al. (2008) for each state.

Intuition for Time Consistency

I have been stressing the importance of positive correlation between risk and time horizon when rationalizing declining discount rates. In expectation, the one-periodic discount rate will tend to repeat itself, as Table 3 shows. This is not saying that the models predict time inconsistency, but rather that the passage of time gives us new information - it changes the risk. When period t becomes period $t-1$, the risk of that date is not as risky as it used to be. A simple illustration follows in Table 3:

Year of payoff	2040	2070	2100	2130
Evaluation year				
2011	r = 3,5 %	r = 3 %	r = 2,5 %	r = 2 %
2040	r = 0 %	r = 3,5 %	r = 3 %	r = 2,5 %
2070	- %	r = 0 %	r = 3,5 %	r = 3 %
2100	- %	- %	r = 0 %	r = 3,5 %

Table 3. Expected repetitive pattern of declining discount rates [19]

This pattern is not deterministic, but it is how we expect the pattern to evolve seen from year 2011. The pattern is repetitive simply because risk correlates with length of time. In the case of hyperbolic preferences, this pattern would be deterministic and lead to time inconsistency.

One main point is that I am not convinced that the particular model presented by Gollier et al. (2008), proves the non-existence of time inconsistency when discount rates decline due to increasing risk in time. I believe such a proof would be helpful for a clear understanding of why declining discount rates are time consistent, in

[19]The numerical discount rates are examples of my own imagination.

spite of the agent‘s tendency to regret initial plans in expectation. Note that such a concept of regret is equally irrelevant for optimal policies as regretting an ex-ante optimal insurance contract because the good state was materialized[20].

In order to see clearly that declining discount rates is fully compatible with time consistent behavior, the analogy to expected utility theory is useful. But in such a framework, the agent will of course expect overinvestment ex-post. This cost is, however, internalized when decisions are made. It is the optimal level of "insurance premium". Expected utility is maximized, rather than expected consumption. To see this, consider the Weitzman-Gollier approach:

$$R^W(t) = -\frac{1}{t} ln \left(\sum_{s=1}^{n} q_s^W e^{-r_s t} \right) \tag{3.78}$$

Looking farther into the future from the point of time 0^* when decision is made and the true r is yet to be revealed, the expectation of r_t does not change, but our *valuation* changes with the time interval because of the increasing per-periodic risk. Thus, our expectation at date 0^* remains

$$E_{0^*}[r_t] = \bar{r}, \forall t \tag{3.79}$$

$$R^W(t) < E_{0^*}[r_t], \forall t > 0 \tag{3.80}$$

3.6 Summary of Declining Discount Rates

3.6.1 Propositions

Based on section 3.2 and 3.3 we are now able to list the conditions for obtaining declining discount rates based on external factors of risk. For illustrative purposes we assume that δ is constant, that the expected change in growth $E[\Delta g]$ is zero, and that the representative agent has power utility. I will now *based on this thesis‘ presentation* argue that these are the conditions:

1. Assumptions on preferences: $u'(c) >, u''(c) < 0, u'''(c) > 0$

[20] This line of reasoning is equivalent to the expected utility theory‘s assumption that the agent has no orientation of regret.

2. Assumptions on the stochastic process: a purely persistent shock stochastic process

3. Assumptions on our degree of knowledge: parametric uncertainty

Proposition 3.1

Condition 1 is a necessary condition to obtain declining discount rates.
I showed in section 3.2.1 that this class of preferences is necessary to induce a motive to increase investments when the future becomes more risky. Therefore, when per-periodic risk is increasing with time, prudence is necessary to get a relatively stronger motive to save between two dates that are more distant.

Proposition 3.2

The assumption of an i.i.d. stochastic process for log consumption or a log-normal distribution for consumption is sufficient to ensure a constant long-term discount rate.
As we have seen, this type of risk reduces the discount rate equally for each period. Prudence is necessary to get a lower discount rate when risk is introduced, but it does not imply a declining pattern under these assumptions of risk.

Proposition 3.3

Condition 1 and 2 combined are sufficient to ensure declining discount rates.
A non mean reverting persistent shock stochastic process in combination with Condition 1 is sufficient to provide declining discount rates, as both the Weitzman-Gollier approach of section 3.3 and the model of persistent shock stochastic process in section 3.2.3 showed us.

Proposition 3.4

Condition 1 and 3 combined are sufficient to ensure declining discount rates.
We have seen in section 3.2.4, that prudence combined with parametric uncertainty yield declining discount rates. If we add the fact that a mean reversion process with

parametric uncertainty is itself a non mean reversion process, ((Weitzman, 1998) and (Gollier and Weitzman, 2010)), it makes a strong case for declining discount rates, as parametric uncertainty is quite realistic and perhaps more so compared to a persistent shock stochastic processes[21]. The Weitzman-Gollier approach relies heavily on a stochastic process with persistent shocks, and this process to be non mean reverting. The model of a mean reversion process of Gollier (unpublished manuscript) exhibits a form of declining discount rates, except for the fact that the long-term discount rate does not approach the lowest possible interest rate or zero; it is lower bounded. Relying on a mean reversion process does not ensure strictly declining discount rates. But in combination with parametric uncertainty the Weitzman-Gollier approach still holds.

It seems fair to conclude that parametric uncertainty is the strongest case for theoretically defending declining discount rates. Firstly, these assumptions also ensure any persistent shock stochastic process to be non mean reverting, and the Weitzman-Gollier approach holds in addition to the model of parametric uncertainty. Secondly, it is likely that parametric uncertainty is a perfectly realistic assumption, at least for the longer time horizon. It is hard to believe that people 200 years ago were able to predict the magnitude of economic growth up to this point of time. Today, our knowledge and understanding of economic growth are much more sophisticated, of course, and we are equipped with computing power to analyze incredibly large sets of data. But still, economists are having a hard time predicting even business cycles and the longer the time horizon, the less do we know about the parameters governing the stochastic process of our economy.

3.6.2 Discussion of Timing Issues

It is consensus that a declining pure rate of time preference (δ) leads to time inconsistent behavior. Gollier et al. (2008) attempts to prove formally that time inconsistency is not a problem when the rationale for declining discount rates is

[21] The existence of a persistent shock stochastic process is to a large extent an empirical question. See Gollier et al. (2008) and Newell and Pizer (2003).

external factors of risk. This proof operates in a deterministic world and is based on declining discount rates due to date-specific events. It seems as if the authors behind the proof slightly miss the target.

