Optimal Fiscal Policy in a Climate-Economy Model with Heterogeneous Households[*]

Thomas Douenne [†‡] Albert Jan Hummel [†§] Marcelo Pedroni [†¶]

September 9, 2022

Disclaimer: This document is a preliminary draft.

Link to most recent version

Abstract

This paper studies optimal fiscal policy in a climate-economy model with heterogeneous households. When individualized lump-sum taxation is not available, distortionary taxes on labor and capital income are levied to provide redistribution. In contrast to the representative-agent setting, the second-best pollution tax is then on average Pigouvian, with fiscal distortions driving only temporary deviations from Pigou. In a quantitative analysis where the climate model is calibrated to DICE and the fiscal system to the one of the U.S., we show that these temporary deviations are negligible, so that the optimal carbon tax is approximately equal to the social cost of carbon (SCC). Economic inequalities do not call for deviations from the Pigouvian principle, but they affect the Pigouvian rate itself: in our baseline experiment, the inequalities that remain after the planner sets income taxes optimally reduce the SCC by 4%. Optimal carbon taxation also leads to a more progressive tax system: contrary to the double-dividend hypothesis, with heterogeneous households it is optimal to use only half of the carbon tax revenue to reduce distortionary taxes, the rest being used to increase transfers.

JEL classification: E62, H21, H23, Q5

Keywords: Climate policy; Carbon taxes; Social cost of carbon; Optimal taxation; Double dividend; Heterogeneous agents; Redistribution.

---

[*]We would like to thank Reyer Gerlagh, Bas Jacobs, Mehrab Kiarsi, Rick van der Ploeg, Christian Stoltenberg, and Sweder van Wijnbergen and seminar participants at EAERE, EEA, IIPF, PET, SMS, and SURED conferences, and at TU Berlin, University of Amsterdam, University of Nantes, and University of Strasbourg for helpful comments and discussions.

[†]University of Amsterdam, Amsterdam School of Economics. Present address: University of Amsterdam, Roetersstraat 11, 1018 WB Amsterdam, Netherlands.

[‡]email: t.r.g.r.douenne@uva.nl

[§]email: a.j.hummel@uva.nl

[¶]email: m.pedroni@uva.nl

1 Introduction

Economic inequality and environmental degradation are certainly two of the most critical issues facing societies today. In order to address these two problems, economists have long argued for the use of fiscal instruments: labor and capital taxes can be used to provide redistribution, and following the Pigouvian principle a pollution tax can be used to internalize environmental externalities. However, pollution taxes also have distributional implications as they heterogeneously impact households' purchasing power. Conversely, capital and labor taxes also affect the costs and benefits of improving the environment by reducing incentives to work and invest. The goal of this study is to analyze how these instruments should be jointly optimized if society wishes to tackle both inequality and environmental degradation.

We address this question from both a theoretical and a quantitative perspective. To do so, this paper presents a dynamic fiscal climate-economy model with heterogeneous agents. We use a technique introduced by Werning (2007) to extend the climate-economy model of Barrage (2019) to heterogeneous agents. In our model, households derive utility from consumption, leisure, and environmental quality. The final consumption good is produced using energy as one of its inputs. Energy production is polluting, and pollution leads to environmental degradation that affects productivity and households' utility. As in Barrage (2019), energy producers can reduce the emission intensity of their output by engaging in costly abatement activities. Because of these costs, positive abatement will occur only if producers also need to pay for their pollution, for example through a pollution tax. The government thus faces multiple tasks at once: mitigating the pollution externality, providing redistribution, and financing some exogenous government spending.

We model this as a Ramsey problem in which the government chooses the level of linear taxes—in particular, taxes on labor and capital income, energy, and pollution—and a uniform lump-sum transfer to maximize aggregate welfare. Because agents are heterogeneous but tax instruments are anonymous, the government must rely on distortionary instruments to provide redistribution. We analytically characterize optimal tax formulas and study the implications of heterogeneity for optimal pollution taxation. We then use our model to examine how inequalities and distortionary taxation affect the social cost of carbon (SCC) and the optimal carbon tax. We calibrate our climate model following DICE 2016 (Nordhaus, 2017). On the economic side, we calibrate the fiscal system and household heterogeneity (first in productivity, later in wealth and preferences for energy consumption) to match U.S. data. Conceptually, our quantitative analysis examines the optimal fiscal policy of the U.S. if they accounted for the negative global impact of their emissions.[1]

Theoretically, we find that the optimal pollution tax is a modified Pigouvian rule that accounts for tax distortions via the marginal cost of public funds (MCF). However, because uniform lump-sum taxation is available, the MCF is no longer higher than 1 as in representative-agent settings (see for instance Bovenberg and Goulder, 1996; Barrage, 2019). In fact, we show that when households have balanced-growth preferences, the MCF is on average equal to 1, so the optimal pollution tax may only temporarily lie above or below the Pigouvian level. These temporary tax distortions are driven by the

[1]Specifically, we consider the problem of the U.S. government with its emissions scaled up to the global level. An equivalent interpretation is that the world consists of a number of U.S. economies coordinating on their climate policies.

costs the planner faces when implementing its preferred allocation. While these costs are on average null in the presence of lump-sum taxation, they may not be in each period. We provide conditions under which these costs are always null, and discuss the determinants of temporal variations in tax distortions when they are not. Our theoretical results also highlight the role of consumption inequalities. When the MCF is equal to unity, the second-best pollution tax is Pigouvian, but the Pigouvian tax is evaluated at the second-best allocation. We show that consumption inequality affects the Pigouvian tax ambiguously through the opportunity cost of emission abatement. On the one hand, consumption is valued less in the presence of inequalities because it disproportionately goes to richer households with lower marginal utilities of consumption. On the other hand, consumption inequalities increase the average marginal utility of consumption, and thus the opportunity cost of abatement. We show that with balanced-growth preferences, the latter effect dominates if and only if the intertemporal elasticity of substitution is lower than 1, in which case inequalities reduce the value of the Pigouvian tax.

Quantitatively, we find that the MCF plays an insignificant role. The second-best carbon tax starts at about 0.5% below the SCC, and then fluctuates at about 0.2% above or below. The SCC is, however, significantly affected by the presence of inequalities: consumption inequality—after income taxes are set optimally—reduces the SCC by 3.9% in our baseline calibration. We then compare our optimal policy to the one of a "climate skeptic" planner who optimizes fiscal instruments assuming climate change is exogenous, thus setting the carbon tax to zero. We find that the additional revenue raised by the carbon tax is about equally split between increasing transfers and reducing the labor income tax. Turning to welfare, we find that the optimal carbon tax policy has progressive effects in the 21$^{\text{st}}$ century—owing to the higher progressivity of the tax system—and very large positive but regressive effects afterwards, as richer households value environmental improvements proportionally more relative to consumption.

We also examine the following extensions to our benchmark model:

1. We show that the roles of tax distortions and inequalities are robust to alternative calibrations, such as more severe damages or different levels of government spending. The role of inequality increases about proportionally with productivity heterogeneity and with the share of direct utility damages, and increases more than proportionally with the intertemporal elasticity of substitution (IES): when the IES is 2, inequalities reduce the carbon tax by 16.2% instead of 3.9% in our baseline where the IES is 1.45.

2. We theoretically characterize and quantitatively compute third-best fiscal policies, *i.e.* optimal fiscal policies when either the labor or the capital income tax is exogenously fixed at its current level. The effect of inequalities on the social cost of carbon remains similar to our benchmark, although it becomes larger when the planner cannot reduce inequalities as much as it would like to. Tax distortions still play an insignificant role through the MCF, but the additional constraint on fiscal instruments now generates a new fiscal interaction term which enters additively into the pollution tax formula. When the labor or capital income tax is exogenously fixed below (resp. above) its optimal value, this term is negative (resp. positive) and the third-best tax rule is set below (resp. above) the second-best level.

3. We show that when wealth is initially unequally distributed, the optimal capital income tax is set such that all wealth is expropriated in the initial period. When the planner cannot choose the initial capital tax, however, wealth inequality is costly for the planner. As this cost depends on the level of aggregate consumption, wealth inequality also affects the opportunity cost of carbon taxation. In particular, we find that wealth inequality significantly reduces the optimal level of the carbon tax in the initial period, but does not affect the tax in subsequent periods.

4. We present a version of the model where households consume an additional dirty good that uses energy as its only input. In order to capture heterogeneous budget shares that vary with income, we model this good as a necessity. We show that as long as agents' needs are identical, the optimal tax formulas are unaffected. When agents have heterogeneous needs, the planner simply adds a subsidy on the dirty good if the agents it values relatively more have higher needs. We find that this heterogeneity does not significantly affect carbon taxes and that the subsidy is quantitatively negligible, as heterogeneous energy needs between and within income groups do not strongly co-vary with households' welfare.

5. We introduce heterogeneous sensitivity to environmental degradation in the utility function. We show that if utility is strongly separable in environmental preferences, heterogeneous environmental damages have no impact on the optimal pollution tax in the utilitarian case, but lead to higher pollution taxes when the planner values more the agents more impacted by pollution. A Rawlsian planner would therefore tax pollution at a higher level if poorer agents are also disproportionately impacted by environmental damages.

Our paper contributes to two strands of the literature. First, it contributes to the literature on the optimal taxation of pollution. In a pioneering work, Pigou (1920) established that the first-best policy response to an externality is to implement a tax equal to its social cost. An extensive literature has then investigated optimal pollution taxation in a second-best environment. In a representative-agent framework, when the government does not have access to lump-sum transfers to finance public expenditures, distortionary taxes typically raise the MCF above 1, and it becomes optimal to set the pollution tax below the Pigouvian level (see *e.g.*, Sandmo, 1975; Bovenberg and de Mooij, 1994; Bovenberg and van der Ploeg, 1994; Bovenberg and Goulder, 1996).[2] More recently, other papers have considered this problem with heterogeneous agents and uniform lump-sum taxation (see *e.g.*, Jacobs and de Mooij, 2015; Jacobs and van der Ploeg, 2019), arguing that in this set-up the MCF is equal to 1 and the second-best tax is set at the Pigouvian level.[3] While these papers focus on static settings and model the pollution externality in a stylized manner, the recent work of Barrage (2019) creates a critical bridge between the climate-economy literature and the dynamic public finance literature. Her framework integrates a climate-economy model in the spirit of Golosov et al. (2014) into the representative-agent Ramsey model (see Chari and Kehoe (1999) for a review). In this setting, tax

[2]For further references on second-best pollution taxation in representative-agents models, see Barrage (2019).

[3]Other papers jointly studying optimal pollution taxation and redistribution include, among others, Pirttilä and Tuomala (1997), Cremer et al. (1998, 2003), Micheletto (2008), and Kaplow (2012).

distortions again call for lower taxes on carbon emissions. Our main innovation relative to Barrage (2019) is to introduce heterogeneous agents, which we see as critical for two reasons. First, this allows us to jointly study environmental and equity issues. In addition of the importance of equity in normative analysis, recent experience has shown that the distributional effects of environmental policies were also critical to ensure their public support.[4] Second, agent heterogeneity provides a sound foundation for the study of second-best policies. In representative-agent settings, the second-best environment arises because lump-sum transfers are assumed unfeasible: governments therefore *need to* rely on distortionary taxes to finance their expenditures. Yet, in practice lump-sum transfers are feasible as they simply correspond to the intercept on a tax scheme.[5] With heterogeneous agents, lump-sum transfers are no longer excluded as long as they do not discriminate between agents. Although this non-distortionary source of public income is available, governments now *want to* use distortionary taxes to provide redistribution. While our optimal tax formulas resemble the ones in Barrage (2019), this significantly changes the implications of tax distortions. In particular, we find that the MCF averages to 1 over time and that its temporal variations are quantitatively insignificant, so the optimal pollution tax is approximately Pigouvian. Our results also show that unlike in representative-agent models, the weak double-dividend hypothesis—according to which it is optimal to use the pollution tax revenue exclusively to reduce distortionary taxes (see *e.g.*, Goulder, 1995)—does not hold with heterogeneous agents. At the optimum, the welfare gain from a marginal reduction in tax distortions is equal to the marginal cost from increasing inequalities, hence the optimal policy divides the carbon tax revenue about equally between reducing tax distortions and providing redistribution.[6] While tax distortions do not call for significant deviations from the Pigouvian principle as the double-dividend literature suggests, we show that inequalities matter for the taxation of pollution: by increasing the opportunity cost of abatement, inequalities reduce the social cost of pollution. In our quantitative analysis, this effect appears to be significant although not very large, reducing the carbon tax by about 4% in the baseline. Intuitively, inequalities are more effectively addressed using income taxes than the carbon tax. Still, the carbon tax is affected by the residual inequality, *i.e.* the level of inequality that remains after the planner has optimally set income taxes.

Second, this paper contributes to the analysis of the distributional effects of environmental taxes in general equilibrium. An extensive literature has analyzed the distributional effects of environmental taxes through the consumption channel (for a recent survey, see Pizer and Sexton, 2019), generally pointing to regressive effects since the consumption share of polluting goods tends to decrease with

[4]Public protests against policy-induced increases in energy prices have recently occurred in many countries worldwide. For instance, in France the Yellow Vests movement strongly opposed carbon tax increases due to the expected impact on households' purchasing power, leading to the abandonment of the scheduled carbon tax reforms (Douenne and Fabre, 2022).

[5]Recent policy proposals—such as the carbon tax and dividend advocated by the Climate Leadership Council and signed by 3,354 American economists—even call for using such instruments to redistribute the carbon tax revenue (Economists Statement on Carbon Dividends, 2019).

[6]This result echoes the recent findings of Fried et al. (2021) who study the optimal recycling policy for an exogenous carbon tax introduced in a sub-optimal tax system. In their model with heterogeneity between and within generations, they find that two-third of the carbon tax revenue should be used to reduce taxes on capital income, one third to provide redistribution.

income (Levinson and OBrien, 2019). More recently, several authors have also analyzed the heterogeneous incidence of environmental taxes on households' income. While a number of papers found progressive effects due to the larger negative impact of the policy on capital income relative to labor income and transfers (see *e.g.* Rausch et al., 2011; Fullerton and Monti, 2013; Williams et al., 2015; Goulder et al., 2019), the recent work of Känzig (2021) shows—exploiting exogenous shocks to the EU-ETS price—that carbon taxation has a larger impact on poor households' income. Many papers have also shown that the incidence of carbon taxation largely depends on how the tax revenue is recycled. In particular, Fried et al. (2018) study the economic impact of introducing a carbon tax with three alternative revenue-recycling schemes in a quantitative OLG model with heterogeneity within-generations. They show that while a uniform lump-sum rebate is more costly than reductions of the labor or capital tax rates in steady state, it is more favorable to the current generation and leads to less adverse distributional effects.[7] Finally, a few papers have considered the heterogeneous environmental benefits of climate change mitigation, between generations (*e.g.*, Leach, 2009; Kotlikoff et al., 2021) or between regions (*e.g.*, Hassler and Krusell, 2012; Krusell and Smith, 2015; Cruz and Rossi-Hansberg, 2021). In this paper, we jointly study the financial and environmental impacts from optimal pollution taxation, both over time and between heterogeneous households who differ in income, wealth, and energy budget share. We find that accounting for environmental benefits, current rich households lose the most from carbon taxation, but future rich households win the most provided they are not less exposed to environmental damages.

The rest of the paper is organized as follows. Section 2 presents the model, and Section 3 the optimal tax formulas. Section 4 describes our calibration and Section 5 presents our main quantitative exercise. Extensions of our main framework are provided in Section 6. Section 7 concludes.

2 Model

The model builds on Barrage (2019): one sector of the economy produces a final good using capital, labor, and energy, which is itself produced in the second sector. Energy production generates pollution that leads to environmental degradation, which in turn affects productivity and households' utility. The government finances an exogenous stream of expenditures using taxes on labor income, capital income, energy, and pollution, as well as a lump-sum tax. The key departures from Barrage (2019) are that, in our model, households are heterogeneous and the government has access to a (non-individualized) lump-sum tax or transfer. Consequently, although the government has access to a non-distortionary source of revenue, it uses distortionary taxes for redistributive purposes.

[7] Leach (2009), Rausch (2013), and Rausch and Yonezawa (2018) also quantitatively investigate the distributional effects from revenue recycling across generations, with a representative agent for each generation. Other papers use a dynamic model to compute the incidence of carbon tax reforms, and simulate the distributional effects across heterogeneous agents in the initial period (Williams et al., 2015) or over different time intervals (Goulder et al., 2019). All these papers consider exogenous reforms and—with the exception of Leach (2009)—ignore environmental effects.

2.1 Households

We consider an economy populated by a continuum of infinitely-lived agents divided into types $i \in I$ of size π_i. The total population size in period t is N_t. Each agent, or dynasty of type $i \in I$ ranks streams of consumption of a final good $c_{i,t}$, labor supply $h_{i,t}$, and environmental degradation Z_t according to the preferences

$$\sum_{t=0}^{\infty} \beta^t N_t u\left(c_{i,t}, h_{i,t}, Z_t\right). \tag{1}$$

In our benchmark, agents are assumed to differ in two ways: their productivity levels, e_i, and their initial asset holdings, $a_{i,0}$. Productivity levels are normalized such that $\sum_i \pi_i e_i = 1$. Agents' assets are composed of government debt and capital and we denote respectively $b_{i,t}$ and $k_{i,t}$ the number of units of these assets held by agents of type i between periods $t-1$ and t, with $a_{i,t} = b_{i,t} + k_{i,t}$. Aggregates are denoted without the subscript i: $C_t = N_t \sum_i \pi_i c_{i,t}$, $H_t = N_t \sum_i \pi_i e_i h_{i,t}$, $B_t = N_t \sum_i \pi_i b_{i,t}$, and $K_t = N_t \sum_i \pi_i k_{i,t}$. In addition, per period average consumption and hours worked are denoted by $c_t \equiv C_t/N_t$ and $h_t \equiv H_t/N_t$.

Let p_t denote the price of the consumption good in period t in terms of consumption in period 0 (so that $p_0 = 1$), w_t and r_t denote the real wage and the rental rate of capital in period t, and R_t its gross return (between $t-1$ and t). Finally, let $\tau_{H,t}$ and $\tau_{K,t}$ represent the labor and capital income taxes, and T_t the *aggregate* uniform lump-sum transfers received by all households in period t. Given $k_{i,0}$, $b_{i,0}$, prices $\{p_t, w_t, R_t\}_{t=0}^{\infty}$ and policies $\{\tau_{H,t}, \tau_{K,t}, T_t\}_{t=0}^{\infty}$, agents of type i choose $\{c_{i,t}, h_{i,t}, k_{i,t+1}, b_{i,t+1}\}_{t=0}^{\infty}$ to maximize (1) subject to the budget constraint

$$\sum_{t=0}^{\infty} p_t N_t \left(c_{i,t} + k_{i,t+1} + b_{i,t+1}\right) \leq \sum_{t=0}^{\infty} p_t N_t \left(\left(1-\tau_{H,t}\right) w_t e_i h_{i,t} + R_t \left(k_{i,t} + b_{i,t}\right) + T_t/N_t\right),$$

where $R_t \equiv 1 + \left(1-\tau_{K,t}\right)\left(r_t - \delta\right)$, for $t \geq 0$. Here, we use the convention that the capital income tax is levied on the rate of return net of depreciation, but none of our results depend on it. No arbitrage requires $p_t = R_{t+1} p_{t+1}$, and defining $T \equiv \sum_{t=0}^{\infty} p_t T_t$ as the present value of lump-sum transfers, the budget constraint can equivalently be written as

$$\sum_{t=0}^{\infty} p_t N_t \Big(c_{i,t} - \left(1-\tau_{H,t}\right) w_t e_i h_{i,t}\Big) \leq R_0 N_0 a_{i,0} + T. \tag{2}$$

From the first order conditions of agent i's problem we have

$$\beta^t \frac{u_{c,i,t}}{u_{c,i,0}} = p_t, \quad \forall\, t \geq 0,$$
$$\frac{u_{h,i,t}}{u_{c,i,t}} = -\left(1-\tau_{H,t}\right) e_i w_t, \quad \forall\, t \geq 0,$$

which holds across all agents. To reduce notations, we use subscripts x, i, t to denote partial derivatives with respect to argument x for agent of type i at time t, and we keep the arguments of the derivatives implicit.

2.2 Final good sector

As in Barrage (2019), there are two production sectors. In the final good sector, indexed by 1, a consumption-capital good is produced with a concave, constant returns to scale technology, $F(K_{1,t}, H_{1,t}, E_t)$, that uses capital $K_{1,t}$, labor $H_{1,t}$, and energy E_t. The total factor productivity is given by $A_{1,t}$ and the function $D(Z_t)$ controls the damages to production implied by environmental degradation, with $D'(Z_t) > 0$. The output $Y_{1,t}$ is given by

$$Y_{1,t} = (1 - D(Z_t))A_{1,t}F(K_{1,t}, H_{1,t}, E_t).$$

The first order conditions for the firm problem are:

$$r_t = (1 - D(Z_t))\, A_{1,t}F_{K,t}, \quad \forall\, t \geq 0, \tag{3}$$

$$w_t = (1 - D(Z_t))\, A_{1,t}F_{H,t}, \quad \forall\, t \geq 0, \tag{4}$$

$$p_{E,t} = (1 - D(Z_t))\, A_{1,t}F_{E,t}, \quad \forall\, t \geq 0. \tag{5}$$

Here, $p_{E,t}$ denotes the price of energy in period t. Because there are constant returns to scale and inputs are paid according to their marginal productivity, final goods producers make zero profits.

2.3 Energy sector

The energy sector, indexed by 2, produces energy E_t using capital $K_{2,t}$ and labor $H_{2,t}$ with a constant returns to scale technology so that

$$E_t = A_{2,t}G(K_{2,t}, H_{2,t}), \quad \forall\, t \geq 0. \tag{6}$$

Energy producers can provide a fraction μ_t of energy from clean technologies, at additional cost $\Theta_t(\mu_t, E_t)$, which satisfies $\Theta_{\mu,t}, \Theta_{E,t}, \Theta_{\mu\mu,t} > 0$, $\Theta_{EE,t} \geq 0$ and $\Theta_t(0, E_t) = \Theta_t(\mu_t, 0) = 0$. Convexity in $\Theta_t(\cdot,\cdot)$ captures decreasing returns to abatement. This function nests the one used in Barrage (2019), where $\Theta_t(\mu_t, E_t) = \Theta_t(\mu_t E_t)$, and in Nordhaus (2017), where it is equivalent to $\Theta_t(\mu_t, E_t) = \Theta_t(\mu_t)\, E_t$. In our calibration, we opt for the latter specification in order to follow DICE as closely as possible. Total profits from energy production are given by

$$\Pi_t = (p_{E,t} - \tau_{I,t})\, E_t - \tau_{E,t}(1 - \mu_t)\, E_t - w_t H_{2,t} - r_t K_{2,t} - \Theta_t(\mu_t, E_t),$$

where $\tau_{I,t}$ denotes the excise intermediate-goods tax on total energy and $\tau_{E,t}$ denotes the excise tax on pollution emissions $E_t^M = (1 - \mu_t)\, E_t$. Firms maximize profits subject to the technology constraint given by equation (6) by choosing the abatement term μ_t, capital $K_{2,t}$, and labor $H_{2,t}$. The first order conditions are

$$r_t = \big(p_{E,t} - \tau_{I,t} - \tau_{E,t}(1 - \mu_t) - \Theta_{E,t}\big)A_{2,t}G_{K,t}, \quad \forall\, t \geq 0, \tag{7}$$

$$w_t = \big(p_{E,t} - \tau_{I,t} - \tau_{E,t}(1 - \mu_t) - \Theta_{E,t}\big)A_{2,t}G_{H,t}, \quad \forall\, t \geq 0, \tag{8}$$

$$\tau_{E,t} = \frac{\Theta_{\mu,t}}{E_t}, \quad \forall\, t \geq 0. \tag{9}$$

If there is positive abatement and $\Theta_t(\cdot,\cdot)$ is convex in its second argument, profits in the energy sector will be positive. For simplicity, we assume that these profits are taxed at a confiscatory rate $\tau_{\pi,t} = 1$. Doing so is typically optimal, as taxing pure profits does not generate distortions and income from shareholdings tends to be unequally distributed. In our calibration in Section 4, the abatement cost function is strictly convex in its first argument and linear in the second, hence profits are null.

Capital and labor are mobile across sectors, so market clearing requires

$$K_{1,t} + K_{2,t} = K_t, \quad \forall\, t \geq 0, \tag{10}$$

$$H_{1,t} + H_{2,t} = H_t, \quad \forall\, t \geq 0. \tag{11}$$

2.4 Government

Each period the government finances its expenses G_t and lump sum transfers T_t with proportional income taxes on capital $\tau_{K,t}$ and labor $\tau_{H,t}$, total energy taxes $\tau_{I,t}$, and emissions taxes $\tau_{E,t}$. In addition, profits are taxed at a confiscatory rate: $\tau_{\pi,t} = 1$. The government's budget constraint is

$$R_0 B_0 + T + \sum_t p_t G_t = \sum_t p_t \left(\tau_{H,t} w_t H_t + \tau_{K,t} \left(r_t - \delta \right) K_t + \tau_{I,t} E_t + \tau_{E,t} E_t^M + \Pi_t \right). \tag{12}$$

Although the instruments levied are proportional, the tax system is progressive when transfers are positive. As shown in Piketty and Saez (2013) and Dyrda and Pedroni (2022), an affine tax system provides a good approximation of actual tax systems such as the one of the U.S.

2.5 Environmental degradation

The environmental variable is affected by the history of pollution emissions $E_t^M = (1 - \mu_t) E_t$, initial conditions S_0, and the history of exogenous shifters η_t according to

$$Z_t = J \left(S_0, E_0^M, ..., E_t^M, \eta_0, ..., \eta_t \right), \quad \forall\, t \geq 0. \tag{13}$$

In our calibration below, Z represents the global mean temperature that is the outcome of the climate model J. In this section and the next, we do not further specify this function and our theoretical results can apply to any kind of pollution externality affecting production and households' utility.

2.6 Competitive equilibrium

Definition 1 *Given a distribution of assets $\{a_{i,0}\}$, aggregate capital K_0 and aggregate bond holdings B_0, a competitive equilibrium is a policy $\{\tau_{H,t}, \tau_{K,t}, \tau_{I,t}, \tau_{E,t}, T_t\}_{t=0}^{\infty}$, a price system $\{p_t, w_t, r_t, p_{E,t}\}_{t=0}^{\infty}$ and an allocation $\left\{ (c_{i,t}, h_{i,t})_i, Z_t, E_t, K_{1,t}, K_{2,t}, K_{t+1}, H_{1,t}, H_{2,t}, H_t \right\}_{t=0}^{\infty}$ such that: (i) agents choose $\left\{ (c_{i,t}, h_{i,t})_i \right\}_{t=0}^{\infty}$ to maximize utility subject to budget constraint (2) taking policies and prices (that satisfy $p_t = R_{t+1} p_{t+1}$) as given; (ii) firms maximize profits; (iii) the government's budget constraint (12) holds; (iv) markets clear: the resource constraints (6), (10), (11), and (13) hold, and*

$$N_t c_t + G_t + K_{t+1} + \Theta_t \left(\mu_t, E_t \right) = \left(1 - D \left(Z_t \right) \right) A_{1,t} F \left(K_{1,t}, H_{1,t}, E_t \right) + (1 - \delta) K_t, \quad \forall\, t \geq 0. \tag{14}$$

3 Optimal tax rules

In this section, we use the technique introduced by Werning (2007) to express agents' equilibrium allocations as a function of aggregate variables, and solve the Ramsey problem as a function of aggregates instead of their full distributions.

3.1 Ramsey problem

A simple characterization of equilibrium Because the government sets linear tax rates, all agents face the same marginal rate of substitution between consumption and leisure. Consequently, the distribution of individual allocations $(c_{i,t}, h_{i,t})$ is efficient *given* aggregates (c_t, h_t, Z_t), where $c_t = C_t/N_t$ and $h_t = H_t/N_t$ denote the average consumption and hours worked in period t. Following Werning (2007), it is therefore possible to split up the optimal tax problem in two steps. The first is to determine individual allocations given aggregates, and the second is to determine the aggregates. Starting with the first step, denote by $\varphi \equiv \{\varphi_i\}$ a set of market weights with $\varphi_i \geq 0$. Using the property that individual allocations are efficient given aggregates, we can characterize these allocations by solving the following static sub-problem for each period t:

$$\begin{aligned} U\left(c_t, h_t, Z_t; \varphi\right) &\equiv \max_{c_{i,t}, h_{i,t}} \sum_i \pi_i \varphi_i u\left(c_{i,t}, h_{i,t}, Z_t\right), \\ \text{s.t.} \quad \sum_i \pi_i c_{i,t} &= c_t \quad \text{and} \quad \sum_i \pi_i e_i h_{i,t} = h_t. \end{aligned} \tag{15}$$

Here, $U\left(c_t, h_t, Z_t; \varphi\right)$ denotes the indirect aggregate utility function, computed using market weights and aggregates. To reduce the notation burden and ease tractability, we assume that utility is additively separable in Z, *i.e.* that we can write

$$u\left(c_{i,t}, h_{i,t}, Z_t\right) \equiv \tilde{u}\left(c_{i,t}, h_{i,t}\right) + \hat{u}(Z_t).$$

Implementability condition Applying the envelope theorem to problem (15) and using consumers' first order conditions we get

$$\frac{U_{h,t}}{U_{c,t}} = \frac{u_{h,i,t}}{u_{c,i,t} e_i} = -w_t \left(1 - \tau_{H,t}\right),$$

and

$$\frac{U_{c,t}}{U_{c,0}} = \frac{u_{c,i,t}}{u_{c,i,0}} = \frac{p_t}{\beta^t}.$$

Using these relationships to substitute out for prices in agents' budget constraints, for any agent i we can derive an implementability condition that depends only on the aggregates c_t and h_t, and market weights φ

$$U_{c,0}\big(R_0 N_0 a_{i,0} + T\big) = \sum_{t=0}^{\infty} N_t \beta^t \Big(U_{c,t} c_{i,t}^m\big(c_t, h_t; \varphi\big) + U_{h,t} e_i h_{i,t}^m\big(c_t, h_t; \varphi\big)\Big), \ \ \forall\, i, \tag{16}$$

with $c_{i,t}^m(c_t, h_t; \varphi)$ and $h_{i,t}^m(c_t, h_t; \varphi)$ solutions to problem (15). The following Proposition follows immediately from the arguments above.

Proposition 1 *An aggregate allocation* $\{c_t, H_{1,t}, H_{2,t}, K_{1,t}, K_{2,t}, E_t, Z_t, \mu_t\}_{t=0}^{\infty}$ *can be supported by a competitive equilibrium if and only if the market clearing conditions (10), and (11) hold, the resource constraints* (6), (13), (14) *hold and there exist market weights* φ *and a lump-sum tax* T *such that the implementability conditions (16) hold for all* $i \in I$*. Individual allocations can then be computed using functions* $c_{i,t}^m$ *and* $h_{i,t}^m$*, prices and taxes can be computed using the usual equilibrium conditions.*

Problem Let $\lambda \equiv \{\lambda_i\}$ be the planner's welfare weight on type i, with $\sum_i \pi_i \lambda_i = 1$. The Ramsey problem is

$$\max_{\substack{\{c_t, H_{1,t}, H_{2,t}, K_{1,t}, K_{2,t}, \\ E_t, Z_t, \mu_t\}_{t=0}^{\infty}, T, \varphi}} \sum_{t,i} N_t \beta^t \pi_i \lambda_i u\Big(c_{i,t}^m(c_t, h_t; \varphi), h_{i,t}^m(c_t, h_t; \varphi), Z_t\Big), \tag{17}$$

subject to

$$\begin{aligned}
U_{c,0}\big(R_0 N_0 a_{i,0} + T\big) &= \sum_{t=0}^{\infty} N_t \beta^t \left(U_{c,t} c_{i,t}^m(c_t, h_t; \varphi) + U_{h,t} e_i h_{i,t}^m(c_t, h_t; \varphi)\right), \quad \forall\, i, \\
F_{K,t} G_{H,t} &= G_{K,t} F_{H,t}, \quad \forall\, t \geq 0, \\
N_t c_t + G_t + K_{t+1} + \Theta_t(\mu_t, E_t) &= (1 - D(Z_t)) A_{1,t} F(K_{1,t}, H_{1,t}, E_t) + (1-\delta) K_t, \quad \forall\, t \geq 0, \\
E_t &= A_{2,t} G(K_{2,t}, H_{2,t}), \quad \forall\, t \geq 0, \\
Z_t &= J\left(S_0, E_0^M, ..., E_t^M, \eta_0, ..., \eta_t\right), \quad \forall\, t \geq 0, \\
K_{1,t} + K_{2,t} &= K_t, \quad \forall\, t \geq 0, \\
H_{1,t} + H_{2,t} &= N_t h_t, \quad \forall\, t \geq 0.
\end{aligned}$$

The first of these constraints is the implementability condition, which must hold for each agent i. It is written solely in terms of aggregate variables and states that the present value of consumption equals the present value of labor income, initial assets and lump-sum transfers. The second constraint states that the marginal rate of technical substitution between capital and labor is the same in both sectors, a restriction associated with the fact that the government does not use sector-specific instruments and factors are mobile across sectors. The other constraints reflect market clearing for capital, labor and goods, and technological constraints.

To simplify the exposition, we assume for now that there is no initial wealth inequality, that is $a_{i,0} = a_{j,0}$ for all i and j. An equivalent interpretation is that there is initial wealth inequality, but that all wealth is expropriated by the planner. This can be done by taxing it directly, $R_0 = 0$, or through a combination of consumption and labor taxes: see Werning (2007) for a discussion.[8] We relax the assumption that there is no initial wealth inequality, or equivalently that all wealth can be expropriated, and study the implications for optimal taxes in Section 6.2. Without initial wealth inequality and with the ability to adjust lump-sum transfers, the optimal level of $\tau_{K,0}$ is indeterminate. We therefore assume that $\tau_{K,0}$ is taken as given by the Ramsey planner.[9]

[8] Levying a confiscatory tax on all initial wealth is generally optimal if assets and productivity are positively correlated. In that case, taxing wealth reduces inequality without generating any distortions.

[9] If there is initial wealth inequality and the government can adjust a lump-sum transfer, the level of $\tau_{K,0}$ is no longer

3.2 General formulas

Capital and labor income taxes From the planner's first order conditions, the labor and capital income taxes are determined by

$$\tau_{H,t} = 1 - \frac{U_{h,t}}{U_{c,t}} \frac{W_{c,t}}{W_{h,t}},$$

and

$$\frac{R_{t+1}}{R^*_{t+1}} = \frac{W_{c,t+1}}{W_{c,t}} \frac{U_{c,t}}{U_{c,t+1}},$$

where the pseudo-utility function W is defined as

$$W\left(c_t, h_t, Z_t; \varphi, \theta, \lambda\right) \equiv V\left(c_t, h_t, Z_t; \varphi, \lambda\right) + \sum_i \pi_i \theta_i IC_i(c_t, h_t, \varphi),$$

with

$$V\left(c_t, h_t, Z_t; \varphi, \lambda\right) \equiv \sum_i \pi_i \lambda_i u\left(c^m_{i,t}\left(c_t, h_t; \varphi\right), h^m_{i,t}\left(c_t, h_t; \varphi\right), Z_t\right), \tag{18}$$

the aggregate utility based on the planner's weights,

$$IC_i(c_t, h_t, \varphi) \equiv U_{c,t} c^m_{i,t}\left(c_t, h_t; \varphi\right) + U_{h,t} e_i h^m_{i,t}\left(c_t, h_t; \varphi\right), \tag{19}$$

the difference between agent i spending on consumption and labor income in period t as it appears in its implementability constraint, and $\pi_i \theta_i$ the Lagrange multiplier on the implementability constraint of agent i in the Ramsey problem. These optimal tax formulas are the same as the ones obtained in Werning (2007). The reason is that the environmental variable enters additively to the problem and does not *directly* affect the labor and capital tax rules.

