Les Cahiers du GERAD ISSN: 0711–2440

Non-constant discounting, social welfare and endogenous growth with pollution externalities

F. Cabo, G. Martín-Herrán,
M. P. Martínez-García

G–2017–85

November 2017

La publication de ces rapports de recherche est rendue possible grâce au soutien de HEC Montréal, Polytechnique Montréal, Université McGill, Université du Québec à Montréal, ainsi que du Fonds de recherche du Québec – Nature et technologies.

Dépôt légal – Bibliothèque et Archives nationales du Québec, 2017
– Bibliothèque et Archives Canada, 2017

The publication of these research reports is made possible thanks to the support of HEC Montréal, Polytechnique Montréal, McGill University, Université du Québec à Montréal, as well as the Fonds de recherche du Québec – Nature et technologies.

Legal deposit – Bibliothèque et Archives nationales du Québec, 2017
– Library and Archives Canada, 2017

GERAD HEC Montréal
3000, chemin de la Côte-Sainte-Catherine
Montréal (Québec) Canada H3T 2A7

Tél. : 514 340-6053
Téléc. : 514 340-5665
info@gerad.ca
www.gerad.ca

Non-constant discounting, social welfare and endogenous growth with pollution externalities

Francisco Cabo [a]
Guiomar Martín-Herrán [a,b]
María Pilar Martínez-García [c]

[a] *IMUVa, Universidad de Valladolid, Spain*

[b] *GERAD, HEC Montréal, Montréal (Québec), Canada*

[c] *Universidad de Murcia, Spain*

pcabo@eco.uva.es
guiomar@eco.uva.es
pilarmg@um.es

November 2017
Les Cahiers du GERAD
G–2017–85

Abstract: We analyze the effect of non-constant discounting on economic growth and social welfare in an endogenous growth model with pollution externalities. For time-consistent agents, the existence of a balanced growth equilibrium is characterized and compared with the equilibrium under standard exponential discounting. A decaying instantaneous discount rate leads to slower growth in a centralized economy, while its effect for a competitive economy is ambiguous. Interestingly, when comparing the planned and the competitive equilibria, the assumption of non-constant discounting may imply greater social welfare under the market equilibrium. This counterintuitive result requires two conditions. First, pollution externalities should lead the central planner to slow down growth to below the growth rate in the market economy. Secondly, individuals' degree of impatience should decrease sharply with the time distance from the present. Contrarily, when the centralized economy welfare dominates the market economy, introducing policy instruments is less effective than under constant discounting.

Keywords: Non-constant discounting, endogenous growth, Social welfare, sustainability, environmental policies, time-consistent solutions

Acknowledgments: The authors have been partially supported by MEC under project ECO2014-52343-P, co-financed by FEDER funds.

1 Introduction

The optimal decisions of economic agents when confronted with intertemporal problems depend fundamentally on the way these agents discount the future. Ever since Samuelson (1937), the standard assumption in intertemporal decision making in economics is to discount at a constant rate, the so-called exponential discounting.[1] However, a recent branch of the literature moves away from this assumption, considering individuals with a decreasing impatience more appropriate. This idea is highlighted in Laibson (1997) who pointed out that individuals are highly impatient about consumption in the near future but much more patient when confronted with decisions in the distant future.

An important strand of the literature supports the appropriateness of time-varying discount rates based on theoretical and empirical grounds (see, Frederick et al. (2002) and DellaVigna (2009) for a survey). However, this opinion is not unanimous, with some criticisms mostly supported by empirical evidence (see, for example, Andersen et al. (2014)). Besides this debate about the suitability of this hypothesis to represent reality, there exists an additional difficulty linked to the use of non-constant discounting. As already pointed out by Ramsey (1928) and later by Strotz (1956), Pollak (1968) and Goldman (1980), with non-constant discounting, preferences change with time and therefore decisions taken at the present time will not necessarily be optimal later on. This time-inconsistency problem has been an important drawback, impeding the generalization of the use of non-constant discounting. The literature on non-constant discounting distinguishes two types of agents: Naïve individuals, who continuously recompute their dynamic optimization problem, and therefore their optimal decisions are time inconsistent, and, on the other hand, sophisticated agents, who are aware that their preferences will vary as time goes by. In consequence, they play a game against their future selves, and their optimal decisions become time consistent. There is no clear consensus in the literature on which type of agent more truthfully resembles real individuals. Some authors (like Caliendo and Aadland (2007), Findley and Caliendo (2014) or Farzin and Wedner (2014)) support the hypothesis that individuals are indeed unaware of their future impatience. In contrast, we align ourselves with others (like Barro (1999), Karp (2007), or Karp and Tsur (2011)) assuming sophisticated consumers, whose optimal consumption paths are time consistent.[2]

With non-constant time preferences, when commitment is not feasible, the solution to the game played by a sophisticated central planner against his future selves need not be Pareto-improving with respect to a decentralized or market solution. This result has been highlighted in the literature of Neoclassical growth by Krusell et al. (2002), based on the idea that the planner acknowledges his effect on the returns to savings, while these returns are considered constant by price-taker consumers. Thus, the rate of return is decreasing for the planner and constant for competitive agents. In consequence, the latter save more and the competitive economy grows faster (towards the steady-state). Therefore the decentralized economy ends up providing greater social welfare.[3] While this result is obtained considering quasi-geometric discounting, Hiraguchi (2014) proves the robustness of this result, which remains valid for a general non-constant discounting function.[4] The main contribution of our paper is to show that the inefficiency of the central planner's solution with respect to the market economy under non-constant discounting also emerges in an endogenous growth model, based on different grounds: the existence of environmental externalities which lead the central planner to slow down growth with respect to the market economy.

In this paper the externalities arise from considering that agents are consumption-driven individuals and also give entrance to the amenities stemming from the environmental quality. Thus, the analysis revisits the seminal paper by Smulders and Gradus (1996), which studies the conditions for sustained economic growth preserving the environmental quality. Pollution is a by-product of production which can be reduced by abatement activities. The environmental quality is associated with two externalities: it increases utility as well as the productivity of the direct factors of production. When considering individuals with a decreasing

[1] Later, Strotz (1956) proved that exponential discounting is the only discount function which guarantees time consistency.

[2] A method to find a time-consistent solution with non-constant discounting is proposed in Karp (2007) for an infinite time horizon and in Marín-Solano and Navas (2009) for a finite time problem.

[3] Along the same line, this result of group inefficiency can be also obtained in the case of asymmetric players (with different discount rates), as shown in Marín-Solano (2015) and Castañer et al. (2016).

[4] However, more recently, Hiraguchi (2016) proves that the competitive economy never outperforms the centralized economy in a two-sector endogenous growth.

rate of impatience, we distinguish the decentralized market economy from the equilibrium associated with a central planner who, acting as a representative agent, internalizes pollution externalities. The conditions for the existence of a balanced growth path with decreasing pollution, the actual growth rate of the economy and social welfare are computed and compared for the two different scenarios.

The literature on economic growth and non-constant discounting was initiated by Barro (1999). For a Neoclassical growth model he shows that non-constant (quasi-hyperbolic) discounting is observationally equivalent to exponential discounting when a log-utility function and sophisticated agents are considered. This equivalence is not necessarily true, as proved first in Farzin and Wedner (2014) when the intertemporal elasticity of substitution differs from one (assuming naïve consumers), and secondly, in Cabo et al. (2016), when a non-constant discount function other than the quasi-hyperbolic is considered (assuming sophisticated consumers). Similarly, for Ak-type endogenous growth models, Strulik (2015) proves that hyperbolic and exponential discounting are observationally equivalent for log-utility and naïve consumers. Moreover he proves strong equivalence, which was defined as having the same growth rate under the assumption of an identical overall impatience. However, as proved in Cabo et al. (2015), isoelastic, although not logarithmic, utility preserves observational equivalence but not strong equivalence, regardless of whether individuals are naive or sophisticated. This literature, which highlights the importance of the intertemporal elasticity of substitution, does not focus on welfare comparisons. Indeed, observational equivalence does not preclude differences in welfare, as already commented by Krusell *et al.* (2002) or Hiraguchi (2014) for a Neoclassical growth model.

The literature on environmental economics and natural resource management has already shown an interest in the rationale for the use of non-constant discounting. Groom et al. (2005) revise the different reasons that justify the use of declining discount rates, and the main criticisms of this assumption for environmental problems, in particular for global warming. The pros and cons of hyperbolic discounting are also reviewed in Hepburn *et al.* (2010), which highlights the time inconsistency problem for a model of natural resource management. A seminal work studying non-constant discounting in an endogenous growth model with a renewable natural resource is presented in Li and Lofgren (2000), which analyzes the sustainability of economic growth based on the Chichilnisky criterion (Chichilnisky 1996). For the particular case of global warming, the effect of hyperbolic discounting on the optimal policies has been analyzed from various perspectives. The time consistent equilibria are characterized for a discrete time formulation in Karp (2005) and Fuji and Karp (2008), and for continuous time in Karp and Tsur (2011). More recently, Gerlagh and Liski (2017) revisit the implications of non-constant discounting for climate policies. Our analysis is a contribution to the literature on environmental economics and non-constant discounting by considering an endogenous growth model in which the quality of the environment affects productivity and consumers' utility.

The analysis of the endogenous growth model with environmental quality by Smulders and Gradus (1996), under the assumption of non-constant discounting, leads us to study three main research questions. First, when a balanced growth path which preserves the environmental quality exists, how is it affected by the assumption of non-constant discounting, both in the central planner and the market economy? Second, is it possible to find conditions under which the market economy welfare dominates the central planner's economy? And lastly, in the opposite case, when policies to approach the market economy to the central planner's equilibrium are appropriate, how efficient are these policies compared to the case of constant discounting?

We characterize conditions for the existence of a balanced growth path under the empirically relevant case when the intertemporal elasticity of substitution remains below or equal to one. As is standard in this literature, in order to compare the results under constant and non-constant discounting, we control the experiment so that both discounting methods show identical overall impatience. Under this assumption, individuals with a decaying level of impatience value the short run lower and the long run higher than individuals with exponential discounting. Thus, the former undervalue the effect of current savings on distant future utility. For this reason a central planner's solution with a decreasing rate of time preference will lead to a lower growth rate and higher investments in abatement than the solution with a constant rate of time preference.

Our main finding is related to the comparison of the social welfare under the central planner and the market economy. Without the possibility of commitment, a sophisticated central planner, anticipating his future inconsistent behavior, plays Markov perfect feedback strategies in a game against his future selves, which will be accepted by all future cohorts. Given this lack of commitment the central planner can be seen as a representative consumer who internalizes the two externalities caused by pollution. If the externality from the producers' denial of the effect of their own generated pollution on production is relatively small in comparison with the negative externality of pollution on utility, then the central planner would welcome a deceleration in economic growth to slow down pollution. This is not taken into consideration in the market economy which, in consequence, might grow faster than the centralized economy. Under these circumstances, and if the discount rate decreases rapidly with the time distance from the present, that is, if the individuals' level of impatience decreases sharply, then a fast growing market economy might provide higher welfare than a benevolent central planner. The intuition is that assuming an identical overall level of impatience, the faster the decay in the instantaneous discount rate, the greater the weight given to the distant future with respect to the near future. Hence, given these preferences, if the market economy grows faster, social welfare is also higher. This result has been proven for a log-utility function.

Finally, when the economy grows faster under the planned economy, or the decay in the discount rate (with the time distance from the present) is slow, then the central planner's solution welfare dominates the decentralized market economy. Therefore, it is appropriate to levy taxes on production and pollution, or to grant a subsidy on abatement activities to bring the market toward the central planner's equilibrium. In thit case, when individuals discount the future at a non-constant rate we prove that these policies are less effective than when consumers discount at a constant rate, at least initially for low values of these taxes or subsidy.

The paper is organized as follows. Section 2 summarizes the main ingredients of the well-established model of pollution abatement and endogenous growth by Smulders and Gradus (1996) and characterizes the balanced path equilibrium for the centralized economy under non-constant discounting. The results are contrasted with the standard assumption of constant discounting. The decentralized economy and the appropriate environmental policies are presented in Section 3. Section 4 analyzes the conditions under which the market economy welfare dominates the central planner's solution, and provides an example for a specific discount function. In contrast, when the centralized solution welfare dominates the market economy, the effectiveness of taxes/subsidy is also compared for constant and non-constant discounting. Finally, Section 5 concludes. Most proofs are presented in the Appendix.

2 A central planner with non-constant discounting

This paper revisits the endogenous growth model with pollution abatement presented in Smulders and Gradus (1996). The economy produces a final output, Y, using capital, K, understood as a broad measure of physical and human capital, knowledge and technology. Productivity of capital is negatively affected by net pollution, P, which is the result of two opposite forces. On the one hand, pollution is a by-product of the economic activity, and hence it shows a positive relation with capital stock. On the other hand, the economy can carry out abatement activities, A, to reduce pollution. Smulders and Gradus (1996) consider the following Cobb-Douglas production and pollution functions.

$$Y(P(A,K),K) = BP^{-\alpha}K^{\beta}, \quad P(A,K) = A^{-\gamma}K^{\lambda}, \quad \alpha,\beta,\gamma,\lambda > 0. \tag{1}$$

Households utility, U(C,P), is an isoelastic utility function which depends positively on consumption and negatively on pollution, that is,

$$U(C,P) = \begin{cases} \dfrac{\left(CP^{-\phi}\right)^{1-1/\sigma} - 1}{1-1/\sigma} & \text{if} \quad \sigma \neq 1, \\ \ln C - \phi \ln P & \text{if} \quad \sigma = 1 \end{cases} \tag{2}$$

Parameter $\sigma > 0$ is the constant intertemporal elasticity of substitution and $\phi > 0$ can be interpreted as a measure of the environmental concern. The intratemporal elasticity of substitution between consumption and pollution equals one. Assuming constant population normalized to one, all variables are in per capita terms.

On a balanced growth path, consumption, capital, abatement and output grow at the same constant rate. Hence, $s_{\rm C} = C/Y$, $s_{\rm A} = A/Y$, and $y = Y/K$ remain constant and a balanced growth path is feasible if the following conditions are satisfied:

$$\beta + \alpha(\gamma - \lambda) = 1, \tag{3}$$

$$\lambda \leq \gamma. \tag{4}$$

Note that, according to (1), the output per unit of capital $y = Y/K$ can be written as a function of s_{A} as

$$y(s_{\text{A}}) \doteqdot \frac{Y}{K} = (Bs_{\text{A}}^{\alpha\gamma})^{\frac{1}{1-\alpha\gamma}}. \tag{5}$$

As in Smulders and Gradus (1996), we assume that $1 - \alpha\gamma > 0$, implying a positive relationship between y and s_{A}.

2.1 The central planner's solution

Smulders and Gradus (1996) study this model under the standard assumption of individuals discounting the future at a constant rate. In contrast, in the present paper, this stylized model is analyzed assuming that individuals discount the future at a decreasing rate: their degree of impatience decreases with the time distance from the present. As Krusell et al. (2002) acknowledge, when successive cohorts disagree it is difficult to determine a suitable notion for central planner. However, we follow Krusell *et al.* (2002) and Hiraguchi (2014), and assume that the central planner's preferences are the same as those of the current self, who plays a game with his future selves in order to maximize his welfare. That is, he behaves sophisticatedly in order to maximize discounted utility, considering a time-consistent solution that will also be optimal for his future selves.

Thus, at each time t the central planner maximizes the lifetime utility of the representative consumer (t-agent), subject to the dynamic evolution of the capital stock from this instant on:

$$\max_{C_t(\tau), A_t(\tau)} \int_t^{\infty} U\left(C_t(\tau), P_t(\tau)\right) \theta(\tau - t)\, d\tau, \tag{6}$$

$$\text{s.t.:} \quad \dot{K}_t(\tau) = BP_t^{-\alpha}(\tau)K_t^{\beta}(\tau) - C_t(\tau) - A_t(\tau), \quad K_t(t) = K_t, \tag{7}$$

where the utility function is given by (2) and $P_t(\tau) = A_t^{-\gamma}(\tau)K_t^{\lambda}(\tau)$. In this problem t represents current date, and $j = \tau - t$ the time distance from the present.

