Be Careful What You Calibrate For: Social Discounting in General Equilibrium

Lint Barrage*

July 27, 2016

Abstract

At what rate should policy-makers discount the future? One influential view posits that the social discount rate should be set below the market rate of return. This idea has risen to renewed prominence in climate change economics (e.g., Stern, 2006). This paper formalizes the broader policy implications of this view by characterizing jointly optimal environmental and fiscal policies in a dynamic general equilibrium climate-economy model with differential planner-household discounting. First, I show that decentralizing the first-best allocation requires not only high carbon prices but also fundamental changes to tax policy: If the government discounts the future less than households, implementing the optimal allocation requires (i) capital income subsidies, and/or (ii) decreasing labor income taxes, and/or (iii) decreasing consumption taxes. Second, for a 'Sternian environmental planner' who can set carbon prices but cannot change income taxes, the *constrained-optimal* carbon tax is up to 50% *below* the present value of marginal damages (the social cost of carbon) due to the general equilibrium effects of climate policy on household savings. Third, given the choice to optimize either carbon or capital taxes, the Sternian planner's welfare ranking is ambiguous and depends critically on the intertemporal elasticity of substitution. Overall, in general equilibrium, a policy-maker's choice to adopt differential social discounting may thus overturn conventional recommendations for both environmental and fiscal policy.

Keywords: Discounting, social discount rate, carbon taxes, Ramsey taxation, distortionary taxes, general equilibrium climate-economy model

*Brown University, Department of Economics and Institute at Brown for Environment and Society (IBES). Contact Email: lint_barrage@brown.edu. I thank Jesse Shapiro for the inspiration for this paper's research question and for detailed comments and support. I also thank Cihan Artunc, Laura Bakkensen, Neil Mehrotra, Mark Roberts, Kieran Walsh, and seminar participants at the Yale Energy and Environment Day, the Columbia Sustainable Development Seminar, ICMAIF 2016, and AERE 2016 for their comments.

1 Introduction

How should policy-makers discount the future? Economists have long debated the appropriate social rate of discount.[1] Recently, in the context of the economics of climate change, social discounting has again risen to the forefront of academic and policy debates (e.g., Arrow et al., 1995; 2012). As the impacts of greenhouse gas emissions occur over long time horizons, optimal climate policy depends critically on discounting parameters (Nordhaus, 2007, 2008; Interagency Working Group, 2010). One influential view posits that it is "ethically indefensible" to discount the utility of future generations, as famously stated by Frank Ramsey in his original work on optimal savings (Ramsey, 1928). Economists embracing this view have incorporated near-zero pure rates of social time preference into their models, yielding aggressive climate policy recommendations (Stern, 2006; Cline, 1992). This approach has been controversial for several reasons, including the fact that it does not align with households' intertemporal preferences as revealed through their savings behavior (Nordhaus, 2007; Dasgupta, 2008). Indeed, Ramsey himself noted that "the rate of saving which the [zero-discounting optimality] rule requires is greatly in excess of that which anyone would normally suggest" (Ramsey, 1928). While Ramsey's ethical critique is widely cited in support of near-zero discounting, his work also reminds us that such discounting may not match household behavior. Several studies have thus proposed frameworks where agents and the planner discount utility differently (e.g., Farhi and Werning, 2005, 2007, 2010; Kaplow, Moyer, and Weisbach, 2010; Goulder and Williams, 2012; von Below, 2012; Belfiori, 2016).

This paper formalizes and quantifies the broader policy implications of differential discounting in a dynamic general equilibrium climate-economy model. While it has been noted that social discounting should affect, e.g., capital allocations (e.g., Manne, 1995; Caplin and Leahy, 2004; Goulder and Williams, 2012), few studies have formalized these effects. I build on related prior work (Farhi and Werning, 2005, 2007, 2010; von Below, 2012, Belfiori, 2016) in three ways:

First, I show that differential discounting fundamentally alters optimal tax policy prescriptions. In a first-best setting, decentralizing the optimal allocation requires both high carbon taxes and capital income subsidies. Intuitively, this is because households are too impatient from the planner's perspective, and consequently fail to invest sufficiently in the economy's capital stocks. While this insight was previously formalized by von Below (2012) in the first-best,[2] I further

[1] This vast literature ranges from debates on the fundamental principles of what discounting should capture (e.g., Baumol, 1968; Sen, 1982; Lind, 1982) to modern treatments about the implications of factors such as heterogeneity (e.g., Gollier and Zeckhauser, 2005; Heal and Millner, 2013), uncertainty (e.g., Newell and Pizer, 2003; Gollier and Weitzman, 2010), the structure of preferences (e.g., Caplin and Leahy, 2004), and commitment (Sleet and Yeltekin, 2006), inter alia.

[2] Belfiori (2016) also considers energy policy under differential discounting in the first best, but focuses on a setting with a consumption-production good (oil) that is a scarce resource. She shows that this fossil resource is depleted too quickly from the planner's perspective when he values future generations differently than households, analogous to insufficient capital investment here. She further shows that Pigouvian carbon

show that it extends to a more realistic fiscal setting where government revenues must be raised from distortionary taxes. If the government values the future more than households, I show that decentralizing the optimal allocation requires (i) capital income subsidies, (ii) labor income taxes that *decrease* over time, and/or (iii) consumption taxes that *decrease* over time. On the one hand, these policies stand in stark contrast to the classic prescriptions with standard discounting: zero capital income taxes, constant (smooth) labor income taxes, and constant (uniform) consumption taxes (see, e.g., Chari and Kehoe, 1999; Atkeson, Chari, and Kehoe, 1999). On the other hand, the result on capital income subsidies aligns with Farhi and Werning (2005, 2010), who study differential discounting in a Mirrleesian economy[3] with overlapping generations. They show that the constrained-efficient allocation can be decentralized by negative marginal estate taxes, indicating that bequests should be subsidized. Even though the Mirrleesian and Ramsey settings can generally lead to different tax policy recommendations, with differential discounting, the results of this paper thus suggest that both frameworks imply the desirability of (effective) savings subsidies.

Second, I consider the more realistic policy problem of an environmental planner who cannot create capital income subsidies. This setup is motivated by the observation that climate and fiscal policy are often set separately, and that capital income is taxed in many countries. I find that an environmental planner with "Sternian" preferences may not want to impose a "Sternian" carbon tax (i.e., a tax equal to the socially discounted marginal damages of emissions) if he cannot subsidize savings at the same time. More formally, the constrained-optimal carbon tax must be adjusted for the general equilibrium effects of climate policy on households' incentives to save: (i) If climate change decreases the returns to capital investments, a component must be added to the optimal carbon tax. (ii) If energy and capital are complements in production, a component must be subtracted from the optimal carbon tax. Intuitively, without a capital income subsidy, private savings are too low. To the extent that climate policy decreases (increases) the returns to capital, it therefore exacerbates (mitigates) this inefficiency, and must be adjusted accordingly.[4] Theoretically, the net sign of these adjustments is ambiguous.

Third, I therefore empirically quantify optimal policies and welfare by integrating differential discounting into modified versions of two climate-economy models. In order to match the

---

taxes are no longer sufficient to decentralize the optimal allocation, and presents several decentralizations that address both the depletion and emissions externalities, including in a cap-and-trade setting.

[3] In this setting, tax policy is designed to trade off insurance and incentives for agents that experience unobservable productivity shocks, and, with differential discounting, face the risk of being born into families with more or less productive parents. Distortions arise as the planner can only observe households' output, but not their work effort or productivity. In contrast, in the Ramsey setting considered here, distortions arise as the planner must raise a set amount of revenues but is assumed to be unable to impose lump-sum taxes.

[4] One could also ask how capital should be taxed when carbon cannot be priced. For a treatment of this issue in a decentralized economy with altruistic households that invest in dirty capital to protect their offspring against climate change (but thereby exacerbate the externality), see Asheim and Nesje (2016).

benchmark theoretical framework, I first employ the generalized numerical implementation of Golosov, Hassler, Krusell, and Tsyvinski's model (2014) by Barrage (2014). In order to consider the extended theoretical framework with endogenous labor supply and distortionary taxes, I then integrate differential discounting into the COMET model (Barrage, 2015). The COMET builds on the seminal DICE framework of Nordhaus (see, e.g., 2008, 2010) by incorporating government expenditures, fiscal policy choice, endogenous labor supply, and a separate accounting of climate damages affecting production (e.g., agriculture) versus direct utility losses (e.g., biodiversity existence value, damage to archaeological sites, etc.).

The central quantitative results are as follows. Adopting the pure rate of social time preference advocated by Stern (2006) (0.1% per year) dramatically increases the optimal carbon price, as expected. Depending on the intertemporal elasticity of substitution, optimal carbon taxes range from \$30-\$62 per metric ton (mtC, \$2005) with standard discounting, but increase to \$42-\$486/mtC with social discounting. This policy limits global temperature change to $2°C$, compared to $3°C$ optimal peak temperature change when the planner adopts households' preferences. Decentralizing these allocations also requires an investment subsidy of $\sim$ 15% in the first-best, and capital income subsidies beginning at 30% and increasing to 65% by 2100 in the COMET.[5] With distortionary fiscal policy, social discounting further changes optimal labor taxes from a constant rate of 41% to a high initial rate of 53% that decreases to 36% by the end of the century. Overall, these results indicate that differential social discounting at the parameters advocated in the literature would imply quantitatively large departures from current policy practice and standard recommendations.

Next, solving the "environmental planner's problem," I find that optimal carbon taxes are 5-50% lower when the planner cannot subsidize savings. Intuitively, this is because higher energy prices reduce the marginal product of capital, thereby reducing the return to savings even further away from its social optimum. Since the planner cares about the overall level of assets given to future generations - both natural and man-made - he adjusts carbon taxes downward to mitigate their undesirable effects on capital accumulation. While the size of this adjustment is sensitive to the parameters, the results imply that a constrained fiscal environment may significantly decrease optimal carbon taxes even from the perspective of a socially discounting planner.

Finally, I compare the welfare gains from optimizing carbon versus capital income taxes from the perspective of a planner with a near-zero pure rate of social time preference. Perhaps surprisingly, the relative importance of these policies is ambiguous, and depends critically on the intertemporal elasticity of substitution. A "Sternian" social planner with logarithmic preferences would indeed achieve significantly larger welfare gains from setting climate policy instead of fiscal

[5] The investment subsidy considered in the first-best also falls on undepreciated capital and is thus not directly comparable to the standard capital income subsidy considered in the COMET.

policy. However, for a consumption elasticity of $\sigma = 1.5$ (e.g., Nordhaus, 2008), the welfare gains from capital income subsidies are estimated to be higher than those from increasing carbon taxes (from privately discounted to socially discounted levels). While these calculations abstract from many important complexities, and are thus subject to numerous caveats, they do illustrate that climate protection is only one of various investments society can make to benefit future generations (e.g., health, education, private capital, etc.).[6] A government that weighs the utility of future households more highly than its citizens should thus incentivize larger investments in all such assets according to their real rates of return.

The remainder of this paper proceeds as follows. Section 2 sets up the benchmark model and describes policies that decentralize the optimal allocation in the first-best. Section 3 characterizes constrained-optimal climate policy for an environmental planner who cannot subsidize capital income. Section 4 presents the extension to a setting with distortionary taxes. Section 5 summarizes the calibration and the quantitative results. Finally, Section 6 concludes.