It is not easy to get a solid grip on the issues of time inconsistency when socially efficient discount rates are declining. Most literature on time inconsistency are concerned with hyperbolic preferences, and not so much with discounting of goods. Along with the recent literature, there seems to be some confusion about the different implications of the different rationales. In the foreword of the book "Discounting and Intergenerational Equity" (Portney and Weyant, 1999), Robert M. Solow writes, (referring to both Weitzman's argument and to the traditional hyperbolic preferences argument), that: "But this suggestion [Declining Discount Rates] solves one problem by creating another. Unless the discount rate is constant, the policy path is subject to 'time inconsistency'"

Additionally, in the commentary section of Gollier et al. (2008), one of the issues raised is the need for exploring problems of timing issues when discount rates are declining.

Gollier et al. (2008) is in stark contrast with the above statements, and is supported by others. Hyperbolic preferences has received greater attention on social level discounting, and recently within the context of global climate change mitigation and intergenerational games [22] Winkler (2009) is perfectly clear on the issue of time consistency: "In fact, if declining discount rates stem from uncertainty over future states of the world there is no issue of time-inconsistency if plans are updated as soon as better information is available."

When the motivation to invest for the future is risk or uncertainty that is reduced when approaching the future, the degree of reversibility becomes important. If investments were fully reversible or we were freely able to delay projects, there would be little need to invest for the long-distant future. We would not even be able to bind capital for the future.

On the other hand, if projects are irreversible or the benefits of a project are

[22] Alongside with Karp (2005), Winkler (2006, 2009) has continued the tradition of Phelps and Pollak (1968) with applications in climate change mitigation.

time-lagged or endogenous in the delay time[23], there might be a potential welfare gain in planning for the long-distant future. These considerations apply whether we are using a flat or declining term structure of the discount rates of course. But the theory in this thesis starts becoming really interesting when considering the latter category of projects.

[23]Such as climate change mitigation projects.

Part 4

Conclusion

What follows in section 4.1 is a list of the main insights that have been drawn from the investigations of this thesis, complemented by a few summarizing remarks in section 4.2. The last section is not part of the conclusion in a strict sense, but is intended to point out two important interfaces to the topics I have focused on, and which could become part of further investigations in the future.

4.1 List of Main Conclusions

1. There is no general consensus that the short-term socially efficient discount rate is equal to the long-term socially efficient discount rate.

2. Quite the contrary: There is good theoretical evidence that external factors to the agent tend to give a lower discount rate for the distant future.

3. In an optimality-seeking framework: Reasonable assumptions on preferences combined with parametric uncertainty or a purely persistent shock stochastic process driving growth, are sufficient to rationalize discount rates that decline with the time horizon.

4. Declining discount rates do not imply time inconsistency. This holds as long as plans are updated to new information or changes in the external environment, and not merely because time is passing.

5. If we for some reason rationalize declining discount rates on the basis of hyperbolic preferences (one way of reflecting relatively greater care for future generations), time inconsistency becomes a problem. The research within commitment strategies and intergenerational games seems important under the assumption of hyperbolic preferences.

6. The theory of declining discount rates do not give answers to ethical questions, questions regarding our care for future generations or questions about the numerical values of the discount rate. Further research on this part is important in order to operationalize the theory, especially empirical research on interest rates and the stochastic process of the long-term economy.

4.2 Summarizing Remarks

In this thesis, we have seen how important the legacy of Frank Ramsey is when analyzing long-term discounting. In the shorter term, the opportunity cost of capital is an appropriate discount rate as long as it is greater than or equals the social rate of time preference. In the longer term however, the answer is far less obvious. The simple intuition of the s.s. Ramsey equation is a good organizing concept in which to think about intergenerational equity in an optimality-seeking framework. As discussed in part 2, in the long run, we are not so sure of which utility function to use, and in the case of power utility, which values for θ that is reasonable. Furthermore, there is no deep consensus regarding the value of δ. Moreover, there is the debate of to what extent the intertemporal framework inherited from Ramsey (1928) is appropriate at all in the long run. After all it assumes that we care for our descendants in a quite specific way. There is always a matter of subjective opinions and ethical aspects when determining the present‘s preferences for the future as well as future consumers‘ preferences. Such information can hardly be revealed or determined properly.

As we developed the framework of Ramsey to account for risk or uncertainty, we saw that plausible assumptions on the stochastic process and preferences rationalize declining discount rates. The distinction between this rationale for a declining pattern and a declining pure rate of time preference δ_t is important when it comes to dynamical consistency. While it is clear that a declining δ_t will result in time inconsistency, it has been weakly proven that declining discount rates are time-consistent when the rationale builds on external factors of risk (Gollier et al., 2008), and I have made an attempt to slightly complement on this.

The total theoretical evidence that the socially efficient discount rate declines with time is quite overwhelming. For practical purposes however, there is a great challenge to estimating numerical values. We need to estimate the specific parameters of the stochastic process or estimate subjective probabilities in the case of parametric uncertainty. The theory presented in this thesis focuses on the shape of the term structure and do not imply any specific level of the discount rate.

When evaluating long-term projects, it seems difficult to reach consensus or some solid ground for analysis. Still, when considering the alternative which is giving up on the problem at the look of it, the ongoing research (both theoretical and empirical) brings us forward into a less uninformed future. And for the comfort of it: If we judge by the debates on climate change in society in general, there is some evidence that people are interested in questions about the distant future and long-term investments.

4.3 Interesting Neighboring Topics

This is a small section in which I briefly comment on two (in my opinion) interesting neighboring topics, which I have chosen not to include in this thesis in spite of their special relevance to the theory of long-term socially efficient discounting.

Consumption Smoothing or Insurance Policy?