Time preference is measured by the discount function $\theta(j) \geq 0$, which is not a function of current time, but of the time distance from the present, j. The discount function satisfies $\theta(j) > 0$, $\dot{\theta}(j) < 0$, $\forall j \geq 0$ and $\theta(0) = 1$. The instantaneous discount rate $\rho(j) \doteqdot -\dot{\theta}(j)/\theta(j)$ is not a constant, but decreases with the time distance from the present: $\rho(j) > 0$, $\dot{\rho}(j) < 0, \forall j \geq 0$. Finally, $\lim_{t\to+\infty} \rho(j)$ can be strictly positive (quasi-hyperbolic discounting) or null (hyperbolic discounting).

In what follows, and in order to compare the results under non-constant discounting with the standard results with exponential discounting, we need to guarantee that the dissimilarities are not due to different degrees of impatience (see, for example, Strulik 2015 and Cabo et al. 2015). Therefore, both discounting methods have to be controlled to show identical overall impatience. That is, discounting at the constant rate, $\hat{\rho} > 0$, has to be equivalent to using the discount function $\theta(j)$:

$$\int_0^{\infty} e^{-\hat{\rho} j}\, dj = \int_0^{\infty} \theta(j)\, dj, \quad \Longleftrightarrow \quad \hat{\rho} = \left[\int_0^{\infty} \theta(j)\, dj\right]^{-1}. \tag{8}$$

That is, the constant discount rate has to be equal to the inverse of the integral $\int_0^{\infty} \theta(j)\, dj$, which can be interpreted as a measure of the overall impatience. This condition can only hold true if this integral is convergent (i.e. $\theta(j)$ decreases with j faster than $1/j$), assumed henceforth. Under non-constant discounting the instantaneous discount rate is initially large and decreases with the time distance from the present. Then, from the assumption of identical overall impatience, the short run is less valued and the long run more valued than under exponential discounting (see Figure 1).

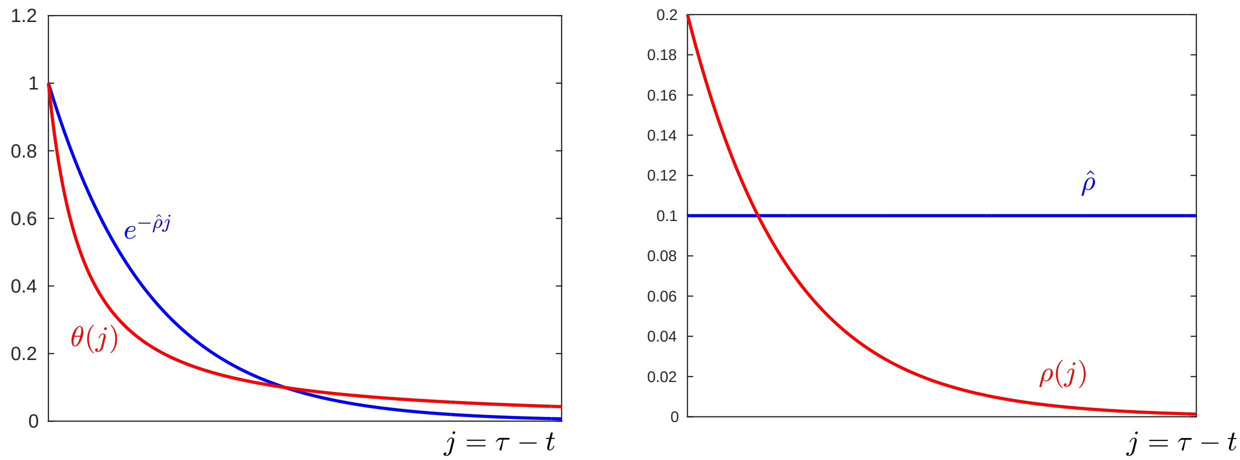

Figure 1: Constant/non-constant discount functions & instantaneous discount rates

A sophisticated t-agent looking for a time-consistent solution, must solve problem (6)–(7) playing a game against his future selves. Following Karp (2007), the optimality of a stationary solution of this problem means satisfaying the following Bellman equation,

$$\int_t^{\infty} U\left(C_t^*(\tau), P_t^*(\tau)\right) \theta(\tau - t)\, \rho(\tau - t)\, d\tau = \max_{C_t, A_t} \left\{ U(C_t, P(A_t, K_t)) + W'(K_t)[Y(P(A_t, K_t), K_t) - C_t - A_t] \right\}, \tag{9}$$

where $C_t^*(\tau)$ is the consumption and $P_t^*(\tau) = P^*(A_t^*(\tau), K_t^*(\tau))$ the pollution along the optimal path for the t-agent, and $W(K_t) = \int_t^{\infty} U\left(C_t^*(\tau), P_t^*(\tau)\right) \theta(\tau - t)\, d\tau$, denotes the value function. Subscript t relates to the t-agent, who solves the maximization problem starting at current time t. In the RHS of (9) the argument for these variables is this current time t. For the ease of presentation we omit this argument when no confusion can arise.

First-order conditions for optimality establish that

$$U'_{\mathrm{C}_t} = W'(K_t), \tag{10}$$

$$U'_{\mathrm{P}_t} P'_{\mathrm{A}_t} = W'(K_t)\left(1 - Y'_{\mathrm{P}_t} P'_{\mathrm{A}_t}\right). \tag{11}$$

And hence: $U'_{\mathrm{P}_t} P'_{\mathrm{A}_t} = U'_{\mathrm{C}_t}(1 - Y'_{\mathrm{P}_t} P'_{\mathrm{A}_t})$, which can be rewritten as

$$A_t = \frac{s_{\mathrm{A}} \phi \gamma}{s_{\mathrm{A}} - \alpha \gamma} C_t \quad \Leftrightarrow \quad s_{\mathrm{A}} = \gamma(\phi s_{\mathrm{C}} + \alpha), \tag{12}$$

establishing a positive relation between the consumption and the abatement expenditure shares, provided that $s_{\mathrm{A}} > \alpha\gamma$, which is assumed henceforth.

Moreover, as the following proposition proves, the optimality of a balanced growth path requires satisfying a modified Ramsey rule. This rule is obtained under the assumption that the intertemporal elasticity of substitution σ is lower than or equal to one. This assumption of an income effect stronger than or equal to the substitution effect is the empirically relevant case.[5]

Proposition 1 *For $\sigma < 1$, the growth rate of the central planner's economy along a balanced path equilibrium for problem (6)–(7) must satisfy the modified Ramsey rule:*

$$g \doteq \frac{\dot{C}}{C} = \frac{\left(\beta - s_{A} \frac{\lambda}{\gamma}\right) y(s_{A}) - \Delta(g, \eta)}{1 + \eta}, \tag{13}$$

[5] As will be shown in the next proposition, this is a sufficient condition for the convergence of the improper integrals which collect the present value of the stream of utility within an infinite time horizon. Additional conditions should be imposed when the intertemporal elasticity of substitution is greater than one to guarantee convergence. However, this is outside the scope of this paper.

with

$$\eta = \frac{1-\sigma}{\sigma}[1+\phi(\gamma-\lambda)] > 0 \tag{14}$$

and

$$\Delta(g,\eta) = \int_0^\infty \rho(j)\omega(j)\,dj, \text{ with } \omega(j) = \frac{\theta(j)e^{-\eta g j}}{\int_0^\infty \theta(i)e^{-\eta g i}di} \in (0,1). \tag{15}$$

Correspondingly, for $\sigma = 1$, the modified Ramsey rule particularizes as:[6]

$$g^1 = \left(\beta - s_A\frac{\lambda}{\gamma}\right) y(s_A) - \hat{\rho} \tag{16}$$

Proof. See the Appendix □

Remark 1 *Note that for $\sigma = 1$, then $\eta = 0$ and $\Delta(g,0) = \int_0^\infty \rho(j)\omega_0(j)\,dj = \hat{\rho}$, with $\omega_0(j) = \theta(j)/\int_0^\infty \theta(i)di \in (0,1)$. Then, (16) is a particular case of (13) when $\sigma = 1$.*

Remark 2 *Note that if $g > 0$, the expression of $\Delta(g,\eta)$ is well defined if $\sigma \leq 1$. For $\sigma > 1$ additional conditions on $\theta(j)$ would be required. However, this case is outside the scope of this paper.*

In the standard case of constant discounting, that is, $\rho(j) \equiv \hat{\rho}$, the integral $\Delta(g,\eta)$ boils down to $\hat{\rho}$. Therefore, the growth rate along the balanced path in expressions (13) or (16) reduce to the Ramsey rule found in Smulders and Gradus (1996):

$$g_{\hat{\rho}} = \frac{\left(\beta - s_A\frac{\lambda}{\gamma}\right) y(s_A) - \hat{\rho}}{1+\eta}, \quad g^1_{\hat{\rho}} = \left(\beta - s_A\frac{\lambda}{\gamma}\right) y(s_A) - \hat{\rho}, \tag{17}$$

where subscript $\hat{\rho}$ refers to the standard exponential discounting.

While the Ramsey rule with constant discounting explicitly defines the growth rate as a function of the abatement expenditure ratio, the equivalent modified Ramsey rule in the case with non-constant discounting in (13) defines the growth rate of the economy implicitly. The implicit relationship is proven in Proposition 10 in the Appendix.

The social rate of return of the economy is given by $\left(\beta - s_A\frac{\lambda}{\gamma}\right) y(s_A)$, as a function of the abatement ratio. This function is identical regardless of whether the future is discounted at a constant or at a decreasing rate:

$$r(s_A) = \left(\beta - s_A\frac{\lambda}{\gamma}\right) y(s_A). \tag{18}$$

On comparing (13) and (17), the role played by the constant temporal discount rate in (13) is played by $\Delta(g,\eta)$ in the case of non-constant discounting. Therefore, the latter is henceforth denoted as the effective rate of time preference.

In what follows, we compare the modified Ramsey curve under non-constant discounting (13) with the standard Ramsey curve (17) under exponential discounting, which corresponds to the comparison between $\Delta(g,\eta)$ and $\hat{\rho}$. This comparison is based on three findings. In the following subsections, it is shown that if the intertemporal elasticity of substitution is lower than one and the economy grows at a positive constant rate, then the central planner underestimates the effect of current savings on future utility. Second, for our modified Ramsey rule in (13) or (16) this is equivalent to having an effective discount rate greater than $\hat{\rho}$, which is "as if" the central planner discounts the future at a higher but constant rate, $\Delta(g,\eta) > \hat{\rho}$. Finally, this immediately establishes that non-constant discounting is associated with a modified Ramsey curve placed below the standard Ramsey under constant discounting. That is, a lower willingness to grow for any abatement output ratio.

[6] Here and henceforth, the particular case $\sigma = 1$ is highlighted with a superscript.

2.2 Effect of current decisions on ongoing utility

Following standard reasoning, as for example in Barro (1999), the discounted utility from the viewpoint of a t-agent, $\int_t^\infty U(C_t(\tau), P_t(\tau))\theta(\tau - t)d\tau$, with $P_t(\tau) = A_t^{-\gamma}(\tau)K_t^{\lambda}(\tau)$, can be decomposed in the short-run and the long-run terms, assuming a current interval $[t, t+\varepsilon]$, of length ε:

$$U(C_t(t), P_t(t))\varepsilon + \int_{t+\varepsilon}^{\infty} U(C_t(\tau), P_t(\tau))\theta(\tau - t)d\tau. \tag{19}$$

Along the balanced path, where consumption, abatement and the capital stock all grow at a constant rate g, utility at any future time τ from the viewpoint of a t-agent can be written as a function of consumption, abatement and capital at present time t, the constant growth rate of the economy, the time distance from the present, $\tau - t$, and parameters:

$$U(C_t(t), P_t(t), \tau) = \begin{cases} \dfrac{\left[C_t(t)P_t^{-\phi}(t)\right]^{1-\frac{1}{\sigma}} e^{-\eta g(\tau-t)} - 1}{1-\frac{1}{\sigma}} & \text{if} \quad \sigma < 1, \\ \ln(C_t(t)) - \phi \ln(P_t(t)) + g^1[1+\phi(\gamma-\lambda)](\tau - t) & \text{if} \quad \sigma = 1. \end{cases} \tag{20}$$

It is now clear that although growing at a positive constant rate, the assumption of $\sigma < 1$ implies a decreasing utility as indicated by the term $e^{-\eta g(\tau-t)}$. Given this expression, the decomposition in (19) can be rewritten as

$$U(C_t(t), P_t(t), t)\varepsilon + DU(C_t(t+\varepsilon), P_t(t+\varepsilon), t+\varepsilon),$$

where $DU(C_t(t+\varepsilon), P_t(t+\varepsilon), t+\varepsilon)$ is the value, at time t, of ongoing utility from $t+\varepsilon$ on, discounted by $\theta(\tau - t)$, in (19).

Optimum consumption requires the marginal utility from current consumption (instantaneous marginal increment in utility with a rise in consumption at present time t, times ε) to equate the discounted gains in future utility from a current increment in savings. To compute this, note that a marginal increment in savings at present time t produces an immediate increment in the capital stock at the subsequent time, $K(t+\varepsilon)$. This greater capital stock worsens pollution, so reducing utility. Moreover, it also has two other effects on utility: It rises consumption $C(t+\varepsilon)$, which directly increases utility, and it raises abatement $A(t+\varepsilon)$, which increases utility by reducing pollution:

$$-\frac{\partial K_t(t+\varepsilon)}{\partial C_t(t)} \left\{ \frac{\partial C_t(t+\varepsilon)}{\partial K_t(t+\varepsilon)} \frac{\partial DU(\bullet, t+\varepsilon)}{\partial C_t(t+\varepsilon)} + \frac{\partial A_t(t+\varepsilon)}{\partial K_t(t+\varepsilon)} \frac{\partial DU(\bullet, t+\varepsilon)}{\partial A_t(t+\varepsilon)} + \frac{\partial DU(\bullet, t+\varepsilon)}{\partial K_t(t+\varepsilon)} \right\}. \tag{21}$$

Along a balanced growth path we conjecture constant propensities to consume and to abate out of wealth,[7] Λ^{C} and Λ^{A}. Moreover, from the optimality condition in (12), it immediately follows that $\Lambda^{\mathrm{A}} = s_{\mathrm{A}}\phi\gamma/(s_{\mathrm{A}} - \alpha\gamma)\Lambda^{\mathrm{C}}$. To compute the effect of increments in consumption, abatement and capital at time $t+\varepsilon$ on ongoing utility along the balanced path, we first compute their marginal effects on the utility at a specific later time $\tau \geq t+\varepsilon$:

$$\frac{\partial U(\bullet, \tau)}{\partial C_t(t+\varepsilon)} = \left[C_t(t+\varepsilon)P_t^{-\phi}(t+\varepsilon)\right]^{-\frac{1}{\sigma}} P_t^{-\phi}(t+\varepsilon)e^{-\eta g(\tau-t-\varepsilon)} = MU(t+\varepsilon)e^{-\eta g(\tau-t-\varepsilon)},$$
$$\frac{\partial U(\bullet, \tau)}{\partial A_t(t+\varepsilon)} = \frac{s_{\mathrm{A}} - \alpha\gamma}{s_{\mathrm{A}}} MU(t+\varepsilon)e^{-\eta g(\tau-t-\varepsilon)}, \quad \frac{\partial U(\bullet, \tau)}{\partial K_t(t+\varepsilon)} = -\phi\lambda\Lambda^{\mathrm{C}} MU(t+\varepsilon)e^{-\eta g(\tau-t-\varepsilon)},$$

with $MU(t+\varepsilon)$ the instantaneous marginal effect of consumption on utility at time $t+\varepsilon$. Since isolated increments of consumption, abatement and capital at time $t+\varepsilon$ continue growing at rate g from this time on, a one unit variation of utility at the present time induces a modification of $e^{-\eta g(\tau-t-\varepsilon)}$, $\tau - t - \varepsilon$ instants in the future. This term defines how current effects are transmitted into the future along a balanced growth path. The assumption of $\sigma < 1$ (together with $g > 0$) implies that the ever growing increments in consumption, abatement and capital have, nonetheless, a decreasing marginal effect on utility due to the strong effect of satiation.

