Working Paper 21/2015

Investment Policy for Time-Inconsistent Discounters

Bård Harstad

The CREE Centre acknowledges financial support from The Research Council of Norway, University of Oslo and user partners.
ISBN: 978-82-7988-216-9
ISSN: 1892-9680
http://www.cree.uio.no

Abstract in Norwegian:

CREE - Working Paper 21/2015

Investeringspolitikk og såkallt hyperbolsk diskontering

Bård Harstad

Denne artikkelen kombinerer adferdsøkonomi og miljøøkonomi for å studere hva slags politikk myndighetene kan føre dersom politikere/velgere har såkalte tids-inkonsistente preferanser. Slike preferanser betyr typisk at når man sammenligner fremtidige goder/kostnader med kostnader/goder som blir realisert i dag, så vil man legge langt mer vekt på nåtid enn framtid enn det man ville gjort på et tidligere tidspunkt der begge de to datoene låg langt fram i tid. Med slike preferanser vil man for eksempel ha et ønske om å føre en bærekraftig politikk med lave utslipp i nær framtid, la oss si om 10-20 år, siden dette vil være viktig for miljø og klima om 100-200 år. Men når de 10-20 neste årene har gått vil man heller prioritere det som da er nåtid, og utsette den bærekraftige politikken.

Når denne stadige utsettelsen tas med i beregningen vil det være optimalt med en politikk i dag som på en eller annen måte binder framtidige politikere til å velge en mer bærekraftig politikk. En slik binding kan oppnås gjennom teknologi. Dersom man nå investerer i såkalt grønn teknologi, som fornybare energikilder eller renseteknologi, så vil framtidige beslutningstakere selv finne det optimalt å kutte på utslipp da, siden kostnaden har blitt redusert ved hjelp av denne teknologien. Bedrifter og private investorer vil ikke ta hensyn til denne ekstragevinsten ved å binde opp framtidige beslutningstakere, så det vil være optimalt for dagens politikere å subsidiere slike investeringer---på tross av at det ikke finnes noen spillover-effekter i resonnementet ovenfor. Med andre ord vil tidsinkonsistende preferanser kunne forsvare subsidier til miljøvennlig teknologi utover det tradisjonelle argumentet som viser til eksternaliteter og spillover-effekter.

Argumentet blir motsatt når det gjelder såkalt «brun teknologi» som er komplementært med å forurense eller forbruke/utvinne fossil energi: Slik teknologi må skattlegges dersom dagens politikere skal kunne påvirke framtidige beslutningstakere i bærekraftig retning.

Artikkelen studerer også flere nivåer av teknologiske investeringer, og det vises at for såkalt grønn teknologi så vil optimale subsidier være høyere jo mer «fundamental» teknologien er. Grunnforskning burde subsidieres mer enn selve innstalleringen av vindmøller, for eksempel, siden mer grunnleggende teknologi/forskning påvirker framtidige beslutninger gjennom et større antall ledd (grunnforskning påvirker både investeringer i fornybar energi så vel som innstalleringen og den senere bruken av disse).

Artikkelen er nå sendt inn til vurdering i et internasjonalt tidsskrift.

Investment Policy for Time-Inconsistent Discounters*

Bård Harstad

October 8th, 2015

Abstract

Standard analyses of economic policy assume exponential discounting, even though empirical and experimental evidence shows that preferences are time-inconsistent and discounting is hyperbolic. When policy makers—or the voters they must satisfy—apply smaller discount rates for long-term than for short-term decisions, they benefit from investing in infrastructure and technologies that will influence future decisions. This paper analyzes the equilibrium investment strategy and policy as a function of the technology's type and position in the production chain. The strategic concern can be measured by the subsidy a sophisticated decision maker would impose on a naive agent, or on a perfect market. Two main results are provided. First, I derive a formula for how the optimal investment subsidy depends on the investment lags and the technology's complementarity with future investments. When applied to climate change, it implies that investments in "green" technology should be subsidized while adaptation and "brown" technology should be taxed, even when laissez faire is first best under exponential discounting. Second, I show that fundamental technologies (i.e., those further upstream in the production chain) should be invested in and subsidized to a larger extent. This result also reveals that quasi-hyperbolic discounting is a poor approximation for strictly decreasing discount rates.

Key words: Time-inconsistency, hyperbolic discounting, quasi-hyperbolic discounting, green technology, production chain, research and development.

*I have benefitted from the comments of Geir Asheim, Christian Gollier, Matti Liski, Paolo Piacquadio, Alessia Russo, Tony Smith, Daniel Spiro, and several seminar audiences. Please contact the author at bard.harstad@econ.uio.no

The author have been associated with CREE – Oslo Centre for Research on Environmentally Friendly Energy. CREE acknowledges financial support from the Research Council of Norway, University of Oslo and user partners.

1. Introduction

The right way is to adopt policies that spur investment in the new technologies needed to reduce greenhouse gas emissions more cost effectively in the longer term without placing unreasonable burdens on American consumers and workers in the short term.

President Bush's Speech on Climate Change, April 16, 2008

Cutting emissions today in order to improve the future quality of life is the wrong way of approaching climate change, according to President Bush's 2008 speech. The right way was to invest in technology that could be used to cut tomorrow's emissions instead.[1]

Many projects generate costs and benefits for future years and generations. Reducing emissions 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 generate future benefits; and research is costly today but creates knowledge we can build on later.

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 philosophers as well as economists have struggled with it for centuries.

Over the last decades, our profession has settled on employing exponential discounting—not because of its normative or positive justifications—but due to its elegance, tractability, and resemblance to private investors' present-discounted value formula. Furthermore, exponential discounting leads to stationary or time-consistent preferences. However, apart from the tractability of exponential discounting, there are few reasons to impose it as a reasonable model of individual or political behavior. The lack of empirical and theoretical foundations for exponential discounting will be reviewed in the next section. That review supports the conclusion reached by Frederick at al. (2002:361) that "the collective evi-

[1] In his 2008 speech, President Bush also said: "there is a wrong way and a right way to approach reducing greenhouse gas emissions.... The wrong way is to...demand sudden and drastic emissions cuts that have no chance of being realized and every chance of hurting our economy. The right way is to set realistic goals for reducing emissions consistent with advances in technology."

dence...seems overwhelmingly to support hyperbolic discounting." When individuals and citizens have hyperbolic discount rates, then policy makers will also behave as they do, both because they are such individuals themselves, and also because they are accountable to voters with hyperbolic discount rates.

The purpose of this paper is to analyze the implications of time-inconsistent preferences for technology choices, investments, and investment policies. Any action today, whether it concerns investments in technology, capital, or knowledge, will inevitably affect future investment decisions. The current sophisticated decision maker will thus have an incentive to distort current investments in order to influence the choices made by her future self. The optimal strategic distortion, which can be implemented or measured by a subsidy or a tax on investments, will depend on the type of capital or technology to be invested in, and its position in the production chain. The analysis derives two main results.

First, I show how investments in technology and capital that are complementary to future investments should be subsidized, and how investments in strategic substitutes for future investments should be taxed. An important policy implication is that so-called "green" technology (which reduces the cost of pollution abatement) should be subsidized, while so-called "brown" technology (e.g., drilling technology or investments in fossil-fuel-dependent industries) should be taxed. This result holds even if we abstract from public good problems, externalities, or technological spillovers.

Second, the investment policy also depends on the technology's position in the production hierarchy. If technologies are strategic complements, technologies that are further upstream should be subsidized at a higher rate because they will have a multiplicative impact on the subsequent steps in the production chain. In other words, basic research should be subsidized at a higher rate than should investments in infrastructure, for example.

These results hinge on the discount factors in interesting ways. Under exponential discounting, the optimal subsidies are *always* zero (this will follow from the envelope theorem). Although exponential discounting leads to laissez faire as the normative policy recommendation (if there is no market failure), this conclusion fails when discount rates depend on the time horizon. Furthermore, the second result does *not* hold under quasi-hyperbolic discounting—which is therefore a poor approximation for hyperbolic discounting.

The next section discusses the background and the literature on discounting, including its foundations, empirical evidence, and critiques. Section 3 presents a simple model which describes how the optimal policy depends on the *type* of technology (e.g., green vs. brown technology), while Section 4, which contains most of the propositions, reveals how the investment policy also depends on the technology's position in the production chain. Section 5 concludes.

2. Background and Literature

In the nineteenth century, the debate regarding how to evaluate future utility losses and gains included a large number of factors, some psychological and many of them were conflicting (Rae, 1834; Senior, 1836; Jevons, 1871; and Böhm-Bawerk, 1889). Ramsey (1928) suggested maximizing a weighted sum of future utilities,

$$v_t = \sum_{\tau=t}^{\infty} D(\tau - t)\, u_\tau,$$

where $D(0) = 1$ and $D(\tau)$ measures the weight of utility u_τ, τ periods ahead, relative to utility right now. Although the discount factor $D(\tau)$ was left unspecified, Paul Samuelson (1937) suggested the now familiar formula for exponential discounting:

$$D(t) = \delta^t = \left(\frac{1}{1+\rho}\right)^t \approx e^{-\rho t},$$

where δ is the corresponding constant discount factor between subsequent periods and ρ is the constant discount rate. With Koopman's (1960) axiomatic foundation, exponential discounting became the standard way of evaluating future gains and losses in economics.

To many, the appeal of exponential discounting is not that its assumptions regarding individual behavior are reasonable but that it simplifies the analysis.[2] In a seminal paper,

[2] Paul Samuelson himself had reservations when suggesting the exponential 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 advice for a public planner ("any connection between utility as discussed here and any welfare concept is disavowed," p. 161). Nevertheless, and "despite Samuelson's manifest reservations, the simplicity and elegance of this [exponential] formulation was irresistible" according to Frederick et al. (2002: 355-6).

Strotz (1955-1956) explained why preferences are likely to be time-inconsistent and that we, as a consequence, had to search for the best plan that would actually be followed. Since then, we have seen an explosion of empirical and experimental evidence which "seems overwhelmingly to support hyperbolic discounting."[3] After all, our basic human senses perceive *relative* differences: When two sound sources are both located nearby, the closer sound is easier to hear; when they are both further away, it is easier to hear the louder sound. From a distance, the larger of two mountains does indeed look larger than the smaller one, but when they are both near, the mountain that looks larger is the one that is closer. If our sense for time has the same characteristic, as experimental evidence suggests that it does,[4] then only the relative difference (the difference between t and, say, t', relative to t) will be important. In this case, the discount factor must be "hyperbolic" in that utility at time t will be weighted by the discount factor:

$$D(t) = \frac{1}{1+\alpha t}, \tag{2.1}$$

where $\alpha > 0$ is a constant that can measure either impatience or the scale of time. With this, the discount factor between period t and $t-1$ is:

$$\delta_t \equiv \frac{D(t)}{D(t-1)} = 1 - \frac{\alpha}{1+\alpha t},$$

which is concave and increasing in t, and approaching one as t grows.