When looking far into the future the projects evaluated seem to be of greater impact, while short-term policies seem to be of a marginal character. It is not controversial to propose that willingness to consider a project that is profitable in a longer time horizon, tends to increase with the economic significance of the problem. The models considered in this thesis are mostly appropriate for evaluating projects not aiming to overcome threats of disasters. The qualitative insight though, that greater risk of macroeconomic conditions increases the valuation of projects, still holds. This insurance perspective is fundamental in financial theory. If per-periodic risk increases with time, then we will tend to attach greater economic value on long-term insurance policies compared to an i.i.d. process. Using a smooth discount rate function might be inappropriate for long-term projects seeking to avoid the consequences of a disaster, even though the qualitative insight of our models still works in the same direction. When for instance evaluating global climate change mitigations, a model considering catastrophic events could prove a better framework for CBA. In Stern and Treasury (2007), Stern used the threat of a disaster as an argument for using a low discount rate. Qualitatively it makes sense that we should invest more to avoid catastrophes. But the method of simply lowering the constant discount rate is hardly a quantitatively reasonable approach when considering catastrophes.

Intrinsic Project Risks

A discussion of how intrinsic project risks should be treated in a long-term perspective has not been within the scope of this thesis. However, the work of the different approaches could potentially be combined, as illustrated in Weitzman (2007). The argument relies on the public economist‘s version of Weighted Average Cost of Capital (WACC). Combining the premise that discount factors should be averaged when there is uncertainty about discount rates, and a framework to incorporate the intrinsic risks of projects into the discount rate function, he obtains a similar result as in Weitzman (1998). In a CAPM-model of the economy, the expected return of a project r^e is higher than the risk-free rate r^f because of the risk premium of systematic risk. As Weitzman points out, the investment βs are correlation coefficients applied to discount factors and not discount rates. Define

$$\beta = corr(R^e, P) \tag{4.1}$$

, where R^e is the gross return of the economy as a whole, and P is the gross return on the project to be evaluated. Then an appropriate discount rate is:

$$r(t) = -\frac{1}{t} ln \left(\beta e^{-r^e t} + (1-\beta) e^{-r^f t} \right) \tag{4.2}$$

This has similar properties to eq. (3.37) concerning risk of future interest rates:

$$r(0) = \beta r^e + (1-\beta) r^f \tag{4.3}$$

$$\frac{dr(t)}{dt} < 0 \tag{4.4}$$

$$r(\infty) = r^f \tag{4.5}$$

This model does *not* tell us that incorporating risk of projects reduces the socially efficient discount rate compared to safe projects.

In the very long run there is much uncertainty about the discount rate, as the reader by now presumably will agree on. There are however, many other aspects of long run policy evaluation that seem very challenging with respect to valuation, such as intrinsic project risks. How to make appropriate expectations of a certain project‘s consumption-equivalent more than hundred years into the future seems at least as difficult as finding appropriate discount rates. There are most certainly large uncertainties with respect to the numerator in the NPV-function in the long run. If there are risks or great uncertainties in the project, this will tend to reduce the certainty equivalent benefits and increase the certainty-equivalent costs. Both effects work in the direction of reducing the profitability of the project. If the uncertainties of the project tends to increase with time, the result of declining discount rates doesn‘t necessarily imply that a vast number of long-term projects become profitable.

Bibliography

Arrow, K., W. Cline, K. Maler, M. Munasinghe, R. Squitieri, and J. Stiglitz (1996). Intertemporal equity, discounting, and economic efficiency. *Climate Change 1995: economic and social dimensions of climate change*, 125–44.

Arrow, K., P. Dasgupta, L. Goulder, G. Daily, P. Ehrlich, G. Heal, S. Levin, K. M "aler, S. Schneider, D. Starrett, et al. (2004). Are we consuming too much? *Journal of Economic Perspectives 18*(3), 147–172.

Barro, R. and X. Sala-i Martin (2004). Economic Growth, 2nd-ed.

Beckerman, W. and C. Hepburn (2007). Ethics of the discount rate in the Stern Review on the economics of climate change. *World Economics 8*(1), 187–210.

Broome, J. (1994). Discounting the future. *Philosophy & Public Affairs 23*(2), 128–156.

Buchholz, W. and J. Schumacher (2008). Discounting the Long-Distant Future: A Simple Explanation for the Weitzman-Gollier-Puzzle. *CESifo Working Paper No. 2357*.

Chichilnisky, G. (1996). An axiomatic approach to sustainable development. *Social Choice and welfare 13*(2), 231–257.

Commonwealth of Australia (2006, January). Handbook of cost benefit analysis. `http://www.finance.gov.au/publications/finance-circulars/2006/docs/Handbook_of_CB_analysis.pdf`.

Dasgupta, P. (2005). Three conceptions of intergenerational justice. *Ramsey's Legacy*, 149–169.

Dasgupta, P. (2007). The Stern Review's economics of climate change. *National Institute Economic Review 199*(1), 4.

Dixit, A. (1990). *Optimization in economic theory.* Oxford University Press, USA.

Encyclopedia Britannica. `http://www.britannica.com/EBchecked/topic/133572/consol`.

Finansdepartementet (2005, September). Veileder i samfunnsøkonomiske analyser. `http://www.regjeringen.no/upload/FIN/Vedlegg/okstyring/Veileder_i_samfunnsokonomiske_analyser.pdf`.

Frederick, S., G. Loewenstein, and T. O'donoghue (2002). Time discounting and time preference: A critical review. *Journal of Economic Literature 40*(2), 351–401.

Gollier, C. (2002). Time horizon and the discount rate. *Journal of Economic Theory 107*(2), 463–473.

Gollier, C. (2004). Maximizing the expected net future value as an alternative strategy to gamma discounting* 1. *Finance Research Letters 1*(2), 85–89.

Gollier, C. (2006). An Evaluation of Stern's Report on the Economics of Climate Change. *IDEI Working*.

Gollier, C. (2007). The consumption-based determinants of the term structure of discount rates. *Mathematics and Financial Economics 1*(2), 81–101.

Gollier, C., P. Koundouri, and T. Pantelidis (2008). Declining discount rates: Economic justifications and implications for long-run policy. *Economic Policy 23*(56), 757–795.

Gollier, C. and M. Weitzman (2010). How should the distant future be discounted when discount rates are uncertain? *Economics Letters*.

Gollier C. (unpublished manuscript). Pricing the future: The economics of discounting and sustainable development. *unpublished*.

Hepburn, C. and B. Groom (2007). Gamma discounting and expected net future value. *Journal of Environmental Economics and Management 53*(1), 99–109.

HM Treasury (2003). The Green Book: Appraisal and Evaluation in Central Government. http://www.hm-treasury.gov.uk/d/green_book_complete.pdf.

Investopedia. http://www.investopedia.com/ask/answers/06/100yearbond.asp.