[7] Because labor is ignored, individuals' wealth is exclusively defined by their assets, which in aggregate terms are equal to the capital stock.

Conversely, with log-utility, $\eta = 0$, and the growing increments in consumption, abatement and capital are exactly offset by satiation. In consequence, current marginal increments at time $t + \varepsilon$ have a constant effect on future utility.[8] Likewise, with a zero growth rate we observe a step effect and current changes again have a permanent constant effect on future utility.

Adding from $t+\varepsilon$, on, the total ongoing effect of a marginal increment in savings at time t in (21) can be rewritten as

$$\varepsilon \Lambda^{C}[1+\phi(\gamma-\lambda)]MU(t+\varepsilon)\int_{t+\varepsilon}^{\infty} e^{-\eta g(\tau-t-\varepsilon)}\theta(\tau-t)d\tau. \tag{22}$$

Taking the limit when ε tends to 0, factor in (22) denoted by $\Omega(g,\eta)$ (for non-constant discounting) or $\Omega_{\hat{\rho}}(g,\eta)$ (for exponential discounting), can be written as

$$\Omega(g,\eta) = \int_0^{\infty} e^{-\eta g j}\theta(j)dj, \quad \Omega_{\hat{\rho}}(g,\eta) = \int_0^{\infty} e^{-(\eta g+\hat{\rho})j}dj = \frac{1}{\eta g+\hat{\rho}} \le \frac{1}{\hat{\rho}}. \tag{23}$$

This term collects information on how a marginal rise in consumption, abatement and capital at the present time influences utility in all future times along the balanced path, valued at the present time. If the income and the substitution effects are identical (under log-utility), or there is no growth, current marginal effects remain constant at any future time. Therefore, the global effect would be exclusively determined by the overall level of impatience, $\Omega(g,0) = \Omega(0,\eta) = \Omega_{\hat{\rho}}(g,0) = \Omega_{\hat{\rho}}(0,\eta) = 1/\hat{\rho}$.

The value a t-agent assigns to the marginal effect of current increments in consumption on future utility, $e^{-\eta g j}\theta(j)$, decreases faster, the lower the intertemporal elasticity of substitution below one, or the faster the growth rate along the balanced path. Hence, the accumulated effect is lower, the lower the individuals' willingness to substitute present for future consumption and the faster the growth rate of the economy, as stated in next lemma.

Lemma 1 *If $\sigma < 1$ and $g > 0$, when the future is discounted at a non-constant rate, $\Omega(g,\eta)$ in (23) satisfies*

$$\frac{d\Omega}{dg} = -\eta \bar{J}\Omega(g,\eta) < 0, \qquad \frac{d\Omega}{d\eta} = g\bar{J}\Omega(g,\eta) \ge 0 \; (> 0 \text{ if } g > 0), \tag{24}$$

with $\bar{J} = \int_0^{\infty} j\omega(j)\,dj \in (0,\infty), \quad \omega(j) = \dfrac{\theta(j)e^{-\eta g j}}{\Omega(g,\eta)} \in (0,1).$

From the definitions in (23), $\Omega(g,\eta)$ and $\Omega_{\hat{\rho}}(g,\eta)$ can be compared taking into account two further facts. First, the marginal effect of current increments in consumption, abatement and capital on future utility, $e^{-\eta g j}$, decreases with the time distance from the present. And second, due to the assumption of an identical overall level of impatience, individuals who discount the future at a non-constant discount rate give less value to the short run and more value to the long run than standard individuals with constant discounting (see Figure 1). These two facts lead to the following result:

Corollary 1 *A central planner with an intertemporal elasticity of substitution lower than one in a growing economy who discounts the future at a non-constant rate underestimates the effect of current savings on ongoing utility. This underestimation disappears if the growth rate of the economy tends to zero or the intertemporal elasticity of substitution tends to one:*

$$\Omega(g,\eta) < \Omega_{\hat{\rho}}(g,\eta), \quad g,\eta > 0, \quad \text{and} \quad \lim_{\sigma\to 1^-}\Omega(g,\eta) = \lim_{g\to 0^+}\Omega(g,\eta) = \frac{1}{\hat{\rho}}. \tag{25}$$

2.3 Comparing constant and non-constant discounting

It is now possible to give an interpretation of the effective rate of time preference, $\Delta(g,\eta)$, defined in (15) as a weighted mean of the decaying instantaneous discount rates, $\rho(j)$, with weights $\omega(j)$ satisfying $\int_0^{\infty}\omega(j)dj = 1$.

[8] For individuals with a very high intertemporal elasticity of substitution, $\sigma > 1$: as the increments in consumption, abatement and capital grow with the time distance from the present, they are not fully counterbalanced by satiation and therefore produce a growing effect on utility.

The weight $\omega(j)$ represents the value that the marginal rise in current consumption has on the utility j instants from now on (in relative terms to the accumulated effect on ongoing utility at all future times). Moreover, integrating by parts in the expression in (15) makes it possible to re-write the effective rate of time preference as the inverse of the present value of the effect of a marginal increment in current utility on all future utilities, diminished by the speed of decay of the marginal effect of current changes on future utility:

$$\Delta(g,\eta) = \frac{1}{\Omega(g,\eta)} - \eta g. \tag{26}$$

From this definition and the expressions in (25), the next proposition compares the effective discount rate and the propensity to consume out of wealth under constant and non-constant discounting. It distinguishes between the extreme cases $\sigma = 1$ or $g = 0$, and the general case with $\sigma < 1$ and $g > 0$.

Proposition 2 *For $\sigma < 1$ and $g > 0$:*

$$\Delta(g,\eta) > \hat{\rho}, \quad \textit{and} \quad \Lambda^c(g,\eta) > \Lambda^c_{\hat{\rho}}(g,\eta). \tag{27}$$

For $\sigma = 1$ or $g = 0$:

$$\Delta(g,0) = \Delta(0,\eta) = \hat{\rho}, \ \textit{and}\ \Lambda^c(g,0) = \Lambda^c(0,\eta) = \Lambda^c_{\hat{\rho}}(g,0) = \Lambda^c_{\hat{\rho}}(0,\eta) = \frac{\hat{\rho}}{1+\phi(\gamma-\lambda)}. \tag{28}$$

Proof. Results $\Delta(g,\eta) > \hat{\rho}$ and $\Delta(g,0) = \Delta(0,\eta) = \hat{\rho}$ follows from (25) and (26).

The propensity to consume out of wealth, $\Lambda^{\mathrm{C}}(g,\eta)$, can be computed by equating the instantaneous marginal utility of consumption to the marginal effect of savings on future utility in (22) and computing the limit when ε tends to zero:

$$\Lambda^{\mathrm{C}}(g,\eta) = \frac{1}{[1+\phi(\gamma-\lambda)]\Omega(g,\eta)}, \quad \Lambda^{\mathrm{C}}_{\hat{\rho}}(g,\eta) = \frac{1}{[1+\phi(\gamma-\lambda)]\Omega_{\hat{\rho}}(g,\eta)}. \tag{29}$$

Then, the results $\Lambda^{\mathrm{C}}(g,\eta) > \Lambda^{\mathrm{C}}_{\hat{\rho}}(g,\eta)$ and $\Lambda^{\mathrm{C}}(g,0) = \Lambda^{\mathrm{C}}(0,\eta) = \hat{\rho}/([1+\phi(\gamma-\lambda)])$ follow straightforwardly from (25) and (29). □

In the extreme cases of $\sigma = 1$ or $g = 0$, there is no decay in the marginal effect of current consumption on future utility,[9] and taking into account the assumption of identical overall impatience, the effective discount rate coincides with the constant discount rate $\hat{\rho}$. Likewise, in these two extreme cases the propensity to consume out of wealth coincides under constant and non-constant discounting.

In the general case of $\sigma < 1$ and a positive growth rate, current effects of present decisions on future utility decrease with the time distance from the present. Therefore, the future is valued less under non-constant than under exponential discounting, and hence, the effective rate of time preference is higher, $\Delta(g,\eta) > \hat{\rho}$. With non-constant discounting, the marginal effect of current savings on future consumption is under-valued and, hence, the propensity to consume out of wealth would be higher than under exponential discounting, $\Lambda^C(g,\eta) > \Lambda^C_{\hat{\rho}}(g,\eta)$. From (29) and Lemma 1 it follows that the propensity to consume out of wealth increases with the growth rate of the economy, driven by a stronger income than substitution effect. The increment in $\Lambda^{\mathrm{C}}(g,\eta)$ is more intense, the higher η is, i.e. the lower below one σ is .

Consequently, when the discount rate for future utility decreases with the time distance from the present the modified Ramsey rule lies below the Ramsey curve with constant discounting. Likewise, as in the standard case of constant discounting, the following proposition shows that the modified Ramsey rule displays an inverted U-shape as a function of the abatement expenditure ratio. Furthermore, the curve is softer and lies below the Ramsey curve with constant discounting.

[9] The weights $\omega(j)$ only depend on the discount function $\theta(j)/\int_0^\infty \theta(j)dj$.

Proposition 3 *When $\sigma < 1$ and $g > 0$, the growth rate of the economy in (13) is an inverted U-shaped curve with a maximum for $\hat{s}_A = \alpha\beta\gamma^2/\lambda$ independently of the rate of decay of the instantaneous discount rate (indeed, $\hat{s}_A$ is the same as in the case with constant discounting). Moreover, for every feasible s_A the growth rate of the economy is lower and the marginal effect of s_A on g, either positive or negative, is always softer than under exponential discounting.*

For $\sigma = 1$, the Ramsey curves under constant and non-constant discounting coincide.

Proof. See the Appendix □

These results are displayed in Figure 2. The standard Ramsey curve under constant discounting is displayed by the inverted U-shaped dashed blue curve. The modified Ramsey rule under non-constant discounting, in solid red, lies below and although still inverse U-shaped, it shows gentler increments and decrements with s_{A}.

To fully characterize the equilibrium growth rate, Equation (7) should also be taken into account. Replacing (12) into (7) the growth rate of the capital stock can be written as

$$g = (1 - s_{\text{C}} - s_{\text{A}})y(s_{\text{A}}) = \left(1 + \frac{\alpha}{\phi} - \frac{1 + \phi\gamma}{\phi\gamma} s_{\text{A}}\right) y(s_{\text{A}}). \tag{30}$$

This curve (dashed-dotted black curve in Figure 2) represents the balanced budget or the feasible set of balanced paths in the terminology of Smulders and Gradus (1996). A balanced growth path equilibrium must satisfy the modified Ramsey rule (13), as well as the constraint inherited from the evolution of the capital stock as expressed in (30). The equilibrium defined by these two curves is displayed by the red dot, (s^*_{A}, g^*), in Figure 2. Likewise, the figure also plots the equilibrium under constant discounting, $(s^*_{\text{A}\hat{\rho}}, g^*_{\hat{\rho}})$, as the cross between the budget constraint in (30) and the Ramsey rule in (16). These equilibria can lie in the upward- or the downward-sloping part of the Ramsey curves. In either case, the assumption of non-constant discounting implies a balanced growth path characterized by a lower growth rate and a greater abatement-output ratio. This result is established in next proposition.

Proposition 4 *For $\sigma < 1$, under the assumption of identical overall impatience in (8), if the future is discounted at a decreasing rate, the central planner's optimal growth rate would be lower and the abatement spending ratio higher than if constant discounting applies:*

$$g^* < g^*_{\hat{\rho}}, \qquad s^*_A > s^*_{A\hat{\rho}}. \tag{31}$$

For $\sigma = 1$ the two equilibria coincide, $g^{1*} = g^{1*}_{\hat{\rho}}$ and $s^{1*}_A = s^{1*}_{A\hat{\rho}}$.

Proof. See the Appendix □

Remark 3 *Following the notation in Strulik (2015), we have proved that if $\sigma < 1$ the models with constant and non-constant discounting are observationally equivalent in the sense that under both discounting methods along the balanced path the growth rate of the economy is constant, and abatement, consumption, capital and output grow at the same constant rate, while pollution decreases at a constant rate. However, the models are not strongly equivalent because they lead to different economic growth rates, despite the assumption of identical overall impatience.*

2.4 Effect of greener preferences on the optimal solution

Next we analyze the effect of greener preferences, as reflected by a higher ϕ, on the growth rate of the economy, and compare this effect under constant and non-constant discounting. The equilibria under both scenarios come from the intersection of the standard and the modified Ramsey curves in (17) and (13) with the curve in (30), as shown in Figure 2.

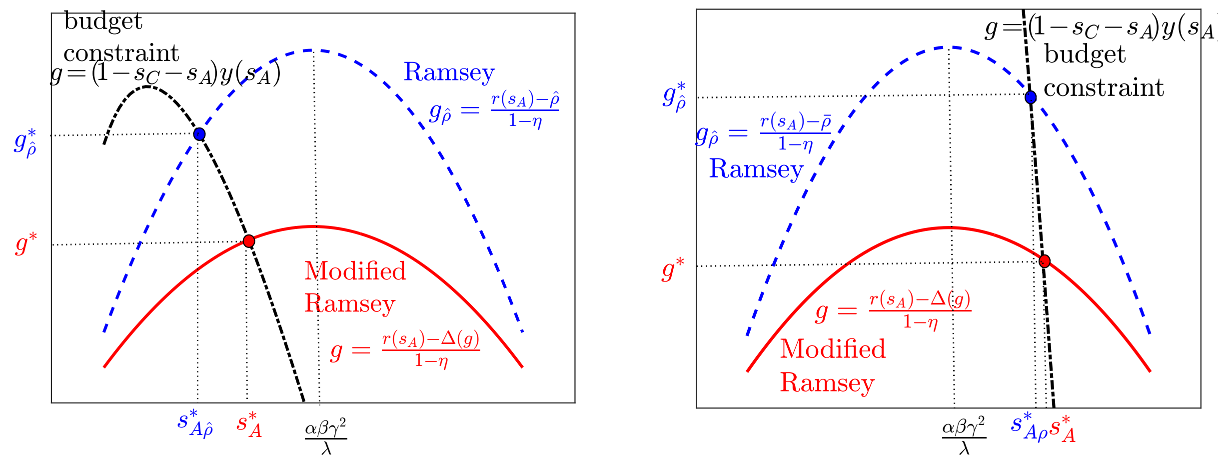

Figure 2: Equilibria under constant & non-constant discounting

In general, a rise in ϕ provokes a downward shift in the Ramsey curves and an up-right shift in the feasible set or budget constraint. Therefore, a rise in ϕ when the equilibria are located in the downward sloping part of the Ramsey curves inevitably leads to a reduction in the growth rate of the economy under both discounting methods. By contrast, when the equilibria are in the upward sloping part of the Ramsey curves, the growth rate might increase, although it is generally not possible to characterize the circumstances which lead to this result.

In general, when equilibria occur in the downward sloping part of the Ramsey curves, we have not succeeded in characterizing under which scenario growth is more strongly reduced. In contrast, when the equilibria lie in the upward sloping part, the slope of the standard Ramsey curve under constant discounting at $s^*_{\mathrm{A}\hat{\rho}}$ is steeper than the slope of the modified Ramsey curve under non-constant discounting at $s^*_{\mathrm{A}\hat{\rho}}$, which, in its turn, is steeper than its slope at s^*_{A}. Thus, the up-right shift of the feasible set induces a stronger rise in the growth rate of the economy in the case of constant discounting. Nevertheless, when $\gamma > \lambda$, greener preferences would shift the Ramsey curves downward, and more strongly than under constant discounting. Therefore, the net effect is unclear and it is not possible to conclude whether greener preferences lead to a faster or a slower growth, or whether the effect is stronger under constant or non-constant discounting.