Excise taxes on energy and emissions The planner's first order conditions together with firms equilibrium conditions give

$$\tau_{I,t} = 0.$$

Thus, as long as labor, capital, profits and pollution can be taxed, there is no point in distorting production decisions. This result can also be found in Bovenberg and Goulder (1996) and Barrage (2019) and goes back to the production efficiency theorem of Diamond and Mirrlees (1971). Turning to the pollution tax we have

$$\tau_{E,t} = \sum_{j=0}^{\infty} \beta^j \left(\frac{V_{c,t+j} + \sum_i \pi_i \theta_i IC_{c,i,t+j}}{V_{c,t} + \sum_i \pi_i \theta_i IC_{c,i,t}} D'_{t+j} A_{1,t+j} F_{t+j} - \frac{N_{t+j} V_{Z,t+j}}{V_{c,t} + \sum_i \pi_i \theta_i IC_{c,i,t}} \right) J_{E^M_t, t+j}, \tag{20}$$

where the arguments to the production function F_t have been dropped to simplify notations. The term $V_{c,t} + \sum_i \pi_i \theta_i IC_{c,i,t}$ appears from the substitution of $W_{c,t} = \nu_{1,t}$, where $\nu_{1,t}$ is the Lagrange multiplier on the planner's resource constraint. When the pollution tax increases, abatement increases, which

indeterminate. However, when studying the impact of initial wealth inequality on optimal taxes in Section 6.2, we also treat $\tau_{K,0}$ as given. The reason for doing so is that optimizing over $\tau_{K,0}$ allows the planner to confiscate all initial wealth, which immediately gets rid of all initial wealth *inequality* as well.

increases the scarcity of the final good. The opportunity cost of increasing the pollution tax therefore corresponds to the marginal cost of increasing the final good's scarcity, which is equal to the marginal utility from consumption as computed using the planner's weights ($V_{c,t}$) plus a term which captures the marginal reduction in the planner's implementation cost from an increase in aggregate consumption ($\sum_i \pi_i \theta_i IC_{c,i,t}$). Intuitively, θ_i represents the shadow cost of transferring one unit of consumption to households i, and $IC_{i,t}$ the difference between i's consumption and labor income in period t. Therefore, $\sum_i \pi_i \theta_i IC_{i,t}$ represents the cost for the planner to implement its preferred allocation in period t. The degree to which this cost depends on the scarcity of the final good, c_t, is captured by the term $IC_{c,i,t}$.

This formula holds both for the first and second-best. Still, the optimal pollution tax may differ between these two fiscal environments for three reasons: the value of the marginal implementation cost, the path of aggregate variables, and the distribution of individual allocations all depend on fiscal policies.

3.3 Comparison with first-best

The role of tax distortions The first potential difference between the first and second-best pollution tax lies in the value of the marginal reduction in implementation cost, $\sum_i \pi_i \theta_i IC_{c,i,t}$. In the first-best, the first order conditions with respect to individualized lump-sum transfers give

$$\theta_i = 0, \quad \forall i.$$

It follows that the planner can achieve its preferred allocation at no cost, and the optimal pollution tax simplifies to

$$\tau_{E,t}^{FB} = \sum_{j=0}^{\infty} \beta^j \left(\frac{V_{c,t+j}}{V_{c,t}} D'_{t+j} A_{1,t+j} F_{t+j} - \frac{N_{t+j} V_{Z,t+j}}{V_{c,t}} \right) J_{E_t^M,t+j}.$$

This formula illustrates the well-known Pigouvian principle according to which the optimal corrective tax is equal to the social cost of the externality: the tax corresponds to the discounted sum of marginal (utility and production) damages valued at the marginal utility of consumption.

Turning to the second-best case, where only a uniform lump-sum transfer is available in each period, the first order condition with respect to the transfer gives

$$\sum_i \pi_i \theta_i = 0,$$

hence

$$\sum_i \pi_i \theta_i IC_{c,i,t} = \text{cov}\Big(\theta_i, IC_{c,i,t}\Big).$$

Thus, at the second-best, the sum of the multipliers associated with the implementability conditions is zero, but the marginal cost for the planner to implement its preferred allocation in a given period is not necessarily zero. The definitions below lead to Proposition 2 which states how the second-best pollution tax deviates from the Pigouvian principle when the covariance term above deviates from 0.

Definitions (Pigouvian tax) *From the first-best tax formula, we can decompose the Pigouvian tax into a production component ($\tau_{E,t}^{Pigou,Y}$) and a utility damage component ($\tau_{E,t}^{Pigou,U}$),*

$$\tau_{E,t}^{Pigou,Y} \equiv \sum_{j=0}^{\infty} \beta^j \frac{V_{c,t+j}}{V_{c,t}} D'_{t+j} A_{1,t+j} F_{t+j} J_{E_t^M,t+j},$$

$$\tau_{E,t}^{Pigou,U} \equiv (-1) \sum_{j=0}^{\infty} \beta^j \frac{N_{t+j} V_{Z,t+j}}{V_{c,t}} J_{E_t^M,t+j},$$

with the total Pigouvian tax $\tau_{E,t}^{Pigou} \equiv \tau_{E,t}^{Pigou,Y} + \tau_{E,t}^{Pigou,U}$, the share of marginal utility damages at time t,

$$\omega_t^U \equiv \frac{\tau_{E,t}^{Pigou,U}}{\tau_{E,t}^{Pigou}},$$

and the share of marginal production damages occurring at time $t+s$ due to a marginal change in emissions at time t,

$$\Delta_{t+s} \equiv \frac{\beta^s V_{c,t+s} D'_{t+s} A_{1,t+s} F_{t+s} J_{E_t^M,t+s}}{\sum_{j=0}^{\infty} \beta^j V_{c,t+j} D'_{t+j} A_{1,t+j} F_{t+j} J_{E_t^M,t+j}} .$$

Definition (Marginal cost of funds) *Let us define the marginal cost of funds (MCF) as the ratio of the public to the private marginal utility of consumption,*[10]

$$\mathrm{MCF}_t \equiv \frac{\nu_{1,t}}{V_{c,t}}.$$

Definition (Balanced-growth preferences) *An agent has balanced-growth preferences if its utility function can be expressed as*

$$u\left(c_i, h_i, Z\right) = \frac{\left(c_i (1 - \varsigma h_i)^{\gamma}\right)^{1-\sigma}}{1-\sigma} + \hat{u}(Z), \tag{21}$$

with $1/\sigma$ the intertemporal elasticity of substitution (IES).

Proposition 2 *Let $\left.\tau_{E,t}^{Pigou}\right|_{SB}$ denote the Pigouvian tax evaluated at the second-best allocation. When the planner has only access to a uniform lump-sum transfer, the optimal pollution tax formula is a modified Pigouvian rule adjusted for the marginal cost of funds,*

$$\tau_{E,t} = \left.\tau_{E,t}^{Pigou}\right|_{SB} \left(\sum_{j=0}^{\infty} \frac{\mathrm{MCF}_{t+j}}{\mathrm{MCF}_t} \Delta_{t+j} (1 - \omega_t^U) + \frac{\omega_t^U}{\mathrm{MCF}_t} \right) , \tag{22}$$

[10] Jacobs and de Mooij (2015) and Jacobs and van der Ploeg (2019) use a definition of the marginal cost of funds that takes into account fiscal externalities resulting from income effects. They find that the marginal cost of funds equals 1 at the optimal tax system, owing to the fact that the government can optimize a lump-sum transfer (see also Jacobs, 2018). However, because as in Barrage (2019) we optimize over the allocation variables directly rather than over tax instruments, it is more convenient to define the marginal costs of funds as the ratio between the multiplier on the government budget constraint and the average marginal utility of consumption computed using Pareto weights, which Jacobs and de Mooij (2015) refer to as the traditional measure of the MCF.

with

$$\mathrm{MCF}_t = 1 + \frac{\operatorname{cov}(\theta_i, IC_{c,i,t})}{V_{c,t}}. \tag{23}$$

If agents have balanced-growth preferences, then from period 0 the welfare-weighted average MCF is 1, i.e. with $V_t \equiv V\left(c_t, h_t, Z_t; \varphi, \lambda\right)$*,*

$$\frac{\sum_{t=0}^{\infty} N_t \beta^t V_t \times \mathrm{MCF}_t}{\sum_{t=0}^{\infty} N_t \beta^t V_t} = 1.$$

In the limit case where the IES tends to 1, then for $t \geq 0$*,* $\mathrm{MCF}_t = 1$*.*

The proof of Proposition 2 is provided in Appendix A.5. The optimal pollution tax reflects the arbitrage between the marginal benefits of pollution abatement and the opportunity cost from reductions in aggregate consumption. In the first-best, this opportunity cost is given by the marginal utility of consumption, $V_{c,t}$. In the second-best, the planner also accounts for the fiscal costs associated with a reduction in consumption. At any time $t \geq 0$, the shadow cost of the consumption good is given by $V_{c,t} \times \mathrm{MCF}_t$, hence the opportunity cost of abatement is higher than in the first best if and only if the MCF is above 1. Fiscal distortions also affect the marginal benefits of pollution abatement through the value of future production damages. Thus, when the MCF decreases over time, fiscal distortions operate as a form of discounting: consumption is valued relatively more in the present than in the future, hence future production damages are relatively under-internalized. We show in Appendix A.5 that the ratio of MCFs can be expressed as

$$\frac{\mathrm{MCF}_{t+j}}{\mathrm{MCF}_t} = \prod_{k=1}^{j} \frac{R_{t+k}}{R^*_{t+k}} \ ,$$

from which we see that the MCF is constant if the capital tax is null for all future periods where current emissions generate production damages. Thus, as in Barrage (2019), the optimal tax on production damage is not distorted as long as, going forward, the capital income tax is optimally set to zero. Intuitively, in this situation fiscal distortions affect future marginal abatement benefits proportionally to current marginal abatement costs. Production damages are then perfectly internalized, and the optimal tax can be expressed as

$$\tau_{E,t} = \tau_{E,t}^{Pigou,Y}\Big|_{SB} + \frac{\tau_{E,t}^{Pigou,U}\Big|_{SB}}{\mathrm{MCF}_t}.$$

Proposition 2 additionally provides an expression for the MCF as a function of individual allocations given by θ_i and $IC_{c,i,t}$. The first term, θ_i, represents the shadow cost for the planner of providing an additional unit of lump-sum transfer to agent i. While θ_i is on average null at the optimum, it typically takes a positive value for rich agents and a negative value for poor agents.[11] The second term, $IC_{c,i,t}$,

[11] As shown in Appendix A.4.2, with balanced-growth preferences we have

$$\theta_i = \sum_j \frac{\pi_j \lambda_j}{\varphi_j} - \frac{\lambda_i}{\varphi_i}, \quad \forall i,$$

hence $-\theta_i$ represents how much the agent is valued by the planner relative to the market, as compared to an average agent.

represents how the difference between the agent current consumption and current labor income changes when more resources are available for consumption. We show in Appendix A.5 that this term is in fact driven by two mechanisms: a volume and a price effect. When less resources are used for pollution abatement, consumption increases and labor supply adjusts, which also affects prices and wages. When households have balanced-growth preferences and the IES tends to 1, these two effects exactly offset each other, such that the present value of the lump-sum transfer necessary to satisfy households' budget constraints remains unchanged. In this situation, taxing pollution does not affect tax distortions, the MCF is equal to 1, and the second-best tax is exactly Pigouvian. When the IES is below (resp. above) 1, the price (resp. volume) effect dominates and an increase in aggregate consumption reduces (resp. increases) the total amount of transfers needed to satisfy agents' implementability constraints. If these changes are heterogeneous across households and correlate with their type, the MCF differs from 1. At the optimum, agents' binding implementability conditions imply

$$\sum_{t=0}^{\infty} N_t \beta^t IC_{i,t} = U_{c,0}(R_0 a_{i,0} + T),$$

hence with no initial wealth inequality (or equivalently, full expropriation of initial wealth) the discounted sum of $IC_{i,t}$ is invariant across types. Intuitively, this condition means that with a uniform lump-sum transfer, the discounted sum of expenditures minus labor income must be the same for everyone. We show in Appendix A.5 that this condition implies that with balanced-growth preferences the covariance term in (23) averages to 0 over time, hence the MCF is on average equal to 1 and the optimal pollution tax is on average equal to the Pigouvian level. Still, in any period $t \geq 0$, this covariance term may differ from 0, hence temporary deviations from the Pigouvian principle may occur. In particular, we show in Appendix A.5 that with balanced-growth preferences, the covariance is positive when IES is below 1 and aggregate labor supply is high relative to its long-run value. In this situation, increasing aggregate consumption makes it relatively easier to satisfy the budget constraint of richer agents for whom transfers are costly for the planner ($\theta_i > 0$), hence the opportunity cost of pollution taxation is higher because of fiscal motives, the MCF is above 1, and the optimal tax is (temporarily) below the Pigouvian level.

The role of inequalities When the marginal cost of funds is 1, the first and second-best tax formulas coincide, and they are both equal to the social cost of pollution. Still, the actual tax levels may differ for two reasons.

The first reason is that, when the tax system is different, aggregate variables generally take different values. When capital and labor are taxed, labor supply and investments are expected to be lower, hence output, consumption, and pollution are also expected to be lower along the optimal path. Since the pollution tax level is determined by the trade-off between the marginal utility of consumption and the marginal utility of pollution abatement, if both pollution and consumption are lower, the optimal tax will generally be set at a lower level since utility is concave in consumption and convex in pollution.[12]

[12]This result also depends on the law of motion of environmental degradation: if each additional unit of pollution emitted increases degradation by less than the previous unit, the marginal abatement benefits could be lower for higher

The second reason is that the distribution of individual allocations also differs depending on the fiscal environment. Because individualized lump-sum transfers are not feasible in the second-best, there are generally more consumption inequalities. The welfare gains from leaving more resources available for agents' consumption by decreasing the pollution tax may then be higher or lower compared to the first-best depending on the curvature of agents' utility function.

Proposition 3 *The social cost of pollution from utility damages is inversely related to the social marginal utility of consumption $V_{c,t}$. If agents have balanced-growth preferences, $V_{c,t}$ can be expressed as*

$$V_{c,t} = \sum_i \pi_i \lambda_i u_{c,i,t} + \mathrm{cov}\Big(\lambda_i u_{c,i,t}, \frac{c_{i,t}}{c_t}\Big),$$

and holding aggregate variables constant, consumption inequalities affect $V_{c,t}$ in two opposite ways: i) they increase it by increasing the average value of households' marginal utility of consumption, and ii) they reduce it because a larger share of additional consumption ($c_{i,t}/c_t$) is attributed to households with lower marginal utilities of consumption ($u_{c,i,t}$). In the limit case where the IES tends to 1, the two effects exactly offset each other and consumption inequalities do not affect the tax level.

The proof of Proposition 3 is provided in Appendix A.5. In the presence of inequalities, an increase in aggregate consumption is valued more to the extent that households' marginal utilities are higher on average (by convexity of the marginal utility function), but it is valued less to the extent that the inflow in consumption disproportionately goes to richer households with lower marginal utilities. An increase in the pollution tax reduces every households' consumption proportionally. In the limiting case in which IES tends to 1, the planner is indifferent between a proportional increase in consumption for a rich or a poor agent, so inequalities do not affect the planner's marginal valuation of aggregate consumption.[13] When utility is more concave, the first mechanism becomes relatively stronger and inequalities lead to a higher social marginal utility of consumption, thereby increasing the opportunity cost associated with raising pollution taxes.

4 Calibration

In this section, we explain how we calibrate the model to explore quantitatively the implications of heterogeneity in productivity for the optimal taxation of carbon, capital income, and labor income. As in Barrage (2019), we consider a climate-economy model based on Nordhaus' DICE model. While Barrage (2019) considers a planner setting taxes for the global economy, we adopt a slightly different approach: we consider a global economy with the economic features of the U.S. economy, *i.e.* we

levels of pollution.

[13]In the simpler case where agents have logarithmic utility on consumption only, it is straightforward to see that the distribution of households' consumption has no effect on the planner's valuation of an homogeneous increase in consumption

$$\sum_i \pi_i \lambda_i \ln((1+x)c_i) - \sum_i \pi_i \lambda_i \ln(c_i) = \ln(1+x).$$

parametrize the income per capita, the productivity distribution, and the fiscal system to match U.S. data, but we scale our economy so that output and emissions match global data. The objective is to determine how an economy with important inequalities and responsible for a significant share of global emissions like the U.S. should design its fiscal system if it were to internalize the global impact of its emissions, assuming that the rest of the world would behave identically.

4.1 Climate model

The calibration of the climate model is based on the 2016 version of DICE, presented for example in Nordhaus (2017). The initial period is 2015, and each period lasts 5 years. The climate model is composed of three sets of equations describing the carbon cycle, radiative forcing, and climate change.

Carbon cycle The carbon cycle is represented by three reservoirs. S^{At}, S^{Up}, and S^{Lo} represent the level of carbon concentration in the atmosphere, the upper oceans and biosphere, and the deep oceans respectively. These stocks evolve according to the following laws of motion:

$$S_t^j = b_{0,j}(E_t^M + E_t^{\text{land}}) + \sum_{i=1}^{3} b_{i,j} S_{t-1}^i,$$

where the three reservoirs j are ranked as above and with E_t^{land} the exogenous land emissions. The coefficient $b_{0,j}$ is 1 for the first reservoir (S^{At}) and 0 for the others: industrial and land emissions directly flow into the atmosphere, and later affect the other two reservoirs through the communication between the carbon stocks captured by the parameters $b_{i,j}$.

Radiative forcing The accumulation of carbon in the atmosphere increases radiative forcing, *i.e.* the net radiation received by the earth. This mechanism is captured by the following equation

$$\mathcal{F}_t = \kappa\big(\ln(S_t^{At}/S_{1750}^{At})/\ln(2)\big) + \mathcal{F}_t^{\text{ex}}.$$

where $\mathcal{F}_t^{\text{ex}}$ is exogenous forcing. A positive radiative forcing means that the earth receives more energy from the sun than it emits back to space, hence the climate warms.

Climate change The change in temperature is modeled through two equations for the mean temperature of the atmosphere (Z_t^{At}) and deep oceans (Z_t^{Lo}) that interact as follows

$$Z_t^{At} = Z_{t-1}^{At} + \zeta_1\big(\mathcal{F}_t - \zeta_2 Z_{t-1}^{At} - \zeta_3(Z_{t-1}^{At} - Z_{t-1}^{Lo})\big),$$

$$Z_t^{Lo} = Z_{t-1}^{Lo} + \zeta_4(Z_{t-1}^{At} - Z_{t-1}^{Lo}).$$

All the parameters of the climate model are taken from DICE 2016, and reported in Table VI in the appendix.

4.2 Damages

We also model production damages as in DICE 2016, with

$$D(Z) = a_1 Z + a_2 Z^{a_3}, \tag{24}$$

As in DICE, we assume that $D(Z)$ is a simple quadratic function with $a_1 = 0$ and $a_3 = 2$. The relevant Z that enters this formula, in each period t, is the atmospheric temperature Z_t^{At}. Since DICE does not distinguish between production and utility damages, we follow Barrage (2019) to decompose the damages from DICE into a production and a utility component. We apply her decomposition and assign 74% of damages at 2.5°C warming to output, and 26% to utility. This provides an adjusted value for the parameter a_2 in equation (24), and enables us to calibrate utility damages (specifically the preference parameter α_0 described below).

To examine the robustness of our quantitative results to the level of damages, we also consider an alternative "high damage" specification. Instead of assuming quadratic damages, we consider a cubic function ($a_1 = 0$, $a_3 = 3$) and we adjust the coefficient a_2 such that damages are identical to the baseline scenario at current warming. This high damages scenario therefore assumes that the damage function in DICE correctly captures current damages, but mis-estimates damages at higher levels of warming because of the high uncertainties surrounding the impacts of climate change at these higher temperatures (see *e.g.*, Weitzman, 2009; Pindyck, 2013).

4.3 Households

We assume households have balanced-growth preferences as defined in (21) with

$$\hat{u}(Z) = \frac{\left(1 + \alpha_0 Z_t^2\right)^{-(1-\sigma)}}{1-\sigma},$$

as in Barrage (2019). Using market weights, the intertemporal aggregate utility is

$$\sum_t \beta^t N_t U\left(c_t, h_t, Z_t, \varphi\right) = \sum_t \beta^t N_t \left(\frac{\left(c_t (1 - \varsigma h_t)^{\gamma}\right)^{1-\sigma}}{1-\sigma} + \Gamma \frac{\left(1 + \alpha_0 Z_t^2\right)^{-(1-\sigma)}}{1-\sigma} \right),$$

with $\Gamma \equiv \sum_i \pi_i \varphi_i$ and where $Z_t \equiv Z_t^{At}$ is the atmospheric temperature (see Appendix A.4.1). To ensure that aggregate emissions remain consistent with DICE, we calibrate the growth rate of population accordingly. Because we also want to match the GDP per capita of the U.S., we set the population levels as U.S. population multiplied by the ratio of world to U.S. GDP in 2011-2015, the first period of the model.

Following DICE, we calibrate the utility discount factor to $\beta = 1/(1 + 0.015)$ per year, and the inverse of the IES to $\sigma = 1.45$. The parameters γ and ς are set in order to match a Frisch elasticity of labor supply of 0.75 (see Chetty et al., 2011) and an average per capita labor supply of $h_{2015} = 0.277$ in the initial period (computed from the Survey of Consumer Finances, see Appendix F.2).

We calibrate the ability distribution on the basis of hourly wage data that we obtain from the Survey of Consumer Finances (SCF). To be consistent with the initial period in DICE (2011-2015), we use the

SCF 2013. We divide the sample of working households into ten groups of hourly wage deciles (*i.e.*, $I = 10$, and for all i, $\pi_i = 0.1$), with an hourly wage of \$6.44 for the bottom productivity group and \$101.35 for the top productivity group, and normalize productivity levels such that $\sum_i \pi_i e_i = 1$. The full procedure is described in Appendix F.1. While we calibrate productivity levels directly instead of targeting a specific *ex post* distribution of inequalities, the model correctly predicts consumption inequalities, with a consumption Gini of 0.33, very close to the value of 0.32 observed in the data (see Heathcote et al., 2010). Thus, although our model abstract from idiosyncratic income risk, we still correctly capture lifetime economic inequalities.

4.4 Production

We model production using a Cobb-Douglas technology for both sectors. We have

$$F(K_{1,t}, H_{1,t}, E_t) = K_{1,t}^{\alpha} H_{1,t}^{1-\alpha-\nu} E_t^{\nu}$$

with $\alpha = 0.3$, and $\nu = 0.04$ (from Golosov et al., 2014), and

$$G(K_{2,t}, H_{2,t}) = K_{2,t}^{1-\alpha_E} H_{2,t}^{\alpha_E}.$$

with $\alpha_E = 0.403$ (from Barrage, 2019). The initial total factor productivities $A_{1,2015}$ and $A_{2,2015}$ are set such that output in sectors one and two match world GDP (2011-2015 average from the World Bank) and aggregate industrial emissions (from DICE 2016) respectively, and their growth rate are taken from DICE 2016.[14] Our abatement cost function is also taken from DICE, with the following specification

$$\Theta(\mu_t, E_t) = c_{1,t} \mu_t^{c_2} E_t,$$

where $c_{1,t} c_2 = P_t^{\text{backstop}}$ represents the backstop price, *i.e.* the price at which it becomes economical to abate 100% of emissions. As in DICE 2016, we assume that this price is \$550/tCO$_2$ in the initial period, and declines at a rate of 0.5% per year. We also calibrate the exponent $c_2 = 2.6$ as in DICE.

4.5 Government

We calibrate the fiscal part of the model to match data on U.S. fiscal policy. Here we deviate from Barrage (2019) who sets tax rates, government spending, and debt to match their empirical counterparts at the *global* level. The reason for targeting the U.S. rather than the global economy is that the degree of inequality is calibrated to match the U.S. income and wealth distribution and, more importantly, in our framework and in reality fiscal policy is typically decided on at the national level. To make the model consistent with the (global) evolution of the climate, we subsequently scale up the economy such that GDP and total emissions are consistent with their global levels. By doing so, rather than ignoring negative effects from emissions on other countries, we assume that U.S. fiscal policy is set to fully internalize the negative global effects from their carbon emissions.

[14] To calibrate the initial values of $K_{1,0}$ and $K_{2,0}$, we assume that the economy is in a balanced-growth path in which temperature remains constant at the current level.

To calibrate fiscal policy, we first require the empirical counterparts of taxes. In the model, there are four taxes: a tax $\tau_{K,t}$ on capital income, a tax $\tau_{H,t}$ on labor income, an excise (intermediate-goods) tax $\tau_{I,t}$ on total energy and a tax $\tau_{E,t}$ on pollution emissions. We set the tax rates on capital and labor income in line with Trabandt and Uhlig (2012), who conduct a detailed analysis of fiscal policies in the U.S. and a number of European countries. Using a comprehensive measure of taxes on capital income, they find that on average, capital income in the U.S. is taxed at a rate of 41,4%, hence we set a time-invariant $\tau_K = 0.411$ in our baseline.[15] They find that labor income in turn, is taxed at a rate of 22.1%. Combined with a tax rate on consumption of 4.6%, this translates into a consumption-labor wedge of 25.5%, or $\tau_H = 1 - (1 - 0.221)/(1 + 0.046) = 0.255$. Turning to energy taxes, we follow Barrage (2019) and set the intermediate-goods tax at $\tau_I = 0$. Regarding the tax on pollution emissions τ_E, we set it at a level so that, in our calibrated economy, 3% of total energy is obtained from clean technologies (Nordhaus, 2017). This requires $\tau_E = 2.01\$/\text{tCO}_2$ in 2015.

To calibrate initial, outstanding debt B_0 at the start of the economy, we calculate the difference between total liabilities and financial assets from the U.S. government's balance sheet, both as a percentage of GDP.[16] Following Barrage (2019) and in order to facilitate reproducing results for other countries, these data are obtained from the IMF Government Finance Statistics. This gives an average debt-to-GDP ratio of approximately 111% over the period 2011–2015. Because in our model a period corresponds to five years, we set $B_0/Y_{1,0} = 1.11/5 = 0.222$.

Lastly, we require an empirical counterpart of government spending. In our model, G_t denotes government consumption of the final good, while T captures the present value of all lump-sum transfers households receive from the government. To better align the model with the data and to analyze business-as-usual scenarios, we follow Barrage (2019) and split up total government spending into final good spending G_t^C and *exogenous* transfers G_t^T that are provided to households. The total transfers households receive thus consist of this exogenous component G_t^T and the endogenous component T.[17] To obtain the empirical counterparts of G_t^C and G_t^T, we proceed as in Barrage (2019) and collect data on U.S. government expenses from the IMF Government Finance Statistics. Averaging over the years 2011–2015, government consumption is $G_0^C/Y_{1,0} = 0.158$, while government transfers are $G_0^T/Y_{1,0} = 0.145$.[18,19] To keep the sizes comparable to GDP going forward, both government consumption and

[15] Specifically, to obtain a comprehensive measure of capital tax rates, Trabandt and Uhlig (2012) adjust the personal income tax rate to account for income, profit and capital gains taxes of corporations, taxes on financial and capital transactions and recurrent taxes on immovable property. Similarly, to calculate labor income taxes, personal income taxes are adjusted to account for payroll taxes and social security contributions.

[16] The numbers are calculated at the "General Government" level.

[17] The endogenous component is set to $T = 0$ in Barrage (2019) and many other Ramsey tax models. The reason is that without heterogeneity, optimal policy would be to finance all spending through lump-sum taxes (*i.e.*, negative transfers), in which case tax distortions become irrelevant. In our model with heterogeneity, we do not have to impose this restriction.

[18] With these expenditure levels and the current tax system the intertemporal government budget constraint is not balanced. To balance the budget, taxes need to be raised in the future. We also consider an alternative calibration with the level of G_t rescaled to balance the budget with status-quo policies. All results discussed below remain unchanged, with only the average level of lump-sum transfers being affected.

[19] As in Barrage (2019), we include the following categories from the expense breakdown in G_t^C: compensation of

exogenous transfers grow at the sum of technological progress and population growth.

5 Quantitative results

We now present the optimal policy obtained under a utilitarian welfare criterion (*i.e.*, $\lambda_i = 1$ for all i), and the associated welfare effects compared to a "climate skeptic" planner scenario in which the planner ignores the anthropogenic origin of climate change and consequently sets the carbon tax to zero.[20]

5.1 Optimal policy

Optimal tax paths Figure 1 shows the path of optimal taxes on capital and labor income in our baseline scenario. The labor income tax roughly doubles in the first period, from 25% to about 50%, and stabilizes at this level. Rebating the revenue from these taxes via lump-sum transfers achieves most of the redistribution implied by the optimal tax system. Because lump-sum taxes are available and there is no initial wealth inequality, the only reason to tax capital income is to mitigate intertemporal distortions associated with labor income taxation. Since optimal labor income taxes are close to constant, the optimal capital income tax converges to zero quickly after the second period.[21] The next section examines scenarios with further constraints on policy instruments leading to deviations from this result.

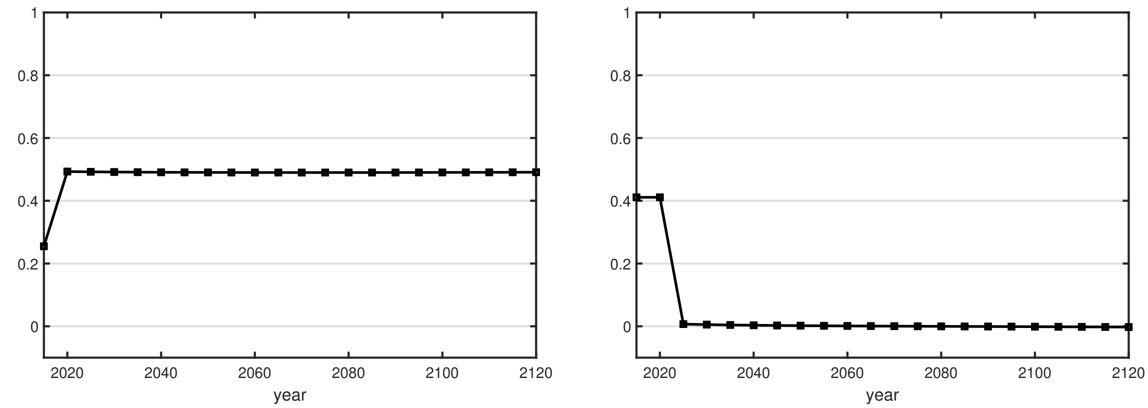
Figure 2 shows the optimal path of carbon taxes: in the baseline scenario, the tax starts at 21.7\$/tCO_2 in 2020 and goes up to reach 229.2\$/tCO_2 a century later. These tax levels are consistent with the ones found in Barrage (2019) and Nordhaus (2017, 2018), but are too low to contain climate change to a level consistent with the +2°C objective of the Paris agreement. In our "high damages" scenario, the optimal income taxes remain almost the same, but the carbon tax is roughly four times as large (see Appendix G.1).

Carbon tax decomposition Figure 3 compares the second-best pollution tax normalized to 1 (black line) to what it would be if the MCF was 1 in all periods (red line)—which also corresponds to the Pigouvian tax evaluated at the second-best allocation—and to what it would be ignoring inequalities (blue line). The MCF appears to play an insignificant role: the social cost of carbon is only 0.5% above the second-best carbon tax in the initial period, a difference that becomes even smaller in subsequent periods. Thus, even in the presence of distortionary taxation, it is optimal to set the carbon tax approximately at the social cost of carbon (*i.e.* at the Pigouvian level). However, the discrepancy between the blue and red lines indicates that the social cost of carbon itself is significantly affected by the presence of inequalities. The reason is that the social cost of carbon represents the monetary value of climate damages, and is determined by the arbitrage between reducing damages and increasing

employees, use of goods and services, subsidies, grants and other expense. For transfers G_t^T, we include social benefits.

[20] Details on the algorithm used to compute the Ramsey policy can be found in Appendix H.

[21] Notice that, because we have lump-sum taxation, the reason for zero long-run capital income taxation is different from the usual Chamley (1986) and Judd (1985), and is not subject to the criticism in Straub and Werning (2020).



(a) Optimal labor income taxes

(b) Optimal capital income taxes

Figure 1: Optimal Income taxes.

Notes: Figures show the path of second-best labor and capital income taxes for the baseline calibration. Initial tax rates (for 2015) are set exogenously to their current levels obtained from Trabandt and Uhlig (2012).

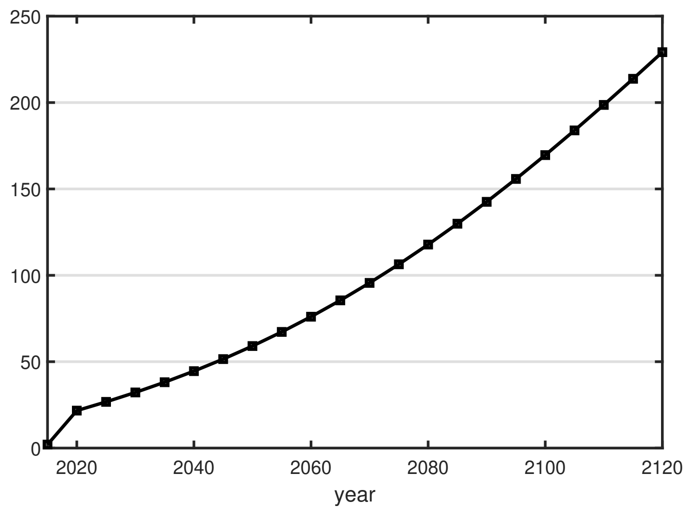


Figure 2: Optimal Carbon Taxes (\$/tCO_2).

Notes: Figure shows the path of second-best carbon taxes for the baseline calibration expressed in dollars per ton of CO_2. Initial level (for 2015) is set exogenously to its current level obtained from Nordhaus (2017).

aggregate consumption. As stated in Proposition 3, a marginal unit of aggregate consumption is valued more in the presence of inequalities if the marginal utility is sufficiently declining in consumption. Intuitively, an increase in aggregate consumption is valued less to the extent that it disproportionately goes in the hand of richer households, but it is valued more to the extent that the average marginal utility becomes higher if some people have relatively low consumption levels. In particular, with $\sigma = 1.45$, the IES is below unity, and ignoring consumption inequalities would lead to a social cost of carbon higher by on average 3.9% over the next century.