2 Benchmark Theoretical Model

2.1 Decentralized Economy

Households

An infinitely-lived,[7] representative household has well-behaved preferences over an aggregate consumption good C_t and the stock of carbon concentrations S_t, with lifetime utility:

$$U_0 \equiv \sum_{t=0}^{\infty} \beta_H^t U(C_t, S_t) \tag{1}$$

The household's pure rate of social time preference ρ_H defines the private discount factor $\beta_H \equiv 1/(1+\rho_H)$. This specification differs from GHKT and von Below (2012) by allowing carbon concentrations to affect both utility and output. Utility impacts represent non-production effects such as damages to cultural heritage sites. Section 3 shows that this distinction is important for constrained-optimal climate policy design. The household earns income from labor L_t (inelastically supplied in the benchmark but endogenized in Section 4), return r_{t+1} on capital holdings K_t (net of depreciation δ and capital income tax τ_{kt}), and profits Π_t from the energy production

6 Goulder and Williams (2012) make this point theoretically. Here, I provide quantitative evidence from a well-known climate economy model (GHKT) to support their assertion.

7 As noted by Farhi and Werning (2005), there is an equivalence between a dynastic household with a lower discount factor than the planner - as in this paper - and an overlapping generations setting where the planner values each generation both directly and through current households' valuation of their descendants. See also Belfiori (2016) for a formal mapping between a planner's Pareto weights on future generations and the social discount factor.

sector, with associated flow budget constraint:

$$C_t + K_{t+1} \leq w_t L_t + \{1 - \delta + r_t\} (1 - \tau_{kt}) K_t + \Pi_t + G_t \tag{2}$$

where G_t denotes government lump-sum transfers. This benchmark version of the model abstracts from consumption and labor taxes as they are not needed to decentralize the first-best allocation. That is, as the only distortions in the economy are the climate externality and the difference in planner and household discount rates, the two tax instruments considered - capital and carbon taxes - are sufficient to decentralize the optimal allocation.[8] The household seeks to maximize (1) subject to (2), yielding the following savings optimality condition:

$$\frac{U_{ct}}{U_{ct+1}} = \beta_H \{1 - \delta + r_{t+1}\} (1 - \tau_{kt+1}) \tag{3}$$

where U_{ct} denotes the marginal utility of consumption in period t.

Production

Aggregate output Y_t is produced competitively using capital K_t, labor L_t^Y, and energy E_t inputs in a constant returns to scale (CRS) technology that satisfies the usual Inada conditions. Output further depends on carbon concentrations S_t and technology A_t :

$$Y_t = F(A_t(S_t); K_t, L_t^Y, E_t) \tag{4}$$

Profit-maximizing firms equate marginal products with factor prices in equilibrium:

$$\begin{aligned} F_{Lt} &= w_t \\ F_{Kt} &= r_t \\ F_{Et} &= p_{Et} \end{aligned} \tag{5}$$

where F_{jt} denotes the marginal product of input j at time t, and p_{Et} is the price of energy. Final output can be consumed or invested, with associated aggregate resource constraint:

$$C_t + K_{t+1} = Y_t + (1 - \delta) K_t \tag{6}$$

The energy input E_t is assumed to be produced with labor inputs L_t^E and CRS technology:[9]

$$E_t = G_t(L_t^E) \tag{7}$$

[8] Of course, other combinations of taxes could also decentralize the optimal allocation.

[9] See Belfiori (2016) for a detailed treatment of non-renewable resource dynamics with differential discounting.

Labor is perfectly mobile across sectors ($L_t = L_t^Y + L_t^E$), implying that wages will be equated in equilibrium. Finally, the statutory incidence of excise carbon fuel taxes τ_{Et} is on energy producers, whose profit maximization problem and condition in each period t are given by:

$$\begin{aligned} &\max_{L_t^E} (p_{Et} - \tau_{Et}) G_t(L_t^E) - w_t L_t^E \\ \Rightarrow \quad &(p_{Et} - \tau_{Et}) G_{Lt} = w_t \end{aligned} \tag{8}$$

Environment

Atmospheric carbon concentrations S_t are a function $\widetilde{S}_t$ of initial concentrations S_0 and emissions E_t dating back to the start of industrialization at time $-T$:

$$S_t = \widetilde{S}_t \left(S_0, E_{-T}, E_{-T+1}, ..., E_t \right) \tag{9}$$

with $\frac{\partial \widetilde{S_{t+j}}}{\partial E_t} > 0 \; \forall j \geq 0$. Note that energy inputs E_t are measured in tons of carbon-equivalent.

2.2 Planner's Problem: First-Best Policy

The government seeks to maximize households' lifetime utility (1), but can choose to discount the future at a different rate than agents. Let the government's pure rate of social time preference ρ_G define the *social utility discount factor* via $\beta_G \equiv 1/(1+\rho_G)$. In the first-best setting the planner has access to an unrestricted set of policy instruments, and maximizes socially discounted utility subject to aggregate production (4), the resource constraint (6), the climate change constraint (9), energy production (7), and the labor resource constraint (see Appendix). The optimal allocation can then be characterized as follows. First, combine the planner's first-order conditions (FOCs) with respect to energy E_t, carbon concentrations S_t, and consumption C_t to obtain the standard optimality condition that the social marginal costs and benefits of carbon energy usage be equated:

$$\underbrace{\frac{\partial Y_t}{\partial E_t}}_{\substack{\text{Marginal} \\ \text{product of} \\ \text{energy}}} - \underbrace{\sum_{j=0}^{\infty} \beta_G^j \left[\frac{U_{ct+j}}{U_{ct}} \frac{\partial Y_{t+j}}{\partial S_{t+j}} + \frac{U_{S_{t+j}}}{U_{ct}} \right] \frac{\partial S_{t+j}}{\partial E_t}}_{\substack{\text{SDV of marginal} \\ \text{damages from} \\ \text{carbon emissions}}} = \underbrace{\frac{w_t}{\frac{\partial G_t}{\partial L_t^E}}}_{\substack{\text{MC of} \\ \text{energy production} \\ \text{(in } C_t \text{ units)}}} \tag{10}$$

Next, combining the planner's FOCs with respect to consumption C_t and capital K_{t+1} yields the planner's Euler equation:

$$[F_{Kt+1} + (1-\delta)] = \frac{1}{\beta_G} \frac{U_{ct}}{U_{ct+1}} \tag{11}$$

2.2.1 Decentralization and Optimal Tax Rates

Comparing the planner's optimality conditions with those governing the behavior of households (3) and energy producers (8) in decentralized equilibrium, it is easy to show the following:

Result 1: *The first-best allocation can be decentralized by the combination of a carbon tax set equal to the socially discounted value (SDV) of marginal damages of emissions at the optimum,*

$$\tau^*_{Et} = -\sum_{j=0}^{\infty}\beta_G^j\left[\frac{U_{ct+j}}{U_{ct}}\frac{\partial Y_{t+j}}{\partial S_{t+j}}+\frac{U_{S_{t+j}}}{U_{ct}}\right]\frac{\partial S_{t+j}}{\partial E_t} \tag{12}$$

$$= \textit{SDV[Output Damages+Utility Damages]} \tag{13}$$

and a capital income subsidy to equate the social and private marginal returns on savings:

$$\tau^*_{Kt+1} = 1-\frac{\beta_G}{\beta_H} \tag{14}$$

Proof: See Appendix A. With standard or "descriptive" discounting where the government adopts household preferences ($\beta_G = \beta_H$), the optimal capital income tax is thus zero. If, however, the planner is more patient than the household, $\beta_G > \beta_H$, then the optimal capital income tax is negative, $\tau^*_{Kt+1} < 0$. This theoretical need for a capital income subsidy with differential discounting was previously formally demonstrated by von Below (2012)[10] and also mentioned by e.g., GHKT. However, this prescription stands in stark contrast to the common policy practice of taxing capital income. The next two sections thus introduce two layers of realism into the analysis: (1) separation of policy powers, and (2) government revenue requirements that must be met through distortionary taxes. Both model features lead to fundamentally different environmental and tax policy prescriptions with social discounting compared to classic results in the literature.

3 Constrained-Optimal Climate Policy

The previous section showed that, if the government discounts the future less than households, then achieving the first-best allocation theoretically requires capital income subsidies. In reality, however, both academic debates and policy decisions surrounding carbon pricing are often made separately from fiscal policy. Consequently, this section characterizes *constrained-optimal*

[10] Specifically, expression (14) is the same as derived by von Below (2012). However, condition (12) differs in allowing for utility damages and more general preferences (i.e., non-logarithmic), but abstracting from Hotelling dynamics in energy production. Both the extensions to constrained and distortionary fiscal settings presented below are also novel in this paper.

climate policy from the perspective of a "Sternian environmental planner" who believes in social discounting but cannot subsidize savings. More formally, in the primal approach, the no-capital subsidy constraint can be formalized as follows:

$$\frac{U_{ct}}{\beta_H U_{ct+1}} = F_{Kt+1}(1-\tau_k) \leq F_{Kt+1} \tag{15}$$

In addition to (15), I now assume that carbon and consumption enter the household's utility separably. The Online Appendix formally shows that non-separability does not change the results presented below, but may add another term to the optimal carbon tax formulation.[11]

Proposition 1 *The constrained-optimal carbon tax not only internalizes the socially discounted value (SDV) of marginal damages of carbon emissions, but must also account for the effects of climate change and energy taxes on capital returns and thus incentives to save:*

$$\begin{aligned} \tau^*_{Et} &= \textit{SDV[Output Damages+Utility Damages]} \\ &\quad + \textit{Climate Investment Incentives Adjustment} \\ &\quad + \textit{Energy Use Investment Incentives Adjustment} \end{aligned}$$

In particular:

(i) If climate damages decrease the returns to capital investments, a component must be added to the optimal carbon tax τ^*_{Et} *:*

$$\begin{aligned} &\textit{Climate Investment Incentives Adjustment} \\ \propto\ &\sum_{j=0}^{\infty} \beta_G^j \cdot \underbrace{\frac{\partial S_{t+j}}{\partial E_t}\left(-\frac{\partial Y_{Kt+j}}{\partial S_{t+j}}\right)}_{\substack{\textit{Climate damages to} \\ \textit{marginal product of capital}}} \end{aligned} \tag{16}$$

(ii) If energy and capital are complements in production, a component must be subtracted

[11] Non-separability specifically implies that the planner must account for the effects of climate change on the household's intertemporal marginal rate of substitution if the shadow value of capital subsidies is changing over time. Both for parsimony and given the limited empirical evidence on how direct utility impacts of climate change (e.g., damages to cultural heritage sites) enter households' preferences, Proposition 1 thus focuses on the benchmark case with separability.

from the optimal carbon tax τ^*_{Et} :

$$\propto \underbrace{\left(-\frac{\partial Y_{Kt}}{\partial E_t}\right)}_{\substack{\text{Energy use effect on} \\ \text{marginal product of capital}}}$$

Energy Use Investment Incentives Adjustment

(iii) The constrained-optimal tax can be higher or lower than the SDV of marginal damages (the social cost of carbon).