We can, however, compare the effect of a higher ϕ in the limiting case when the environment has almost no effect on utility. For an initially low ϕ, greener preferences imply a rise in the growth rate of the economy, and this rise is less sharp when individuals discount the future at a decreasing rate.

Proposition 5 *In a society with a low environmental concern, ϕ close to zero, a higher ϕ leads to a rise in both $g^*_{\hat{\rho}}$ and g^*. Moreover, the effect is softer when the future is discounted at a decreasing rate:*

$$\lim_{\phi\to 0} \frac{dg^*_{\hat{\rho}}}{d\phi} > \lim_{\phi\to 0} \frac{dg^*}{d\phi} > 0.$$

Proof. See the Appendix □

3 The market economy

This section presents the market or decentralized economy considering that time preferences are identical to those previously explained for the central planner. Recall that the central planner has been interpreted as a representative agent who internalizes the two externalities caused by pollution. Consumption decisions are taken by dynamic consumers whose time preferences decrease with the time distance from the present. Correspondingly, abatement is chosen by static maximizing firms. Once the decentralized equilibrium is characterized, we compare the growth rate and the social welfare with those associated with the central planner's solution. When appropriate, we also analyze the effect of the environmental policies put into practice by a benevolent government and compare them against the case of exponential discounting.

We consider an economy populated by a large number of identical firms of measure one. Firm i uses only capital as input.[10] In the absence of any effect of pollution on production, output would be exclusively dependent on the capital stock. Nevertheless, production activities (measured by the capital involved) of firm i generates pollution, $P_i(t)$, which can be reduced by the abatement activities carried out by this firm $A_i(t)$:

$$P_i(t) = A_i^{-\gamma}(t) K_i^{\lambda}(t).$$

Following Smulders and Gradus (1996), the aggregate pollution generated by all firms, $P(t) = \int_0^1 P_i(t)di$ negatively affects each firm's productivity. However, each particular firm only internalizes part of its own pollution. Therefore, firm's i production is not exclusively a function of the capital input, but also of global pollution, and in particular its own pollution which is partially internalized:[11]

$$Y_i = BK_i^{\beta} P_i^{-\tilde{\alpha}} P^{-(\alpha-\tilde{\alpha})}.$$

Elasticity $\tilde{\alpha} < \alpha$ measures to what extent firm i internalizes the effects of its own pollution on production. In this decentralized setting, Smulders and Gradus (1996) introduce taxes in production and pollution, and subsidies to abatement in order to bring the decentralized equilibrium towards the central planner's solution. Thus, firm's profit can be written as

$$(1 - T_{\text{Y}})Y_i - rK_i - wL_i - (1 - T_{\text{A}})A_i - T_{\text{P}}P_i,$$

where r is the rental price to capital, T_{A} is an abatement subsidy rate, T_{P} is a pollution tax rate and T_{Y} is a value-added tax rate. Under perfect competition, each firm chooses the abatement level as well as the amount of capital in order to maximize its profits. The first-order conditions for optimality are

$$\tilde{s}_{\text{A}} = \frac{(1 - T_{\text{Y}})\tilde{\alpha}\gamma + \gamma\hat{T}_{\text{P}}}{1 - T_{\text{A}}}, \tag{32}$$

$$\tilde{r}(s_{\text{A}}) = [(1 - T_{\text{Y}})(\beta - \tilde{\alpha}\lambda) - \lambda\hat{T}_{\text{P}}] y_i(s_{\text{A}}), \tag{33}$$

where $\hat{T}_{\text{P}} = T_{\text{P}}P_i/Y_i$. Note that the behavior of firms is independent of consumers' discounting. Therefore, the value of s_{A}, as a function of taxes and parameters, is the same as in the case of constant discounting.

In a pure market economy with $T_{\text{Y}} = T_{\text{P}} = T_{\text{A}} = 0$, the spending ratio on abatement is lower than in a centralized solution (this is straightforward from the comparison of Equations (12) and (32)). As pointed out by Smulders and Gradus (1996), in the decentralized scenario producers disregard to some extent the effect of abatement on production as well as the effect of pollution on consumers' utility. The gap between the abatement spending ratios can be narrowed by the subsidies to abatement or the tax on pollution, but it is contrarily widened with the tax on production.

Since all firms are identical, they take the same decisions and, being of measure one, $K = \int_0^1 K_i di = K_i$ and $A = \int_0^1 A_i di = A_i$. Consequently, $P = P_i = A^{-\gamma}K^{\lambda}$, and the aggregate production function reads $Y(t) = \int_0^1 Y_i(t)di = BK^{\beta}P^{-\alpha}$, as given in (1).

The economy is populated by a large number of identical consumers of measure one. The representative n-th consumer, $n \in [0, 1]$ maximizes, at each t, his lifetime utility subject to his budget constraint:

[10] Alternatively, we could also include labor as an input. Endogenous growth could be attained assuming learning by doing and knowledge spillovers as, for example, explained in Barro and Sala-i-Martin (2004). To have a better insight on the results of the model, we analyze a simpler version without labor, hence ignoring the knowledge externality and focusing only on the two externalities caused by pollution (on production and utility).

[11] Henceforth we omit the time argument when no confusion can arise. A tilde refers to a result in the market economy as opposed to the central planner's solution.

$$\max_{c_{nt}(\tau)} \int_t^{\infty} U\left(c_{nt}(\tau), P_t(\tau)\right) \theta(\tau - t)\, d\tau, \tag{34}$$

$$\text{s.t.:} \qquad \dot{k}_{nt}(\tau) = r k_{nt}(\tau) - c_{nt}(\tau), \quad k_{nt}(t) = k_{nt}, \tag{35}$$

where $k_{nt}(\tau)$ denotes the accumulated capital for the n-th consumer, and r denotes the market interest rate. Note that with no labor, consumers wealth is exclusively determined by the capital stock they own and lend to firms.

The consumer's optimization problem is solved along the balanced path equilibrium. For such an equilibrium, it easily follows from (32) that the rate of return remains constant (like the marginal productivity of capital). This balanced path, characterized by a constant abatement expenditure ratio, requires a constant $\hat{T}_{\text{P}}$, or equivalently, a tax on pollution increasing at the same speed as the ratio pollution per unit of output decreases.

Assuming that consumers are sophisticated (time-consistent), the optimal growth rate of consumption along a balanced path equilibrium is governed by a modified Ramsey rule, similar to that obtained for the central planner's scenario.

Proposition 6 *For $\sigma < 1$, the growth rate of a market economy along a balanced path equilibrium for problem (34)–(35) must satisfy the modified Ramsey rule:*

$$\tilde{g} = \frac{r - \Delta(\tilde{g}, \eta)}{1 + \eta}, \tag{36}$$

where η and $\Delta(g, \eta)$ are given by (14) and (15).

Correspondingly, for $\sigma = 1$, the modified Ramsey rule reads

$$\tilde{g}^1 = r - \hat{\rho}. \tag{37}$$

Proof. See the Appendix □

Provided that all firms behave identically, the rate of return is given by (33), and the modified Ramsey rule in the decentralized economy reads

$$\tilde{g} = \frac{\tilde{r}(s_{\text{A}}) - \Delta(\tilde{g}, \eta)}{1 + \eta}, \qquad \tilde{g}^1 = \tilde{r}(s_{\text{A}}) - \hat{\rho}. \tag{38}$$

Correspondingly, the standard Ramsey curves under constant discounting would be

$$\tilde{g}_{\hat{\rho}} = \frac{\tilde{r}(s_{\text{A}}) - \hat{\rho}}{1 + \eta}, \qquad \tilde{g}^1_{\hat{\rho}} = \tilde{r}(s_{\text{A}}) - \hat{\rho}. \tag{39}$$

Note that from (33) an increase in abatement monotonously raises the market rate of return $\tilde{r}'(s_{\text{A}}) > 0$, but it also has an effect on $\Delta(\tilde{g}, \eta)$. Therefore, the marginal effect of a greater abatement spending ratio on the growth rate implicitly defined by the modified Ramsey rule in (38) is positive, but softer than the corresponding effect in the case of constant discounting. This is shown by applying the implicit function theorem to the expressions in (38), taking into account Lemma 1 and the standard Ramsey curves in (39) under exponential discounting:

$$0 < \frac{d\tilde{g}}{ds_{\text{A}}} = \frac{1}{1 + \eta \frac{\bar{J}}{\Omega(g,\eta)}} \tilde{r}'(s_{\text{A}}) < \frac{d\tilde{g}_{\hat{\rho}}}{ds_{\text{A}}} = \frac{1}{1 + \eta} \tilde{r}'(s_{\text{A}}). \tag{40}$$

The upward sloping curves $\tilde{g}$ and $\tilde{g}_{\hat{\rho}}$ are depicted as the dotted red and blue lines in Figures 3 and 4. Nonetheless, the abatement ratio in the market economy is known and given by $\tilde{s}_{\text{A}}$ in (32). Therefore, the implicit equation which defines the growth rate of the economy is no longer a function of s_A, and it is fully characterized by substituting s_A by $\tilde{s}_A$ in (38) and (39).

Note too that in a pure market economy, the abatement ratio, $\tilde{s}_A$, in (32) particularizes to $\tilde{\alpha}\gamma$. Provided that the firm does not internalize the whole effect of its own pollution on production, $\tilde{\alpha} < \alpha$, and that the abatement ratio in the centralized scenario cannot surpass $\alpha\gamma$, it follows that $\tilde{s}_A < s_A$. In contrast, the growth rate of the market economy can be above or below the growth rate in the centralized economy.

4 Comparing social welfare in the centralized and the market economies

Under non-constant discounting, a sophisticated central planner can be interpreted, as already mentioned, as a representative consumer who internalizes the external effects associated with pollution. He takes consumption and abatement decisions to maximize discounted welfare and at the same time, due to the lack of commitment, he must also induce his future selves to behave in his best interest. As a result of this game with his future selves, the balanced path equilibrium differs from the market economy equilibrium. However, as this section will show, contrary to the case of exponential discounting, when individuals present a decreasing degree of impatience the central planner's solution does not necessarily lead to higher social welfare. This result is in line with Krusell et al. (2002), although the underlying mechanism is different. In the Neoclassical growth model proposed in Krusell et al. (2002), the central planner acknowledges that he affects the returns to savings, while the price-taker individuals in a decentralized setting perceive these returns as constant. Thus, a central planner approaching the steady-state equilibrium from below has a decreasing marginal propensity to save, while it is constant for the competitive agent. In consequence, the latter saves more and the competitive economy grows faster, allowing for a higher welfare. In contrast, this section presents a similar result, but for an endogenous growth model and based on different grounds. In particular, two conditions are needed: a strong externality of pollution on utility, which induces the central planner to slow down growth, and a strong decay in the instantaneous discount rate, that is agents who strongly value the long run.

When comparing the growth rates in the market and the centralized economies, we can also distinguish the two possible situations commented on by Smulders and Gradus (1996). If the representative firm internalizes a small part of the effect of its emissions on production, (i.e. $\tilde{\alpha}$ and hence $\tilde{s}_A = \tilde{\alpha}\gamma$ are small), then the central planner's solution characterized by $s_A^* > \tilde{s}_A$ will almost certainly lead to a faster growth, as displayed by equilibria E and $\tilde{E}$ in Figure 3. In contrast, if the firm internalizes a large part of its externality on production and, at the same time, pollution has a strong impact on utility, (i.e. large ϕ), then the market equilibrium would be characterized by high abatement efforts and fast growth, while the centralized solution, in which the strong externality on utility is internalized, would be associated with a slower growth. This situation is shown in Figure 4.

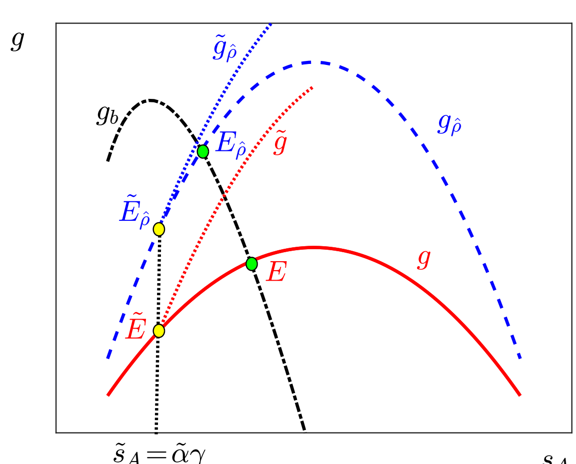

Figure 3: $\alpha - \tilde{\alpha}$ **large w.r.t.** ϕ

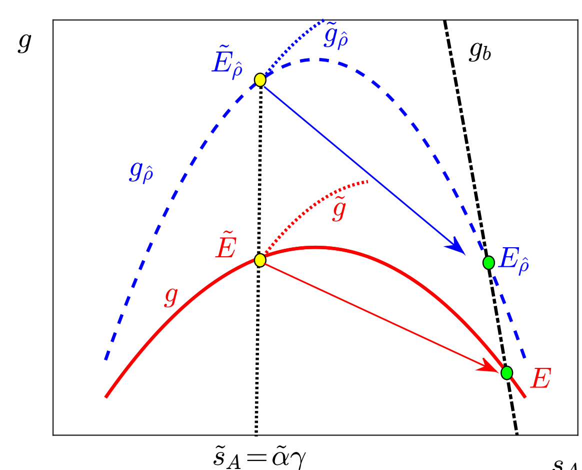

Figure 4: ϕ **large w.r.t.** $\alpha - \tilde{\alpha}$

The social welfare can only be greater in the decentralized economy than in the central planner's solution when growth is faster in the market economy, i.e. when the externality of pollution on utility is relatively strong (as represented in Figure 4).

If we restrict our analysis to the particular case with log-utility ($\sigma = 1$), the growth rate of the economy is independent of whether individuals discount the future at a constant rate, $\hat{\rho}$, or at a decreasing rate (as long as the same overall impatience is assumed). These rates in the centralized and the market economy read:

$$g_{\hat{\rho}}^{1*} = g^{1*} = r(s_{\text{A}}^{1*}) - \hat{\rho} = \left(\beta - s_{\text{A}}^{1*}\frac{\lambda}{\gamma}\right) y(s_{\text{A}}^{1*}) - \hat{\rho} \quad \text{with } s_{\text{A}}^{1^*} = s_{\text{A}\hat{\rho}}^{1*} > \tilde{\alpha}\gamma, \tag{41}$$

$$\tilde{g}_{\hat{\rho}}^{1*} = \tilde{g}^{1*} = r(\tilde{s}_{\text{A}}) - \hat{\rho} = (\beta - \tilde{\alpha}\lambda) y(\tilde{s}_{\text{A}}) - \hat{\rho} \quad \text{with } \tilde{s}_{\text{A}} = \tilde{s}_{\text{A}\hat{\rho}} = \tilde{\alpha}\gamma. \tag{42}$$

In these circumstances, the following proposition proves that although the growth rate is identical, individuals with non-constant discounting experience higher welfare than standard agents under exponential discounting. This result immediately follows if the growth rate is positive, given that the discounting method values the long run more strongly than exponential discounting, as clearly shown in Figure 1.