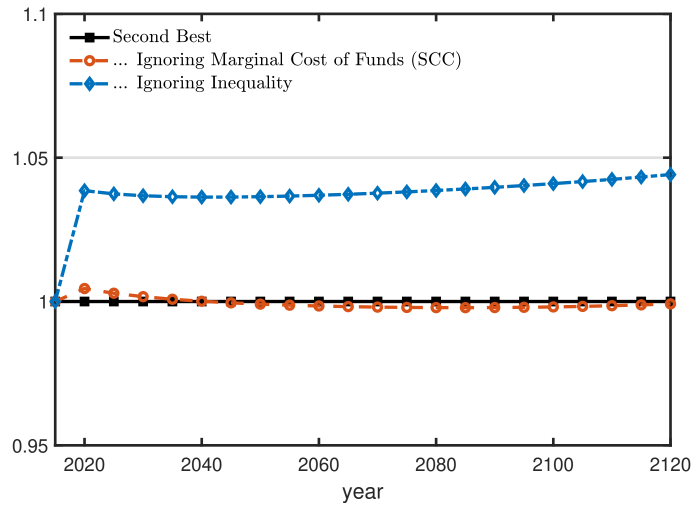


Figure 3: Carbon Tax Decomposition.

Notes: The black line represents the second-best carbon tax normalized to 1. The red line shows what this tax would be if the MCF was set to 1 in all periods, holding aggregates constant (see Proposition 2). The blue line shows what this tax would be absent consumption inequalities, again holding aggregates constant (see Proposition 3). All taxes are computed under the baseline calibration.

Sensitivity to calibration choices The level of government expenditures does not significantly affect the results. When choosing government expenditures such that current taxes are sustainable—at 22.5% instead of 30.3% of GDP—the effects of the MCF and inequalities are unaffected. In Appendix G.1, we also show that with a more severe calibration of climate damages leading to a SCC about four times higher, the role of the MCF remains negligible while the effect of inequalities decreases, at 2.6% instead of 3.9% in our baseline. This lower value is due to the lower share of utility damages at lower levels of warming (that result from higher carbon taxes). Figure 21b in Appendix G.4 illustrates this intuition: the figure plots the effect of inequalities on the optimal carbon tax for alternative values of the share of utility *vs.* production damages. When climate change impacts production only, inequalities have no effect on the optimal carbon tax. As the share of utility damages increases, the effect of inequalities rises, although at a decreasing rate. For instance, if 10% of damages were directly impacting utility at 2.5°C warming instead of 26% in the baseline, the effect of inequalities on the carbon tax would be 1.8% instead of 3.9%. If the share of utility damages was 40%, the effect of inequality would increase to 5.2%. Finally, we also consider different levels of heterogeneity in productivity (see Figure 21a in Appendix G.4). The effect on the optimal carbon tax appears relatively linear: it would be twice smaller if inequalities were twice lower than currently observed in the U.S.

As highlighted in Proposition 3, the effect of inequalities is sensitive to the value of σ, which in our dynamic framework with heterogeneous agents captures both the IES and the degree of inequality aversion of the planner. Figure 21c in Appendix G.4 plots the effect of inequalities on the optimal carbon tax for different values of σ. As stated in Proposition 3, the effect is null when σ tends to 1. For higher degrees of inequality aversion however (*i.e.* higher values of σ), the effect goes up non-linearly: with $\sigma = 2$, inequalities reduce the optimal carbon tax by 16.2%, instead of 3.9% with our baseline value of $\sigma = 1.45$ taken from DICE. So, credible alternative calibrations could lead to significantly stronger effects of inequality.

Fiscal adjustments relative to a climate skeptic planner Table I below reports the adjustments made to the government budget between our baseline scenario and a "climate skeptic" planner scenario in which the planner ignores the anthropogenic origin of climate change. Specifically, this climate skeptic planner sets all taxes optimally but behaves as if the climate variable was exogenous and not driven by human-made emissions. The objective of this experiment is to see how the planner should adjust the fiscal system once it acknowledges the necessity to address climate change. As shown in the table, the additional revenue provided by the carbon tax is split about equally between reducing distortionary taxes, with the present value of the labor tax decreasing by 0.7% of GDP, and increasing transfers, whose present value increases by 0.8% of GDP.[22] This finding violates the weak double-dividend hypothesis (for a review, see Goulder, 1995) according to which it is optimal to use the proceeds of the carbon tax to reduce distortionary taxes. With heterogeneous agents, distortionary taxes serve a redistributive purpose, hence it is not desirable to reduce them unless additional transfers can be provided through another mean. This result also gives some grounds to the popular carbon tax and dividend policy (see Economists Statement on Carbon Dividends, 2019) that calls for redistributing the proceeds of the tax lump-sum to address redistributive concerns, although we find that only half of the tax revenue should serve that purpose, the rest being aimed at improving economic efficiency.

Table I: Government Budget Adjustment.

	Revenue Source			Revenue Use		
	Labor	Capital	Carbon	Gov. Cons.	Transfer	Interest
No Carbon Tax	33.5%	0.6%	0.0%	17.2%	14.6%	2.3%
Optimal Carbon Tax	32.9%	0.6%	1.2%	16.9%	15.4%	2.3%
Change	−0.7%	0.0%	1.2%	−0.3%	0.8%	0.0%

Notes: Numbers represent the present value of each component of the government budget constraint divided by the present value of GDP, in the scenarios without carbon taxes (first row) and with carbon taxes (second row). The third row displays the difference between the two scenarios.

5.2 Welfare effects

Figure 4 displays the percentage increase in consumption that would be necessary in the climate skeptic scenario to make households as well-off as in the optimal scenario in each period and for each productivity group. While the average long run gains are positive for all productivity groups (the average

[22] The -0.3% change in government consumption expenditures reported in Table I results from the effect of carbon taxation on the present value of GDP since the expenditures are exogenous.

discounted gain is 5.8% with baseline damages), the period welfare gains are heterogeneous over time and between groups. Overall, welfare gains increase dramatically after the 21$^{\text{st}}$ century.[23] While they are initially progressively distributed, this pattern eventually reverses. The reason why the optimal carbon tax is progressive initially is that the revenue gains from carbon taxation are rebated through both a higher lump-sum transfer and a reduction in the labor income tax rate (see Table I). This contributes to an increase in the progressivity of the overall tax system, which makes poorer households benefit more (or suffer less) from the initial increase in carbon taxes. In the long run, richer households are the ones who benefit more from carbon taxation. A significant share of the welfare gains from a lower temperature come from reduced utility damages. Richer households care relatively more about those damages in the sense that, when the IES is below 1, they are willing to give up a higher *share* of their consumption for a reduction in temperature. This explains why in the long run, the welfare gains from carbon taxation are regressive when expressed in consumption units. Naturally, this exercise abstracts from heterogeneity in climate damages, an extension that we theoretically investigate in Section 6.4.

Income Decile \ Generation	2020	2030	2040	2050	2060	2070	2080	2090	2100	2110	2120	2130	2140	2150	2160	2170
1st	2.7	1.8	1.3	0.9	0.7	0.8	1.3	2.2	3.6	5.7	8.6	12.4	17.4	24.1	32.6	42.8
2nd	2	1.1	0.5	0.2	0	0.1	0.6	1.5	2.9	5	7.9	11.7	16.7	23.5	31.9	42.3
3rd	1.7	0.7	0.2	-0.2	-0.3	-0.2	0.3	1.2	2.6	4.7	7.6	11.5	16.6	23.4	31.9	42.4
4th	1.4	0.5	-0.1	-0.4	-0.6	-0.4	0	1	2.4	4.6	7.5	11.4	16.5	23.4	32.1	42.7
5th	1.2	0.3	-0.3	-0.6	-0.7	-0.6	-0.1	0.8	2.3	4.5	7.4	11.4	16.6	23.6	32.4	43.1
6th	1	0.1	-0.4	-0.8	-0.9	-0.8	-0.3	0.7	2.2	4.4	7.4	11.4	16.7	23.8	32.7	43.6
7th	0.8	-0.1	-0.6	-0.9	-1.1	-0.9	-0.4	0.5	2.1	4.3	7.4	11.5	16.8	24.1	33.2	44.3
8th	0.7	-0.2	-0.8	-1.1	-1.2	-1.1	-0.6	0.4	2	4.3	7.4	11.6	17.1	24.6	33.9	45.4
9th	0.5	-0.4	-1	-1.3	-1.4	-1.3	-0.7	0.3	2	4.4	7.6	11.9	17.7	25.3	35.1	47.1
10th	0	-0.9	-1.4	-1.7	-1.8	-1.5	-0.9	0.4	2.3	5.1	8.9	13.9	20.7	29.8	41.4	56.1

Figure 4: Period Welfare Gains (%).

Notes: For each decade and each income decile the table shows the welfare gains, in percentage of consumption, from optimal carbon taxation relative to a scenario without carbon taxation. Numbers are computed under the baseline calibration.

While carbon tax policies are often considered unpopular because of their potentially regressive effects, we find that an optimal carbon policy—*i.e.*, combined with optimal adjustments of income taxes and transfers—is actually progressive. Thus, although the expected gains from carbon taxation will disproportionately benefit future generations, the optimal carbon tax policy still benefits poor households in the present, which could make the policy more attractive to a government concerned with redistribution and increase public support in the first stages of the policy implementation.

[23] As shown in Dietz et al. (2021), the DICE model features too much thermal inertia, *i.e.*, the temperature response to an impulse in emissions is delayed too much compared to what climate science models predict. If this response was more immediate, welfare gains from carbon taxation could become positive earlier.

6 Extensions

In this section, we consider different extensions of our baseline results. First, we explore fiscal environments with additional constraints on the set of available instruments. Then, we study in turn the effects of introducing inequality in wealth, energy consumption, and sensitivity to environmental damages.

6.1 Third-best policies

We have considered a Ramsey problem in which the government faces two key constraints: only linear and anonymous instruments can be used. Still, this set of fiscal instruments confers a lot of power to the government, arguably more than what most governments have. When introducing an environmental tax policy, a government may not have complete freedom to adjust labor or capital income taxes.

6.1.1 Third-best tax formulas

Exogenous labor income tax Let us assume that the planner cannot choose the labor income tax, that is exogenously fixed at a level $\bar{\tau}_H$ in all periods $t \geq 0$. The planner now faces additional constraints: in every period $t \geq 0$, it must ensure that

$$\frac{U_{h,t}}{U_{c,t}} = -\left(1 - \bar{\tau}_H\right)\left(1 - D_t\right) A_{1,t} F_{H,t}, \tag{25}$$

which pins down the wedge between the marginal rate of substitution between consumption and leisure and the marginal product of labor. For a given value of $\bar{\tau}_H$, equation (25) puts a restriction on the implementable allocations that the planner must satisfy. Let $\beta^t \Lambda_t^H$ denote the multiplier on the constraint (25). The latter is proportional to the welfare impact of raising the exogenous $\bar{\tau}_H$ in a particular period. The multiplier Λ_t^H will be positive (resp. negative) on average if the labor income tax is fixed at a sub-optimally high (resp. low) level. With the additional constraint (25) in each period t, the expression for the optimal pollution tax becomes[24]

$$\tau_{E,t} = \frac{1}{\nu_{1,t}} \sum_{j=0}^{\infty} \beta^j \left(\nu_{1,t+j} D'_{t+j} A_{1,t+j} F_{t+j} - N_{t+j} V_{Z,t+j} + \Lambda_{t+j}^H \left(1 - \bar{\tau}_H\right) D'_{t+j} A_{1,t+j} F_{H,t+j} \right) J_{E_t^M, t+j}. \tag{26}$$

where, as in Section 3, $\nu_{1,t}$ is the multiplier on the aggregate resource constraint in period t, which measures the scarcity of consumption goods and hence, the opportunity costs of reducing emissions. Compared to equation (20), the main modification is the final component, which Barrage (2019) refers to as the *fiscal interaction* term. It reflects another reason for deviating from the Pigouvian tax rule. By reducing production damages, a higher pollution tax $\tau_{E,t}$ raises the marginal product of labor and hence, the before-tax wage. If τ_H is fixed at a sub-optimally low level, a further increase in the before-tax wage is welfare-reducing. The pollution tax then amplifies the costs of having a tax on labor income

[24] Without constraint (25), it is optimal to equalize the marginal rate of technical substitution between capital and labor across both sectors: the government does not wish to distort production decisions. In the third best, with constraint (25), this is no longer the case, and it is optimal to deviate from zero excise energy taxes, $\tau_{I,t}$. See Appendix B for more details.

that is below the welfare-maximizing level. Consequently, the optimal pollution tax is reduced. The fiscal interaction term thus calls for a lower pollution tax when the labor income tax is fixed at a sub-optimally low level and *vice versa* if the labor income tax is fixed at a sub-optimally high level.

Exogenous capital income tax Let us now assume that the planner cannot choose the capital income tax, that is exogenously fixed at a level $\bar{\tau}_K$ in all periods $t \geq 0$. The new constraints faced by the planner are such that in every period $t \geq 0$,

$$\frac{U_{c,t}}{U_{c,t+1}} = \beta\Big(1 + (1 - \bar{\tau}_K)\left((1 - D_{t+1})\, A_{1,t+1} F_{K,t+1} - \delta\right)\Big), \tag{27}$$

which links the marginal rate of substitution between consumption in periods t and $t+1$ (on the left-hand side) to the after-tax interest rate (on the right-hand side). As with an exogenous labor income tax, equation (27) restricts the set of implementable allocations for a given value of $\bar{\tau}_K$. Let $\beta^t \Lambda^K_{t+1}$ be the multiplier on this constraint in period t. The multiplier is positive (negative) if the capital income tax rate is fixed at a sub-optimally high (low) level, so that raising $\bar{\tau}_K$ in a particular period lowers welfare. With the additional constraint (27), the expression for the optimal pollution tax is modified to:

$$\tau_{E,t} = \frac{1}{\nu_{1,t}} \sum_{j=0}^{\infty} \beta^j \Big(\nu_{1,t+j} D'_{t+j} A_{1,t+j} F_{t+j} - N_{t+j} W_{Z,t+j} + \Lambda^K_{t+j} (1 - \bar{\tau}_K)\, D'_{t+j} A_{1,t+j} F_{K,t+j}\Big) J_{E^M_t,t+j}, \tag{28}$$

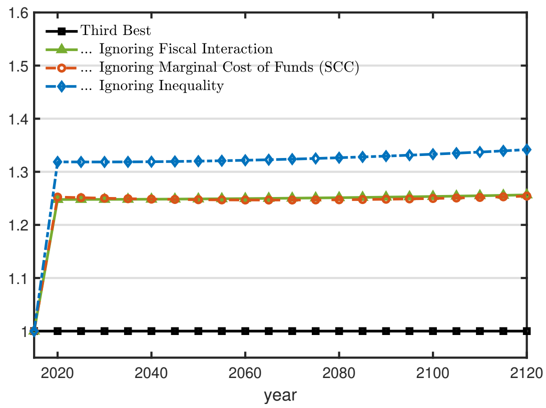
where again the last component captures the fiscal interaction term. The intuition is similar as before. A higher pollution tax raises the marginal product of capital by lowering production damages. The latter is beneficial if the capital income tax is fixed at a sub-optimally high level. A higher pollution tax then alleviates the savings distortion by raising the before-tax interest rate. If, by contrast, the capital income tax is fixed at a level below the one that maximizes welfare, a pollution tax amplifies the savings distortion and the fiscal interaction term reduces the optimal pollution tax.

6.1.2 Quantitative analysis

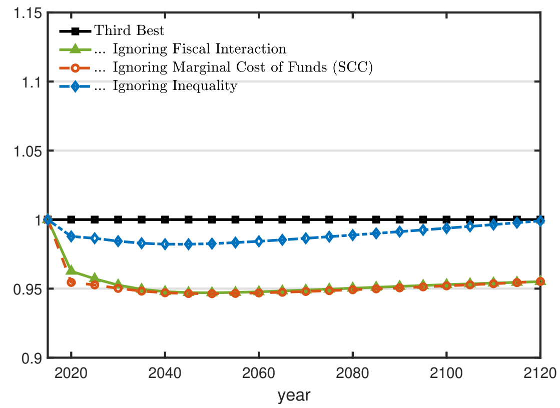
Figure 5 below compares the third-best pollution tax normalized to 1 (black line) with what it would be ignoring the new fiscal interaction term (green line), ignoring the MCF (red line), and ignoring inequalities (blue line).[25] As in our benchmark scenario, the MCF plays an insignificant role but inequalities push the carbon tax downward. The effect of inequalities is slightly larger when the labor income tax is fixed: ignoring inequalities would increase the tax by around 6% in this scenario instead of about 4% in the second-best and in the scenario where the capital tax is fixed. Indeed, since $\bar{\tau}_H$ is set to 25.5%, *i.e.* below the second-best tax rate, there are more consumption inequalities than in the second-best and the opportunity cost of emission abatement is higher.

While the MCF still plays a negligible role, fiscal interactions now drive the carbon tax away from its Pigouvian level through the additional constraints that arise in the third-best environment.

[25] Appendix G.2 presents figures for the optimal path of income and carbon taxes in the third-best scenarios.



(a) Given labor income tax



(b) Given capital income tax

Figure 5: Third-Best Carbon Tax Decomposition.

Notes: The black line represents the second-best carbon tax normalized to 1. The green line shows what this tax would be without the fiscal interaction term, holding aggregates constant—for (a) this is the last term in equation (26) and for (b) the last term in equation (28). As in Figure 3, the red and blue lines display the effects of the MCF and inequalities respectively, relative to the green line. All taxes are computed under the baseline calibration.

Interestingly, the fiscal interaction term lowers the optimal carbon tax when the labor income tax is fixed, whereas it raises the optimal carbon tax when the capital income tax is fixed. Recall that a carbon tax, by reducing production damages, increases *both* the marginal product of labor and the marginal product of capital and hence, the before-tax wage and interest rate. A higher before-tax wage, in turn, lowers welfare because the labor income tax is set at a sub-optimally low level (*i.e.*, $\bar{\tau}_H = 25.5\%$ instead of around 49% at the optimum), whereas a higher before-tax interest rate raises welfare because the capital income tax is set at a sub-optimally high level (*i.e.*, $\bar{\tau}_K = 41.1\%$ instead of virtually 0% at the optimum). A higher carbon tax thus alleviates the savings distortion, whereas it amplifies the costs of taxing labor income at a sub-optimally low level. This explains why quantitatively we find that the fiscal interaction term is positive when the capital income tax is fixed, and negative when the labor income tax is fixed.

Appendix G.2 also provides the government budget adjustments and welfare gains in these third-best policy scenarios. These results suggest that the general pattern of the distribution of welfare gains from carbon taxation does not strongly depend on the fiscal policies currently in place, but the optimal use of the carbon tax revenue does. While this revenue is split about equally between increasing transfers and reducing the labor income tax in our baseline scenario, with additional constraints on instruments this is not the case anymore. In particular, when the government is forced to redistribute "too little" because labor income taxes are set below the optimum, the carbon tax revenue is mostly targeted towards redistribution, leading to more progressive effects.

6.2 Initial wealth inequality

In this section, we consider the effect of initial wealth inequality on the optimal tax system. When the planner is allowed to set the initial tax on capital income, it is optimal to fully expropriate initial wealth (if less productive households are also less wealthy). To study the implications of wealth inequality on optimal fiscal policy, we therefore assume that the planner is unable to set the capital income tax in the first period, *i.e.* $\tau_{K,0}$ is exogenous. We discuss the optimal rules and investigate the quantitative effects given the levels of wealth inequality observed in the U.S. In Appendix C.1, we also discuss the implications of initial wealth inequality for the time-consistency of Ramsey policies.

6.2.1 Optimal tax rules

For $t \geq 1$, the optimal tax rules are not affected by the presence of initial wealth inequality.[26] However, if $\tau_{K,0}$ cannot be chosen to eliminate initial wealth inequality, there is another reason for deviating from Pigouvian taxation in period 0. Let Δ denote the shadow cost of wealth inequality,

$$\Delta \equiv \sum_i \pi_i \theta_i a_{i,0},$$

then, the optimal period-0 pollution tax is given by (see Appendix C):

$$\tau_{E,0} = \frac{1}{\nu_{1,0}} \left(\sum_{j=0}^{\infty} \beta^j \left(\nu_{1,j} D'_j A_{1,j} F_j - N_j W_{Z,j} \right) J_{E_0^M,j} - N_0 U_{c,0} \Delta \left(1 - \tau_{K,0}\right) D'_0 A_{1,0} F_{K,0} J_{E_0^M,0} \right), \tag{29}$$

where

$$\nu_{1,0} = W_{c,0} - U_{cc,0} R_0 \Delta$$

is the planner's multiplier on the aggregate resource constraint.

Notice that wealth inequality, through Δ, affects pollution taxation in period zero via two mechanisms: (1) it implies an additional term, the last one in equation (29); and (2) it affects the planner's valuation of a unit of consumption in period 0. First, the additional term has to do with the fact that higher damages reduce interest rates which, as a side-effect, mitigates wealth inequality, calling for lower pollution taxes. This is a very subtle effect and quantitatively this term is small. Second, the effect on $\nu_{1,0}$ is a result of the fact that we do not allow full expropriation of initial wealth, which could be achieved by increasing $\tau_{K,0}$ so that $R_0 = 0$. We instead fix $\tau_{K,0}$ and this constraint is equivalent to having the planner expropriate all initial wealth and then return the amount assigned to each household. When more productive households have higher wealth, this is costly for the planner, so $\Delta > 0$. The opportunity cost of abatement given by $\nu_{1,0}$ is then higher to the extent that increasing aggregate consumption would lower the initial price and exacerbate this cost. This effect leads to a substantial reduction in period-0 pollution taxes. Similarly to inequalities in productivity, wealth inequalities reduce the optimal pollution tax, although the effect is concentrated in the first period.

[26] The exception is the tax rule for $\tau_{K,1}$. See Appendix C for details.

6.2.2 Quantitative analysis of the effect of wealth inequality

We calibrate the joint distribution of productivity and initial wealth from the SCF. We divide households into 10 productivity groups, and 10 wealth groups within each productivity group, for a total of 100 different groups of equal size. The full procedure is described in Appendix F.1. We fix $\tau_{K,0}$ to be at the same level as in the current tax system, at 41.1%.

Figure 6 below provides a decomposition similar to the one shown in Figure 3 above.[27] While the effects of the MCF and consumption inequalities remain similar to the baseline, wealth inequalities call for a significant reduction of the optimal tax in the first period (green line). This effect is fully driven by the second mechanism described above, *i.e.* the higher value to the planner of an extra unit of consumption in period 0, $\nu_{1,0}$. It should however be noted that this temporary decrease in the optimal carbon tax is accompanied by an equivalent increase in the energy tax, τ_I, that mitigates losses in productive efficiency.

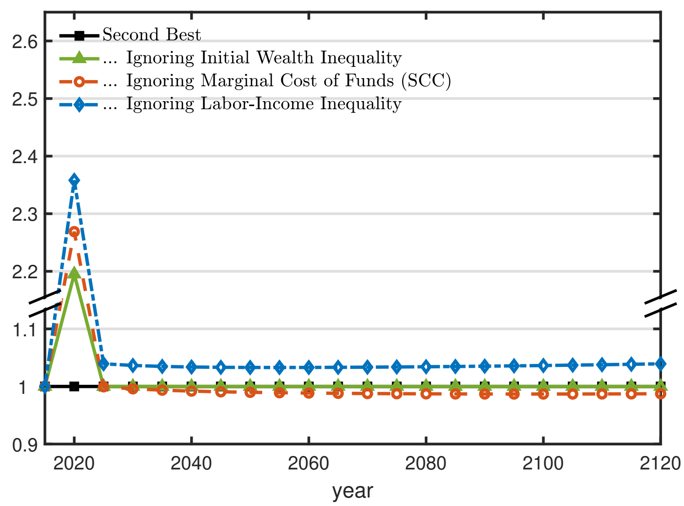


Figure 6: Carbon Tax Decomposition, Initial Wealth Heterogeneity and Exogenous Initial Capital Tax.

Notes: The black line represents the second-best carbon tax normalized to 1. The green line shows what this tax would be without wealth inequality, holding aggregates constant—more precisely, it shows what happens to $\tau_{E,t}$ in equation (29) if Δ is set to zero. As in Figure 3, the red and blue lines display the effects of the MCF and inequalities respectively, relative to the green line. All taxes are computed under the baseline calibration.

[27] Appendix G.3 includes figures for the optimal path of income and carbon taxes with initial wealth heterogeneity when the initial capital tax is fixed at its current level. The appendix also contains a table showing the government budget adjustments made relative to the climate skeptic planner and a figure that displays the distribution of the lifetime welfare gains for each of the 100 groups. These gains are U-shaped with respect to income, but strictly increasing with initial wealth.

6.3 Energy consumption inequality

6.3.1 Optimal tax rules

Our benchmark model considers heterogeneous households who differ in productivity and initial asset holdings. To further explore the role of households' heterogeneity on optimal fiscal policy, we now introduce into our benchmark model a second *dirtier* consumption good modeled as a necessity.

Two-goods economy Formally, we assume that a household of type i derives utility from the consumption of a final good $c_{i,t}$, a dirtier good $d_{i,t}$, labor supply $h_{i,t}$, and environmental degradation Z_t according to a utility function

$$\sum_{t=0}^{\infty} N_t \beta^t u_i \left(c_{i,t}, d_{i,t}, h_{i,t}, Z_t\right),$$

where the second dirtier good d is produced from a linear technology that uses energy as its only input. To further simplify notations, we assume that energy produced in the energy sector (E_t) is now used in the final good sector or directly consumed by households, such that

$$E_t = E_{1,t} + N_t d_t,$$

with $E_{1,t}$ the quantity of energy used as an input in the final good sector and $d_t = \sum_i \pi_i d_{i,t}$ the households' average per period energy consumption. In order to match empirically observed budget shares for energy (or alternatively, polluting goods) for different income groups, we assume households' utility can be represented by the following period utility function

$$u_i \left(c_i, d_i, h_i\right) = \frac{\left(c_i (d_i - \bar{d}_i)^{\epsilon} (1 - \varsigma h_i)^{\gamma}\right)^{1-\sigma}}{1-\sigma} + \frac{\left(1 + \alpha_0 Z^2\right)^{-(1-\sigma)}}{1-\sigma}. \tag{30}$$

Thus, in line with previous studies in this literature (*e.g.* Fried et al., 2018; Klenert et al., 2018; Aubert and Chiroleu-Assouline, 2019; Jacobs and van der Ploeg, 2019) preferences for consumption are modeled with a Stone-Geary utility function, so that an agent of type i experiences positive utility from energy consumption only after consuming its first $\bar{d}_i$ units of energy. $\bar{d}_i$ therefore denotes the subsistence consumption level of energy for an agent of type i, which we allow to be type (and time) specific. This specification allows us to consider households with non-homothetic preferences to better capture the heterogeneous impact of pollution taxes on households' budgets. Assuming type-specific values for $\bar{d}_i$, this specification also allows us to potentially consider non-linear *aggregate* Engel curves as well as horizontal heterogeneity.[28,29]

[28]With Stone-Geary preferences, agents' Engel curves are linear. When preferences are heterogeneous, the aggregate distribution of expenditures may however be a non-linear function of income.

[29]Horizontal heterogeneity arises when households with the same income do not consume goods in the same proportions. Recent studies have shown the importance of horizontal heterogeneity on the distributional impacts of energy taxes in the U.S. (Cronin et al., 2019; Pizer and Sexton, 2019), and their implications for the design of tax reforms (Sallee, 2019).

Solution method Because there is an additional consumption good, the planner uses an additional instrument: it levies an excise tax $\tau_{D,t}$ on households' consumption of energy. The budget constraint of agents of type i can thus be expressed as

$$\sum_{t=0}^{\infty} p_t N_t \Big(c_{i,t} + d_{i,t}(p_{E,t} + \tau_{D,t}) - (1 - \tau_{H,t})\, w_t e_i h_{i,t} \Big) \leq R_0 N_0 a_0 + T. \tag{31}$$

To focus on the additional sources of heterogeneity, we assume here that there is no initial wealth inequality, so that $a_{i,0} = a_0$ for all i. We apply the same solution method as in our benchmark model. Following Werning (2007), we can express individual allocations as a function of aggregate variables and market weights. These expressions allow us to write the aggregate utility function $U(c_t, d_t, h_t, Z_t, \varphi)$ and individual implementability conditions necessary to solve the Ramsey problem based on aggregate variables and market weights only.

Optimal tax formulas Propositions 4 and 5 below state the role of preferences for the additional polluting commodity on the optimal taxation of pollution and energy consumption respectively.

Proposition 4 *If agents' utility is given by (30), the optimal pollution tax can be expressed as (22), i.e. a modified Pigouvian rule that accounts for the MCF given by*

$$\mathrm{MCF}_t = 1 + \frac{\mathrm{cov}(\theta_i, IC_{c,i,t})}{V_{c,t}},$$

with

$$IC_{c,i,t} = (1-\sigma)U_{c,t}\left((1+\gamma+\epsilon)\omega_i - \gamma \frac{e_i}{1-\varsigma h_t} + \epsilon \frac{\bar{d}_{i,t}}{d_t - \bar{d}_t} \right).$$

From period 0, the welfare-weighted average MCF is 1. In the limit case where the IES tends to 1, then for $t \geq 0$, $\mathrm{MCF}_t = 1$.

Proposition 4 (see proof in Appendix D.4) states that the additional dirty good affects the optimal pollution tax only through the MCF: temporal variations in energy needs affect households' budget constraint, thereby affecting the planner's implementation cost over time. This mechanism might cause temporal fluctuations in the MCF, but it does not affect its the long-term average value that remains equal to 1 as in the benchmark. In addition, as in Proposition 2 above, when the IES tends to 1, the price and volume effects from an increase in aggregate consumption exactly offset each other, so households' expenditures net of income remain unaffected by a marginal increase in the pollution tax and the MCF is equal to 1 in all periods.

Proposition 5 *If agents' utility is given by (30), then*

$$\tau_{D,t} = \frac{\Lambda_t \epsilon \frac{c_t}{(d_t - \bar{d}_t)^2}}{\Phi + \frac{\Psi \varsigma \gamma(\sigma-1)}{(1-\varsigma h_t)} - \frac{\Lambda_t \epsilon(\sigma-1)}{(d_t - \bar{d}_t)}},$$

with

$$
\begin{aligned}
\Phi &= \sum_j \pi_j \frac{\lambda_j}{\varphi_j} + \Big(1 - (1+\epsilon+\gamma)(1-\sigma)\Big)\mathrm{cov}(\lambda_i/\varphi_i, \omega_i),\\
\Psi &= -\frac{\mathrm{cov}(\lambda_i/\varphi_i, e_i)}{\varsigma},\\
\Lambda_t &= -\mathrm{cov}(\lambda_i/\varphi_i, \bar{d}_{i,t}),
\end{aligned}
$$

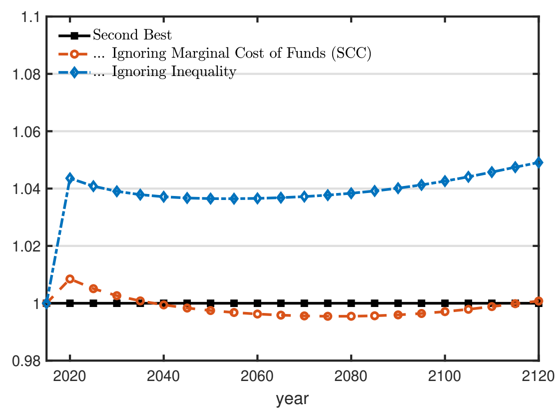
hence the energy good is subsidized if and only if the agents who are valued relatively more by the planner compared to the market (higher λ_i/φ_i) have higher energy needs ($\bar{d}_{i,t}$).

A corollary of Proposition 5 (see proof in Appendix D.4) is that, when preferences for the energy good are homogeneous, the optimal excise tax on this good is zero. We also show in Appendix D.4 that in this case, the explicit formulas for labor, capital, and energy input taxes are unchanged relative to the benchmark model. Thus, although poor households spend a larger share of their budget in the polluting energy necessity, the optimal tax formulas are the same as in the benchmark model. This result is reminiscent of Jacobs and van der Ploeg (2019) who show that as long as Engel curves are linear—which is the case with Stone-Geary utility—corrective taxation should not serve to address redistributive objectives, even when non-linear income taxation is not available. Still, the optimal tax levels might differ from the benchmark due to differences in allocations: having a second good modeled as a necessity generates a fixed-cost to households' utility, which exacerbates inequalities.

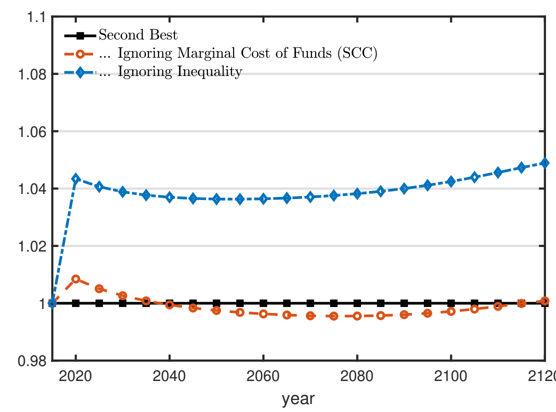
In the general case where preferences differ between agents, the *aggregate* Engel curves are non-linear, hence commodity taxes offer an additional levy for redistribution. When the agents who are valued relatively more by the planner also have higher energy needs, the planner can target these agents by subsidizing the energy good. The sign and magnitude of this mechanism therefore depend on the distribution of $\{\bar{d}_i\}_{i\in I}$, both between and within productivity types. First, as less productive types tend to have higher marginal utilities of consumption, the relative planner's weights are generally higher for these agents. The excise tax will therefore be higher to the extent that more productive agents have on average higher energy needs. Second, for a given productivity level, agents with higher energy needs will also tend to have higher marginal utilities of consumption because of the higher fixed cost that they incur. This horizontal heterogeneity will therefore drive the value of the excise tax downward. Our quantitative analysis below uses data on U.S. households' energy consumption to illustrate the impact of these two sources of heterogeneity.

6.3.2 Quantitative analysis of the extended model

Calibration choices To calibrate this extended model, we choose our parameters to meet two additional targets: the share of households' expenditures on the energy good, and the share of aggregate emissions coming from households' energy consumption. Using the model's first order conditions, we show in Appendix F.1 that ϵ can be expressed as a coefficient in a regression where households' energy and total expenditures are the only variables to observe. The distribution of these variables is obtained from the Consumer Expenditure Surveys (CEX), where energy expenditures corresponds to the sum of households' expenditures on energy used for transport and housing. We first use this data to compute the value of ϵ, and set the initial value of $\bar{d}$ to target an average energy expenditure share of 10.8% as observed in the CEX. We then use the value of ϵ to compute 15 type-specific subsistence levels $\bar{d}_i$ to match the observed distribution of energy expenditure shares across and within income quintiles. For the share of emissions coming from households' energy consumption, we target 30%, which represents the share of emissions coming from the residential sector and households' transportation.[30] To do so, we adjust the energy share in the final good production function ν from 0.04 to 0.17. Although this may seem like a significant change compared to our benchmark, we confirm that using this higher value of ν would not affect our results in the benchmark model. To remain consistent with our previous targets for the initial labor supply, Frisch elasticity, initial emissions, share of utility damages, and capital share relative to labor, we also adjust the values of ς, γ, $A_{2,2015}$, α_0, and α. The full procedure is described in Appendix F.1, where we also report the value of the adjusted parameters and the distribution of households' energy consumption shares.



