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

Thomas Douenne † Albert Jan Hummel † Marcelo Pedroni †

November 13, 2025

Abstract

We study optimal fiscal policy to address climate change and inequality. We theoretically characterise optimal carbon and income taxes and quantify them for the US economy with a climate model calibrated to DICE. In contrast to the representative-agent setting, we find that (i) tax distortions have a negligible effect on the optimal carbon tax; (ii) inequality only slightly reduces it; (iii) the revenue from carbon taxes is optimally split about equally between reducing tax distortions and increasing transfers. Unlike the double-dividend policy, optimal carbon taxation has progressive welfare effects and low-income households benefit even in the short run.

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 Adrien Fabre, Reyer Gerlagh, Bas Jacobs, Mehrab Kiarsi, Rick van der Ploeg, Christian Stoltenberg, Sweder van Wijnbergen, and seminar participants at EAERE, EEA, IIPF, NED, PET, SED, SMS, and SURED conferences, and at Einaudi, FGV-EPGE, Georgetown University in Qatar, IDB, Paris School of Economics, Stockholm School of Economics, Toulouse School of Economics, TU Berlin, University of Alicante, University of Amsterdam, University of Bologna, University of Bristol, University of Nantes, University of Oxford, University of Potsdam, University of Strasbourg, and University of Surrey for helpful comments and discussions.

†University of Amsterdam, Amsterdam School of Economics. Address: University of Amsterdam, Roetersstraat 11, 1018 WB Amsterdam, Netherlands. Correspondence: t.r.g.r.douenne@uva.nl, a.j.hummel@uva.nl, m.pedroni@uva.nl.

1 Introduction

Economic inequality and environmental degradation are certainly two of the most critical issues facing societies today. Economists have long argued for the use of fiscal instruments to address these problems. Labour and capital taxes can provide redistribution, and, following the Pigouvian principle, a pollution tax can internalise environmental externalities. However, pollution taxes also have distributional implications as they heterogeneously impact households' purchasing power. At the same time, capital and labour taxes affect the costs and benefits of improving the environment by reducing incentives to work and invest. This study aims to analyse how these instruments should be jointly optimised to tackle both inequality and environmental degradation. Specifically, we ask: Do inequalities and redistributive taxation call for more or less ambitious environmental policies? How do environmental policies affect inequalities?

We address these questions 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 (2020) 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, which leads to environmental degradation that affects both economic productivity and households' utility. As in Barrage (2020), energy producers can reduce the emission intensity of their output by engaging in costly abatement activities. Because of these costs, positive abatement occurs 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 maximises social welfare by choosing the level of linear taxes—in particular, taxes on labour and capital income, energy, and pollution—and a uniform lump-sum transfer to adjust the progressivity of the tax system. Because agents are heterogeneous but tax instruments are anonymous, the government must rely on distortionary instruments to provide redistribution.[1] We derive 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 energy demand) to match US data. Conceptually, our quantitative analysis examines the optimal fiscal policy of the US if they accounted for the negative global impact of their emissions, assuming emissions from the rest of the world evolve proportionally.

Theoretically, we find that the optimal pollution tax is a modified Pigouvian rule that accounts

[1]In a representative-agent economy, if the government could, it would obtain all necessary revenue via undistortive lump-sum *taxes*. With inequality, the government typically prefers to provide lump-sum *transfers* for redistributive purposes resorting to distortionary taxation. In this setting, there is no need to restrict the government from choosing the level of lump-sum transfers or taxes.

for tax distortions via the marginal cost of public funds (MCF). However, because the government can optimally choose the level of lump-sum transfers (or taxes), the MCF is not higher than 1 as in representative-agent settings (see for instance Bovenberg and Goulder, 1996; Barrage, 2020).[2] In fact, we show that, when households have balanced-growth preferences, the MCF is on average equal to 1. A direct implication is that the optimal pollution tax may only temporarily lie above or below the Pigouvian level. These temporary tax distortions are driven by the costs associated with implementing the second-best allocation. While these costs are on average null because the planner can optimally choose the level of lump-sum transfers, they need 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 1, 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. This result contrasts with an influential finding from representative-agent settings that tax distortions lower the optimal tax on carbon, cf. Bovenberg and Goulder (1996) and Barrage (2020). Perhaps surprisingly, we find that the role of the MCF remains very small even if the government cannot optimally choose the level of lump-sum transfers. The reason is that, in the presence of inequality, the optimal level of distortionary taxes is much closer to the one that would be necessary to meet revenue requirements than in a representative-agent economy, so the cost associated with this constraint is significantly lowered. The SCC is, however, affected by the presence of inequalities, although the effect is not very large: consumption inequality—after income taxes are set optimally—reduces the SCC by 3.9% in our baseline calibration. The reason why the effect is not large is that 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.

We compare our optimal policy to the one of a "climate-skeptic" planner who optimises 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 labour income tax. Regarding the impact on 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—

[2] An MCF above 1 means that transferring resources from the private to the public sector is costly, in which case it is optimal to tax below the Pigouvian levy.

and very large positive but regressive effects afterwards, as richer households value environmental improvements proportionally more relative to consumption. Importantly, we show that this result stands in contrast to the representative-agent policy where the government uses the carbon tax revenue exclusively to reduce the level of distortionary taxes: the overall policy would then have strongly regressive effects. These results highlight the importance of considering household heterogeneity in climate policy, as distributional implications have welfare consequences, and also impact the political feasibility of climate policy.

We study the robustness of our main results to alternative calibrations and alternative settings. We show that the effects of inequality and tax distortions on optimal carbon taxes remain close to our baseline results when opting for a much more severe calibration of climate damages than the one in DICE-2016. Similarly, changing the amount of fiscal pressure has only minor effects on our quantitative results. The effect of inequality on optimal carbon taxes increases roughly proportionally with productivity heterogeneity and with the share of damages that directly affect utility. Furthermore, the effect of inequality changes more than proportionally with the intertemporal elasticity of substitution (IES): when the IES is reduced from our benchmark level of 1/1.45 (from DICE) to 1/2, inequalities lead to a reduction in carbon taxes of 16.2% instead of 3.9%. We also consider alternative policy scenarios: we study third-best policies, *i.e.* optimal fiscal policies when either the labour or the capital income tax is exogenously fixed at its current level. While the role of inequalities and the MCF remain very similar, these additional constraints lead to an additional fiscal interaction effect that has an ambiguous effect on the optimal carbon tax: we show that when these taxes are fixed at a sub-optimally low level, the optimal carbon tax should be reduced relative to its second-best level, and vice-versa. Finally, we show that our results hold under a more general specification of preferences than the one used in our benchmark, where environmental damages enter the utility function in an additively separable way.

To further explore the role of household inequality, we additionally consider households who differ in wealth, energy demand, and exposure to environmental damages. First, if the planner is unable to expropriate initial wealth, the optimal carbon tax is significantly affected to correct for this missing instrument, but only in the first period. Quantitatively, the roles of inequality and tax distortions remain very similar to the benchmark case. Second, we present a version of the model where households consume an additional dirty good that uses energy as its only input. To account for the fact that households are heterogeneously exposed to carbon taxes depending on their energy consumption, we consider households with non-homothetic utility and we use the Consumer Expenditure Survey (CEX) to match the observed distribution of energy budget shares in the US, both between and within income groups. We show that this additional source of heterogeneity does not significantly affect the role of inequality and tax distortions on the optimal carbon tax. While theoretically we find that it becomes optimal for the planner to subsidise the energy good if the households it values relatively more have higher energy needs, this additional instrument is quantitatively negligible. Naturally, this additional source of inequality also implies that ignoring household inequality in climate policy leads to even larger distributional effects. Third, we introduce heterogeneous sensitivity to environmental degradation. We theoretically show that if environmental damages enter the utility function in an additively separable way, heterogeneous environmental damages have no effect on the optimal pollution tax in the utilitarian case. However, they do lead to higher pollution taxes when the planner directly values more those households that are more affected by pollution. A Rawlsian planner would therefore tax pollution at a higher level if poorer households are also disproportionately affected by environmental damages.

Related literature 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 cannot optimally choose the level of lump-sum taxes or 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).[3] More recently, other papers have considered this problem with heterogeneous agents and a uniform lump-sum transfer (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.[4]

While these papers focus on static settings and model the pollution externality in a stylised manner, the recent work of Barrage (2020) 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 a representative-agent Ramsey model (see Chari and Kehoe, 1999, for a review). In this setting, tax distortions again call for lower taxes on carbon emissions. Our main innovation relative to Barrage (2020) is to introduce heterogeneous agents, which we see as important for two reasons. First, this allows us to jointly study optimal environmental and distributional policies, both theoretically and quantitatively in a dynamic, general equilibrium framework. In addition to 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. Second, agent heterogeneity provides a different rationale for distortionary taxation. In representative-agent settings, the government *needs to* rely on distortionary taxes to finance its expenditures, since lump-sum taxes are ruled out. In our setting, by contrast, the government *wants to* use distortionary taxes beyond what is needed for revenue in order to provide redistribution by rebating the proceeds through *uniform* lump-sum transfers.[5] The uniformity constraint plays the same functional role as the no lump-sum tax assumption in the representative-agent case: it limits the set of available instruments and rules out the first best.

[3]For further references on second-best pollution taxation in representative-agents models, see Barrage (2020).

[4]Other studies that analyse the joint design of pollution and non-linear income taxes conclude that deviations from the Pigouvian principle may be optimal—*e.g.*, Pirttilä and Tuomala (1997), Cremer et al. (1998), Cremer and Gahvari (2001) and Micheletto (2008). This is the case if doing so allows for a relaxation of incentive compatibility constraints. Kaplow (2012) also analyses a setting with pollution and non-linear income taxes, and shows that it is possible to construct Pareto improvements if pollution taxes deviate from their Pigouvian level.

[5]Unlike lump-sum taxes, lump-sum transfers are feasible in practice. Recent policy proposals even call for using such instruments to redistribute the carbon tax revenue (see the Economists Statement on Carbon Dividends signed by 3,354 American economists, Akerlof et al., 2019).