Proposition 7 *Assuming log-utility ($\sigma = 1$) and g^{1*}, $\tilde{g}^{1*} > 0$, the consumer welfare is higher in the case of non-constant discounting than under the standard exponential discounting, both for the centrally planned and for the market economy. Moreover,*[12]

$$W^1(K_t) = W_{\hat{\rho}}^1(K_t) + \frac{g^{1*}\left[1 + \phi(\gamma - \lambda\right]}{\hat{\rho}}\left(\bar{J}_0 - \frac{1}{\hat{\rho}}\right) > W_{\hat{\rho}}^1(K_t), \tag{43}$$

$$\widetilde{W}^1(K_t) = \widetilde{W}_{\hat{\rho}}^1(K_t) + \frac{\tilde{g}^{1*}\left[1 + \phi(\gamma - \lambda)\right]}{\hat{\rho}}\left(\bar{J}_0 - \frac{1}{\hat{\rho}}\right) > \widetilde{W}_{\hat{\rho}}^1(K_t), \tag{44}$$

with $\bar{J}_0 = \hat{\rho}\int_0^\infty j\theta(j)dj > 1/\hat{\rho}$.

Proof. See the Appendix □

Under log-utility, despite an identical growth rate as well as the same overall impatience, agents who discount the future at a decreasing rate assign a higher weight to the long run. In consequence, with positive growth they attain higher welfare. The gap $\bar{J}_0 - 1/\hat{\rho}$ in (43) and (44) measures how different from exponential discounting consumers' time preferences are. Under constant discounting $\bar{J}_0 - 1/\hat{\rho} = 0$, and the gap widens with the speed of decay of the instantaneous discount rates with the time distance from the present.

If economic growth in the central planner's solution is faster than or equally rapid as in the market economy ($g^{1*} \geq \tilde{g}^{1*}$ as in Figure 3) then, since the central planner's solution is Pareto optimal under constant discounting, Equations (43) and (44) imply that the central planner's solution is also Pareto optimal under non-constant discounting:

$$W^1(K_t) - \widetilde{W}^1(K_t) \geq W_{\hat{\rho}}^1(K_t) - \widetilde{W}_{\hat{\rho}}^1(K_t) \geq 0.$$

However, if the market economy grows faster ($\tilde{g}^{1*} > g^{1*}$ as in Figure 4), next proposition shows that, contrary to the case of exponential discounting, if individuals discount the future at a non-constant rate they could, under certain condition, derive higher welfare in the market equilibrium than in the central-planner solution.

Proposition 8 *Assuming log-utility ($\sigma = 1$), if the growth rate in the market economy is higher than the growth rate in the central-planner solution, $\tilde{g}^{1*} > g^{1*}$, then $\widetilde{W}^1(K_t) > W^1(K_t)$ if and only if*

$$\bar{J}_0 > \frac{1}{\hat{\rho}} - \hat{\rho}\frac{W_{\hat{\rho}}^1(K_t) - \widetilde{W}_{\hat{\rho}}^1(K_t)}{g^{1*} - \tilde{g}^{1*}}. \tag{45}$$

Proof. The proof is straightforward from the expressions of $\widetilde{W}^1(K_t)$ and $W^1(K_t)$ in (43) and (44). □

[12] Because all consumers are identical, $k_{nt} = k_t = \int_0^1 k_t dn = K_t$. Thus, in short notation, $\widetilde{W}^1(k_{nt}, K_t)$ and $\widetilde{W}_{\hat{\rho}}^1(k_{nt}, K_t)$ are simply denoted as $\widetilde{W}^1(K_t)$ and $\widetilde{W}_{\hat{\rho}}^1(K_t)$.

Under log-utility, $g^{1*} = g^{1*}_{\hat\rho}$ and $\tilde g^{1*} = \tilde g^{1*}_{\hat\rho}$ and, in consequence, the RHS in condition (45) is invariant to considering constant or non-constant discounting.[13] In contrast, the LHS of this inequation, $\bar J_0$, which can be interpreted as the mean of the probability distribution $\omega_0(j) = \theta(j)/\int_0^\infty \theta(i)di = \hat\rho\theta(j)$, can take a value as large as is needed by considering a sufficiently fast decay in the instantaneous discount rate is. Thus, there always exists a sufficiently fast speed of decay in the instantaneous discount rate above which condition (45) is satisfied.

The balanced path equilibrium associated with a decentralized economy can provide a higher social welfare than the central planner's solution under two conditions. It is first necessary that competitive agents save more and grow faster than the central planner. As already pointed out by Smulders and Gradus (1996) this is not based on the way individuals discount the future, but on the relative size of the pollution externalities. If the pollution externality on utility is relatively large, with respect to the pollution externality on production, the central planner would prefer a slower growth rate along the balanced path. Secondly, if the individuals' instantaneous discount rate decreases sharply with the time distance from the present, they value the future more strongly, and hence, can attain a higher social welfare in a decentralized economy which grows indefinitely at a faster rate.

Example 1 *To illustrate the previous result we assume a hybrid exponential-hyperbolic discounting as in Tsoukis et al. (2017), that is,*

$$\theta(j) = (1+\delta j)^{-\varphi/\delta} e^{-\rho j}, \tag{46}$$

where $\rho > 0$, $0 < \varphi < 1$, $\delta > 0$ and $\varphi/\delta < 1$. The instantaneous discount rate $\rho(j) = \rho + \varphi/(1+\delta j)$ decreases with j. Note that if φ tends to 0, then the discount function converges to exponential discounting. Applying the properties of the Laplace transformation to (46) as in Dyke (2014), the value of $\bar J_0$ follows:

$$\bar J_0 = \frac{1-\varphi/\delta}{\rho}.$$

For this discount function Tables 1 and 2 report a constellation of parameters values for which $\tilde g^{1*} > g^{1*}$. For these parameters values Figures 5 and 6 display the regions within the plane ρ-φ/δ at which condition (45) is satisfied, which guarantees a higher social welfare under the decentralized solution.

Table 1 and Figure 5 for a fixed ϕ, analyze different values of $\tilde\alpha$, which represents to what degree the firms internalize the effect of their on pollution on production. The shaded regions in Figure 5 represent the pairs $(\rho, \varphi/\delta)$ for which, agents with non-constant discounting attain higher welfare in the market than in the centralized economy. An identical analysis is made in Table 2 and Figure 6 for a fixed $\tilde\alpha$, and different utility elasticity of pollution, ϕ.[14]

Table 1: $\hat\rho = .05$, $\beta = .8$, $\lambda = .3$, $\gamma = .5$, $\alpha = 1$

ϕ	$\tilde\alpha$	$W^1_{\hat\rho} - \widetilde W^1_{\hat\rho}$	$g^{1*} - \tilde g^{1*}$
0.5	0.95	55.2353	-0.0134
0.5	0.9	56.8798	-0.0109
0.5	0.85	58.6904	-0.0081
0.5	0.8	60.6709	-0.0050

Table 2: $\hat\rho = .05$, $\beta = .8$, $\lambda = .3$, $\gamma = .5$, $\alpha = 1$

ϕ	$\tilde\alpha$	$W^1_{\hat\rho} - \widetilde W^1_{\hat\rho}$	$g^* - \tilde g^*$
0.44	0.95	55.1664	-0.0028
0.45	0.95	55.2943	-0.0042
0.5	0.95	55.2353	-0.0134
0.52	0.95	54.8129	-0.0182

[13]Indeed, the gap $W^1_{\hat\rho}(K_t) - \widetilde W^1_{\hat\rho}(K_t)$ is not even dependent on the actual stock of capital, K_t, as shown in the proof of Proposition 7 in the Appendix.

[14]As proven by Tsoukis et al. (2017), for any pair ρ and φ/δ in the shaded regions of Figures 5 or 6, the value of δ will be determined by the assumption of identical overall level of impatience:

$$\frac{1}{\hat\rho} = \frac{\Gamma(1-\varphi/\delta)}{\delta^{\varphi/\delta}\rho^{1-\varphi/\delta}}.$$

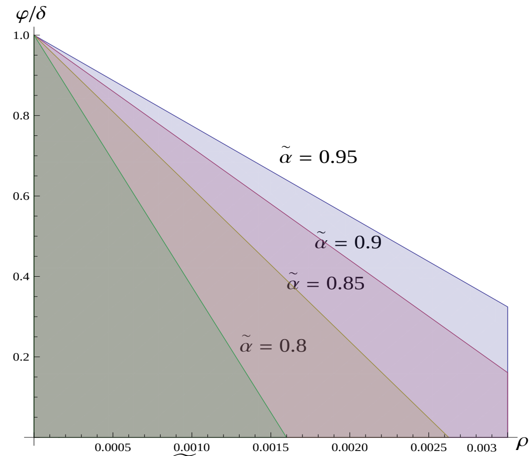

Figure 5: $\widetilde{W}^1(K_t) > W^1(K_t)$ **for different** $\tilde{\alpha}$

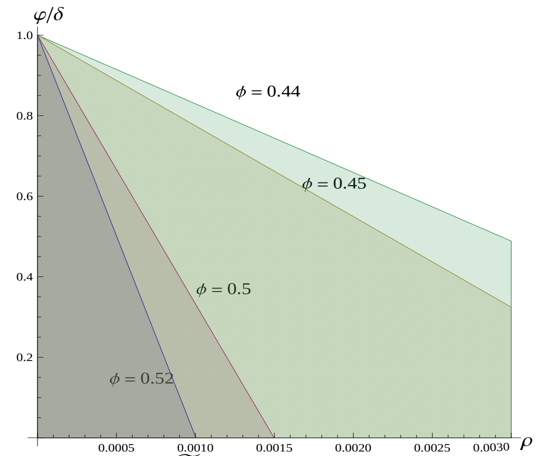

Figure 6: $\widetilde{W}^1(K_t) > W^1(K_t)$ **for different** ϕ

The counterintuitive result that the competitive economy can provide a higher social welfare than the central planer solution is obtained for the particular case of a log-utility. The hypothesis of an intertemporal elasticity of substitution equal to one is made for simplicity. For the more general case of $\sigma < 1$, the analysis is more cumbersome. A condition similar to (45) can be found. Whether this condition can or cannot be satisfied remains to be solved both analytically or numerically. At most it can be conjectured that the main result presented in Proposition 8 also holds for σ sufficiently close to one.

Under the conditions in Proposition 8 illustrated by Example 1, policy measures are inappropriate. However, when the market economy grows faster but condition (45) is not satisfied, then the social welfare in the market economy falls below the social welfare attained by a central planner. Likewise, the central planner's solution is also welfare enhancing when the production externality is large in relative terms to the externality on utility, so the growth rate in a decentralized economy remains lower than when a central planner internalizes the two externalities. In both situations, and regardless of whether the market economy grows too fast or too slow, it is pertinent to move the market equilibrium towards the planner's solution. As stated by Smulders and Gradus (1996), this gap can be closed by using either a single pollution tax, $\hat{T}_{\rm P}$, or a combination of an income tax, $T_{\rm Y}$, together with an abatement subsidy, $T_{\rm A}$. The immediate question that arises is whether these policies are more or less effective under non-constant than under constant discounting. As Figures 3 and 4 show, the gap between the centralized and the decentralized growth rates is not the same if individuals discount at a constant rate or their discount rate decreases with the time distance from the present. Thus, to analyze effectiveness, the next proposition computes and compares the elasticity of this gap for non-constant and for constant discounting to changes in either a tax or the subsidy.

Proposition 9 *The elasticity of the gap between the growth rates under a central planner and the market economy with no taxes (NT), with respect to increments in any of the taxes/subsidy is lower under non-constant discounting than in the standard case of exponential discounting, that is,*

$$\left| \frac{d(g^* - \tilde{g}^{NT*})}{dT} \frac{T}{g^* - \tilde{g}^{NT*}} \right| < \left| \frac{d(g^*_{\hat{\rho}} - \tilde{g}^{NT*}_{\hat{\rho}})}{dT} \frac{T}{g^*_{\hat{\rho}} - \tilde{g}^{NT*}_{\hat{\rho}}} \right|, \quad \text{with } T \in \{T_A, T_Y, \hat{T}_P\},$$

with $\tilde{g}^{NT*}$ and $\tilde{g}^{NT*}_{\hat{\rho}}$ the equilibrium growth rates of the market economy when $T_Y = \hat{T}_P = T_A = 0$, under non-constant and constant discounting respectively.

Proof. See the Appendix □

This result is valid for a log-utility, and any iso-elastic utility with $\sigma < 1$. According to this proposition the imposition of taxes/subsidy is less effective in narrowing the absolute gap in growth rates when individuals discount at a non-constant rate. More properly, this lower effectiveness refers to a marginal rise starting from an initial situation of no taxes/subsidy in the market economy. Nevertheless, this does not imply a lower effectiveness when raising the tax/subsidy from an already positive level.

5 Conclusions

This paper introduces non-constant discounting within the framework of the model of pollution abatement and endogenous growth presented by Smulders and Gradus (1996). In this model pollution is a by-product of production and can be abated. It is associated with two externalities, negatively affecting both the utility of consumers and the marginal productivity of capital. Within this context, the optimal consumption and abatement decision, when taken by dynamic individuals, will be dependent on whether their rate of time preference remains constant or decreases with the time distance from the present.

Under the standard assumption that exponential discounting decisions taken at a given point in time will continue to be optimal if recomputed at any later time. The time consistency of the solution implies that the central planner's solution coincides with the social optimum. Therefore, under the hypothesis of constant discounting and regardless of whether the market economy is characterized by under/over growth, bringing the market equilibrium towards the central-planner equilibrium is welfare enhancing. It may be, then, appropriate to impose taxes or subsidies to move the decentralized economy towards the social optimum. This is not necessarily the case if individuals discount the future at a decreasing rate.

When dealing with non-constant discounting, we have made two important assumptions. First, we neglect the possibility of commitment, and consider that individuals are time-consistent sophisticated agents. They know that their future selves will not stick to current planned decisions, and to prevent this time inconsistency, they play a game against their future selves considering Markov perfect feedback strategies which will be accepted by all future cohorts. This sophisticated behavior is followed by competitive agents in the decentralized solution, and also by the central planner. Given the lack of commitment, the central planner can be interpreted as a representative consumer who internalizes the two externalities caused by pollution, and takes consumption and abatement decisions as the solution to the inter-generations game.

Second, the analysis is controlled to have an identical overall level of impatience. This assumption guarantees that the comparison of the results under constant and non-constant discounting depends only on the way individuals discount the future, and not on different degrees of impatience. From this assumption it immediately follows that the individuals with a decreasing rate of impatience value relatively softly the short run and relatively strongly the long run, compared to standard agents with constant discounting. With positive growth and an intertemporal elasticity of substitution lower than one, current increments in consumption provoke increments in current and ulterior utility, although the marginal effect weakens with the time distance from the present. In contrast, the marginal effect on ongoing utility is constant under exponential discounting. Therefore, individuals with non-constant discounting undervalue the effect of current savings on ongoing utility, and hence invest less. This leads to our first result: under the empirically relevant hypothesis of an intertemporal elasticity of substitution lower than one, the fact that the households discount the future at a decreasing rate leads to a balanced growth path characterized by a lower growth rate and higher abatement efforts. Although the equilibrium diverges from the standard balanced growth path under exponential discounting, the conditions for its existence are identical. In contrast, with log-utility the equilibrium remains unchanged regardless of whether we consider time-declining or constant discounting.

Our main result lies in the comparison of social welfares in the centralized and the decentralized economy under non-constant discounting. Having no commitment power, the planner's solution stemming from the game between subsequent generations does not necessarily reach the social optimum. More importantly, we prove that the social welfare in the market economy surpasses the central planner's solution when certain conditions are met. This is consistent with the symmetric result in Krusell et al. (2002) for a Neoclassical growth model, although the underlying reasoning diverges. A central planner who internalizes the two externalities caused by pollution might be inclined to reduce the growth rate of the economy. As already stated by Smulders and Gradus (1996) for the case with constant discounting, this can occur when the externality of pollution on utility is relatively strong with respect to the externality on production. A faster growth is welcomed by those individuals who discount the future at a non-constant rate, and value the short run softer and the long run stronger than the standard individuals with constant discounting. Thus, for log-utility, and if the market economy grows faster than the centralized economy, we prove that this faster growth leads to a higher social welfare if the individuals' degree of impatience decays rapidly with the time distance from the present, and hence, the value assigned by current individuals to the distant future is very high. The result is valid for a general discount function satisfying a continuous decay in the instantaneous discount rate. However, a numerical example is presented for a specific discount function to illustrate this result.