Proof: See Appendix A. The central implication of Proposition 1 is that an environmental planner with "Sternian" preferences may not actually want to impose a "Sternian" carbon tax (i.e., a tax equal to the socially discounted marginal damages of emissions) if he cannot subsidize savings at the same time. Intuitively, this is because the planner cares about the overall level of assets given to future generations - both natural and man-made - and consequently takes into account the general equilibrium effects of carbon taxes on households' savings behavior. Without a capital income subsidy, private savings are too low. Consequently, climate policy's impacts on private investment can now have first-order welfare effects, analogous to settings with pre-existing distortions from income taxes (e.g., Bovenberg and Goulder, 1996) or imperfect competition (e.g., Ryan, 2012). On the one hand, carbon taxes should be increased (all else equal) to the extent that climate change will decrease the marginal product of capital. Intuitively, this is because these damages reduce private returns to savings even further away from the social optimum.[12] On the other hand, carbon taxes should be reduced (all else equal) to the extent that energy price increases lower the marginal product of capital. Intuitively, this is again because households are already insufficiently incentivized to save, implying that any further decrease in the marginal product of capital due to higher energy costs will exacerbate this social inefficiency. In seeking to optimize future generations' overall welfare, a Sternian environmental planner should adjust climate policy to take these effects into account.

4 Distortionary Fiscal Setting

The analysis presented so far has implicitly assumed that governments can impose lump-sum taxes. In reality, however, public funds must be raised through distortionary instruments, such as labor income taxes. This section thus adds government revenue requirements and linear tax

[12] A first-order importance of the *nature* of climate damages - holding levels constant - has also been documented in the presence of pre-existing tax distortions (e.g., Williams, 2002; Barrage, 2015).

instruments in the Ramsey tradition (see, e.g., Chari and Kehoe, 1999) to the model. The objectives of this extension are to study both (i) the robustness of the desirability of capital income subsidies to a setting where they must be financed with distortionary taxes, and (ii) the broader fiscal policy implications of differential social discounting. The latter question has been of independent interest in the public finance literature (e.g., Farhi and Werning, 2005, 2007, 2010), to which this paper adds an analysis in the Ramsey setting.

First, I introduce preferences for leisure, changing household utility to:

$$U_0 \equiv \sum_{t=0}^{\infty} \beta_H^t U(C_t, L_t, S_t) \tag{17}$$

Second, the household is now subject to labor and consumption taxes as well as more realistic capital income taxes (falling only on returns net-of-depreciation):

$$(1+\tau_{ct})C_t + \rho_t B_{t+1} + K_{t+1} \leq w_t(1-\tau_{lt})L_t + \{1+(r_t-\delta)(1-\tau_{kt})\} K_t + B_t + \Pi_t \tag{18}$$

Here, τ_{ct} denotes the consumption tax in period t, B_{t+1} is purchases of one-period government bonds at price ρ_t, K_{t+1} denotes the household's capital holdings in period $t+1$, τ_{lt} the linear labor income tax, τ_{kt} the linear net-of-depreciation capital income tax, B_t repayments of government bond holdings, and the remaining variables are as defined above. The household's optimality conditions for its labor-consumption and savings decisions in this setup become:

$$\frac{-U_{lt}}{U_{ct}} = \frac{w_t(1-\tau_{lt})}{(1+\tau_{ct})} \tag{19}$$

$$\frac{U_{ct}}{U_{ct+1}} = \beta_H \{1+(r_{t+1}-\delta)(1-\tau_{kt+1})\} \frac{(1+\tau_{ct})}{(1+\tau_{ct+1})} \tag{20}$$

The government must finance an exogenously given sequence of revenue requirements $\{G_t\}_{t=0}^{\infty}$. As is common, G_t is first modeled as wasteful government consumption. The quantitative version of the model further adds social transfers to households. The government can raise revenues by issuing bonds and levying taxes, with corresponding flow budget constraint:

$$G_t + B_t^G = \tau_{lt} w_t L_t + \tau_{ct} C_t + \tau_{kt}(r_t-\delta)K_t + \tau_{Et} E_t + \rho_t B_{t+1}^G \tag{21}$$

It should be noted that not all of these tax instruments are needed to form a "complete" tax system (Chari and Kehoe, 1999). That is, the same allocation can be decentralized by many different tax systems that correspond to the same overall wedges between the relevant marginal rates of substitution, as can be readily seen for the use of either τ_{lt} or τ_{ct} to create a wedge in the consumption-leisure tradeoff (19). I thus consider both a version of the model where the untaxed numeraire is the consumption good or capital investments, respectively.

Competitive equilibrium in this economy is defined in the standard way, with the addition of the carbon cycle constraint and pre-industrial concentrations S_0 as initial condition. The government's objective is to implement the competitive equilibrium that yields the highest household lifetime utility (17) - discounted at the social rate β_G - for a given set of initial conditions (K_0, B_0, S_0, $\overline{\tau_{k0}}$, $\overline{\tau_{c0}}$). As is standard, the initial capital income tax $\overline{\tau_{k0}}$ and consumption tax $\overline{\tau_{c0}}$ are assumed to be exogenously given as they can otherwise be used as effective lump-sum taxes. I also assume that the government can commit to a sequence of capital income taxes.[13]

The set of allocations that can be decentralized as a competitive equilibrium can be characterized by two sets of constraints: feasibility and an "implementability" constraint that captures the optimizing behavior of households and firms. In the present setting, it is straightforward to show (see Chari and Kehoe, 1999, for a general discussion, and Barrage, 2015, for a derivation in a climate-economy model very similar to the present framework) that the implementability constraint that must be added to the planner's problem is given by:

$$+\phi\left[\sum_{t=0}^{\infty}\beta_H^t\underbrace{(U_{ct}C_t+U_{lt}L_t)}_{\equiv W_t}-\left\{\frac{U_{c0}}{1+\overline{\tau_{c0}}}\left[K_0\left\{1+(F_{K0}-\delta)(1-\overline{\tau_{k_0}})\right\}\right]\right\}\right] \tag{IMP}$$

where ϕ denotes the Lagrange multiplier on (IMP). The main difference between (IMP) and the standard setting is that one must be careful to employ the household's discount factor β_H rather than the planner's β_G in the derivation of (IMP) as it captures the optimizing behavior of *households*. Letting $W_t \equiv (U_{ct}C_t + U_{lt}L_t)$ denote the bracketed term on the left-hand side of (IMP), the planner's problem leads to the following characterization of optimal tax policy:

4.1 Capital Income Taxes

This section formally shows that capital income subsidies remain desirable if the government values the future more than households, even if they must be financed through distortionary taxes. First, note that the planner's FOCs (for $t > 0$) imply that the net return on investment

13 This assumption is common (see discussion in Chari and Kehoe, 1999) in the Ramsey taxation literature, but is known not to be without loss of generality (e.g., Klein and Rios-Rull, 2003; Benhabib and Rustichini, 1997). However, focusing on carbon taxes alongside other distortionary taxes, Schmitt (2013) finds that relaxing the assumption of commitment has only a minor quantitative effect on optimal policy.

should evolve according to:

$$\beta_H \left[(1-\delta)+F_{Kt+1}\right] = \frac{\left[\left(\frac{\beta_G}{\beta_H}\right)^t U_{ct} + \phi W_{ct}\right]}{\left[\left(\frac{\beta_G}{\beta_H}\right)^{t+1} U_{ct+1} + \phi W_{ct+1}\right]} \tag{22}$$

where W_{ct} denotes the partial derivative of expression W_t in (IMP) with respect to consumption at time t. Conversely, in the decentralized economy, the representative household's optimality condition for savings (when consumption is the untaxed numeraire) is given by:

$$\beta_H \left[1 + (F_{Kt+1}-\delta)(1-\tau_{kt+1})\right] = \frac{U_{ct}}{U_{ct+1}} \tag{23}$$

Whether and what kind of capital income tax τ_{kt+1} is needed to align private (23) and public (22) savings rules thus depends on W_t and hence the utility function. I proceed considering two commonly used constant elasticity of substitution preferences that satisfy consistency with balanced growth (see King, Plosser, and Rebelo, 2001):

$$U(C_t, L_t, S_t) = \log C_t + v(L_t) + \vartheta(S_t) \tag{A}$$

$$U(C_t, L_t, S_t) = \frac{(C_t L_t{}^{-\gamma})^{1-\sigma}}{1-\sigma} + \vartheta(S_t) \tag{B}$$

where $v(L_t)$ is an increasing and concave function of leisure $(1-L_t)$. With preferences of the form (A), it is straightforward to show that $W_{ct} = 0$. Consequently, the planner's optimality condition for investment (22) becomes:

$$\beta_H \left[(1-\delta)+F_{Kt+1}\right] = \frac{U_{ct}}{U_{ct+1}}\left(\frac{\beta_H}{\beta_G}\right) \tag{24}$$

Comparing (23) and (24) demonstrates that the capital income tax required to decentralize the optimal allocation for $t > 1$ is defined by:

$$\tau^*_{kt+1} = \left(\frac{\beta_H - \beta_G}{\beta_H}\right)\frac{(F_{Kt+1}-\delta+1)}{(F_{Kt+1}-\delta)}. \tag{25}$$

It immediately follows from expression (25) that, if the government is more patient than households ($\beta_G > \beta_H$), the optimal capital income tax is negative, or a subsidy. Conversely, with standard discounting ($\beta_G = \beta_H$), the optimal tax term (25) reduces to zero, in line with the classic policy prescription that capital distortions are undesirable in a wide range of settings (Atkeson, Chari, and Kehoe, 1999; Judd,1999; Chari and Kehoe, 1999, etc.).

For preferences of the form (B), one can easily show that $W_{ct} = U_{ct}[1-\sigma-\gamma(1-\sigma)]$.

Substituting into the planner's optimality condition for savings (22) and comparing it with the household's (23), Appendix B show that the optimal capital income tax is defined by:

$$\tau^{*}_{kt+1} = \left(\frac{\omega_t - 1}{\omega_t}\right)\frac{(F_{Kt+1} - \delta + 1)}{(F_{Kt+1} - \delta)} \tag{26}$$

where

$$\omega_t \equiv \frac{\left[\left(\frac{\beta_G}{\beta_H}\right)^{t} + \phi[1 - \sigma - \gamma(1-\sigma)]\right]}{\left[\left(\frac{\beta_G}{\beta_H}\right)^{t+1} + \phi[1 - \sigma - \gamma(1-\sigma)]\right]} \tag{27}$$

As before, with standard discounting ($\beta_G = \beta_H$), we see that $\omega_t = 1$ and that the optimal capital income tax (26) reduces to zero. In contrast, if the planner values the future more than households ($\beta_G > \beta_H$), Appendix B shows that $0 < \omega_t < 1$, and that capital income taxes (26) must be negative. Proposition 2 summarizes these findings:

Proposition 2 *If the social planner values the future more than households ($\beta_G > \beta_H$), and if preferences are of the forms (A) or (B) and consistent with balanced growth, then the optimal tax policy requires a capital income subsidy for all periods* $t + 1 > 1$, *even when the necessary revenues must be raised through distortionary taxes.*

Proof: See Appendix B.