Recently, a few papers have investigated optimal carbon taxation with incomplete markets, therefore introducing both inequality and imperfect insurance (Le Grand et al., 2022; Belfiori et al., 2024; Wöhrmüller, 2024).[6] To maintain tractability and reduce computational complexity, these studies consider either more restrictive welfare objectives or constrained policy instruments. By contrast, our approach allows us to provide analytic expressions for a very general Ramsey problem focusing on the role of inequality. In a follow-up paper, Douenne, Dyrda, Hummel, and Pedroni (2025), we also address the distinct issue of idiosyncratic uninsurable risk and show that the key elements driving optimal deviations from Pigou in the present paper extend to this framework. While the follow-up study provides a richer environment to explore these elements quantitatively—building on the computational approach introduced by Dyrda and Pedroni (2023)—the present paper allows us to analytically characterise their determinants. In doing so, our paper also relates to a large public finance literature that studies redistribution while abstracting from insurance. This includes not only the Ramsey literature (for a recent example, see Straub and Werning, 2020), but also most of the Mirrleesian literature (e.g., Scheuer and Werning, 2017; Sachs et al., 2020; Ferey et al., 2024). Moreover, our model without risk endogenously matches, without targeting, the level of consumption inequality observed in the US, which is the key determinant of how inequality shapes the optimal carbon tax.

Although our optimal tax formulas resemble the ones in Barrage (2020), taking agent heterogeneity into account 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, similar to Jacobs and de Mooij (2015). At the optimum, the welfare gain from a marginal reduction in tax distortions is equal to the marginal cost from increasing inequalities, hence in our quantitative analysis the optimal policy divides the carbon tax revenue about equally between reducing tax distortions and providing redistribution.

Second, this paper contributes to the analysis of the distributional effects of environmental taxes in general equilibrium. An extensive literature has analysed 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 income (Levinson and OBrien, 2019). More recently, several authors have also analysed 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 labour 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 (2023) exploits exogenous shocks to the EU-ETS price to show that carbon taxation has a larger negative 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, with lump-sum transfers typically leading to less inequality and income tax cuts to higher levels

[6] A related literature studies the distributional effect of climate policy with incomplete markets (Benmir and Roman, 2022; Kuhn and Schlattmann, 2024), but not optimality.

of aggregate efficiency (*e.g.*, Williams et al., 2015; Fried et al., 2018, 2024; Goulder et al., 2019; van der Ploeg et al., 2022). Finally, a few papers have considered the heterogeneous environmental benefits of climate change mitigation between generations (*e.g.*, Leach, 2009; Kotlikoff et al., 2021), between countries/regions (*e.g.*, Hassler and Krusell, 2012; Krusell and Smith Jr, 2022; Cruz and Rossi-Hansberg, 2024; Bourany, 2024), or both (Kotlikoff et al., 2024).

We contribute to this literature by jointly analysing the economic and environmental impacts from optimal pollution taxation, both over time and between heterogeneous households who differ in income, wealth, and energy demand. We find that optimally introducing a carbon tax, while accounting for redistribution, leads to progressive welfare effects, in contrast with the representative-agent policy.

The rest of the paper is organised as follows. Section 2 presents the baseline model, and Section 3 the optimal tax formulas. Section 4 describes our calibration and Section 5 presents our main quantitative exercise. Section 6 considers the role of additional sources of inequality. Section 7 concludes.

2 Model

The model builds on Barrage (2020). One sector of the economy produces a final good using capital, labour, and energy, which is itself produced in the second sector. Energy production generates pollution that leads to environmental degradation, which affects productivity and households' utility. The government finances an exogenous stream of expenditures and lump-sum transfers using taxes on labour income, capital income, energy, and pollution. The key departures from Barrage (2020) are that, in our model, households are heterogeneous and the government can optimally choose the level of a (non-individualised) lump-sum transfer or tax. Consequently, although the government has access to a non-distortionary source of revenue, it uses distortionary taxes for redistributive purposes.

2.1 Households

We consider an economy populated by a continuum of infinitely-lived agents, or dynasties 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 per-capita consumption of a final good $c_{i,t}$, per-capita labour 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 heterogeneous in two dimensions: their productivity levels, e_i, and their initial asset holdings, $a_{i,0}$. Productivity levels are normalised such that $\sum_i \pi_i e_i = 1$. To focus on the effects of inequality without studying the implications of risk, we assume throughout that productivity levels are time-invariant. Hence, all inequality in labour income is permanent and does not result from idiosyncratic shocks to labour productivity. An equivalent interpretation is that ex-ante heterogeneous households face idiosyncratic earnings risk but can trade in complete asset markets. This modelling approach allows us to derive theoretical results as in Barrage (2020), Straub and Werning (2020), and

Chari et al. (2020), while at the same time capturing the fact that differences in labour income are, to a large degree, persistent (see, for example, Storesletten et al., 2004).

Agents' assets are composed of government debt and capital, and we denote $b_{i,t}$ and $k_{i,t}$ as 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: $C_t \equiv N_t \sum_i \pi_i c_{i,t}$, $H_t \equiv N_t \sum_i \pi_i e_i h_{i,t}$, $B_t \equiv N_t \sum_i \pi_i b_{i,t}$, and $K_t \equiv N_t \sum_i \pi_i k_{i,t}$. In addition, per capita 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 labour and capital income taxes, and T_t the *aggregate* uniform lump-sum transfers received by households in period t, which, importantly, is non-individualised and hence the same for all households.[7] 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 maximise (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

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

which holds across all agents. To simplify the exposition, we use subscripts x, i, t to denote partial derivatives with respect to argument x for agent of type i at time t, and suppress function arguments when there is no risk of confusion.

2.2 Final good sector

As in Barrage (2020), 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, $F\left(K_{1,t}, H_{1,t}, E_t\right)$, that uses capital $K_{1,t}$, labour $H_{1,t}$, and energy E_t. The total factor productivity

[7] In practice, while *individualised* lump-sum transfers are infeasible, transfers can be conditioned on some household characteristics. Letting the planner optimise over transfers that are "in between" uniform and individualised would bring the economy closer to first-best.

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

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

Firms in the final goods sector maximise profits by choosing capital, labour and energy. The first-order conditions are:

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

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

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

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 labour $H_{2,t}$ with a constant returns to scale technology so that

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

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 (as in Nordhaus, 2017). We choose this general specification because it nests the one used in Barrage (2020), where $\Theta_t(\mu_t, E_t) = \Theta_t(\mu_t E_t)$, and the one 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 $E_t^M = (1 - \mu_t)\, E_t$ which will be the carbon tax in our quantitative analysis. Firms maximise profits subject to the technology constraint given by equation (6) by choosing capital $K_{2,t}$, labour $H_{2,t}$, and the abatement share μ_t. The first-order conditions are

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

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

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

Profits in the energy sector are positive if there is positive abatement and $\Theta_t(\cdot,\cdot)$ is strictly convex in its second argument. To prevent the need to specify firm ownership, we assume these profits, if there are any, are taxed at a confiscatory rate $\tau_{\pi,t} = 1$.[8] As in DICE, these profits are zero in our quantitative analysis.

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

$$K_{1,t} + K_{2,t} = K_t, \tag{10}$$

$$H_{1,t} + H_{2,t} = H_t. \tag{11}$$

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 labour $\tau_{H,t}$, total energy taxes $\tau_{I,t}$, and emissions taxes $\tau_{E,t}$. The government's budget constraint is

$$R_0 B_0 + T + \sum_t p_t G_t = \sum_t p_t \left(\tau_{H,t} w_t H_t + \tau_{K,t} \left(r_t - \delta\right) K_t + \tau_{I,t} E_t + \tau_{E,t}(1-\mu_t) E_t + \Pi_t\right). \tag{12}$$

Although the instruments levied are proportional, the tax system is progressive when transfers are positive. The reason for focusing on linear instruments is twofold. First, as shown in Piketty and Saez (2013) and Dyrda and Pedroni (2023), an affine tax system provides a good approximation of the actual US tax system.[9] Second, with linear taxes the optimal tax problem can be formulated as a standard Ramsey problem despite the fact that individuals are heterogeneous in terms of their productivity levels and asset holdings—see Werning (2007). If instead the government could optimise non-linear taxes (for instance, on labour or capital income), the impact of inequality and tax distortions on optimal carbon taxes would likely be smaller, because the resulting allocation would be closer to first-best.

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

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

In our calibration below, J is a climate model that determines Z_t, the global mean temperature change relative to pre-industrial levels. 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 household utility.

[8] Doing so is typically optimal, as taxing pure profits does not generate distortions and income from shareholdings tends to be unequally distributed.

[9] As income rises, transfers are phased out and income taxes phased in, yielding an overall tax-transfer system that can be approximated by an affine function. A model with uniform transfers and linear income taxes, like ours, can thus serve as a reduced-form representation of the more complex system with targeted transfers and progressive taxes observed in the US. Normatively, the restriction is also not too severe: starting from the US system, Heathcote and Tsujiyama (2021) estimate that the welfare gain from moving to the optimal affine system is about two-thirds of that from implementing the fully nonlinear Mirrleesian optimum.

2.6 Competitive equilibrium

Given the resource constraint for the final good, which states that total output net of abatement costs can be used for private consumption, public consumption, and investment,

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

we define a competitive equilibrium as follows.

Definition *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 $\left\{(c_{i,t}, h_{i,t})_i, K_{1,t}, K_{2,t}, K_{t+1}, H_{1,t}, H_{2,t}, H_t, E_t, Z_t, \mu_t\right\}_{t=0}^{\infty}$ such that: (i) agents choose $\left\{(c_{i,t}, h_{i,t})_i\right\}_{t=0}^{\infty}$ to maximise utility subject to budget constraint (2) taking policies and prices (with $p_t = R_{t+1} p_{t+1}$) as given; (ii) firms maximise profits; (iii) the government's budget constraint (12) holds; (iv) markets clear: (6), (10), (11), and (14) hold; (v) the environmental variable follows (13).*

3 Optimal tax rules

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

3.1 Ramsey problem

A simple characterisation 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 $c_t \equiv C_t/N_t$ and $h_t \equiv H_t/N_t$ denote the average consumption and hours worked in period t. Following Werning (2007), it is therefore possible to split the optimal tax problem into two steps.[10] 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 $\varphi_i \geq 0$. These weights determine how aggregate consumption and labour supply are distributed among agents. Using the property that individual allocations are efficient given aggregates, we can characterise these individual allocations by solving the following static sub-problem for each period t, given weights φ:

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

Here, $U(c_t, h_t, Z_t; \varphi)$ denotes the indirect aggregate utility function, computed using market weights and aggregates. For simplicity, in what follows we assume that utility is additively separable in Z_t, so

[10] This would not be possible, for instance, if the government levied *non-linear* taxes on labour or capital income. In that case, not all households face the same marginal rate of substitution between consumption and leisure, or between consumption goods at different points in time. See also the discussion in Section 2.4.

that we can write

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

We show, in Appendix F, how our results generalise to non-separable preferences.