When the above conditions are not satisfied, and hence the decentralized market economy does not welfare dominate the central planner (because the market economy grows too slowly, or because the degree of impatience decreases too slowly), then appropriate taxes on production and pollution, and/or a subsidy on abatement activities which help to approach the market toward the central planner's equilibrium, will be welfare enhancing. In that situation, we prove that starting from an initial situation of no intervention, the effectiveness of introducing these policies is lower when individuals discount at a decreasing rate than under constant discounting.

A Appendix

Proof of Proposition 1. Isoelastic utility with $\sigma < 1$

Along a balanced path consumption, abatement and output are proportional to the capital stock. Thus, given the utility function and the dynamics of the capital stock we conjecture a value function of the form:

$$W(K_t) = W_0 + W_1 K_t^{-\eta}, \tag{47}$$

with W_0, W_1 and η unknowns to be determined. From (12) and (47), the first-order condition (10) reads

$$\left\{ C_t(\tau) \left[K_t(\tau)^{\lambda} \left(\frac{s_{\text{A}}\phi\gamma}{s_{\text{A}} - \alpha\gamma} C_t(\tau) \right)^{-\gamma} \right]^{-\phi} \right\}^{1-\frac{1}{\sigma}} = -C_t(\tau)\eta W_1 K_t(\tau)^{-\eta-1}. \tag{48}$$

For consumption to be proportional to capital, the value of η must be as defined in (14). Constant W_0 can be derived from the definition of the value function, conjecture (47) and the overall impatience in (8):

$$W_0 = \frac{\sigma}{(1-\sigma)\hat{\rho}}. \tag{49}$$

Likewise, constant W_1 can be obtained from (48) and (12):

$$W_1 = -\frac{\left(\frac{s_{\text{A}}-\alpha\gamma}{\phi\gamma}\right)^{-\frac{1}{\sigma}} s_{\text{A}}^{-\frac{1-\sigma}{\sigma}\phi\gamma} y(s_{\text{A}})^{-\frac{1}{\sigma}[1+(1-\sigma)\phi\gamma]}}{\eta}. \tag{50}$$

From the optimality conditions and the conjectured value function, the RHS of the Bellman equation in (9) along the balanced path can be rewritten as

$$\frac{\sigma}{1-\sigma} - \eta W_1 \left(g - \frac{\sigma s_{\text{C}}}{1-\sigma} y(s_{\text{A}}) \right) K_t^{-\eta}. \tag{51}$$

On the other hand, from the definition of $\rho(j)$ it is immediately obvious that $\int_t^\infty \rho(\tau-t)\theta(\tau-t)ds = 1$, and, therefore, the LHS of the Bellman equation can be written as

$$\frac{\sigma}{1-\sigma} + \frac{\sigma s_{\text{C}}\eta W_1}{1-\sigma} y(s_{\text{A}}) \int_t^\infty K_t(\tau)^{-\eta}\theta(\tau-t)\rho(\tau-t)\, d\tau. \tag{52}$$

Along a balanced growth path, the capital stock at the running time τ, for the t-agent is $K_t(\tau) = K_t e^{g(\tau-t)}$. Therefore, equating expressions (51) and (52) it follows:

$$-\frac{\sigma s_{\text{C}}}{1-\sigma} y(s_{\text{A}}) \int_t^\infty e^{-\eta g(\tau-t)}\theta(\tau-t)\rho(\tau-t)\, d\tau = g - \frac{\sigma s_{\text{C}}}{1-\sigma} y(s_{\text{A}}). \tag{53}$$

The assumption $\sigma < 1$ implies $\eta > 0$ and, hence, it guarantees the convergence of this integral when the growth rate of the economy is a positive constant. Integrating by parts one gets

$$\int_t^\infty e^{-\eta g(\tau-t)}\theta(\tau-t)\rho(\tau-t)\,d\tau = 1-\eta g\int_t^\infty e^{-g\eta(\tau-t)}\theta(\tau-t)\,d\tau. \tag{54}$$

And from (53) and (54) it follows:

$$\frac{\int_t^\infty e^{-g\eta(\tau-t)}\theta(\tau-t)\rho(\tau-t)\,d\tau}{\int_t^\infty e^{-g\eta(\tau-t)}\theta(\tau-t)\,d\tau} = \left(\frac{\sigma s_{\rm C}}{1-\sigma}y(s_{\rm A})-g\right)\eta.$$

The ratio of integrals in the LHS defines the effective rate of time preference, $\Delta(g,\eta)$, in (15). Then,

$$\Delta(g,\eta) = \left(\frac{\sigma s_{\rm C}}{1-\sigma}y(s_{\rm A})-g\right)\eta. \tag{55}$$

By adding and subtracting $g=(1-s_{\rm C}-s_{\rm A})y(s_{\rm A})$ in the RHS, and taking into account the constant returns to scale hypothesis in (3) together with the consumption-abatement ratio in (12) and the expression for η in (14), Equation (55) can be rewritten as

$$\Delta(g,\eta) = \left[\beta-\frac{\lambda}{\gamma}s_{\rm A}\right]y(s_{\rm A})-(1+\eta)g,$$

which leads to the Ramsey equation[15] in (13).

Logarithmic utility: $\sigma = 1$

The value function along a balanced growth path can be expressed as

$$\begin{aligned} W^1(K_t) = \int_t^\infty \ln\left[C_tP_t^{-\phi}e^{g^1[1+\phi(\gamma-\lambda)](\tau-t)}\right]\theta(\tau-t)d\tau = \\ \ln\left[C_tP_t^{-\phi}\right]\int_t^\infty\theta(\tau-t)d\tau + g^1\left[1+\phi(\gamma-\lambda)\right]\int_t^\infty(\tau-t)\theta(\tau-t)d\tau = \\ \frac{[1+\phi(\gamma-\lambda)]}{\hat\rho}\ln K_t + \frac{1}{\hat\rho}\ln\left[s_{\rm C}s_{\rm A}^{\gamma\phi}y(s_{\rm A})^{1+\phi\gamma}\right] + \frac{g^1\bar J_0\left[1+\phi(\gamma-\lambda)\right]}{\hat\rho}, \end{aligned} \tag{56}$$

with $\bar J_0 = \hat\rho\int_0^\infty j\theta(j)$. Then, the value function is defined as

$$W^1(K_t) = W_0^1 + W_1^1\ln K_t, \tag{57}$$

with

$$W_0^1 = \frac{g^1\left[1+\phi(\gamma-\lambda)\right]\bar J_0 + \ln s_{\rm C}s_{\rm A}^{\gamma\phi}y(s_{\rm A})^{1+\phi\gamma}}{\hat\rho}, \qquad W_1^1 = \frac{1+\phi(\gamma-\lambda)}{\hat\rho}. \tag{58}$$

On the other hand, the first-order condition (10) reads

$$\frac{1}{C_t} = \frac{W_1^1}{K_t}. \tag{59}$$

Then, from (7), (12), (58) and (59), the growth rate of the economy can be written as

$$g^1 = (1-s_{\rm A})y(s_{\rm A}) - \frac{\hat\rho}{1+\phi(\gamma-\lambda)} = \left(\beta - s_{\rm A}\frac{\lambda}{\gamma}\right)y(s_{\rm A}) - \hat\rho.$$

□

[15] A necessary condition for a positive growth rate is $\beta-\lambda s_{\rm A}/\gamma > 0$. Since we are assuming $\alpha\gamma < 1$, this necessary condition holds for any $s_{\rm A}\in(\alpha\gamma,1]$.

Proposition 10 *Under condition $\sigma < 1$ and $\hat{\rho} < r\left(\arg\min_{s_A} r(s_A)\right)$, the modified Ramsey rule (13) implicitly defines a positive growth rate, g, as a function of the abatement spending ratio, s_A.*

Proof. Integrating by parts in (15), one gets expression (26), $\Delta(g,\eta) = 1/\Omega(g,\eta) - \eta g$. Replacing this expression into Equation (13) and rearranging terms we get

$$g + \frac{1}{\Omega(g,\eta)} = r(s_{\rm A}). \tag{60}$$

Since, from (24) $d\Omega(g,\eta)/dg < 0$, the LHS in (60) can be regarded as a monotonously increasing function of g. Moreover, $\lim_{g\to 0}\Omega(g,\eta) = 1/\hat{\rho}$ and $\lim_{g\to+\infty}\Omega(g,\eta) = 0$. Then, the LHS runs from $\hat{\rho}$ towards ∞ as g goes from zero to ∞. Moreover, for any $s_{\rm A} \in (\alpha\gamma, 1]$, $r(s_{\rm A}) > 0$, and hence, under condition $\hat{\rho} < r\left(\arg\min_{s_{\rm A}} r(s_{\rm A})\right)$, there always exists a unique $g > 0$ satisfying Equation (60). □

Lemma 2 $Cov(j,\rho(j)) = \dfrac{\Omega(g,\eta) - \bar{J}}{\Omega(g,\eta)} < 0$ *and therefore, $\Omega(g,\eta) < \bar{J}$.*

Proof.

$$\begin{aligned} Cov(j,\rho(j)) &= \int_0^\infty \rho(j)\omega(j)[j-\bar{J}]\,dj = \int_0^{\bar{J}} \rho(j)\omega(j)[j-\bar{J}]\,dj + \int_{\bar{J}}^\infty \rho(j)\omega(j)[j-\bar{J}]\,dj \\ &< \rho(\bar{J})\int_0^{\bar{J}} \omega(j)[j-\bar{J}]\,dj + \rho(\bar{J})\int_{\bar{J}}^\infty \omega(j)[j-\bar{J}]\,dj = \rho(\bar{J})\int_0^\infty \omega(j)[j-\bar{J}]\,dj = 0. \end{aligned}$$

Moreover, since $\rho(j) = -\dot{\theta}(j)/\theta(j)$, then

$$Cov(j,\rho(j)) = \frac{-1}{\Omega(g,\eta)}\int_0^\infty \dot{\theta}(j)(j-\bar{J})e^{-\eta g j}dj = \frac{-1}{\Omega(g,\eta)}\left\{\int_0^\infty j\dot{\theta}(j)e^{-\eta g j}dj - \bar{J}\int_0^\infty \dot{\theta}(j)e^{-\eta g j}dj\right\}.$$

Integrating by parts, after some calculus it follows that

$$0 > Cov(j,\rho(j)) = \frac{\Omega(g,\eta) - \bar{J}}{\Omega(g,\eta)}.$$

□

Proof of Proposition 3. Writing the modified Ramsey equation as

$$g(s_{\rm A}) = r(s_{\rm A}) - \frac{1}{\Omega(g,\eta)},$$

and differentiating implicitly with respect to $s_{\rm A}$, taking into account (17) and (24) one gets

$$g'(s_{\rm A}) = \frac{r'(s_{\rm A})}{1 + \eta\frac{\bar{J}}{\Omega(g,\eta)}}, \quad g'_{\hat{\rho}}(s_{\rm A}) = \frac{r'(s_{\rm A})}{1+\eta}. \tag{61}$$

Because $\eta > 0$, and from Lemma 2 $\Omega(g,\eta) < \bar{J}$, then the effect of $s_{\rm A}$ on the growth rate in the case of non-constant discounting is softer than under exponential discounting:

$$|g'(s_{\rm A})| < |g'_{\hat{\rho}}(s_{\rm A})| \quad \text{for all} \quad s_{\rm A} \geq 0.$$

From (18) the marginal effect of $s_{\rm A}$ on the social rate of return reads

$$r'(s_{\rm A}) = \frac{\alpha\gamma\beta - s_{\rm A}\dfrac{\lambda}{\gamma}}{(1-\alpha\gamma)s_{\rm A}}\, y(s_{\rm A}).$$

Therefore, for $s_{\rm A} < \hat{s}_{\rm A} = \alpha\beta\gamma^2/\lambda$, $r'(s_{\rm A}) > 0$; and for $s_{\rm A} > \hat{s}_{\rm A}$, $r'(s_{\rm A}) < 0$. Then, from (61) the abatement spending ratio $\hat{s}_{\rm A}$ maximizes the modified Ramsey rule both under constant and non-constant discounting.

Taking into account that $\Omega(g,\eta) < \Omega_{\hat{\rho}}(g,\eta)$, then,

$$g_{\hat{\rho}}(s_{\rm A}) = \frac{r(s_{\rm A}) - \hat{\rho}}{1+\eta} = r(s_{\rm A}) - \frac{1}{\Omega_{\hat{\rho}}(g,\eta)} > r(s_{\rm A}) - \frac{1}{\Omega(g,\eta)} = g(s_{\rm A}),$$

which proves that the modified Ramsey curve under non-constant discounting lies below the standard Ramsey curve under exponential discounting. □

Proof of Proposition 4. Equation (30) is independent of time preferences and, in consequence identical under both discounting behaviors. In contrast, the modified Ramsey rule with non-constant discounting given in (13) differs from the standard Ramsey rule in the case of constant discounting in (17). Taking $s_{\rm A}$ as given, the gap between the two curves can be written as

$$g_{\hat{\rho}} - g = \frac{\Delta(g,\eta) - \hat{\rho}}{1+\eta}.$$

It is known that $\Delta(g,\eta) > \hat{\rho}$ for any $g > 0$ and $\eta > 0$. Hence, $g_{\hat{\rho}} > g$, which proves that the modified Ramsey curve with non-constant discounting is located below the standard Ramsey curve under exponential discounting. Furthermore, since the curve defined in (30) is downward sloping, in a centralized economy the assumption of non-constant discounting implies a lower growth rate and a higher abatement spending ratio. □

Proof of Proposition 5. From Proposition 3 the standard Ramsey curve in the case of constant discounting and the modified Ramsey rule in the case of non-constant discounting, can be re-written as

$$D_{\hat{\rho}}(g,s_{\rm A}) \doteqdot \frac{r(s_{\rm A}) - \hat{\rho}}{1+\eta} - g = 0, \qquad D(g,s_{\rm A}) \doteqdot \frac{r(s_{\rm A}) - \Delta(g,\eta)}{1+\eta} - g = 0 \tag{62}$$

or simply, $D_{\hat{\rho}}$ and D. Likewise, the feasible set of equilibria can be expressed as

$$F(g,s_{\rm A}) \doteqdot \left(1 + \frac{\alpha}{\phi} - \frac{1+\phi\gamma}{\phi\gamma}s_{\rm A}\right) y(s_{\rm A}) - g = 0, \tag{63}$$

or simply, F.

