CESifo

# Working Papers

# Investment Policy for Time-Inconsistent Discounters

## Bård Harstad

CESifo
Center for Economic Studies & Ifo Institute

# Investment Policy for Time-Inconsistent Discounters

## Abstract

This paper explores how a principal with time-inconsistent preferences invests optimally in technology or capital. If the current principal prefers her future self to save more, she can increase current investments complementary to future savings and decrease investments in the strategic substitutes, for example. To characterize the principal's choices they are compared to a market equilibrium where the investors are private agents. Each investing agent applies the same discount factors as do the principal and he obtains full property rights to the future returns. With geometric discounting, there would be no need to regulate (subsidize/tax) these agents. With time-inconsistent preferences, however, the current principal benefits from subsidizing investments in "green" capital (complementary to future savings) and tax investments in substitute capital such as "brown" technology and even adaptation technology. The paper can thus compare policies for different types of investments at the same level in the production hierarchy, but investments at different levels are also compared. With quasi-hyperbolic discounting, the optimal subsidy is unrelated to this level. With discount rates that are strictly decreasing in relative time, however, upstream investments (needed for downstream investments) will optimally be subsidized at a higher rate. When applied to environmental policy, the paper provides a new rationale for subsidizing green (and taxing brown) technology unrelated to the traditional motivation emphasizing public good aspects.

JEL-Code: D990, H430, H230, Q550.

Keywords: time inconsistency, hyperbolic discounting, commitment, investments, R&D, green technology, investment policy, environmental policy, climate change.

*Bård Harstad*
*University of Oslo*
*Oslo / Norway*
*bard.harstad@econ.uio.no*

# 1. Introduction

Many projects generate costs and benefits for future years and generations: Reducing emission today generates a cleaner environment in the future; conserving nature now makes it available for future users; extracting resources today reduces the amount available later; investments in public infrastructure creates future benefits; and research is costly today but builds knowledge we can later learn from.

When evaluating whether such projects are worthwhile, we are faced with the fundamental question of how to compare costs and benefits that occur at different points in time. This question is a deep and difficult one, and scholars have struggled with it both in the past and in recent times.

Over the last decades, economists have settled on employing "geometric" or "exponential" discounting - not because of its normative or positive justifications - but due to its elegance, tractability, and its resemblance of private investors' present-discounted value formula. Furthermore, exponential discounting leads to stationary or time-consistent preferences. Thus, assuming exponential discounting often leads to laissez-faire as the normative result, since there is no desire to distort current of future choices as a way of pre-committing to paths that seem optimal today but not later on.

However, besides the tractability of geometric discounting, there are few reasons to impose it as a reasonable model of individual or political behavior. The next section reviews the literature by discussing the lack of empirical and theoretical foundation for geometric/exponential discounting. Based on this, I conclude that "the collective evidence outlined above seems overwhelmingly to support hyperbolic discounting," as also concluded by earlier reviewers (Frederick et al 2002: 361).

The rest of the paper is deriving the implications of time-inconsistent preferences for investments and investment policies. Any action today, whether it regards investments in capital, technology or knowledge, will inevitably affect future choices. The current principal will thus have an incentive to distort current investments in order to influence the choices made by the future decision-makers. The optimal distortion, which can be implemented by a subsidy or a tax on investments, will depend on the type of capital/technology

to be invested in as well as how "upstream" the technology is: i.e., whether the technology is used as an input to another type of investment. The more upstream the technology is, the larger is the optimal subsidy level. For investments in "green" technology (complementary to future abatement or savings), the optimal subsidy is positive but it is optimal to instead tax investments in "brown" technology (i.e., technology which are strategic substitutes to future abatement or savings).

## 2. Discounting: Background and Foundation

In the nineteenth century, the debate regarding how to evaluate future costs and benefits included a large number of factors, some psychological and some of them were conflicting (e.g., one may have a pleasure from anticipating future benefits or derive utility from being rich today relative to the past: these early philosophers included Rae, Jevons, Senior and Bohm-Bawerk).

It was Ramsey (1928) who first suggested that one may want to maximize a weighted sum of future utilities:

$$v_t = \sum_{\tau=t}^{\infty} d_{\tau-t} u_\tau,$$

where the discount factor $d_\tau$ measures how important utility $\tau$ periods ahead is relative to utility right now (by this definition, $d_0 = 1$). The discount factor $d_\tau$ was not specified by Ramsey (although he argued that, normatively, $d_\tau$ ought to be a constant independent of $\tau$, implying $d_\tau = 1$ when $d_0 = 1$). Note that the corresponding discount factor for utility in period $t$ relative to utility a period before is,

$$\delta_t = \frac{d_t}{d_{t-1}} \Leftrightarrow d_t = \prod_{\tau=1}^{t} \delta_\tau. \tag{2.1}$$

*Theoretical foundation.* The "breakthrough" first came when Paul Samuelson (1937) *suggested* the now so familiar formula for $d_t$:

$$d_t = \delta^t = \left(\frac{1}{1+r}\right)^t \approx e^{-rt},$$

where $\delta$ is the corresponding constant discount factor between subsequent periods and $r$ is the associated constant discount rate. Later, Koopman (1960) provided an axiomatic

foundation for this formulae, and this firmly established geometric discounting as the standard way of evaluating future gains and losses in economics.

*Critique.* To many, the appeal of geometric discounting is that it is simplifying a difficult problem rather than it being a reasonable assumptions regarding individual behavior. Paul Samuelson was himself full of reservations when suggesting the geometric formulation, both as a representation of an individual's preference ("It is completely arbitrary to assume that the individual behaves so as to maximize an integral of [this] form," Samuelson, 1937: 159), or as an advice for a public planner ("any connection between utility as discussed here and any welfare concept is disavowed," p. 161). Nevertheless, "despite Samuelson's manifest reservations, the simplicity and elegance of this formulation was irresistible" (Frederick et al 2002: 355-6) and the criterion became "dominant... largely due to its simplicity and resemblance to the familiar compound interest formula and not as a result of empirical research demonstrating its validity" (Frederick et al, 2002: 352-3).