David Laibson (1997) adopted a simpler approximation of (2.1), 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 welfare at time t is:

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

[3] The quote is from the survey by Frederick et al. (2002: 361). For empirical evidence, see the survey by Angeletos et al. (2001), or more recent research such as Paserman (2008), who find estimates of the short-run annualized discount rate that range from 11 to 91 percent and a long-run discount rate of only 0.1 percent. Laibson et al. (2007) find that the short-term discount rate is 15 percent but the long-term discount rate is 3.8 percent. In lab experiments, individuals often prefer a smaller benefit today to a larger benefit tomorrow, but reverse the ranking if the two consecutive days are further into the future. See, for example, Thaler (1981), Ainslie (1992), Benhabib et al. (2010), or Halevy (2015).

[4] There are also evolutionary arguments suggesting that humans may evolve and survive better if they have so-called hyperbolic discounting functions (Dasgupta and Maskin, 2005).

Other names for this formula are (β, δ)-discounting, quasi-geometric discounting, quasi-exponential discounting, and sometimes, simply but misleadingly, hyperbolic discounting.

Although individuals apparently apply discount rates that decline in time, does this imply that governments ought to do the same? There are four reasons for an affirmative answer. First, the government consists of individual policy makers who share these preferences regarding the future, so it is inevitable that policy-makers *will* act in a time-inconsistent way. Second, to be re-elected, the government *must* be accountable and apply the same discount rates as the voters.[5] Third, one can argue that a government *should*—also from a normative perspective—discount future utility by using a discount factor that increases in relative time: the extent to which future generations are important, also morally, is already taken into account by the voters (Galperti and Strulovici, 2015). In fact, the formula for quasi-hyperbolic discounting, $D(t) = \beta\delta^t$, was first suggested by Phelps and Pollak (1968), who argued that it may represent "imperfect altruism" between generations. More generally, time-inconsistent preferences follow naturally when the current generation cares about the grandchildren's welfare in addition to the children's.[6] Finally, even if each individual had time-consistent preferences, collective decisions would necessarily be time-inconsistent as long as the discount rates differed among the individuals (Gollier and Zeckhauser, 2005; Jackson and Yariv, 2014; 2015).

There is a growing literature on policies in the presence of time inconsistency. For example, hyperbolic discounters may retire too early (Diamond and Kőszegi, 2003), or save too little, so the government can help the decision makers to commit by subsidizing saving (Krusell et al., 2009 and 2010).[7] Or, since hyperbolic discounters find it hard

[5] However, citizens may prefer that the government apply a lower discount rate than the citizens themselves would (Caplin and Leahy, 2004).

[6] If parents are "thoughtful" (as in Barro, 1974), then the welfare of a generation is a weighted sum of its own utility and the next generation's welfare. We can then write welfare recursively as a weighted sum of all future utilities, and the discount factor will be constant over time (leading to exponential discounting). However, if today's parents also care about the welfare of its grandchildren, then stationarity will be violated and the effective discount rate will indeed decline in time (Harstad, 1999; Saez-Marti and Weibull, 2005; Galperti and Strulovici, 2015). Note that the fact that the pure time preference rate depends on the time horizon is orthogonal to the arguments by Gollier and Weitzman (2010) and Weitzman (2001), who have shown that if the growth rate of consumption is uncertain, then it is optimal to discount future consumption at a rate that is decreasing in time in order to reflect risk aversion and the accelerating level of risk.

[7] See also Laibson and Harris (2001). In Bisin et al. (2015), it would be optimal to ban illiquid assets or require balanced budget rules. On climate change, see Karp (2005), or Gerlagh and Liski (2013), who

to quit smoking, the government could tax tobacco more (Gruber and Kőszegi, 2001). But also individuals may try to commit their future selves by limiting the future choice set (Gul and Pesendorfer, 2001), by signing up for saving plans which are costly to end (Thaler and Benartzi, 2004), or by paying today the cost of attending the gym tomorrow (DellaVigna and Malmendier, 2006).[8]

The present paper does not allow for any such pre-commitment to future policies. When I refer to a subsidy, it is to one that is set today by the decision maker of today as a simple way to account for how today's investment will influence future investments, and as a measure of this strategic concern. More importantly, I allow for a general class of technology, and focus on how the *type* of that technology and its *position* in the production chain determine the optimal (and equilibrium) investment strategy and policy. By allowing discount factors to depend on time in a general way, the model encompasses exponential discounting, hyperbolic discounting, and quasi-hyperbolic discounting as special cases.

3. Preliminary Results: Investments in Capital

3.1. Notation and Measures of Strategic Investments

Consider a single planner or decision maker playing a dynamic game against her future self. If u_t measures the momentary utility t periods from now, the objective today is to maximize $v_0 \equiv \sum_{t=0}^{\infty} D(t)\, u_t$, where $D(0) = 1$, while $D(t)$ measures the weight on utility at time $t \geq 0$ compared to utility today. The discount factor between period $t-1$ and t is

$$\delta_t \equiv \frac{D(t)}{D(t-1)} \Leftrightarrow D(t) = \prod_{\tau=1}^{t} \delta_\tau .$$

I will assume that $\delta_t \in (0,1)$ is strictly increasing in t unless otherwise stated. (For example, $\delta_t = \delta$ is constant when considering exponential discounting.)

It is obvious that any action that increases every u_t will be taken. The interesting decisions are those that require the decision maker to trade off future gains against current

derive the optimal carbon tariffs in a setting with quasi-hyperbolic discounting.

[8] Nonsophisticated hyperbolic discounters may also be taken advantage of in the market: see Heidhues and Kőszegi (2010) for an analysis of the credit card market, or, for a more general survey, Kőszegi (2014).

losses or, equivalently, vice versa. If an action a is costly today, it may nevertheless be worthwhile if it increases future utility. If we assume differentiable utility functions, the necessary first-order condition for an optimal interior a is:

$$-\frac{du_0}{da} = \frac{d}{da}\sum_{t=1}^{\infty} D(t)\, u_t. \tag{3.1}$$

If other decisions or investments might be made in the future, it is useful to distinguish between the *total* derivatives and the *partial* derivatives. The total derivative $d(\cdot)/da$ in (3.1) reflects the fact that when taking decision a, a sophisticated decision maker may take into account the fact that the choice of a today may influence other, future choices, that may in turn also influence utilities. If, in contrast, the decision maker did not try to influence future choices, then the choice of a would instead solve:

$$-\frac{du_0}{da} = \frac{\partial}{\partial a}\sum_{t=1}^{\infty} D(t)\, u_t. \tag{3.2}$$

Of course, if the decision maker were time-consistent, then (3.1) and (3.2) would be equivalent, since future choices would be optimal also from today's point of view, and thus there would be no reason to account for the fact that a will influence these future choices (this follows from the envelope theorem). But when preferences are time-inconsistent, then we can measure the strategic consideration when choosing a in the following way:

$$s^* \equiv \frac{\sum_{t=1}^{\infty} D(t)\, du_t/da}{\sum_{t=1}^{\infty} D(t)\, \partial u_t/\partial a} - 1. \tag{3.3}$$

That is, when $s^* > 0$, the investment level that is chosen according to (3.1) is strategically large when the decision maker takes into account the fact that a influences future choices. If $s^* < 0$, the investments are instead strategically small when the effect on future decisions is taken into account. In either case, s^* measures the extent to which the optimal choice of a is distorted *because* of the decision maker's desire to influence future decisions, i.e., because of the time-inconsistent preference.

The starred superscript reflects the interesting feature that s^* can also be interpreted as the *optimal subsidy* if the actual investment is decentralized to a "naive agent," or, a "perfect market," as long as the decision maker's discount factors are shared. To see this, note that a naive agent, with no desire to influence future choices, would invest according to (3.2) in the absence of any subsidy.[9] Alternatively, (3.2) would measure investments in a "perfect market," defined as a market in which investors obtain full property rights to the direct revenues of their investments.[10] If there were a large number of price-taking investors, they would take as given the future price $\partial u_t/\partial a$ and discount those revenues by $D(t)$; if there were a single investor able to price-discriminate perfectly, the investor would capture the marginal revenue $\partial u_t/\partial a$ for every t and discount them by $D(t)$. However, if the investment cost were subsidized by $\underline{s}$, then either the market or the naive agent would invest according to:

$$\begin{aligned} -\left(1-\underline{s}\right)\frac{du_0}{da} &= \frac{\partial}{\partial a}\sum_{t=1}^{\infty} D(t)\,u_t \Leftrightarrow \\ -\frac{du_0}{da} &= (1+s)\,\frac{\partial}{\partial a}\sum_{t=1}^{\infty} D(t)\,u_t, \text{ if} \\ s &\equiv \frac{1}{1-\underline{s}}-1. \end{aligned} \tag{3.4}$$

Here, s is equivalent to a subsidy on future revenues. Obviously, a subsidy on investment costs is equivalent to a subsidy on future revenues. Also, we can let an investment-cost subsidy $\underline{s}$ be *measured* by $s \equiv 1/\left(1-\underline{s}\right)-1$, so that we can write the equilibrium condition as (3.4). The decision maker of today can implement her preferred a by ensuring that (3.4) coincides with (3.1). This requires $s = s^*$, as it is given by (3.3). Furthermore, this choice of s^* is preferred by the decision maker if she considers the subsidies to be simply transfers within the society with no real cost, except for the fact that the subsidy may affect the choice of a.

LEMMA 1: *If the investment is made in a perfect market, or by a naive agent, the decision maker implements her preferred decision with the subsidy s^* given by (3.3).*

REMARK 1: *On commitment.* Note that there is no commitment to any future subsidies. Instead of setting s^*, the decision maker can implement the same a with the

[9] A naive agent is, in contrast to the sophisticated decision maker, not aware of the time inconsistency problem, and will thus (by the envelope theorem) not seek to influence any future decision. The distinction between naive and sophisticated is further discussed by O'Donoghue and Rabin (1999).