Karp, L. (2005). Global warming and hyperbolic discounting. *Journal of Public Economics 89*(2-3), 261–282.

Laibson, D. (1997). Golden Eggs and Hyperbolic Discounting*. *Quarterly Journal of Economics 112*(2), 443–477.

Laibson, D. (1998). Life-cycle consumption and hyperbolic discount functions. *European Economic Review 42*(3-5), 861–871.

Li, C. and K. Löfgren (2000). Renewable Resources and Economic Sustainability: A Dynamic Analysis with Heterogeneous Time Preferences* 1. *Journal of Environmental Economics and Management 40*(3), 236–250.

Newell, R. and W. Pizer (2003). Discounting the distant future: how much do uncertain rates increase valuations? *Journal of Environmental Economics and Management 46*(1), 52–71.

Nordhaus, W. (2007). A Review of the" Stern Review on the Economics of Climate Change". *Journal of Economic Literature 45*(3), 686–702.

Office of Management and Budget (1992, October). Circular No. A-94 Revised:Guidelines and Discount Rates for Benefit-Cost Analysis of Federal Programs. http://www.whitehouse.gov/omb/circulars_a094.

OXERA, A. (2002). Social Time Preference Rate for Use in Long-term Discounting.

Pearce, D., B. Groom, C. Hepburn, and P. Koundouri (2003). Valuing the future. *World Economics 4*(2), 121–141.

Phelps, E. and R. Pollak (1968). On second-best national saving and game-equilibrium growth. *The Review of Economic Studies*, 185–199.

Portney, P. and J. Weyant (1999). *Discounting and intergenerational equity.* Resources for the Future.

Ramsey, F. (1928). A mathematical theory of saving. *The Economic Journal 38*(152), 543–559.

Rapport Lebégue (2005, January). RÉvision du taux d'actualisation des investissements publics. `http://www.plan.gouv.fr/intranet/upload/actualite/Rapport%20Lebegue%20Taux%20actualisation%2024-01-05.pdf`.

Rasmusen, E. (2008). Some Common Confusions about Hyperbolic Discounting. *Working Papers*.

Schelling, T. (1995). Intergenerational discounting. *Energy policy 23*(4-5), 395–401.

Sen, A. (1967). Isolation, assurance and the social rate of discount. *The Quarterly Journal of Economics 81*(1), 112–124.

Sen, A., A. Sen, B. Williams, and B. Williams (1982). *Utilitarianism and beyond.* Cambridge Univ Pr.

Stern, N. and G. Treasury (2007). *The economics of climate change: the Stern review.* Cambridge Univ Pr.

Strotz, R. (1955). Myopia and inconsistency in dynamic utility maximization. *The Review of Economic Studies 23*(3), 165–180.

Weitzman, M. (1998). Why the far-distant future should be discounted at its lowest possible rate. *Journal of environmental economics and management 36*(3), 201–208.

Weitzman, M. (2001). Gamma discounting. *American Economic Review 91*(1), 260–271.

Weitzman, M. (2007). A review of the Stern Review on the economics of climate change. *Journal of Economic Literature 45*(3), 703–724.

Winkler, R. (2006). Does [] better'discounting lead to [] worse'outcomes in long-run decisions? The dilemma of hyperbolic discounting. *Ecological Economics 57*(4), 573–582.

Winkler, R. (2009). Now or never: environmental protection under hyperbolic discounting. *Economics: The Open-Access, Open-Assessment E-Journal 3*.

World Commission on Environment and Development (1987). Our common future. `http://www.un-documents.net/ocf-02.htm`.

Appendix A

Deriving the Ramsey Model

A.1 The Consumers' Problem

First, we want to analyze the problem using per capita effective labor units variables denoted by small letter and a cap. The steady state growth of consumption and capital will be $\dot{\hat{c}} = \dot{\hat{k}} = 0$. Define:

$$\frac{dA(t)}{dt} = rA(t) + wL(t) - C(t) \tag{A.1}$$

$$\dot{a} = \frac{d\frac{A(t)}{L(t)}}{dt} \tag{A.2}$$

Then, using the chain rule

$$\dot{a} = \frac{dA(t)}{dt}\frac{1}{L(t)} - A(t)(L(t))^{-2}\frac{dL(t)}{dt} \tag{A.3}$$

$$= \frac{dA(t)}{dt}\frac{1}{L} - na \tag{A.4}$$

$$\tag{A.5}$$

By (A.1),

$$\dot{a} = (rA(t) + wL(t) - C(t))\frac{1}{L(t)} - na \tag{A.6}$$

$$\implies \dot{a} = w(t) - c(t) - a(t)\big(n - r(t)\big) \tag{A.7}$$

Define lifetime utility:

$$U = \int_0^\infty u[c(t)]e^{-(\delta-n)}dt \tag{A.8}$$

Define $v(t)$ = "present value shadow price of income". Then we can set up the Hamiltonian, which defines, for each t, the present value optimization problem.

$$H = u[c(t)]e^{-(\delta-n)t} + v(t)\big(w(t) - c(t) - a(t)\big(n - r(t)\big)\big) \tag{A.9}$$

The first term is the instantaneous present value of period t consumption, while the last term in the brackets is the law of income flow multiplied with present value shadow price; the marginal increase in present value of utility of a marginal increase of income.

$$\left.\begin{aligned} \frac{\partial H}{\partial c} &= 0 \\ -\frac{\partial H}{\partial a} &= \dot{v} \end{aligned}\right\} \qquad \text{FOC} \tag{A.10}$$

The transversality condition:

$$\lim_{t\to\infty} \left(v(t) \cdot a(t)\right) = 0 \tag{A.11}$$

In this case, the problem in (A.9) is stated in order of the transversality condition not to be violated for reasonable parameter values. From (A.9) and (A.10):

$$v = u'(c)e^{-(\delta-n)t} \tag{A.12}$$

$$\dot{v} = -(r-n)v \tag{A.13}$$

Combining these two first-order conditions gives us the Euler equation:

$$\ln(v) = \ln(u'(c)) - (\delta - n)t \tag{A.14}$$

$$\frac{\dot{v}}{v} = \frac{d}{dt}\left(\ln(v)\right) \tag{A.15}$$

$$= \frac{u''(c)}{u'(c)}\dot{c}\frac{c}{c} - (\delta - n) \tag{A.16}$$

From (A.13),

$$\frac{\dot{v}}{v} = -(r-n) \tag{A.17}$$

$$\implies -(r-n) = \left(\frac{u''(c)}{u'(c)}c\right)\frac{\dot{c}}{c} - (\delta - n) \tag{A.18}$$