(a) Equal energy necessity



(b) Unequal energy necessity

Figure 7: Carbon Tax Decomposition With and Without Energy Necessity Inequality.

Notes: The black line represents the second-best carbon tax normalized to 1. The red and blue lines display the effects of the MCF and inequalities (both in productivity and in energy necessity) respectively.

Results Figure 7 shows a decomposition similar to the one shown for the baseline in Figure 3 above, for the case where (a) households have identical energy necessity levels and (b) heterogeneous energy necessity levels. We see that in both cases the MCF has again a negligible impact on the second-best carbon tax. For the two scenarios, the role of inequalities is also very comparable to the baseline of 3.9%, both at 4.1%. While we could expect that the presence of a necessity—which is akin to a fixed-cost to households' consumption—would further increase the effect of inequalities on the carbon tax, this is in fact mitigated by an increase in transfers financed by a higher labor tax.

[30]The U.S. EPA reports that, in 2013 (our initial period), 17% of U.S. emissions were due to the residential sector, 11% to passenger cars, and 5% to light-duty trucks such as pickups, minivans, and SUVs (see EPA, 2017, Tables 2-12 and 2-13). Assuming households' are directly responsible for the largest part of these emissions, the emissions coming from households' energy consumption represent about 30% of U.S. aggregate emissions.

Introducing heterogeneous necessity levels has negligible effects on the optimal carbon tax. From our calibration, we see that on the one hand the necessity level is on average higher for richer households, which reduces inequalities, but on the other hand horizontal heterogeneity (*i.e.* differences in necessity levels within productivity groups) increases inequalities. As a result, households' necessity levels do not strongly co-vary with their marginal utility of consumption, so heterogeneity in necessity barely affects the carbon tax rates. For the same reason, optimal excise taxes on energy consumption are very small, amounting to about -0.4% of energy prices in every period—it is exactly zero when there is no heterogeneity in necessity levels.

Figure 8 displays the inter-temporal welfare gains from carbon taxation for each category. Between income groups, we observe a U-shape pattern: while the poorest households benefit relatively more from the increase in tax progressivity, the richest households benefit relatively more from future environmental improvements that they value proportionally more. Within income groups, we see that households with lower energy needs benefit relatively more, as they pay relatively less of the carbon tax while still enjoying the revenue-recycling and mitigation benefits.

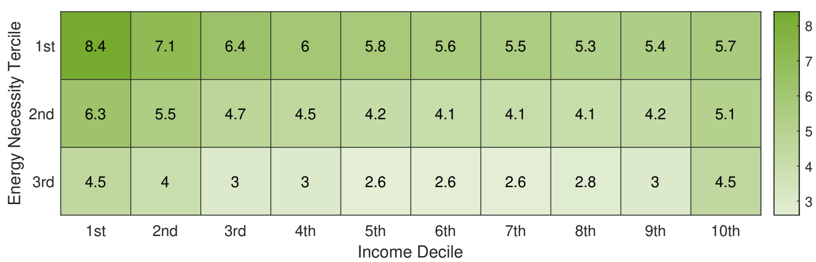


Figure 8: Welfare Gains (%), Energy Necessity Inequality.

Notes: For each income decile and expenditure share tercile, the table shows the discounted welfare gains, in percentage of consumption, from optimal carbon taxation relative to a scenario without carbon taxation.

6.4 Heterogeneous sensitivity to environmental damages

Several recent studies have highlighted that the impact of environmental degradation is heterogeneous across individuals, and likely more negative for those more financially deprived (for recent reviews, see Banzhaf et al., 2019; Hsiang et al., 2020). In the case of climate change, higher exposure to extreme temperatures and weaker adaptation means make poorer households on average more vulnerable (see *e.g.*, Dell et al., 2012; Ricke et al., 2018). While heterogeneity in income, wealth, and consumption patterns are key to explain the unequal burden from environmental policies, we now introduce heterogeneous exposure to environmental degradation to account for the unequal benefits from pollution mitigation. Formally, we again assume that utility is strongly separable in Z and additionally consider households with heterogeneous sensitivity to environmental degradation, so that agent's i utility function can be expressed as

$$u_i\left(c_{i,t}, h_{i,t}, Z_t\right) \equiv \tilde{u}\left(c_{i,t}, h_{i,t}\right) + \hat{u}_i(Z_t). \tag{32}$$

While production damages still arise at the aggregate level, households are heterogeneously affected by environmental degradation directly through their utility. In the context of climate change, this may capture heterogeneous effects on people' health, exposure to conflicts, forced re-settlement, or losses in various forms of amenity values. Although these types of damages may also affect households' productivity, we abstract from heterogeneous impacts of Z on agents' productivity e_i to keep the problem sufficiently tractable.

When environmental degradation Z heterogeneously affects households' utility, the optimal pollution tax can still be expressed as the modified Pigouvian rule stated in Proposition 2, but the term $V_{Z,t}$ entering the Pigouvian formula—that captures the marginal dis-utility from environmental degradation for the planner—now depends on the joint distribution of utility damages and the planner's welfare weights,

$$V_{Z,t} = \sum_i \pi_i \hat{u}_i'(Z_t) + \mathrm{cov}\big(\lambda_i, \hat{u}_i'(Z_t)\big).$$

Proposition 6 and corollary 1 state the role of heterogeneous utility damages on the optimal pollution tax.

Proposition 6 *If environmental utility is strongly separable from consumption and leisure as in (32), heterogeneity in the marginal pollution damages to utility increases the pollution tax if and only if the planner's weights are positively correlated with marginal utility damages.*

Corollary 1 *If environmental utility is strongly separable from consumption and leisure as in (32) and the planner is utilitarian, heterogeneity in the marginal pollution damages to utility has no effect on the optimal pollution tax. If the planner is Rawlsian, the pollution tax is higher if and only if agents with the lowest welfare experience larger marginal utility damages from pollution.*

The logic behind Proposition 6 (see proof in Appendix E) is that, as long as environmental welfare is additively separable from consumption and leisure, marginal utility losses from environmental damage for the rich and the poor are perfect substitutes for the planner. Corollary 1 additionally states that if the planner is utilitarian, then an extra unit of utility is valued equivalently for all households, hence the distribution of environmental damage is inconsequential for the planner when setting the optimal pollution tax. When the planner gives a higher direct value to households who are worse-off however, the environmental utility damages experienced by these households are valued more by the planner. In the extreme case where the planner cares only about the household with the lowest-welfare, the planner determines the optimal pollution tax to internalize pollution on that household only, and sets it to a higher level if this household is more exposed.

The previous results rely on an important separability assumption. If this assumption was relaxed, then even in the utilitarian case heterogeneous climate damages could call for a higher pollution tax if households with lower consumption levels were more impacted by pollution. Similarly, we have abstracted from heterogeneous productivity damages that could reinforce the case for higher pollution taxes even in the utilitarian case. The theoretical and quantitative exploration of these mechanisms goes beyond the scope of this paper, but we see this topic as critical for future research.[31]

[31] As explained in Hsiang et al. (2020), heterogeneous impacts from climate change are difficult to measure as they are

7 Conclusion

Should environmental policies be less stringent in the presence of inequalities? Do inequalities increase when optimal environmental policies are implemented? This paper attempts to shed light on these questions. We develop a climate-economy model where environmental degradation generates both production and utility externalities. Our model features heterogeneous agents, which provides a micro-foundation for the use of distortionary taxes on labor and capital income. We study both theoretically and quantitatively how different sources of heterogeneity and a concern for redistribution affect the optimal carbon tax.

We show that when households are heterogeneous but individualized lump-sum taxation is not available, the optimal carbon tax is approximately equal to the social cost of carbon (SCC), but the SCC is lower than it would be absent inequalities. Indeed, tax distortions do not significantly matter for carbon taxation when distortionary taxes are optimally chosen to provide redistribution, and the optimal carbon tax is approximately Pigouvian. However, inequalities call for lower carbon taxes owing to the fact that the presence of poor households increases the marginal value of consumption and increases the opportunity cost of pollution abatement. We also re-examine the double-dividend hypothesis, and show that at the optimum the carbon tax revenue is divided about equally between increasing transfers and reducing distortionary taxes. This revenue recycling increases the progressivity of the tax system, making the carbon tax policy relatively more beneficial for poorer households. In the long run however, rich households experience larger welfare gains from climate change mitigation because their willingness to pay for environmental improvement is higher.

Our paper includes numerous extensions to study the implications of inequality for optimal carbon taxation. We analyze alternative policy scenarios, and multiple sources of household heterogeneity, including heterogeneous budget shares, unequal initial assets, and differences in the sensitivity to environmental damages. Still, there are other relevant aspects that deserve further investigation. In particular, we have left for future research the role of risk—on the economic or climate side—which could interact with inequalities and be an important determinant of fiscal policies. In addition, it would be interesting to further explore theoretically and quantitatively the role of heterogeneous damages of climate change.

determined by several sources of heterogeneity that are hard to disentangle: heterogeneity in initial climatic conditions, in their evolution, in individuals' response to these changes, as well as in other individuals' characteristics influencing their welfare impacts.

References

Aubert, Diane and Mireille Chiroleu-Assouline (2019) "Environmental tax reform and income distribution with imperfect heterogeneous labour markets," *European Economic Review*, 116, 60–82, https://doi.org/10.1016/j.euroecorev.2019.03.006.

Banzhaf, Spencer, Lala Ma, and Christopher Timmins (2019) "Environmental justice: The economics of race, place, and pollution," *Journal of Economic Perspectives*, 33 (1), 185–208.

Barrage, Lint (2019) "Optimal Dynamic Carbon Taxes in a ClimateEconomy Model with Distortionary Fiscal Policy," *The Review of Economic Studies*, 87 (1), 1–39, 10.1093/restud/rdz055.

Bovenberg, A. Lans and Lawrence H. Goulder (1996) "Optimal Environmental Taxation in the Presence of Other Taxes: General- Equilibrium Analyses," *The American Economic Review*, 86 (4), 985–1000, `http://www.jstor.org/stable/2118315`.

Bovenberg, Lans and Ruud de Mooij (1994) "Environmental Levies and Distortionary Taxation," *American Economic Review*, 84 (4), 1085–89.

Bovenberg, Lans and Frederick (Rick) van der Ploeg (1994) "Environmental policy, public finance and the labour market in a second-best world," *Journal of Public Economics*, 55 (3), 349–390, `https://EconPapers.repec.org/RePEc:eee:pubeco:v:55:y:1994:i:3:p:349-390`.

Chamley, Christophe (1986) "Optimal taxation of capital income in general equilibrium with infinite lives," *Econometrica: Journal of the Econometric Society*, 607–622.

Chari, V.V. and Patrick J. Kehoe (1999) "Chapter 26 Optimal fiscal and monetary policy," 1 of Handbook of Macroeconomics, 1671–1745: Elsevier, https://doi.org/10.1016/S1574-0048(99)10039-9.

Chetty, Raj, Adam Guren, Day Manoli, and Andrea Weber (2011) "Are micro and macro labor supply elasticities consistent? A review of evidence on the intensive and extensive margins," *American Economic Review*, 101 (3), 471–75.

Cremer, Helmuth, Firouz Gahvari, and Norbert Ladoux (1998) "Externalities and optimal taxation," *Journal of Public Economics*, 70 (3), 343–364, https://doi.org/10.1016/S0047-2727(98)00039-5.

——— (2003) "Environmental taxes with heterogeneous consumers: an application to energy consumption in France," *Journal of Public Economics*, 87 (12), 2791–2815.

Cronin, Julie Anne, Don Fullerton, and Steven Sexton (2019) "Vertical and Horizontal Redistributions from a Carbon Tax and Rebate," *Journal of the Association of Environmental and Resource Economists*, 6 (S1), 169–208.

Cruz, José-Luis and Esteban Rossi-Hansberg (2021) "The Economic Geography of Global Warming," Working Paper 28466, National Bureau of Economic Research, 10.3386/w28466.

Dell, Melissa, Benjamin F Jones, and Benjamin A Olken (2012) "Temperature shocks and economic growth: Evidence from the last half century," *American Economic Journal: Macroeconomics*, 4 (3), 66–95.

Diamond, Peter A and James A Mirrlees (1971) "Optimal taxation and public production I: Production efficiency," *The American economic review*, 61 (1), 8–27.

Dietz, Simon, Frederick van der Ploeg, Armon Rezai, and Frank Venmans (2021) "Are economists getting climate dynamics right and does it matter?" *Journal of the Association of Environmental and Resource Economists*, 8 (5), 895–921.

Douenne, Thomas and Adrien Fabre (2022) "Yellow Vests, Pessimistic Beliefs, and Carbon Tax Aversion," *American Economic Journal: Economic Policy*, 14 (1), 81–110, 10.1257/pol.20200092.

Dyrda, Sebastian and Marcelo Pedroni (2022) "Optimal Fiscal Policy in a Model with Uninsurable Idiosyncratic Income Risk," *The Review of Economic Studies*.

EPA (2017) "Inventory of U.S. Greenhouse Gas Emissions and Sinks: 1990-2015," report, U.S. Environmental Protection Agency.

Fried, Stephie, Kevin Novan, and William Peterman (2018) "The Distributional Effects of a Carbon Tax on Current and Future Generations," *Review of Economic Dynamics*, 30, 30–46, https://EconPapers.repec.org/RePEc:red:issued:16-217.

——— (2021) "Recycling Carbon Tax Revenue to Maximize Welfare," working paper, FEDS Working Paper.

Fullerton, Don and Holly Monti (2013) "Can pollution tax rebates protect low-wage earners?" *Journal of Environmental Economics and Management*, 66 (3), 539–553.

Golosov, Mikhail, John Hassler, Per Krusell, and Aleh Tsyvinski (2014) "Optimal Taxes on Fossil Fuel in General Equilibrium," *Econometrica*, 82 (1), 41–88, http://www.jstor.org/stable/24029171.

Goulder, Lawrence H (1995) "Environmental taxation and the double dividend: a reader's guide," *International tax and public finance*, 2 (2), 157–183.

Goulder, Lawrence H, Marc AC Hafstead, GyuRim Kim, and Xianling Long (2019) "Impacts of a carbon tax across US household income groups: What are the equity-efficiency trade-offs?" *Journal of Public Economics*, 175, 44–64.

Hassler, John and Per Krusell (2012) "Economics and climate change: integrated assessment in a multi-region world," *Journal of the European Economic Association*, 10 (5), 974–1000.

Heathcote, Jonathan, Fabrizio Perri, and Giovanni L Violante (2010) "Unequal we stand: An empirical analysis of economic inequality in the United States, 1967–2006," *Review of Economic dynamics*, 13 (1), 15–51.

Hsiang, Solomon, Paulina Oliva, and Reed Walker (2020) "The distribution of environmental damages," *Review of Environmental Economics and Policy*.

Jacobs, Bas (2018) "The marginal cost of public funds is one at the optimal tax system," *International Tax and Public Finance*, 25 (4), 883–912.

Jacobs, Bas and Ruud A. de Mooij (2015) "Pigou meets Mirrlees: On the irrelevance of tax distortions for the second-best Pigouvian tax," *Journal of Environmental Economics and Management*, 71, 90–108, https://doi.org/10.1016/j.jeem.2015.01.003.

Jacobs, Bas and Frederick (Rick) van der Ploeg (2019) "Redistribution and pollution taxes with non-linear Engel curves," *Journal of Environmental Economics and Management*, 95 (C), 198–226, https://EconPapers.repec.org/RePEc:eee:jeeman:v:95:y:2019:i:c:p:198-226.

Judd, Kenneth L (1985) "Redistributive taxation in a simple perfect foresight model," *Journal of public Economics*, 28 (1), 59–83.

Känzig, Diego (2021) "The unequal economic consequences of carbon pricing," working paper, London Business School.

Kaplow, Louis (2012) "Optimal Control of Externalities in the Presence of Income Taxation," *International Economic Review*, 53 (2), 487–509, http://www.jstor.org/stable/23251596.

Klenert, David, Gregor Schwerhoff, Ottmar Edenhofer, and Linus Mattauch (2018) "Environmental Taxation, Inequality and Engels Law: The Double Dividend of Redistribution," *Environmental & Resource Economics*, 71 (3), 605–624, 10.1007/s10640-016-0070-y.

Kotlikoff, Laurence, Felix Kubler, Andrey Polbin, Jeffrey Sachs, and Simon Scheidegger (2021) "Making carbon taxation a generational win win," *International Economic Review*, 62 (1), 3–46.

Krusell, Per and Anthony A Smith (2015) "Climate change around the world," in *Work in progress, Walras-Bowley Lecture. Econometric Society 2015 World Congress, Montréal*.

Kuhn, Moritz and José-Víctor Ríos-Rull (2016) "2013 Update on the US earnings, income, and wealth distributional facts: A View from Macroeconomics," *Federal Reserve Bank of Minneapolis Quarterly Review*, 37 (1), 2–73.

Leach, Andrew J (2009) "The welfare implications of climate change policy," *Journal of Environmental Economics and Management*, 57 (2), 151–165.

Levinson, Arik and James OBrien (2019) "Environmental Engel curves: Indirect emissions of common air pollutants," *Review of Economics and Statistics*, 101 (1), 121–133.

Micheletto, Luca (2008) "Redistribution and optimal mixed taxation in the presence of consumption externalities," *Journal of Public Economics*, 92 (10-11), 2262–2274.

Nordhaus, William (2018) "Projections and uncertainties about climate change in an era of minimal climate policies," *American economic journal: economic policy*, 10 (3), 333–60.

Nordhaus, William D. (2017) "Revisiting the social cost of carbon," *Proceedings of the National Academy of Sciences*, 114 (7), 1518–1523, 10.1073/pnas.1609244114.

Pigou, A.C. (1920) *The Economics of Welfare*: Macmillan.

Piketty, Thomas and Emmanuel Saez (2013) "Optimal labor income taxation," in *Handbook of public economics*, 5, 391–474: Elsevier.

Pindyck, Robert S (2013) "Climate change policy: what do the models tell us?" *Journal of Economic Literature*, 51 (3), 860–72.

Pirttilä, Jukka and Matti Tuomala (1997) "Income tax, commodity tax and environmental policy," *International Tax and Public Finance*, 4 (3), 379–393.

Pizer, William A and Steven Sexton (2019) "The Distributional Impacts of Energy Taxes," *Review of Environmental Economics and Policy*, 13 (1), 104–123.

Rausch, Sebastian (2013) "Fiscal consolidation and climate policy: An overlapping generations perspective," *Energy Economics*, 40, S134–S148.

Rausch, Sebastian, Gilbert E Metcalf, and John M Reilly (2011) "Distributional impacts of carbon pricing: A general equilibrium approach with micro-data for households," *Energy economics*, 33, S20–S33.

Rausch, Sebastian and Hidemichi Yonezawa (2018) "The intergenerational incidence of green tax reform," *Climate Change Economics*, 9 (01), 1840007.

Ricke, Katharine, Laurent Drouet, Ken Caldeira, and Massimo Tavoni (2018) "Country-level social cost of carbon," *Nature Climate Change*, 8 (10), 895–900.

Sallee, James M (2019) "Pigou Creates Losers: On the Implausibility of Achieving Pareto Improvements from Efficiency-Enhancing Policies," Working Paper 25831, National Bureau of Economic Research, 10.3386/w25831.

Sandmo, Agnar (1975) "Optimal Taxation in the Presence of Externalities," *The Swedish Journal of Economics*, 77 (1), 86–98, `http://www.jstor.org/stable/3439329`.

Straub, Ludwig and Iván Werning (2020) "Positive long-run capital taxation: Chamley-Judd revisited," *American Economic Review*, 110 (1), 86–119.

Trabandt, Mathias and Harald Uhlig (2012) "How do Laffer curves differ across countries?" *Fiscal policy after the financial crisis*, 211–249.

Weitzman, Martin L (2009) "On modeling and interpreting the economics of catastrophic climate change," *The review of economics and statistics*, 91 (1), 1–19.

Werning, Iván (2007) "Optimal Fiscal Policy with Redistribution," *The Quarterly Journal of Economics*, 122 (3), 925–967, 10.1162/qjec.122.3.925.

Williams, Roberton, Hal Gordon, Dallas Burtraw, Jared Carbone, and Richard D. Morgenstern (2015) "The Initial Incidence of a Carbon Tax Across Income Groups," *National Tax Journal*, 68 (1), 195–214.

Appendices

A Optimal tax rules in the benchmark model

A.1 Implementability conditions

Let $\varphi \equiv \{\varphi_i\}$ be the market weights with $\varphi_i \geq 0$. Then, given aggregate levels c_t, h_t and Z_t, the individual levels can be found by solving the following static subproblem for each period t:

$$U\left(c_t, h_t, Z_t; \varphi\right) \equiv \max_{c_{i,t}, h_{i,t}} \sum_i \pi_i \varphi_i u\left(c_{i,t}, h_{i,t}, Z_t\right), \quad \text{s.t.} \sum_i \pi_i c_{i,t} = c_t, \quad \text{and} \quad \sum_i \pi_i e_i h_{i,t} = h_t. \tag{33}$$

The Lagrangian for this problem is

$$L = \sum_i \pi_i \varphi_i u\left(c_{i,t}, h_{i,t}, Z_t\right) + \theta_t^c \left(c_t - \sum_i \pi_i c_{i,t}\right) - \theta_t^h \left(h_t - \sum_i \pi_i e_i h_{i,t}\right),$$

where θ_t^c and θ_t^h are Lagrange multipliers. Applying the envelope theorem to problem (33), we get

$$U_{c,t} = \theta_t^c, \quad \text{and} \quad U_{h,t} = -\theta_t^h.$$

From the first order conditions of problem (33), we also have

$$\varphi_i u_{c,i,t} = \theta_t^c, \quad \text{and} \quad \varphi_i u_{h,i,t} = -e_i \theta_t^h.$$

It follows that

$$U_{c,t} = \varphi_i u_{c,i,t}, \tag{34}$$

$$U_{h,t} = \frac{\varphi_i u_{h,i,t}}{e_i}. \tag{35}$$

In any competitive equilibrium these optimality conditions must hold for every agent i. Hence, using (34), (35), and agents' first order conditions given by

$$\beta^t \frac{u_{c,i,t}}{u_{c,i,0}} = p_t, \quad \forall\, t \geq 0,$$

$$\frac{u_{h,i,t}}{u_{c,i,t}} = -\left(1 - \tau_{H,t}\right) e_i w_t, \quad \forall\, t \geq 0,$$

we obtain

$$\frac{U_{h,t}}{U_{c,t}} = \frac{u_{h,i,t}}{u_{c,i,t} e_i} = -w_t \left(1 - \tau_{H,t}\right), \tag{36}$$

and

$$\frac{U_{c,t}}{U_{c,0}} = \frac{u_{c,i,t}}{u_{c,i,0}} = \frac{p_t}{\beta^t}. \tag{37}$$

Given the relationships above we can derive the implementation condition which relies only on the aggregates c_t, h_t, and market weights φ. Let $c_{i,t}^m\left(c_t, h_t; \varphi\right)$ and $h_{i,t}^m\left(c_t, h_t; \varphi\right)$ be the arg max of problem (33). The budget constraint of agent i implies

$$\sum_{t=0}^{\infty} N_t p_t \left(c_{i,t}^m\left(c_t, h_t; \varphi\right) - \left(1 - \tau_{H,t}\right) w_t e_i h_{i,t}^m\left(c_t, h_t; \varphi\right)\right) \leq R_0 N_0 a_{i,0} + T,$$

which using (36) and (37) can be restated as

$$U_{c,0}\big(R_0 N_0 a_{i,0} + T\big) = \sum_{t=0}^{\infty} N_t \beta^t \Big(U_{c,t} c_{i,t}^m \big(c_t, h_t; \varphi\big) + U_{h,t} e_i h_{i,t}^m \big(c_t, h_t; \varphi\big)\Big), \quad \forall\ i. \tag{38}$$

A.2 Ramsey problem

A.2.1 Problem

Let $\lambda \equiv \{\lambda_i\}$ be the planner's welfare weight on type i, with $\sum_i \pi_i \lambda_i = 1$. Define

$$\begin{aligned} W\left(c_t, h_t, Z_t; \varphi, \theta, \lambda\right) \equiv & \sum_i \pi_i \lambda_i u \left(c_{i,t}^m \left(c_t, h_t; \varphi\right), h_{i,t}^m \left(c_t, h_t; \varphi\right), Z_t\right) \\ & + \sum_i \pi_i \theta_i \left[U_{c,t} c_{i,t}^m \left(c_t, h_t; \varphi\right) + U_{h,t} e_i h_{i,t}^m \left(c_t, h_t; \varphi\right)\right] \end{aligned}$$

where $\pi_i \theta_i$ is the Lagrange multiplier on the implementability constraint of agent i, and $\theta \equiv \{\theta_i\}$. The Ramsey problem can be written as

$$\max_{\substack{\{C_t, H_{1,t}, H_{2,t}, K_{1,t}, K_{2,t}, \\ E_t, Z_t, \mu_t\}_{t=0}^{\infty}, T, \varphi}} \sum_{t,i} N_t \beta^t W\left(c_t, h_t, Z_t; \varphi, \theta, \lambda\right) - U_{c,0} \sum_i \pi_i \theta_i \left(R_0 N_0 a_{i,0} + T\right) \tag{39}$$

subject to

$$\begin{aligned} N_t c_t + G_t + K_{t+1} + \Theta_t\left(\mu_t, E_t\right) &= \left(1 - D\left(Z_t\right)\right) A_{1,t} F\left(K_{1,t}, H_{1,t}, E_t\right) + \left(1 - \delta\right) K_t, \quad \forall\ t \geq 0, \\ E_t &= A_{2,t} G\left(K_{2,t}, H_{2,t}\right), \quad \forall\ t \geq 0, \\ Z_t &= J\left(S_0, E_0^M, ..., E_t^M, \eta_0, ..., \eta_t\right), \quad \forall\ t \geq 0, \\ F_{K,t} G_{H,t} &= G_{K,t} F_{H,t}, \quad \forall\ t \geq 0, \\ K_{1,t} + K_{2,t} &= K_t, \quad \forall\ t \geq 0, \\ H_{1,t} + H_{2,t} &= N_t h_t, \quad \forall\ t \geq 0, \end{aligned}$$

where $\beta^t \nu_{jt}$ for $j \in \{1, 2, 3\}$ are the Lagrange multipliers on the feasibility constraints in the order above. When using a functional form for households' utility below, it will also be convenient to add an additional constraint from the normalization of market weights. Because this constraint is a simple normalization, it has no impact on the resulting allocations.

In what follows, we assume that there is no initial wealth inequality, that is $a_{i,0} = a_{j,0}$ for every i and j. We relax this assumption in Appendix C.

A.2.2 First order conditions

The first order conditions are

$$[c_t] : W_{c,t} - \nu_{1,t} = 0, \quad \forall\, t \geq 0, \tag{40}$$

$$[H_{1,t}] : W_{h,t} + \nu_{1,t} \left(1 - D_t\right) A_{1,t} F_{H,t} = 0, \quad \forall\, t \geq 0, \tag{41}$$

$$[H_{2,t}] : W_{h,t} + \nu_{2,t} A_{2,t} G_{H,t} = 0, \quad \forall\, t \geq 0, \tag{42}$$

$$[K_{1,t+1}] : -\nu_{1,t} + \left[\left(1 - D_{t+1}\right) A_{1,t+1} F_{K,t+1} + \left(1 - \delta\right)\right] \beta \nu_{1,t+1} = 0, \quad \forall\, t \geq 0, \tag{43}$$

$$[K_{2,t+1}] : -\nu_{1,t} + A_{2,t+1} G_{K,t+1} \beta \nu_{2,t+1} + \left(1 - \delta\right) \beta \nu_{1,t+1} = 0, \quad \forall\, t \geq 0, \tag{44}$$

$$[E_t] : -\nu_{1,t} \left(\Theta_{E,t} - \left(1 - D_t\right) A_{1,t} F_{E,t}\right) - \nu_{2,t} - \sum_{j=0}^{\infty} \beta^j \nu_{3,t+j} J_{E_t^M,t+j} \left(1 - \mu_t\right) = 0, \quad \forall\, t \geq 0, \tag{45}$$

$$[Z_t] : N_t W_{Z,t} - \nu_{1,t} D_t' A_{1,t} F_t + \nu_{3,t} = 0, \quad \forall\, t \geq 0, \tag{46}$$

$$[\mu_t] : -\nu_{1,t} \Theta_{\mu,t} \left(\mu_t, E_t\right) + \sum_{j=0}^{\infty} \beta^j \nu_{3,t+j} J_{E_t^M,t+j} E_t = 0, \quad \forall\, t \geq 0, \tag{47}$$

$$[T] : \sum_i \pi_i \theta_i = 0, \tag{48}$$

and at $t = 0$,

$$\left[\tau_0^k\right] : U_{c,0} \left(\left(1 - D_0\right) A_{1,0} F_{K,0} - \delta\right) N_0 \sum_i \pi_i \theta_i a_{i,0} = 0, \tag{49}$$

$$[K_{1,0}] : \left[\left(1 - D_0\right) A_{1,0} F_{K,0} + \left(1 - \delta\right)\right] \nu_{1,0} - \kappa = 0, \tag{50}$$

$$[K_{2,0}] : A_{2,0} G_{K,0} \nu_{2,0} + \left(1 - \delta\right) \nu_{1,0} - \kappa = 0, \tag{51}$$

where κ is the Lagrange multiplier on the constraint $K_{1,0} + K_{2,0} = K_0$, and it follows that

$$(1 - D_0) A_{1,0} F_{K,0} \nu_{1,0} = A_{2,0} G_{K,0} \nu_{2,0},$$

which together with (41) and (42), implies that

$$\frac{F_{K,0}}{F_{H,0}} = \frac{G_{K,0}}{G_{H,0}}.$$

As in any other period, in $t = 0$ the requirement that the marginal rates of technical substitution are equated between sectors is satisfied at the second-best allocation. Therefore, in most of what follows we ignore the multiplier on this constraint.

A.3 Optimal taxes

A.3.1 Capital and Labor income taxes

From (40) and (41) we obtain

$$(1 - D(Z_t)) A_{1,t} F_{H,t} = -\frac{W_{h,t}}{W_{c,t}}, \quad \forall\, t \geq 0, \tag{52}$$

and using the intertemporal condition (43) we get

$$R^*_{t+1} \equiv 1 + r_{t+1} - \delta = \frac{1}{\beta}\frac{W_{c,t}}{W_{c,t+1}}, \quad \forall\, t \geq 0, \tag{53}$$

These two equations can be used to back out the optimal taxes on labor and capital income. Plugging (52) into (36) implies

$$\frac{U_{h,t}}{U_{c,t}} = \frac{W_{h,t}}{W_{c,t}}(1 - \tau_{H,t}),$$

which can be rearranged into

$$\tau_{H,t} = 1 - \frac{U_{h,t}}{U_{c,t}}\frac{W_{c,t}}{W_{h,t}}. \tag{54}$$

In any competitive equilibrium (37) holds, which together with $p_t = R_{t+1}p_{t+1}$ implies

$$\frac{U_{c,t+1}}{U_{c,t}}\beta R_{t+1} = 1.$$

Substituting this into (53), it follows that

$$\frac{R_{t+1}}{R^*_{t+1}} = \frac{W_{c,t+1}}{W_{c,t}}\frac{U_{c,t}}{U_{c,t+1}}. \tag{55}$$

A.3.2 Excise taxes of energy and emissions

From the abatement first-order condition (47) and the energy firm abatement decision (9) we have that

$$\tau_{E,t} = \frac{\Theta_{\mu,t}}{E_t} = \frac{1}{\nu_{1,t}}\sum_{j=0}^{\infty}\beta^j \nu_{3,t+j} J_{E_t^M,t+j}.$$

From the climate variable first-order condition (46) we have that

$$\nu_{3,t} = \nu_{1,t} D'_t A_{1,t} F(K_{1,t}, H_{1,t}, E_t) - N_t W_{Z,t},$$

hence the pollution tax is given by

$$\tau_{E,t} = \frac{1}{\nu_{1,t}}\sum_{j=0}^{\infty}\beta^j \left(\nu_{1,t+j} D'_{t+j} A_{1,t+j} F_{t+j} - N_{t+j} W_{Z,t+j}\right) J_{E_t^M,t+j}. \tag{56}$$

From the energy first-order condition (45) we have that

$$-\nu_{1,t}\Big(\Theta_{E,t} + (1 - \mu_t)\frac{\Theta_{\mu,t}}{E_t} - (1 - D(Z_t))A_{1,t}F_{E,t}\Big) = \nu_{2,t}, \tag{57}$$

and combining the first-order conditions for sectoral labor supplies (41) and (42), it follows that

$$\frac{\nu_{2,t}}{\nu_{1,t}} = \frac{\left(1 - D\left(Z_t\right)\right) A_{1,t} F_{H,t}}{A_{2,t} G_{H,t}}.$$

From (4) and (8) we also have

$$\frac{\left(1 - D\left(Z_t\right)\right) A_{1,t} F_{H,t}}{A_{2,t} G_{H,t}} = p_{E,t} - \tau_{I,t} - \tau_{E,t}(1-\mu_t) - \Theta_{E,t}.$$

Hence, using (5), (9), and (57) we have

$$-\Theta_{E,t} - (1-\mu_t)\tau_{E,t} + p_{E,t} = p_{E,t} - \tau_{I,t} - \tau_{E,t}(1-\mu_t) - \Theta_{E,t},$$

and therefore

$$\tau_{I,t} = 0. \tag{58}$$

A.4 Explicit formulas

A.4.1 Characterization of equilibrium

Let us consider the following balanced-growth utility function

$$u\left(c_i, h_i, Z\right) = \frac{\left(c_i(1-\varsigma h_i)^{\gamma}\right)^{1-\sigma}}{1-\sigma} + \frac{\left(1+\alpha_0 Z_t^2\right)^{-(1-\sigma)}}{1-\sigma}. \tag{59}$$

To obtain explicit formulas, it is convenient to normalize market weights as follows

$$\sum_j \pi_j \left(\varphi_j e_j^{\gamma(\sigma-1)}\right)^{\frac{1}{\sigma-(1-\sigma)\gamma}} = 1.$$

Using the period utility function defined in (59), the Lagrangian for the characterization problem defined by (15) is

$$L = \sum_i \pi_i \varphi_i \left[\frac{\left(c_{i,t}\left(1-\varsigma h_{i,t}\right)^{\gamma}\right)^{1-\sigma}}{1-\sigma} + \frac{\left(1+\alpha_0 Z_t^2\right)^{-(1-\sigma)}}{1-\sigma}\right] + \theta_t^c \left(c_t - \sum_i \pi_i c_{i,t}\right) - \theta_t^h \left(h_t - \sum_i \pi_i e_i h_{i,t}\right),$$

The first order conditions are

$$\begin{aligned}
&[c_{i,t}] : \varphi_i \left(c_{i,t}\left(1-\varsigma h_{i,t}\right)^{\gamma}\right)^{1-\sigma} c_{i,t}^{-1} = \theta_t^c, \quad \forall\, t \geq 0, \\
&[h_{i,t}] : \varphi_i \left(c_{i,t}\left(1-\varsigma h_{i,t}\right)^{\gamma}\right)^{1-\sigma} \gamma\varsigma \left(1-\varsigma h_{i,t}\right)^{-1} = e_i \theta_t^h, \quad \forall\, t \geq 0,
\end{aligned}$$

rearranging yields

$$c_{i,t} = \frac{\theta_t^h}{\theta_t^c} \frac{e_i\left(1-\varsigma h_{i,t}\right)}{\gamma\varsigma},$$

so that

$$\begin{aligned}
c_{i,t} &= \left(\frac{\theta_t^c}{\varphi_i}\left(\frac{\theta_t^h}{\theta_t^c}\frac{e_i}{\gamma\varsigma}\right)^{\gamma(1-\sigma)}\right)^{-\frac{1}{\sigma-(1-\sigma)\gamma}} \\
1-\varsigma h_{i,t} &= \frac{\theta_t^c}{\theta_t^h}\frac{\gamma\varsigma}{e_i}\left(\frac{\theta_t^c}{\varphi_i}\left(\frac{\theta_t^h}{\theta_t^c}\frac{e_i}{\gamma\varsigma}\right)^{\gamma(1-\sigma)}\right)^{-\frac{1}{\sigma-(1-\sigma)\gamma}},
\end{aligned}$$

and summing across types (given that $c_t = \sum_i \pi_i c_{i,t}$, and $h_t = \sum_i \pi_i e_i h_{i,t}$)

$$c_t = \left(\theta_t^c \left(\frac{\theta_t^h}{\theta_t^c}\frac{1}{\gamma\varsigma}\right)^{\gamma(1-\sigma)}\right)^{-\frac{1}{\sigma-(1-\sigma)\gamma}} \sum_i \pi_i \left(\frac{e_i^{\gamma(1-\sigma)}}{\varphi_i}\right)^{-\frac{1}{\sigma-(1-\sigma)\gamma}}$$

$$1-\varsigma h_t = \frac{\theta_t^c}{\theta_t^h}\gamma\varsigma\left(\theta_t^c \left(\frac{\theta_t^h}{\theta_t^c}\frac{1}{\gamma\varsigma}\right)^{\gamma(1-\sigma)}\right)^{-\frac{1}{\sigma-(1-\sigma)\gamma}} \sum_i \pi_i \left(\frac{e_i}{\varphi_i}^{\gamma(1-\sigma)}\right)^{-\frac{1}{\sigma-(1-\sigma)\gamma}}$$

It follows that

$$c_{i,t}^m\left(c_t, h_t; \varphi\right) = \omega_i c_t, \tag{60}$$

$$1-\varsigma h_{i,t}^m\left(c_t, h_t; \varphi\right) = \frac{\omega_i}{e_i}\left(1-\varsigma h_t\right), \tag{61}$$

where

$$\omega_i = \frac{\left(\varphi_i\left(e_i\right)^{\gamma(\sigma-1)}\right)^{\frac{1}{\sigma-(1-\sigma)\gamma}}}{\sum_i \pi_i\left(\varphi_j e_j^{\gamma(\sigma-1)}\right)^{\frac{1}{\sigma-(1-\sigma)\gamma}}} = \left(\varphi_i\left(e_i\right)^{\gamma(\sigma-1)}\right)^{\frac{1}{\sigma-(1-\sigma)\gamma}}. \tag{62}$$

Thus, we can write aggregate indirect utility $U\left(c_t, h_t, Z_t; \varphi\right)$ in terms of the aggregates c_t, h_t, and Z_t

$$\begin{aligned} U\left(c_t, h_t, Z_t, \varphi\right) =& \sum_j \pi_j\varphi_j\left(\frac{\omega_j^{1+\gamma}}{e_j^\gamma}\right)^{1-\sigma}\frac{\left(c_t\left(1-\varsigma h_t\right)^\gamma\right)^{1-\sigma}}{1-\sigma} + \sum_i \pi_i\varphi_i\frac{\left(1+\alpha_0 Z_t^2\right)^{-(1-\sigma)}}{1-\sigma} \\ =& \frac{\left(c_t\left(1-\varsigma h_t\right)^\gamma\right)^{1-\sigma}}{1-\sigma} + \Gamma\frac{\left(1+\alpha_0 Z_t^2\right)^{-(1-\sigma)}}{1-\sigma}, \end{aligned} \tag{63}$$

since from the normalization of market weights we have

$$\sum_j \pi_j\varphi_j\left(\frac{\omega_j^{1+\gamma}}{e_j^\gamma}\right)^{1-\sigma} = \sum_j \pi_j\left(\varphi_j e_j^{\gamma(\sigma-1)}\right)^{\frac{1}{\sigma-(1-\sigma)\gamma}} = 1,$$

and with $\Gamma \equiv \sum_i \pi_i\varphi_i$.