[10] A complete set of competitive markets would be perfect in this sense. In fact, the statements referring to a perfect market continue to hold as long as each investor is the full residual claimant to all direct costs or benefits of her investment. With exponential discounting, the first welfare theorem implies that the market equilibrium would be first best and there would be no need for any regulation.

corresponding investment-cost subsidy $\underline{s} = 1 - 1/(1 + s^*)$. This subsidy is set today and I assume it is impossible to commit to any *future* subsidies or policies. The only way to partially commit is to take today's decision a in such a way as to influence future choices. Whether the decision maker sets a directly or by regulating the market or a naive agent, we can let s^* measure the equilibrium level of a and how it differs from the choice of a in the absence of any strategic considerations.

3.2. A Simple Investment

To illustrate the notation and derive a benchmark comparison, consider a simple once-and-for-all investment or action a generating a future benefit $b(a)$ at the cost $c(a)$ today. If the benefit is realized Δ_a periods from now, it is discounted by $D(\Delta_a)$. Thus, a decision maker maximizes $v = -c(a) + D(\Delta_a) b(a)$. The necessary first-order condition is:

$$c'(a) \equiv \frac{dc}{da} = D(\Delta_a) \frac{db}{da}. \tag{3.5}$$

To ensure that we have interior solutions and that the second-order condition is satisfied, let $c(\cdot)$ be increasing and convex and $b(\cdot)$ increasing and concave.

As a comparison, investors in a perfect market can invest today and earn the marginal revenue $\partial b(a)/\partial a$ tomorrow. With the subsidy s_a, the first-order condition is:

$$\frac{c'(a)}{1 + s_a} = D(\Delta_a) \frac{\partial b(a)}{\partial a}. \tag{3.6}$$

A naive agent would invest in the same way.

As long as a does not influence any other future choices, then $db/da = \partial b/\partial a$, so (3.5) and (3.6) are equivalent for $s_a = 0$. A naive agent or investors in a perfect market are making the same decision that the decision maker would, so laissez faire works perfectly fine.

Proposition 0: *For the simple investment or action a, the decision maker invests according to (3.5), or, equivalently, by satisfying (3.6) with s_a given by*

$$s_a^* = 0.$$

3.3. Investments in Capital

The investment or action a can have a large number of interpretations. The investment can be in health, education, infrastructure, or pollution abatement, to mention some examples. For some investments, it is reasonable that the *cost* of investing depends on the level of capital or infrastructure. In other cases, it may be the later *benefit* of a that depends on the level of capital or infrastructure. To capture the importance of such capital, k, let the cost of investing a be written as $c(a;k)$ and the benefit as $b(a,k)$.[11]

To better explain and motivate the importance of k, it is useful to revert to the application in which a measures pollution abatement. For this application, k may represent 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, $\partial c(a;k)/\partial a > 0$ decreases in k, implying that $\partial^2 c(a;k)/\partial a \partial k < 0$. The green capital does not (by assumption) affect the environmental harm directly, so $\partial b(a,k)/\partial k = 0$.

Brown capital refers to drilling technologies or investments in industries that pollute. Such capital may be beneficial in the sense that it increases the utility ($\partial c(a;k)/\partial k < 0$), but a larger level of k also makes it costly to cut back on pollution. Thus, $\partial^2 c(a;k)/\partial a \partial k > 0$, meaning that a and k are strategic substitutes. The brown capital does not (by assumption) affect the environmental harm directly, so again $\partial b(a,k)/\partial k = 0$.

Adaptation capital refers to investments that enhance the economy's ability to deal with pollution. 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 capital not only increases the future benefit $b(a,k)$, but also reduces the marginal environmental harm; in other words, a larger level of k reduces the value of a so that $\partial^2 b(a,k)/\partial a \partial k < 0$. Such adaptation capital does not (by assumption) affect the cost of abating, so $\partial c(a;k)/\partial k = 0$.

[11] The use of semicolon in $c(a;k)$ reflects the fact that k is sunk when a is chosen. When the benefit $b(a,k)$ is experienced, both variables are sunk.

The level of k is given when a is decided upon. By differentiating (3.5), we can see how the decision on a varies with the level of k:

$$\begin{aligned}
\frac{da}{dk} &= \frac{D(\Delta_a)\,\partial^2 b/\partial a \partial k - \partial^2 c/\partial a \partial k}{\partial^2 c/(\partial a)^2 - D(\Delta_a)\,\partial^2 b/(\partial a)^2} \Rightarrow \qquad (3.7)\\
\text{sign}\left(\frac{da}{dk}\right) &= \text{sign}\left(D(\Delta_a)\frac{\partial^2 b(a,k)}{\partial a \partial k} - \frac{\partial^2 c(a;k)}{\partial a \partial k}\right).
\end{aligned}$$

Thus, $da/dk > 0$ for green capital, while $da/dk < 0$ for adaptation and brown capital.

Let Δ_k measure the number of periods between the (one-shot) decision on k and the decision on a. That is, Δ_k may be the time it takes for the capital to be ready or built. Further, let $c_k(k)$ be the cost of k. When k is decided upon, the decision maker takes into account that the level of k affects c and b not only directly, but also indirectly through the choice of a.

As a comparison, the effect on a would not be taken into account by a naive agent or by a perfect market investing in k. In these cases, the choice of k would satisfy the following first-order condition:

$$\frac{c_k'(k)}{1+s_k} = -D(\Delta_k)\frac{\partial c(a;k)}{\partial k} + D(\Delta_k + \Delta_a)\frac{\partial b(a,k)}{\partial k}, \qquad (3.8)$$

where s_k represents a subsidy on k. The decision maker can implement her preferred level of k by setting the appropriate s_k. Even when the decision maker decides on k directly, there exists some s_k, referred to as s_k^*, such that the decision maker's preferred level of k satisfies (3.8) when $s_k = s_k^*$. So, as mentioned above, s_k^* can measure how much the decision maker invests in k relative to a situation with time consistency, a naive agent, or a perfect market.

Proposition 1: *The equilibrium capital level k satisfies (3.8) with s_k given by:*

$$s_k^* = \left(\prod_{t=0}^{\Delta_a}\frac{\delta_{t+\Delta_k}}{\delta_t} - 1\right)\frac{\partial c(a;k)/\partial a}{-\partial c(a;k)/\partial k + (D(\Delta_k+\Delta_a)/D(\Delta_k))\,\partial b(a;k)/\partial k}\cdot\frac{da}{dk} \qquad (3.9)$$

Proof: The decision maker's objective is to maximize $v_k(k) \equiv -c_k(k) - D(\Delta_k)\,c(a;k) + D(\Delta_k+\Delta_a)\,b(a,k)$. By taking the total derivative w.r.t. k, we get the first-order condition:

$$\begin{aligned} v_k'(k) &\equiv -c_k'(k) - D(\Delta_k)\left[\frac{\partial c(a;k)}{\partial k} + \frac{\partial c(a;k)}{\partial a}\frac{da}{dk}\right] \\ &+ D(\Delta_k + \Delta_a)\left[\frac{\partial b(a,k)}{\partial k} + \frac{\partial b(a,k)}{\partial a}\frac{da}{dk}\right] = 0 \Leftrightarrow \\ c_k'(k) &= -D(\Delta_k)\frac{\partial c(a;k)}{\partial k} + D(\Delta_k + \Delta_a)\frac{\partial b(a,k)}{\partial k} \\ &\quad + \left[D(\Delta_k + \Delta_a)\frac{\partial b(a,k)}{\partial a} - D(\Delta_k)\frac{\partial c(a;k)}{\partial a}\right]\frac{da}{dk}. \end{aligned}$$

When substituting in for (3.5), we get:

$$c_k'(k) = -D(\Delta_k)\frac{\partial c(a;k)}{\partial k} + D(\Delta_k + \Delta_a)\frac{\partial b(a,k)}{\partial k} + \left[\frac{D(\Delta_k + \Delta_a)}{D(\Delta_a)} - D(\Delta_k)\right]\frac{\partial c(a;k)}{\partial a}\frac{da}{dk},$$

and when also (3.8) must hold, s_k must be given by (3.9). The second-order condition $v_k''(k) < 0$ holds when $c_k(k)$ is sufficiently convex (see the proof of Proposition 2). *Q.E.D.*

The contribution of Proposition 1 is best illustrated by stating a number of corollaries.

COROLLARY 1:

(i) *With exponential discounting,* $s_k^* = 0$.

(ii) *If either* $\Delta_k = 0$ *or* $\Delta_a = 0$, $s_k^* = 0$.

(iii) *Suppose* $\Delta_k \Delta_a > 0$. *With quasi-hyperbolic discounting, (3.9) simplifies to:*

$$s_k^* = \left(\frac{1}{\beta} - 1\right)\frac{\partial c(a;k)/\partial a}{-\partial c(a;k)/\partial k + \delta^{\Delta_a}\partial b(a;k)/\partial k} \cdot \frac{da}{dk}.$$

(iv) *It is optimal to subsidize investments in green capital:*

$$s_k^* = \left(\prod_{t=0}^{\Delta_a}\frac{\delta_{t+\Delta_k}}{\delta_t} - 1\right) \cdot \frac{\partial c(a;k)/\partial a}{-\partial c(a;k)/\partial k} \cdot \frac{da}{dk} > 0.$$

(v) *It is optimal to tax investments in brown capital:*

$$s_k^* = \left(\prod_{t=0}^{\Delta_a}\frac{\delta_{t+\Delta_k}}{\delta_t} - 1\right) \cdot \frac{\partial c(a;k)/\partial a}{-\partial c(a;k)/\partial k} \cdot \frac{da}{dk} < 0.$$

(vi) *It is optimal to tax investments in adaptation capital:*

$$s_k^* = \left(1 - \prod_{t=0}^{\Delta_a}\frac{\delta_t}{\delta_{t+\Delta_k}}\right) \cdot \frac{\partial b(a;k)/\partial a}{\partial b(a;k)/\partial k} \cdot \frac{da}{dk} < 0.$$

(vii) *For green, brown, or adaptation capital,* $|s_k^*|$ *increases in* Δ_k *and in* Δ_a*. For any given sum* $\Delta_k + \Delta_a$*, if* δ_t *is concave in* t*,* $|s_k^*|$ *is maximized when* $\Delta_k = \Delta_a$*.*

The following numbered points discuss the corresponding parts of Corollary 1.

(i) In traditional settings where the decision maker has time-consistent preferences, there is no need today for the decision maker to distort the choices of her future self. So, if an investor captures the full future return of the investment, there is no need for regulation. This confirms Proposition 0, which suggested that laissez faire is just fine.

(ii) Furthermore, if $\Delta_k = 0$, it takes no time to build the capital. It is then the same decision maker selecting k and a and there is obviously no need to distort either decision. Alternatively, if $\Delta_a = 0$, the decision maker choosing a gets the benefit herself immediately and the level of a does not influence any future utility which the two selves would evaluate differently.