For intergenerational settings, one can easily make the case that stationary as an axiom must be violated: If the future self represents the next generation, it may be reasonable to assume that parents are "thoughtful," as in, for example, Barro (1974): If each generation's welfare is a weighted sum of its own utility as well as the next generation's welfare, then one can write this generation's welfare recursively as a weighted sum of all future utilities, where the discount factor is constant over time (leading to geometric discounting). However, if today's parents also care about the welfare of its grandchildren, then stationarity must be violated and the effective discount rate declines in time (Harstad, 1999, Saez-Marti and Weibull, 2005). There are also evolutionary arguments suggesting that humans may evolve and survive better if they have so-called hyperbolic discounting functions; see Dasgupta and Maskin (2005).[1]

*Empirical evidence.* There is now a vast literature showing that individual preferences

---

[1]Christian Gollier and Martin Weitzman have shown that if the growth rate $\dot{c}$ of consumption is uncertain, then one it is optimal to discount future consumption at a rate which is decreasing in time to reflect risk aversion and the accelerating level of risk (see for instance Gollier and Weitzman, 2010, Weitzman 2001). This paper abstracts from such uncertainty and discusses a complementary but more fundamental reason for why discount rates ought to be a decreasing function of time, which is that the pure time preference discount rate does decrease in time.

do not satisfy stationarity. That is, individuals may be impatient in the short-run, but the difference between time t and time t+1 tends to vanish as t becomes large, implying a discount rate between the two periods to approach zero as t grows (see Eisenhauer and Ventura (2006), Frederick et al (2002), Angeletos et al (2001)). To be specific, Paserman (2008) obtains identification from heterogeneity in unemployment durations and reservation-wages to find estimates of the short-run annualized discount rate that range from 11% to 91% and a long-run discount rate of only 0.1%. Based on this, he rejects the exponential discounting null hypothesis. Also, based on individual's participation in retirement accounts (IRAs), O'Donoghue and Rabin (1999) find that hyperbolic individuals would imply exactly the low IRA participation which we actually observe in reality. In another study, Laibson et al (2007) find that the short-term discount rate is 15% and the long-term discount rate is 3.8%.

*Experimental evidence.* There is also a large experimental literature suggesting that individuals discount less between period t and t+1 if t is far into the future (see, for example, Viscusi and Huber (2006), Kirby and Marakovic (1995), Benhabib, Bisin and Schollter (2010), Ainslie (1992), Kirby and Herrnstein (1995), Thaler (1981)). The traditional experiment offers the choice between a benefit ("one apple") today or a larger benefit tomorrow ("two apples"). While individuals may prefer today's benefit to tomorrow's, the ranking is reversed if the choice instead is between a benefit in period t, where t is further into the future. After reviewing the literature, Frederick et al (2002: 361) conclude that "the collective evidence outlined above seems overwhelmingly to support hyperbolic discounting."

*Hyperbolic discounting.* It may be quite natural to think that, although there may be a big difference between a gain realized today and a gain realized in ten years, the difference between year 100 and year 110 is smaller (this intuition is better explained by Strotz, 1956). In other words, of importance may be only the relative difference (the difference in time divided by the length from today's point of view). Our basic human senses are developed to care about relative differences: The closest sound is easiest to hear when two sources are both located nearby, but when they are both further away, it is easiest to hear the sound that is in fact the loudest. From a distance, the largest

mountain does indeed look like it is larger than the smaller one, but when they are both near, the mountain that is located closest tend to look the biggest. If our sense for time follows the same path, as experimental evidence suggest that it does, then the discount factor must be "hyperbolic" so that utility at time $t$ will be weighted by the discount factor:

$$d_t = \frac{1}{1 + \alpha t}, \tag{2.2}$$

where $\alpha > 0$ is a constant measuring impatience or the scale of time. By inserting in (2.1) we find the discount factor between period $t$ and $t - 1$:

$$\delta_t = 1 - \frac{\alpha}{1 + \alpha t} = 1 - \alpha d_t,$$

which is approaching one as $t$ grows.

*Quasi-hyperbolic discounting.* David Laibson (1997) adopted a simpler approximation of (2.2) often referred to as quasi-hyperbolic discounting:

$$d_t = \beta \delta^t \text{ if } t > 0,$$

where both $\beta < 1$ and $\delta < 1$. With such discount factors, the objective at time $t$ is to maximize:

$$v_t = u_t + \beta \sum_{\tau=t+1}^{\infty} \delta^{\tau-t} u_\tau.$$

This formula was first presented by Phelps and Pollak (1968) who argued that it may represent "imperfect altruism" between generations. Other names for the same discount rate function are $(\beta, \delta)$-discounting, quasi-geometric discounting, quasi-exponential discounting, and sometimes misleadingly but simply hyperbolic discounting.

*Governments and normative justification.* Although individuals apparently apply discount rates that decline in time, is it clear that governments ought to do the same? Normatively, it seems hard to justify giving future utilities less weight than present utilities. Utilitarianism states that the normative objective is to sum the utility of all agents with equal weight. With this criterion, then individuals living at different times will still be counted equally, implying a discount rate of zero. Ramsey strongly advocated the use of zero discount rate from a normative point of view. The Stern review (2007) applied a

6

discount rate of 0.001, i.e., 0.1%, reflecting the slim chance that the future may cease to exist if a catastrophe occurs.

While there may be no normative justification for a positive discount rate, such an "ethical" government will fail to be reelected at the next election stage, if the voters do not share such a perfect altruistic preference. To be reelected, the government will inevitably apply the same discount rates as the voters. Furthermore, the government is, of course, consisting of individual policy-makers, sharing these preferences regarding the future. For both reasons, the individual's preference for a positive discount rate which decline in time will and must be applied also by the government (although, the citizens may prefer that the government apply a lower discount rate than the citizens themselves; Caplin and Leahy 2004).

Rather than insisting that the government use normative criteria or apply a zero discount rate, this paper takes as given that a democratic government will and must apply the same discount rates as the citizens, and, when these rates decline in time, we need to look at the consequences for investment policy.

*The benefit of committing.* There is some evidence that individuals fear they will retire too early (Diamond and Koszegi 2003) and that hyperbolic discounting is a reason for obesity (Scharff 2009). Some scholars have, based on such reasoning, suggested that there is a role for governments in regulating individual behavior. For example, since hyperbolic discounters save too little, the government can help the decision-makers to commit by subsidizing saving.[2] Or, since hyperbolic discounters find it hard to quit smoking, one may tax tobacco more (Gruber and Koszegi 2001). There is also evidence in that individuals try to commit their future self's by signing up to saving plans which makes it costly to quit (Thaler and Benartzi (2004)), or one may subsidize the future self by paying now the future cost of good behavior, like attending the gym (DellaVigna and Malmendier 2006). For climate change, see Karp (2005), or Gerlagh and Liski (2013) who derive optimal carbon tariffs in a setting with quasi-hyperbolic discounting.

---

[2]Krusell et al. (2003, 2009, and 2010). See also Laibson and Harris (2001).