A.4.2 Explicit tax formulas

From (38), substituting the derivatives of $U\left(c_t, h_t, Z_t; \varphi\right)$ into the definition of $W\left(c_t, h_t, Z_t; \varphi, \theta, \lambda\right)$ we get

$$\begin{aligned} W\left(c_t, h_t, Z_t; \varphi, \theta, \lambda\right) =& \sum_i \pi_i\lambda_i\left(\frac{\omega_i}{\varphi_i}\frac{\left(c_t\left(1-\varsigma h_t\right)^\gamma\right)^{1-\sigma}}{1-\sigma} + \frac{\left(1+\alpha_0 Z_t^2\right)^{-(1-\sigma)}}{1-\sigma}\right) \\ &+ \sum_i \pi_i\theta_i\left[\left(c_t\left(1-\varsigma h_t\right)^\gamma\right)^{1-\sigma}\omega_i - \gamma\left(c_t\left(1-\varsigma h_t\right)^\gamma\right)^{1-\sigma}\left(1-\varsigma h_t\right)^{-1}\left(e_i-\omega_i\left(1-\varsigma h_t\right)\right)\right] \end{aligned}$$

Collecting terms and simplifying we obtain

$$W\left(c_t, h_t, Z_t; \varphi, \theta, \lambda\right) = \Phi\frac{\left(c_t\left(1-\varsigma h_t\right)^\gamma\right)^{1-\sigma}}{1-\sigma} + \frac{\left(1+\alpha_0 Z_t^2\right)^{-(1-\sigma)}}{1-\sigma} + \Psi U_{h,t}. \tag{64}$$

where

$$\Phi \equiv \sum_i \pi_i \omega_i \left(\frac{\lambda_i}{\varphi_i} + (1-\sigma)(1+\gamma)\theta_i \right),$$
$$\Psi \equiv \sum_i \frac{\pi_i \theta_i e_i}{\varsigma}.$$

Substituting the derivatives into equation (54) we get

$$\tau_{H,t} = \frac{\Psi \varsigma (1-\varsigma h_t)^{-1}}{\Phi + \Psi \varsigma (1-\gamma(1-\sigma))(1-\varsigma h_t)^{-1}},$$

substituting the derivatives into (55) yields

$$\frac{R_{t+1}}{R^*_{t+1}} = \frac{\Phi - \Psi \varsigma \gamma (1-\sigma)(1-\varsigma h_{t+1})^{-1}}{\Phi - \Psi \varsigma \gamma (1-\sigma)(1-\varsigma h_t)^{-1}}.$$

and substituting the derivatives into (56) we get

$$\tau_{E,t} = \frac{1}{\nu_{1,t}} \sum_{j=0}^{\infty} \beta^j \Big(\nu_{1,t+j} D'_{t+j} A_{1,t+j} F_{t+j} - N_{t+j} V_Z(Z_{t+j}) \Big) J_{E_t^M, t+j}, \tag{65}$$

with $\nu_{1,t}$ the multiplier of the resource constraint which we can express as

$$\nu_{1,t} = W_{c,t} = V_{c,t} + \sum_i \pi_i \theta_i IC_{c,i,t}. \tag{66}$$

If we add—without loss of generality—the normalization of market weights as a constraint into the Ramsey problem, we obtain the following first order conditions with respect to market weights

$$\sum_t \beta^t N_t W_{\varphi_i,t} - \frac{\zeta}{\sigma - (1-\sigma)\gamma} \frac{\pi_i \omega_i}{\varphi_i} = 0, \quad \forall\ i.$$

From this equation we have that

$$\sum_{t=0}^{\infty} N_t \beta^t \frac{(c_t (1-\varsigma h_t)^{\gamma})^{1-\sigma}}{1-\sigma} \frac{(1-\sigma)(1+\gamma)}{\sigma - (1-\sigma)\gamma} \frac{\pi_i \omega_i}{\varphi_i} \left(\frac{\lambda_i}{\varphi_i} + \theta_i \right) - \frac{\zeta}{\sigma - (1-\sigma)\gamma} \frac{\pi_i \omega_i}{\varphi_i} = 0, \quad \forall\ i,$$

and therefore

$$\frac{\lambda_i}{\varphi_i} + \theta_i = \frac{\zeta}{(1-\sigma)(1+\gamma) \sum_{t=0}^{\infty} N_t \beta^t \tilde{U}(c_t, h_t)}, \quad \forall\ i,$$

with

$$\tilde{U}(c_t, h_t) = \frac{(c_t (1-\varsigma h_t)^{\gamma})^{1-\sigma}}{1-\sigma}.$$

Using the fact that

$$\sum_i \pi_i \theta_i = 0, \quad \sum_i \pi_i \omega_i = 1, \quad \text{and} \quad \sum_i \pi_i e_i = 1,$$

it follows that

$$\sum_j \frac{\pi_j \lambda_j}{\varphi_j} = \frac{\zeta}{(1-\sigma)(1+\gamma) \sum_{t=0}^{\infty} N_t \beta^t \tilde{U}(c_t, h_t)},$$

and, therefore

$$\theta_i = \sum_j \frac{\pi_j \lambda_j}{\varphi_j} - \frac{\lambda_i}{\varphi_i}. \tag{67}$$

This allows us to rewrite

$$\begin{aligned}
\Phi &= \sum_i \pi_i \omega_i \left(\frac{\lambda_i}{\varphi_i} + (1-\sigma)(1+\gamma) \left(\sum_j \frac{\pi_j \lambda_j}{\varphi_j} - \frac{\lambda_i}{\varphi_i} \right) \right) \\
&= \sum_j \pi_j \frac{\lambda_j}{\varphi_j} + \big(1 - (1+\gamma)(1-\sigma)\big) \mathrm{cov}(\lambda_i/\varphi_i, \omega_i), \qquad (68) \\
\Psi &= \frac{1}{\varsigma} \sum_j \pi_j \frac{\lambda_j}{\varphi_j} (1 - e_j) = -\frac{\mathrm{cov}(\lambda_i/\varphi_i, e_i)}{\varsigma}, \qquad (69)
\end{aligned}$$

where the last result is obtained using the normalization of productivity levels, $\sum_i \pi_i e_i = 1$.

Notice that labor and capital income taxes are zero whenever $\Psi = 0$, which, according to equation (69), occurs in three special cases: *(i)* when there is no agent heterogeneity, *(ii)* when the planner's and the market's weights are perfectly aligned, and *(iii)* when agents' productivity are uncorrelated with the relative social weights. Intuitively, the first case corresponds to the outcome of a representative-agent model in which lump-sum taxation is allowed: since there is no need to redistribute, the government can rely only on non-distortionary taxes to finance its expenditures. The second case corresponds to the situation in which the market allocation happens to be the one preferred by the planner: although there might be inequalities due to differences in productivity and asset holdings, they are consistent with the relative weight the planner gives to each type of individual. The third situation encompasses the two previous ones, but also includes situations in which the planner would want to redistribute but faces a targeting problem, *i.e.* it cannot reach a better allocation than the market one using anonymous linear instruments due to the absence of correlation between the source of inequalities and its relative preference over agents' types.

The implementability conditions can be rewritten as

$$\omega_i = \frac{U_{c,0}\left(R_0 N_0 a_{i,0} + T\right) + M e_i}{(1-\sigma)(1+\gamma)\sum_{t=0}^{\infty} N_t \beta^t \tilde{U}(c_t, h_t)}, \quad \forall\ i,$$

with

$$M \equiv \sum_{t=0}^{\infty} N_t \beta^t \gamma \left(c_t (1 - \varsigma h_t)^\gamma\right)^{1-\sigma} (1 - \varsigma h_t)^{-1}. \tag{70}$$

Since $\sum_{i=1}^{n} \omega_i = 1$, it follows that

$$\omega_i = 1 + \frac{U_{c,0} R_0 N_0 (a_{i,0} - A_0) + M(e_i - 1)}{(1-\sigma)(1+\gamma)\sum_{t=0}^{\infty} N_t \beta^t \tilde{U}(c_t, h_t)}, \quad \forall\ i. \tag{71}$$

Moreover, since

$$\omega_i = \left(\varphi_i e_i^{\gamma(\sigma-1)}\right)^{\frac{1}{\sigma - (1-\sigma)\gamma}},$$

we can express market weights as

$$\varphi_i = \frac{\omega_i^{\sigma-(1-\sigma)\gamma}}{e_i^{\gamma(\sigma-1)}} = \frac{1}{e_i^{\gamma(\sigma-1)}} \left(1 + \frac{U_{c,0} R_0 N_0 (a_{i,0} - A_0) + M(e_i - 1)}{(1-\sigma)(1+\gamma)\sum_{t=0}^{\infty} N_t \beta^t \tilde{U}(c_t, h_t)}\right)^{\sigma-(1-\sigma)\gamma}. \tag{72}$$

A.5 Comparison with Pigou

First-best pollution tax To compare our second-best results with the first-best, we solve the same Ramsey problem except that we now allow for individualized lump-sum transfers. All first order conditions remain the same except for the one with respect to T given by (48): since we now have individualized instruments T_i, we obtain

$$\theta_i = 0, \quad \forall i,$$

hence for all t, $\sum_i \pi_i \theta_i IC_{c,i,t} = 0$. From (67), this also implies that

$$\frac{\lambda_i}{\varphi_i} = \sum_j \frac{\pi_j \lambda_i}{\varphi_i}, \quad \forall i,$$

and as a consequence we have $\Psi = 0$, so that for all t, $\tau_{H,t} = 0$ and $\tau_{K,t} = 0$. Substituting for $\nu_{1,t}$ in (56), we can express the first-best tax as

$$\tau_{E,t}^{FB} = \sum_{j=0}^{\infty} \beta^j \left(\frac{V_{c,t+j}}{V_{c,t}} D'_{t+j} A_{1,t+j} F_{t+j} - \frac{N_{t+j} V_{Z,t+j}}{V_{c,t}} \right) J_{E_t^M, t+j} \ .$$

The first-best tax is equal to the social cost of the externality—*i.e.*, to the Pigouvian tax—evaluated at the first-best allocation.

Proof of Proposition 2: Recall the following definitions from Section 3.3:

$$\begin{aligned}
\text{MCF}_t &\equiv \frac{\nu_{1,t}}{V_{c,t}}, \\
\tau_{E,t}^{Pigou,Y} &\equiv \sum_{j=0}^{\infty} \beta^j \frac{V_{c,t+j}}{V_{c,t}} D'_{t+j} A_{1,t+j} F_{t+j} J_{E_t^M, t+j}, \\
\tau_{E,t}^{Pigou,U} &\equiv (-1) \sum_{j=0}^{\infty} \beta^j \frac{N_{t+j} V_{Z,t+j}}{V_{c,t}} J_{E_t^M, t+j}, \\
\tau_{E,t}^{Pigou} &\equiv \tau_{E,t}^{Pigou,Y} + \tau_{E,t}^{Pigou,U}, \\
\omega_t^U &\equiv \frac{\tau_{E,t}^{Pigou,U}}{\tau_{E,t}^{Pigou}}, \\
\Delta_{t+s} &\equiv \frac{\beta^s V_{c,t+s} D'_{t+s} A_{1,t+s} F_{t+s} J_{E_t^M, t+s}}{\sum_{j=0}^{\infty} \beta^j V_{c,t+j} D'_{t+j} A_{1,t+j} F_{t+j} J_{E_t^M, t+j}} \ .
\end{aligned}$$

Substituting into equation (65), we obtain equation (22) stated in Proposition 2,

$$\tau_{E,t} = \tau_{E,t}^{Pigou}\Big|_{SB} \Big(\sum_{j=0}^{\infty} \frac{\text{MCF}_{t+j}}{\text{MCF}_t} \Delta_{t+j} (1 - \omega_t^U) + \frac{\omega_t^U}{\text{MCF}_t} \Big).$$

Using equation (66) to substitute in the definition of the MCF, we also obtain equation (23) stated in Proposition 2,

$$\text{MCF}_t = 1 + \frac{\text{cov}(\theta_i, IC_{c,i,t})}{V_{c,t}}.$$

With balanced-growth preferences, using the expression of $U(c_t, h_t, Z_t, \varphi)$ given by (63) and the solutions for $c_{i,t}^m$ and $h_{i,t}^m$ given by (60) and (61), we can show that $IC_{i,t}$ defined by (19) can be expressed as

$$IC_{i,t} = \Big(c_t(1-\varsigma h_t)^\gamma\Big)^{(1-\sigma)} \left(\omega_i + \gamma\Big(\omega_i - \frac{e_i}{(1-\varsigma h_t)}\Big)\right), \tag{73}$$

from which we obtain

$$IC_{c,i,t} = (1-\sigma)\frac{IC_{i,t}}{c_t}.$$

Using the fact that $\sum_i \pi_i \theta_i = 0$, we can re-write the marginal reduction in implementation cost as

$$\sum_i \pi_i \theta_i IC_{c,i,t} = (1-\sigma)\frac{\mathrm{cov}(\theta_i, IC_{i,t})}{c_t}. \tag{74}$$

This term is equal to 0 when either σ tends to 1 or θ_i and $IC_{i,t}$ are uncorrelated. Thus, when the IES tends to 1, the MCF is equal to 1 in all periods. Using the binding implementability conditions, we can also express the discounted sum of $IC_{i,t}$,

$$\sum_{t=0}^{\infty} N_t \beta^t IC_i(c_t, h_t, \varphi) = U_{c,0}(R_0 a_{i,0} + T).$$

When there is no initial wealth inequality, or when initial wealth is expropriated—which, as we have shown, is optimal as long as initial wealth and productivity are positively correlated—then for any i, j, $R_0 a_{i,0} = R_0 a_{j,0}$. We can then write the discounted sum of $IC_{i,t}$ as a constant κ_{IC} that does not depend on agents' type,

$$\sum_{t=0}^{\infty} N_t \beta^t IC_i(c_t, h_t, \varphi) = \kappa_{IC}.$$

Using this expression, we can show that the welfare-weighted average MCF from period 0 onward is equal to 1, since with $V_t \equiv V(c_t, h_t, Z_t; \varphi, \lambda)$,

$$\begin{aligned}
\frac{\sum_{t=0}^{\infty} N_t \beta^t V_t \times \mathrm{MCF}_t}{\sum_{t=0}^{\infty} N_t \beta^t V_t} &= \frac{1}{\sum_{t=0}^{\infty} N_t \beta^t V_t}\left(\sum_{t=0}^{\infty} N_t \beta^t V_t + \sum_{t=0}^{\infty} N_t \beta^t V_t \sum_i \pi_i \theta_i \frac{(1-\sigma) IC_{i,t}}{V_{c,t} c_t}\right) \\
&= \frac{1}{\sum_{t=0}^{\infty} N_t \beta^t V_t}\left(\sum_{t=0}^{\infty} N_t \beta^t V_t + \sum_i \pi_i \theta_i \sum_{t=0}^{\infty} N_t \beta^t IC_{i,t}\right) \\
&= \frac{1}{\sum_{t=0}^{\infty} N_t \beta^t V_t}\left(\sum_{t=0}^{\infty} N_t \beta^t V_t + \kappa_{IC} \sum_i \pi_i \theta_i\right) \\
&= 1,
\end{aligned}$$

with $\sum_i \pi_i \theta_i = 0$. ■

Link with the capital income tax From (18), using balanced-growth utility, we can show that

$$V(c_t, h_t, Z_t; \varphi, \lambda) = \sum_i \pi_i \omega_i \frac{\lambda_i}{\varphi_i} \frac{(c_t(1-\varsigma h_t)^\gamma)^{1-\sigma}}{1-\sigma} + \frac{(1+\alpha_0 Z_t^2)^{-(1-\sigma)}}{1-\sigma},$$

hence using the explicit expression of $U(c_t, h_t, Z_t, \varphi)$ given by (63) and taking derivatives we have

$$V_{c,t} = \sum_i \pi_i \omega_i \frac{\lambda_i}{\varphi_i} U_{c,t}, \tag{75}$$

from which we can show that

$$\frac{V_{c,t+j}}{V_{c,t}} = \frac{U_{c,t+j}}{U_{c,t}}.$$

From (55), and using the fact that

$$\mathrm{MCF}_t = \frac{W_{c,t}}{V_{c,t}}, \tag{76}$$

we can write the ratio of MCFs as

$$\frac{\mathrm{MCF}_{t+j}}{\mathrm{MCF}_t} = \prod_{k=1}^{j} \frac{R_{t+k}}{R^*_{t+k}}.$$

Thus, from Proposition 2 we see that production damages are perfectly internalized if the capital tax is optimally set to zero for all future periods where current emissions generate production damages.

Price effect To understand the role of the IES, it is useful to go back to the origin of the term $IC_{i,t}$. This term comes from households' budget constraint (2) in which we have substituted for the price and real wage using (36) and (37). From these equations, it appears that when making more resources available to households, the price goes down since

$$p_t = \beta^t \Big(\frac{c_t}{c_0}\Big)^{-\sigma} \Big(\frac{1 - \varsigma h_t}{1 - \varsigma h_0}\Big)^{\gamma(1-\sigma)}.$$

When σ tends to 1, the price effect exactly offsets the volume effect so that households' expenditures and nominal income remain unchanged after an inflow of aggregate consumption, hence the planner does not need to change the value of the lump-sum transfer and the implementation cost remains constant.

Labor supply effect To determine the sign of the covariance term driving the MCF, we can examine the ratio of the period implementation cost for two agents i and j such that $e_i > e_j$. From (73), we have

$$\frac{IC_{i,t}}{IC_{j,t}} = \frac{\omega_i + \gamma\Big(\omega_i - \frac{e_i}{(1-\varsigma h_t)}\Big)}{\omega_j + \gamma\Big(\omega_j - \frac{e_j}{(1-\varsigma h_t)}\Big)}.$$

Although the discounted sum of $IC_{i,t}$ is invariant across type, in period t this ratio may be below or above 1 depending on the value of the aggregate labor supply. In particular, we have

$$\frac{\partial \frac{IC_{i,t}}{IC_{j,t}}}{\partial h_t} = \frac{\varsigma\gamma(1+\gamma)\big(e_j\omega_i - e_i\omega_j\big)}{(1-\varsigma h_t)^2\big(\omega_j(1+\gamma) - \frac{\gamma e_j}{(1-\varsigma h_t)}\big)^2}. \tag{77}$$

From (71), we can also show that with homogeneous initial wealth (or full expropriation of initial wealth), when transfers plus initial assets are positive (as they are in our quantitative analysis) then

ω_i/e_i is strictly declining in e_i, hence for $e_i > e_j$, the derivative in (77) is negative. This result means that when h_t is high relative to its average value, the relative labor supply of highly productive households compared to less productive households is higher, hence more productive households need lower transfers to satisfy the planners' allocation at that period. If the more productive also have a lower marginal utility of consumption (hence a higher θ_i), then $\text{cov}(\theta_i, IC_{i,t}) < 0$. Thus, when the IES is less than unity so that the price effect dominates, an increase in aggregate consumption reduces the planner's implementation cost and the MCF is higher than 1 in a given period if and only if the labor supply is relatively high compared to its long-run value.

Proof of Proposition 3: From our characterization problem, we know that market weights are determined by the following expression,

$$\varphi_i u_{c,i,t} = U_{c,t}, \quad \forall i,$$

hence substituting into equation (75) and using the fact that for any period t, $\omega_i = c_{i,t}/c_t$ and $\sum_i \pi_i (c_{i,t}/c_t) = 1$, we have

$$\begin{aligned} V_{c,t} &= \sum_i \pi_i \lambda_i \frac{u_{c,i,t} c_{i,t}}{c_t} \\ &= \sum_i \pi_i \lambda_i u_{c,i,t} + \text{cov}\Big(\lambda_i u_{c,i,t}, \frac{c_{i,t}}{c_t}\Big). \end{aligned} \tag{78}$$

Thus, between the first-best and the second-best case, the marginal utility of consumption will differ due to the path of aggregate consumption, as well as the distribution of individual allocations. Holding aggregate consumption constant, we see that an increase in the variance of $c_{i,t}$ has ambiguous effects. On the one hand, since u_c is convex in c for $\sigma > 0$, from Jensen's inequality the average marginal utility is increasing with consumption inequalities. On the other hand, higher marginal utilities are weighted by lower consumption levels, hence increasing consumption dispersion reduces the relative weight given to high marginal utilities. The net effect depends on the curvature of the utility function. From (62), when σ tends to 1 we have $\omega_i = \varphi_i$, hence from (75) and using the normalization of the planner's weights we have

$$V_{c,t} = U_{c,t}.$$

Thus, when σ tends to 1, the two previous effects cancel each other and the distribution of individual allocations has no incidence on the marginal utility of consumption. ■

B Optimal tax rules in a third-best environment

B.1 Given labor taxes

Suppose that labor taxes are given, $\tau_{H,t} = \bar{\tau}_H$. Then, the planner's problem described in Appendix A.2.1 has the following additional constraints,

$$\frac{U_{h,t}}{U_{c,t}} = -\left(1 - \bar{\tau}_H\right)\left(1 - D_t\right) A_{1,t} F_{H,t}, \tag{79}$$

$$F_{H,t} G_{K,t} = F_{K,t} G_{H,t}. \tag{80}$$

Although the second of these two constraints is already required in the benchmark model, it happens to be endogenously satisfied in that case. With an additional constraint on instruments, this is not necessarily the case anymore. Let $\beta^t \Lambda_t^H$ and $\beta^t \Gamma_t^H$ be the multipliers on these constraints. Then, the first-order conditions of the planner's problem become

$$[c_t] : W_{c,t} - \nu_{1,t} + \Lambda_t^H \vartheta_{c,t} = 0, \quad \forall\, t \geq 0, \tag{81}$$

$$\begin{aligned}[H_{1,t}] : W_{h,t} &+ \nu_{1,t} (1 - D_t) A_{1,t} F_{H,t} + \Lambda_t^H \Big(\vartheta_{h,t} + (1 - \bar{\tau}_H) (1 - D_t) A_{1,t} F_{HH,t} \Big) \\ &+ \Gamma_t^H \Big(F_{HH,t} G_{K,t} - F_{KH,t} G_{H,t} \Big) = 0, \quad \forall\, t \geq 0, \end{aligned} \tag{82}$$

$$[H_{2,t}] : W_{h,t} + \nu_{2,t} A_{2,t} G_{H,t} + \Lambda_t^H \vartheta_{h,t} + \Gamma_t^H \left(F_{H,t} G_{KH,t} - F_{K,t} G_{HH,t} \right) = 0, \quad \forall\, t \geq 0, \tag{83}$$

$$\begin{aligned}[K_{1,t+1}] : -\nu_{1,t} &+ \Big((1 - D_{t+1}) A_{1,t+1} F_{K,t+1} + (1 - \delta) \Big) \beta \nu_{1,t+1} + \beta \Lambda_{t+1}^H \Big((1 - \bar{\tau}_H) (1 - D_{t+1}) A_{1,t+1} F_{HK,t+1} \Big) \\ &+ \beta \Gamma_{t+1}^H \Big(F_{HK,t+1} G_{K,t+1} - F_{KK,t+1} G_{H,t+1} \Big) = 0, \quad \forall\, t \geq 0, \end{aligned} \tag{84}$$

$$\begin{aligned}[K_{2,t+1}] : -\nu_{1,t} &+ A_{2,t+1} G_{K,t+1} \beta \nu_{2,t+1} + (1 - \delta) \beta \nu_{1,t+1} \\ &+ \beta \Gamma_{t+1}^H \left(F_{H,t+1} G_{KK,t+1} - F_{K,t+1} G_{HK,t+1} \right) = 0, \quad \forall\, t \geq 0, \end{aligned} \tag{85}$$

$$\begin{aligned}[E_t] : -\nu_{1,t} &\left(\Theta_{E,t} - (1 - D_t) A_{1,t} F_{E,t} \right) - \nu_{2,t} - \sum_{j=0}^{\infty} \beta^j \nu_{3,t+j} J_{E_t^M, t+j} (1 - \mu_t) \\ &+ \Lambda_t^H \left((1 - \bar{\tau}_H) (1 - D_t) A_{1,t} F_{HE,t} \right) + \Gamma_t^H \left(F_{HE,t} G_{K,t} - F_{KE,t} G_{H,t} \right) = 0, \quad \forall\, t \geq 0, \end{aligned} \tag{86}$$

$$[Z_t] : N_t W_{Z,t} - \nu_{1,t} D'_t A_{1,t} F_t + \nu_{3,t} - \Lambda_t^H (1 - \bar{\tau}_H) D'_t A_{1,t} F_{H,t} = 0, \quad \forall\, t \geq 0, \tag{87}$$

$$[\mu_t] : -\nu_{1,t} \Theta_{\mu,t} + \sum_{j=0}^{\infty} \beta^j \nu_{3,t+j} J_{E_t^M, t+j} E_t = 0, \quad \forall\, t \geq 0, \tag{88}$$

$$[T] : \sum_i \pi_i \theta_i = 0, \tag{89}$$

$$[\varphi_i] : \sum_t \beta^t N_t W_{\varphi_i,t} - \frac{\zeta}{\sigma - (1 - \sigma) \gamma} \frac{\pi_i \omega_i}{\varphi_i} = 0, \tag{90}$$

where

$$\vartheta_{c,t} \equiv \frac{U_{ch,t} U_{c,t} - U_{h,t} U_{cc,t}}{N_t U_{c,t}^2},$$

$$\vartheta_{h,t} \equiv \frac{U_{hh,t} U_{c,t} - U_{h,t} U_{ch,t}}{N_t U_{c,t}^2}.$$

B.1.1 Capital income taxes and multipliers on new constraints

From (82) and (81) we obtain

$$(1 - D_t) A_{1,t} F_{H,t} = - \frac{W_{h,t} + \Lambda_t^H \Big(\vartheta_{h,t} + (1 - \bar{\tau}_H) (1 - D_t) A_{1,t} F_{HH,t} \Big) + \Gamma_t^H \Big(G_{K,t} F_{HH,t} - G_{H,t} F_{KH,t} \Big)}{W_{c,t} + \Lambda_t^H \vartheta_{c,t}}, \quad \forall\, t \geq 0, \tag{91}$$

and using the intertemporal condition (84) we get

$$R^*_{t+1} \equiv 1 + r_{t+1} - \delta$$
$$= \frac{1}{\beta}\frac{W_{c,t} + \Lambda_t \vartheta_{c,t} - \beta\Lambda^H_{t+1}\Big(\left(1-\bar{\tau}_H\right)\left(1-D_{t+1}\right)A_{1,t+1}F_{HK,t+1}\Big) - \beta\Gamma_{t+1}\left(F_{HK,t+1}G_{K,t+1} - F_{KK,t+1}G_{H,t+1}\right)}{W_{c,t+1} + \Lambda^H_{t+1}\vartheta_{c,t+1}}, \tag{92}$$

Solving (82) and (83), and (84) and (85) for $\nu_{2,t}/\nu_{1,t}$, and equating both equations, using (80), yields

$$\Gamma^H_t = \zeta_t \Lambda^H_t,$$

where

$$\zeta_t \equiv \frac{\left(1-\bar{\tau}_H\right)\left(1-D_t\right)A_{1,t}\left(G_{K,t}F_{HH,t} - G_{H,t}F_{KH,t}\right)}{\left\{\begin{array}{c} G_{H,t}\Big(\left(F_{KH,t}G_{K,t} - F_{KK,t}G_{H,t}\right) - \left(F_{H,t}G_{KK,t} - F_{K,t}G_{KH,t}\right)\Big) \\ -G_{K,t}\Big(\left(F_{HH,t}G_{K,t} - F_{KH,t}G_{H,t}\right) - \left(F_{H,t}G_{KH,t} - F_{K,t}G_{HH,t}\right)\Big)\end{array}\right\}}, \quad \forall\, t \geq 0.$$

Combining this equation with (91) we can then solve for

$$\Lambda^H_t = -\frac{W_{h,t} + \left(1-D_t\right)A_{1,t}F_{H,t}W_{c,t}}{\vartheta_{h,t} + \left(1-D_t\right)A_{1,t}F_{H,t}\vartheta_{c,t} + \left(1-\bar{\tau}_H\right)\left(1-D_t\right)A_{1,t}F_{HH,t} + \zeta_t\left(G_{K,t}F_{HH,t} - G_{H,t}F_{KH,t}\right)}.$$

In any competitive equilibrium (37) holds, which together with $p_t = R_t p_{t+1}$ implies

$$\frac{U_{c,t+1}}{U_{c,t}}\beta R_{t+1} = 1.$$

Substituting this into (92), it follows that

$$\frac{R_{t+1}}{R^*_{t+1}} = \frac{W_{c,t+1} + \Lambda^H_{t+1}\vartheta_{c,t+1}}{\left\{\begin{array}{c} W_{c,t} + \Lambda^H_t\vartheta_{c,t} - \beta\Lambda^H_{t+1}\Big(\left(1-\bar{\tau}_H\right)\left(1-D_{t+1}\right)A_{1,t+1}F_{HK,t+1}\Big) \\ -\beta\Gamma^H_{t+1}\Big(F_{HK,t+1}G_{K,t+1} - F_{KK,t+1}G_{H,t+1}\Big)\end{array}\right\}}\frac{U_{c,t}}{U_{c,t+1}}.$$

B.1.2 Excise taxes of energy and emissions

From (9) and the abatement first-order condition (88) we have that

$$\tau_{E,t} = \frac{\Theta_{\mu,t}}{E_t} = \frac{1}{\nu_{1,t}}\sum_{j=0}^{\infty}\beta^j \nu_{3,t+j} J_{E^M_t,t+j}. \tag{93}$$

From the climate variable first-order condition (87) we have that

$$\nu_{3,t} = \nu_{1,t}D'_t A_{1,t}F_t - W_{Z,t} + \Lambda^H_t\left(1-\bar{\tau}_H\right)D'_t A_{1,t}F_{H,t}. \tag{94}$$

Substituting (94) into (93) we obtain the optimal pollution tax

$$\tau_{E,t} = \frac{1}{\nu_{1,t}}\sum_{j=0}^{\infty}\beta^j\Big(\nu_{1,t+j}D'_{t+j}A_{1,t+j}F_{t+j} - N_{t+j}W_{Z,t+j} + \Lambda^H_{t+j}\left(\left(1-\bar{\tau}_H\right)D'_{t+j}A_{1,t+j}F_{H,t+j}\right)\Big)J_{E^M_t,t+j}.$$

From the energy first-order condition (86) we have that

$$(1-D_t)\,A_{1,t}F_{E,t}-\frac{\nu_{2,t}}{\nu_{1,t}}=\big(\Theta_{E,t}+(1-\mu_t)\tau_{E,t}\big)-\frac{\Lambda_t^H}{\nu_{1,t}}\,(1-\bar{\tau}_H)\,(1-D_t)\,A_{1,t}F_{HE,t}-\frac{\Gamma_t^H}{\nu_{1,t}}\,(F_{HE,t}G_{K,t}-F_{KE,t}G_{H,t})\,.$$