(iii) When $\Delta_k \Delta_a > 0$, assumed from now on, a time-inconsistent decision maker is not satisfied with the future choice of a. Today's decision maker would prefer a larger investment a than the level that will actually be implemented by her future self, and the choice of a can be influenced by s_k. In general, the disagreement between the two selves, and thus the optimal level of s_k, will depend on every relevant δ_t. With quasi-hyperbolic discounting, however, $\delta_t = \delta$ for $t > 1$ and the formula for s_k^* simplifies.

(iv) Regardless of whether discounting is quasi-hyperbolic, or whether the δ_t's are instead strictly increasing in t, it is optimal with $s_k^* > 0$ for so-called green capital. For this type of capital, a increases in k, and thus the decision maker prefers a strategically large k in order to motivate a larger a in the future. When the decision maker herself is deciding on k, then $s_k^* > 0$ has the interpretation that she prefers a larger k than she would have chosen either if she had taken the future choice of a as given, or if she did not want or try to influence a. If the choice of k is left to the perfect market or to a naive agent, $s_k^* > 0$ means that it is optimal to subsidize today's investment in k.

(v) For brown capital, a decreases in k. To motivate a larger a, which the decision maker would prefer, it is necessary to reduce the investment in k today. Thus, the decision maker benefits from investing strategically little in so-called brown capital, and she benefits from taxing these kinds of investments. Note that the level of s_k^* is always proportional to da/dk, has the same sign as da/dk, and is zero when $da/dk = 0$.

(vi) The result for adaptation may seem provocative. Adaptation can certainly be a good thing, in that it may be that $\partial b(a,k)/\partial k > 0$. However, even a naive agent or private investor will account for the value $\partial b(a,k)/\partial k$, so this creates no reason to strategically distort k. On the contrary, more investments in adaptation will reduce the cost of polluting, and the level of abatement will thus be reduced as well. The decision maker of today prefers a larger a in the future, and this can be achieved by strategically reducing the level of adaptation capital.

(vii) When discount factors are strictly increasing in relative time, the decision maker's disagreement with her future self is larger if the next decision is made at a much later point in time. Thus, the parenthesis in (3.9) is increasing in both investment lags. (This is not true for quasi-hyperbolic discounting, however, since the disagreement then is not increasing in the lags, and the parenthesis simplifies to $1/\beta - 1 > 0$.)

If δ_t is a concave function of t, the disagreements are increasing at a decreasing rate in t. Thus, conditional on the sum of the lags being the same, the optimal choice of $|s_k^*|$ is at the largest when the two lags are equal.

Remark 2: *Long-lasting stocks and investments in every period.* For simplicity, the decisions on a and k were treated above as one-shot decisions. It is straightforward, however, to let an action a_t and an investment k_t be decided on in *every* period t, if just the cost and benefit of a_t depend on an earlier choice of capital, say, k_{t-1}, and not on k_t. Then, the choice of a_t will still be given by (3.5) and da_t/dk_{t-1} by (3.7). Further, if a fraction $q \in [0,1]$ of k_t survives to the next period, the choice of k_t remains independent of k_{t-1} if the cost of upgrading qk_{t-1} to k_t is additively separable and given by $c_k(k_t) - h(qk_{t-1})$, for some function h. To see this, let $\widetilde{c}(a_t; k_{t-\tau})$ be the actual cost of action a_t, define $c(a_t; k_{t-\tau}) \equiv \widetilde{c}(a_t; k_{t-\tau}) - h(qk_{t-1})$, and note that the analysis stays unchanged. Remark 3 at the end of Section 4 discusses all this in more detail.

4. Main Results: Investments in Technologies

The previous section made a distinction between different *types* of investments at the same *stage* in the production chain: while some capital types were complementary to the abatement decision, others could be strategic substitutes. The type of capital turned out to be crucial for how the investments were strategically chosen so as to influence future decisions. Green capital should be subsidized, according to Proposition 1, while adaptation and brown capital should be taxed.

The second goal of this paper is to investigate how the strategic choice of investment (or subsidy) also depends on the *stage* in the production chain. For example, while a larger number of windmills will make it cheaper to reduce pollution, the production cost of each windmill will depend on the amount of technology, knowledge, or basic research. The fact that distinguishing between the stages may be important is evident when comparing the decision on capital (Proposition 1) to the downstream decision on, say, abatement (Proposition 0).

The first subsection below takes us another step upstream by analyzing investments in technology. The second subsection generalizes by investigating a production chain of arbitrary length and by showing how the strategic considerations (or the equilibrium subsidy) depends on the investment's location in the production chain. The final subsection discusses so-called "stepping stone technologies" and derives a simple formula for how such technologies are optimally chosen (or subsidized).

4.1. Investments in Technology

The production chain now has three stages. First, technology is invested upstream. Second, it is used to produce capital. And third, that capital is used to invest in future utility.

To recognize the similarity between the stages, I now switch notation by referring to k_1 instead of a, with $c_1(k_1; k_2)$ as the investment cost (instead of $c(k; a)$). Capital is referred to as k_2 (instead of simply k) and the capital investment cost is $c_2(k_2; k_3)$. The cost of investing in technology, k_3, is $c_3(k_3)$.

To focus on the chain, it is assumed that (i) $\Delta_k = \Delta_a = 1$, and (ii) while k_3 influences only the cost of investing in k_2, k_2 influences only the cost of investing in k_1 (thus, $\partial b / \partial k_2 = 0$, so we can write $b' \equiv \partial b / \partial a$). Both assumptions (i)-(ii) can easily be relaxed (see footnote 13).

If we rewrite Propositions 0 and 1 using the new notation, we get:

$$\begin{aligned} s_1^* &= 0, \text{ and} \\ s_2^* &= \left(\frac{\delta_2}{\delta_1} - 1\right)\left[-\frac{\partial c_1(k_1;k_2)/\partial k_1}{\partial c_1(k_1;k_2)/\partial k_2}\right]\frac{dk_1}{dk_2}, \text{ where} \\ \frac{dk_1}{dk_2} &= -\frac{\partial^2 c_1(k_1;k_2)}{\partial k_1 \partial k_2}\left(\frac{1}{\partial^2 c_1/(\partial k_1)^2 - \delta_1 \partial^2 b''}\right), \end{aligned}$$

and the term in the brackets is simply the slope of the iso-cost curve.

When deciding on k_3, a naive agent or the perfect market would invest as follows:

$$\frac{c_3'(k_3)}{1+s_3} = -\delta_1 \frac{\partial c_2(k_2;k_3)}{\partial k_3}. \tag{4.1}$$

With time-inconsistent preferences, today's decision maker is not satisfied with the future choices of k_2 and k_1 and, in order to influence these choices, it may be optimal to distort today's investments in k_3. To see how k_3 influences k_2, we can simply differentiate the first-order condition for k_2 to show that the cross-derivative is, again, crucial:

$$\frac{dk_2}{dk_3} = -\frac{\partial^2 c_2(k_2;k_3)}{\partial k_2 \partial k_3}\left(\frac{1}{-v_2''}\right),$$

where $v_2'' < 0$ is the second-order condition when k_2 is chosen.[12] The influence of k_3 on k_1 is given by the product of dk_2/dk_3 and dk_1/dk_2.

Just as in the previous section, we can measure the decision maker's decision on k_3, relative to her choice in the absence of the strategic concerns, by deriving the level of s_3 which would ensure that (4.1) is in line with the decision maker's preferred level.

PROPOSITION 2: *The equilibrium technology level k_3 satisfies (4.1) with s_k given by:*

$$s_3^* = \left(\frac{\delta_2}{\delta_1} - 1\right)\left[-\frac{\partial c_2(k_2;k_3)/\partial k_2}{\partial c_2(k_2;k_3)/\partial k_3}\right]\frac{dk_2}{dk_3} + \delta_2(\delta_3 - \delta_2)\left[-\frac{\partial b/\partial k_1}{\partial c_2(k_2;k_3)/\partial k_3}\right]\frac{dk_1}{dk_2}\frac{dk_2}{dk_3}. \tag{4.2}$$

PROOF: The decision maker prefers the k_3 solving the total derivative:

$$\begin{aligned} c_3'(k_3) &= -D(1)\frac{\partial c_2(k_2;k_3)}{\partial k_3} - D(1)\frac{\partial c_2(k_2;k_3)}{\partial k_2}\frac{dk_2}{dk_3} - D(2)\frac{\partial c_1(k_1;k_2)}{\partial k_2}\frac{dk_2}{dk_3} \\ &\quad -D(2)\frac{\partial c_1(k_1;k_2)}{\partial k_1}\frac{dk_1}{dk_2}\frac{dk_2}{dk_3} + D(3)\,b'\frac{dk_1}{dk_2}\frac{dk_2}{dk_3}. \end{aligned}$$

[12] From the proof of Proposition 1, we can simplify to:

$$v_2'' \equiv \frac{\partial v_2(k_2;k_3)}{(\partial k_2)^2} \equiv -\frac{\partial c_2(k_2;k_3)}{(\partial k_2)^2} - \delta_1\left[\frac{\partial c_1(k_1;k_2)}{\partial k_2} + \frac{\partial c_1(k_1;k_2)}{\partial k_1}\frac{dk_1}{dk_2}\right] + D(2)\left[b'\frac{dk_1}{dk_2}\right] < 0.$$

This can be combined with the first-order condition for k_1, $\partial c_1(k_1;k_2)/\partial k_1 = \delta_1 b'$, and the first-order condition for k_2,

$$\begin{aligned}
\frac{\partial c_2(k_2;k_3)}{\partial k_2} &= -D(1)\frac{\partial c_1(k_1;k_2)}{\partial k_2} - D(1)\frac{\partial c_1(k_1;k_2)}{\partial k_1}\frac{dk_1}{dk_2} + D(2)b'\frac{dk_1}{dk_2} \Rightarrow \\
\frac{\partial c_1(k_1;k_2)}{\partial k_2} &= -\frac{\partial c_2(k_2;k_3)}{D(1)\partial k_2} - \frac{\partial c_1(k_1;k_2)}{\partial k_1}\frac{dk_1}{dk_2} + \frac{D(2)}{D(1)}b'\frac{dk_1}{dk_2}.
\end{aligned}$$

in order to get:

$$\begin{aligned}
c_3'(k_3) &= -D(1)\frac{\partial c_2(k_2;k_3)}{\partial k_3} - D(1)\frac{\partial c_2(k_2;k_3)}{\partial k_2}\frac{dk_2}{dk_3} - D(2)\frac{\partial c_1(k_1;k_2)}{\partial k_1}\frac{dk_1}{dk_2}\frac{dk_2}{dk_3} \\
&\quad -D(2)\left[-\frac{\partial c_2(k_2;k_3)}{D(1)\partial k_2} - \frac{\partial c_1(k_1;k_2)}{\partial k_1}\frac{dk_1}{dk_2} + \frac{D(2)}{D(1)}b'\frac{dk_1}{dk_2}\right]\frac{dk_2}{dk_3} + D(3)b'\frac{dk_1}{dk_2}\frac{dk_2}{dk_3} \\
&= -\delta_1\frac{\partial c_2(k_2;k_3)}{\partial k_3} + (\delta_2-\delta_1)\frac{\partial c_2(k_2;k_3)}{\partial k_2}\frac{dk_2}{dk_3} + (\delta_3-\delta_2)D(2)b'\frac{dk_1}{dk_2}\frac{dk_2}{dk_3}.
\end{aligned}$$

Together with (4.1), we get (4.2). *Q.E.D.*

Note that the term in the first bracket in (4.2) is simply the slope of the iso-cost curve. The essence of the proposition is illustrated in the following corollary. To exemplify the result, it is natural to define "green technology" as technology that is complementary to the investment in green capital, and "brown technology" as technology that is complementary to the investment in brown capital.