## 3. Investments in Capital

### 3.1. A model

The principal, planner, or person $P$ making investments at time $t \geq 0$ is referred to as $P_t$. Her momentary utility is $u_t$ and objective is to maximize $v_t = \sum_{\tau=t}^{\infty} d_{t-\tau} u_\tau$, where $d_{t-\tau} \in (0,1)$ measures the discount factor on utilities realized $t - \tau$ periods ahead. Thus, utilities at time $t$ relative to utilities at $t-1$ are then discounted by the factor $\delta_t \equiv d_t/d_{t-1}$. I will assume that $\delta_t$ is strictly increasing in $t$ unless otherwise stated (for example, $\delta_t = \delta$ is constant when considering geometric discounting).

The following subsections describe the investments that can be taken, how these affect payoffs, and characterize the optimal investment level. To provide an intuitive description of $P$'s investment decision, and to characterize it in terms of policy, it is useful to compare the planner's choice to the "market equilibrium" benchmark. In reality, the government may not make the investments directly. Instead the investments into capital or technology stock are made by private firms or agents. Thus, as a benchmark, we may consider an agent $A_t$ making the investment decision at $t$.

The agent ("he") applies the same discount factors as the principal ("she") so there is no disagreement over how to evaluate future gains and losses. Further, the investment cost faced by the agent is identical to the investment cost faced by the principal. Finally, it is assumed that the agent obtains perfect property rights over the returns to the investments, so that $A_t$ can sell or lease the returns at the later stages. Thus, in most traditional settings, and in my own benchmark examples, $A_t$ makes the correct decision from $P_t$'s point of view and there is no need for the principal to regulate the agent.

Nevertheless, in the following section I do allow the principal to subsidize the agent by paying a fraction $\underline{s}$ of his investment cost (or taxing such investments, if $\underline{s} < 0$). The principal considers subsidies simply as transfers within the society with no real cost, but the subsidy does affect the agent's decision. The agent invests up to the point where the marginal investment cost, multiplied with $(1 - \underline{s})$, equals the marginal return, appropriately discounted. Alternatively, if the principal could set a subsidy $s$ (or tax $-s$) on the agent's revenues from the investment, then the agent would invest up to the point where

8

the marginal investment cost would equal the discounted marginal return multiplied with $(1 + s)$. The two alternative investment-policies are obviously equivalent if

$$1 - \underline{s} = \frac{1}{1+s} \Leftrightarrow \underline{s} = 1 - \frac{1}{1+s}.$$

In the analysis below it will be most convenient to represent the subsidy by $s$ (which is a tax if $s < 0$) even if the principal $P_t$ actually implements $s$ by subsidizing the cost of investments at time $t$ by $\underline{s} = 1 - 1/(1+s)$.

## 3.2. A simple investment

Consider an action $a \in \Re$ to be made at time $t_a$. The level of $a$ affects the utility at time $t_a$ and also the utility $\Delta_a$ periods later, i.e. at time $t_a + \Delta_a \geq t_a$. In general, one needs to trade off current utility with future utility when choosing $a$. Without much loss of generality, let a higher level of action $a$ be costly at time $t_a$ but beneficial at time $t_a + \Delta_a$, so that increasing $a$ can indeed be considered as an investment. The investment can be in health, education, infrastructure or pollution abatement. In order to introduce some intuitive notation, define $u_{t_a}(a; k) \equiv -C(a; k)$ and $u_{t_a + \Delta_a}(a; k) \equiv B(a; k)$ so that $C(\cdot)$ is the investment cost and $B(\cdot)$ is the benefit or the returns of the investment. For now, $k$ is an exogenously or historically given capital stock.

The principal would like to maximize $v_{t_a} = -C(a; k) + d_{\Delta_a} B(a; k)$, giving:

$$C_a = d_{\Delta_a} B_a \tag{3.1}$$

To get interior solutions and satisfy second-order conditions, let $C(\cdot)$ be increasing and convex in $a$ while $B(\cdot)$ is increasing and concave in $a$: $C_a > 0$, $C_{aa} \geq 0$, $B_a > 0$, $B_{aa} \leq 0$.

If the agent makes the investment, he pays the cost $(1 - \underline{s}_a) C(a; k)$ at time $t_a$ where $\underline{s}_a$ is an investment subsidy (or tax if $\underline{s}_a < 0$) set by the principal. After investing, the agent sells the returns and receives $B(a; k) - B(0; k)$ at time $t_a + \Delta_a$. Thus, the agent's problem is:

$$\max_a - (1 - \underline{s}_a) C(a; k) + d_{\Delta_a} [B(a; k) - B(0; k)]$$
$$\Rightarrow \frac{C_a}{1 + s_a} = d_{\Delta_a} B_a,$$

where $1 + s_a \equiv 1/(1 - \underline{s}_a)$. The following benchmark-result follows straightforwardly.

**Proposition 0.** *At time $t_a$ the principal invests according to* (3.1) *either directly or indirectly by setting $s_a = 0$.*

Of course, the desired subsidy $s_a = 0$ corresponds to, and can be implemented by, a subsidy on the current investment cost equal to $\underline{s}_a \equiv 1 - 1/(1 + s_a) = 0$.

### 3.3. Investments in capital

The payoffs above are depending not only on action $a$ but also on the capital stock $k \in \Re$. Let $t_k \leq t_a$ be the time at which the decision $k$ is taken, so that $\Delta_k \equiv t_a - t_k \geq 0$ is the time it takes for the capital to mature or be built.

As with action $a$, the choice of $k$ is considered to be an investment and it is thus costly at time $t_k$. Define $u_{t_k}(k) \equiv -D(k)$ so that $D(\cdot)$ measures the cost or disutility of investing in capital $k$. Let $D_k > 0$ and $D_{kk} \geq 0$. To capture the benefit of the investment, let either $du_{t_a}/dk = -C_k \geq 0$ or $du_{t_a + \Delta_a}/dk = B_k \geq 0$ or both.

In the market, the agent makes the investment at cost $D(\cdot)$ and he sells the return at times $t_a$ and $t_a + \Delta_a$. If $s^k$ represents the subsidy on investments in $k$, it is straightforward to see that the agent invests so that the following first-order condition is satisfied:

$$\frac{D_k}{1 + s_k} = -d_{\Delta_k} C_k + d_{\Delta_k + \Delta_a} B_k. \tag{3.2}$$

Just as before, the principal can induce the same decision by subsidizing the cost of investing at time $t_k$ by $\underline{s}_k \equiv 1 - 1/(1 + s_k)$.

It is straightforward to show that the principal prefers $s_k = 0$ if she agrees with how $a$ is chosen at time $t_a$: to see this, we can simply apply the Envelope theorem when maximizing $v_{t_k}$ with respect to $k$ to get $D_k = -d_{\Delta_k} C_k + d_{\Delta_k + \Delta_a} B_k$.