Combining the first-order conditions for sectoral labor supplies (82) and (83), it follows that

$$\begin{aligned}\frac{\nu_{2,t}}{\nu_{1,t}}=&\frac{(1-D_t)\,A_{1,t}F_{H,t}}{A_{2,t}G_{H,t}}+\frac{\Lambda_t^H}{\nu_{1,t}}\frac{(1-\bar{\tau}_H)\,(1-D_t)\,A_{1,t}F_{HH,t}}{A_{2,t}G_{H,t}}\\&+\frac{\Gamma_t^H}{\nu_{1,t}}\frac{(F_{HH,t}G_{K,t}-F_{KH,t}G_{H,t})-(F_{H,t}G_{KH,t}-F_{K,t}G_{HH,t})}{A_{2,t}G_{H,t}},\end{aligned}$$

and, therefore

$$\begin{aligned}(1-D_t)\,A_{1,t}F_{E,t}=&\big(\Theta_{E,t}+(1-\mu_t)\tau_{E,t}\big)+\frac{(1-D_t)\,A_{1,t}F_{H,t}}{A_{2,t}G_{H,t}}+\frac{\Lambda_t^H}{\nu_{1,t}}\,(1-\bar{\tau}_H)\,(1-D_t)\,A_{1,t}\left(\frac{F_{HH,t}}{A_{2,t}G_{H,t}}-F_{HE,t}\right)\\&+\frac{\Gamma_t^H}{\nu_{1,t}}\left(\frac{(F_{HH,t}G_{K,t}-F_{KH,t}G_{H,t})-(F_{H,t}G_{KH,t}-F_{K,t}G_{HH,t})}{A_{2,t}G_{H,t}}-(F_{HE,t}G_{K,t}-F_{KE,t}G_{H,t})\right).\end{aligned}$$

Then, from (4), (5), and (8) we have that

$$(1-D_t)\,A_{1,t}F_{H,t}=\Big((1-D_t)\,A_{1,t}F_{E,t}-\tau_{I,t}-\big(\Theta_{E,t}+(1-\mu_t)\tau_{E,t}\big)\Big)A_{2,t}G_{H,t},$$

and therefore

$$\begin{aligned}\tau_{I,t}=&\frac{\Lambda_t^H}{\nu_{1,t}}\,(1-\bar{\tau}_H)\,(1-D_t)\,A_{1,t}\left(\frac{F_{HH,t}}{A_{2,t}G_{H,t}}-F_{HE,t}\right)\\&+\frac{\Gamma_t^H}{\nu_{1,t}}\left(\frac{(F_{HH,t}G_{K,t}-F_{KH,t}G_{H,t})-(F_{H,t}G_{KH,t}-F_{K,t}G_{HH,t})}{A_{2,t}G_{H,t}}-(F_{HE,t}G_{K,t}-F_{KE,t}G_{H,t})\right).\end{aligned}$$

B.2 Given capital taxes

Now suppose that labor capital are given, $\tau_{K,t}=\bar{\tau}_K$. Then, the planner's problem has the following additional constraints,

$$\frac{U_{c,t}}{U_{c,t+1}}=\beta\Big(1+(1-\bar{\tau}_K)\,((1-D_{t+1})\,A_{1,t+1}F_{K,t+1}-\delta)\Big),\tag{95}$$

$$F_{H,t}G_{K,t}=F_{K,t}G_{H,t}.\tag{96}$$

Let $\beta^t \Lambda_{t+1}^K$ and $\beta^t \Gamma_t^K$ be the multipliers on these constraints. Then the first-order conditions of the planner's problem become

$$[c_t] : W_{c,t} - \nu_{1,t} - \frac{\Lambda_{t+1}^K}{N_t} \frac{U_{cc,t}}{U_{c,t+1}} + \frac{\Lambda_t^K}{N_t} \frac{U_{c,t-1} U_{cc,t}}{\beta U_{c,t}^2} = 0, \quad \forall\, t \geq 0, \tag{97}$$

$$\begin{aligned} [H_{1,t}] : W_{h,t} + \nu_{1,t} \left(1 - D_t\right) A_{1,t} F_{H,t} - \frac{\Lambda_{t+1}^K}{N_t} \frac{U_{ch,t}}{U_{c,t+1}} + \frac{\Lambda_t^K}{N_t} \frac{U_{c,t-1} U_{ch,t}}{\beta U_{c,t}^2} \\ + \Lambda_t^K \left(1 - \bar{\tau}_K\right) \left(1 - D_t\right) A_{1,t} F_{KH,t} + \Gamma_t^K \Big(F_{HH,t} G_{K,t} - F_{KH,t} G_{H,t} \Big) = 0, \quad \forall\, t \geq 0, \end{aligned} \tag{98}$$

$$\begin{aligned} [H_{2,t}] : W_{h,t} + \nu_{2,t} A_{2,t} G_{H,t} - \frac{\Lambda_{t+1}^K}{N_t} \frac{U_{ch,t}}{U_{c,t+1}} + \frac{\Lambda_t^K}{N_t} \frac{U_{c,t-1} U_{ch,t}}{\beta U_{c,t}^2} \\ + \Gamma_t^K \Big(F_{H,t} G_{KH,t} - F_{K,t} G_{HH,t} \Big) = 0, \quad \forall\, t \geq 0, \end{aligned} \tag{99}$$

$$\begin{aligned} [K_{1,t+1}] : -\nu_{1,t} + \Big(\left(1 - D_{t+1}\right) A_{1,t+1} F_{K,t+1} + (1 - \delta) \Big) \beta \nu_{1,t+1} + \beta \Lambda_{t+1}^K \Big(\left(1 - \bar{\tau}_K\right) \left(1 - D_{t+1}\right) A_{1,t+1} F_{KK,t+1} \Big) \\ + \beta \Gamma_{t+1}^K \Big(F_{HK,t+1} G_{K,t+1} - F_{KK,t+1} G_{H,t+1} \Big) = 0, \quad \forall\, t \geq 0, \end{aligned} \tag{100}$$

$$\begin{aligned} [K_{2,t+1}] : -\nu_{1,t} + A_{2,t+1} G_{K,t+1} \beta \nu_{2,t+1} + (1 - \delta) \beta \nu_{1,t+1} \\ + \beta \Gamma_{t+1}^K \Big(F_{H,t+1} G_{KK,t+1} - F_{K,t+1} G_{HK,t+1} \Big) = 0, \quad \forall\, t \geq 0, \end{aligned} \tag{101}$$

$$\begin{aligned} [E_t] : -\nu_{1,t} \left(\Theta_{E,t} - (1 - D_t) A_{1,t} F_{E,t} \right) - \nu_{2,t} - \sum_{j=0}^{\infty} \beta^j \nu_{3,t+j} J_{E_t^M, t+j} \left(1 - \mu_t\right) \\ + \Lambda_t^K \Big(\left(1 - \bar{\tau}_K\right) \left(1 - D_t\right) A_{1,t} F_{KE,t} \Big) + \Gamma_t^K \Big(F_{HE,t} G_{K,t} - F_{KE,t} G_{H,t} \Big) = 0, \quad \forall\, t \geq 0, \end{aligned} \tag{102}$$

$$[Z_t] : N_t W_{Z,t} - \nu_{1,t} D_t' A_{1,t} F_t + \nu_{3,t} - \Lambda_t^K \left(\left(1 - \bar{\tau}_K\right) D_t' A_{1,t} F_{K,t} \right) = 0, \quad \forall\, t \geq 0, \tag{103}$$

$$[\mu_t] : -\nu_{1,t} \Theta_{\mu,t} + \sum_{j=0}^{\infty} \beta^j \nu_{3,t+j} J_{E_t^M, t+j} E_t = 0, \quad \forall\, t \geq 0, \tag{104}$$

$$[T] : \sum_i \pi_i \theta_i = 0, \tag{105}$$

$$[\varphi_i] : \sum_t \beta^t W_{\varphi_i, t} - \frac{\zeta}{\sigma - (1 - \sigma) \gamma} \frac{\pi_i \omega_i}{\varphi_i} = 0. \tag{106}$$

B.2.1 Capital income taxes and multipliers on new constraints

From(98) and (97) we obtain

$$\left(1 - D_t\right) A_{1,t} F_{H,t} = \frac{ - \left\{ \begin{array}{c} W_{h,t} - \frac{\Lambda_{t+1}^K}{N_t} \frac{U_{ch,t}}{U_{c,t+1}} + \frac{\Lambda_t^K}{N_t} \frac{U_{c,t-1} U_{ch,t}}{\beta U_{c,t}^2} + \Lambda_t^K \left(1 - \bar{\tau}_K\right) \left(1 - D_t\right) A_{1,t} F_{KH,t} \\ + \Gamma_t^K \left(F_{HH,t} G_{K,t} - F_{KH,t} G_{H,t} \right) \end{array} \right\} }{ W_{c,t} - \frac{\Lambda_{t+1}^K}{N_t} \frac{U_{cc,t}}{U_{c,t+1}} + \frac{\Lambda_t^K}{N_t} \frac{U_{c,t-1} U_{cc,t}}{\beta U_{c,t}^2} }, \tag{107}$$

and using the intertemporal condition (100) we get

$$R_{t+1}^{*} \equiv 1+r_{t+1}-\delta$$
$$=\frac{1}{\beta}\frac{\left\{\begin{array}{c} W_{c,t}-\frac{\Lambda_{t+1}^{K}}{N_t}\frac{U_{cc,t}}{U_{c,t+1}}+\frac{\Lambda_t^K}{N_t}\frac{U_{c,t-1}U_{cc,t}}{\beta U_{c,t}^2}-\beta\Lambda_{t+1}^{K}\left(1-\bar{\tau}_K\right)\left(1-D_{t+1}\right)A_{1,t+1}F_{KK,t+1} \\ -\beta\Gamma_{t+1}^{K}\left(F_{HK,t+1}G_{K,t+1}-F_{KK,t+1}G_{H,t+1}\right) \end{array}\right\}}{W_{c,t+1}-\frac{\Lambda_{t+2}^{K}}{N_{t+1}}\frac{U_{cc,t+1}}{U_{c,t+2}}+\frac{\Lambda_{t+1}^{K}}{N_{t+1}}\frac{U_{c,t}U_{cc,t+1}}{\beta U_{c,t+1}^2}}. \tag{108}$$

Solving (82) and (83), and (84) and (85) for $\nu_{2,t}/\nu_{1,t}$, and equating both equations, using (96), yields

$$\Gamma_t^K=\zeta_t\Lambda_t^K,$$

where

$$\zeta_t \equiv \frac{\left(1-\bar{\tau}_K\right)\left(1-D_t\right)A_{1,t}\left(G_{K,t}F_{KH,t}-G_{H,t}F_{KK,t}\right)}{\left\{\begin{array}{c} G_{H,t}\Big(\left(F_{KH,t}G_{K,t}-F_{KK,t}G_{H,t}\right)-\left(F_{H,t}G_{KK,t}-F_{K,t}G_{KH,t}\right)\Big) \\ -G_{K,t}\Big(\left(F_{HH,t}G_{K,t}-F_{KH,t}G_{H,t}\right)-\left(F_{H,t}G_{KH,t}-F_{K,t}G_{HH,t}\right)\Big) \end{array}\right\}}, \quad \forall\, t\geq 0.$$

In any competitive equilibrium (37) holds, which together with $p_t=R_tp_{t+1}$ implies

$$\frac{U_{c,t+1}}{U_{c,t}}\beta R_{t+1}=1.$$

Substituting this into (108), it follows that

$$\beta R_{t+1}^{*}=\frac{\left\{\begin{array}{c} W_{c,t}+\frac{\left(R_t\Lambda_t^K-\beta R_{t+1}\Lambda_{t+1}^K\right)}{N_t}\frac{U_{cc,t}}{U_{c,t}}-\beta\Lambda_{t+1}^{K}\left(1-\bar{\tau}_K\right)\left(1-D_{t+1}\right)A_{1,t+1}F_{KK,t+1} \\ -\beta\Lambda_{t+1}^{K}\zeta_{t+1}\left(F_{HK,t+1}G_{K,t+1}-F_{KK,t+1}G_{H,t+1}\right) \end{array}\right\}}{W_{c,t+1}+\frac{\left(R_{t+1}\Lambda_{t+1}^K-\beta R_{t+2}\Lambda_{t+2}^K\right)}{N_{t+1}}\frac{U_{cc,t+1}}{U_{c,t+1}}}.$$

Plugging (107) into (36) implies

$$\frac{U_{h,t}}{U_{c,t}}=\frac{\left\{\begin{array}{c} W_{h,t}-\frac{\Lambda_{t+1}^{K}}{N_t}\frac{U_{ch,t}}{U_{c,t+1}}+\frac{\Lambda_t^K}{N_t}\frac{U_{c,t-1}U_{ch,t}}{\beta U_{c,t}^2}+\Lambda_t^K\left(1-\bar{\tau}_K\right)\left(1-D_t\right)A_{1,t}F_{KH,t} \\ +\Gamma_t^K\left(F_{HH,t}G_{K,t}-F_{KH,t}G_{H,t}\right) \end{array}\right\}}{W_{c,t}-\frac{\Lambda_{t+1}^{K}}{N_t}\frac{U_{cc,t}}{U_{c,t+1}}+\frac{\Lambda_t^K}{N_t}\frac{U_{c,t-1}U_{cc,t}}{\beta U_{c,t}^2}}\left(1-\tau_{H,t}\right),$$

which can be rearranged into

$$\tau_{H,t}=1-\frac{U_{h,t}}{U_{c,t}}\frac{W_{c,t}-\frac{\Lambda_{t+1}^{K}}{N_t}\frac{U_{cc,t}}{U_{c,t+1}}+\frac{\Lambda_t^K}{N_t}\frac{U_{c,t-1}U_{cc,t}}{\beta U_{c,t}^2}}{\left\{\begin{array}{c} W_{h,t}-\frac{\Lambda_{t+1}^{K}}{N_t}\frac{U_{ch,t}}{U_{c,t+1}}+\frac{\Lambda_t^K}{N_t}\frac{U_{c,t-1}U_{ch,t}}{\beta U_{c,t}^2} \\ +\Lambda_t^K\left(1-\bar{\tau}_K\right)\left(1-D_t\right)A_{1,t}F_{KH,t}+\Gamma_t^K\left(F_{HH,t}G_{K,t}-F_{KH,t}G_{H,t}\right) \end{array}\right\}}. \tag{109}$$

B.2.2 Excise taxes of energy and emissions

From (9) and the abatement first-order condition (104) we have that

$$\tau_{E,t} = \frac{\Theta_{\mu,t}}{E_t} = \frac{1}{\nu_{1,t}} \sum_{j=0}^{\infty} \beta^j \nu_{3,t+j} J_{E_t^M,t+j}. \tag{110}$$

From the climate variable first-order condition (103) we have that

$$\nu_{3,t} = \nu_{1,t} D_t' A_{1,t} F_t - N_t W_{Z,t} + \Lambda_t^K \left(1 - \bar{\tau}_K\right) D_t' A_{1,t} F_{K,t}. \tag{111}$$

Substituting (111) into (110) we obtain the optimal pollution tax

$$\tau_{E,t} = \frac{1}{\nu_{1,t}} \sum_{j=0}^{\infty} \beta^j \Big(\nu_{1,t+j} D_{t+j}' A_{1,t+j} F_{t+j} - N_{t+j} W_{Z,t+j} + \Lambda_{t+j}^K \left(\left(1 - \bar{\tau}_K\right) D_{t+j}' A_{1,t+j} F_{K,t+j} \right) \Big) J_{E_t^M,t+j}.$$

From the energy first-order condition (102) we have that

$$\left(1 - D_t\right) A_{1,t} F_{E,t} - \frac{\nu_{2,t}}{\nu_{1,t}} = \left(\Theta_{E,t} + (1-\mu_t)\tau_{E,t}\right) - \frac{\Lambda_t^K}{\nu_{1,t}} \left(1 - \bar{\tau}_K\right) \left(1 - D_t\right) A_{1,t} F_{KE,t} - \frac{\Gamma_t^K}{\nu_{1,t}} \left(F_{HE,t} G_{K,t} - F_{KE,t} G_{H,t}\right).$$

Combining the first-order conditions for sectoral labor supplies (98) and (99), it follows that

$$\begin{aligned}\frac{\nu_{2,t}}{\nu_{1,t}} =& \frac{\left(1 - D_t\right) A_{1,t} F_{H,t}}{A_{2,t} G_{H,t}} + \frac{\Lambda_t^K}{\nu_{1,t}} \frac{\left(1 - \tau_K\right) \left(1 - D_t\right) A_{1,t} F_{KH,t}}{A_{2,t} G_{H,t}} \\ &+ \frac{\Gamma_t^K}{\nu_{1,t}} \frac{\left(F_{HH,t} G_{K,t} - F_{KH,t} G_{H,t}\right) - \left(F_{H,t} G_{KH,t} - F_{K,t} G_{HH,t}\right)}{A_{2,t} G_{H,t}},\end{aligned}$$

and, therefore

$$\begin{aligned}\left(1 - D_t\right) A_{1,t} F_{E,t} =& \left(\Theta_{E,t} + (1 - \mu_t)\tau_{E,t}\right) + \frac{\left(1 - D_t\right) A_{1,t} F_{H,t}}{A_{2,t} G_{H,t}} + \frac{\Lambda_t^K}{\nu_{1,t}} \left(1 - \bar{\tau}_K\right) \left(1 - D_t\right) A_{1,t} \left(\frac{F_{KH,t}}{A_{2,t} G_{H,t}} - F_{KE,t} \right) \\ &+ \frac{\Gamma_t^K}{\nu_{1,t}} \left(\frac{\left(F_{HH,t} G_{K,t} - F_{KH,t} G_{H,t}\right) - \left(F_{H,t} G_{KH,t} - F_{K,t} G_{HH,t}\right)}{A_{2,t} G_{H,t}} - \left(F_{HE,t} G_{K,t} - F_{KE,t} G_{H,t}\right) \right).\end{aligned}$$

Then, from (4), (5), and (8) we have that

$$\left(1 - D_t\right) A_{1,t} F_{H,t} = \Big(\left(1 - D_t\right) A_{1,t} F_{E,t} - \tau_{I,t} - \left(\Theta_{E,t} + (1 - \mu_t)\tau_{E,t}\right) \Big) A_{2,t} G_{H,t},$$

and therefore

$$\begin{aligned}\tau_{I,t} =& \frac{\Lambda_t^K}{\nu_{1,t}} \left(1 - \bar{\tau}_K\right) \left(1 - D_t\right) A_{1,t} \left(\frac{F_{KH,t}}{A_{2,t} G_{H,t}} - F_{KE,t} \right) \\ &+ \frac{\Gamma_t^K}{\nu_{1,t}} \left(\frac{\left(F_{HH,t} G_{K,t} - F_{KH,t} G_{H,t}\right) - \left(F_{H,t} G_{KH,t} - F_{K,t} G_{HH,t}\right)}{A_{2,t} G_{H,t}} - \left(F_{HE,t} G_{K,t} - F_{KE,t} G_{H,t}\right) \right).\end{aligned}$$

C Optimal tax rules with initial wealth inequality

In Appendix A.2.1, we describe the Ramsey problem with wealth inequality.

C.1 Time-inconsistency

The tax rules we have described in our benchmark apply unchanged for every period including period 0. This is the result of two features of the benchmark model. The first is the ability of the Ramsey planner to choose lump-sum transfers (or taxes), and the second is the assumption that the planner can set the initial capital tax to expropriate initial wealth, thereby eliminating any initial wealth inequality. To see this, notice that the planner's problem (see equation (39)) is symmetric with respect to time except for the last term in the objective function of the Ramsey planner, which we denote here by W_0,

$$W_0 \equiv U_{c,0} \sum_i \pi_i \theta_i \left(N_0 R_0 a_{i,0} + T \right).$$

As argued above, the optimality condition associated with the choice of T implies that $\sum_i \pi_i \theta_i = 0$. Thus, if there is no initial wealth inequality, *i.e.* if $a_{i,0} = a_0$ for every i, it follows that $W_0 = 0$ and that the tax rules are time invariant. Moreover, if there is initial wealth inequality, the planner can set $\tau_{K,0}$ such that $R_0 = 0$, and we again have $W_0 = 0$.

This does not mean that the tax rules are time-consistent: if the Ramsey planner was allowed to re-optimize in a future period, they would want to deviate from the choices made by the planner in period 0. The reason for the time-inconsistency is, however, different from the one in the usual representative-agent version of the Ramsey problem in which the planner cannot choose lump-sum transfers. In that case, in general $\sum_i \pi_i \theta_i \neq 0$, and $W_0 \neq 0$ regardless of initial wealth inequality, which leads to the usual reason for time-inconsistent Ramsey policies; initial capital income taxes mimic the unavailable and undistortive lump-sum taxes. In our setup, the reason for time inconsistency has to do instead with the use of capital income taxes to redistribute unequal asset income. Since asset inequality evolves endogenously over time, starting the Ramsey problem in a future period would mean having a different initial asset distribution.

There is a sense in which the time-inconsistency problem in our setup is less severe than in the usual representative agent case. If there is no initial wealth inequality, and the optimal Ramsey policy was such that the economy was in a balanced-growth path starting from period 0, then there would still be no wealth inequality in every future period and the Ramsey policy would be time-consistent. In any case, in this section we address how the Ramsey policy changes in the presence of initial wealth inequality.

C.2 First order conditions

Here we consider the problem of the planner assuming that $\tau_{K,0}$ is taken as given. For $t \geq 1$, the conditions are exactly the same as the ones derived above, in particular, we have that $\sum_i \pi_i \theta_i = 0$, which we use to simplify the equations below. The period-0 marginal rate of technical substitution constraint is no longer automatically satisfied, so let Γ_0 denote the Lagrange multiplier on this constraint. The first order conditions for period 0 are

$$[c_0]: W_{c,0} - \nu_{1,0} - U_{cc,0} \sum_i \pi_i \theta_i R_0 a_{i,0} = 0, \tag{112}$$

$$\begin{aligned}[H_{1,0}]: &W_{h,0} + \nu_{1,0} (1 - D_0) A_{1,0} F_{H,0} - U_{ch,0} \sum_i \pi_i \theta_i R_0 a_{i,0} \\ &- N_0 U_{c,0} \sum_i \pi_i \theta_i a_{i,0} (1 - \tau_{K,0}) (1 - D_0) A_{1,0} F_{KH,0} + \Gamma_0 (F_{HH,0} G_{K,0} - F_{KH,0} G_{H,0}) = 0,\end{aligned} \tag{113}$$

$$[H_{2,0}]: W_{H,0} + \nu_{2,0} A_{2,0} G_{H,0} - U_{ch,0} \sum_i \pi_i \theta_i R_0 a_{i,0} + \Gamma_0 (F_{H,0} G_{KH,0} - F_{K,0} G_{HH,0}) = 0, \tag{114}$$

$$\begin{aligned}[K_{1,0}]: &((1 - D_0) A_{1,0} F_{K,0} + (1 - \delta)) \nu_{1,0} - N_0 U_{c,0} \sum_i \pi_i \theta_i a_{i,0} (1 - \tau_{K,0}) (1 - D_0) A_{1,0} F_{KK,0} - \kappa \\ &+ \Gamma_0 (F_{HK,0} G_{K,0} - F_{KK,0} G_{H,0}) = 0,\end{aligned} \tag{115}$$

$$[K_{2,0}]: A_{2,0} G_{K,0} \nu_{2,0} + (1 - \delta) \nu_{1,0} - \kappa + \Gamma_0 (F_{H,0} G_{KK,0} - F_{K,0} G_{HK,0}) = 0, \tag{116}$$

$$\begin{aligned}[E_0]: &- (\Theta_{E,0} - (1 - D_0) A_{1,0} F_{E,0}) \nu_{1,0} - \nu_{2,0} - \sum_{j=0}^{\infty} \beta^j \nu_{3,j} J_{E_0^M,j} (1 - \mu_0) \\ &- N_0 U_{c,0} \sum_i \pi_i \theta_i a_{i,0} (1 - \tau_{K,0}) (1 - D_0) A_{1,0} F_{KE,0} + \Gamma_0 (F_{HE,0} G_{K,0} - F_{KE,0} G_{H,0}) = 0,\end{aligned} \tag{117}$$

$$[Z_0]: N_0 W_{Z,0} - \nu_{1,0} D_0' A_{1,0} F_0 + \nu_{3,0} + N_0 U_{c,0} \sum_i \pi_i \theta_i a_{i,0} (1 - \tau_{K,0}) D_0' A_{1,0} F_{K,0} = 0, \tag{118}$$

$$[\mu_0]: -\nu_{1,0} \Theta_{\mu,0} + \sum_{j=0}^{\infty} \beta^j \nu_{3,j} J_{E_0^M,j} E_0 = 0. \tag{119}$$

C.3 Multiplier on period-0 marginal rate of technical substitution constraint

From (115) and (116), it follows that

$$\begin{aligned}\frac{\nu_{2,0}}{\nu_{1,0}} = &(1 - D_0) \frac{A_{1,0} F_{K,0}}{A_{2,0} G_{K,0}} - \frac{N_0 U_{c,0}}{\nu_{1,0}} \sum_i \pi_i \theta_i a_{i,0} (1 - \tau_{K,0}) (1 - D_0) \frac{A_{1,0} F_{KK,0}}{A_{2,0} G_{K,0}} \\ &+ \frac{\Gamma_0}{\nu_{1,0}} \frac{((F_{HK,0} G_{K,0} - F_{KK,0} G_{H,0}) - (F_{H,0} G_{KK,0} - F_{K,0} G_{HK,0}))}{A_{2,0} G_{K,0}}.\end{aligned}$$

From (113) and (114), it follows that

$$\begin{aligned}\frac{\nu_{2,0}}{\nu_{1,0}} = &(1 - D_0) \frac{A_{1,0} F_{H,0}}{A_{2,0} G_{H,0}} - \frac{N_0 U_{c,0}}{\nu_{1,0}} \sum_i \pi_i \theta_i a_{i,0} (1 - \tau_{K,0}) (1 - D_0) \frac{A_{1,0} F_{KH,0}}{A_{2,0} G_{H,0}} \\ &+ \frac{\Gamma_0}{\nu_{1,0}} \frac{((F_{HH,0} G_{K,0} - F_{KH,0} G_{H,0}) - (F_{H,0} G_{KH,0} - F_{K,0} G_{HH,0}))}{A_{2,0} G_{H,0}}.\end{aligned}$$

Hence, putting these two equations together, we obtain

$$\Gamma_0 = \frac{N_0 U_{c,0} \sum_i \pi_i \theta_i a_{i,0} (1 - \tau_{K,0}) (1 - D_0) A_{1,0} (G_{K,0} F_{KH,0} - G_{H,0} F_{KK,0})}{\left\{ \begin{array}{c} G_{K,0} ((F_{HH,0} G_{K,0} - F_{KH,0} G_{H,0}) - (F_{H,0} G_{KH,0} - F_{K,0} G_{HH,0})) \\ -G_{H,0} ((F_{HK,0} G_{K,0} - F_{KK,0} G_{H,0}) - (F_{H,0} G_{KK,0} - F_{K,0} G_{HK,0})) \end{array} \right\}}.$$

C.4 Labor income taxes

From (113) and (112) we obtain

$$(1-D_0)\,A_{1,0}F_{H,0}=\frac{\left\{\begin{array}{c}-W_{h,0}+U_{ch,0}\sum_i\pi_i\theta_iR_0a_{i,0}\\+N_0U_{c,0}\sum_i\pi_i\theta_ia_{i,0}\left(1-\tau_{K,0}\right)\left(1-D_0\right)A_{1,0}F_{KH,0}-\Gamma_0\left(F_{HH,0}G_{K,0}-F_{KH,0}G_{H,0}\right)\end{array}\right\}}{W_{c,0}-U_{cc,0}\sum_i\pi_i\theta_iR_0a_{i,0}} \tag{120}$$

Plugging (120) into (36) implies

$$\frac{U_{h,0}}{U_{c,0}}=\frac{\left\{\begin{array}{c}W_{h,0}-U_{ch,0}\sum_i\pi_i\theta_iR_0a_{i,0}\\-N_0U_{c,0}\sum_i\pi_i\theta_ia_{i,0}\left(1-\tau_{K,0}\right)\left(1-D_0\right)A_{1,0}F_{KH,0}+\Gamma_0\left(F_{HH,0}G_{K,0}-F_{KH,0}G_{H,0}\right)\end{array}\right\}}{W_{c,0}-U_{cc,0}\sum_i\pi_i\theta_iR_0a_{i,0}}\left(1-\tau_{H,0}\right),$$

which can be rearranged into

$$\tau_{H,0}=1-\frac{U_{h,0}}{U_{c,0}}\frac{W_{c,0}-U_{cc,0}\sum_i\pi_i\theta_iR_0a_{i,0}}{\left\{\begin{array}{c}W_{h,0}-U_{ch,0}\sum_i\pi_i\theta_iR_0a_{i,0}\\-N_0U_{c,0}\sum_i\pi_i\theta_ia_{i,0}\left(1-\tau_{K,0}\right)\left(1-D_0\right)A_{1,0}F_{KH,0}+\Gamma_0\left(F_{HH,0}G_{K,0}-F_{KH,0}G_{H,0}\right)\end{array}\right\}}.$$

C.5 Capital income taxes

From (43) and (112) we obtain

$$R_1^*\equiv1+r_1-\delta=\frac{1}{\beta}\frac{W_{c,0}-U_{cc,0}\sum_i\pi_i\theta_iR_0a_{i,0}}{W_{c,1}}.$$

In any competitive equilibrium (37) holds, which implies

$$\frac{U_{c,1}}{U_{c,0}}\beta R_1=1.$$

Substituting this into (53), it follows that

$$\frac{R_1}{R_1^*}=\frac{W_{c,1}}{W_{c,0}-U_{cc,0}\sum_i\pi_i\theta_iR_0a_{i,0}}\frac{U_{c,0}}{U_{c,1}}. \tag{121}$$

C.6 Excise taxes of energy and emissions

From (9) and the abatement first-order condition (119) we have that

$$\tau_{E,0}=\frac{\Theta_{\mu,0}}{E_0}=\frac{1}{\nu_{1,0}}\sum_{j=0}^{\infty}\beta^j\nu_{3,j}J_{E_0^M,j}. \tag{122}$$

From the climate variable first-order condition (118) we have that

$$\nu_{3,0}=\nu_{1,0}D_0'A_{1,0}F_0-N_0W_{Z,0}-N_0U_{c,0}\sum_i\pi_i\theta_ia_{i,0}\left(1-\tau_{K,0}\right)D_0'A_{1,0}F_{K,0}. \tag{123}$$

Substituting (123) into (122) we obtain the initial pollution tax

$$\tau_{E,0} = \frac{1}{\nu_{1,0}} \sum_{j=0}^{\infty} \beta^j \left(\nu_{1,j} D_j' A_{1,j} F_j - N_j W_{Z,j}\right) J_{E_0^M,j} - N_0 \frac{U_{c,0}}{\nu_{1,0}} \sum_i \pi_i \theta_i a_{i,0} \left(1 - \tau_{K,0}\right) D_0' A_{1,0} F_{K,0} J_{E_0^M,0}.$$

From the energy first-order condition (117) we have that

$$\begin{aligned}\left(1 - D_0\right) A_{1,0} F_{E,0} - \frac{\nu_{2,0}}{\nu_{1,0}} = \big(\Theta_{E,0} + (1-\mu_0)\tau_{E,0}\big) + N_0 \frac{U_{c,0}}{\nu_{1,0}} \sum_i \pi_i \theta_i a_{i,0} \left(1 - \tau_{K,0}\right) \left(1 - D_0\right) A_{1,0} F_{KE,0} \\ - \frac{\Gamma_0}{\nu_{1,0}} \left(F_{HE,0} G_{K,0} - F_{KE,0} G_{H,0}\right).\end{aligned}$$

Combining the first-order conditions for sectoral labor supplies (113) and (114), it follows that

$$\begin{aligned}\frac{\nu_{2,0}}{\nu_{1,0}} = \left(1 - D_0\right) \frac{A_{1,0} F_{H,0}}{A_{2,0} G_{H,0}} - \frac{N_0 U_{c,0}}{\nu_{1,0}} \sum_i \pi_i \theta_i a_{i,0} \left(1 - \tau_{K,0}\right) \left(1 - D_0\right) \frac{A_{1,0} F_{KH,0}}{A_{2,0} G_{H,0}} \\ + \frac{\Gamma_0}{\nu_{1,0}} \frac{\left(\left(F_{HH,0} G_{K,0} - F_{KH,0} G_{H,0}\right) - \left(F_{H,0} G_{KH,0} - F_{K,0} G_{HH,0}\right)\right)}{A_{2,0} G_{H,0}},\end{aligned}$$

and, therefore

$$\begin{aligned}\left(1 - D_0\right) A_{1,0} F_{E,0} &= \big(\Theta_{E,0} + (1-\mu_0)\tau_{E,0}\big) + \left(1 - D_0\right) \frac{A_{1,0} F_{H,0}}{A_{2,0} G_{H,0}} \\ &+ N_0 \frac{U_{c,0}}{\nu_{1,0}} \sum_i \pi_i \theta_i a_{i,0} \left(1 - \tau_{K,0}\right) \left(1 - D_0\right) \left(A_{1,0} F_{KE,0} - \frac{A_{1,0} F_{KH,0}}{A_{2,0} G_{H,0}}\right) \\ &+ \frac{\Gamma_0}{\nu_{1,0}} \left(\frac{\left(F_{HH,0} G_{K,0} - F_{KH,0} G_{H,0}\right) - \left(F_{H,0} G_{KH,0} - F_{K,0} G_{HH,0}\right)}{A_{2,0} G_{H,0}} - \left(F_{HE,0} G_{K,0} - F_{KE,0} G_{H,0}\right)\right).\end{aligned}$$

Then, from (4), (5), and (8) we have that

$$\left(1 - D_0\right) A_{1,0} F_{H,0} = \Big(\left(1 - D_0\right) A_{1,0} F_{E,0} - \tau_{I,0} - \big(\Theta_{E,0} + (1-\mu_0)\tau_{E,0}\big)\Big) A_{2,0} G_{H,0},$$

and therefore

$$\begin{aligned}\tau_{I,0} = N_0 \frac{U_{c,0}}{\nu_{1,0}} \sum_i \pi_i \theta_i a_{i,0} \left(1 - \tau_{K,0}\right) \left(1 - D_0\right) \left(A_{1,0} F_{KE,0} - \frac{A_{1,0} F_{KH,0}}{A_{2,0} G_{H,0}}\right) \\ + \frac{\Gamma_0}{\nu_{1,0}} \left(\frac{\left(F_{HH,0} G_{K,0} - F_{KH,0} G_{H,0}\right) - \left(F_{H,0} G_{KH,0} - F_{K,0} G_{HH,0}\right)}{A_{2,0} G_{H,0}} - \left(F_{HE,0} G_{K,0} - F_{KE,0} G_{H,0}\right)\right).\end{aligned}$$

D Optimal tax rules with Stone-Geary utility

The derivation of optimal tax rules in this extended version of the model closely follows the method applied to solve the benchmark model. This appendix highlights the differences with the benchmark presented in Appendix A.