4.2 Consumption Taxes

While the previous section treated the consumption good as the untaxed numeraire, the optimal allocation can also be decentralized by a fiscal system of consumption, labor, and emissions taxes. In particular, consider a decentralized economy where capital investments are the untaxed numeraire. The household's optimal savings condition (20) then becomes:

$$\frac{U_{ct}}{U_{ct+1}} = \beta_H\left[(1-\delta) + F_{Kt+1}\right]\frac{(1+\tau_{ct})}{(1+\tau_{ct+1})} \tag{28}$$

For separable preferences of the form (A), comparison of (28) with the planner's optimality condition for savings (24) shows that, on order to decentralize the optimal allocation, consumption taxes for $t > 0$ must satisfy:

$$\frac{(1+\tau^{*}_{ct+1})}{(1+\tau^{*}_{ct})} = \frac{\beta_H}{\beta_G} \tag{29}$$

Condition (29) immediately implies that consumption taxes must be decreasing over time (for $t > 0$) if the government values the future more than households ($\beta_G > \beta_H$). The intuition for this result is straightforward: Capital income taxes are equivalent to ever-increasing consumption taxes, as shown by Judd (1999). Consequently, the need for capital income subsidies - an incentive to increase savings and delay consumption - can be met through ever-decreasing consumption taxes. For example, given the pure rate of social time preference advocated by Stern ($\beta_G = 0.999$) versus the private rate adopted by Nordhaus ($\beta_H = 0.985$),[14] condition (29) implies that the real after-tax price of consumption should decrease by 1.4% per year.

Similarly, for preferences (B), comparison of the household's Euler Equation (28) with the planner's optimality condition shows that the optimal consumption tax sequence must satisfy:

$$\frac{(1+\tau^*_{ct+1})}{(1+\tau^*_{ct})} = \frac{\left[\left(\frac{\beta_G}{\beta_H}\right)^t + \phi[1-\sigma-\gamma(1-\sigma)]\right]}{\left[\left(\frac{\beta_G}{\beta_H}\right)^{t+1} + \phi[1-\sigma-\gamma(1-\sigma)]\right]} \tag{30}$$

If the government discounts the future less than households ($\beta_G > \beta_H$), then the denominator on the right-hand side of (30) is larger than the numerator. Consequently, the optimal consumption tax sequence must be decreasing over time. Proposition 3 summarizes these results:

Proposition 3 *If the social planner values the future more than households ($\beta_G > \beta_H$), and if preferences are of the form (A) or (B), then the optimal consumption tax is decreasing over time for all $t > 1$.*

In contrast, with standard discounting ($\beta_G > \beta_H$), it is clear from (29) and (30) that the optimal consumption tax (for $t > 0$) is constant over time. Intuitively, the standard case can be understood both through the classic uniform commodity taxation result applied to consumption over time (see, e.g., Chari and Kehoe, 1999), and in terms of the equally classic result that intertemporal distortions are undesirable in a wide range of settings (see, e.g., Judd, 1985, 1999; Atkeson, Chari, and Kehoe, 1999; Acemoglu, Golosov, and Tsyvinski, 2011). Social discounting thus overturns both well-known public finance recommendations.

4.3 Labor Income Taxes

This section shows that social discounting can imply that labor income taxes should be decreasing over time, in contrast with the standard policy prescription that tax distortions should be smoothed. First, the planner's FOCs imply the following social optimality condition for the

[14] It should be noted that Nordhaus (e.g., 2008) calibrates $\beta_H = 0.985$ jointly with an inverse intertemporal elasticity of substitution $\sigma = 1.5$ to match observed rates of return as per the Ramsey equation. For logarithmic preferences ($\sigma = 1$) as employed by Stern (2006), the descriptive approach would thus in principle require an even lower private discount factor (higher pure rate of social time preference).

consumption-labor tradeoff for $t > 0$:

$$\frac{(-U_{lt}/U_{ct})}{F_{Lt}} = 1 + \phi \left[\frac{W_{ct}}{U_{ct}} + \frac{W_{lt}}{U_{ct} F_{Lt}} \right] \left(\frac{\beta_H}{\beta_G} \right)^t \tag{31}$$

In the decentralized economy, going back to the case where consumption is the untaxed numeraire, the household's intratemporal optimality condition is given by:

$$\frac{-U_{lt}/U_{ct}}{F_{Lt}} = (1 - \tau_{lt}) \tag{32}$$

Comparing (31) and (32) reveals that the optimal allocation can be decentralized by a labor tax implicitly defined by:

$$\tau_{lt}^* = (-\phi) \left[\frac{W_{ct}}{U_{ct}} + \frac{W_{lt}}{U_{ct} F_{Lt}} \right] \left(\frac{\beta_H}{\beta_G} \right)^t \tag{33}$$

For a given allocation, differential discounting with $\beta_G > \beta_H$ would thus imply decreasing labor income taxes compared to the standard case where $\beta_G = \beta_H$. However, as the optimal allocation depends on β_G, one must again consider specific utility functions to formally assess the implications of social discounting. Consider the following functional form for $v(L_t)$ in (A):

$$U(C_t, L_t, S_t) = \log C_t - v_1 L_t^{v_2} + \vartheta(S_t) \tag{A'}$$

It is easy to show that (A') is increasing and concave in leisure as long as $v_1 > 0$ and $v_2 > 1$. If preferences are of the form (A'), $W_{lt} = U_{lt} v_2$ and $W_{ct} = 0$. Substituting into the planner's optimality condition (31), rearranging, and comparing with the household's labor supply condition (32) demonstrates that the optimal labor income tax (for $t > 0$) is implicitly defined by:

$$\tau_{lt}^* = 1 - \left[1 + \phi v_2 \left(\frac{\beta_H}{\beta_G} \right)^t \right]^{-1} \tag{34}$$

With standard discounting ($\beta_G = \beta_H$), it is clear from (34) that the optimal labor income tax would be constant over time, and depends only on leisure preference parameter v_2 and the tightness with which the implementability constraint binds ϕ. However, if the planner is more patient than households ($\beta_G > \beta_H$), this tax smoothing result is overturned:

Proposition 4 *If the social planner values the future more than households ($\beta_G > \beta_H$), and if preferences are of the form (A'), then optimal labor income taxes are decreasing over time for $t > 1$.*

Proof: This result follows directly from the optimal labor tax expression (34) given the assumption that $v_2 > 1$ and the fact that the Lagrange multiplier on (IMP) is $\phi > 0$.

For preferences that are non-separable in consumption and leisure, the quantitative results in Section 5 suggest that optimal labor income taxes are also decreasing for the specification:[15]

$$U(C_t, L_t, S_t) = \frac{(C_t(1-L_t)^v)^{1-\sigma}}{1-\sigma} + \vartheta(S_t) \tag{B'}$$

Analytically, however, it is difficult to evaluate the rate of change in τ_{lt}^* for (B') as the optimal tax depends on the optimal labor allocation *at time t*, which is itself endogenous. Overall, however, both Proposition 4 and the quantitative results demonstrate that differential social discounting can overturn the classic policy prescription that tax distortions should be smoothed across time.

4.4 Distortionary Fiscal Setting Theory Results Summary

The central theoretical finding is that the adoption of different social than private discount rates fundamentally alters optimal policy prescriptions in the Ramsey optimal taxation framework. In the benchmark setting with descriptive discounting - where the planner adopts the household's discount rate ($\beta_G = \beta_H$) - the optimal allocation can be decentralized by zero capital income taxes, constant labor taxes, and/or constant consumption taxes. In contrast, if the planner decides to value the future more than households ($\beta_G > \beta_H$), then for some commonly used utility specifications I show that the optimal tax system features capital income subsidies, labor income taxes that are decreasing over time, and/or consumption taxes that are decreasing over time. To the best of my knowledge, these implications of differential discounting have not been previously formalized in this setting.

5 Quantitative Analysis

5.1 Calibration

This section assesses the quantitative importance of differential social discounting by extending two dynamic general equilibrium climate-economy model calibrations of the presented framework:

GHKT-B: First, in order to match the benchmark theoretical model, I build on the generalized numerical implementation of Golosov, Hassler, Krusell, and Tsyvinski's model (2014) (GHKT) by Barrage (2014), which allows for general CRRA preferences ($U(C_t) = \frac{C_t^{1-\sigma}-1}{1-\sigma}$). Here,

[15] I now focus on specification (B') as opposed to (B) since the latter has some undesirable implications with regards to labor income taxes. In particular, with preferences (B) the optimal labor income tax would be zero for $t > 0$, implying that all revenues should be raised through labor income taxes at $t = 0$ and period $t = 1$ capital income taxes. While a similar result has been discussed, e.g., by Jones, Manuelli, and Rossi (1997) in a setting with human capital accumulation and with regards to long-run taxes, here this stark result is simply driven by formulation (B). Consequently, I focus on the more common (B'), which has been used in classic studies such as Kydland and Prescott (1992) or Christiano and Eichenbaum (1992).

I introduce two new elements. First, I allow for differential discounting between households and the planner. Second, I introduce a nested CES production structure to consider capital-energy elasticities of substitution below unity. In contrast, the standard GHKT framework focuses on Cobb-Douglas technology. In particular, based on a review of the literature by van der Werf (2008), I consider the following structure:

$$Y_t = (1 - D(S_t)) \cdot A_t \left[\varkappa \left[K_t^{\alpha} L_t^{1-\alpha} \right]^{\frac{\varepsilon-1}{\varepsilon}} + (1-\varkappa) \left[E_t \right]^{\frac{\varepsilon-1}{\varepsilon}} \right]^{\frac{\varepsilon}{\varepsilon-1}}$$

where ε denotes the elasticity of substitution between energy inputs E_t and the capital-labor aggregate. In addition to these changes, I also simplify the energy sector representation compared to GHKT. Appendix C describes these changes and calibration details.

COMET: Second, in order to quantify optimal policies in the extended theoretical framework with endogenous labor supply and distortionary taxes, I integrate differential discounting into the COMET model (Barrage, 2015). The COMET builds upon the model structure presented above and on the seminal DICE framework of Nordhaus (see, e.g., 2008, 2010), which is one of three models currently used by the U.S. Government to value the social cost of carbon (Interagency Working Group, 2010). The COMET expands upon DICE in several ways, including preferences for leisure and the climate, energy production, tax policy choice, and government spending requirements (see Barrage, 2015). In particular, household utility in the COMET mirrors specification (B'):

$$U(C_t, L_t, T_t) = \left\{ \frac{\left[C_t \cdot (1 - \varsigma L_t)^{v} \right]^{1-\sigma}}{1-\sigma} \right\} + \frac{\left(1 + \alpha_0 T_t^2\right)^{-(1-\sigma)}}{1-\sigma} \tag{35}$$

The additional parameter ς is introduced in order to ensure that the calibration can simultaneously match (i) a desired intertemporal elasticity of substitution ($\sigma = 1.5$), (ii) a Frisch elasticity of labor supply of 0.78, and (iii) and to rationalize base year (2005) labor supply as estimated from OECD data (see Barrage (2015) for details). The variable T_t denotes mean atmospheric surface temperature change over pre-industrial levels.[16] Government revenue requirements and expenditure patterns are calibrated based on IMF Government Finance Statistics.[17] On the macroeconomic side, the COMET further adopts the productivity, clean energy technology, and population growth rate projections of the DICE/RICE model family (Nordhaus, 2008, 2010).

[16] The DICE and thus COMET damages are calibrated to global temperature change T_t, whereas GHKT damages are a function of carbon concentrations S_t.

[17] In the model base year 2005, the PPP-adjusted GDP-weighted average share of government expenditures is $\sim 33\%$ of GDP. Expenditures are further broken down into government consumption ($G_t^C \sim 57\%$) and social transfers ($G_t^T \sim 43\%$). In line with other fiscal computable general equilibrium and optimization models (e.g., Jones, Manuelli, and Rossi, 1993; Goulder, 1995), government expenditure G_t grows at the rates of labor productivity and population growth, with G_t^C and G_t^T evolving at constant shares proportional to G_t.