In general, however, the principal may not agree with how $a$ is chosen. A larger $a$ is an investment for the future and $P_{t_k}$ may disagree with $P_{t_a}$ regarding how future benefits at time $t_a + \Delta_a$ compare to the cost at time $t_a$. When $P_{t_k}$ prefers that $P_{t_a}$ pays more attention to future benefits, $P_{t_k}$ would like $P_{t_a}$ to invest more. If $k$ and $a$ are complementary, it is indeed possible to encourage a larger future investment $a$ by increasing $k$ above the level

that would have been optimal otherwise. If $k$ and $a$ are strategic substitutes, the benefit of $a$ is reduced by $k$, and to encourage $a$ the principal has to reduce (or tax) investments in $k$. To see how $k$ affects $a$ we can simply differentiate (3.1) to get:

$$\frac{da}{dk} = \frac{d_{\Delta_a} B_{ak} - C_{ak}}{C_{aa} - d_{\Delta_a} B_{aa}} \Rightarrow \tag{3.3}$$

$$\text{sign}\left(\frac{da}{dk}\right) = \text{sign}\left(d_{\Delta_a} B_{ak} - C_{ak}\right).$$

To appreciate the problems of the principal, consider the following important application of the model: Let the level of $a$ measure pollution abatement while $k$ represents one of (at least) three different types of capital:

*Green capital* is assumed to be complementary to pollution abatement. Such technology can be cleaning technology or alternative energy sources: in either case, a larger stock of green technology is a strategic complement to reducing pollution, and the marginal cost of abating. So, $C_a(a;k)$ decreases in $k$: $C_{ak} < 0$. The green capital does not (by assumption) affect the cost of pollution directly, so $B_{ak} = B_k = 0$. Substituted into (3.3), we see that $da/dk > 0$.

*Brown capital* refers to drilling technologies or investments in industries that pollute. Such capital may be beneficial in the sense that $du_{t_a}/dk = -C_k > 0$, but a larger level of $k$ is also making it costly to cut back on pollution. Thus, $C_{ak} > 0$, meaning that $a$ and $k$ are strategic substitutes. The brown capital does not (by assumption) affect the cost of pollution directly, so again $B_k = 0$. Consequently, $da/dk < 0$.

*Adaptation capital* refers to investments that enhance the possibility to deal with pollution at time $t_a + \Delta_a$. For example, one can invest in agricultural products that can cope with pollution or climate change, or one can build infrastructure that is robust to pollution, climate change, or sea-level rises. Such investments are not only increasing the benefit at time $t_a + \Delta_a$ but they also reduce the marginal cost of pollution: in other words, a larger level of $k$ reduces the value of $a$ so that $B_{ak} < 0 < B_k$. Such adaptation capital does not (by assumption) affect the cost of abating at time $t_a$, so $C_{ak} = 0$. Substituted into (3.3), $da/dk < 0$. (As a fourth special case, one could of course consider the possibility $B_{ak} > 0$ while $C_{ak} = 0$).

**Proposition 1.** *At time $t_k$ the principal ensures that $k$ satisfies (3.2) with:*

$$s_k = \left( \frac{d_{\Delta_k + \Delta_a}}{d_{\Delta_a}} - d_{\Delta_k} \right) \frac{d_{\Delta_a} B_a}{d_{\Delta_k + \Delta_a} B_k - d_{\Delta_k} C_k} \cdot \frac{da}{dk},$$

*where the second-order is:*

$$\Omega \equiv -D_{kk} + d_{\Delta_k + \Delta_a} \left( B_{kk} + B_{ka} \frac{da}{dk} \right) - d_{\Delta_k} \left( C_{kk} + C_{ka} \frac{da}{dk} \right) \qquad (3.4)$$

$$+ \left[ d_{\Delta_k + \Delta_a} B_a - d_{\Delta_k} C_a \right] \frac{d^2 a}{dk^2}$$

$$+ \left[ d_{\Delta_k + \Delta_a} \left( B_{ak} + B_{aa} \frac{da}{dk} \right) - d_{\Delta_k} \left( C_{ak} + C_{aa} \frac{da}{dk} \right) \right] \frac{da}{dk} \leq 0.$$

*Proof.* The problem of $P_{t_k}$ is to maximize $v_{t_k}$,

$$-D(k) - d_{\Delta_k} C(a; k) + d_{\Delta_k + \Delta_a} B(a; k),$$

giving the f.o.c.

$$\begin{aligned} D_k &= -d_{\Delta_k} \left[ C_a \frac{da}{dk} + C_k \right] + d_{\Delta_k + \Delta_a} \left[ B_a \frac{da}{dk} + B_k \right] \\ &= -d_{\Delta_k} C_k + d_{\Delta_k + \Delta_a} B_k + \left[ d_{\Delta_k + \Delta_a} B_a - d_{\Delta_k} C_a \right] \frac{da}{dk} \\ &= -d_{\Delta_k} C_k + d_{\Delta_k + \Delta_a} B_k + \left[ d_{\Delta_k + \Delta_a} - d_{\Delta_k} d_{\Delta_a} \right] B_a \frac{da}{dk}, \end{aligned}$$

when inserting (3.1). Since $A_{t_k}$ invests according to (3.2), incentives are aligned if

$$s_k = \frac{\left[ d_{\Delta_k + \Delta_a} - d_{\Delta_k} d_{\Delta_a} \right] B_a}{-d_{\Delta_k} C_k + d_{\Delta_k + \Delta_a} B_k} \frac{da}{dk}.$$

*QED*

I will now discuss and explain the proposition by making a number of corollaries.

**Corollary 1.** *With exponential discounting, $s_k = 0$.*

In traditional settings where the decision-makers have time-consistent preferences, there is no need for the principal today to distort the choices of her future self. So, since an investor captures the full future return to his effort, his decision will be optimal and there is no need to regulate him. This confirms the presumption at the beginning as well as Proposition 0 where the principal would never like to regulate investments in $a$.

**Corollary 2.** *If either $\Delta_k = 0$ or $\Delta_a = 0$, then $s_k = 0$.*

If $\Delta_k = 0$, it takes no time to build the capital to be used at time $t_a$. The principal making decision $k$ and $a$ have the same preference and there is thus no need to distort the decision on $k$. Alternatively, if $\Delta_a = 0$, the principal deciding on $k$ is immediately getting the benefit herself and there is no future utility which the two self's evaluate differently.

From now I will assume $\Delta_k \Delta_a > 0$ unless otherwise stated. Although Corollary 2 is not a limit result, it does suggest that if the time to develop capital or the pollution lag is quite short, then the optimal subsidy on capital is small. This may change when the lags are long, as we will now discuss.