Implementability condition Applying the envelope theorem to problem (15) and combining the first-order conditions with those from the consumers' problem we get

$$\frac{U_{h,t}}{U_{c,t}} = \frac{u_{h,i,t}}{u_{c,i,t} e_i} = -w_t \left(1 - \tau_{H,t}\right), \qquad \text{and} \qquad \frac{U_{c,t}}{U_{c,0}} = \frac{u_{c,i,t}}{u_{c,i,0}} = \frac{p_t}{\beta^t}.$$

Using these relationships to substitute out for after-tax prices in each agent's lifetime budget constraint, 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:

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

with $c^m_{i,t}(c_t, h_t; \varphi)$ and $h^m_{i,t}(c_t, h_t; \varphi)$ 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), and (14) hold and there exist market weights φ and a lump-sum transfer T such that the implementability conditions (16) hold for all $i \in I$. Individual allocations can then be computed using functions $c^m_{i,t}$ and $h^m_{i,t}$, prices and taxes can be computed using the firms' and agents' first-order conditions.*

Optimal tax problem Let λ_i be the planner's welfare weight on type i, with $\sum_i \pi_i \lambda_i = 1$. Together with the concavity of the individual utility function, these weights determine the government's preferences for redistribution. The Ramsey problem is

$$\max_{\substack{\{c_t, H_{1,t}, H_{2,t}, K_{1,t}, K_{2,t},\\ E_t, Z_t, \mu_t\}_{t=0}^{\infty}, T, \varphi}} \sum_{t,i} \beta^t N_t \pi_i \lambda_i \Big(\tilde{u}\Big(c^m_{i,t}\big(c_t, h_t; \varphi\big), h^m_{i,t}\big(c_t, h_t; \varphi\big) \Big) + \hat{u}(Z_t) \Big), \tag{17}$$

subject to

$$\begin{aligned}
U_{c,0}\big(R_0 N_0 a_{i,0} + T\big) &= \sum_{t=0}^{\infty} \beta^t N_t \left(U_{c,t} c^m_{i,t}(c_t, h_t; \varphi) + U_{h,t} e_i h^m_{i,t}\left(c_t, h_t; \varphi\right)\right), \quad \forall\ i, \\
F_{K,t} G_{H,t} &= G_{K,t} F_{H,t}, \quad \forall\ t \geq 0, \\
N_t c_t + G_t + K_{t+1} + \Theta_t\left(\mu_t, E_t\right) &= \left(1 - D\left(Z_t\right)\right) A_{1,t} F\left(K_{1,t}, H_{1,t}, E_t\right) + (1-\delta) K_t, \quad \forall\ t \geq 0, \\
E_t &= A_{2,t} G\left(K_{2,t}, H_{2,t}\right), \quad \forall\ t \geq 0, \\
Z_t &= J\left(S_0, E^M_0, ..., E^M_t, \eta_0, ..., \eta_t\right), \quad \forall\ t \geq 0, \\
K_{1,t} + K_{2,t} &= K_t, \quad \forall\ t \geq 0, \\
H_{1,t} + H_{2,t} &= N_t h_t, \quad \forall\ t \geq 0.
\end{aligned}$$

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 labour income, initial assets and lump-sum transfers. The second constraint states that the marginal rate of technical substitution between capital and labour is the same in both sectors, a restriction the government needs to satisfy because it does not use sector-specific instruments and factors are mobile across sectors. The other constraints reflect market clearing for capital, labour, and goods, as well as 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.[11] An equivalent interpretation is that there is initial wealth inequality, but that all initial wealth can be expropriated by the planner. This would be optimal and can be achieved by taxing it directly, setting $R_0 = 0$, or through a combination of consumption and labour taxes—see Werning (2007) for a discussion.[12] In Section 6.1, we relax the assumption that there is no initial wealth inequality, or equivalently that all initial wealth can be expropriated, and study the implications for optimal taxes. 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.[13]

3.2 General formulas

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

$$\tau_{H,t} = 1 - \frac{U_{h,t}}{U_{c,t}} \frac{W_{c,t}}{W_{h,t}}, \qquad \text{and} \qquad \frac{R_{t+1}}{R^*_{t+1}} = \frac{W_{c,t+1}}{W_{c,t}} \frac{U_{c,t}}{U_{c,t+1}},$$

where $R^*_{t+1} \equiv 1 + (1 - D(Z_{t+1}))A_{1,t+1}F_{K,t+1} - \delta$ is the social return to capital, and the pseudo-utility function W is defined as

$$W(c_t, h_t, Z_t; \varphi, \theta, \lambda) \equiv V(c_t, h_t, Z_t; \varphi, \lambda) + \sum_i \pi_i \theta_i \mathcal{I}_i(c_t, h_t, \varphi),$$

with

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

denoting the aggregate utility based on the planner's weights,

$$\mathcal{I}_i(c_t, h_t, \varphi) \equiv U_{c,t} c^m_{i,t}(c_t, h_t; \varphi) + U_{h,t} e_i h^m_{i,t}(c_t, h_t; \varphi), \tag{19}$$

[11]Notice that the absence of initial wealth inequality does not mean there is no wealth inequality in future periods: in our model, persistent differences in labour productivity ultimately lead to differences in wealth.

[12]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.

[13]If there is initial wealth inequality, the level of $\tau_{K,0}$ is no longer indeterminate. However, when studying the impact of initial wealth inequality on optimal taxes in Section 6.1, we also treat $\tau_{K,0}$ as given. The reason for doing so is that optimising over $\tau_{K,0}$ allows the planner to confiscate all initial wealth, which immediately gets rid of all initial wealth *inequality* as well.

denoting the difference between agent i spending on consumption and labour income in period t as it appears in their implementability constraint, and $\pi_i\theta_i$ denoting 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), and hold under more general utility specifications—Appendix F shows how they generalise when utility is not separable with respect to Z_t.

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

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

As long as labour, capital, profits, and pollution can be taxed, there is no point in distorting production decisions. This result is also found in Bovenberg and Goulder (1996) and Barrage (2020), and goes back to the production efficiency theorem of Diamond and Mirrlees (1971). Turning to the pollution tax, we obtain

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

When the pollution tax increases, abatement increases, which in turn increases the scarcity of the final good. The opportunity cost of reducing emissions by increasing the pollution tax, therefore, corresponds to the marginal cost of increasing the final good's scarcity, captured by the term $V_{c,t} + \sum_i \pi_i \theta_i \mathcal{I}_{c,i,t}$. The latter is equal to the marginal utility of raising aggregate consumption as computed using the planner's weights, $V_{c,t}$, plus a term that captures the marginal reduction in the planner's implementation cost from an increase in aggregate consumption, $\sum_i \pi_i \theta_i \mathcal{I}_{c,i,t}$. Intuitively, θ_i represents the shadow cost of transferring one unit of consumption to household i, and $\sum_i \pi_i \theta_i \mathcal{I}_{i,t}$ 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 $\mathcal{I}_{c,i,t}$.

Importantly, equation (20) holds both in the first-best (with $\theta_i = 0$ for all i) and in the second-best. Still, the optimal pollution tax may differ between these two fiscal environments for three reasons: (i) the path of aggregate variables, (ii) the value of the marginal implementation cost, and (iii) given aggregates, the distribution of individual allocations. The first of these captures differences in the size of the economy between first-best, where distortionary taxes are equal to zero and redistribution is achieved through individualised lump-sum taxes and transfers, and second-best, where distortionary taxes are used to provide redistribution. The second and third capture the effect of tax distortions and the effect of inequality, respectively. We explain these forces, which we attempt to quantify in Section 5, in more detail below.

3.3 Comparison with first-best

The role of tax distortions The first potential difference between the first- and the second-best pollution tax lies in the value of the marginal reduction in implementation cost, $\sum_i \pi_i \theta_i \mathcal{I}_{c,i,t}$. The first-best allocation—the one that maximises welfare subject to only resource constraints—is achieved when the planner has access to individualised lump-sum transfers. The corresponding first-order conditions imply

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

In words, the government sets the individualised lump-sum transfer in such a way that the implementability condition for each agent is not binding. It follows that the planner can achieve full redistribution 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 is equal to the discounted sum of marginal (utility and production) damages valued at the marginal utility of consumption.

The second-best allocation maximises welfare subject to resource constraints *and* implementability conditions (16). The first-order condition with respect to the uniform, non-individualised transfer gives

$$\sum_i \pi_i \theta_i = 0.$$

The non-individualised transfer is set in such a way that the implementability condition is not binding *on average*. From this it follows that the impact of raising aggregate consumption on the implementation costs is

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

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,

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

Notice that $\tau_{E,t}^{Pigou}$ denotes the Pigouvian tax *rule*, which can be evaluated at different allocations. In particular, $\tau_{E,t}^{FB}$ corresponds to a special case where $\tau_{E,t}^{Pigou}$ is evaluated at the first-best allocation.

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,*[14]

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

where $\nu_{1,t}$ denotes the Lagrange multiplier on the final-goods resource constraint.

Recall that $V_{c,t}$ measures the marginal utility of raising aggregate consumption, c_t, computed using the planners weights. This is used to convert the welfare costs of environmental damages and tax distortions into consumption units in the computation of both the Pigouvian tax and the MCF. Without inequality, $V_{c,t}$ coincides with the marginal utility of consumption of the representative agent. With inequality, the welfare benefit of raising aggregate consumption generally depends on the distribution of marginal utilities.

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 MCF,*

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

with

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

If agents have balanced-growth preferences, then from period 0 the welfare-weighted average MCF is 1,

$$\frac{\sum_{t=0}^{\infty} N_t \beta^t V_t \times \mathrm{MCF}_t}{\sum_{t=0}^{\infty} N_t \beta^t V_t} = 1, \tag{24}$$

with $V_t \equiv V\left(c_t, h_t, Z_t; \varphi, \lambda\right)$. If the IES is equal to 1, then $\mathrm{MCF}_t = 1$ for all $t \geq 0$.