Solving for $\frac{\dot{c}}{c}$ gives the Euler equation:

$$\frac{\dot{c}}{c} = (\delta - r)\frac{1}{c}\left(\frac{u'(c}{u''(c)}\right) \tag{A.19}$$

Rearranging gives the famous Ramsey equation:

$$r = \delta - \left(c\frac{u''(c)}{u'} \right) \frac{\dot{c}}{c} \tag{A.20}$$

It is easy to show that in the case of power utility, $u(c) = \left(\frac{c^{1-\theta}-1}{1-\theta}, \quad \theta > 1 \right)$ we will get

$$\text{Euler equation: } \frac{\dot{c}}{c} = \frac{1}{\theta}(r-\delta) \tag{A.21}$$

$$\text{Ramsey equation: } r = \delta + \theta\frac{\dot{c}}{c} \tag{A.22}$$

A.2 The Firms‘ Problem

To ensure that $\dot{\hat{k}} = 0$ in steady state, we assume that technology is labor-augmenting only. With technology function $G(t) = e^{gt}$, we get

$$Y(t) = F[K(t), L(t) \cdot G(t)] = F[K, \hat{L}] \quad \text{Divide by } \hat{L} \tag{A.23}$$

$$\hat{y} = F[\frac{K}{\hat{L}}, 1] = f(\hat{k}) \quad \text{made possible by CRS} \tag{A.24}$$

$$\left(\hat{y} \overset{\text{def}}{=} \frac{Y}{\hat{L}} \quad \hat{k} \overset{\text{def}}{=} \frac{K}{\hat{L}} \right)$$

In order to maximize profits we need to differentiate w.r.t. K and L. Define

$$Y = \hat{L} \cdot f(\hat{k}) \quad \text{and differentiate} \tag{A.25}$$

$$\frac{\partial Y}{\partial K} = \hat{L} f'(\hat{k}) \frac{1}{\hat{L}} = f'(k) \tag{A.26}$$

$$\frac{\partial Y}{\partial L} = \frac{\partial Y}{\partial \hat{L}} \cdot \frac{d\hat{L}}{dL} \tag{A.27}$$

$$= \frac{d\hat{L}}{dL} \left(f(\hat{k}) - \hat{L} f'(\hat{k}) K L^{-2} \right) \tag{A.28}$$

$$\Longrightarrow \quad \frac{\partial Y}{\partial L} = e^{gt} \left(f(\hat{k}) - f'(\hat{k})\hat{k} \right) \tag{A.29}$$

The firms optimize profit in each period, and markets is assumed to clear. The profit function:

$$\pi = F[K, L] - rK - wL \tag{A.30}$$

Combining (A.30) with (A.26) and (A.29), we get the following first-order conditions:

$$\left.\begin{aligned}\frac{\partial \pi}{\partial K} &= 0 \\ f'(\hat{k}) &= r\end{aligned}\right\} \qquad \text{FOC capital} \tag{A.31}$$

$$\left.\begin{aligned}\frac{\partial \pi}{\partial L} &= 0 \\ e^{gt}\left(f(\hat{k}) - f'(\hat{k})\hat{k}\right) &= w\end{aligned}\right\} \qquad \text{FOC labor} \tag{A.32}$$

A.3 Equilibrium

We assume that the economy is closed, therefore we have $a = k, \quad \forall t$. Using (A.7), and substituting a for k we have that

$$\dot{k} = w - c - k(n - r) \tag{A.33}$$

By definition $\hat{k} = ke^{-gt}$. Differentiating gives w.r.t t gives $e^{gt}\left(\dot{\hat{k}} + \hat{k}g\right)$ If we use (A.31) and (A.32) and substitute for r and w we get

$$\dot{k} = e^{gt}\left(\dot{\hat{k}} + \hat{k}g\right) = e^{gt}\left(f(\hat{k}) - \hat{k}f'(\hat{k})\right) - c - k(n - f'(\hat{k})) \tag{A.34}$$

$$\dot{\hat{k}} + \hat{k}g = \left(f(\hat{k}) - \hat{k}f'(\hat{k})\right) - \hat{c} - \hat{k}(n - f'(\hat{k})) \tag{A.35}$$

$$\dot{\hat{k}} = f(\hat{k}) - \hat{c} - \hat{k}(n + g) \tag{A.36}$$

Now all variables are in the form of per effective capita units. Same procedure for consumption, using (A.21) and (A.31) and the definition $\hat{c} = ce^{-gt}$. Differentiating w.r.t. t gives $\dot{\hat{c}} = e^{-gt}\left(\dot{c} - gc\right)$. Dividing by $\hat{c}$ gives

$$\frac{\dot{\hat{c}}}{\hat{c}} = \frac{\dot{c}}{c} - g \tag{A.37}$$

$$\frac{\dot{\hat{c}}}{\hat{c}} = \frac{1}{\theta}\left(f'(\hat{k}) - \delta - \theta g\right) \tag{A.38}$$

The last equation is derived using (A.36) and (A.21). Equation (A.36) and (A.38) determines equilibrium path of capital and consumption, when both firms and consumers optimize. We also need the transversality condition and initial value of capital to say something about the viable optimal path.

A.4 The Steady State

For the purpose of this thesis I will simply state by intuition that the transversality condition rules out the possibility of $r < g+n$. If this does not hold, the representative consumer would be able to finance increasing debt through technology process and population growth only.

We have also simply stated that $\dot{\hat{c}} = \dot{\hat{k}} = 0$ is the only equilibrium, even though it can be proven formally. Intuitively, if $\hat{c}$ were to grow in steady state, the growth couldn't come from technology or population growth, so $\hat{k}$ must have been negative and this is ruled out by the transversality condition. If $\hat{k} > 0$, the same argument applies. Optimizing agents and the transversality condition give us $\dot{\hat{c}} = \dot{\hat{k}} = 0$ in steady state. Using this in the model gives us steady state. From (A.36) and (A.38)

$$\dot{\hat{k}} = 0 \implies \hat{c} = f(\hat{k}) - (g+n)\hat{k} \tag{A.39}$$

$$\frac{\dot{\hat{c}}}{\hat{c}} = 0 \implies f'(\hat{k}) = \delta + \theta g \tag{A.40}$$

This equation (A.40) is the *steady state Ramsey equation*, famously used in debates on long-term discount rates. From this equation (given Inada conditions) we can find a value for $\hat{k}$ and use this value to find a unique value for $\hat{c}$.