**Corollary 3.** *With quasi-linear discounting,*

$$s_k = (1 - \beta) \frac{\delta^{\Delta_a} B_a}{\delta^{\Delta_a} B_k - C_k} \cdot \frac{da}{dk}.$$

With time-inconsistent preferences, the principal is not satisfied with the future choice of $a$. So, if $k$ and $a$ are complementary, for example, $P_{t_k}$ has the point of view that $P_{t_a}$ invests too little and is willing to pay too little for $k$. Investors are then not sufficiently motivated to invest, so $P_{t_k}$ prefers them to invest more. A larger $k$, after all, will induce $P_{t_a}$ to increase $a$. It is for this reason $P_{t_k}$ finds it necessary to regulate investors and subsidize investments in $k$ when $a$ and $k$ are strategic complements. If $a$ and $k$ are strategic substitutes, however, the same logic implies that it is optimal to tax investments in $k$.

**Corollary 4.** *With strictly increasing discount factors,*

$$sign \ (s_k) \ = \ sign \ \left( \frac{da}{dk} \right) = sign \ (d_{\Delta_a} B_{ak} - C_{ak}).$$

This corollary can be exemplified by referring to the three types of capital discussed above.

**Corollary 5.**

*(i) It is optimal to subsidize investments in green capital.*

*(ii) It is optimal to tax investments in brown capital.*

13

*(iii) It is optimal to tax investments in adaptation capital.*

The third part of the corollary may seem provocative. Adaptation may certainly be a good thing, in that $B_k > 0$, but note that the investor will himself capture this value so this creates no reason to subsidize $k$. On the contrary, more investments in adaptation technology will reduce the cost of polluting at time $t_a$ and the level of abatement will thus be reduced as well. The principal at time $t_k$, however, is instead preferring a larger $a$ and this can be achieved by reducing the level of abatement capital $k$.

## 4. Investments in Technology

The previous subsection made a "horizontal" distinction between different types of capital at the same level in the production hierarchy: while some technologies were complementary to the abatement decision, others were strategic substitutes.

One may also distinguish technologies "vertically" by their level in the production hierarchy. For example, while the number of windmills will make it cheaper to reduce pollution, the production cost of each windmill will depend on the level of technology, knowledge, or basic research. The fact that distinguishing between the levels may be important is evident when comparing the decision on $s_k$ to the downstream decision $s_a = 0$. This section is taking us another step upstream by analyzing investments in technology.

Let now the cost of investing in $k$ be given by $D(k; r)$, where $r$ is the level of technology that exists at time $t_k$. Introducing $r$ into the formulae does not change anything in the analysis above since $r$ must be taken as exogenously given in the stages studied so far. While the capital $k$ can, as before, be of any type (complementary or a substitute to $a$), I seek a clean comparison between these cases by keeping fixed one type of technology investment: $r$ and $k$ are assumed to be complementary, in that more technology reduces the cost of building capital: $D_r < 0$ and $D_{kr} < 0$. It is also assumed that, for any given level of $k$ and $a$, the technology level $r$ does not influence directly the costs or benefits of action $a$. Thus, $C$ and $B$ are not functions of $r$ directly.

Let the time at which $r$ is chosen by $t_r \leq t_k$. Thus, $\Delta_r \equiv t_k - t_r \geq 0$ is the time required for the technology to mature or be developed after the process has been initiated. An agent investing in $r$ will be able to capture the full value of reducing the cost $D$. If $u_{t_r}(r) \equiv -E(r)$, so that an increasing and convex $E(r)$ represents the cost of investing in technology, then the agent invests up to the point where the marginal investment cost equals the marginal benefit:

$$\frac{E_r}{1 + s_r} = d_{\Delta_r}(-D_r), \qquad (4.1)$$

where $s_r$ represents the subsidy on technology. Just as before, the principal can induce the same decision by subsidizing the cost of investing in $r$ by $\underline{s}_r \equiv 1 - 1/(1 + s_r)$.

It is, as before, straightforward to show that if the principal's time preferences are time consistent, such that she finds optimal her own later choices on $k$ and $a$, then there is again no need to regulate the current investor. By the Envelope theorem, the principal sets $E_r = -d_{\Delta_r} D_r$ or $s_r = 0$.

With time-inconsistent preferences, however, the principal is not satisfied with the future choices of $k$ and $a$ and, in order to influence these choices, it may be optimal to distort today's investments in $r$. To see how $r$ influences $k$ we can simply differentiate the first-order condition for $k$ to get:

$$\frac{dk}{dr} = \frac{-D_{kr}}{(-\Omega)} > 0,$$

where $\Omega < 0$ is the expression for the second-order condition when $v_{t_k}$ is maximized with respect to $k$, see Equation (3.4).

Despite the apparent analogy to the situation where $k$ could influence $a$, there might be quite different forces at work when setting $r$ and $s_r$ than when setting $k$ and $s_k$. This is evident from the following result and the explanation that follows it.

**Proposition 2.** *At time $t_r$ the principal ensures that $r$ satisfies (4.1) with:*

$$
\begin{aligned}
s_r &= s_r^k + s_r^a, \text{ where} \\
s_r^k &= \frac{D_k}{-D_r}\left[\frac{d_{\Delta_r + \Delta_k}}{d_{\Delta_r} d_{\Delta_k}} - 1\right]\frac{dk}{dr} > 0 \text{ and} \\
s_r^a &= \frac{D_k}{-D_r}\left[\frac{d_{\Delta_r + \Delta_k + \Delta_a} d_{\Delta_k} - d_{\Delta_r + \Delta_k} d_{\Delta_k + \Delta_a}}{d_{\Delta_k} d_{\Delta_r}(d_{\Delta_k + \Delta_a} - d_{\Delta_k} d_{\Delta_a})}\right]\left(B_k + B_a \frac{da}{dk}\right)\frac{dk}{dr}.
\end{aligned}
$$

*Proof.* The principal $P_{t_r}$ prefers to maximize $v_{t_r}$ with respect to $r$, giving the f.o.c.:

$$E_r = -d_{\Delta_r}\left(D_r + D_k\frac{dk}{dr}\right) - d_{\Delta_r+\Delta_k}\left(C_k\frac{dk}{dr} + C_a\frac{da}{dk}\frac{dk}{dr}\right)$$
$$+ d_{\Delta_r+\Delta_k+\Delta_a}\left(B_k\frac{dk}{dr} + B_a\frac{da}{dk}\frac{dk}{dr}\right).$$

When substituting in for the first-order conditions for $a$ and $k$, one can compare the expression to (4.1) and derive $s_r$. *QED*

In the following I will explain and discuss Proposition 2 step-wise by making a number of corollaries referring to special cases.