[14] Jacobs and de Mooij (2015) and Jacobs and van der Ploeg (2019) use an alternative definition of the MCF that takes into account fiscal externalities resulting from income effects. They find that the MCF equals 1 at the optimal tax system, owing to the fact that the government can optimise a lump-sum transfer (see also Jacobs, 2018). However, because as in Barrage (2020) we optimise over the allocation variables directly rather than over tax instruments, it is more suitable to define the marginal costs of funds as the ratio between the multiplier on the government budget constraint and the social welfare impact of raising aggregate consumption. This definition is closer to what Jacobs and de Mooij (2015) refer to as the traditional measure of the MCF, and, as we show below, captures the difference between optimal pollution taxes in first-best and second-best.

The proof of Proposition 2 is provided in Appendix A.5. The optimal pollution tax balances the marginal benefits of pollution abatement against the opportunity cost from reductions in aggregate consumption. In the first-best, this opportunity cost is given by the marginal utility of aggregate consumption, $V_{c,t}$. In the second-best, see equation (22), 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 $V_{c,t} \times \text{MCF}_t$, hence the opportunity cost of abatement is higher than in the first-best if and only if the MCF is above 1. Tax distortions also affect the marginal benefits of pollution abatement through the valuation of future production damages. In particular, when the MCF decreases over time, tax distortions operate as a form of discounting: consumption is valued relatively more in the present than in the future, hence future production damages receive a lower weight than in the first-best Pigouvian rule. Put differently, in this case, taxing in the present is more costly than taxing in the future. We show in Appendix A.5 that the ratio of MCFs can be expressed as

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

from which we see that the MCF is constant if the capital tax is null for all future periods. Thus, as in Barrage (2020), the optimal tax on production damages is not distorted as long as, going forward, the capital income tax is optimally set to zero. Intuitively, in this situation, tax distortions affect future marginal abatement benefits proportionally to current marginal abatement costs. Production damages are then perfectly internalised, and the optimal pollution tax can be expressed as

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

In this case, optimal carbon taxes are affected by the MCF in proportion to the share of utility damages.

Proposition 2 additionally provides an expression for the MCF as a function of the covariance between θ_i and $\mathcal{I}_{c,i,t}$, see equation (23). 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 zero on average at the optimum, in the typical case where the government wishes to redistribute from rich to poor agents, θ_i is positive for the rich and negative for the poor.[15] The second term, $\mathcal{I}_{c,i,t}$, represents how the difference between a household's current consumption expenditure and labour 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 fewer resources are used for pollution abatement, consumption increases, and labour supply adjusts, which also affects prices and wages. When households

[15] As shown in Appendix A.4.2, with balanced-growth preferences,

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

The ratio λ_i/φ_i captures how much the planner values agent i relative to the market; θ_i is the difference between this ratio and its population average. It is negative when the planner values the agent more than the market does. Since the market weight, φ_i, reflects how rich the agent is (in present value), a redistributive planner assigns higher ratios to poorer agents and lower ones to richer agents, yielding negative θ_i for the poor and positive for the rich.

have balanced-growth preferences and the IES is equal to 1, these two effects exactly offset each other, such that the present value of resources necessary to satisfy households' budget constraints remains unchanged. In this situation, taxing pollution does not affect the implementation costs that result from distortionary taxes; the MCF is equal to 1, and the second-best tax is exactly Pigouvian. When the IES is below 1, the price effect dominates, and an increase in aggregate consumption reduces the total amount of transfers needed to satisfy agents' implementability constraints. If, at a particular point in time, these changes are heterogeneous across households and correlate with their type, the MCF differs from 1.

From the households' implementability conditions, we have that

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

and it follows that, with no initial wealth inequality (or equivalently, with full expropriation of initial wealth), the discounted sum of $\mathcal{I}_{i,t}$ is invariant across types. Intuitively, with a uniform lump-sum transfer and no wealth inequality, the discounted sum of expenditures minus labour 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, see equation (24). Hence, the MCF is *on average* equal to 1, and the optimal pollution tax is *on average* equal to the Pigouvian level.[16] Still, in any period $t \geq 0$, the MCF may differ from 1, and 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 labour 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 MCF 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: because individualised lump-sum transfers are not feasible in the second-best, even with optimal redistribution there is some *residual inequality*. It is this inequality that affects the welfare gains from leaving more resources available for agents' consumption by decreasing the pollution tax. The effect of this residual inequality on the optimal pollution tax depends on the curvature of agents' utility function.[17]

[16] This result is related to Jacobs and de Mooij (2015), who, using a different definition of the MCF (see footnote 14), show in a static model that the MCF is equal to 1 provided the government optimises a lump-sum transfer. Consequently, tax distortions do not call for a lower pollution tax. Kaplow (2012) also argues that concerns about distortionary effects from taxes on labour supply are independent of the question of how to tax externalities. According to Proposition 2, a similar logic holds *on average* in dynamic environments as well.

[17] In addition, the presence of distortionary taxation affects the path of aggregate variables, and thereby the level of environmental taxes.

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*

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

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, $u_{c,i,t}$. If the IES is equal to 1, the two effects exactly offset each other and consumption inequalities do not affect the pollution tax.

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 increase in consumption disproportionately goes to richer households with lower marginal utilities. An increase in the pollution tax reduces every households' consumption proportionally. When the IES is equal 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.[18] 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. Hence, if the IES is below one, inequality lowers the optimal pollution tax.

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 labour income. Because fiscal policy is typically decided on at the national level, we calibrate the economic features of our model based on one country. We assume that country takes into account the global impact of its emissions, leaving strategic considerations aside. Specifically, our baseline calibration adopts the economic features of the US and the climate model from DICE-2016. For consistency, we scale the 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 US should design its fiscal system to internalise the global effect of its externalities under the assumption that global emissions are proportional to its emissions.[19]

[18]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 a proportional increase in all agents' consumption:

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

[19]US GDP is roughly one-quarter of world GDP. So, having matched world GDP, we set the initial population to that of the US multiplied by four, in order to also match US GDP per capita. We then separately calibrate emissions to

4.1 Climate model

The calibration of the climate model is based on the 2016 version of DICE, presented 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_t^{At}, S_t^{Up}, and S_t^{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 $j \in \{At, Up, Lo\}$ are ranked as above and with E_t^{land} denoting exogenous land emissions. The coefficient $b_{0,j}$ is 1 for the first reservoir, S_t^{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, and hence the climate warms.

Climate change The temperature is modelled through two equations for the mean temperature change 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 Appendix G.

match global levels. One interpretation is that the world economy consists of a single global economy, equivalent to four US economies, setting income and climate taxes to address US-like inequality and global climate change (Figure 21a in Appendix H.4 shows how our results vary with inequality). Alternatively, and equivalently, our preferred interpretation is that the model describes the optimal policy for the US under the assumption that the rest of the worlds emissions are proportional to its own, and that global damages are internalised as if they affect others analogously to how they affect the US. This is consistent with treating each unit of US emissions as scaled to the global level, with global-level damages then taken into account.

4.2 Damages

We also model production damages as in DICE-2016, with

$$D(Z_t) = a_1 Z_t + a_2 Z_t^{a_3}, \tag{25}$$

We assume that $D(Z_t)$ is a simple quadratic function with $a_1 = 0$ and $a_3 = 2$. The relevant Z_t that enters this formula is the atmospheric temperature change, Z_t^{At}. Since DICE does not distinguish between production and utility damages, we follow Barrage (2020) 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 (25) relative to DICE and enables us to calibrate utility damages, specifically, the preference parameter α_0 described below.

To examine to what extent our results depend on the underlying SCC, which could depend on the choice of climate model, 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 underestimates 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 utility damage from temperature increases given by

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

as in Barrage (2020). Using market weights, the intertemporal aggregate utility can be written as

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

with $\Gamma \equiv \sum_i \pi_i \varphi_i$, and where $Z_t \equiv Z_t^{At}$ is the atmospheric temperature change (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 US, we set the population levels as US population multiplied by the ratio of world GDP to US GDP in 2011-2015, the first period of the model.

Following DICE, we set the 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 labour supply of 0.75 (see Chetty et al., 2011) and an average per capita labour supply of $h_{2015} = 0.277$ in the initial period (computed from the Survey of Consumer Finances, see Appendix G.4).

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 normalise productivity levels such that $\sum_i \pi_i e_i = 1$. The full procedure is described in Appendix G.2. While we calibrate productivity levels directly instead of targeting a specific *ex post* distribution, 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 abstracts 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, 2020). 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.[20] 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 set the exponent $c_2 = 2.6$ as in DICE.

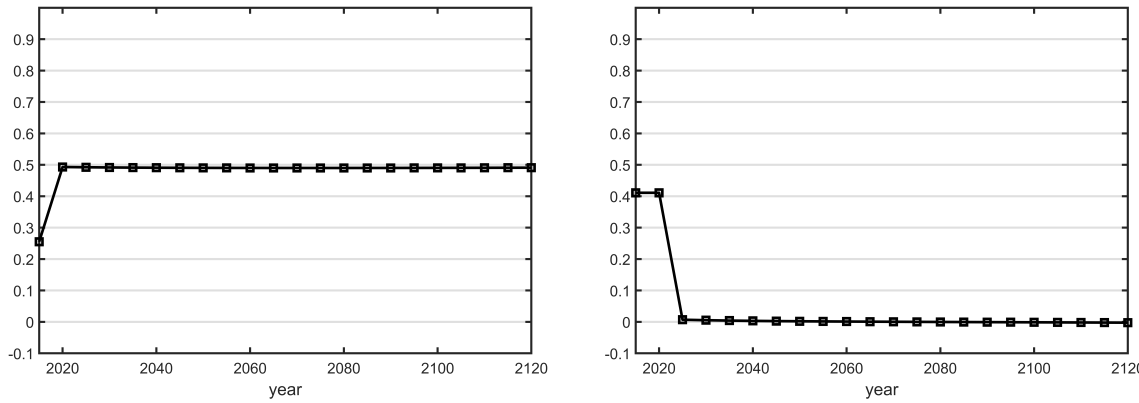
4.5 Government

We set the tax rates on capital and labour income in line with effective rates computed by Trabandt and Uhlig (2012), at $\tau_K = 0.411$ and $\tau_H = 0.255$ in our baseline (see Appendix G.1). We follow Barrage (2020) and set the intermediate-goods tax to $\tau_I = 0$. The tax on carbon emissions, τ_E, is set at a level so that, in our calibrated economy, 3% of total energy is obtained from clean technologies (Nordhaus, 2017). This requires $\tau_E = 2.01$\$/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 US government's balance sheet, both as a percentage of GDP.[21] Following Barrage (2020) and in order to facilitate reproducing results for other

[20] 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.