At last, from the transversality condition, $f'(\hat{k}) > g + n$, we see from (A.39) and (A.40), that

$$\delta + \theta g > g + n \tag{A.41}$$

$$\delta > n + (1-\theta)g \tag{A.42}$$

The transversality condition rules out the possibility that δ is zero for positive (and often realistic) values for n, θ and g.

Appendix B

Models of Section 3.2

B.1 When Log-consumption is Normally Distributed

$$r = \delta - \frac{1}{t} ln \left(\frac{E[u'(c_t)]}{u'(c_0)} \right) \tag{B.1}$$

Define x as the increase in the logarithm of consumption.

$$x \sim N(\mu, \sigma^2) \tag{B.2}$$

$$c_1 = c_0 e^x \tag{B.3}$$

Using the mathematical formula of log-normal distribution:

$$E[e^x] = e^{\mu + 0.5\sigma^2} \tag{B.4}$$

Then,

$$E[c_1] = c_0 e^{\mu + \sigma^2} \tag{B.5}$$

$$ln(E[c_1]) = ln(c_0) + \mu + 0.5\sigma^2 \tag{B.6}$$

Then $E[g]$, the expected growth rate of consumption can be written:

$$E[g] = ln(\frac{E[c_1]}{c_0}) = ln(E[c_1]) - ln(c_0) = \mu + 0.5\sigma^2 \tag{B.7}$$

Power utility is assumed, so $u'(c) = c^{-\theta}$. Using eq. (B.3)

$$E[u'(c_1)] = E[c_1^{-\theta}] \tag{B.8}$$

$$= (c_0 e^x)^{-\theta} \tag{B.9}$$

$$= c_0^{-\theta} e^{-\theta x} \tag{B.10}$$

$$\Longrightarrow \frac{E[u'(c_1)]}{u'(c_0)} = E[e^{-\theta x}] \tag{B.11}$$

Using the mathematical formula for log-normal distributions:

$$E[e^{-\theta x}] = e^{-\theta(\mu - 0.5\theta\sigma^2)} \tag{B.12}$$

and equation (B.1):

$$r = \delta - ln(E[e^{-\theta x}]) \tag{B.13}$$

$$= \delta + \theta(\mu - 0.5\theta\sigma^2) \tag{B.14}$$

$$= \delta + \theta\mu - 0.5\theta^2\sigma^2 \tag{B.15}$$

$$= \delta + \theta(\mu + 0.5\sigma^2) - \theta 0.5\sigma^2 - \theta^2 0.5\sigma^2 \tag{B.16}$$

$$r = \delta + \theta g - \frac{1}{2}\theta(\theta + 1)\sigma^2 \tag{B.17}$$

B.2 The General Result of an i.i.d. Stochastic Process

Define

$$c_{t+1} = c_t e^x \tag{B.18}$$

where x is i.i.d. This property means that we can write eq. (B.1) like this:

$$r_t = \delta - \frac{1}{t} ln \left(\frac{E[u'(c_0 \prod_{\tau=0}^{t-1} e^{x_\tau})]}{u'(c_0)} \right) \tag{B.19}$$

We use the power utility function, so $u'(c) = c^{-\theta}$. Then,

$$r_t = \delta - \frac{1}{t} \sum_{\tau=0}^{t-1} ln \left(E[e^{-\theta x_\tau}] \right) \tag{B.20}$$

Note that it is the i.i.d. property that enables us to put the logarithm of the product as a sum of the logarithms of factors. Recognizing that the expectation of x_τ is the same for each τ we can rewrite eq. (B.20) as

$$r_t = \delta - ln\left(E[e^{-\theta x}]\right) \tag{B.21}$$

B.3 A Persistent Shock Stochastic Process

Define the technology of consumption:

$$c_{t+1} = c_t e^{x_t} \tag{B.22}$$

Define the stochastic variable:

$$x_t = x_{t-1} + \epsilon_t \tag{B.23}$$

$$\epsilon_t \sim N(0, \sigma^2) \tag{B.24}$$

The general efficient discount rate function:

$$r_t = \delta - \frac{1}{t} ln\left(\frac{E[u'(c_t)]}{u'(c_0)}\right) \tag{B.25}$$

Rewrite eq. (B.22)

$$ln(c_{t+1}) = ln(c_t) + x_t \tag{B.26}$$

By forward induction,

$$ln(c_1) = ln(c_0) + x_0 \tag{B.27}$$

$$= ln(c_0) + x_{-1} + \epsilon_0 \tag{B.28}$$

$$ln(c_2) = ln(c_1) + x_1 \tag{B.29}$$

$$= ln(c_0) + x_{-1} + \epsilon_0 + x_0 + \epsilon_1 \tag{B.30}$$

$$= ln(c_0) + x_{-1} + \epsilon_0 + x_{-1} + \epsilon_0 + \epsilon_1 \tag{B.31}$$

$$ln(c_3) = ln(c_0) + x_{-1} + \epsilon_0 + x_{-1} + \epsilon_0 + x_{-1} + \epsilon_0 + \epsilon_1 + \epsilon_1 + \epsilon_2 \tag{B.32}$$

We end up with:

$$ln(c_t) - ln(c_0) = tx_{-1} + t\epsilon_0 + (t-1)\epsilon_1 + (t-2)\epsilon_2 + \cdots + \epsilon_{t-1} \tag{B.33}$$

This may be written as

$$ln(c_t) - ln(c_0) = tx_{-1} + \sum_{\tau=0}^{t-1} (t-\tau)\epsilon_\tau \tag{B.34}$$

Since the ϵs are uncorrelated, the variance is:

$$Var\left(ln(c_t) - ln(c_0)\right) = \sigma^2 \sum_{\tau=0}^{t-1} (t-\tau)^2 \tag{B.35}$$

Note that $ln(c_t) - ln(c_0)$ is nothing else than g, the relative consumption growth and it is normal-distributed because the ϵs are. In order to derive the socially efficient discount rate under the specific assumptions of persistence, we use power utility and the formula for log-normal distributions in eq. (B.12):

$$r_t = \delta - \frac{1}{t} ln\left(E[e^{-\theta(ln(c_t) - ln(c_0))}]\right) \tag{B.36}$$