Corollary 2:

(i) *With exponential discounting,* $s_3^* = 0$.

(ii) *With quasi-hyperbolic discounting, the second term in (4.2) is zero, so* s_3^* *can be written analogously to* s_2^*:

$$s_n^* = \left(\frac{\delta_2}{\delta_1} - 1\right)\left[-\frac{\partial c_{n-1}(k_{n-1};k_n)/\partial k_{n-1}}{\partial c_{n-1}(k_{n-1};k_n)/\partial k_n}\right]\frac{dk_{n-1}}{dk_n},\ n \in \{2,3\}.$$

(iii) *For green technology, both terms in (4.2) are positive, so* $s_3^* > 0$.

(iv) *For brown technology, the first term in (4.2) is positive, the second negative, and*

$$s_3^* > 0 \quad \textit{if and only if} \quad \frac{dk_1}{dk_2}(\delta_3-\delta_2) < -\frac{\partial c_2(k_2;k_3)/\partial k_2}{b'}\frac{\delta_2-\delta_1}{\delta_1\delta_2}.$$

Just as before, the contribution of the proposition is best illustrated by discussing its corollaries. The following four numbers refer to the corresponding numbered parts of Corollary 2:

(i) It is easy to check that with exponential discounting, both parts in (4.2) are zero. Intuitively, if the decision maker were time-consistent, she would be perfectly satisfied with her own future choices of k_2 and k_1. She would have no desire to distort these choices and thus (by the envelope theorem) she would prefer an amount of technology which took into account only the direct cost-savings. The perfect market would then invest optimally and there would be no need for regulation.

(ii) With time-inconsistent preferences, the decision maker disagrees with the future choice of k_2. Thus, k_3 may be chosen, or distorted, in order to influence and increase the investment in k_2. If the cross-derivative of $c_2(k_2; k_3)$ is negative, so that k_3 is a strategic complement to the investment in k_2, then the current decision maker has an incentive to invest strategically more in k_3 in order to motivate a larger investment in k_2. The optimal investment in k_3 is larger if the current decision maker disagrees strongly with her future self. With quasi-hyperbolic discounting, this disagreement is larger if β is small. For this case, note the similarity between s_3^* and s_2^*; we see exactly the same forces at work. For example, if technology k_3 is complementary to k_2, then k_3 requires a subsidy just as k_2 did when k_2 was complementary to k_1.

Interestingly, when we derive s_3^* for the case with quasi-hyperbolic discounting, it is only important whether k_2 increases or decreases in k_3. It is irrelevant whether the capital k_2 is itself green or brown (i.e., whether k_2 increases or decreases k_1). The explanation for the irrelevance of the capital type is the following. Although the current decision maker disagrees with her future self regarding the appropriate level of investments k_2, these two selves agree perfectly when trading off utilities between two later dates. With quasi-hyperbolic discounting, the discount factor of utility at time $t+1$ relative to time t is δ whenever $t > 1$. Thus, the decision maker choosing k_3 agrees with the decision maker choosing k_2 regarding the need to influence the decision maker selecting k_1.

(iii) When the discount factor δ_t is strictly increasing in t, however, the conclusions are quite different. Then, the decision maker investing in k_3 also seeks to raise k_1, and this can be done by strategically deciding on k_3 or s_3. In particular, for green technology, complementary to green capital, the decision maker invests strategically more in k_3 for two reasons, and the expression for s_3^* thus consists of two positive terms.

(iv) For brown technology, however, we know that k_1 decreases when the level of brown capital, k_2, increases. Therefore, the second term of s_3^* is negative, while the first term is positive. It is certainly possible that $s_3^* < 0$ if the second term dominates the first, positive term. This will be the case, for example, when the degree of substitutability between k_2 and k_1 is particularly large (thus, when dk_1/dk_2 is large). In this case, the motivation to subsidize investments in technology in order to motivate larger capital investments is outweighed by the fear that the capital stock will subsequently lead to more emissions.

Note that for both green and brown technology, the second term of s_3^* has the same sign as s_2^* if technology and capital are strategic complements (i.e., when $dk_2/dk_3 > 0$).[13]

4.2. The Supply Chain of Technologies

The analysis above suggests that for investment policies it is crucial to determine the technology's position in the production hierarchy. While the final investment stage before consumption did not need any regulation, investments in complementary green capital are subsidized. Furthermore, the investment in green technology will be subsidized at a rate which consists of two positive terms rather than just one, and its first term corresponds to the optimal subsidy on investments in capital. These comparisons suggest that the optimal subsidy for complementary investments further upstream may have a tendency to be larger and more complex.

To generalize the analysis in Section 4.1, assume now that there are N technology stages, indexed by $n \in \{1,..,N\}$. The investment cost for technology n is given by $c_n(k_n; k_{n+1})$, if we take k_{N+1} as exogenously given. To streamline notation, we may also take k_0 as given when defining $c_0(k_0; k_1) \equiv -b(k_1)$, so that the decision maker investing

[13]The result can easily be generalized to a setting in which the investment k_n takes time Δ_n to be developed, $n \in \{1,2,3\}$, and if k_2 influences the benefit of k_1, through $b(k_1, k_2)$. In this case, the optimal technology investment is given by:

$$
\begin{aligned}
s_3^* &= \underline{s}_3 + \overline{s}_3, \text{ where} \\
\underline{s}_3 &\equiv \frac{\partial c_2/\partial k_2}{-\partial c_2/\partial k_3}\left[\frac{D(\Delta_3+\Delta_2)}{D(\Delta_3)\,D(\Delta_2)} - 1\right]\frac{dk_2}{dk_3}, \text{ and} \\
\overline{s}_3 &\equiv \frac{\partial c_2/\partial k_2}{-\partial c_2/\partial k_3}\left[\frac{D(\Delta_3+\Delta_2+\Delta_1)\,D(\Delta_2) - D(\Delta_3+\Delta_2)\,D(\Delta_2+\Delta_1)}{D(\Delta_2)\,D(\Delta_3)\left[D(\Delta_2+\Delta_1) - D(\Delta_2)\,D(\Delta_1)\right]}\right]\left(\frac{\partial b}{\partial k_2} + \frac{\partial b}{\partial k_1}\frac{dk_1}{dk_2}\right)\frac{dk_2}{dk_3}.
\end{aligned}
$$

k_n solves the following problem:

$$\max_{k_n} \sum_{j=0}^{n} -D(n-j)\, c_j(k_j; k_{j+1}). \tag{4.3}$$

While I here will simplify and assume that the decision maker invests in only one k_n, $n \in \{1, ...N\}$, at each point in time, the analysis is unchanged if an entire vector $(k_{1,t}, ..., k_{N,t})$ is chosen at each time t (see Remark 3 at the end of this section).

To simplify notation, let p_n refer to the willingness to pay for k_n in the next period:

$$p_n \equiv -\frac{\partial c_{n-1}(k_{n-1}; k_n)}{\partial k_n}.$$

With this definition, we can write $c_n'(k_n; k_{n+1}) \equiv \partial c_n(k_n; k_{n+1}) / \partial k_n$ to simplify notation further. Inserted into the results derived already, we can write:

$$\begin{aligned} s_1^* &= 0, \\ s_2^* &= (\delta_2 - \delta_1) \frac{p_1}{p_2} \frac{dk_1}{dk_2}, \text{ and} \\ s_3^* &= (\delta_2 - \delta_1) \frac{p_2}{p_3} \frac{dk_2}{dk_3} + \left[(\delta_2 - \delta_1)^2 + \delta_2(\delta_3 - \delta_2)\right] \frac{p_1}{p_3} \frac{dk_1}{dk_3}. \end{aligned}$$

To solve the model with n stages, note that with a subsidy s_n, the market, or a naive agent, will invest according to:

$$\frac{c_n'(k_n; k_{n+1})}{1 + s_n} = \delta_1 p_n. \tag{4.4}$$

The decision maker, however, will take into account that the choice of k_n influences the next choice of k_{n-1}, and so on. In other words, the decision maker's preferred level of k_n may satisfy (4.4) only for some $s_n \neq 0$.