**Corollary 6.** *With exponential discounting, $s_r = s_r^k = s_r^a = 0$.*

It is easy to check that with exponential discounting, the brackets in both $s_r^k$ and $s_r^a$ are zero. Intuitively, when the principal is time consistent, she is perfectly satisfied with her own future choices of $k$ and $a$. She has no desire to distort these choices and thus (by the Envelope theorem) she prefers a level of technology which takes into account only the saved cost when investing in capital. The perfect agent invests optimally and there is no need to regulate him.

**Corollary 7.** *(i) If $\Delta_r = 0$, then $s_r = s_r^k = s_r^a = 0$. (ii) If $\Delta_k = 0$, then $s_r^k = 0$ so $s_r = s_r^a$. (iii) If $\Delta_a = 0$, then $s_r^a = 0$ so $s_r = s_r^k$.*

Part (i) is saying that if technology takes no time to invest, it needs no subsidy. The reason is, of course, that when time preferences are not any different today as when the decision on $k$ is to be taken, there is no reason to distort the future decision on $k$. I will return to parts (ii) and (iii) below. In the following I will assume that $\Delta_r\Delta_k\Delta_a > 0$ unless stated otherwise.

With time-inconsistent preferences, $P_{t_a}$ may disagree with the future choices of $k$ and $a$ and thus $r$ may be chosen, or distorted, in order to influence the future choices. With quasi-hyperbolic discounting, the bracket in the expression for $s_r^a$ is zero so $s_r$ simplifies.

**Corollary 8.** *With quasi-hyperbolic discounting, $s_r^a = 0$ and*

$$s_r = s_r^k = \frac{D_k}{-D_r}\left(\frac{1}{\beta} - 1\right)\frac{dk}{dr} > 0.$$

It is optimal to always subsidize investments in technology since this will induce more investments in $k$ at time $t_k$. The optimal investment is larger if the current principal disagrees strongly with the future principal ($\beta$ is small) and if increasing $r$ is an efficient way of increasing $k$ (in that $dk/dr$ is large). Note the similarity between $s_r^k$ and $s_k$ for the special case of "green" or complementary capital: it is exactly the same forces at work. Technology $r$ is by assumption complementary to $k$ and thus $r$ requires a subsidy just as $k$ did when $k$ were complementary to $a$. In special symmetric settings, $s_r^k = s_k$.[3]

When deriving $s_r^k$ and $s_r$ for the case with quasi-hyperbolic discounting, it is irrelevant whether the capital $k$ is itself green or brown: investments in $r$ are optimal to subsidize regardless.

The explanation for the irrelevance of the capital type is the following. Although $P_{t_r}$ disagrees with $P_{t_k}$ regarding the appropriate level of investments $k$ at time $t_k$, these two self's agree perfectly when trading off utilities between two future dates. With quasi-linear discounting, the discount factor of utility at time $t + \tau$ relative to time $t$ is $\delta^\tau$ for every decision-maker $P_{t'}$ whenever $t' < t$. Thus, given that $P_{t_k}$ distorts $k$ in order to influence the future choice of $a$, $P_{t_a}$ cannot improve on this decision.

When discount factor $\delta_t$ is strictly increasing in $t$, however, the conclusions are quite different. Then, $s_r^a$ is strictly positive if and only if $dB(a(k), k)/dk > 0$. That is, when a larger $k$ at time $t_a$ increases utility also at time $t_a + \Delta_a$. While $k$ may affect $B$ directly when $B_k \neq 0$, $k$ also affects $B$ through the choice of $a$. When $k$ represents green technology, more $k$ leads to more abatement and thus a larger benefit $B$. The subsidy $s_r$ is thus the sum of two positive terms.

**Corollary 9.** *For green technology, $s_r^k > 0$ and $s_r^a > 0$ so the optimal subsidy is positive.*

For brown technology, however, we know that $a$ decreases when the level of brown capital increases since this lead to more pollution. Thus, the second term of $s_r$ is negative.

---

[3]To see this, substitute in for the first-order condition for $a$, set $B_k = 0$ and rewrite $s_k$ to

$$s_k = \frac{[d_{\Delta_k + \Delta_a}/d_{\Delta_a} d_{\Delta_k} - 1] C_a}{-C_k} \frac{da}{dk}.$$

If $C(x; y) = D(x; y)$, $da/dk = dk/dr$ and $\Delta_r = \Delta_k = \Delta_a$, then the expression for $s_r^k$ coincides with the expression for $s_k$. The next section is making the analogy clearer.

It is certainly possible that $s_r < 0$ if the second term dominates the first, positive term. This will be the case when the degree of substitutability between $k$ and $a$ is large (in that $C_{ak}$ and thus $da/dk$ is large). In this case, the results from Corollary 8 would be reversed.

**Corollary 10.** *For brown technology, $s_r^k > 0$ but $s_r^a < 0$ so the optimal subsidy $s_r$ can be positive or negative: if $\Delta_k = 0$, $s_r < 0$ but if $\Delta_a = 0$, $s_r > 0$.*

For both green and brown technology, the second term of $s_r$, $s_r^a$, has the same sign as $s_k$ since $B_k = 0$. For adaptation technology, however, we know that $B_k > 0$ at the same time as $da/dk < 0$. It is thus unclear whether the second term is positive or negative when it comes to adaptation technology. As shown, the answer hinges on the sign of $dB\left(a\left(k\right), k\right)/dk$.[4] This is quite intuitive.

**Corollary 11.** *For adaptation technology, $s_r^k > 0$ but $s_r^a$ can be either positive or negative, and so can the optimal subsidy. If $\Delta_a = 0$, $s_r > 0$.*

### 4.1. A Hierarchy of Technologies

The analysis above suggests that for setting investment policy it is crucial to determine the level of the technology in the production hierarchy. While the final investment step before consumption $(a)$ did not need any regulation, investments in complementary capital or technology are subsidized. Furthermore, the investment in technology will be subsidized at a rate which consisted of two positive terms rather than just one as for the investment in capital (where the first term corresponds to the optimal subsidy on investments in capital). This does not prove, but it does suggest, that the optimal subsidy at investments further upstream is higher.

---

[4]Note that $B\left(a\left(k\right); k\right)$ increases in $k$ if the direct effect $B_k > 0$ is large enough:

$$\frac{dB\left(a\left(k\right), k\right)}{dk} = B_k \left(1 - \frac{B_a}{B_k}\frac{d_{\Delta_a}\left(-B_{ak}\right)}{C_{aa} - d_{\Delta_a}B_{aa}}\right),$$

which is positive indeed if, for example, $B_a/B_k$ decreases in $a$: i.e., if the demand (willingness to pay) for abatement (relative to the willingness to pay for adaptation capital) decreases in the level of abatement. To see this statement, note that the derivative of $B_a/B_k$ with respect to $a$ is $\left(B_{aa}B_k - B_kB_{ak}\right)/\left(B_k\right)^2$, and when this is negative then $\left(-B_{aa}\right)B_k > B_k\left(-B_{ak}\right)$. Substituted into the above formula we get that the parenthesis is positive if $C_{aa} \geq 0$.