[21] The numbers are calculated at the "General Government" level. We also explore the sensitivity of our results to different levels of initial government debt.



(a) Optimal labour income taxes

(b) Optimal capital income taxes

Figure 1: Optimal Income Taxes.

Notes: Figures show the path of second-best labour 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).

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.

Turning to government spending, in our model G_t denotes government consumption, while T captures the present value of all lump-sum transfers. To better align the model with the data, we follow Barrage (2020) and split total government spending into final good spending G_t^C, and *exogenous* transfers G_t^T. The total transfers that households receive consist of the exogenous component, G_t^T, and the endogenous component, T. As in Barrage (2020), empirical counterparts of G_t^C and G_t^T are obtained from the IMF Government Finance Statistics. The initial value of government consumption, G_0^C, is 15.8% of GDP, while the exogenous transfers, G_0^T, are set at 14.5% of GDP. See Appendix G.1 for details.

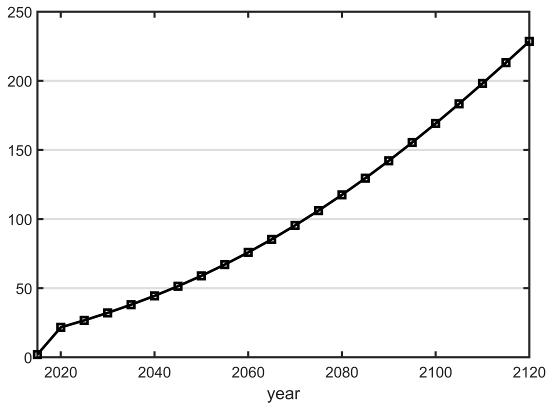
5 Quantitative results

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.[22]

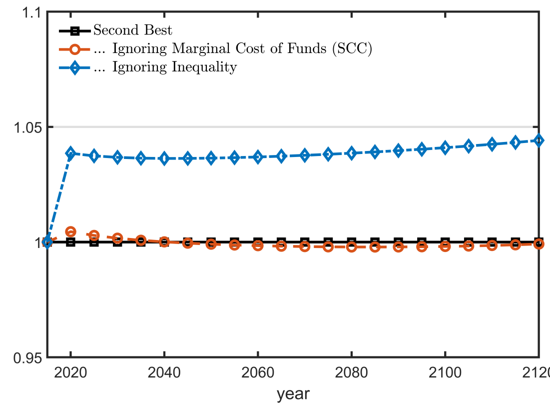
5.1 Optimal policy

Optimal tax paths Figure 1 shows the path of optimal taxes on capital and labour income in our baseline scenario. Compared to their calibrated levels, the labour income tax roughly doubles in the first period, from 25% to about 50%, and stabilises 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 transfers are available and there is no initial wealth inequality, the only reason to tax capital

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



(a) Optimal carbon taxes ($\$/tCO_2$)



(b) Carbon tax decomposition

Figure 2: Carbon Taxation.

Notes: (a) Figure shows the path of second-best carbon taxes for the baseline calibration expressed in dollars per ton of CO_2. The initial level (for 2015) is set exogenously to its current level obtained from Nordhaus (2017). (b) The black line represents the second-best carbon tax normalised 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.

income is to mitigate intertemporal distortions associated with labour income taxation. Since optimal labour income taxes are close to constant, the optimal capital income tax converges to zero quickly after the second period.[23] Appendix E examines scenarios with further constraints on policy instruments leading to deviations from this result.

Figure 2a 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 (2020) 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 the "high damages" scenario, introduced in Section 4.2, it is optimal to stay close to the Paris objective. Then, the optimal income taxes remain almost the same, while the carbon tax is roughly four times as large (see Appendix H.1).

Carbon tax decomposition Figure 2b shows the implications of tax distortions and inequality on optimal carbon taxes, for given paths of aggregate variables.[24] The black line plots the second-best pollution tax, normalised to 1. The red line, in turn, shows what the pollution tax would be if the MCF is set to 1 in each period. This level corresponds to the SCC. Recall from Proposition 2 that the MCF

[23]Notice that, because the government has access to lump-sum taxes, 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).

[24]To focus exclusively on the effects of tax distortions and inequality, we thus abstract from differences in the size of the economy between first-best and second-best. Put differently, our focus is not on comparing the optimal carbon tax in a first-best versus a second-best world. Rather, we are interested in assessing how far off a planner would be if they applied the first-best formula—thus ignoring the role of inequality and tax distortions—in a second-best setting.

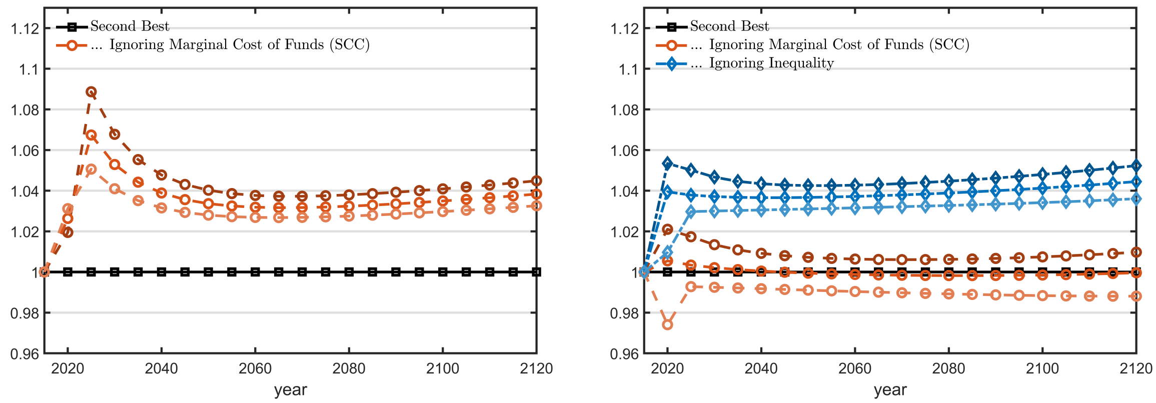
is 1 *on average*. As it turns out, temporary deviations of the MCF from unity play an insignificant role: the second-best carbon tax is initially only 0.5% below the SCC, a difference that becomes even smaller in subsequent periods. Thus, even in the presence of distortionary taxation, it is optimal to set the carbon tax approximately equal to the SCC in every period (*i.e.*, at the Pigouvian level). This result contrasts with Barrage (2020), who, in a representative-agent setting, finds that the MCF exceeds one, which leads to optimal carbon taxes lower than the SCC.[25] The reason is that in her model, the government *has* to rely on distortionary taxes to generate revenue. By contrast, in our framework, the government optimally *chooses* to levy distortionary taxes for redistributive purposes. Tax distortions do not call for deviations from the SCC to the extent that they are used to provide redistribution.

To study the role of inequality, the blue line plots what happens to the carbon tax when we abstract from household inequality (*i.e.*, with all households' consumption and labour supply equal to the aggregate values). The discrepancy between the blue and red lines indicates the degree to which inequality matters for optimal carbon taxes, through its impact on the SCC. The SCC is determined by the trade-off between reducing damages and increasing aggregate consumption. As explained below Proposition 3, the residual inequality that remains after redistributive taxes have been optimised raises the value of increasing aggregate consumption if $\sigma > 1$. With $\sigma = 1.45$, ignoring inequality leads to an SCC that is on average 3.9% higher over the next century. Hence, taking inequality into account has a moderately negative impact on optimal carbon taxes.

Role of lump-sum transfers A potential explanation for why the MCF has such a small impact on optimal carbon taxes in our setting is that the government has access to a non-distortionary source of revenue. Perhaps surprisingly, this is not the case: with inequality, the MCF is close to 1 even if we remove the planner's ability to adjust lump-sum transfers. To make this point clear we consider the following two experiments. In the first, we shut down inequality and the ability of the planner to change the level of lump-sum transfers. Our setting then simplifies to a representative-agent environment and the experiment confirms the main results obtained in Barrage (2020). Figure 3a shows the effect of the MCF, which is sizeable when compared to our benchmark results; tax distortions call for lower carbon taxes. We also show how the results change when we vary the amount of fiscal pressure by reducing the initial level of net government debt to zero (light-shaded lines), and by doubling the benchmark value (dark-shaded lines). Naturally, the MCF calls for a larger reduction of carbon taxes when fiscal pressure is higher.

In the second experiment, we maintain the constraint that the planner cannot adjust lump-sum transfers but reintroduce labour income inequality. Figure 3b shows that, unlike in the representative-agent case, the effect of the MCF is reduced to virtually zero—similarly to the benchmark results from Figure 2b. The reason behind this perhaps surprising result is that preventing the government from

[25] Specifically, Barrage (2020) finds that, in the 21$^{\text{st}}$ century, optimal carbon taxes are about 8% lower when there are distortionary taxes. Unlike our approach, which compares formulas evaluated at the same allocation, this number also reflects residual differences in aggregates after the economy is re-calibrated to resemble the second-best setting more closely. Figure 2 of Barrage (2020) suggests that the majority of this effect is driven by differences in the MCF, especially in earlier periods.



(a) Representative-agent economy

(b) Labour income inequality

Figure 3: Fixing Lump-Sum Transfers.