D.1 Characterization of equilibrium

Let $\varphi \equiv \{\varphi_i\}$ be the market weights normalized so that

$$\sum_j \pi_j \left(\varphi_j e_j^{\gamma(\sigma-1)}\right)^{\frac{1}{1-(1+\epsilon+\gamma)(1-\sigma)}} = 1,$$

with $\varphi_i \geq 0$. Then, given aggregate levels c_t, d_t, h_t and Z_t, the individual levels can be found by solving the following static subproblem for each period t:

$$\begin{gathered} U\left(c_t, d_t, h_t, Z_t; \varphi\right) \equiv \max_{c_{i,t}, d_{i,t}, h_{i,t}} \sum_i \pi_i \varphi_i u_i\left(c_{i,t}, d_{i,t}, h_{i,t}, Z_t\right), \\ \text{s.t.} \ \sum_i \pi_i c_{i,t} = c_t, \quad \text{and} \quad \sum_i \pi_i d_{i,t} = d_t, \quad \text{and} \quad \sum_i \pi_i e_i h_{i,t} = h_t. \end{gathered} \tag{124}$$

Using the utility function defined by equation (30) and following the same steps as in Appendix A, we obtain the following solutions for this problem

$$\begin{aligned} c_{i,t}^m\left(c_t, d_t, h_t; \varphi\right) &= \omega_i c_t, \\ d_{i,t}^m\left(c_t, d_t, h_t; \varphi\right) &= \bar{d}_{i,t} + \omega_i\big(d_t - \bar{d}_t\big), \\ 1 - \varsigma h_{i,t}^m\left(c_t, d_t, h_t; \varphi\right) &= \frac{\omega_i}{e_i}(1 - \varsigma h_t), \end{aligned}$$

with $\bar{d}_t = \sum_i \pi_i \bar{d}_{i,t}$, and where

$$\omega_i = \left(\varphi_i e_i^{\gamma(\sigma-1)}\right)^{\frac{1}{1-(1+\epsilon+\gamma)(1-\sigma)}},$$

which enables us to write the aggregate indirect utility in terms of the aggregates and market weights

$$U\left(c_t, d_t, h_t, Z_t\right) = \frac{\left(c_t (d_t - \bar{d}_t)^{\epsilon}\left(1 - \varsigma h_t\right)^{\gamma}\right)^{1-\sigma}}{1-\sigma} + \Gamma \frac{\left(1 + \alpha_0 Z_t^2\right)^{-(1-\sigma)}}{1-\sigma},$$

with $\Gamma \equiv \sum_i \pi_i \varphi_i$.

D.2 Implementability condition

From the first order conditions of problem (124) and applying the envelope theorem we have

$$\begin{aligned} U_{c,t} &= \varphi_i u_{c,i,t}, \\ U_{d,t} &= \varphi_i u_{d,i,t}, \\ U_{h,t} &= \frac{\varphi_i u_{h,i,t}}{e_i}, \end{aligned}$$

which together with the first order conditions of individual agents' problems give

$$\frac{U_{h,t}}{U_{c,t}} = \frac{u_{h,i,t}}{u_{c,i,t} e_{i,t}} = -w_t\left(1 - \tau_{H,t}\right), \tag{125}$$

$$\frac{U_{d,t}}{U_{c,t}} = \frac{u_{d,i,t}}{u_{c,i,t}} = p_{E,t} + \tau_{D,t}, \tag{126}$$

and

$$\frac{U_{c,t}}{U_{c,0}} = \frac{u_{c,i,t}}{u_{c,i,0}} = \frac{p_t}{\beta^t}. \tag{127}$$

Using (125), (126), and (127) to substitute in households' budget constraint (31), we obtain the implementability conditions

$$U_{c,0}\left(R_0 N_0 a_{i,0} + T\right) = \sum_{t=0}^{\infty} N_t \beta^t \Big(U_{c,t} c_{i,t}^m\big(c_t, d_t, h_t; \varphi\big) + U_{d,t} d_{i,t}^m\left(c_t, d_t, h_t; \varphi\right) + U_{h,t} e_i h_{i,t}^m\left(c_t, d_t, h_t; \varphi\right) \Big), \ \ \forall i.$$

D.3 Ramsey problem

Let again $\lambda \equiv \{\lambda_i\}$ be the planner's welfare weight on type i, with $\sum_i \pi_i \lambda_i = 1$. Define the pseudo-utility function

$$\begin{aligned} W\big(c_t, d_t, h_t, Z_t; \varphi, \theta, \lambda\big) \equiv & \sum_i \pi_i \lambda_i u_i\big(c_{i,t}^m\big(c_t, d_t, h_t; \varphi\big), d_{i,t}^m\big(c_t, d_t, h_t; \varphi\big), h_{i,t}^m\big(c_t, d_t, h_t; \varphi\big), Z_t\big) \\ & + \sum_i \pi_i \theta_i \Big[U_{c,t} c_{i,t}^m\big(c_t, d_t, h_t; \varphi\big) + U_{d,t} d_{i,t}^m\big(c_t, d_t, h_t; \varphi\big) + U_{h,t} e_{i,t} h_{i,t}^m\big(c_t, d_t, h_t; \varphi\big) \Big], \end{aligned}$$

where $\pi_i \theta_i$ is the Lagrange multiplier on the implementability constraint of agent i, and $\theta \equiv \{\theta_i\}$. The new Ramsey problem can be written as

$$\max_{\substack{\{c_t, H_{1,t}, H_{2,t}, K_{1,t}, K_{2,t}, \\ d_t, E_{1,t}, Z_t, \mu_t\}_{t=0}^{\infty}, T, \varphi}} \sum_{t,i} N_t \beta^t W\left(c_t, d_t, h_t, Z_t; \varphi, \theta, \lambda\right) - U_{c,0} \sum_i \pi_i \theta_i \left(R_0 N_0 a_{i,0} + T\right),$$

subject to

$$\begin{aligned} N_t c_t + G_t + K_{t+1} + \Theta_t\left(\mu_t, E_t\right) &= \left(1 - D\left(Z_t\right)\right) A_{1,t} F\left(K_{1,t}, H_{1,t}, E_{1,t}\right) + (1-\delta) K_t, \quad \forall\, t \geq 0, \\ E_t &= A_{2,t} G\left(K_{2,t}, H_{2,t}\right), \quad \forall\, t \geq 0, \\ Z_t &= J\left(S_0, E_0^M, ..., E_t^M, \eta_0, ..., \eta_t\right), \quad \forall\, t \geq 0, \\ F_K(K_{1,t} H_{1,t}, E_{1,t}) G_H(K_{2,t} H_{2,t}) &= F_H(K_{1,t} H_{1,t}, E_{1,t}) G_K(K_{2,t} H_{2,t}), \quad \forall\, t \geq 0, \\ K_{1,t} + K_{2,t} &= K_t, \quad \forall\, t \geq 0, \\ H_{1,t} + H_{2,t} &= N_t h_t, \quad \forall\, t \geq 0, \\ N_t d_t + E_{1,t} &= E_t, \quad \forall\, t \geq 0, \\ \sum_j \pi_j \left(\varphi_j e_j^{\gamma(\sigma-1)}\right)^{\frac{1}{1-(1+\epsilon+\gamma)(1-\sigma)}} &= 1, \end{aligned}$$

where $N_t d_t + E_{1,t} = E_t$ is the only additional constraint compared to the benchmark problem.

D.4 Optimal taxes

Tax formulas From the first order conditions of the Ramsey problem, and using the same steps as in Appendix A, we can show that

$$\tau_{H,t} = 1 - \frac{U_{h,t}}{U_{c,t}} \frac{W_{c,t}}{W_{h,t}}, \tag{128}$$

$$\frac{R_{t+1}}{R^*_{t+1}} = \frac{W_{c,t+1}}{W_{c,t}} \frac{U_{c,t}}{U_{c,t+1}}, \tag{129}$$

$$\tau_{E,t} = \frac{1}{\nu_{1,t}} \sum_{j=0}^{\infty} \beta^j \left(\nu_{1,t+j} D'_{t+j} A_{1,t+j} F_{t+j} - N_{t+j} W_{Z,t+j} \right) J_{E_t^M, t+j}, \tag{130}$$

and

$$\tau_{I,t} = 0.$$

Using the first order conditions with respect to d_t, $E_{1,t}$ and c_t we have

$$W_{d,t} = W_{c,t}(1 - D(Z_t)) A_{1,t} F_{E,t},$$

which together with (126) and the final good firm's first order condition with respect to $E_{1,t}$ (given by (5) in the benchmark model) gives

$$\tau_{D,t} = \frac{U_{d,t}}{U_{c,t}} - \frac{W_{d,t}}{W_{c,t}}. \tag{131}$$

Proof of Proposition 4: The proof follows the same steps as the proof of Proposition 2. If we define

$$V\big(c_t, d_t, h_t, Z_t; \varphi, \theta, \lambda\big) \equiv \sum_i \pi_i \lambda_i u_i \big(c^m_{i,t}\big(c_t, d_t, h_t; \varphi\big), d^m_{i,t}\big(c_t, d_t, h_t; \varphi\big), h^m_{i,t}\big(c_t, d_t, h_t; \varphi\big), Z_t\big),$$

and

$$IC_i(c_t, d_t, h_t, \varphi) \equiv U_{c,t} c^m_{i,t}\big(c_t, d_t, h_t; \varphi\big) + U_{d,t} d^m_{i,t}\big(c_t, d_t, h_t; \varphi\big) + U_{h,t} e_{i,t} h^m_{i,t}\big(c_t, d_t, h_t; \varphi\big), \tag{132}$$

we can express the marginal cost of funds as

$$\text{MCF}_t = 1 + \frac{\text{cov}(\theta_i, IC_{c,i,t})}{V_{c,t}},$$

and we can re-write the optimal pollution tax given by (130) as

$$\tau_{E,t} = \sum_{j=0}^{\infty} \beta^j \left(\frac{V_{c,t+j} + \sum_i \pi_i \theta_i IC_{c,i,t+j}}{V_{c,t} + \sum_i \pi_i \theta_i IC_{c,i,t}} D'_{t+j} A_{1,t+j} F_{t+j} - \frac{N_{t+j} V_{Z,t+j}}{V_{c,t} + \sum_i \pi_i \theta_i IC_{c,i,t}} \right) J_{E_t^M, t+j}.$$

Using the definitions of $\tau^{Pigou}_{E,t}$, Δ_t, and ω^U_t stated in Section 3.3, and substituting for the MCF, we can write

$$\tau_{E,t} = \tau^{Pigou}_{E,t}\Big|_{SB} \left(\sum_{j=0}^{\infty} \frac{\text{MCF}_{t+j}}{\text{MCF}_t} \Delta_{t+j} (1 - \omega^U_t) + \frac{\omega^U_t}{\text{MCF}_t} \right).$$

With balanced-growth preferences, substituting into (132) we obtain

$$IC_{c,i,t} = (1 - \sigma) U_{c,t} \left((1 + \gamma + \epsilon) \omega_i - \gamma \frac{e_i}{1 - \varsigma h_t} + \epsilon \frac{\bar{d}_{i,t}}{d_t - \bar{d}_t} \right). \quad \blacksquare$$

Proof of Proposition 5: Using our functional form assumption, we can rewrite the pseudo-utility function as follows

$$W\left(c_t, d_t, h_t, Z_t; \varphi, \theta, \lambda\right) = \Phi \frac{(c_t (d_t - \bar{d}_t)^{\epsilon} (1 - \varsigma h_t)^{\gamma})^{1-\sigma}}{1 - \sigma} + \frac{(1 + \alpha_0 Z_t^2)^{-(1-\sigma)}}{1 - \sigma} + \Psi U_{h,t} + \Lambda_t U_{d,t} \ ,$$

where

$$
\begin{aligned}
\Phi &\equiv \sum_i \pi_i \omega_i \Big(\frac{\lambda_i}{\varphi_i} + \big(1-\sigma\big)\big(1+\epsilon+\gamma\big)\theta_i \Big), \\
\Psi &\equiv \frac{1}{\varsigma} \sum_i \pi_i \theta_i e_i, \\
\Lambda_t &\equiv \sum_i \pi_i \theta_i \bar{d}_{i,t}.
\end{aligned}
$$

We can use the first order conditions with respect to market weights to obtain

$$
\theta_i = \sum_j \frac{\pi_j \lambda_j}{\varphi_j} - \frac{\lambda_i}{\varphi_i},
$$

from which we can rewrite

$$
\begin{aligned}
\Phi &= \sum_i \pi_i \omega_i \Big(\frac{\lambda_i}{\varphi_i} + \big(1-\sigma\big)\big(1+\epsilon+\gamma\big) \Big(\sum_j \frac{\pi_j \lambda_j}{\varphi_j} - \frac{\lambda_i}{\varphi_i} \Big) \Big) \\
&= \sum_j \pi_j \frac{\lambda_j}{\varphi_j} + \Big(1 - (1+\epsilon+\gamma)(1-\sigma) \Big) \mathrm{cov}(\lambda_i/\varphi_i, \omega_i), \\
\Psi &= \frac{1}{\varsigma} \sum_i \pi_i \Big(\sum_j \frac{\pi_j \lambda_j}{\varphi_j} - \frac{\lambda_i}{\varphi_i} \Big) e_i \\
&= -\frac{\mathrm{cov}(\lambda_i/\varphi_i, e_i)}{\varsigma}, \\
\Lambda_t &= \sum_i \pi_i \Big(\sum_j \frac{\pi_j \lambda_j}{\varphi_j} - \frac{\lambda_i}{\varphi_i} \Big) \bar{d}_{i,t} \\
&= -\mathrm{cov}(\lambda_i/\varphi_i, \bar{d}_{i,t}).
\end{aligned}
$$

Substituting the derivatives of U into equation (131), we get

$$
\tau_{D,t} = \frac{\Lambda_t (d_t - \bar{d}_t)^{-1} U_{d,t}}{\Phi U_{c,t} + \Psi U_{hc,t} + \Lambda_t U_{dc,t}} = \frac{\Lambda_t \frac{\epsilon c_t}{(d_t - \bar{d}_t)^2}}{\Phi + \frac{\Psi \varsigma \gamma (\sigma - 1)}{(1 - \varsigma h_t)} - \frac{\Lambda_t \epsilon (\sigma - 1)}{(d_t - \bar{d}_t)}}. \quad \blacksquare
$$

Explicit income tax formulas We can additionally obtain expressions for the other tax rates. In particular, substituting the derivatives of U and W into equations (128) and (129), we have

$$
\tau_{H,t} = 1 - \frac{\Phi + \Psi \frac{U_{ch,t}}{U_{c,t}} + \Lambda_t \frac{U_{cd,t}}{U_{c,t}}}{\Phi + \Psi \frac{U_{hh,t}}{U_{h,t}} + \Lambda_t \frac{U_{dh,t}}{U_{h,t}}} = \frac{\Psi \varsigma (1 - \varsigma h_t)^{-1}}{\Phi + \Psi \frac{\varsigma \big(1 - \gamma(1-\sigma)\big)}{(1 - \varsigma h_t)} + \Lambda_t \frac{\epsilon(1-\sigma)}{(d_t - \bar{d}_t)}},
$$

$$
\frac{R_{t+1}}{R^*_{t+1}} = \frac{\Phi + \Lambda_{t+1} \frac{U_{cd,t+1}}{U_{c,t+1}} + \Psi \frac{U_{ch,t+1}}{U_{c,t+1}}}{\Phi + \Lambda_t \frac{U_{cd,t}}{U_{c,t}} + \Psi \frac{U_{ch,t}}{U_{c,t}}} = \frac{\Phi + \Lambda_{t+1} \frac{\epsilon(1-\sigma)}{(d_{t+1} - \bar{d}_{t+1})} - \Psi \frac{\varsigma \gamma (1-\sigma)}{(1 - \varsigma h_{t+1})}}{\Phi + \Lambda_t \frac{\epsilon(1-\sigma)}{(d_t - \bar{d}_t)} - \Psi \frac{\varsigma \gamma (1-\sigma)}{(1 - \varsigma h_t)}},
$$

and following the same steps as in Appendix A.4.2 we can obtain an expression for market weights

$$
\varphi_i = \frac{1}{e_i^{\gamma(\sigma-1)}} \left(1 + \frac{U_{c,0} R_0 N_0 (a_{i,0} - A_0) + \sum_t N_t \beta^t \Big(\frac{U_{h,t}}{\varsigma}(e_i - 1) - U_{d,t}(\bar{d}_{i,t} - \bar{d}_t) \Big)}{(1-\sigma)(1+\epsilon+\gamma) \sum_t N_t \beta^t \tilde{U}(c_t, d_t, h_t)} \right)^{1-(1+\epsilon+\gamma)(1-\sigma)}.
$$

Comparison with the benchmark formula The previous expression is the same as the one found in our benchmark, and the optimal tax will again be equal to the social cost of pollution when the marginal reduction in implementation cost ($\sum_i \pi_i \theta_i IC_{c,i,t}$) is null, which is the case in the first-best. Compared to our benchmark, the marginal implementation cost now includes an additional term from the derivative of U_d with respect to consumption. In particular, we again have

$$\sum_i \pi_i \theta_i IC_{c,i,t} = (1-\sigma)\frac{\text{cov}(\theta_i, IC_{i,t})}{c_t},$$

but now the ratio of the period implementation cost for two agents i and j is

$$\frac{IC_{i,t}}{IC_{j,t}} = \frac{(1+\epsilon+\gamma)\omega_i + \frac{\epsilon \bar{d}_{i,t}}{(d_t - \bar{d}_t)} - \frac{\gamma e_i}{(1-\varsigma h_t)}}{(1+\epsilon+\gamma)\omega_j + \frac{\epsilon \bar{d}_{j,t}}{(d_t - \bar{d}_t)} - \frac{\gamma e_j}{(1-\varsigma h_t)}}.$$

Thus, the sign of the marginal implementation cost depends on a price effect through σ, and on an energy demand and labor supply effects from $\text{cov}(\theta_i, IC_{i,t})$. The covariance term is higher in periods when richer households (higher θ_i) work relatively less, or when they have higher energy needs relative to poor households compared to an average period.

E Heterogeneous climate damages

Proof of Proposition 6: When utility is separable in consumption-leisure and environmental quality, the benchmark model presented in Appendix A can easily be generalized to the case where households experience heterogeneous climate damages on their utility. If we write agents' utility function as

$$u_i(c_{i,t}, h_{i,t}, Z_t) \equiv \tilde{u}(c_{i,t}, h_{i,t}) + \hat{u}_i(Z_t),$$

and apply the same steps as in Appendix A, we can again show that

$$\tau_{E,t} = \sum_{j=0}^{\infty} \beta^j \left(\frac{V_{c,t+j} + \sum_i \pi_i \theta_i IC_{c,i,t+j}}{V_{c,t} + \sum_i \pi_i \theta_i IC_{c,i,t}} D'_{t+j} A_{1,t+j} F_{t+j} - \frac{N_{t+j} V_{Z,t+j}}{V_{c,t} + \sum_i \pi_i \theta_i IC_{c,i,t}} \right) J_{E_t^M, t+j}.$$

The only difference with the benchmark model is the expression of $V_{Z,t}$, the marginal dis-utility from environmental degradation for the planner, which now writes

$$V_{Z,t} = \sum_i \pi_i \lambda_i \hat{u}'_i(Z_t).$$

Using the fact that $\sum_i \pi_i \lambda_i = 1$, we obtain

$$V_{Z,t} = \sum_i \pi_i \hat{u}'_i(Z_t) + \text{cov}\big(\lambda_i, \hat{u}'_i(Z_t)\big),$$

hence heterogeneity in marginal utility damages matters for the optimal pollution tax if and only if these marginal damages correlate with the planner's weight. ■

Corollary 1 is a straightforward application of Proposition 6.

F Calibration

F.1 Household heterogeneity

Productivity distribution We calibrate the ability distribution on the basis of hourly wage data that we obtain from the Survey of Consumer Finances (SCF). For each of the 6,015 households in the 2013 wave of the survey, we sum the hours worked on their main job and potential additional job(s) in a normal week. Annual labor supply of the respondent and their partner is then calculated by multiplying weekly hours worked by 52 minus the number of weeks they have spent unemployed during the past 12 months minus the number of weeks spent on holidays (which we assume to be equal to 3 for each worker). The household's hourly wage is then obtained as the household's annual income from wages and salaries before taxes, divided by the household's total annual labor supply (*i.e.*, the sum of the respondent and their partner's labor supply). This number reflects how much households members were paid on average for each hour of work they supplied in the past year.

To obtain the hourly wage distribution, we make a few additional adjustments. We first drop all households with an hourly wage below \$1 or above \$1,000. We also restrict the sample to households who have worked at least 1 week over the past 12 months, who work at least 1 hour on a normal week, and with no member working above 100 hours. Finally we restrict the sample to households whose respondent is at least 18 years old, and at most 65 years old. Using this sub-sample, we divide households in ten groups of hourly wage deciles. These correspond to $I = 10$ groups with size $\pi_i = 0.10$. For each group, we compute the average hourly wage.

Asset distribution For each of the ten productivity groups, we divide again households in ten deciles of net worth. For each sub-group, we compute the average net worth. This provides a table in which households are split in 100 groups of equal size, with for each of these groups the average hourly wage and the net worth.[32]

Because agents in our model are infinitely lived but hourly wage and asset holdings are positively correlated with age, we control for generational heterogeneity. To do so, we divide households in ten generations based on the age of the respondent, and compute the average hourly wage and net worth of each of the 100 groups within each generation. We then obtain the average hourly wage and net worth for each group as the average of that group over all generations. Table II below provides the results.

Distribution of energy consumption Our objective is to estimate households' subsistence level ($\bar{d}$) and relative preference for the dirty good (ϵ). From the households' first order conditions, we have

$$\frac{u_{d,h,t}}{u_{c,h,t}} = \frac{\epsilon c_{h,t}}{d_{h,t} - \bar{d}_{h,t}} = p_{E,t} + \tau_{D,t},$$

[32]On the sub-sample of households from whom we compute the productivity distribution, we find a correlation coefficient of 0.60 between income and wealth, a figure consistent with the 0.58 found by Kuhn and Ríos-Rull (2016) on the general population.

Table II: Distribution of Households Hourly Wages and Net Worth by Productivity Deciles (Rows) and Net Worth Deciles (Columns).

		Net worth deciles										
		1st	2nd	3rd	4th	5th	6th	7th	8th	9th	10th	Hourly wage
Productivity deciles	1st	-4.59e+04	-7.00e+03	1.22e+03	7.45e+03	1.79e+04	3.25e+04	6.44e+04	1.12e+05	2.18e+05	1.10e+06	6.44e+00
	2nd	-2.99e+04	-1.97e+03	4.89e+03	1.23e+04	2.50e+04	3.97e+04	6.46e+04	1.03e+05	1.83e+05	1.04e+06	1.11e+01
	3rd	-4.13e+04	-6.00e+03	3.72e+03	1.29e+04	2.76e+04	4.47e+04	7.69e+04	1.09e+05	2.01e+05	7.19e+05	1.42e+01
	4th	-4.56e+04	-2.65e+03	1.44e+04	3.31e+04	5.38e+04	7.48e+04	1.01e+05	1.50e+05	2.67e+05	7.64e+05	1.73e+01
	5th	-4.94e+04	-2.15e+03	1.55e+04	3.58e+04	6.72e+04	9.53e+04	1.40e+05	2.07e+05	2.98e+05	1.10e+06	2.05e+01
	6th	-3.82e+04	1.21e+04	3.94e+04	7.26e+04	1.14e+05	1.60e+05	2.13e+05	2.88e+05	4.60e+05	1.75e+06	2.41e+01
	7th	-2.41e+04	3.79e+04	6.75e+04	1.03e+05	1.54e+05	2.06e+05	2.63e+05	3.58e+05	5.32e+05	1.23e+06	2.86e+01
	8th	-2.93e+04	3.00e+04	7.10e+04	1.34e+05	2.11e+05	2.80e+05	3.90e+05	5.04e+05	6.94e+05	2.57e+06	3.48e+01
	9th	4.38e+03	6.86e+04	1.44e+05	2.11e+05	3.07e+05	4.20e+05	5.53e+05	7.45e+05	1.08e+06	3.50e+06	4.47e+01
	10th	-8.53e+04	1.40e+05	2.77e+05	4.43e+05	6.38e+05	8.55e+05	1.29e+06	2.14e+06	3.45e+06	1.00e+07	1.01e+02

Note: The rows correspond to productivity (*i.e.* hourly wage) decile groups. The last column corresponds to the average hourly wage in dollars for each productivity group. Columns 1 to 10 correspond to net worth decile groups within productivity groups. The number reported in these columns are the average net worth for each group in dollars. All groups are defined for a given generation, and values correspond to the weighted average across ten generation groups. Example: 1.10e+06 in the 1st row, 10th column, means that among the people that belong to the bottom 10% of the hourly wage distribution of their generation, the 10% wealthiest have an average net worth of $1.10e+06.

with $u_{d,h,t}$, $u_{c,h,t}$ the marginal utility of energy and final good consumption of household h at time t. Rearranging the previous equation, we obtain for each household h, and for each period t,

$$d_{h,t}(p_{E,t} + \tau_{D,t}) = \bar{d}_{h,t}(p_{E,t} + \tau_{D,t}) + \epsilon c_{h,t}. \tag{133}$$

We quantitatively investigate two scenarios: one where all households share the same subsistence level $\bar{d}$, and one where different groups share different subsistence levels. Under the assumption that households all face the same subsistence level $\bar{d}$, we can write the following regression from equation (133),

$$d_h(p_E + \tau_D) = \beta_d + \beta_\epsilon c_h + \mu_h, \tag{134}$$

where μ_h is the error term and β_d and β_ϵ are the empirical counterparts to $\bar{d}(p_E + \tau_D)$ and ϵ in the model. These parameters are estimated based on the cross-sectional distribution of energy and non-energy expenditures ($d_h(p_E + \tau_D)$ and c_h) observed in the Consumer Expenditure Surveys (CEX). We estimate regression (134) using OLS and abstracting from endogeneity issues as our aim is simply to inform our structural model so that it is consistent with the observed distribution of energy expenditure shares across groups.

In order to quantify the importance of heterogeneity in subsistence levels, we then use the estimated value of ϵ to compute—from equation (133)—household-specific values for $\bar{d}_h(p_E + \tau_D)$ that we regress against a set of binary variables denoting the subsistence type of different households,

$$\bar{d}_h(p_E + \tau_D) = \sum_{j \in J} \beta_j \mathcal{I}_h^j + \eta_h, \tag{135}$$

Table III: Distribution of Households Energy Expenditure Shares by Productivity Quintiles (Rows) and Expenditure Share Terciles (Columns).

		Expenditure share terciles		
		1st	2nd	3rd
Productivity quintiles	1st	6.39%	10.80%	15.59%
	2nd	6.47%	10.59%	15.21%
	3th	6.08%	9.85%	14.59%
	4th	5.65%	9.00%	13.73%
	5th	5.10%	8.03%	12.86%

Note: The rows correspond to productivity (*i.e.* hourly wage) quintile groups. The columns correspond to energy expenditure share tercile groups within productivity quintile groups. The numbers reported in these columns are the average energy expenditure shares for each group. All groups are defined for a given month and year, and values correspond to the average across all periods. Example: 6.39% in the 1st row, 1st column, means that among the people that belong to the bottom 20% of the hourly wage distribution at the month $\times$ year they were interviewed, the 33.3% with lowest energy shares spend on average 6.39% of their total expenditures in energy. Sample: CEX from 2011 to 2015, only workers included, outliers excluded.

where η_h is the error term, and $\{\mathcal{I}_h^j\}_{j\in J}$ is a set of subsistence-type binary variables defined as

$$\mathcal{I}_h^j = \begin{cases} 1, & \text{if } h \in j, \\ 0, & \text{otherwise.} \end{cases}$$

To be consistent with the timing of DICE, we pool surveys from the 20 quarters between January 2011 and December 2015, for a total of 129,573 observations. Energy expenditures ($d_h(p_E + \tau_D)$) are obtained by summing expenditures on gasoline and motor oil, electricity, natural gas, fuel oil, and other fuels. Non-energy expenditures (c_h) are obtained by subtracting energy expenditures from total expenditures. In order to characterize the joint distribution of productivity and necessity types, we compute the hourly wage following the same procedure as with the SCF. We first restrict our sample to working households. We again compute the household annual wage by summing the income received from salary or wages before taxes. We then compute the annual labor supply of the respondent and its partner: we multiply the number of hours usually worked per week by the number of weeks worked in the past twelve months, minus 3 weeks of imputed holidays. The household hourly wage is then the ratio of the household's annual wage over annual hours. Just like with the SCF data, this number reflects how much households members were paid on average for each hour of work they supplied in the past year.[33]

To avoid extreme values potentially driven by consumers' misreporting of their expenditures, we eliminate outliers that we define as the households whose energy expenditure shares are in the top or bottom 10% of the distribution. Using this sub-sample, we divide households in five groups of hourly

[33]The bottom hourly wage is \$6.59 and the top hourly wage is \$110.12 (without generational adjustments).

Table IV: Estimated Parameters for Energy Preferences with Homogeneous Necessity.

	Dependent variable: energy consumption (d)
β_ϵ	0.0529
	(0.000)
β_d	592.48
	(3.78)
Observations	67,520
adjusted-R^2	0.405

Note: The numbers give the estimated values of the parameters. Standard errors are reported in parentheses. β_ϵ corresponds to the empirical counterpart of ϵ in the model. β_d represents the empirical counterpart of the initial $\bar{d}$ in the model. Sample: CEX from 2011 to 2015, only workers included, outliers excluded.

wage quintiles. For each of the five groups, we divide again households in three terciles of energy expenditure shares, and compute the average energy expenditure share for each group. This provides a table in which households are split into 15 groups of equal size.[34] Since energy consumption shares do not appear to be strongly determined by age among working households, we do not control for generational differences. However, we control for seasonality and yearly variations that could lead to overestimate consumption heterogeneity. We proceed in the same way as with generational controls: we group households based on their ranking relative to the households interviewed in the same month and same year. The resulting distribution of initial energy shares by subsistence type j, $\{X_j\}_{j \in J}$, is presented in Table III, and the outputs of regressions (134) and (135) are given in Tables IV and V, respectively.

The values of β_j reported in Table V provide the initial distribution of $\bar{d}_j(p_E + \tau_D)$. These estimates are in dollars, and need to be normalized in order to target an average expenditure share of 10.8% in the model, as observed in the CEX. Relative to our baseline, we divide each of our ten productivity groups in three necessity types, and impute to each of the 30 groups the value of $\bar{d}_j$ corresponding to its productivity quintile (two deciles pooled together) and necessity tercile. Finally, we set $\bar{d}_{j,t}$ to grow over time following the same trajectory as the other aggregate variables on the balanced-growth path.

F.2 Parameters choice

Baseline hours worked We also use the SCF 2013 to compute the initial labor supply that we impute to the model. To do so, we again restrict the sample to individuals between 18 and 65 years old. However, because our aim is not to compute hourly wages but to look at the average labor supply, we do not eliminate outliers based on their hourly wage or labor supply. In particular, we keep unemployed households for whom the hourly wage is not observed, as dropping them would lead to overestimate

[34] We divide households in only 15 necessity groups to mitigate the potential over-estimation of consumption heterogeneity due to measurement error at the household level in the CEX, and to avoid negative values for the necessity levels.

Table V: Estimated Parameters for Type-Specific Subsistence Levels.

	Dependent variable: energy consumption (d)
$\beta_{1,1}$	128.7
	(5.6)
$\beta_{2,1}$	170.5
	(5.6)
$\beta_{3,1}$	148.1
	(5.5)
$\beta_{4,1}$	111.0
	(5.5)
$\beta_{5,1}$	0.2
	(5.4)
$\beta_{1,2}$	497.3
	(5.6)
$\beta_{2,2}$	599.1
	(5.6)
$\beta_{3,2}$	659.1
	(5.6)
$\beta_{4,2}$	651.8
	(5.5)
$\beta_{5,2}$	617.1
	(5.4)
$\beta_{1,3}$	811.9
	(5.6)
$\beta_{2,3}$	1001.3
	(5.6)
$\beta_{3,3}$	1120.7
	(5.6)
$\beta_{4,3}$	1174.5
	(5.5)
$\beta_{5,3}$	1229.4
	(5.5)
Observations	67,520
adjusted-R^2	0.542

Note: The numbers give the estimated values of the parameters. Standard errors are reported in parentheses. Each parameter $\beta_{a,b}$ represents the empirical counterpart of the initial $\bar{d}_h$ for an agent h belonging to the a^{th} productivity quintile, and the b^{th} energy-share tercile within this productivity quintile. Sample: CEX from 2011 to 2015, only workers included, outliers excluded.

the average labor supply. For all households in the sample, we divide the annual labor supply by the number of working age individuals (individuals between 18 and 65). This yields an average of 1440 hours annually. Assuming a maximum labor supply capacity of 52 weeks per year and 100 hours per week per individual, this yields an average labor supply of 0.277 of the maximum capacity.

Table VI: Calibration Summary: Climate Parameters (from DICE 2016).

Parameter	Description	Value
Carbon stocks		
S_{2015}^{At}	Initial carbon concentration in atmosphere (in GtC)	851
S_{2015}^{Up}	Initial carbon concentration in upper strata (in GtC)	460
S_{2015}^{Lo}	Initial carbon concentration in lower strata (in GtC)	1740
S_{eq}^{At}	Equilibrium carbon concentration in atmosphere (in GtC)	588
E_{2015}^{land}	Initial CO_2 emissions from land ($GtCO_2$ per year)	2.6
$g_{E^{\text{land}}}$	Decline rate of land emissions (per period)	0.115
Carbon cycle transition matrix		
$b_{1,1}$	Carbon cycle coefficient	0.88
$b_{2,1}$	Carbon cycle coefficient	0.047
$b_{3,1}$	Carbon cycle coefficient	0
$b_{1,2}$	Carbon cycle coefficient	0.12
$b_{2,2}$	Carbon cycle coefficient	0.94796
$b_{3,2}$	Carbon cycle coefficient	0.00075
$b_{1,3}$	Carbon cycle coefficient	0
$b_{2,3}$	Carbon cycle coefficient	0.005
$b_{3,3}$	Carbon cycle coefficient	0.99925
Radiative forcing		
κ	Forcings of equilibrium CO_2 doubling (Wm-2)	3.6813
$\mathcal{F}_{2015}^{\text{Ex}}$	Initial forcings of non-CO2 GHG (Wm-2)	0.5
$\mathcal{F}_{2100}^{\text{Ex}}$	2100 forcings of non-CO2 GHG (Wm-2)	1
$g_{\mathcal{F}^{\text{Ex}}}$	Rate of convergence of $\mathcal{F}$	1/17
Temperature		
T_{2015}	Initial atmospheric temperature change (C since 1900)	0.85
T_{2015}^{Lo}	Initial lower stratum temperature change (C since 1900)	0.0068
ζ_1	Climate model coefficient	0.1005
ζ_2	Climate model coefficient	1.1875
ζ_3	Climate model coefficient	0.088
ζ_4	Climate model coefficient	0.025

Table VII: Calibration Summary: Economic Parameters in the Baseline.