When the decision maker selects k_n or, equivalently, s_n, then she may anticipate that her later choices, such as the choice of k_{n-1} or s_{n-1}, will also be optimally selected at *that* stage. However, the following formula for the optimal s_n does *not* require s_j to be optimal for $j < n$, since the formula states the current decision maker's optimal choice of s_n quite generally, regardless of what the downstream investments or subsidies actually are:

$$s_n^* = \sum_{j=1}^{n-1} \left(\frac{\delta_{j+1}}{\delta_1} - 1 - s_{n-j} \right) D(j) \frac{p_{n-j}}{p_n} \frac{dk_{n-j}}{dk_n}. \tag{4.5}$$

PROPOSITION 3: *For any* $n \in \{1, ..., N\}$, *(4.5) defines:*

(i) *the optimal* s_n *for arbitrary* s_j, $j \in \{1, .., n-1\}$, *and*

(ii) *the optimal* s_n *recursively, if* s_j, $j \in \{1, ..., n-1\}$, *are also optimally chosen.*

PROOF: Maximizing (4.3) with respect to k_n is directly giving the first-order condition:

$$
\begin{aligned}
v_n'(k_n; k_{n+1}) &\equiv \frac{d}{dk_n} \sum_{j=0}^{n} -D(n-j)\, c_j(k_j; k_{j+1}) = 0 \Leftrightarrow \\
c_n'(k_n; k_{n+1}) &= -\sum_{j=0}^{n-1} D(n-j) \frac{d}{dk_n} c_j(k_j; k_{j+1}) \\
= \quad & -\delta_1 p_n - \sum_{j=1}^{n-1} D(n-j) \frac{\partial}{\partial k_j} \left[\delta_{n-j+1} c_{j-1}(k_{j-1}; k_j) + c_j(k_j; k_{j+1})\right] \frac{dk_j}{dk_n}. \qquad (4.6)
\end{aligned}
$$

If we replace n with j in (4.4) and substitute into the above equation, we get:

$$
\begin{aligned}
c_n'(k_n; k_{n+1}) &= -\delta_1 \frac{\partial c_{n-1}(k_{n-1}; k_n)}{\partial k_n} \\
&- \sum_{j=1}^{n-1} D(n-j) \left[\delta_{n-j+1} \frac{\partial}{\partial k_j} c_{j-1}(k_{j-1}; k_j) + \left(-\delta_1 (1+s_j) \frac{\partial}{\partial k_j} c_{j-1}(k_{j-1}; k_j)\right)\right] \frac{dk_j}{dk_n} \\
&= \delta_1 p_n + \sum_{j=1}^{n-1} D(n-j) \left[\delta_{n-j+1} - \delta_1 (1+s_j)\right] p_j \frac{dk_j}{dk_n}.
\end{aligned}
$$

With this, s_n^* can be derived by a comparison to (4.4), or directly from (3.3):

$$
\begin{aligned}
s_n^* &= \frac{\delta_1 p_n + \sum_{j=1}^{n-1} D(n-j) \left[\delta_{n-j+1} - \delta_1 (1+s_j)\right] p_j dk_j / dk_n}{\delta_1 p_n} - 1 \\
&= \sum_{j=1}^{n-1} D(n-j) \left[\frac{\delta_{n-j+1}}{\delta_1} - (1+s_j)\right] \frac{p_j}{p_n} \frac{dk_j}{dk_n} \\
&= \sum_{i=1}^{n-1} D(i) \left[\frac{\delta_{i+1}}{\delta_1} - (1+s_{n-i})\right] \frac{p_{n-i}}{p_n} \frac{dk_{n-i}}{dk_n}.
\end{aligned}
$$

The second-order condition is $v_n'' \equiv \partial v_n'(k_n; k_{n+1}) / \partial k_n < 0$, which must hold, and it does hold if c_n is sufficiently convex in k_n. By differentiating the first-order condition $v_n'(k_n; k_{n+1}) = 0$ w.r.t. k_{n+1}, we get:

$$
\frac{dk_n}{dk_{n+1}} = -\frac{\partial^2 c_n(k_n; k_{n+1})}{\partial k_n \partial k_{n+1}} \left(\frac{1}{-v_n''}\right), \qquad (4.7)
$$

and $dk_{n-i}/dk_n = \prod_{j=n-i+1}^{n} (dk_{j-1}/dk_j)$. *Q.E.D.*

Thus, the expression for s_n^* is the sum of $n-1$ terms. The terms inside the parentheses are zero if discounting is exponential and if $s_j = 0$ for every $j < n$; so, in this case, $s_n^* = 0$, as well. If we had $s_j = 0$ for $j < n$ and discount factors increased in t, then every parenthesis would be strictly positive. Each parenthesis is multiplied with the positive discount factor $D(j)$, and the positive price ratio p_{n-j}/p_{n-1}, so the sign of each term depends on the sign of dk_{n-j}/dk_n. If all technologies are strategic complements (in that a larger k_n reduces the cost of k_{n-1}) then $dk_{n-j}/dk_n > 0$. In this case, s_n^* would be the sum of $n-1$ positive terms, suggesting that s_n^* may have a tendency to increase in n. This conjecture will be further explored in the rest of this section.

In equilibrium, s_j for $j < n$ will be given by the same formula, (4.5). In this case, (4.5) is a recursive formula that pins down every s_n and thus every investment level.[14]

An interesting corollary can be derived by assuming that at least the next s_{n-1} is set according to Proposition 3. This s_{n-1} can then be substituted into (4.5). For this corollary, we do not need to assume that the s_j's further downstream ($j < n-1$) are also optimally chosen (although, of course, they may be, and, in equilibrium, they will be).

COROLLARY 3: *Suppose s_{n-1} is given by (4.5). Regardless of whether s_j, $j \in \{1, ..., n-2\}$, is optimally or arbitrarily chosen, the following obtain:*

(i) *With exponential discounting, $s_n^* = 0$.*

(ii) *With quasi-hyperbolic discounting, s_n^* consists of the single term accounting for the effect on k_{n-1}:*

$$s_n^* = \left(\frac{\delta_2}{\delta_1} - 1\right)\left[-\frac{\partial c_{n-1}/\partial k_{n-1}}{\partial c_{n-1}/\partial k_n}\right]\frac{dk_{n-1}}{dk_n}.$$

(iii) *With strictly increasing discount factors, s_n^* is the sum of $n-1$ terms:*

$$\begin{aligned} s_n^* &= \left(\frac{\delta_2}{\delta_1} - 1\right)\left[-\frac{\partial c_{n-1}/\partial k_{n-1}}{\partial c_{n-1}/\partial k_n}\right]\frac{dk_{n-1}}{dk_n} \qquad (4.8) \\ &+ \sum_{i=2}^{n-1}\left[\frac{\delta_i}{\delta_1}(\delta_{i+1} - \delta_2) - (\delta_i - \delta_2)(1 + s_{n-i})\right] D(i-1)\frac{p_{n-i}}{p_n}\frac{dk_{n-i}}{dk_n}. \end{aligned}$$

[14]The equilibrium levels can be found by first applying the formula to k_1, or s_1, which will give us $s_1^* = 0$. By substituting in for this value of s_1, we can derive s_2^*, and so on.

PROOF: If we replace n with $n-1$ in (4.6) and rewrite, the f.o.c. for k_{n-1} becomes:

$$\begin{aligned} -\frac{\partial c_{n-2}(k_{n-2};k_{n-1})}{\partial k_{n-1}} &= \frac{c'_{n-1}(k_{n-1};k_n)}{\delta_1} \\ &+\sum_{j=1}^{n-2}\frac{D(n-j-1)}{\delta_1}\left[\delta_{n-j}\frac{\partial c_{j-1}(k_{j-1};k_j)}{\partial k_j}+\frac{\partial c_j(k_j;k_{j+1})}{\partial k_j}\right]\frac{dk_j}{dk_{n-1}}. \end{aligned} \tag{4.9}$$

Also, note that we can rewrite (4.6) to:

$$\begin{aligned} c'_n(k_n;k_{n+1}) &= -\delta_1\frac{\partial c_{n-1}(k_{n-1};k_n)}{\partial k_n}-\left[D(2)\frac{\partial c_{n-2}(k_{n-2};k_{n-1})}{\partial k_{n-1}}+\delta_1\frac{\partial c_{n-1}(k_{n-1};k_n)}{\partial k_{n-1}}\right]\frac{dk_{n-1}}{dk_n} \\ &-\sum_{j=1}^{n-2}\left[D(n-j+1)\frac{\partial c_{j-1}(k_{j-1};k_j)}{\partial k_j}+D(n-j)\frac{\partial c_j(k_j;k_{j+1})}{\partial k_j}\right]\frac{dk_j}{dk_n}. \end{aligned}$$

This equation becomes, after substituting in with (4.9):

$$\begin{aligned} c'_n(k_n;k_{n+1}) &= -\delta_1\frac{\partial c_{n-1}(k_{n-1};k_n)}{\partial k_n}-\left[-D(2)\frac{c'_{n-1}(k_{n-1};k_n)}{\delta_1}+\delta_1\frac{\partial c_{n-1}(k_{n-1};k_n)}{\partial k_{n-1}}\right]\frac{dk_{n-1}}{dk_n} \\ &-\sum_{j=1}^{n-2}\left[\begin{array}{c}\left(D(n-j+1)-\frac{D(2)D(n-j)}{D(1)}\right)\frac{\partial c_{j-1}(k_{j-1};k_j)}{\partial k_j} \\ +\left(D(n-j)-\frac{D(2)D(n-j-1)}{D(1)}\right)\frac{\partial c_j(k_j;k_{j+1})}{\partial k_j}\end{array}\right]\frac{dk_j}{dk_n} \\ &= \delta_1 p_n+(\delta_2-\delta_1)c'_{n-1}(k_{n-1};k_n)\frac{dk_{n-1}}{dk_n} \\ &+\sum_{j=1}^{n-2}D(n-j-1)\left[\delta_{n-j}(\delta_{n-j+1}-\delta_2)p_j-(\delta_{n-j}-\delta_2)\frac{\partial c_j(k_j;k_{j+1})}{\partial k_j}\right]\frac{dk_j}{dk_n}. \end{aligned}$$

To ensure that also (4.4) holds, s_n must be optimal and given by:

$$\begin{aligned} s_n^* &= \left(\frac{\delta_2}{\delta_1}-1\right)\frac{c'_{n-1}(k_{n-1};k_n)}{p_n}\frac{dk_{n-1}}{dk_n} \\ &+\sum_{j=1}^{n-2}\left[\frac{\delta_{n-j}}{\delta_1}(\delta_{n-j+1}-\delta_2)-(\delta_{n-j}-\delta_2)(1+s_j)\right]D(n-j-1)\frac{p_j}{p_n}\frac{dk_j}{dk_n}, \end{aligned}$$

which we can rewrite to (4.8). *Q.E.D.*

Parts (ii) and (iii) reveal that there is a dramatic difference between quasi-hyperbolic discounting and strictly increasing discount factors. With quasi-hyperbolic discounting, the expression for s_n^* consists of only one single term, and that term is written equivalently for every $n > 1$. The explanation is the following: On the one hand, the decision maker is time-inconsistent and she prefers to subsidize investments that are complementary to the choice of k_{n-1}, in order to influence the decision maker at that next stage to invest more. That decision maker, in turn, disagrees with the decision maker who decides on k_{n-2}. The decision maker deciding on k_n and the decision maker deciding on k_{n-1} both agree on how much more the decision maker deciding on k_{n-2} ought to invest. The disagreement between the first two decision makers is limited to the choice of k_{n-1}, thanks to discount factors that are constant after the first increase from $\beta\delta$ to δ, since $\delta_t = \delta \forall t > 1$. In contrast, when discount factors are strictly increasing in t, then s_n^* consists of $n-1$ terms. This comparison reveals that quasi-hyperbolic discounting is not a good approximation for hyperbolic discounting when studying production-chain investments.