Notes: The black line represents the second-best carbon tax normalised to 1. The red lines show what this tax would be if the MCF was set to 1 in all periods, holding aggregates constant. The blue lines show what this tax would be absent consumption inequalities, again holding aggregates constant. All taxes are computed under the baseline calibration except for the fact that for the light-shaded lines we reduce the initial net government debt to zero, while for the dark-shaded lines, we double the benchmark value.

adjusting the level of lump-sum transfers affects the level of tax distortions only to the extent that this constraint is binding. In the representative-agent economy, the planner would optimally want to set the endogenous transfer T_t to a large negative value, such that the total lump-sum tax $-(G_t^T + T_t)$ finances all its exogenous spending with income taxes optimally set to zero. Imposing $T_t = 0$, for all t, therefore has a large impact on the amount of tax distortions. With inequality, the planner would optimally choose to provide lump-sum transfers for redistribution, hence the constraint over T_t is less binding. In our benchmark, it happens to be optimal to keep lump-sum transfers, $G_t^T + T_t$, close to their calibrated level, G_t^T, hence the constraint that the planner cannot adjust lump-sum transfers is hardly binding. Notice, however, that even with the substantial changes to fiscal pressure associated with eliminating or doubling net government debt, the effect is still subdued. The effect of the MCF can even flip sign when meeting the necessary revenue actually requires lower income taxes compared to what an unconstrained planner would choose. This happens if net government debt is set to zero, in which case ignoring the MCF leads to a small reduction in carbon taxes: see the light-shaded curve in Figure 3b. Finally, notice that the effect of inequality relative to the effect of the MCF, as captured by the distance between the blue and red lines, is fairly robust to changes in fiscal pressure.

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 H.1, we also show that with a more severe calibration of climate damages leading to an SCC about four times higher, the role of the MCF remains negligible while the effect of inequalities decreases, to 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 H.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. In line with Proposition 2, 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 the 26% from 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%.[26] Finally, we also consider different levels of income inequality (see Figure 21a in Appendix H.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 US.

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 H.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 $\sigma = 1$. 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 stronger effects of inequality.

5.2 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 were 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 labour tax decreasing by 0.7% of GDP, and increasing transfers, whose present value increases by 0.8% of GDP.[27,28] Intuitively, because labour taxes and transfers are optimised, it is optimal to use the revenues from carbon taxation

[26] Cruz and Rossi-Hansberg (2024) estimate climate damages on productivity and amenities, and find that in the long run, up to half of the damages impact utility directly through amenities (in which case the effect of inequalities would go up to 5.9%). In the short run, direct utility damages represent a smaller share, more consistent with our calibration obtained from Barrage (2020).

[27] 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.

[28] This result echoes the recent findings of van der Ploeg et al. (2022), who argue that rebating carbon taxes using a combination of higher transfers and lower income taxes is most effective in garnering public support. Fried et al. (2024) study the optimal recycling policy for an exogenous carbon tax introduced in a sub-optimal tax system and find that two-thirds of the carbon tax revenue should be used to reduce taxes on capital income and one-third to provide redistribution.

to *both* reduce labour taxes and raise transfers, rather than doing only one of these. 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 exclusively 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 other means. This result also gives some grounds to the popular carbon tax and dividend policy (see the Economists Statement on Carbon Dividends, Akerlof et al., 2019) that calls for redistributing the proceeds of the tax via lump-sum transfers 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		
	Labour	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.

5.3 Welfare effects

Figure 4a 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 decade and for each productivity group. The average inter-temporal gains are positive for all productivity groups—the average discounted gain is 5.8% with baseline damages.[29] However, the period welfare gains, which determine how different generations are affected, are heterogeneous over time and between groups. Overall, welfare

[29] The comparable number in Barrage (2020) is about 1%. Within this century, our period welfare gains are of similar magnitude but begin to diverge afterward. These differences stem mainly from assumptions about carbon taxation in the next century. Specifically, Barrage (2020) states that "carbon taxes are allowed after 2115 so as to keep the analysis in an appropriate range for the (smooth) damage function," which limits the business-as-usual temperature increase to 4.3 degrees. In contrast, we keep carbon taxes at zero until 2240 (see Appendix I), resulting in temperature increases exceeding 12 degrees by then (see Figure 11). This leads to larger welfare effects, amplified by convex damages and rising consumption levels due to economic growth—which increases the consumption value of improved climate, a luxury good under additively separable preferences. While such long-run projections warrant caution (the IPCC typically limits projections to 2100), their welfare implications can be substantial. Accordingly, our 5.8% average welfare gain should be interpreted with caution, while near-term period welfare gains in Figure 4 are more reliably estimated.

gains increase dramatically after the 21st century.[30] 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 labour income tax rate (see Table I). The overall progressivity of the tax system increases, which makes poorer households benefit more 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 lower temperatures comes from reduced utility damages. Richer households care relatively more about those damages in the sense that, since 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. It is worth emphasising that this exercise abstracts from heterogeneity in climate damages, an extension that we theoretically investigate in Section 6.3.

(a) Optimal policy

(b) SCC and double dividend

Figure 4: Period Welfare Gains (%).

Notes: For each decade and each income decile the table shows the welfare gains, in percentage of consumption, of moving from a scenario without carbon taxation and otherwise optimal income taxes to: (a) optimal carbon and income taxation accounting for inequality; (b) carbon taxes set to the SCC and extra revenue used to reduce labour income taxes. Numbers are computed under the baseline calibration. The large welfare gains in the 22nd century reflect the fact that, as consumption levels grow and environmental quality declines, households are willing to give up more consumption to prevent an increase in temperature.

To better understand the benefits of combining optimal carbon taxation with optimal income taxation, we present, in Figure 4b, the result of another experiment. Starting again from the optimal policy of a climate-skeptic planner, we set carbon taxes to the SCC, ignoring inequalities, and we use all additional revenue exclusively to reduce distortive labour-income taxes—following the prescription associated with the double-dividend hypothesis. In this case, the policy is regressive from the beginning with poorer households bearing the bulk of the costs associated with the introduction of carbon taxation. These results may illustrate why carbon tax policies are often considered unpopular because

[30] As shown in Dietz et al. (2021), the DICE-2016 model features relatively high thermal inertia, *i.e.*, the temperature response to an impulse in emissions is delayed compared to what climate science models predict (note that this issue was addressed in the updated version of the model, DICE-2023, see Barrage and Nordhaus, 2024). If this response was more immediate, welfare gains from carbon taxation could become positive earlier.

of their potentially regressive effects.[31] Our results indicate that combining the implementation of carbon taxation with an appropriate increase in the progressivity of the tax system can, therefore, be important for a more equitable sharing of the gains from carbon taxation. Then, even though these gains 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.[32]

6 Additional sources of heterogeneity

In this section, we study the effects of introducing additional sources of heterogeneity. We consider, in turn, inequality in wealth, energy demand, and sensitivity to environmental damages.

6.1 Initial wealth inequality

With inequality in initial asset holdings, if the planner is allowed to set the tax on capital income in the first period, it is optimal to fully expropriate initial wealth (provided that 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 US. In Appendix B.1, we also discuss the implications of initial wealth inequality for the time-consistency of Ramsey policies.

6.1.1 Optimal tax rules

For $t \geq 1$, the optimal tax rules are not affected by the presence of initial wealth inequality.[33] 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 B.6):

$$\tau_{E,0} = \frac{1}{\nu_{1,0}} \left(\sum_{j=0}^{\infty} \beta^j \left(\nu_{1,j} D'_j A_{1,j} F_j - N_j W_{Z,j} \right) J_{E_0^M,j} - N_0 U_{c,0} \Delta \left(1 - \tau_{K,0}\right) D'_0 A_{1,0} F_{K,0} J_{E_0^M,0} \right), \quad (26)$$

[31] The French Yellow Vests movement offers a good example: following the double-dividend strategy, the French government simultaneously increased the carbon tax while reducing taxes on capital and labour. Concerns about the impact on the purchasing power of (poor) households led to massive social unrest and the ultimate withdrawal of the planned carbon tax increases (see Douenne and Fabre, 2022).

[32] Goulder et al. (2019), among others, argue that carbon taxes can have progressive effects even with revenue recycled through labour income taxes, as inflation indexing raises real transfers for lower-income households. In our model, which abstracts from nominal rigidities, this effect is captured directly through an explicit increase in real transfers.

[33] The exception is the tax rule for $\tau_{K,1}$. See Appendix B.5 for details.

where

$$\nu_{1,0} = W_{c,0} - U_{cc,0} R_0 \Delta \tag{27}$$

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 leads to an additional term, the last one in equation (26); 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 $\tau_{K,0}$, which is equivalent to having the planner expropriate all initial wealth and then partially returning it 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 reducing aggregate consumption exacerbates the cost of initial wealth inequality by increasing the consumption value of initial wealth—see equation (27). 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.

6.1.2 Quantitative analysis of the effect of wealth inequality

We calibrate the joint distribution of productivity and initial wealth using data 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 G.3. We fix $\tau_{K,0}$ to be at the same level as in the current tax system, at 41.1%.

Figure 5 below provides a decomposition similar to the one shown in Figure 2b above.[34] 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, $\nu_{1,0}$. To mitigate the negative impact on productive efficiency, the large temporary decrease in the optimal carbon tax is accompanied by an equivalent increase in the energy tax, τ_I. Despite the large impact of wealth inequality on carbon taxes in the initial period, the effects of tax distortions (red vs. green) and consumption inequalities (blue vs. red) remain very similar to the baseline in subsequent periods.

[34] Appendix H.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.

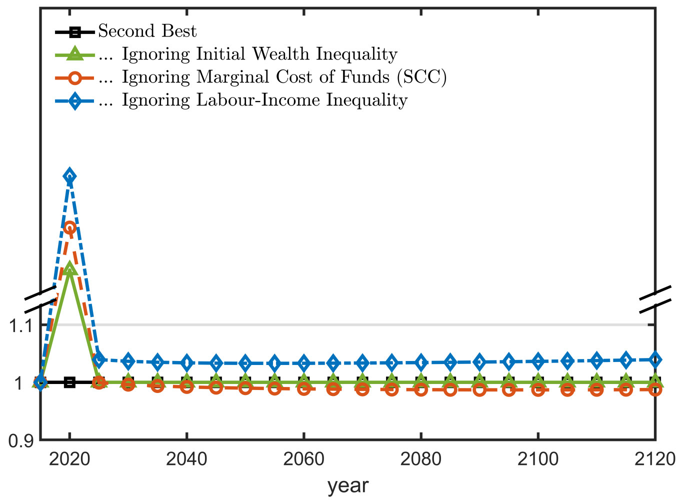


Figure 5: Carbon Tax Decomposition, Initial Wealth Heterogeneity and Exogenous Initial Capital Tax.