5.2 Optimal Policy Results

5.2.1 First-Best Benchmark (GHKT-B)

Figure 1 displays first-best carbon taxes across different values of the social discount rate and households' intertemporal elasticity of substitution. As expected, climate policy is significantly more stringent with prescriptive social discounting (dashed lines) than when the planner adopts households' pure rate of social time preference. Given the way capital income taxes are defined in the benchmark, they are constant and equal to $\tau_k^* = -15.16\%$ for $\beta_G = 0.999$ (as advocated by Stern (2006)) and $\beta_H = 0.985$ (as adopted by Nordhaus (e.g., 2008)).

5.2.2 Constrained Environmental Planner (GHKT-B)

Next, I solve for optimal carbon prices from the perspective of a "Sternian environmental planner" who cannot subsidize savings. The theoretical results indicated that the constrained-optimal carbon tax can be above or below the social cost of carbon due to the countervailing effects of climate policy on household savings. The central *quantitative* result is that the net effect is negative for all parameters considered: the constrained-optimal carbon tax is 5-50% *below* the first-best. Figure 2 displays first-best and constrained-optimal carbon prices for a range of parameter values. The downward adjustment in carbon prices is larger, the higher capital income taxes are constrained to be. As predicted by the theoretical results, the constrained policy is also adjusted downward more strongly when capital and energy are more complementary in production, as this implies a larger negative effect of carbon taxes on the returns to saving (by making energy inputs more expensive).

For example, if energy and the labor-capital aggregate are complements ($\varepsilon_{KL,E} = 0.4$, $\sigma = 1.5$, and $\beta_G = 0.999$), the planner would ideally like to impose a high carbon tax and subsidize capital income so as to achieve an aggregate savings rate of 26.34% (on average in the 21st Century). However, if capital income taxes are fixed at 30% and the planner nonetheless imposes first-best Sternian carbon taxes, the aggregate savings rate falls to 20.01%. By adjusting carbon taxes downward to the constrained-optimal level, the planner can increase the savings rate slightly to 20.05%. The constrained-optimal policy thus strikes a balance between incentivizing investments in environmental and man-made assets for the overall benefit of future generations.

5.2.3 Distortionary Fiscal Setting (COMET)

This section evaluates the quantitative implications of social discounting in a richer fiscal setting. While the theoretical results reveal that capital income subsidies remain desirable even when they must be financed with distortionary taxes, the empirical magnitude of this and other tax policy deviations from standard recommendations remains an open question. Figures 4 and 5 display optimal carbon taxes and temperature change across social discount rates. As expected, prescriptive discounting at rates advocated by authors such as Stern (2006) and Cline (1992) reduces optimal peak global temperature change to below $2°C$ (specifically $1.93C°$). In contrast, with descriptive discounting ($\beta_G = \beta_H = 0.985$), the optimal policy limits climate change to around $3°C$, in line with the DICE/RICE model family (Nordhaus, 2008, 2010, etc.).[18]

While the climate policy implications of social discounting are well-known, Figures 6 and 7 showcase the broader fiscal policy adjustments that this approach also requires. I find that these adjustments are large for the discount factors advocated in the literature: For labor taxes, social discounting implies that labor should first be taxed at a high rate of 53%, followed by a continual decline to 36% by the end of the century. Decentralizing the optimal allocation further requires capital income subsidies ranging from 30% to 65% by the end of the century (dotted line with stars). In contrast, with descriptive discounting ($\beta_G = \beta_H = 0.985$), we obtain the standard results that labor taxes should be constant at 41%, and that capital income should not be taxed after the first period (solid line with asterisks). In reality, many countries tax capital income at high rates. Social discounting thus implies a fundamental departure from both current policy practice and standard policy recommendations in the Ramsey framework.

5.2.4 Welfare (GHKT-B)

The final quantitative exercise employs the GHKT-B model to compare the welfare gains of optimizing carbon versus capital income taxes from the perspective of a social planner with Stern preferences ($\beta_G = .999$). I specifically compare the following scenarios:

1. *Descriptive Discounting Policies*: No capital income subsidies and carbon prices that would be optimal with $\beta_G = \beta_H = 0.985$ (as in Nordhaus, 2008).

2. *First-Best Social Discounting Policies:* Optimized carbon taxes and capital income subsidies for $\beta_G = 0.999$ and $\beta_H = 0.985$.

3. *Constrained-Optimal Climate Policy:* No capital income subsidies but socially discounted ($\beta_G = 0.999$) constrained-optimal carbon prices.

4. *Constrained-Optimal Savings Policy:* "Nordhaus" carbon prices that would be optimal with $\beta_G = \beta_H = 0.985$, but constrained-optimal capital income subsidies.

Table 1 provides the results in terms of the permanent consumption change equivalent to the policy. Perhaps surprisingly, the relative importance of being able to optimize carbon versus capital

[18] These figures all assume an inverse intertemporal elasticity of substitution of $\sigma = 1.5$.

income taxes is ambiguous, and depends critically on the intertemporal elasticity of substitution. A "Sternian" social planner with logarithmic preferences ($\sigma = 1$) would indeed achieve significantly larger welfare gains from setting climate policy instead of fiscal policy. However, capital income subsidies become relatively more important for $\sigma \geq 1.5$. While these calculations abstract from, e.g., downside climate risks through non-convexities in the damage function, and should thus perhaps be taken with a grain of salt, they do illustrate the core point that the climate is only one of many assets that current generations can invest in to benefit future generations. A government that weighs the overall well-being of future generations more highly than its citizens should thus incentivize larger investments in all of those assets according to their rates of return (e.g., health, education, private capital), rather than treating the climate in isolation.

6 Conclusion

The rate at which policy-makers should discount the future has long been one of the most intensely debated questions in public economics. This issue is moreover central to current debates on climate policy, where a fundamental disagreement has emerged between those calling for an ethics-based calibration of the pure rate of social time preference as close to zero (e.g., Stern, 2006), and those advocating that discounting parameters should be calibrated to match households' revealed time preferences, market interest rates, and the opportunity cost of capital (e.g., Nordhaus, 2007). In order to reconcile these perspectives, a number of authors have proposed a differentiated approach that allows the planner to select a different utility discount rate than exhibited by households. While these authors have generally cautioned that this approach would likely have policy implications beyond the climate realm (e.g., Goulder and Williams, 2012), the nature and magnitude of these effects remains an open question.

This paper formalizes the broader policy implications of differential discounting in a dynamic general equilibrium climate-economy model with an extension to fiscal policy in the Ramsey tradition. I empirically quantify the magnitude of the required policy changes and their welfare effects by building on the integrated assessment models of both Golosov, Hassler, Krusell and Tsyvinski's (2014), and an extended version of the seminal DICE model (Nordhaus, 2008) that incorporates fiscal policy and endogenizes key variables such as labor supply (the COMET by Barrage, 2015).

The main result is that a policy-maker's choice to adopt differential social discounting fundamentally alters prescriptions for both environmental and fiscal policy. On the one hand, decentralizing the optimal allocation would require radical changes to tax policy, including a transition to substantial capital income subsidies (~15-65%), and, depending on the fiscal setting, a change to labor income and/or consumption taxes that are decreasing over time. On the other hand, in a more realistic policy setting where the environmental planner cannot subsidize capital income, the constrained-optimal carbon tax may no longer equal the social cost of carbon (the Pigouvian tax). The standard policy prescription to price externalities at the Pigouvian rate must be adjusted for the general equilibrium effects of climate policy on households' incentives to save. These effects depend on the precise nature of climate damages and on the complementarity between energy and capital in production, highlighting the value of future research on both of these parameters. Quantitatively, I find that the constrained-optimal carbon tax is up to 50% below the first-best levy, indicating that even a planner with "Sternian" social preferences may not want to impose a "Sternian" carbon tax due to its potential effects on the overall level of assets that will be provided to future generations. The welfare calculations similarly suggest that, for certain areas of the parameter space, subsidies for private investment may provide a better way of redistributing to the future than tighter climate policy.

One important objection to these results is that they fail to capture the potentially limited substitutability between the climate and man-made assets and goods, which can increase the optimal carbon price as much as Sternian discounting (Sterner and Persson, 2008; Traeger, 2011). Formally accounting for this limited substitutability would moreover not lead to the surprising broader policy implications of social discounting outlined in this paper.

In general it is essential to note that the results of this study rely on a disconnect between the planner's and the household's utility discount factors based on an ethical disagreement. Alternative micro-foundations may well lead to different policy implications. For example, aggregation of heterogeneous rates of time preference could yield a declining representative discount rate while maintaining decentralized real interest rates (Heal and Millner, 2013). As another example, Gerlagh and Liski (2014) show that governments may value climate investments differently than households if delays in the climate system create a commitment value for carbon versus capital investments. On the other hand, many have argued for a disconnect between planner and household discounting on behavioral grounds. The policy implications in this case are likely to be similar. For example, Allcott, Mullainathan, and Taubinsky (2012) show that purchases of energy efficient durables should be subsidized if households discount the returns on these investments excessively.

Overall, the results of this paper thus demonstrate that differential discounting has far-reaching implications for both fiscal and environmental policy. While some of the results - such as the desirability of savings subsidies - align with new research on optimal estate taxation and intergenerational insurance (Farhi and Werning, 2005, 2010), other results - such as declining labor taxes - may be considered troubling, and represent a stark departure from conventional wisdom on optimal taxation. How governments should discount the future, and what the corresponding policy implications are, thus remains an open and essential area of research.

7 Figures and Tables

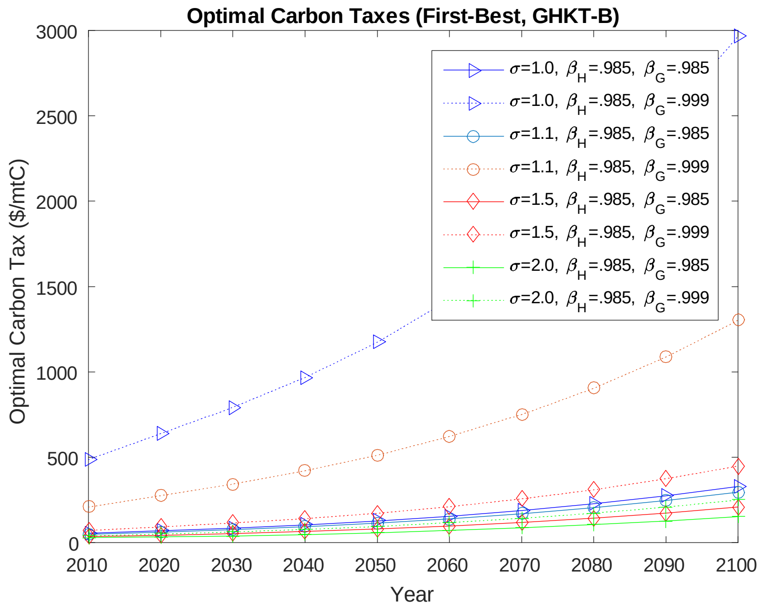

Figure 1: First-best carbon taxes with differential discounting ($\beta_G > \beta_H$) across values of the inverse intertemporal elasticity of substitution (σ).