4.3. Stepping Stone Technologies

As illustrated by Proposition 3, the optimal subsidy consists of a number of terms that equal the technology's rank in the production chain. This does not prove, of course, that the subsidy is larger for more fundamental (or more "upstream") technologies, but there might be such a tendency for complementary technologies.

To investigate this claim, consider now what I will refer to as "stepping stone technologies." For such technologies, each stage is the stepping stone for the next. The larger is one stepping stone, k_{n+1}, the larger is also k_n, for any given investment cost at stage n. Thus, the cost of investing in k_n can be written as $c_n\left(k_n - \phi_{n+1}k_{n+1}\right)$. Without loss of generality, we can let $\phi_j = 1$ for any $j \in \{1, ..., N\}$.[15] With this, technology k_{n+1} becomes a perfect complement to k_n: one more unit of k_{n+1} makes it possible to also raise k_n by one unit, changing neither the cost nor the marginal cost of investing in k_n.

The study of stepping stone technologies can be motivated in several ways. One motivation is that these technologies capture quite well the way in which environmentally friendly technologies enter the production chain. The amount of energy that can be generated by renewable energy sources reduces, one by one, the amount of greenhouse gas

[15] If the true investment costs were $c_n\left(k_n - \phi_{n+1}\widetilde{k}_{n+1}\right)$, and the technology level $\widetilde{k}_{n+1}$ could be invested in at cost $\widetilde{c}_{n+1}\left(\widetilde{k}_{n+1} - \phi_{n+2}\widetilde{k}_{n+2}\right)$, then we could simply define $k_{n+1} \equiv \phi_{n+1}\widetilde{k}_{n+1}$ and let the investment cost for k_{n+1} be $c_{n+1}\left(k_{n+1} - \phi_{n+2}\phi_{n+1}\widetilde{k}_{n+2}\right) \equiv \widetilde{c}_{n+1}\left(k_{n+1}/\phi_{n+1} - \phi_{n+2}\widetilde{k}_{n+2}\right)$. In an analogous way we can eliminate $\phi_{n+2}\phi_{n+1}$ and write $c_{n+1}\left(k_{n+1} - k_{n+2}\right)$ by defining $k_{n+2} \equiv \phi_{n+2}\phi_{n+1}\widetilde{k}_{n+2}$ and redefining $c_{n+2}\left(\cdot\right)$, and so on.

that enters the atmosphere, for any given level of energy consumption. For this reason, stepping stone technologies have already been used in other studies of climate change.[16]

Proposition 4: *For stepping stone technologies, where* $c_n(k_n; k_{n+1}) = c_n(k_n - k_{n+1})$, *the equilibrium* k_n *satisfies (4.4) with the following* $s_n \geq 0$*, increasing in* n*:*

$$s_n^* = \frac{\delta_n}{\delta_1} - 1.$$

Proof: From (4.7) we have $dk_n/dk_{n+1} = c_n''/c_n'' = 1$. Thus, the decision maker's first-order condition simplifies to $c_n' = D(n)$. Combined with (4.4), we get $s_n^* = D(n)/\delta_1 p_n - 1$. But $p_n = -\partial c_{n-1}(k_{n-1} - k_n)/\partial k_n = \partial c_{n-1}(k_{n-1} - k_n)/\partial k_{n-1}$, which equals $D(n-1)$. Thus, $s_n^* = D(n)/\delta_1 D(n-1) - 1 = \delta_n/\delta_1 - 1$. *Q.E.D.*

Just as before, the subsidy is zero at the last stage. If discounting is exponential, the subsidy is zero at every stage. And, as a confirmation of Corollary 3, the subsidy is indeed constant in n under quasi-hyperbolic discounting, but increasing in n if discount factors are strictly increasing in relative time.

Corollary 4:

(i) *With exponential discounting, or if* $n = 1$*, then* $s_n^* = 0$*.*

(ii) *With quasi-hyperbolic discounting,* $s_n^* = 1/\beta - 1 > 0$ *is constant for all* $n > 1$*.*

(v) *With strictly decreasing discount rates,* s_n^* *is strictly increasing in* n*.*

(iv) *With hyperbolic discounting,*

$$s_n^* = \alpha\left(1 - \frac{1+\alpha}{1+\alpha n}\right).$$

Figure 4.1 illustrates Corollary 4: The production stage is measured at the horizontal axis. The solid line measures equilibrium investments, or, in fact, the marginal investment cost, $c_n'(\cdot) = D(n) = \prod_{i=1}^{n} \delta_i$, at each stage in the production chain. The lower dashed line similarly measures investments under laissez faire, or, alternatively, if a naive agent were to invest at all stages: then, $c_n'(\cdot) = \delta_1^n$. The upper dashed line is in the same way

[16] See, for example, Harstad (2012) or Battaglini and Harstad (2016). The term "stepping stone technology" is not used in those papers, even though the technology is a perfect substitute for reducing consumption, as assumed here.

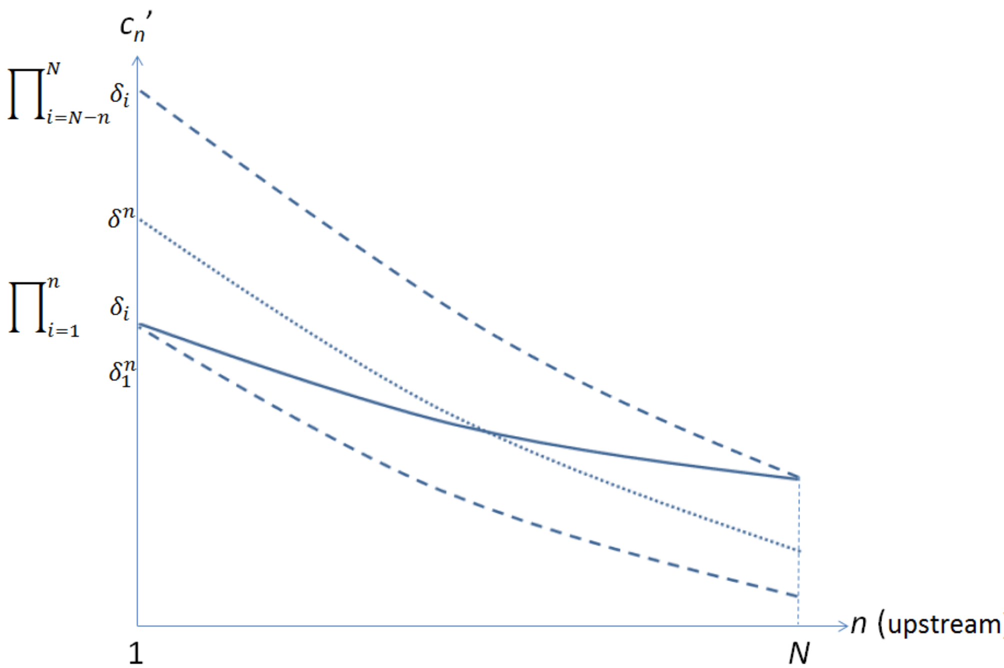

Figure 4.1: *In equilibrium (solid line), more is spent on upstream investments relative to downstream investments, regardless of whether we compare to laissez faire, naive investments, exponential discounting, or investments under commitment.*

measuring investment expenditures at each stage under commitment, assuming that the decision maker deciding on k_N could commit to how much to invest in all future stages: In this case, investments would be larger and given by $c_n'(\cdot) = D(N)/D(N-n) = \prod_{i=N-n}^{N} \delta_i$. Finally, the dotted line measures the investment expenditures under exponential discounting, for some fixed discount factor $\delta \in \left(\delta_1, \sqrt[N]{\delta_1 \cdot \delta_2 \cdot ... \delta_N}\right)$. Relative to all these three benchmarks, the equilibrium investment expenditures are biased toward the investments that are further upstream, and away from the downstream investments. In other words, with time-inconsistent preferences, more of the budget is spent on basic research and the development of fundamental technology, whether the comparison is to a setting with time consistency, commitment, the investments of a naive agent, or the investments in a perfect market under laissez faire.

Remark 3: *Long-lasting stocks and investments in* $(k_{1,t}, k_{2,t}, ..., k_{N,t})$ *in every period* t. In the analysis above, (a) the decision on k_n was, for simplicity, taken before the decision on k_{n-1}, and (b) the stock k_n played no role thereafter (it depreciated completely). The results do not hinge on these assumptions, however, and both of them can be relaxed.

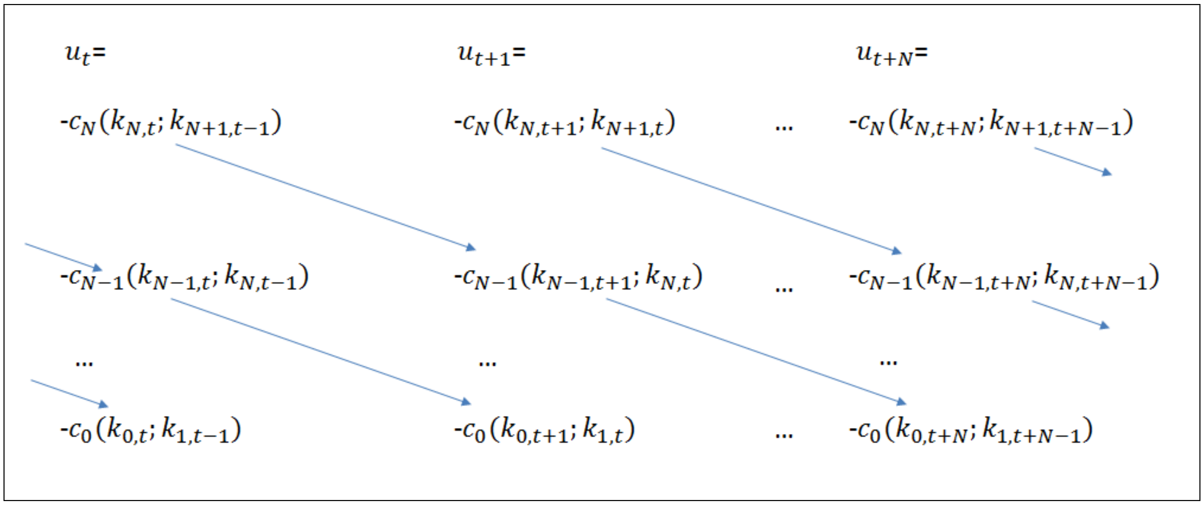

Figure 4.2: *Maximizing the vector* $(k_{1,t}, ..., k_{N,t})$ *can be separated into* N *independent maximization problems.*