Notes: The black line represents the second-best carbon tax normalised 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 (26) if Δ is set to zero. As in Figure 2b, 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.

6.2 Energy consumption inequality

A key concern regarding the distributional impact of carbon taxation is that households' energy budget shares differ, both between and within income groups (see *e.g.*, Pizer and Sexton, 2019). To explore this issue, we now introduce into our benchmark model a second *dirtier* consumption good and heterogeneous preferences over this good.

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 $d_{i,t}$, labour 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, $d_{i,t}$, is produced with a linear technology that uses energy as its only input. For simplicity, 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 $E_{1,t}$ denoting 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' per capita energy consumption. In order to match empirically observed budget shares for energy for different income groups, we assume households' utility can be represented by the following period utility function,

$$u_i\left(c_{i,t}, d_{i,t}, h_{i,t}, Z_t\right) = \frac{\left(c_{i,t}(d_{i,t} - \bar{d}_{i,t})^{\epsilon}(1 - \varsigma h_{i,t})^{\gamma}\right)^{1-\sigma}}{1-\sigma} + \hat{u}(Z_t). \tag{28}$$

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 modelled with a Stone-Geary utility function, so that an agent of type i experiences positive utility from energy consumption only after consuming its first $\bar{d}_{i,t}$ units of energy. Therefore, $\bar{d}_i$ denotes the subsistence consumption level of energy for an agent of type i, which we allow to be type-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.[35,36]

With an additional consumption good, we assume the planner uses an additional instrument to keep the tax system complete: it levies an excise tax $\tau_{D,t}$ on households' consumption of energy.[37] The budget constraint of agents of type i can, thus, be expressed as

$$\sum_{t=0}^{\infty} p_t N_t \Big(c_{i,t} + d_{i,t}(p_{E,t} + \tau_{D,t}) - (1 - \tau_{H,t})\, w_t e_i h_{i,t} \Big) \leq R_0 N_0 a_{i,0} + T. \tag{29}$$

To focus on demand heterogeneity, we assume—as in our baseline—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 and provide derivation details in Appendix C.

6.2.1 Optimal tax rules

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

Proposition 4 *If agents' utility is given by (28), the optimal pollution tax can be expressed as (22), i.e. a modified Pigouvian rule that accounts for the MCF given by*

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

with

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

From period 0, the over-time welfare-weighted average MCF is 1. If the IES is equal to 1, then $MCF_t = 1$ for all $t \geq 0$.

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

[36] 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 US (Cronin et al., 2019; Pizer and Sexton, 2019), and their implications for the design of tax reforms (Sallee, 2019).

[37] Without this instrument, there would be an additional constraint on the set of implementable allocations the planner has to satisfy. To prevent the optimal carbon tax from attempting to correct for this missing instrument, we assume the planner can also levy a tax on households' energy consumption.

The proof of Proposition 4 is provided in Appendix C.4. This result implies that the additional dirty good affects the optimal pollution tax formula only through the MCF: energy needs affect households' budget constraints, thereby affecting the planner's implementation cost over time. This mechanism might cause temporal fluctuations in the MCF, but it does not affect its long-term average value that remains equal to 1 as in the benchmark. In addition, as in Proposition 2 above, when the IES is equal to 1, the price and volume effects from an increase in aggregate consumption exactly offset each other, so households' total expenditures net of labour 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 (28), then the optimal tax on the polluting good is given by*

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

with

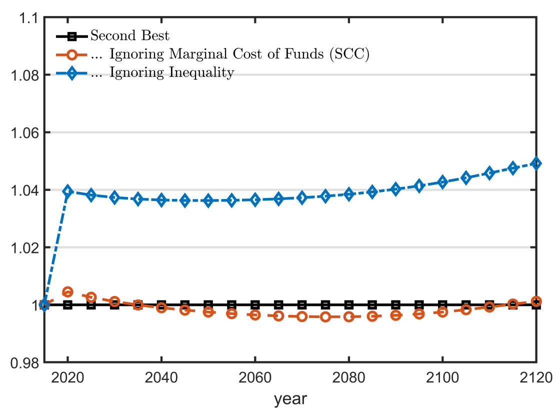
$$\Phi \equiv \sum_j \pi_j \frac{\lambda_j}{\varphi_j} + \Big(1 - (1+\epsilon+\gamma)(1-\sigma)\Big)\mathrm{cov}(\lambda_i/\varphi_i, \omega_i), \quad \Psi \equiv -\frac{\mathrm{cov}(\lambda_i/\varphi_i, e_i)}{\varsigma}, \quad \Lambda_t \equiv -\mathrm{cov}(\lambda_i/\varphi_i, \bar{d}_{i,t}).$$

The proof of Proposition 5 is provided in Appendix C.4. A corollary to this proposition is that, when preferences for the energy good are homogeneous, the optimal excise tax on this good is zero, as $\Lambda_t = 0$. We also show in Appendix C.4 that, in this case, the explicit formulas for labour, capital, and energy input taxes are unchanged relative to the benchmark model. Thus, although poor households spend a larger share of their budget on 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 modelled as a necessity generates a fixed cost to households' utility, which exacerbates inequalities.

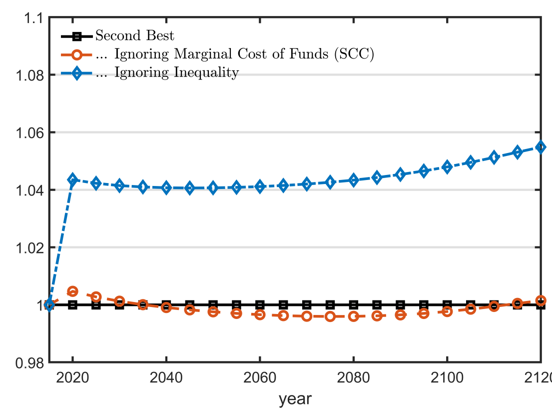
In the general case where preferences differ between agents, the energy good is subsidised, $\tau_D < 0$, if the agents who are valued relatively more by the planner compared to the market (higher λ_i/φ_i) have higher energy needs, $\bar{d}_{i,t}$. In this case, the *aggregate* Engel curves are non-linear, hence subsidies on necessities 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 subsidising the energy good. Heterogeneity in tastes thus calls for a deviation from uniform consumption taxation, as in Saez (2002). The sign and magnitude of this mechanism 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 weights λ_i/φ_i are generally higher for these agents. The excise tax will therefore be higher to the extent that more productive agents have, on average, higher *absolute* 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 costs that they incur. This horizontal heterogeneity will therefore drive the value of the excise tax downward. Our quantitative analysis below uses data on US households' energy consumption to illustrate the impact of these two sources of heterogeneity.

6.2.2 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 G.6, 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 Survey (CEX), where energy expenditures correspond 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 $\bar{d}_t$ to target an average energy expenditure share of 10.8% as observed in the CEX. We then use the value of ϵ to compute type-specific subsistence levels $\bar{d}_{i,t}$ to match the observed distribution of energy expenditure shares across and within income groups. 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. For additional details on the procedure and calibration choices, see Appendix G.6.



(a) Equal energy necessity



(b) Unequal energy necessity

Figure 6: Carbon Tax Decomposition With and Without Energy Necessity Inequality.

Notes: The black line represents the second-best carbon tax normalised 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 6 shows a decomposition similar to the one shown for the baseline in Figure 2b above, for the case where households have (a) 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 both 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 labour tax.

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.[38] For the same reason, optimal excise taxes on energy consumption are very small, amounting to about -0.4% of energy prices in every period—they are exactly zero when there is no heterogeneity in necessity levels.

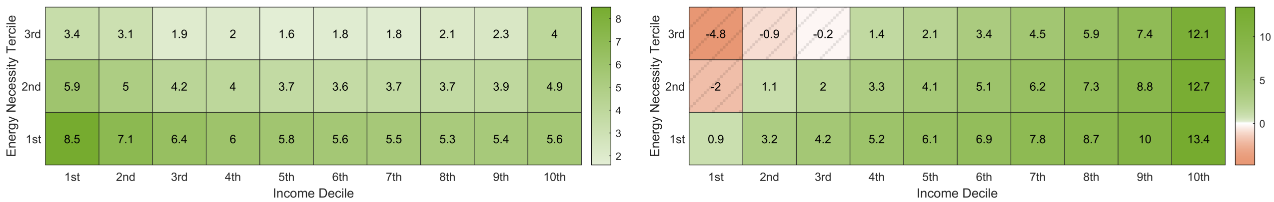
Figure 7a displays the inter-temporal welfare gains from carbon taxation for each category. Between income groups, we observe a U-shaped pattern: while the poorest households benefit relatively more from the increase in tax progressivity, the richest households benefit relatively more from future environmental 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 7b shows the welfare gains from setting carbon taxes to the SCC, ignoring inequalities, and using the additional revenue exclusively to reduce distortive labour taxes. As in the benchmark results presented in Section 5, the resulting distribution of welfare gains is highly regressive. These effects are magnified by the presence of energy necessity inequality. The poorest households with higher energy necessity actually lose significantly from the introduction of carbon taxes, even accounting for future climate mitigation benefits. These results further highlight the importance of adjusting carbon taxation for the presence of inequalities and combining it with the appropriate increase in the progressivity of income taxes.

6.3 Heterogeneous sensitivity to environmental damages

Several recent studies have highlighted that the impact of environmental degradation is heterogeneous across individuals, and is likely more negative for more financially deprived households (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; Cruz and Rossi-Hansberg, 2024). While heterogeneity in income, wealth, and consumption patterns are key to explaining the unequal burden from environ-

[38] This result seems to contrast with Cremer et al. (2003), who find in the context of France that optimal environmental taxes are significantly lower when preference heterogeneity is taken into account. In their model with four groups who differ in their income and tastes, preferences for energy consumption co-vary strongly with marginal utilities of consumption. Hence, there are large distributional benefits from setting lower environmental taxes. These effects are magnified when the planner is more averse to inequality and when environmental damages are lower. In contrast, we employ a more granular partition of households that captures heterogeneity both across and within income groups, allowing for greater flexibility in computing the covariance relevant for determining the optimal tax. Our findings indicate that the resulting correction term is negligible relative to the magnitude of the climate externality.