To illustrate this argument, consider the following simplification of the above model. At time $T - \Delta$, say, the principal (or the agent) can make an investment $a_1$ which increases utility at time $T$ with $b > 0$ units. The cost of investing in $a_1$ is quadratic and given by $c_1 (a_1 - a_2)^2 / 2$, where $c_1 > 0$ is a constant and $a_2$ is the complementary capital when making the investment $a_1$. The level of $a_2$ is determined $\Delta$ periods earlier and, then, the investment cost of $a_2$ is similarly given by $c_2 (a_2 - a_3)^2 / 2$, where $c_2 > 0$ is a constant and $a_3$ is now the complementary technology when making the investment in capital $a_2$.

This model is a special case of the "green capital" model above (where I instead used the labels $a$ for $a_1$, $k$ for $a_2$, $r$ for $a_3$, and now I am further assuming $\Delta_r = \Delta_k = \Delta_a = \Delta$).

Given this simplification, however, it is now straightforward to both derive explicit intuitive formulae and generalize by allowing for many stages or levels of upstream technology: simply let the investment cost at stage $\tau$, as counted from downstream, be given by $-c_\tau (a_\tau - a_{\tau-1})^2 / 2$, where $c_\tau > 0$ is a constant and the current principal must take $a_{\tau-1}$ as exogenously given. Each stage of investment takes $\Delta$ periods, so $a_\tau$ is the investment level at time $T - \tau\Delta$. Investments in more basic research is characterized by a larger $\tau$, but such investments will contribute to increased consumption only $\tau\Delta$ periods later.

Just as before, each investment choice might be taken by the market or an investing agent. The principal can then subsidize the different (vertically organized) types of investments. Let $\underline{s}_\tau$ measure principal $P_{T-\tau\Delta}$'s subsidy on the cost of investing in technology $\tau$ stages upstream. Equivalently, $s_\tau \equiv 1/(1 - \underline{s}_\tau) - 1$ could measure the corresponding subsidy on revenues. The question, then, is how $\underline{s}_\tau$, or equivalently $s_\tau$, depends on the stage $\tau$ in the production technology.

**Proposition 3.** *The optimal subsidy on investments at stage $\tau$ is:*

$$s_\tau = \frac{d_{\tau\Delta}}{d_\Delta d_{(\tau-1)\Delta}} - 1.$$

*Proof.* The proof permits the lag $\Delta_\tau$ to vary in $\tau$. For technology 1 the principal at invests such that

$$a_1 - a_2 = d_{\Delta_1} \frac{b}{c_1}.$$

Note that the investment cost becomes $(d_{\Delta_1} b)^2 / 2c_1$, independent of $a_2$. This makes the principal at stage 1 willing to pay $b d_{\Delta_1}$ for another unit of $a_2$. Thus, when investing in $a_2$

19

the principal realizes that a higher level of $a_2$ is (only) affecting utility $\Delta_1 + \Delta_2$ periods later. She therefore invests up to the point where

$$a_\tau - a_{\tau+1} = (d_{\Delta_1 + \ldots + \Delta_\tau}) \frac{b}{c_\tau},$$

where $\tau = 2$. The same condition holds for any $\tau > 2$.

If investments is made by an agent, he invests in $a_1$ such that

$$\frac{a_1 - a_2}{1 + s_1} = d_{\Delta_1} \frac{b}{c_1},$$

which coincides with the principal's preferred investment when $s_1 = 0$. This agent is thus willing to pay $d_{\Delta_1}$ for another unit of $a_2$. This, in turn, implies that the agent investing in $a_2$ invests such that

$$a_\tau - a_{\tau+1} = \frac{b}{c_\tau} (1 + s_2) \, d_{\Delta_\tau} d_{\Delta_1 + \ldots + \Delta_{\tau-1}},$$

where $\tau = 2$. The same condition holds for any $\tau > 2$. Thus, the principal sets $s_\tau = d_{\Delta_1 + \ldots + \Delta_\tau} / d_{\Delta_\tau} d_{\Delta_1 + \ldots + \Delta_{\tau-1}} - 1$ to induce the optimal choice. Parts (i)-(iii) follow directly. For hyperbolic discounting, we can write

$$
\begin{aligned}
s_\tau &= \frac{(1 + \alpha \Delta_\tau)(1 + \alpha(\Delta_1 + \ldots + \Delta_{\tau-1}))}{1 + \alpha(\Delta_1 + \ldots + \Delta_\tau)} - 1 \\
&= \alpha \Delta_\tau \left[ 1 - \frac{1 + \alpha \Delta_\tau}{1 + \alpha(\Delta_1 + \ldots + \Delta_\tau)} \right],
\end{aligned}
$$

giving (iv) as a special case. To see (v), set $\Delta_\tau = \Delta$ and write

$$s_\tau = \frac{\prod_{i=1}^{\tau\Delta} \delta_i}{\left( \prod_{i=1}^{\Delta} \delta_i \right) \left( \prod_{i=1}^{\Delta(\tau-1)} \delta_i \right)} - 1 = \frac{\prod_{i=\Delta(\tau-1)+1}^{\Delta\tau} \delta_i}{\left( \prod_{i=1}^{\Delta} \delta_i \right)} - 1,$$

where the numerator is a product of $\Delta$ factors, $\delta_{\Delta(\tau-1)+1}$, $\delta_{\Delta(\tau-1)+2}$, ... $\delta_{\Delta(\tau-1)+\Delta}$, etc, each of which is increasing in $\tau$. *QED*

The following corollaries follow straightforwardly:

**Corollary 12.**

*(i) The final investment-stage does not require a subsidy: $s_1 = 0$.*

*(ii) With geometric discounting, $s_\tau = 0$ for all $\tau \geq 1$.*

*(iii) With quasi-hyperbolic discounting, $s_\tau = 1/\beta - 1 > 0$ is constant for all $\tau > 1$.*

*(iv) With hyperbolic discounting,*

$$s_\tau = \alpha\Delta \left[ 1 - \frac{1 + \alpha\Delta}{1 + \alpha\Delta\tau} \right]$$

*so this subsidy is strictly increasing in $\tau$.*

*(v) With strictly decreasing discount rates, $s_\tau$ is always strictly increasing in $\tau$.*

This confirms the previous findings and the intuition discussed already.

## 4.2. Conclusions

There is large amount of evidence indicating that decision-makers have time-inconsistent preferences and behave as if they are more patient regarding long-run decisions than for short-term decisions. With this motivation there is also a literature studying how decision-makers can benefit by committing themselves to certain actions in the future.