Using eq. (B.12):

$$r_t = \delta - \frac{1}{t} ln\left(e^{-\theta\left(tx_{-1} - 0.5\theta\sigma^2 \sum_{\tau=0}^{t-1}(t-\tau)^2\right)}\right) \tag{B.37}$$

This is straightforward rewritten as:

$$r_t = \delta + \theta x_{-1} - \frac{1}{t} 0.5\theta\sigma^2 \sum_{\tau=0}^{t-1} (t-\tau)^2 \tag{B.38}$$

We also rewrite the summation operator:

$$r_t = \delta + \theta x_{-1} - \frac{1}{t} 0.5\theta\sigma^2 \sum_{\tau=1}^{t} \tau^2 \tag{B.39}$$

We see that the socially efficient discount rate declines in time by carrying out the following calculations:

$$r_{t+1} - r_t = \delta + \theta x_{-1} - \frac{1}{t+1}0.5\theta\sigma^2 \sum_{\tau=1}^{t+1} \tau^2 - \left(\delta + \theta x_{-1} - \frac{1}{t}0.5\theta\sigma^2 \sum_{\tau=1}^{t} \tau^2\right) \quad \text{(B.40)}$$

$$= 0.5\theta\sigma^2 \left(\frac{1}{t}\sum_{\tau=1}^{t} \tau^2 - \frac{1}{t+1}\sum_{\tau=1}^{t+1} \tau^2\right) \quad \text{(B.41)}$$

$$= 0.5\theta\sigma^2 \left[\frac{1}{t}\sum_{\tau=1}^{t} \tau^2 - \frac{1}{t+1}\left(\sum_{\tau=1}^{t} \tau^2 + (t+1)^2\right)\right] \quad \text{(B.42)}$$

$$= 0.5\theta\sigma^2 \left[\sum_{\tau=1}^{t} \tau^2 \left(\frac{1}{t} - \frac{1}{t+1}\right) - (t+1)\right] \quad \text{(B.43)}$$

$$= 0.5\theta\sigma^2 \left[\sum_{\tau=1}^{t} \tau^2 \left(\frac{1}{t(t+1)}\right) - (t+1)\right] \quad \text{(B.44)}$$

$$= 0.5\theta\sigma^2 \frac{1}{t(t+1)}\left(\sum_{\tau=1}^{t} \tau^2 - t(t+1)^2\right) \quad \text{(B.45)}$$

$$= 0.5\theta\sigma^2 \frac{1}{t(t+1)}\left(\sum_{\tau=1}^{t} \tau^2 - \sum_{\tau=1}^{t}(t+1)^2\right) \quad \text{(B.46)}$$

$$= 0.5\theta\sigma^2 \frac{1}{t(t+1)}\left(\sum_{\tau=1}^{t}[\tau^2 - (t+1)^2]\right) \quad \text{(B.47)}$$

All terms in the last summation are negative, hence the discount rate is declining.

B.4 Parametric Uncertainty

The task here is to use these equations:

$$r_t = \delta - ln\sum_{\gamma=1}^{n} q_\gamma \frac{E[u'(c_t)|\gamma]}{u'(c_0)} \quad \text{(B.48)}$$

$$r_t(\gamma) = \delta - \frac{1}{t} ln\frac{E[u'(c_t)|\gamma]}{u'(c_0)} \quad \text{(B.49)}$$

to show the result of section 3.2.4. For ease of notation we define:

$$\lambda = \frac{E[u'(c_t)|\gamma]}{u'(c_0)} \quad \text{(B.50)}$$

Then we rearrange, and take exponentials of (B.48):

$$r_t = \delta - ln \sum_{\gamma=1}^{n} q_\gamma \lambda \tag{B.51}$$

$$-(r_t - \delta)t = ln \sum_{\gamma=1}^{n} q_\gamma \lambda \tag{B.52}$$

$$e^{-(r_t-\delta)t} = \sum_{\gamma=1}^{n} q_\gamma \lambda \tag{B.53}$$

We do the same for eq. (B.49):

$$\lambda = e^{-(r_t(\gamma)-\delta)t} \tag{B.54}$$

We substitute for λ:

$$e^{-(r_t-\delta)t} = \sum_{\gamma=1}^{n} q_\gamma \left(e^{-(r_t(\gamma)-\delta)t}\right) \tag{B.55}$$

Note that δ is constant and independent of γ, hence it disappears from both sides:

$$e^{-r_t t} = \sum_{\gamma=1}^{n} q_\gamma e^{r_t(\gamma)t} \tag{B.56}$$

Appendix C

Models of Section 3.3 and 3.5

C.1 Weitzman Properties 1-3

Property 1

$$R^W(0) = \lim_{t\to 0} R^W(t) = \lim_{t\to 0} -\frac{1}{t}\ln\left(\sum_{i=1}^{n} p_i e^{-r_i t}\right) \tag{C.1}$$

$$\overset{\text{l'Hopital}}{=} \lim_{t\to 0} \frac{\sum_{i=1}^{n} r_i p_i e^{-r_i t}}{\sum_{i=1}^{n} p_i e^{-r_i t}} \tag{C.2}$$

$$= \sum_{i=1}^{n} p_i r_i \tag{C.3}$$

because $\sum_{i=1}^{n} p_i = 1$.

Property 2

Let $k > 1$ (constant) be given and introduce the notation $e^{-r_i t} = x_i$. We compute the following difference

$$R^W(kt) - R^W(t) = -\frac{1}{kt}\ln(\sum_{i=1}^{n} p_i e^{-kr_i t}) + \frac{1}{t}\ln(\sum_{i=1}^{n} p_i e^{-r_i t}) \tag{C.4}$$

$$= \frac{1}{kt}\left(k\ln(\sum_{i=1}^{n} p_i x_i) - \ln(\sum_{i=1}^{n} p_i x_i^k)\right) \tag{C.5}$$

$$= \frac{1}{kt}\ln\left(\frac{(\sum_{i=1}^{n} p_i x_i)^k}{(\sum_{i=1}^{n} p_i x_i^k)}\right) \tag{C.6}$$

We now use Jensen‘s inequality to show that this expression is strictly less than zero. Then $R^W(t)$ will be a strictly decreasing function of t and since $K^W(t)$ is differentiable for all $t > 0$, this implies that $\frac{dR^W}{dt} < 0$ for $t > 0$.