First, to relax (a), suppose that at every time t the decision maker decides on the vector $\mathbf{k}_t = (k_{1,t}, k_{2,t}, ..., k_{N,t})$, receives the momentary utility $u_t = -\sum_{n=0}^{N} c_n (k_{n,t}; k_{n+1,t-1})$, and seeks to maximize $v_t = \sum_{t=t}^{\infty} D(\tau - t) u_\tau$. By inserting the expression for u_τ and rearranging, we can write the objective as:

$$\begin{aligned} v_t &= -\sum_{n=0}^{N} \left(\sum_{j=0}^{n} D(n-j) c_j (k_{j,t+n-j}; k_{j+1,t+n-j-1}) \right) \\ &\quad -\sum_{\tau=1}^{\infty} \sum_{j=0}^{N} D(N-j+\tau) c_j (k_{j,t+N-j+\tau}; k_{j+1,t+N-j+\tau-1}). \end{aligned}$$

The final term captures future payoffs that are independent of $\mathbf{k}_t$. In the first term, the effect of $k_{n,t}$ is contained in each parenthesis, and the term in each parenthesis is identical to (4.3), except that time subscripts are added. Thus, the problem of maximizing v_t with respect to $\mathbf{k}_t$ consists of N maximization problems, each identical to one studied above. The intuition for this separation is illustrated in Figure 4.2.

We can also relax (b) and allow k_n to depreciate at rate $1 - q_n \in [0, 1]$. The separation above will fail if the choice of $k_{n,t}$ will influence not only $k_{n-1,t+1}$, but also $k_{n,t+1}$, and therefore $k_{n-1,t+2}$, and so on. These multiple links would vastly complicate the analysis. However, the sign of $dk_{n,t+1}/dk_{n,t}$ is not necessarily positive and it is zero if, as in Remark 2, the cost of upgrading to $k_{n,t}$ is assumed to be additively separable and given by $\widetilde{c}_n (k_{n,t}; k_{n+1,t-1}) - h_n (q_n k_{n,t-1})$, for some function h_n. If we account for the cost-saving

$h_n(q_n k_{n,t-1})$ when defining the cost of investing in $k_{n-1,t}$, then we can leave the analysis above unchanged by using this definition:

$$c_n(k_{n,t}; k_{n+1,t-1}) \equiv \widetilde{c}_n(k_{n,t}; k_{n+1,t-1}) - h_{n+1}(q_{n+1} k_{n+1,t-1}).$$

5. Conclusions

There is a large amount of evidence indicating that individuals have time-inconsistent preferences and behave as if they are more patient regarding long-term decisions than for short-term decisions. Governments and policy makers will also have these preferences, both because they are citizens themselves, and because they must be accountable to voters with time-inconsistent preferences.

To study the public policy consequences of time-inconsistent preferences, this paper analyzes models of investments in capital or technology and the associated investment policies. The current decision maker can influence future investment choices by strategically choosing investments today. A measure of the strategic concern is the subsidy the decision maker would impose on today's investment if the actual decision were (perhaps hypothetically) made in a perfect market or by a naive agent. A time-consistent policy maker would see no need to influence future decisions, and the optimal subsidy would then be zero. With more realistic discount factors that increase in relative time, however, I derive two important results.

First, the subsidy will depend on the *type* of capital or technology to be invested in. In particular, the current decision maker has an incentive to subsidize or invest more in capital or technologies that are complementary to future investments, but to tax or invest less in technologies that are strategic substitutes for future investments. This result has important policy implications for environmental policy, for example. Even when one abstracts from pollution externalities and technological spillovers, it is optimal to subsidize investments in "green" capital or technology but it is optimal to tax investments in "brown" capital or technology. Investments in adaptations to climate change are a strategic substitute to pollution abatement and, therefore, the current decision makers benefit from taxing such investments.

Second, the optimal policy will depend on the position of the technology in the production chain. When upstream technology is a strategic complement to the development of downstream technology or capital, then it is often optimal to provide a larger subsidy to more upstream technologies, i.e., to more basic research. The reason is that upstream technologies have a multiplicative effect on the sequence of future investment decisions which the current policy maker would like to influence.

This paper takes only a few small steps toward an understanding of how governments may want to influence or regulate strategic investments in the presence of time inconsistency. Although the literature on this topic is still limited, I believe it will and should be vastly expanded in the coming years for two reasons. First, the fields of political economics and behavioral economics are rapidly growing; Second, public decisions about long-term problems—such as climate change—are receiving increasing attention, and it is for such long-term decisions that time inconsistency and declining discount rates will have the most dramatic policy consequences.

References

Ainslie, George (1992): *Picoeconomics: The Strategic Interaction of Successive Motivational States*, Cambridge U. Press.

Angeletos, George-Marios; Laibson, David; Repetto, Andrea; Tobacman, Jeremy and Weinberg, Stephen (2001): "The Hyperbolic Consumption Model: Calibration, Simulation and Empirical Evaluation," *Journal of Economic Perspectives* 15: 47-68.

Barro, Robert (1974): "Are Government Bonds Net Wealth?," *Journal of Political Economy* 82: 1095-1117.

Battaglini, Marco and Harstad, Bård (2016): "Participation and Duration of Environmental Agreements," forthcoming, *Journal of Political Economy.*

Benhabib, Jess; Bisin, Alberto and Schotter, Andrew (2010): "Present-bias, quasi-hyperbolic discounting, and fixed costs," *Games and Economic Behavior* (2): 205-23.

Bisin, Alberto; Lizzeri, Alessandro and Yariv, Leeat (2015): "Government Policy with Time Inconsistent Voters," *American Economic Review* 105(6): 1711-37.

Böhm-Bawerk, Eugene v. (1889): *Capital and Interest*, Libertarian Press.

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

Dasgupta, Partha and Maskin, Eric (2005): "Uncertainty and Hyperbolic Discounting." *American Economic Review* 95(4): 1290-99.

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

Diamond, Peter and Kőszegi, Botond (2003): "Quasi-Hyperbolic Discounting and Retirement," *Journal of Public Economics* 87(9-10): 1839-72.

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

Galperti, Simone and Strulovici, Bruno (2015): "Pure Altruism and Time Inconsistency: An Axiomatic Foundation," mimeo, Northwestern University.

Gerlagh, Reyer and Liski, Matti (2013): "Carbon Prices for the Next Thousand Years."

Gollier, Christian and Weitzman, Martin L. (2010): "How Should the Distant Future be Discounted When Discount Rates are Uncertain?," *Economics Letters* 107(3): 350-53.

Gollier, Christian and Zeckhauser, Richard (2005): "Aggregation of Heterogeneous Time Preferences," *Journal of Political Economy* 113(4): 878-96.

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

Gul, Faruk and Pesendorfer, Wolfgang (2001): "Temptation and self-control," *Econometrica* 69(6): 1403-35.

Halevy, Yoram (2015): "Time Consistency: Stationarity and Time Invariance," *Econometrica* 83(1): 335-52.

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

Harstad, Bård (2012): "Climate Contracts: A Game of Emissions, Investments, Negotiations, and Renegotiations," *Review of Economic Studies* 79(4),1527-57.

Heidhues, Paul and Kőszegi, Botond (2010): "Exploiting Naïvete about Self-Control in the Credit Market," *American Economic Review* 100(5): 2279–2303

Jackson, Matthew O. and Yariv, Leeat (2015): “Collective Dynamic Choice: The Necessity of Time Inconsistency,” forthcoming, *AEJ: Microeconomics*.

Jackson, Matthew O. and Yariv, Leeat (2014): “Present Bias and Collective Dynamic Choice in the Lab,” *American Economic Review* 104(12): 4184-4204.

Jevons, William S. (1871): *Theory of Political Economy*, London: Macmillan

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

Koopman, Tjalling (1960): "Stationary Ordinal Utility and Impatience," *Econometrica* 28(2): 287-309.

Kőszegi, Botond (2014): "Behavioral Contract Theory," *Journal of Economic Literature* 52(4): 1075–1118

Krusell, Per; Kurusçu, Burhanettin and Smith, Anthony A. Jr. (2009): "How Much Can Taxation Alleviate Temptation and Self-Control Problems?," Manuscript.

Krusell, Per; Kurusçu, Burhanettin and Smith, Anthony A. (2010): "Temptation and Taxation," *Econometrica* 78(6): 2063-84.

Laibson, David (1997): "Golden eggs and hyperbolic discounting," *Quarterly Journal of Economics* 112(2): 443-78.

Laibson, David and Christopher Harris (2001): "Dynamic Choices of Hyperbolic Consumers," *Econometrica* 69: 935-57.

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

O’Donoghue, Ted and Rabin, Matthew (1999): "Doing It Now or Later," *American Economic Review* 89(1): 103-24.

Paserman, Daniele (2008): "Job Search and Hyperbolic Discounting: Structural Estimation and Policy Evaluation," *Economic Journal* 118(531): 1418-52.

Phelps, Edmund S. and Pollak, Robert A. (1968): "On Second-Best National Saving and Game-Equilibrium Growth," *Review of Economic Studies* 35: 165-199.

Rae, John (1834): “The Sociological Theory of Capital,” London: Macmillan.

Ramsey, Frank P. (1928): "A Mathematical Theory of Saving," *Economic Journal* 38(152): 543-59.

Saez-Marti, Maria and Weibull, Jörgen W. (2005): "Discounting and altruism to future decision-makers,” *Journal of Economic Theory* 122: 254-66.

Samuelson, Paul (1937): "A Note on Measurement of Utility," *Review of Economic Studies* 4: 155-161.

Senior, Nassau W. (1836): *An Outline of the Science of Political Economy*, London: Clowes and Sons.

Strotz, Robert H. (1956): "Myopia and Inconsistency in Dynamic Utility Maximization," *Review of Econonomic Studies* 23: 166-180.

Thaler, Richard H. (1981): "Some Empirical Evidence on Dynamic Inconsistency," *Economic Letters* 8(3): 201-207.

Thaler, Richard H. and Benartzi, Shlomo (2004): "Save More Tomorrow," *Journal of Political Economy* 112(1): 164-87.

Weitzman, Martin L. (2001): "Gamma Discounting," *American Economic Review* 91(1): 260-71.