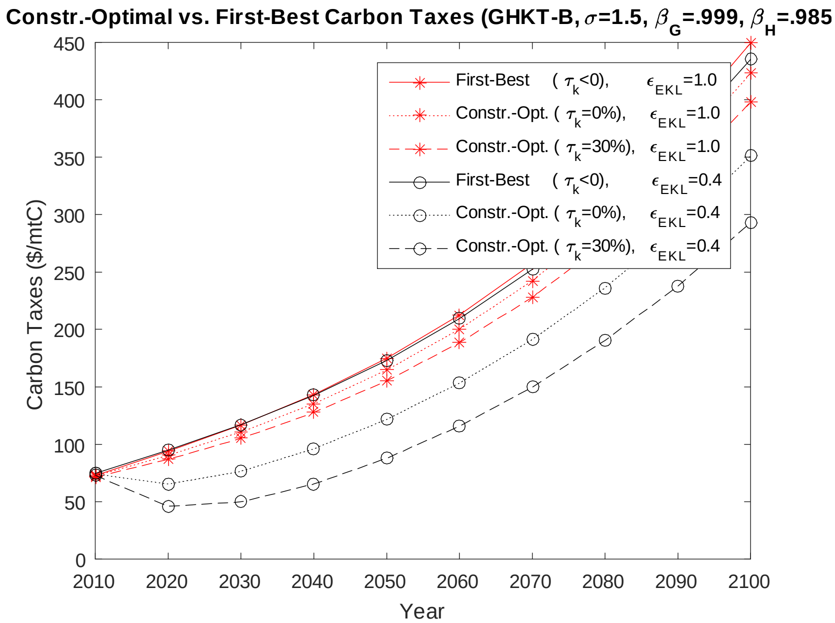

Figure 2: First-best and constrained-optimal carbon taxes for optimized ($\tau_k = \tau_k^* < 0$) and constrained ($\tau_k \geq 0\%$, $\tau_k \geq 30\%$) capital income taxes, respectively, across different values of the elasticity of substitution between energy and the capital-labor aggregate in final goods production ($\varepsilon_{E,KL}$).

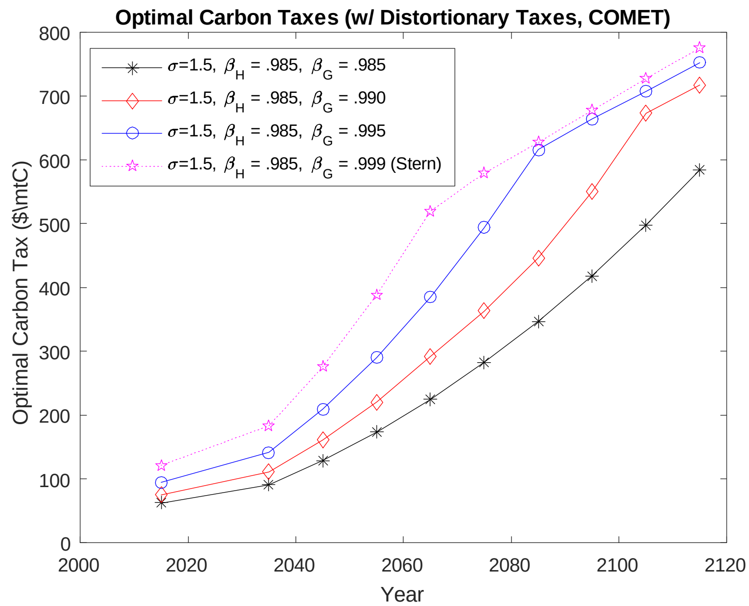

Figure 3: Optimal carbon taxes in the distortionary fiscal setting across different values of the social utility discount factor (β_G).

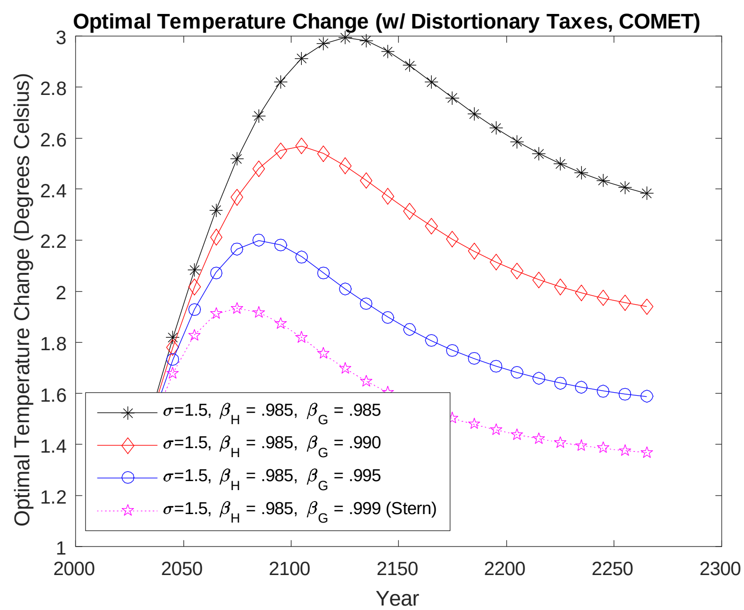

Figure 4: Optimal temperature change in the distortionary fiscal setting across different values of the social utility discount factor (β_G).

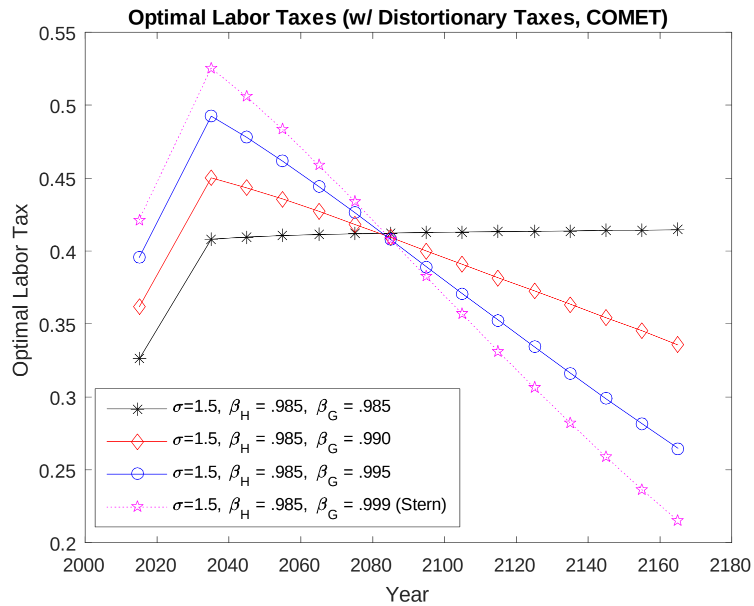

Figure 5: Optimal labor taxes in the distortionary fiscal setting across different values of the social utility discount factor (β_G).

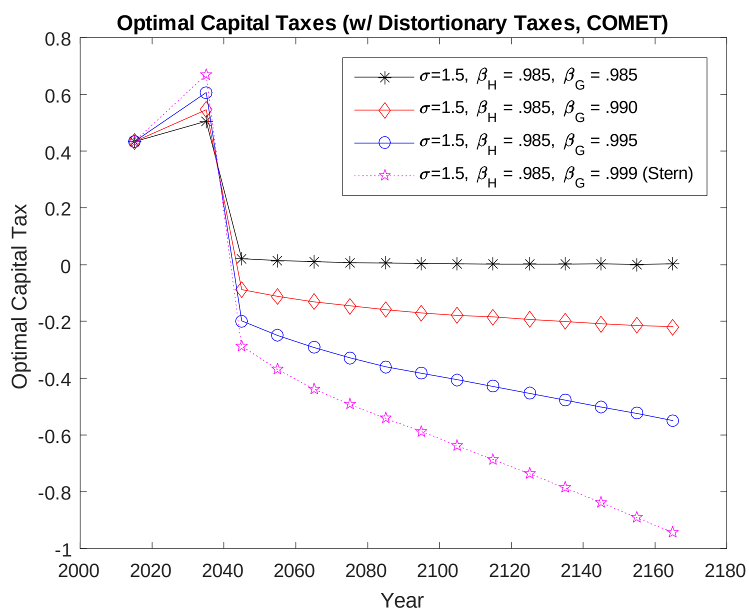

Figure 6: Optimal capital taxes in the distortionary fiscal setting across different values of the social utility discount factor (β_G).

Table 1: Welfare Effects from the Perspective of a Stern ($\beta_G = .999$) Planner					
Policy Scenario:		ΔWelfare‡ ($2010 tril.)			
Carbon Taxes:	Capital Taxes:	$\sigma^{\dagger}= 1$	$\sigma = 1.1$	$\sigma = 1.5$	$\sigma = 2$
1. Standard Discounting (τ_{Et}^{DICE})	None	-	-	-	-
2. Stern Discounting (τ_{Et}^{Stern})	Optimized ($\tau_{kt+1}^{Stern} < 0$)	2.45%	1.80%	0.63%	0.39%
3. Stern Discounting (τ_{Et}^{Stern})	None	1.74%	1.26%	0.17%	0.04%
4. Standard Discounting (τ_{Et}^{DICE})	Optimized ($\tau_{kt+1}^{Stern} < 0$)	0.27%	0.46%	0.44%	0.38%
†Equivalent variation permanent percentage change in consumption. Table displays welfare changes relative to Case 1 (no capital subsidy, standard discounting carbon tax) computed in the GHKT-B model.					

References

[1] Acemoglu, Daron, Mikhail Golosov, and Aleh Tsyvinski (2011) "Political economy of Ramsey taxation" *Journal of Public Economics* 95(7): 467-475.

[2] Allcott, Hunt, Sendhil Mullainathan, and Dmitry Taubinsky (2014) "Energy Policy with Externalities and Internalities" *Journal of Public Economics*, 112: 72–88.

[3] Allcott, Hunt, and Michael Greenstone (2012) "Is There an Energy Efficiency Gap?" *Journal of Economic Perspectives,* 26(1): 3-28.

[4] Arrow, Kenneth J., W.R. Cline, K-G. Maler, M. Munasinghe, R. Squitieri, J.E. Stiglitz (1995) "Intertemporal Equity, Discounting, and Economic Efficiency" in: Intergovernmental Panel on Climate Change (IPCC) Second Assessment Report, Working Group II, 1995.

[5] Arrow, Kenneth J., Maureen L. Cropper, Christian Gollier, Ben Groom, Geoffrey M. Heal, Richard G. Newell, William D. Nordhaus, Robert S. Pindyck, William A. Pizer, Paul R. Portney, Thomas Sterner, Richard S. J. Tol and Martin L. Weitzman (2012) "How Should Benefits and Costs Be Discounted in an Intergenerational Context?" Resources for the Future Discussion Paper 12-42.

[6] Asheim, Geir B., and Frikk Nesje (2016) "Destructive intergenerational altruism," *Journal of the Association of Environmental and Resource Economists* (forthcoming).

[7] Atkeson, Andrew, V.V. Chari, and Patrick J. Kehoe (1999) "Taxing Capital Income: A Bad Idea" *Federal Reserve Bank of Minneapolis Quarterly Review*, 23(3): 3-17.

[8] Barrage, Lint (2015) "Optimal Dynamic Carbon Taxes in a Climate-Economy Model with Distortionary Fiscal Policy" Working Paper.

[9] Barrage, Lint (2014) “Sensitivity Analysis for Golosov, Hassler, Krusell, and Tsyvinski (2013): ‘Optimal Taxes on Fossil Fuel in General Equilibrium”’ Supplementary Material, Econometrica, 82(1): 41-88. [Supplemental Material to: Golosov, Mikhail, John Hassler, Per Krusell, and Aleh Tsyvinski (2014) “Optimal Taxes on Fossil Fuel in General Equilibrium,” Econometrica, 82(1)]

[10] Baumol, William J. (1968) "On the Social Rate of Discount" *American Economic Review*, 58(4): 788-802.