Jensen‘s inequality can be stated for our purposes in the following way. If φ is a strictly convex function and $x_1, x_2, \ldots, x_n$ are numbers in its domain, then

$$\varphi\left(\sum p_i x_i\right) < \sum p_i \varphi(x_i) \tag{C.7}$$

In our case, φ takes the form $\varphi(u) = u^k$. Applying this inequality to (C.6), we find that the argument of the logarithm is less than one, so the logarithm will be negative. This gives the desired result.

Property 3

$$\begin{aligned}
R^W(\infty) = \lim_{t\to\infty} R^W(t) &= \lim_{t\to\infty} -\frac{1}{t}\ln\left(\sum_{i=1}^{n} p_i e^{-r_i t}\right) && \text{(C.8)}\\
&\overset{\text{l'Hopital}}{=} \lim_{t\to\infty} \frac{\sum_{i=1}^{n} r_i p_i e^{-r_i t}}{\sum_{i=1}^{n} p_i e^{-r_i t}} && \text{(C.9)}\\
&= \lim_{t\to\infty} \frac{\frac{\sum_{i=1}^{n} r_i p_i e^{-r_i t}}{e^{-min\{r_i\}t}}}{\frac{\sum_{i=1}^{n} p_i e^{-r_i t}}{e^{-min\{r_i\}t}}} && \text{(C.10)}\\
&= \lim_{t\to\infty} \frac{\sum_{i=1}^{n} r_i p_i e^{-(r_i - min\{r_i\})t}}{\sum_{i=1}^{n} p_i e^{-(r_i - min\{r_i\})t}} && \text{(C.11)}\\
&= \lim_{t\to\infty} \frac{min\{r_i\}p_1 + \sum_{i=2}^{n} r_i p_i e^{-(r_i - min\{r_i\})t}}{p_1 + \sum_{i=2}^{n} p_i e^{-(r_i - min\{r_i\})t}} && \text{(C.12)}\\
&= \frac{min\{r_i\}p_1 + 0}{p_1 + 0} && \text{(C.13)}\\
&= min\{r_i\} && \text{(C.14)}
\end{aligned}$$

We have assumed in (C.12), without loss of generality, that the term containing $min\{r_i\}$ has index $i = 1$.

C.2 A Model on Time Consistency

C.2.1 First-Order Conditions

I use notation as defined in section 3.5.3. Optimization in the first period:

$$max_{c_0,c_1,c_2} L_0 = u(c_0)+e^{-\delta_1}u(c_1)+e^{-2\delta_2}u(c_2)+\lambda_0\left(m-c_0-e^{-r_1}c_1-e^{-r_2}c_2\right) \tag{C.15}$$

$$\frac{\partial L}{\partial c_0} = u'(c_0)-\lambda_0 = 0 \tag{C.16}$$

$$\Longrightarrow \lambda_0 = u'(c_0) \tag{C.17}$$

$$\frac{\partial L}{\partial c_1} = e^{-\delta_1}u'(c_1)-\lambda_0 e^{-r_1} = 0 \tag{C.18}$$

$$\Longrightarrow \lambda_0 = e^{r_1-\delta_1}u'(c_1) \tag{C.19}$$

$$\frac{\partial L}{\partial c_2} = e^{-2\delta_2}u'(c_2)-\lambda_0 e^{-2r_1} = 0 \tag{C.20}$$

$$\Longrightarrow \lambda_0 = e^{2(r_2-\delta_2)}u'(c_2) \tag{C.21}$$

Then we end up with the first-order condition:

$$\lambda_0 = u'(c_0) = e^{r_1-\delta_1}u'(c_1) = e^{2(r_2-\delta_2)}u'(c_2) \tag{C.22}$$

Optimization program, second period:

$$max_{c_1,c_2} L_1 = u(c_1)+e^{-\delta_1}u(c_2)+\lambda_1\left(e^{-r_1}(m-c_1^*)-c_1-e^{-\hat{r}}c_2\right) \tag{C.23}$$

$$\frac{\partial L}{\partial c_1} = u'(c_1)-\lambda_1 = 0 \tag{C.24}$$

$$\Longrightarrow \lambda_1 = u'(c_1) \tag{C.25}$$

$$\frac{\partial L}{\partial c_2} = e^{-\delta_1}u'(c_2)-\lambda_1 e^{-\hat{r}} = 0 \tag{C.26}$$

$$\Longrightarrow \lambda_1 = e^{\hat{r}-\delta_1}u'(c_2) \tag{C.27}$$

Then we end up with the first-order condition:

$$\lambda_1 = u'(c_1) = e^{\hat{r}-\delta_1}u'(c_2) \tag{C.28}$$

C.2.2 Proof of Time Consistency

Substituting for $u'(c_1)$ in eq. (C.22) (combining eq. (C.22) and (C.28)), and using the no-arbitrage condition ($\hat{r} = 2r_2 - r_1$), we obtain:

$$e^{\hat{r}-\delta_1}u'(c_2)e^{r_1-\delta_1} = u'(c_2)e^{2(r_2-\delta_2)} \tag{C.29}$$

$$\implies e^{\hat{r}-\delta_1+r_1-\delta_1} = e^{2(r_2-\delta_2)} \tag{C.30}$$

$$\implies \hat{r} - 2\delta_1 + r_1 = 2(r_2 - \delta_2) \tag{C.31}$$

$$[2r_2 - r_1] - 2\delta_1 + r_1 = 2(r_2 - \delta_2) \tag{C.32}$$

$$\implies 2(r_2 - \delta_1) = 2(r_2 - \delta_2) \tag{C.33}$$

This equation holds if, and only if, $\delta_1 = \delta_2$.

C.2.3 Inequalities

If the discount rates are declining we know that

$$r_2 < r_1 \tag{C.34}$$

In a deterministic world, or in a stochastic world seen from time 0, no-arbitrage condition must hold

$$\hat{r} = 2r_2 - r_1 \tag{C.35}$$

$$\implies \hat{r} + r_1 = 2r_2 \tag{C.36}$$

Since, $r_1 > r_2$,

$$\hat{r} < r_2 < r_1 \tag{C.37}$$