From the implicit function theorem, the derivatives of g^* and $g^*_{\hat{\rho}}$ with respect to ϕ can be written as

$$\frac{dg^*}{d\phi} = \frac{Num(s^*_{\rm A})}{Den(s^*_{\rm A})}, \qquad \frac{dg^*_{\hat{\rho}}}{d\phi} = \frac{Num_{\hat{\rho}}(s^*_{{\rm A}\hat{\rho}})}{Den_{\hat{\rho}}(s^*_{{\rm A}\hat{\rho}})},$$

$$Num(s_{\rm A}) = \begin{vmatrix} -\frac{\partial D}{\partial \phi} & \frac{\partial D}{\partial s_{\rm A}} \\ -\frac{\partial F}{\partial \phi} & \frac{\partial F}{\partial s_{\rm A}} \end{vmatrix}, \quad Den(s_{\rm A}) = \begin{vmatrix} \frac{\partial D}{\partial g} & \frac{\partial D}{\partial s_{\rm A}} \\ \frac{\partial F}{\partial g} & \frac{\partial F}{\partial s_{\rm A}} \end{vmatrix}$$

$$Num_{\hat{\rho}}(s_{\rm A}) = \begin{vmatrix} -\frac{\partial D_{\hat{\rho}}}{\partial \phi} & \frac{\partial D_{\hat{\rho}}}{\partial s_{\rm A}} \\ -\frac{\partial F}{\partial \phi} & \frac{\partial F}{\partial s_{\rm A}} \end{vmatrix}, \quad Den_{\hat{\rho}}(s_{\rm A}) = \begin{vmatrix} \frac{\partial D_{\hat{\rho}}}{\partial g} & \frac{\partial D_{\hat{\rho}}}{\partial s_{\rm A}} \\ \frac{\partial F}{\partial g} & \frac{\partial F}{\partial s_{\rm A}} \end{vmatrix}$$

which can be obtained from (62)–(63) and the expressions for $r(s_{\rm A})$ and $\Delta(g,\eta)$ in (18) and (15). Under constant discounting it is easy to compute

$$Den_{\hat{\rho}}(s_{\rm A}) = \frac{\eta[s_{\rm A} - \alpha^2\gamma^2 + \gamma(s_{\rm A} - \alpha\gamma)(1-\sigma)\phi]y(s_{\rm A})}{s_{\rm A}(1-\sigma)(1-\alpha\gamma)\phi\gamma(1+\eta)} > 0. \tag{64}$$

The positive sign of $Den_{\hat{\rho}}(s_{\rm A})$ can be guaranteed for $s_{\rm A} \geq \alpha\gamma$ as long as $\alpha\gamma < 1$. In contrast, the expression for $Num_{\hat{\rho}}(s_{\rm A})$ is rather cumbersome and of unclear sign.

When ϕ tends to zero, the downward sloping curve $F(g, s_{\text{A}})$ converges towards a vertical line at $s^*_{\text{A}\hat{\rho}} = \alpha\gamma$. Then, after some simplifications, it follows that

$$\lim_{\phi\to 0} \frac{dg^*_{\hat{\rho}}}{d\phi} = \frac{\sigma\lambda(\hat{s}_{\text{A}} - s^*_{\text{A}\hat{\rho}})y(s^*_{\text{A}\hat{\rho}})}{\gamma(s^*_{\text{A}\hat{\rho}} - \alpha^2\gamma^2)} \lim_{\phi\to 0} \frac{s^*_{\text{A}\hat{\rho}} - \alpha\gamma}{\phi}.$$

Because $\alpha\gamma < 1$, from (3) and (4) it follows that $\hat{s}_{\text{A}} > \alpha\gamma$ and in consequence, the first factor is positive. Moreover, from (12), the second factor can be written as $\lim_{\phi\to 0} \gamma s^*_{\text{C}\hat{\rho}}$. To prove that this term is also positive, note that the growth rate of the economy in (17) at $s^*_{\text{A}\hat{\rho}} = \alpha\gamma$ reads

$$g^*_{\hat{\rho}} = \sigma[(1-\alpha\gamma)y(\alpha\gamma) - \hat{\rho}].$$

And equivalently, for this value, $s^*_{\text{A}\hat{\rho}}$, the growth rate can also be written as

$$g^*_{\hat{\rho}} = (1 - \alpha\gamma - \lim_{\phi\to 0} s^*_{\text{C}\hat{\rho}})y(\alpha\gamma).$$

And from these two expressions, the limit follows:

$$\gamma \lim_{\phi\to 0} s^*_{\text{C}\hat{\rho}} = \gamma(1-\alpha\gamma)(1-\sigma) + \gamma\frac{\sigma\hat{\rho}}{y(\alpha\gamma)} > 0.$$

This concludes the proof that $dg^*_{\hat{\rho}}/d\phi > 0$ when ϕ tends to zero.

To characterize the condition under which $\lim_{\phi\to 0} dg^*/d\phi > 0$, we express $Num(s_{\text{A}})$ and $Den(s_{\text{A}})$ as functions of $Num_{\hat{\rho}}(s_{\text{A}})$ and $Den_{\hat{\rho}}(s_{\text{A}})$:

$$\begin{aligned}
Num(s_{\text{A}}) &= Num_{\hat{\rho}}(s_{\text{A}}) + \Xi_N(s_{\text{A}}), \quad Den(s_{\text{A}}) = Den_{\hat{\rho}}(s_{\text{A}}) + \Xi_D(s_{\text{A}}),\\
\Xi_N(s_{\text{A}}) &= -\frac{1 + g\Omega(g,\eta) - \bar{J}g(1+\eta)}{(1+\eta)^2\Omega(g,\eta)}\frac{1-\sigma}{\sigma}(\gamma-\lambda)\frac{\partial F}{\partial\phi},\\
\Xi_D(s_{\text{A}}) &= \frac{\eta(\bar{J} - \Omega(g,\eta))}{(1+\eta)\Omega(g,\eta)} + \frac{s_{\text{A}} - \alpha^2\gamma^2 + \phi\gamma(s_{\text{A}} - \alpha\gamma)}{(1-\alpha\gamma)\phi\gamma s_{\text{A}}} y(s_{\text{A}}) > 0.
\end{aligned}$$

From (64) $lim_{\phi\to 0} Den_{\hat{\rho}}(s^*_{\text{A}\hat{\rho}}) = +\infty$ and since $\lim_{\phi\to 0} dg^*_{\hat{\rho}}/d\phi$, is positive and finite, then $lim_{\phi\to 0} Num_{\hat{\rho}}(s^*_{\text{A}\hat{\rho}}) = +\infty$ must be true. At the same time, since $\Omega(g,\eta)$ and $\bar{J}$ are finite, also $\Xi_N(s^*_{\text{A}\hat{\rho}})$ is finite. Furthermore, curve $F(g, s_{\text{A}})$ (equally valid in the case of non-constant discounting) turns into the vertical $s^*_{\text{A}} = \alpha\gamma$ as ϕ tends to zero. Thus in the limit when ϕ tends to zero, $s^*_{\text{A}} = s^*_{\text{A}\hat{\rho}} = \alpha\gamma$. From all this it follows that

$$\lim_{\phi\to 0} \frac{dg^*}{d\phi} = \lim_{\phi\to 0} \frac{Num_{\hat{\rho}}(s^*_{\text{A}}) + \Xi_N(s^*_{\text{A}})}{Den_{\hat{\rho}}(s^*_{\text{A}}) + \Xi_D(s^*_{\text{A}})} = \lim_{\phi\to 0} \frac{Num_{\hat{\rho}}(s^*_{\text{A}\hat{\rho}})}{Den_{\hat{\rho}}(s^*_{\text{A}\hat{\rho}}) + \Xi_D(s^*_{\text{A}\hat{\rho}})}.$$

Proceeding as in the constant discount case, this limit can be computed:

$$\lim_{\phi\to 0} \frac{dg^*}{d\phi} = \frac{\Omega(g,\eta)\sigma\lambda(\hat{s}_{\text{A}} - s^*_{\text{A}})y(s^*_{\text{A}})}{\bar{J}\gamma(s^*_{\text{A}} - \alpha^2\gamma^2)} \lim_{\phi\to 0} \frac{s^*_{\text{A}} - \alpha\gamma}{\phi}.$$

Therefore, and taking into account Lemma 2,

$$\lim_{\phi\to 0} \frac{dg^*}{d\phi} = \frac{\Omega(g,\eta)}{\bar{J}} \lim_{\phi\to 0} \frac{dg^*_{\hat{\rho}}}{d\phi} < \lim_{\phi\to 0} \frac{dg^*_{\hat{\rho}}}{d\phi}.$$

□

Proof of Proposition 6. Isoelastic utility with $\sigma < 1$

The Bellman equation for the optimization problem (34)–(35) of the n-th consumer's belonging to the t-cohort reads

$$\int_t^{\infty} U\left(c^*_{nt}(\tau), P_t(K_t(\tau))\right) \theta(\tau-t)\, \rho(\tau-t)\, d\tau = \max_{c_{nt}} \left\{ U(c_{nt}, P_t(K_t)) + \widetilde{W}'_{k_{nt}}(k_{nt}, K_t)[rk_{nt} - c_{nt}] + \widetilde{W}'_{K_t}(k_{nt}, K_t)\dot{K}_t \right\}, \quad (65)$$

where $c^*_{nt}(\tau)$ denotes the optimal consumption planned by the t-cohort of the n-th consumer for the elapsed time $\tau > t$. The pollution stock $P_t(K_t(\tau))$ depends on the abatement carried out by all firms[16] and the total capital stock accumulated by all the members of the t-cohort from t and up to the elapsed time τ, $K_t(\tau)$. For ease of notation, pollution for the t-cohort is written as a function of the elapsed time τ:

$$P_t(\tau) = P_t(K_t(\tau)) = [s_{\text{A}} y(s_{\text{A}})]^{-\gamma} K_t(\tau)^{-(\gamma-\lambda)}. \quad (66)$$

From (9) it follows that the value function of the n-th consumer, $\widetilde{W}(k_{nt}, K_t)$, depends on the capital stock he owns at time t, k_{nt}, whose dynamics is determined by this agent's consumption decisions, and also on the global capital stock at this time t, K_t. While this global capital stock affects the consumer's welfare through its effect on the pollution stock, the n-th consumer does not internalize the effect of his capital stock on the global stock, and hence on the ongoing pollution. Therefore his consumption decision does not affect the dynamics of the global capital stock, and the first-order optimality condition for the problem above reads

$$U'_{c_{nt}}(c_{nt}, P_t(K_t)) = \widetilde{W}'_{k_{nt}}(k_{nt}, K_t).$$

Equivalently, at any ulterior time τ, the maximizer of the RHS of Equation (65) reads

$$c^*_{nt}(\tau) = P_t(\tau)^{(1-\sigma)\phi} (\widetilde{W}'_{k_{nt}}(\tau))^{-\sigma},$$

where, for simplicity, $\widetilde{W}'_{k_{nt}}(\tau)$ denotes $\widetilde{W}'_{k_{nt}}(k_{nt}(\tau), K_t(\tau))$.

Plugging this optimal consumption into (2), the utility at the optimal trajectory reads

$$U(c^*_{nt}(\tau), P_t(\tau)) = \frac{\sigma}{1-\sigma}[1 - P_t(\tau)^{(1-\sigma)\phi} (\widetilde{W}'_{k_{nt}}(\tau))^{1-\sigma}].$$

Plugging these last two expressions into (65) and taking into account $\int_t^{\infty} \rho(\tau-t)\theta(\tau-t)ds = 1$, the Bellman equation can be rewritten as

$$-\int_t^{\infty} \frac{\sigma}{1-\sigma} P_t(\tau)^{(1-\sigma)\phi} (\widetilde{W}'_{k_{nt}}(\tau))^{1-\sigma} \theta(\tau-t)\, \rho(\tau-t)\, d\tau = -\frac{\sigma}{1-\sigma} P_t(t)^{(1-\sigma)\phi} (\widetilde{W}'_{k_{nt}}(t))^{1-\sigma} + \widetilde{W}'_{k_{nt}}(t) \left[rk_{nt} - P_t(t)^{(1-\sigma)\phi} (\widetilde{W}'_{k_{nt}}(t))^{-\sigma} \right] + \widetilde{W}'_{K_t}(t)\dot{K}_t. \quad (67)$$

Taking into account in (67) that k_{nt}, K_t and $c_{nt}(t)$ grow at the same constant rate along the balanced path, we conjecture a value function of the form

$$\widetilde{W}(k_{nt}(\tau), K_t(\tau)) = \widetilde{W}_0 + \widetilde{W}_1 k_{nt}^{\eta_1} K_t^{\eta_2}. \quad (68)$$

This conjecture is in the same spirit as Krusell et al. (2002), who consider a logarithmic utility in a discrete setting. Plugging this conjecture into (67), parameters η_1 and η_2 satisfy

$$\eta_1 = -\frac{1-\sigma}{\sigma}, \qquad \eta_2 = -\frac{1-\sigma}{\sigma}\phi(\gamma-\lambda), \qquad \eta_1 + \eta_2 = -\eta. \quad (69)$$

Provided that all consumers are identical, the n-th and the global capital stock grow at the same constant rate, $\tilde{g}$, along the balanced path. Then, the individual and the global capital stocks at the running time τ for the t-cohort can be written as $k_{nt}(\tau) = k_{nt} e^{\tilde{g}(\tau-t)}$ and $K_t(\tau) = K_t e^{\tilde{g}(\tau-t)}$, respectively. From the optimality conditions of firms in (32)–(33), abatement is proportional to output, at the ratio $\tilde{s}_{\text{A}}$, and the rate of return

[16] Note that $A_t(\tau)$ can be written as $s_{\text{A}} y(s_{\text{A}}) K_t(\tau)$. Since $\tilde{s}_{\text{A}}$ is constant from the optimality condition in (32) then, abatement and capital will be proportional in equilibrium.

is proportional to $y(s_{\text{A}})$. Along the balanced path output, capital and abatement grow at the same rate as capital and, hence, s_{A} and r remain constant. Then, the LHS of (67) can be written

$$\Theta K_t^{(1-\sigma)[\eta_2-\phi(\gamma-\lambda)]} k_{nt}^{\eta_1} \int_t^\infty e^{-\tilde{g}\eta(\tau-t)}\theta(\tau-t)\rho(\tau-t)d\tau, \tag{70}$$

where

$$\Theta = -\frac{\sigma}{1-\sigma}[s_{\text{A}} y(s_{\text{A}})]^{-(1-\sigma)\gamma\phi}(\eta_1 \widetilde{W}_1)^{1-\sigma}.$$

Expression (70) can be equivalently written as

$$\int_t^\infty \rho(\tau-t)\frac{\theta(\tau-t)e^{-\tilde{g}\eta(\tau-t)}}{\int_t^\infty \theta(\tau-t)e^{-\tilde{g}\eta(\tau-t)}d\tau}d\tau \int_t^\infty -\frac{\sigma}{1-\sigma}P_t(\tau)^{(1-\sigma)\phi}(\widetilde{W}'_{k_{nt}}(\tau))^{1-\sigma}\theta(\tau-t)\, d\tau,$$

and from (15) and the definition of the value function, the LHS reads

$$\Delta(\tilde{g},\eta)(\widetilde{W}(k_{nt},K_t)-\widetilde{W}_0) \equiv \Delta(\tilde{g},\eta)\widetilde{W}_1 k_{nt}^{\eta_1}K_t^{\eta_2}. \tag{71}$$

Similarly, along the balanced path the RHS of the Bellman equation can be rewritten as

$$-\frac{1}{1-\sigma}\left(\eta_1\widetilde{W}_1\right)^{1-\sigma}[s_{\text{A}} y(s_{\text{A}})]^{-(1-\sigma)\gamma\phi}k_{nt}^{\eta_1}K_t^{\eta_2} + \widetilde{W}_1 k_{nt}^{\eta_1}K_t^{\eta_2}\left\{\eta_1 r + \tilde{g}\eta_2\right\}. \tag{72}$$

Given expressions (71) and (72), we check whether the conjectured value function satisfies the Bellman equation, and compute constants $\widetilde{W}_0, \widetilde{W}_1$. Constant $\widetilde{W}_1$ is given by

$$\widetilde{W}_1 = -\left(\frac{\sigma(-\eta_1)^{1-\sigma}}{(1-\sigma)[\Delta(\tilde{g},\eta)+\tilde{g}\eta][s_{\text{A}} y(s_{\text{A}})]^{(1-\sigma)\gamma\phi}}\right)^{\frac{1}{\sigma}}. \tag{73}$$

Following the same reasoning as in the proof of Proposition 1, $\widetilde{W}_0 = W_0$, given in (49):

$$\widetilde{W}_0 = \frac{\sigma}{(1-\sigma)\hat{\rho}}.$$

Finally, to compute the modified Ramsey rule, we take derivatives with respect to k_{nt} in the RHS of (67) and in (71), and obtain

$$\Delta(\tilde{g},\eta)\widetilde{W}'_{k_{nt}}(k_{nt},K_t) = \widetilde{W}''_{k_{nt}k_{nt}}(k_{nt},K_t)\dot{k}_{nt} + \widetilde{W}'_{k_{nt}}(k_{nt},K_t)r + \widetilde{W}''_{K_t k_{nt}}(k_{nt},K_t)\dot{K}_t$$

Substituting in this expression $\dot{k}_{nt}$ and $\dot{K}_t$ by $\tilde{g}k_{nt}$ and $\tilde{g}K_t$, taking into account (68) and the last equation in (69), the modified Ramsey rule (36) follows.