This paper reports on the preliminary findings of an ongoing project in this line of research. I consider a principal (decision-maker or government) that would like to stick to a certain plan but her future self is going to be impatient and deviate from the plan. To influence the future principal, the current principal has an incentive to subsidize or invest more in capital or technologies which are complementary to future savings, but to tax or invest less in technologies that are strategic substitutes to future savings. In a relatively general model, I derive optimal subsidies/taxes as a function of both the type of technology and the technology's level in the production hierarchy. A finding is that it is optimal to subsidize more upstream technologies at a higher rate.

Although parts of the model are very general, the implications for environmental policy are noteworthy. Even when abstracting from externalities or public good aspects, I find that it is optimal to subsidize investments in "green capital" (like alternative energy

sources) while it is optimal to tax investments in "brown capital" (like drilling technology). Investing in adaptation or adapting to climate change is a strategic substitute to pollution abatement and, therefore, the current decision-makers benefit from taxing such investments.

Needless to say, the above analysis is only a start and the preliminary findings are suggestive. While the model of investments in capital permitted a relatively general production- or utility function, one may learn more about the exact subsidy- or tax-rate by specifying the functional forms. More importantly, the analysis of multiple investment levels (where some types of technologies could be used as inputs when making downstream investments) was based on a particular example and it is important to analyze the extent to which the results generalize.

## References

Ainslie, G. (1992): "Picoeconomics: The strategic interaction of successive motivational states, " Cambridge University Press.

Angeletos, G.-M. D. Laibson, A. Repetto, J. Tobacman, S. Weinberg (2001): "The hyperbolic consumption model: calibration," J. Econ. Perspect. 15: 47–68.

Barro, R. (1974): "Are government bonds net wealth?," J. Polit. Economy 82: 1095–1117.

Benhabib, Jess; Bisin, Alberto; Schotter, Andrew (2010): "Present-bias, quasi-hyperbolic discounting, and fixed costs," Games and Economic Behavior (2): 205-223.

Caplin, Andrew; Leahy, John (2004): "The Social Discount Rate," Journal of Political Economy 112(6): 1257-1268.

Dasgupta, Partha, and Eric Maskin (2005): "Uncertainty and Hyperbolic Discounting." American Economic Review, 95(4): 1290-1299.

DellaVigna,Stefano; Malmendier , Ulrike (2006). "Paying Not to Go to the Gym," American Economic Review 96(3): 694-719.

Diamond, P. and Koszegi, B. (2003): "Quasi-Hyperbolic Discounting and Retirement," Journal of Public Economics 87(9-10): 1839-72.

Eisenhauer, Joseph G.; Ventura, Luigi (2006): "The prevalence of hyperbolic discounting: some European evidence," Applied Economics 38(11): 1223-1234.

Frederick, Shane; Loewenstein, George; O'Donoghue, Ted (2002): "Time Discounting and Time Preference," Journal of Economic Literature 40(2): 351-401.

Gerlagh, Reyer and Liski, Matti (2013): "Carbon prices for the next thousand years," mimeo.

Gollier, Christian; Weitzman, Martin L. (2010): "How should the distant future be discounted when discount rates are uncertain?," Economics Letters 107(3): 350-353.

Gruber, Jonathan; Köszegi, Botond (2001): "Is Addiction "Rational"? Theory and Evidence," The Quarterly Journal of Economics 116(4): 1261-1305.

Harstad, Bård (1999): "Intertemporal preferences and optimal capital management," MS thesis, University of Oslo (May).

Karp, Larry (2005): "Global Warming and Hyperbolic Discounting," Journal of Public Economics 89: 261-82.

Kirby, Kris N.; Herrnstein, R. J. (1995): "Preference Reversals Due to Myopic Discounting of Delayed Reward," Psychological Science 6(2): 83-89.

Koopman, Tjalling (1960): "Stationary ordinal utility and impatience," Econometrica 28(2): 287-309.

Krusell, P.;Kuruşçu, B; Smith, A. A. JR. (2009): "How Much Can Taxation Alleviate Temptation and Self-Control Problems?, " Manuscript.

Krusell, P.;Kuruşçu, B; Smith, A. A. JR. (2010): " Temptation and Taxation, " Econometrica 78(6): 2063-2084.

Krusell, Per; Smith, A. A. JR. (2003): "Consumption-Savings Decisions with Quasi-geometric Discounting, " Econometrica 71(1): 365-375.

Laibson, D. (1997): "Golden eggs and hyperbolic discounting, " The Quarterly Journal of Economics 112(2): 443-478.

Laibson, D. and C. Harris (2001): "Dynamic choices of hyperbolic consumers," Econometrica 69: 935–57.

Laibson, D.; Repetto; A.; Tobacman, J. (2007): "Estimating Discount Functions with Consumption Choices over the Lifecycle, " University of Oxford Economics Series Working Papers 341.

O'Donoghue, Ted, and Matthew Rabin (1999): "Procrastination in Preparing for Retirement." In Henry Aaron, ed., Behavioral Dimensions of Retirement Economics. Washington, D.C.: Brookings Institution.

Paserman, Daniele (2008): "Job Search and Hyperbolic Discounting: Structural Estimation and Policy Evaluation," Economic Journal, Royal Economic Society, vol. 118(531): 1418-1452.

Phelps, E.S., Pollak, R.A. (1968): " On second-best national saving and game-equilibrium growth, " The Review of Economic Studies 35: 165-199.

Ramsey, Frank P. (1928): "A mathematical theory of saving," The Economic Journal, Vol. 38, No. 152: 543-559.

Saez-Marti, Maria; Weibull, Jörgen W. (2003): "Discounting and altruism to future

decision-makers, " Journal of Economic Theory 122: 254-266.

Samuelson, Paul (1937): "A note on measurement of utility," Review of Economic Studies 4: 155–161.

Scharff, Robert L. (2009): "Obesity and Hyperbolic Discounting: Evidence and Implications," Journal of Consumer Policy 32(1): 3-21.

Strotz, R.H. (1956): " Myopia and inconsistency in dynamic utility maximization, " The Review of Econonomic Studies 23: 166-180.

Thaler, R.H. (1981): " Some empirical evidence on dynamic inconsistency, " Economic Letters 8(3): 201-207.

Thaler, R. H. and Benartzi, Shlomo (2004): "Save more tomorrow: using behavioral economics to increase employee saving," Journal of Political Economy 112(1): 164-187.

Viscusi, W. Kip and Huber, Joel (2006): "Hyperbolic discounting of public goods," NBER Wp 11935.

Weitzman, Martin L. (2001): "Gamma Discounting," American Economic Review 91(1): 260-271.