[11] Belfiori, Maria Elisa (2016) "Carbon Trading, Carbon Taxes and Social Discounting" *Working Paper.*

[12] Benhabib, Jess, and Aldo Rustichini (1997) "Optimal Taxes without Commitment" *Journal of Economic Theory* 77: 231-259.

[13] Bovenberg, A. Lans, and Lawrence H. Goulder (1996) "Optimal environmental taxation in the presence of other taxes: general-equilibrium analyses," *American Economic Review,* 86.4: 985-1000.

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

[15] Chari, Varadarajan V., and Patrick J. Kehoe (1999) "Optimal fiscal and monetary policy" *Handbook of Macroeconomics* 1: 1671-1745.

[16] Christiano, Lawrence J., and Martin Eichenbaum (1992) "Current real-business-cycle theories and aggregate labor-market fluctuations," *American Economic Review*: 430-450.

[17] Cline, William R. (1992). *The Economics of Global Warming*, Washington, DC: Institute for International Economics.

[18] Dasgupta, Partha (2008) "Discounting climate change" *Journal of Risk and Uncertainty,* 37.2-3: 141-169.

[19] Farhi, Emmanuel, and Ivan Werning (2005) "Inequality, Social Discounting, and Estate Taxation" *NBER WP 11408.*

[20] Farhi, Emmanuel, and Ivan Werning (2007) "Inequality and Social Discounting" *Journal of Politcal Economy*, 115(3): 365-402.

[21] Farhi, Emmanuel, and Ivan Werning (2010) "Progressive Estate Taxation" *Quarterly Journal of Economics*, 125(2): 635-673.

[22] Gerlagh, Reyer, and Matti Liski (2014) "Carbon prices for the next thousand years" *Working Paper.*

[23] Gollier, Christian, and Martin L. Weitzman (2010) "How should the distant future be discounted when discount rates are uncertain?" *Economics Letters,* 107.3: 350-353.

[24] Gollier, Christian, and Richard Zeckhauser (2005) "Aggregation of heterogeneous time preferences." *Journal of Political Economy* 113, no. 4: 878-896.

[25] Golosov, Mikhail, John Hassler, Per Krusell, and Aleh Tsyvinski (2014) "Optimal taxes on fossil fuel in general equilibrium" *Econometrica*, 82(1): 41-88.

[26] Goulder, Lawrence H. (1995) "Environmental Taxation and the Double Dividend: A Reader's Guide" *International Taw and Public Finance*, 2: 157-183.

[27] Goulder, Lawrence H., and Roberton Williams III (2012) "The choice of discount rate for climate change policy evaluation" *Climate Change Economics* 3(4): 1250024-1-1250024-18.

[28] Heal, Geoffrey, and Antony Millner (2013) "Discounting under disagreement" NBER WP No. w18999.

[29] Interagency Working Group on Social Cost of Carbon (2010) "Social Cost of Carbon for Regulatory Impact Analysis under Executive Order 12866" United States Government.

[30] International Energy Agency (2012) "2012 CO_2 Emissions Overview," available at [last accessed June 2016]: https://www.iea.org/media/statistics/topics/emissions/CO2_Emissions_Overview.pdf

[31] Jones, Larry E., Rodolfo E. Manuelli, and Peter E. Rossi (1993) "Optimal Taxation in Models of Endogenous Growth" *Journal of Political Economy*, 101(3): 485-517.

[32] Jones, Larry E., Rodolfo E. Manuelli, and Peter E. Rossi (1997) "On the optimal taxation of capital income" *Journal of Economic Theory,* 73(1): 93-117.

[33] Judd, Kenneth L. (1985) "Redistributive taxation in a simple perfect foresight model" *Journal of Public Economics,* 28:59-83.

[34] Judd, Kenneth L. (1999) "Optimal taxation and spending in general competitive growth models" *Journal of Public Economics,* 71: 1-26.

[35] Kaplow, Louis, Elisabeth Moyer, and David A. Weisbach (2010) "The social evaluation of intergenerational policies and its application to integrated assessment models of climate change." *The BE Journal of Economic Analysis & Policy* 10(2).

[36] Klein, Paul, and José-Víctor Ríos-Rull (2003) "Time-consistent optimal fiscal policy" *International Economic Review,* 44(4): 1217-1245.

[37] King, Robert G., Charles I. Plosser, and Sergio T. Rebelo (2001) "Production, growth and business cycles: Technical appendix." *Computational Economics* 20(1-2): 87-116.

[38] Kydland, Finn E., and Edward C. Prescott (1982) "Time to build and aggregate fluctuations," *Econometrica*, 50: 1345-1370.

[39] Lind, Robert C., Ed. (1982) *Discounting for Time and Risk in Energy Policy*, Resources for the Future, Washington, DC.

[40] Manne, Alan S. (1995) "The Rate of Time Preference - Implications for the Greenhouse Debate" Energy Policy, 23(4/5): 391-394.

[41] Newell, Richard G., and William A. Pizer (2003) "Discounting the distant future: how much do uncertain rates increase valuations?" *Journal of Environmental Economics and Management,* 46.1: 52-71.

[42] Nordhaus, William D. (2007) "A Review of the" Stern Review on the Economics of Climate Change" *Journal of Economic Literature*, Vol. XLV: 686-702.

[43] Nordhaus, William D. (2008) *A Question of Balance*, Yale University Press, New Haven, CT.

[44] Nordhaus, William D. (2010) "Economic aspects of global warming in a post-Copenhagen environment" *Proceedings of the National Academy of Sciences*, 107(26): 11721-11726.

[45] Nordhaus, William D. (2010) *DICE-2010 model*, available at http://nordhaus.econ.yale.edu/RICEmodels.htm

[46] Nordhaus, William D., and Boyer, Jospeh (2000) *Warming the World: Economic Models of Global Warming*, MIT Press, Cambridge, MA.

[47] Ramsey, Frank P. (1928) "A Mathematical Theory of Saving" *The Economic Journal*, 38(152): 543-559.

[48] Ryan, Stephen P. (2012) "The costs of environmental regulation in a concentrated industry," *Econometrica,* 80.3: 1019-1061.

[49] Schmitt, Alex (2014) "Optimal Carbon and Income Taxation" Working Paper.

[50] Sen, Amartya K. (1982) "Approaches to the Choice of Discount Rates for Social Benefit-Cost Analysis" in: Lind, Robert C. (editor) *Discounting for Time and Risk in Energy Policy*, Resources for the Future, Washington, DC.

[51] Sleet, Christopher, and Şevin Yeltekin (2006) "Credibility and endogenous societal discounting" *Review of Economic Dynamics,* 9(3): 410-437.

[52] Stern, Nicholas H. (2006) *Stern Review: The economics of climate change*. Vol. 30. London: HM Treasury, 2006.

[53] Sterner, Thomas, and U. Martin Persson (2008) "An even sterner review: Introducing relative prices into the discounting debate," *Review of Environmental Economics and Policy,* 2.1: 61-76.

[54] Traeger, Christian (2011), Sustainability, Limited Substitutability and Non-constant Social Discount Rates" *Journal of Environmental Economics and Management,* 62: 215–228.

[55] Van der Werf, Edwin (2008) "Production functions for climate policy modeling: An empirical analysis," *Energy Economics,* 30, no. 6: 2964-2979.

[56] von Below, David (2012) "Optimal carbon taxes with social and private discounting." SURED Conference, Ascona, Switzerland.

[57] Williams III, Roberton C. (2002) "Environmental Tax Interactions When Pollution Affects Health Or Productivity," *Journal of Environmental Economics and Management*, 44: 261-270.

8 Appendix A: Benchmark Model

8.1 First-Best Policies: Result 1

The planner's problem in the first-best setting is as follows:

$$\begin{aligned}
&\max \sum_{t=0}^{\infty} \beta_G^t [U(C_t, S_t)] \\
&+\sum_{t=0}^{\infty} \beta_G^t \lambda_t \left[F(A_t(S_t); K_t, L_t^Y, E_t) + (1-\delta)K_t - C_t - K_{t+1}\right] \\
&+\sum_{t=0}^{\infty} \beta_G^t \xi_t [S_t - \widetilde{S}_t\,(S_0, E_{-T}, E_{-T+1}, ..., E_t)] \\
&+\sum_{t=0}^{\infty} \beta_G^t \chi_t \left[G_t(L_t^E) - E_t\right] \\
&+\sum_{t=0}^{\infty} \beta_G^t \lambda_{lt} \left[L_t - L_t^Y - L_t^E\right]
\end{aligned} \tag{36}$$

where λ_t, ξ_t, χ_t, and λ_{lt} denote the Lagrange multipliers on the final goods resource constraint, the carbon concentrations constraint, the energy production constraint, and the labor resource constraint, respectively.

Decentralization: Carbon Taxes

In order to demonstrate that the proposed policies of Result 1 can decentralize the optimal allocation, consider first the competitive energy producer's profit-maximizing problem for a given excise emissions tax τ_{Et} at each time $t \geq 0$:

$$\begin{aligned}
& \max (p_{Et} - \tau_{Et})E_t - w_t L_t^E \\
\text{s.t. } E_t &= G_t(L_t^E)
\end{aligned}$$

The energy producer's optimality conditions imply that, in equilibrium,

$$p_{Et} - \tau_{Et} = \frac{w_t}{\frac{\partial G_t}{\partial L_t^E}} \tag{37}$$

Next, profit maximization by the final good producer implies that marginal products are equated to factor prices as per (5). Substituting these equilibrium prices into the energy pricing condition (37) yields:

$$\frac{\partial F_t}{\partial E_t} - \tau_{Et} = \frac{\frac{\partial F_t}{\partial L_t^Y}}{\frac{\partial G_t}{\partial L_t^E}} \tag{38}$$

Finally, note that, in the social planner's problem (36), the FOCs with respect to labor supplied to final goods production L_t^Y and energy production L_t^E imply that the private marginal cost of energy production at the optimum is given by

$$\frac{\chi_t}{U_{ct}} = \frac{\chi_t}{\lambda_t} = \frac{\frac{\partial F_t}{\partial L_t^Y}}{\frac{\partial G_t}{\partial L_t^E}} \tag{39}$$

Comparing the energy producer's optimality condition (37) with the planner's optimality condition for energy usage (10), it is thus immediately obvious that setting the carbon tax as noted in Result 1 makes the two optimality conditions coincide, as desired.

Decentralization: Capital Income Taxes

Consider the representative household's Euler Equation (3):

$$\beta_H \left[1 + (r_{t+1} - \delta)(1 - \tau_{kt+1})\right] = \frac{U_{ct}}{U_{ct+1}} \tag{40}$$

Noting that, in equilibrium, $r_{t+1} = F_{kt+1}$ as per the firm's optimality condition (5), it is then straightforward that setting $\tau^*_{Kt+1} = 1 - \frac{\beta_G}{\beta_H}$ makes the household's intertemporal optimality condition coincide with that of the social planner (11).