Logarithmic utility: $\sigma = 1$

The value function along a balanced growth path can be expressed as

$$\begin{aligned}\widetilde{W}^1(k_{nt},K_t) &= \int_t^\infty \ln\left[c_{nt}P_t^{-\phi}e^{\tilde{g}^1[1+\phi(\gamma-\lambda)](\tau-t)}\right]\theta(\tau-t)d\tau = \\ &\ln\left[c_{nt}P_t^{-\phi}\right]\int_t^\infty \theta(\tau-t)d\tau + \tilde{g}^1\left[1+\phi(\gamma-\lambda)\right]\int_t^\infty(\tau-t)\theta(\tau-t)d\tau = \\ &\frac{1}{\hat{\rho}}\ln\left[s_{\text{C}}s_{\text{A}}^{\gamma\phi}y(s_{\text{A}})^{1+\phi\gamma}\right] + \frac{1}{\hat{\rho}}\ln k_{nt} + \frac{\phi(\gamma-\lambda)}{\hat{\rho}}\ln K_t + \frac{\tilde{g}^1\bar{J}_0\left[1+\phi(\gamma-\lambda)\right]}{\hat{\rho}}.\end{aligned}$$

Therefore, the value function is defined as

$$\widetilde{W}^1(k_{nt},K_t) = \widetilde{W}_0^1 + \tilde{\eta}_1\ln k_{nt} + \tilde{\eta}_2\ln K_t = \widetilde{W}_0^1 + \ln\left[k_{nt}^{\tilde{\eta}_1}K_t^{\tilde{\eta}_2}\right], \tag{74}$$

with

$$\widetilde{W}_0^1 = \frac{\tilde{g}^1\left[1+\phi(\gamma-\lambda)\right]\bar{J}_0 + \ln s_{\rm C} s_{\rm A}^{\gamma\phi} y(s_{\rm A})^{1+\phi\gamma}}{\hat{\rho}}, \quad \tilde{\eta}_1 = \frac{1}{\hat{\rho}}, \quad \tilde{\eta}_2 = \frac{\phi(\gamma-\lambda)}{\hat{\rho}}.$$

On the other hand, the LHS of the Bellman equation can be written as

$$\int_t^\infty \left[\ln c_{nt}(\tau) P_t^{-\phi}(\tau)\right] \theta(\tau-t)\,\rho(\tau-t)\,d\tau = \hat{\rho}\widetilde{W}^1(k_{nt},K_t) + \tilde{g}^1\left[1+\phi(\gamma-\lambda)\right]\left\{\frac{1}{\hat{\rho}} - \bar{J}_0\right\}.$$

Plugging this into (65), the Bellman equation can be rewritten as

$$\hat{\rho}\widetilde{W}^1(k_{nt},K_t) + \tilde{g}^1\left[1+\phi(\gamma-\lambda)\right]\left\{\frac{1}{\hat{\rho}} - \bar{J}_0\right\} = \ln c_{nt} - \phi\ln P_t + (\widetilde{W}^1_{k_{nt}})'(t)\left[rk_{nt}-c_{nt}\right] + (\widetilde{W}^1_{K_t})'(t)\dot{K}_t.$$

And differentiating with respect k_{nt}, one gets

$$\hat{\rho}(\widetilde{W}^1_{k_{nt}})'(k_{nt},K_t) = (\widetilde{W}^1_{k_{nt}k_{nt}})''(k_{nt},K_t)\dot{k}_{nt} + (\widetilde{W}^1_{k_{nt}})'(k_{nt},K_t)r + (\widetilde{W}^1_{K_t k_{nt}})''(k_{nt},K_t)\dot{K}_t. \tag{75}$$

Taking into account (74), Equation (75) leads to

$$\hat{\rho} = -\frac{\dot{k}_{nt}}{k_{nt}} + r,$$

and substituting in this expression $\dot{k}_{nt}$ by $\tilde{g}^1 k_{nt}$ expression (37) follows. □

Proof of Proposition 7. In the central-planner balanced path equilibrium, when $\sigma = 1$, $g^{1*} = g^{1*}_{\hat{\rho}}$, $s_{\rm A}^{1*} = s_{{\rm A}\hat{\rho}}^{1*}$ and the associated consumer welfare is

$$W^1(K_t) = \ln\left(c_t^* P_t^{-\phi}\right)\int_t^\infty \theta(\tau-t)d\tau + g^{1*}\left[1+\phi(\gamma-\lambda)\right]\int_t^\infty (\tau-t)\theta(u-t)d\tau =$$
$$\frac{\ln\left(c_t^* P_t^{-\phi}\right)}{\hat{\rho}} + \frac{g^{1*}\left[1+\phi(\gamma-\lambda)\right]\bar{J}_0}{\hat{\rho}} = W^1_{\hat{\rho}}(K_t) + \frac{g^{1*}\left[1+\phi(\gamma-\lambda\right]}{\hat{\rho}}\left(\bar{J}_0 - \frac{1}{\hat{\rho}}\right).$$

Likewise, in the market economy, $\tilde{g}^{1*} = \tilde{g}^{1*}_{\hat{\rho}}$, $\tilde{s}_{\rm A} = \tilde{s}_{{\rm A}\hat{\rho}} = \tilde{\alpha}\gamma$, and then, following the same reasoning

$$\widetilde{W}^1(K_t) = \widetilde{W}^1_{\hat{\rho}}(K_t) + \frac{\tilde{g}^{1*}\left[1+\phi(\gamma-\lambda)\right]}{\hat{\rho}}\left(\bar{J}_0 - \frac{1}{\hat{\rho}}\right).$$

From Lemma 2 it is known that $\Omega(g,\eta) < \bar{J}$, and particularizing for $\sigma = 1$, $\Omega(g,0) < \bar{J}_0$. Given that $\Omega(g,0) = 1/\hat{\rho}$, the last term in the expression above is positive. □

Proof of Proposition 9. The modified Ramsey rule under a central-planner or a market economy can be written as

$$g(s_{\rm A}) = r(s_{\rm A}) - \frac{1}{\Omega(g,\eta)}, \quad \tilde{g}(s_{\rm A}) = \tilde{r}(s_{\rm A}) - \frac{1}{\Omega(g,\eta)}.$$

By linearizing $g(s_{\rm A})$ at $s_{\rm A} = \tilde{s}_{\rm A}^{NT*} \equiv \tilde{\alpha}\gamma$, where superscript NT denotes the market economy with no taxes or subsidy, one gets[17]

$$g(s_{\rm A}) \simeq g(\tilde{\alpha}\gamma) + \frac{r'(\tilde{\alpha}\gamma)}{1+\eta\frac{\bar{J}}{\Omega(g,\eta)}}(s_{\rm A} - \tilde{\alpha}\gamma).$$

Moreover, from (18) and (33) it follows that $g(s_{\rm A}^*) = \tilde{g}(\tilde{s}_{\rm A}^{NT*})$. In consequence, the gap in the growth rates under the centralized and the market economy without taxes can be approximated as

$$g^* - \tilde{g}^{NT*} = g(s_{\rm A}^*) - \tilde{g}(\tilde{\alpha}\gamma) \simeq \frac{r'(\tilde{\alpha}\gamma)}{1+\eta\frac{\bar{J}}{\Omega(g,\eta)}}(s_{\rm A} - \tilde{\alpha}\gamma).$$

[17] Note that the Ramsey curve is inverse U-shaped and $\alpha\gamma$ is lower than $\hat{s}_{\rm A} = \alpha\beta\gamma^2/\lambda$ (where the Ramsey curve reaches its maximum). Hence, the linear approximation is inaccurate for values of $s_{\rm A} > \hat{s}_{\rm A}$.

Applying identical reasoning for the case of constant discounting, the gap between the growth rates reads

$$g^*_{\hat\rho} - \tilde g^{NT*}_{\hat\rho} = g_{\hat\rho}(s^*_{\text{A}}) - \tilde g_{\hat\rho}(\tilde\alpha\gamma) \simeq \frac{r'(\tilde\alpha\gamma)}{1+\eta}(s_{\text{A}\hat\rho} - \tilde\alpha\gamma).$$

The total effect of taxes on the market growth rate is

$$\begin{aligned} \frac{d\tilde g}{dT} = \frac{\partial \tilde g}{\partial s_{\text{A}}}\frac{ds_{\text{A}}}{dT} + \frac{\partial \tilde g}{\partial T} &= \frac{\frac{\partial \tilde r}{\partial s_{\text{A}}}\frac{ds_{\text{A}}}{dT} + \frac{\partial \tilde r}{\partial T}}{1+\eta\frac{\bar J}{\Omega(g,\eta)}} = \frac{\Psi}{1+\eta\frac{\bar J}{\Omega(g,\eta)}}, \\ \frac{d\tilde g_{\hat\rho}}{dT} = \frac{\partial \tilde g_{\hat\rho}}{\partial s_{\text{A}}}\frac{ds_{\text{A}}}{dT} + \frac{\partial \tilde g_{\hat\rho}}{\partial T} &= \frac{\frac{\partial \tilde r}{\partial s_{\text{A}}}\frac{ds_{\text{A}}}{dT} + \frac{\partial \tilde r}{\partial T}}{1+\eta} = \frac{\Psi}{1+\eta}. \end{aligned}$$

Then, since $\bar J > \Omega(g,\eta)$, from Lemma 2 it follows that

$$\left|\frac{d(g^* - \tilde g^{NT*})}{dT}\frac{T}{g^* - \tilde g^{NT*}}\right| < \left|\frac{d(g^*_{\hat\rho} - \tilde g^{NT*}_{\hat\rho})}{dT}\frac{T}{g^*_{\hat\rho} - \tilde g^{NT*}_{\hat\rho}}\right|$$

is equivalent to

$$\frac{\Psi}{s^*_{\text{A}} - \tilde\alpha\gamma} < \frac{\Psi}{s^*_{\text{A}\hat\rho} - \tilde\alpha\gamma} \Leftrightarrow s^*_{\text{A}\hat\rho} - \tilde\alpha\gamma < s^*_{\text{A}} - \tilde\alpha\gamma,$$

which is true from Proposition 4. □

References

[1] Andersen, S., Harrison, G.W., Morten, I.L. and Rutström, E.E. (2014). Discounting behavior: a reconsideration. European Economic Review 71, 15–33.

[2] Barro, R.J. (1999). Ramsey meets Laibson in the Neoclassical growth model. Quarterly Journal of Economics 114, 1125–1152.

[3] Barro, R. and Sala-i-Martin, X. (2004), Economic Growth. second ed., MIT Press Cambridge, Massachusetts.

[4] Cabo, F., Martín-Herrán, G. and Martínez-García, M.P. (2015). Non-constant discounting and Ak-type growth models. Economics Letters 131, 54–58.

[5] Cabo, F., Martín-Herrán, G. and Martínez-García, M.P. (2016). Unbounded growth in the Neoclassical growth model with non-constant discounting. Mathematical Social Sciences 84, 93–104.

[6] Caliendo, F.N. and Aadland, D. (2007). Short-term planning and the life-cycle consumption puzzle. Journal of Economic Dynamics and Control 31, 1392–1415.

[7] CastaÃ±er, A., Marín-Solano, J. and Ribas, C. (2016). Pareto inefficiency and dynamic bargaining in common property resource games with asymmetric players. www.bioecon-network.org/pages/18th$_$2016/Marin-Solano.pdf.

[8] Chichilnisky, G. (1996). An axiomatic approach to sustainable development, Social Choice and Welfare 13, 231–257.

[9] DellaVigna, S. (2009). Psychology and economics: Evidence from the field. Journal of Economic Literature 47, 315–372.

[10] Dyke, P. (2014), An Introduction to Laplace Transforms and Fourier Series, Springer, London, UK.

[11] Farzin, Y.H. and Wendner, R. (2014). The time path of the saving rate: hyperbolic discounting and short-term planning. MPRA Paper No. 54614, posted 19. March 2014.

[12] Findley, T.S. and Caliendo, F.N. (2014). Interacting mechanisms of time inconsistency. Journal of Economics Psychology 41, 68–76.

[13] Frederick, S., Loewnstein, G. and O'Donoghue, T. (2002). Time discounting and time preference: a critical review. Journal of Economics Literature 40, 351–401.

[14] Fujii, T and Karp, L. (2008). Numerical analysis of non-constant pure rate of time preference: a model of climate policy. Journal of Environmental Economics and Management 56, 83–101.

[15] Gerlagh, R and Liski, M. (2017). Climate policies with non-constant discounting. Journal of the European Economic Association doi:10.1093/jeea/jvx010.

[16] Goldman, S.M. (1980). Consistent plans. Review of Economic Studies 47, 533–537.

[17] Groom, B., Hepburn, C., Koundouri, P. and Pearce, D. (2005). Declining Discount Rates: The Long and the Short of it. Environmental and Resource Economics 32, 445–493.

[18] Hepburn, C., Duncan, S. and Papachristodoulou, A. (2010) Behavioural economics, hyperbolic discounting and environmental policy. Environmental and Resource Economics 46, 189–206.

[19] Hiraguchi, R. (2014). On the neoclassical growth model with non-constant discounting. Economics Letters 125, 175–178.

[20] Hiraguchi, R. (2016). On a two-sector endogenous growth model with quasi-geometric discounting. Journal of Mathematical Economics 65, 26–35.

[21] Karp, L. (2005). Global warming and hyperbolic discounting. Journal of Public Economics, 89, 261–282.

[22] Karp, L. (2007). Non-constant discounting in continuous time. Journal of Economic Theory, 132, 557–568.

[23] Karp, L. and Tsur, Y. (2011). Time perspective and climate change policy. Journal of Environmental Economics & Management 62, 1–14.

[24] Krusell, P., Kuruşçu, B. and Smith, A. A. Jr. (2002). Equilibrium welfare and government policy with quasi-geometric discounting. Journal of Economic Theory 105, 42–72.

[25] Laibson, D. (1997). Golden eggs and hyperbolic discounting. Quarterly Journal of Economics 112, 443–477.

[26] Li, C.Z. and Lofgren, K.G. (2000). A dynamic analysis with heterogeneous time preferences. Journal of Environmental Economics & Management, 40, 236–250.

[27] Marín-Solano, J. (2015). Group inefficiency in a common property resource game with asymmetric players. Economic Letters 136, 214–217.

[28] Marín-Solano, J. and Navas, J. (2009). Non-constant discounting in finite horizon: the free terminal time case. Journal of Economic Dynamics & Control 33, 666–675.

[29] Pollak, R.A. (1968). Consistent Planning. Review of Economic Studies 35, 201–208.

[30] Ramsey, F. (1928). A mathematical theory of saving. Economic Journal 38, 543–559.

[31] Samuelson, P.A. (1937). A note on measurement of utility. Review of Economic Studies 4, 155–161.

[32] Smulders, S. and Gradus, R. (1996). Pollution abatement and long-term growth. European Journal of Political Economy 12, 505–532.

[33] Strotz, R.H. (1956). Myopia and inconsistency in dynamic utility maximization. Review of Economic Studies 23, 165–80.

[34] Strulik, H. (2015). Hyperbolic discounting and endogenous growth. Economics Letters 126, 131–134.

[35] Tsoukis, C., Tournemaine, F. and Gillman, M. (2017). Hybrid exponential-hyperbolic discounting and growth without commitment. The Manchester School, doi:10.1111/manc.12197.