Parameter	Description	Value	Source
<u>Preferences</u>			
β	Utility discount rate (per year)	1/(1.015)	DICE 2016
σ	Inverse of IES	1.45	DICE 2016
η^F	Frisch elasticity of labor supply	0.75	Chetty et al (2011)
ς	Labor dis-utility coefficient	1.875	To target η^F and h_{2015}
γ	Labor dis-utility exponent	0.753	To target η^F and h_{2015}
α_0	Relative preference for the environment	7.88e-05	Adapted from Barrage (2019)
<u>Production damages</u>			
a_1	Damage intercept	0	DICE 2016
a_2	Damage coefficient quadratic term	0.00175	DICE 2016 adjusted
a_3	Damage exponent	2	DICE 2016
<u>Production first sector</u>			
α	Return to scale on labor sector 1	0.3	DICE 2016
ν	Return to scale on energy sector 1	0.04	Golosov et al (2014)
δ	Depreciation rate on capital (per year)	0.1	DICE 2016
$r_{2015} - \delta$	Initial net rate of return on capital	0.032	At steady state
Y_{2015}	Initial output (in trillions 2015 USD)	70.807	World Bank (2011-2015)
$hh_{1,2015}$	Initial share of labor in sector 1	0.977	To equate MPL across sectors
$kk_{1,2015}$	Initial share of capital in sector 2	0.928	To equate MPL across sectors
E_{2015}	Initial industrial emissions ($GtCO_2$ per year)	35.85	DICE 2016
h_{2015}	Initial labor supply per capita	0.277	Computed from SCF
$A_{1,2015}$	Initial TFP sector 1	141.9	To target Y_{2015}
<u>Production second sector</u>			
α_E	Return to scale on capital sector 2	0.403	Barrage (2019)
$A_{2,2015}$	Initial TFP sector 2	86.9	To target E_{2015}
<u>Abatement costs</u>			
$P_{2015}^{\text{backstop}}$	Backstop price in 2015 (in \$/tCO_2)	550	DICE 2016
$g_{P^{\text{backstop}}}$	Decline rate backstop price (per period)	2.5%	DICE 2016
c_2	Exponent abatement cost function	2.6	DICE 2016
μ_{2015}	Initial abatement share	0.03	DICE 2016
<u>Government</u>			
G_t/Y_t	Government spending to GDP ratio	0.3030	IMF-GFS
B_{2015}	Initial public debt to GDP ratio	0.2220	IMF-GFS
$\tau_{H,2015}$	Initial tax rate on labor income	0.255	Trabandt & Uhlig (2012)
$\tau_{K,2015}$	Initial tax rate on capital income	0.411	Trabandt & Uhlig (2012)

Calibration Summary: Economic Parameters in the Baseline (continued).

Exogenous growth parameters			
$g_{A_{1,2015}}$	Initial TFP growth rate sector 1 (per period)	0.076	DICE 2016
$gg_{A_{1,t}}$	Decline rate TFP growth sector 1 (per year)	0.005	DICE 2016
$g_{A_{2,2015}}$	Initial TFP growth rate sector 2 (per period)	0.076	DICE 2016
$gg_{A_{2,t}}$	Decline rate TFP growth sector 2 (per year)	0.005	DICE 2016
N_{2015}	Initial population (in millions)	1,309	World bank (2015)
$N_{\max}$	Asymptotic population (in millions)	2,034	DICE 2016 US-adjusted
g_N	Rate of convergence of population	0.134	DICE 2016

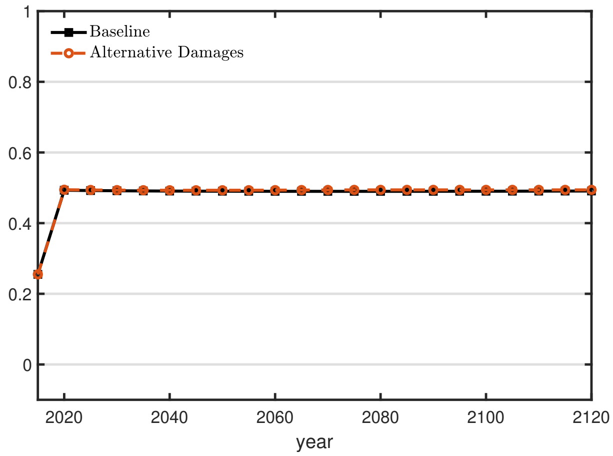
Table VIII: Calibration Summary: Economic Parameters with Stone-Geary Preferences.

Parameter	Description	Value
ϵ	Energy consumption utility exponent	0.053
$\bar{d}$	Initial average energy subsistence ($GtCO_2$ per year)	6.05
ν	Return to scale on energy sector 1	0.169
α	Return to scale on labor sector 1	0.259
ς	Labor dis-utility coefficient	1.881
γ	Labor dis-utility exponent	0.728
$A_{2,2015}$	Initial TFP sector 2	20.4
α_0	Relative preference for the environment	7.92e−05

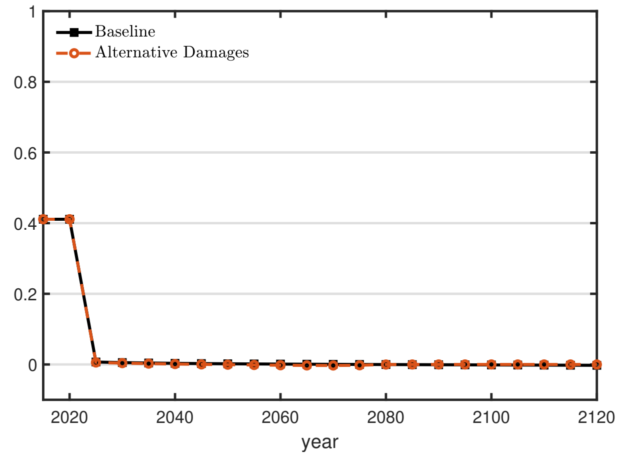
Notes: The table reports the values of the parameters used in the calibration of the extended version of the model with two goods and Stone-Geary utility (see Section 6.3). The parameters are selected to obtain energy expenditure shares consistent with the CEX and a share of aggregate emissions coming from households' energy consumption consistent with EPA's estimates.

G Additional quantitative results

G.1 Alternative damages



(a) Optimal labor income taxes



(b) Optimal capital income taxes

Figure 9: Optimal Income Taxes, Alternative Damages.

Notes: Figures show the path of second-best labor and capital income taxes for the baseline calibration (black) and for the alternative-damages calibration (red). Initial tax rates (for 2015) are set exogenously to their current levels obtained from Trabandt and Uhlig (2012).

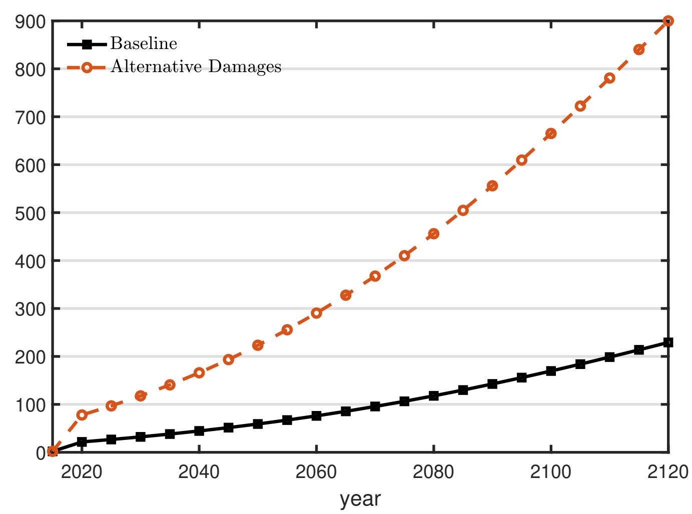


Figure 10: Optimal Carbon Taxes ($\$/tCO_2$), Alternative Damages.

Notes: Figure shows the path of second-best carbon taxes for the baseline calibration (black) and for the alternative-damages calibration (red), expressed in dollars per ton of CO_2. Initial level (for 2015) is set exogenously to its current level obtained from Nordhaus (2017).

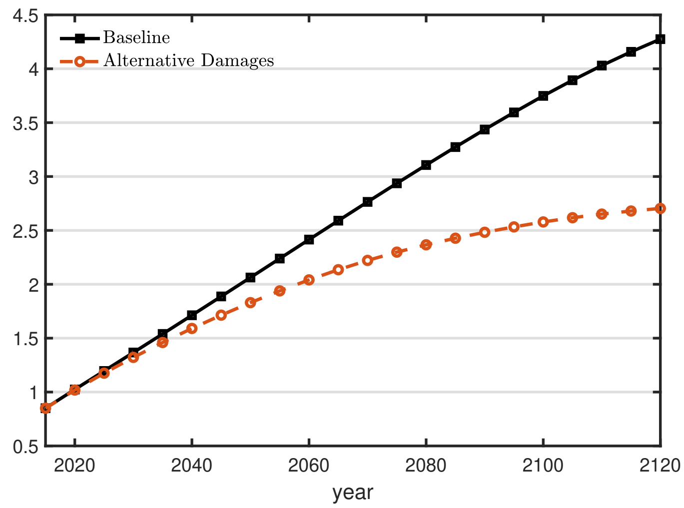


Figure 11: Temperature on the Optimal Path, Alternative Damages.

Notes: Figure shows the path of atmospheric temperature (Z_t^{At}) for the baseline calibration (black) and for the alternative-damages calibration (red), expressed in degree Celsius.

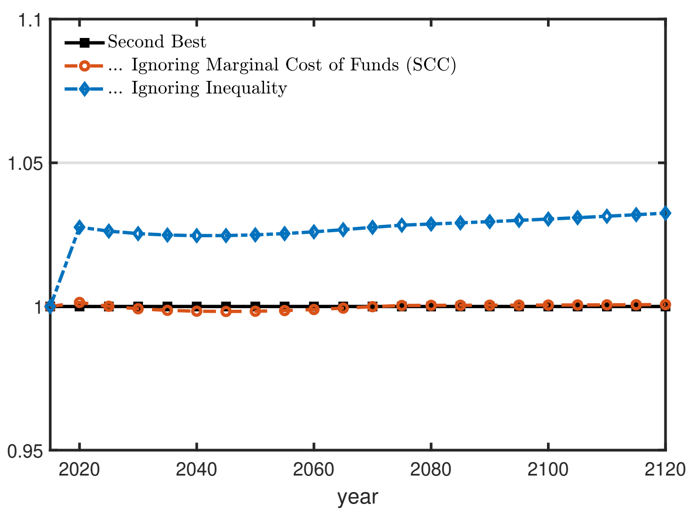
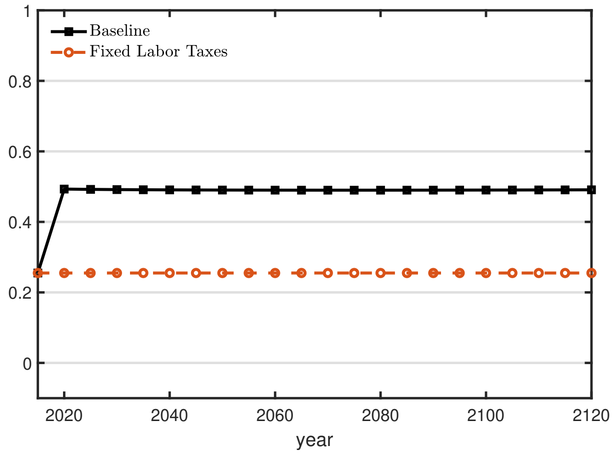


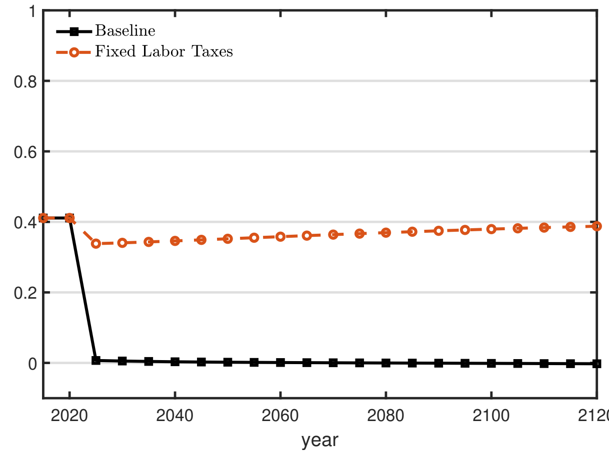
Figure 12: Carbon Tax Decomposition, Alternative Damages.

Notes: The black line represents the second-best carbon tax normalized to 1. The red line shows what this tax would be if the MCF was set to 1 in all periods, holding aggregates constant (see Proposition 2). The blue line shows what this tax would be absent consumption inequalities, again holding aggregates constant (see Proposition 3). All taxes are computed under the alternative-damages calibration.

G.2 Third-best policies



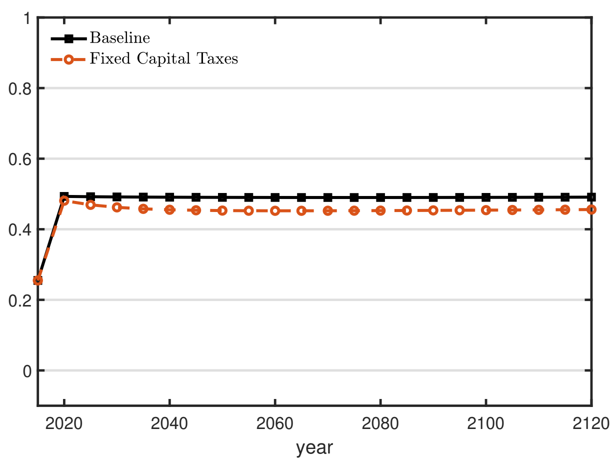
(a) Optimal labor income taxes



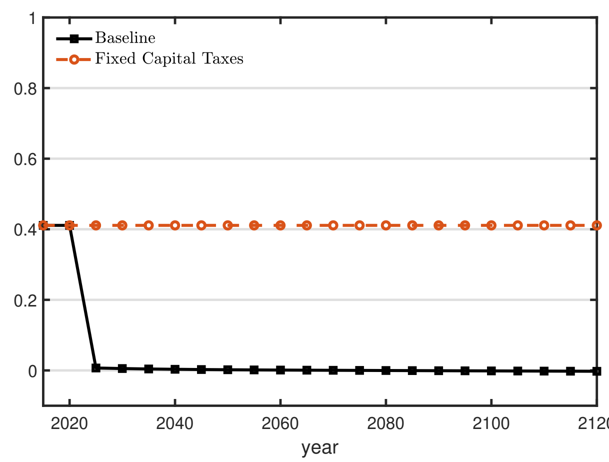
(b) Optimal capital income taxes

Figure 13: Optimal Income Taxes, Given Labor Tax.

Notes: Figures show the path of second-best labor and capital income taxes for the baseline calibration (black) and for the economy with given labor income taxes (red). Initial tax rates (for 2015) are set exogenously to their current levels obtained from Trabandt and Uhlig (2012).



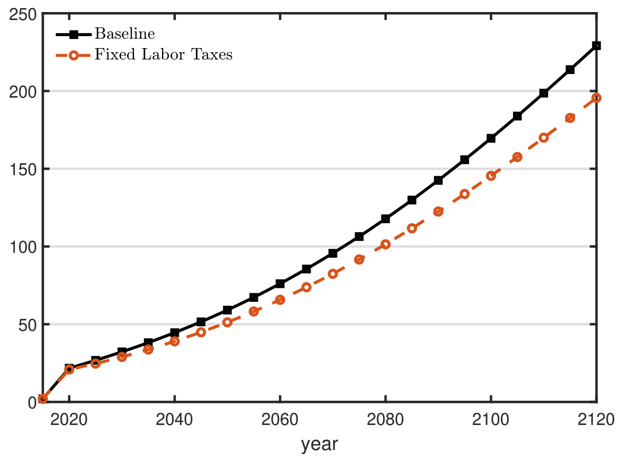
(a) Optimal labor income taxes



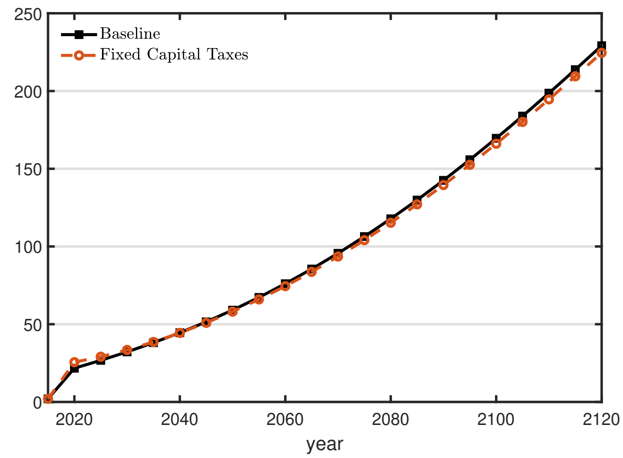
(b) Optimal capital income taxes

Figure 14: Optimal Income Taxes, Given Capital Tax.

Notes: Figures show the path of second-best labor and capital income taxes for the baseline calibration (black) and for the economy with given capital income taxes (red). Initial tax rates (for 2015) are set exogenously to their current levels obtained from Trabandt and Uhlig (2012).



(a) Given labor income tax



(b) Given capital income tax

Figure 15: Optimal Carbon Taxes ($/t$CO_2$), Given Income Taxes.

Notes: Figure shows the path of second-best carbon taxes for the baseline calibration (black) and for the economies with fixed labor and capital income taxes (red), expressed in dollars per ton of CO_2. Initial level (for 2015) is set exogenously to its current level obtained from Nordhaus (2017). Differences with the baseline are due to the change in tax formulas, as well as differences in individual and aggregate allocations.

Table IX: Government Budget Adjustment, Given Labor Income Taxes.

	Revenue Source			Revenue Use		
	Labor	Capital	Carbon	Gov. Cons.	Transfer	Interest
No Carbon Tax	17.2%	5.5%	0.0%	15.9%	5.1%	2.0%
Optimal Carbon Tax	17.0%	5.3%	1.0%	15.7%	6.1%	2.0%
Change	−0.2%	−0.2%	1.0%	−0.2%	1.0%	0.0%

Notes: For given labor income taxes, the numbers represent the present value of each component of the government budget constraint divided by the present value of GDP, in the scenarios without carbon taxes (first row) and with carbon taxes (second row). The third row displays the difference between the two scenarios.

Table X: Government Budget Adjustment, Given Capital Income Taxes.

	Revenue Source			Revenue Use		
	Labor	Capital	Carbon	Gov. Cons.	Transfer	Interest
No Carbon Tax	31.5%	6.2%	0.0%	18.5%	16.6%	2.0%
Optimal Carbon Tax	30.7%	6.0%	1.3%	18.2%	17.4%	2.1%
Change	−0.8%	−0.1%	1.3%	−0.3%	0.8%	0.0%

Notes: For given capital income taxes, the numbers represent the present value of each component of the government budget constraint divided by the present value of GDP, in the scenarios without carbon taxes (first row) and with carbon taxes (second row). The third row displays the difference between the two scenarios.

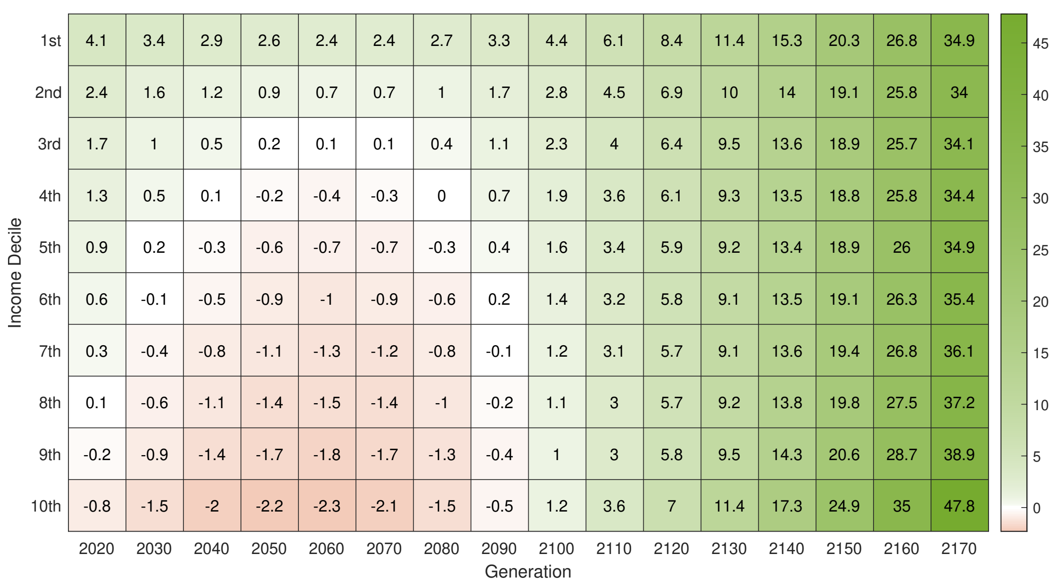


Figure 16: Period Welfare Gains (%), Given Labor Income Taxes.

Notes: For each decade and each income decile the table shows the welfare gains, in percentage of consumption, from optimal carbon taxation relative to a scenario without carbon taxation. Numbers are computed under the baseline calibration with given labor income taxes.

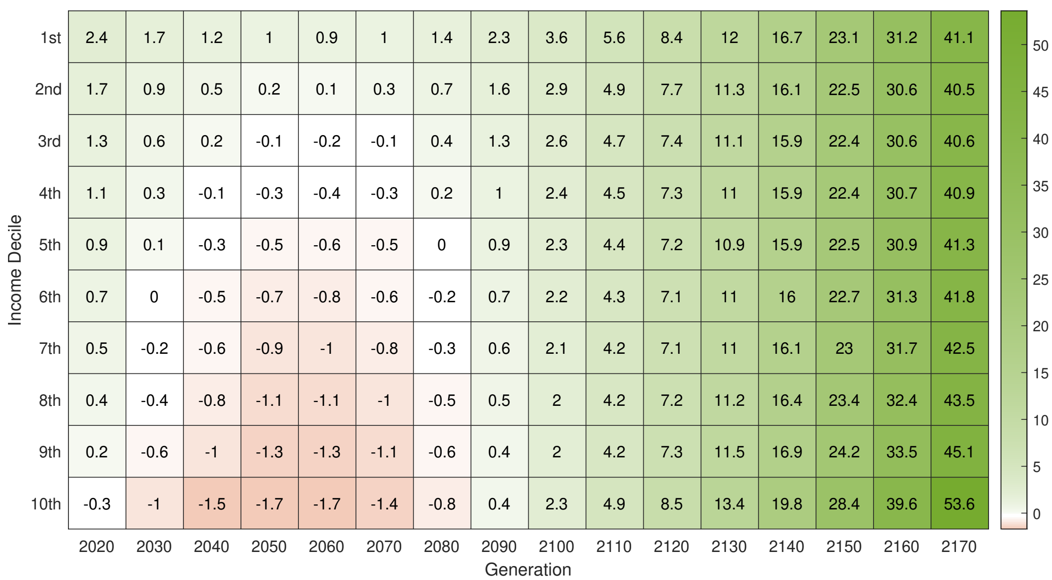
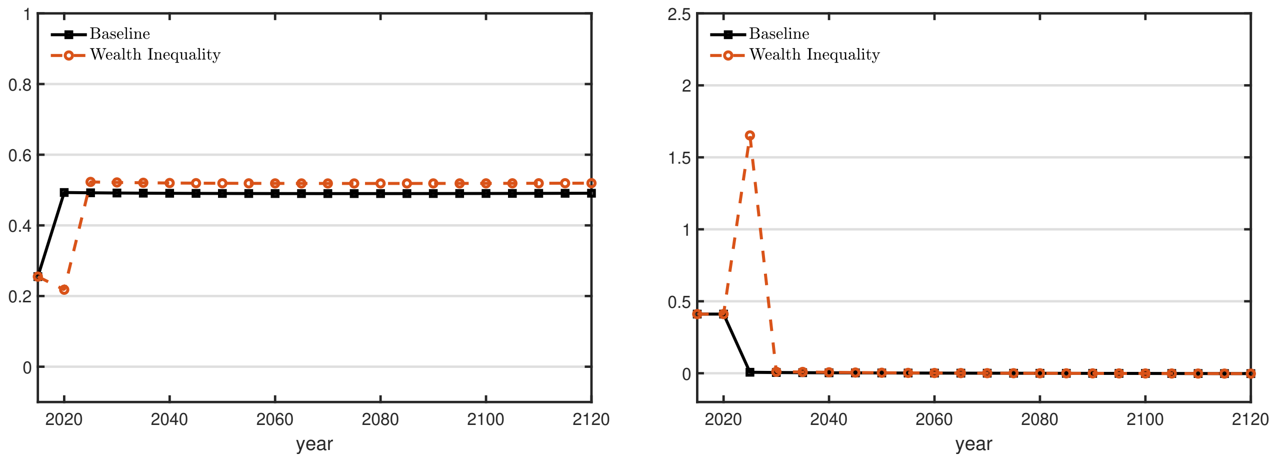


Figure 17: Period Welfare Gains (%), Given Capital Income Taxes.

Notes: For each decade and each income decile the table shows the welfare gains, in percentage of consumption, from optimal carbon taxation relative to a scenario without carbon taxation. Numbers are computed under the baseline calibration with given capital income taxes.

G.3 Initial wealth inequality



(a) Optimal labor income taxes

(b) Optimal capital income taxes

Figure 18: Optimal Income Taxes, Initial Wealth Heterogeneity and Exogenous Initial Capital Tax.

Notes: Figures show the path of second-best labor and capital income taxes for the baseline calibration (black) and for the economy with initial wealth inequality (red). Initial tax rates (for 2015) are set exogenously to their current levels obtained from Trabandt and Uhlig (2012).

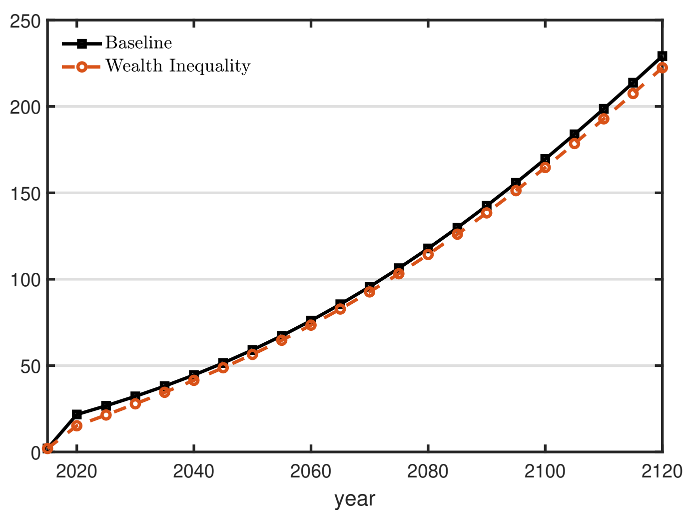


Figure 19: Optimal Carbon Taxes ($/t$CO_2$), Initial Wealth Heterogeneity and Exogenous Initial Capital Tax.

Notes: Figure shows the path of second-best carbon taxes for the baseline calibration (black) and for the economy with initial wealth inequality (red), expressed in dollars per ton of CO_2. Initial level (for 2015) is set exogenously to its current level obtained from Nordhaus (2017). Differences with the baseline are due to the change in tax formulas, as well as differences in individual and aggregate allocations.

Table XI: Government Budget Adjustment, Initial Wealth Heterogeneity.

	Revenue Source			Revenue Use		
	Labor	Capital	Carbon	Gov. Cons.	Transfer	Interest
No Carbon Tax	34.2%	3.2%	0.0%	16.4%	18.0%	1.5%
Optimal Carbon Tax	33.5%	3.2%	1.1%	16.1%	18.8%	1.5%
Change	−0.7%	0.0%	1.1%	−0.3%	0.8%	0.0%

Notes: For the economy with initial wealth inequality and fixed initial capital income tax, the numbers represent the present value of each component of the government budget constraint divided by the present value of GDP, in the scenarios without carbon taxes (first row) and with carbon taxes (second row). The third row displays the difference between the two scenarios.

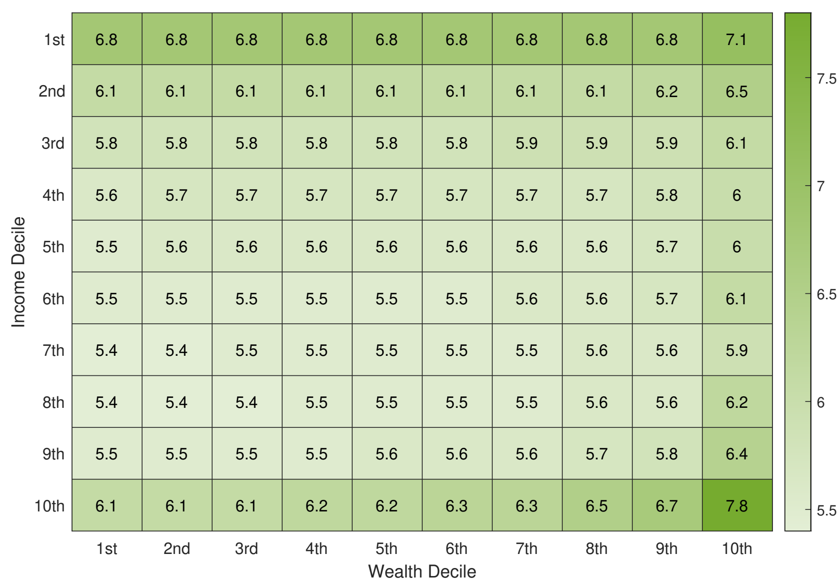
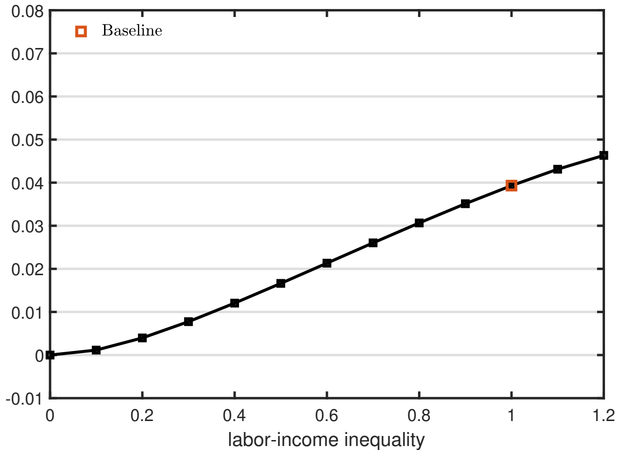


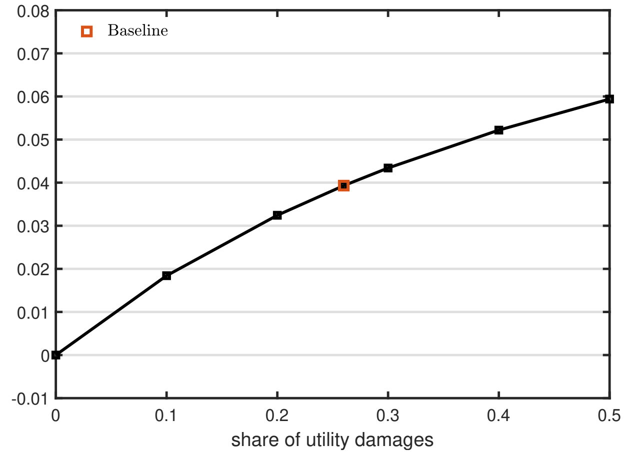
Figure 20: Welfare Gains (%), Initial Wealth Heterogeneity and Exogenous Initial Capital Tax.

Notes: For each income and wealth decile the table shows the discounted welfare gains, in percentage of consumption, from optimal carbon taxation relative to a scenario without carbon taxation. Numbers are computed under the calibration with wealth inequality.

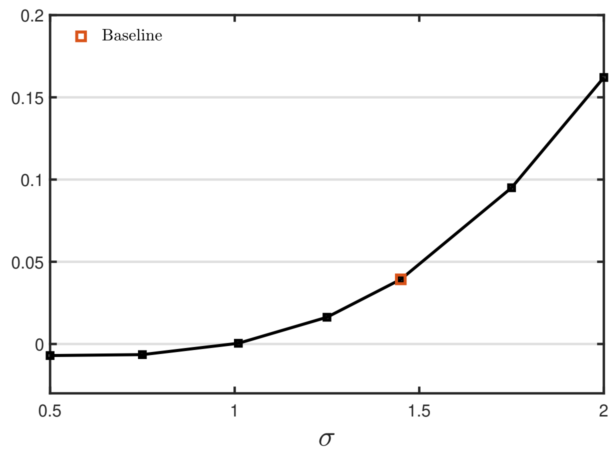
G.4 Sensitivity of inequality effects to calibration choices



(a) Effect of Labor Income Inequality



(b) Effect of Share of Utility Damages



(c) Effect of σ

Figure 21: How Inequality Effects Change with Different Levels of Labor Income Inequality, Different Shares of Utility Damages in Total Damages, and Different σ's

Notes: The y-axis of the three figures represents the average percentage increase in optimal carbon taxes over the next 100 years that would result from ignoring labor income inequality. (a) To obtain different levels of labor income inequality we take a convex combination between the vector of productivities from the baseline economy and a vector with equal productivities. In the x-axis we have the weight put on the baseline vector. A weight of zero implies no labor income inequality, and a weight of one implies the baseline level of inequality. (b) In the baseline calibration, we choose α_0 so that 26% of total damages are utility damages. The x-axis represents different targets for the share of utility damages. (c) In the baseline calibration, we set σ equal to 1.45, following DICE 2016. For each alternative σ, we recalibrate γ, ς, and $A_{2,2015}$ to match the targets described in Table VII.

H Algorithm to compute Ramsey policies

To solve the Ramsey problem numerically we apply an algorithm that directly uses the first-order conditions obtained above. Here, we explain the procedure we used to obtain the benchmark results. The idea behind the algorithm is simple. Given a policy (a sequence of taxes and transfers), standard methods can be used to compute the associated equilibrium aggregate. Given equilibrium aggregates, we can use the optimality conditions derived from the Ramsey problem to update the policy. We then iterate on these two steps until convergence. The steps below explain the algorithm in more detail:

1. *Guess a policy:* $\{\tau_{H,t}, \tau_{K,t}, \tau_{I,t}, \tau_{E,t}\}_{t=0}^{\infty}$ and T.

2. *Compute the associated equilibrium aggregate allocation and prices:* $\{c_t, h_t, K_{1,t}, K_{2,t}, H_{1,t}, H_{2,t}, E_t, \mu_t, Z_t, r_t, w_t, p_{E,t}, R_t\}_{t=0}^{\infty}$. We use a shooting algorithm but different standard methods could be used, so we will not elaborate further on this part.

3. *Compute terms that appear in the optimality conditions of the Ramsey planner:* Compute M using equation (70), then obtain ω_i and φ_i, for all i, from equations (71) and (72)—equations (60) and (61) can be used to obtain individual allocations and welfare. Next, obtain Φ and Ψ using equations (68) and (69). Equation (64) then gives $W_{c,t}$, $W_{h,t}$, and $W_{Z,t}$, for all t.

4. *Update policy:* Use equations (54), (55), (58), and (65) to update $\{\tau_{H,t}, \tau_{K,t}, \tau_{I,t}, \tau_{E,t}\}_{t=0}^{\infty}$. Use the government budget constraint to update T.

5. *Iterate:* If the updated policy differs from the initial guess, return to step 2.