8.2 Constrained-Optimal Climate Policy: Proposition 1

Letting Φ_t denote the Lagrange multiplier on the no-subsidy constraint (15) added to the planner's problem (36), the constrained ('environmental') planner's optimality condition for energy usage E_t yields the following revised version of (10) defining the optimal carbon tax τ^*_{Et} :

$$\underbrace{F_{Et}}_{\substack{\text{Marginal}\\ \text{product of}\\ \text{energy}}} \underbrace{-\underbrace{\sum_{j=0}^{\infty} \beta_G^j \left[\frac{\partial S_{t+j}}{\partial E_t} \frac{\xi_{t+j}}{\lambda_t}\right]}_{\substack{\text{SDV of marginal}\\ \text{impacts from}\\ \text{carbon emissions}}} + \underbrace{\frac{1}{\beta_G} \frac{\Phi_{t-1}}{\lambda_t} \frac{\partial F_{Kt}}{\partial E_t}}_{\substack{\text{Social cost of energy use}\\ \text{investment incentive}\\ \text{change (in } C_t \text{ units)}}}}_{=-\tau^*_{Et}} = \underbrace{\frac{\chi_t}{\lambda_t}}_{\substack{\text{MC of}\\ \text{energy production}\\ \text{(in } C_t \text{ units)}}} \tag{41}$$

Optimal energy usage thus balances its benefits in final goods production (F_{Et}) minus the direct externality cost of the associated climate change ($\frac{1}{\lambda_t} \sum_{j=0}^{\infty} \beta_G^j \frac{\partial S_{t+j}}{\partial E_t} \xi_{t+j}$) and the social cost of the effect of energy input usage on the marginal product of capital and thus investment incentives ($\frac{1}{\beta_G} \frac{\Phi_{t-1}}{\lambda_t} \frac{\partial F_{Kt}}{\partial E_t}$), against the private marginal costs of producing energy ($\frac{\chi_t}{\lambda_t}$). In order to show that the *Energy Use Investment Incentives Adjustment* is negative if energy and capital are complements in production, first note that, if $\beta_G > \beta_H$, then the Lagrange multiplier on the no-subsidy constraint is binding $\Phi_t > 0$ due to the desirability of capital income subsidies demonstrated above. In addition, noting that $\lambda_t > 0$, we thus have that $\frac{\partial F_{Kt}}{\partial E_t} > 0$ implies that the *Energy Use Investment Incentives Adjustment* $-[\frac{1}{\beta_G}\frac{\Phi_{t-1}}{\lambda_t}\frac{\partial F_{Kt}}{\partial E_t}]$ is negative.

Next, the externality cost of carbon is defined by the FOC with respect to the stock S_t :

$$\underbrace{-\xi_t}_{\substack{\text{Shadow}\\ \text{value of carbon}\\ \text{concentrations}}} = \underbrace{U_{St} + \lambda_t F_{St}}_{\substack{\text{Disutility and}\\ \text{production}\\ \text{damages}}} + \underbrace{\frac{\Phi_{t-1}}{\beta_G}\frac{\partial F_{Kt}}{\partial S_t}}_{\substack{\text{Climate impact on}\\ \text{investment}\\ \text{incentive (MPK)}}} \tag{42}$$

Combining (42) with (41) and comparing with the energy producer's profit-maximizing condition, it follows that the climate investment incentives adjustment to the optimal carbon tax is defined by:

$$\begin{aligned}&\textit{Climate Investment Incentives Adjustment}\\ =\ &\sum_{j=0}^{\infty}\beta_g^j \underbrace{\frac{(\Phi_{t-1+j}/\beta_G)}{\lambda_t}}_{\substack{\text{Value of no-subsidy}\\ \text{constraint in}\\ \text{units of } C_t}} \cdot \underbrace{\frac{\partial S_{t+j}}{\partial E_t}\left(-\frac{\partial F_{Kt+j}}{\partial S_{t+j}}\right)}_{\substack{\text{Climate damages}\\ \text{to marginal product}\\ \text{of capital}}}\end{aligned}$$

In order to show that this term is positive if climate change decreases the marginal product of capital $\left(\frac{\partial F_{Kt+j}}{\partial S_{t+j}} < 0\right)$, first note that, if $\beta_G > \beta_H$, then $\Phi_t \geq 0$ (due to the desirability of a capital income subsidy as demonstrated above). Further noting that assumptions about preferences and technology imply that $\lambda_t > 0$ and that concentrations are increasing in past emissions $\left(\frac{\partial S_{t+j}}{\partial E_t} > 0\right)$ we thus have the desired result that a positive component is added to the optimal tax if climate change lowers the marginal return to capital investments (as the overall term $-\frac{\partial F_{Kt+j}}{\partial S_{t+j}}$ is then positive), ceteris paribus.

9 Appendix B: Distortionary Fiscal Setting

9.1 Capital Income Taxes: Proposition 2

First, the optimality of capital income subsidies when $\beta_G > \beta_H$ and preferences are of the form (A) follows directly from equation (25) derived in the text. Second, for preferences (B), it is easy to show that $W_{ct} = (U_{cct}C_t + U_{ct} + U_{lct}L_t) = U_{ct}[1-\sigma-\gamma(1-\sigma)]$. Substituting this expression into the planner's optimality condition for private capital (22) and defining the wedge term

$$\omega_t \equiv \frac{\left[\left(\frac{\beta_G}{\beta_H}\right)^t + \phi[1-\sigma-\gamma(1-\sigma)]\right]}{\left[\left(\frac{\beta_G}{\beta_H}\right)^{t+1} + \phi[1-\sigma-\gamma(1-\sigma)]\right]} \tag{43}$$

the planner's first order condition for capital becomes:

$$\beta_H \left[(1-\delta) + F_{Kt+1}\right] = \frac{U_{ct}}{U_{ct+1}}\omega_t \tag{44}$$

Next, consider the household's Euler equation for a given capital income tax:

$$\beta_H \left[1 + (F_{Kt+1} - \delta)(1 - \tau_{kt+1})\right] = \frac{U_{ct}}{U_{ct+1}}$$

Rearranging terms allows one to express the capital income tax that decentralizes a given allocation as:

$$\tau_{kt+1} = 1 - \frac{\frac{U_{ct}}{\beta_H U_{ct+1}}}{(F_{kt+1} - \delta)} + \frac{1}{(F_{kt+1} - \delta)} \tag{45}$$

Finally, rearrange terms in the planner's optimality condition (44) as follows:

$$\begin{aligned} [1 + F_{kt+1} - \delta]\frac{1}{\omega_t} &= \frac{U_{ct}}{\beta_H U_{ct+1}} \\ 1 - \frac{[1 + F_{kt+1} - \delta]\frac{1}{\omega_t}}{(F_{kt+1} - \delta)} + \frac{1}{(F_{kt+1} - \delta)} &= 1 - \frac{(U_{ct}/\beta_H U_{ct+1})}{(F_{kt+1} - \delta)} + \frac{1}{(F_{kt+1} - \delta)} \end{aligned} \tag{46}$$

Comparing (45) and (46), we obtain the desired result that the capital income tax that decentralizes the optimal allocation at $t+1$ for $t>0$ is defined by equation (26):

$$\begin{aligned} \tau^*_{kt+1} &= 1 - \frac{[1 + F_{kt+1} - \delta]\frac{1}{\omega_t}}{(F_{kt+1} - \delta)} + \frac{1}{(F_{kt+1} - \delta)} \\ &= \left(\frac{\omega_t - 1}{\omega_t}\right)\frac{(F_{kt+1} - \delta + 1)}{(F_{kt+1} - \delta)} \end{aligned} \tag{47}$$

The Online Appendix formally shows that, in order for utility specification (B) to be consistent with balanced growth when $\sigma > 1$, the leisure preference parameter must satisfy $\gamma > \frac{-\sigma}{(1-\sigma)}$. Since the Lagrange multiplier on (IMP), ϕ, is necessarily weakly positive, we have that:

$$\phi[1 - \sigma - \gamma(1-\sigma)] \geq 0 \tag{48}$$

Given (48), it immediately follows that the wedge term (43) is $\omega_t \in (0,1)$ when $\beta_G > \beta_H$. Consequently, it follows that the optimal capital income tax (47) is negative, completing the proof of Proposition 2.[19]

[19] The restriction of this result to $t+1>1$ stems from the fact that the planner's first-order conditions are non-stationary, implying a different optimal capital income tax rule for $t=0$ versus all subsequent periods. However, the optimality conditions are valid for all $t>0$, giving the desired result.

10 Appendix C: GHKT-B Recalibration

Energy Sector: The benchmark numerical framework uses a simplified energy sector representation compared to GHKT, who allow for multiple energy inputs and non-renewable resource dynamics. Whereas GHKT calibrate separate labor productivity parameters for coal and oil production, here we need a representative figure for both. In the benchmark scenario in the base decade 2010, the GHKT model yields a share of oil in total fossil energy consumption (all in tons of carbon-equivalent) of $(33.6/55.2) = 61\%$. Similarly, the IEA reports that oil and gas accounted for 55% of global CO_2 emissions in 2012 (IEA, 2012). GHKT compute model base year prices of oil/gas and coal of $\$606.5/mtC$ and $\$103.5/mtC$, respectively. The weighted average price is thus $\$405.3/mtC$. I thus calculate base year energy sector productivity as:

$$A_{E0}(\$/GtC) = \frac{w_0}{p_{E0}} \sim \frac{(1-\alpha-v)Y_0}{\$405.3 \cdot 10^9} = 1,140$$

CES Production

The GHKT framework focuses on Cobb-Douglas aggregate production, where the elasticity of substitution between energy and capital inputs is equal to unity. In order to illustrate the importance of energy-capital complementarity for the constrained-optimal carbon tax, I introduce an alternative nested CES production function:

$$Y_t = (1-D(S_t)) \cdot A_t \left[\varkappa \left[K_t^{\alpha} L_t^{1-\alpha}\right]^{\frac{\varepsilon-1}{\varepsilon}} + (1-\varkappa)\left[E_t\right]^{\frac{\varepsilon-1}{\varepsilon}}\right]^{\frac{\varepsilon}{\varepsilon-1}}$$

This specification is motivated by van der Werf (2008), who, in reviewing the literature as well as cross-country data, concludes that a (KL)E nesting specification generally provides the best fit for climate-economy models. In line with several other studies, I consider $\varepsilon_{KL,E} = 0.4$ as a benchmark case. Given base period output shares, the distribution parameter is set to $\varkappa = 0.96$. Base period labor supply is normalized to $L_0 = 1$, the initial net damage term can be calculated as $(1 - D(S_0)) = 0.9948$, and base period energy inputs are given from the data and the GHKT calibration at $E_0 = 34.9522\ GtC$ per decade. Note that I retain the initial capital stock measure K_0 from the benchmark model (which is calibrated such that the marginal product of capital in the Cobb-Douglas production framework matches the decadal net return to capital for an annual return equivalent of 5% plus 100% decadal depreciation). Finally, initial TFP A_0 can then be calibrated as:

$$A_0 = \frac{Y_0}{(1-D(S_0))\left[\varkappa \left[K_0^{\alpha} L_0^{1-\alpha}\right]^{\frac{\varepsilon-1}{\varepsilon}} + (1-\varkappa)\left[E_0\right]^{\frac{\varepsilon-1}{\varepsilon}}\right]^{\frac{\varepsilon}{\varepsilon-1}}}$$

Optimization

Given the high discount factor, I extend the model's direct optimization period to 50 periods (500 years), after which savings rates are locked in and iterated forward for 1,000 years for the climate to achieve a steady state, after which a balanced growth path is imposed to compute the continuation value of the allocations over an infinite time horizon. For further details on the computation, see GHKT (2014) and Barrage (2014).