(a) Optimal policy (b) SCC and double dividend

Figure 7: Welfare Gains (%), Energy Necessity Inequality.

Notes: For each income decile and expenditure share tercile, the table shows the welfare gains, in percentage of consumption, of moving from a scenario without carbon taxation and otherwise optimal income taxes to: (a) optimal carbon and income taxation; (b) carbon taxes set to the SCC and extra revenue used to reduce labour income taxes.

mental 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 additively separable in Z_t and additionally consider households with heterogeneous sensitivity to environmental degradation, so that agent i's utility function can be expressed as

$$u_i\left(c_{i,t}, h_{i,t}, Z_t\right) \equiv \tilde{u}\left(c_{i,t}, h_{i,t}\right) + \hat{u}_i(Z_t). \tag{30}$$

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's health, exposure to conflicts, forced re-settlement, or losses in various forms of amenity values.[39]

When environmental degradation Z_t 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 disutility from environmental degradation for the planner—now depends on the joint distribution of utility damages and the planner's welfare weights,

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

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

Proposition 6 *If environmental utility is additively separable from consumption and leisure as in (30), 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 from pollution.*

Corollary 1 *If environmental utility is additively separable from consumption and leisure as in (30) 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 above-average marginal utility damages from pollution.*

[39] Although these types of damages may also affect households' productivity, we abstract from heterogeneous impacts of Z_t on agents' productivity e_i to keep the problem sufficiently tractable.

The proof of Proposition 6 is provided in Appendix D. As long as environmental welfare is additively separable from consumption and leisure, marginal utility losses from environmental damages 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 for each household is valued equally by the planner, hence the distribution of environmental damages is inconsequential for the planner when setting the optimal pollution tax. If 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 internalise pollution on that household only and sets it to a higher level if this household is more exposed.

Naturally, there are other mechanisms through which heterogeneous environmental damages could matter for the optimal carbon tax, even with a utilitarian planner. This could be the case, for instance, if accounting for damage heterogeneity raises the average marginal damage, if households have varying capacities to adapt to these damages, or under certain alternative non-separable preferences between consumption and the environment. These mechanisms are likely to operate mostly through the value of the SCC and are beyond the scope of this paper.[40]

7 Conclusion

Should environmental policies be less stringent in the presence of inequalities and redistributive taxation? 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 labour 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 tax distortions do not significantly affect carbon taxation when distortionary taxes are optimally chosen to provide redistribution: the optimal carbon tax is approximately Pigouvian. However, inequalities call for lower carbon taxes as the presence of poor households raises the marginal value of consumption, increasing the opportunity costs of abatement. Still, in our calibration to the US economy, this effect is quantitatively small, which is robust to reasonable variations in inequality, fiscal pressure, and severity of damages. The carbon tax is not used to address inequality directly but instead is affected by the residual inequality that remains after the planner has optimally set income taxes. 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

[40] Quantitatively, recent progress has been made by Cruz and Rossi-Hansberg (2024) in modelling spatially heterogeneous damages from climate change. Still, as explained in Hsiang et al. (2020), heterogeneous impacts at the individual level are difficult to measure as they are 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.

recycling increases the progressivity of the tax system, making poor households benefit even in the short run. By contrast, we show that following the well-known double-dividend policy—using the carbon tax revenue exclusively to reduce distortionary taxes—would generate strongly regressive effects.

Our paper presents numerous extensions to examine the implications of inequality for optimal carbon taxation. We analyse various policy scenarios and multiple sources of household heterogeneity, including unequal initial assets, energy demand, and sensitivity to environmental damages. Nevertheless, there are additional relevant aspects that warrant further investigation. Specifically, we have left for future research the role of risk—whether economic or climate-related—which could interact with inequalities and influence the optimal policy. Moreover, it would be interesting to further explore theoretically and quantitatively the role of heterogeneous damages of climate change.

References

AKERLOF, G., R. AUMANN, A. DEATON, P. DIAMOND, R. ENGLE, E. FAMA, L. HANSEN, O. HART, B. HOLMSTRÖM, D. KAHNEMAN, ET AL. (2019): "Economists statement on carbon dividends," *Wall Street Journal*, A13.

AUBERT, D. AND M. CHIROLEU-ASSOULINE (2019): "Environmental tax reform and income distribution with imperfect heterogeneous labour markets," *European Economic Review*, 116, 60–82.

BANZHAF, S., L. MA, AND C. TIMMINS (2019): "Environmental justice: The economics of race, place, and pollution," *Journal of Economic Perspectives*, 33, 185–208.

BARRAGE, L. (2020): "Optimal dynamic carbon taxes in a climate–economy model with distortionary fiscal policy," *The Review of Economic Studies*, 87, 1–39.

BARRAGE, L. AND W. NORDHAUS (2024): "Policies, projections, and the social cost of carbon: Results from the DICE-2023 model," *Proceedings of the National Academy of Sciences*, 121, e2312030121.

BELFIORI, M. E., D. CARROLL, AND S. HUR (2024): "Unequal Climate Policy in an Unequal World," *Globalization Institute Working Paper*.

BENMIR, G. AND J. ROMAN (2022): "The distributional costs of net-zero: A HANK perspective," *Work. Pap., Lond. Sch. Econ. Political Sci., London.*

BOURANY, T. (2024): "Optimal Energy Policy and the Inequality of Climate Change," Mimeo.

BOVENBERG, A. L. AND L. H. GOULDER (1996): "Optimal Environmental Taxation in the Presence of Other Taxes: General- Equilibrium Analyses," *American Economic Review*, 86, 985–1000.

BOVENBERG, L. AND R. DE MOOIJ (1994): "Environmental Levies and Distortionary Taxation," *American Economic Review*, 84, 1085–89.

BOVENBERG, L. AND F. VAN DER PLOEG (1994): "Environmental policy, public finance and the labour market in a second-best world," *Journal of Public Economics*, 55, 349–390.

CHAMLEY, C. (1986): "Optimal taxation of capital income in general equilibrium with infinite lives," *Econometrica*, 607–622.

CHARI, V. AND P. J. KEHOE (1999): "Chapter 26 Optimal fiscal and monetary policy," Elsevier, vol. 1 of *Handbook of Macroeconomics*, 1671–1745.

CHARI, V., J. P. NICOLINI, AND P. TELES (2020): "Optimal capital taxation revisited," *Journal of Monetary Economics*, 116, 147–165.

CHETTY, R., A. GUREN, D. MANOLI, AND A. WEBER (2011): "Are micro and macro labor supply elasticities consistent? A review of evidence on the intensive and extensive margins," *American Economic Review*, 101, 471–75.

Cremer, H. and F. Gahvari (2001): "Second-best taxation of emissions and polluting goods," *Journal of Public Economics*, 80, 169–197.

Cremer, H., F. Gahvari, and N. Ladoux (1998): "Externalities and optimal taxation," *Journal of Public Economics*, 70, 343–364.

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

Cronin, J. A., D. Fullerton, and S. Sexton (2019): "Vertical and Horizontal Redistributions from a Carbon Tax and Rebate," *Journal of the Association of Environmental and Resource Economists*, 6, 169–208.

Cruz, J.-L. and E. Rossi-Hansberg (2024): "The economic geography of global warming," *Review of Economic Studies*, 91, 899–939.

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

Diamond, P. A. and J. A. Mirrlees (1971): "Optimal taxation and public production I: Production efficiency," *American Economic Review*, 61, 8–27.

Dietz, S., F. van der Ploeg, A. Rezai, and F. Venmans (2021): "Are economists getting climate dynamics right and does it matter?" *Journal of the Association of Environmental and Resource Economists*, 8, 895–921.

Douenne, T., S. Dyrda, A. J. Hummel, and M. Pedroni (2025): "Optimal Climate Policy with Incomplete Markets," *CEPR Discussion Paper*.

Douenne, T. and A. Fabre (2022): "Yellow Vests, Pessimistic Beliefs, and Carbon Tax Aversion," *American Economic Journal: Economic Policy*, 14, 81–110.

Dyrda, S. and M. Pedroni (2023): "Optimal Fiscal Policy in a Model with Uninsurable Idiosyncratic Income Risk," *The Review of Economic Studies*, 90, 744–780.

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

Ferey, A., B. B. Lockwood, and D. Taubinsky (2024): "Sufficient statistics for nonlinear tax systems with general across-income heterogeneity," *American Economic Review*, 114, 3206–3249.

Fried, S., K. Novan, and W. Peterman (2018): "The Distributional Effects of a Carbon Tax on Current and Future Generations," *Review of Economic Dynamics*, 30, 30–46.

Fried, S., K. Novan, and W. B. Peterman (2024): "Understanding the inequality and welfare impacts of carbon tax policies," *Journal of the Association of Environmental and Resource Economists*, 11, S231–S260.

Fullerton, D. and H. Monti (2013): “Can pollution tax rebates protect low-wage earners?” *Journal of Environmental Economics and Management*, 66, 539–553.

Golosov, M., J. Hassler, P. Krusell, and A. Tsyvinski (2014): “Optimal Taxes on Fossil Fuel in General Equilibrium,” *Econometrica*, 82, 41–88.

Goulder, L. H. (1995): “Environmental taxation and the double dividend: a reader’s guide,” *International Tax and Public Finance*, 2, 157–183.

Goulder, L. H., M. A. Hafstead, G. Kim, and X. 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, J. and P. Krusell (2012): “Economics and climate change: integrated assessment in a multi-region world,” *Journal of the European Economic Association*, 10, 974–1000.

Heathcote, J., F. Perri, and G. L. Violante (2010): “Unequal we stand: An empirical analysis of economic inequality in the United States, 1967–2006,” *Review of Economic Dynamics*, 13, 15–51.

Heathcote, J. and H. Tsujiyama (2021): “Optimal Income Taxation: Mirrlees Meets Ramsey,” *Journal of Political Economy*, 129, 3141–3184.

Hsiang, S., P. Oliva, and R. Walker (2020): “The distribution of environmental damages,” *Review of Environmental Economics and Policy*, 13, 83–103.

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

Jacobs, B. and R. 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.

Jacobs, B. and F. van der Ploeg (2019): “Redistribution and pollution taxes with non-linear Engel curves,” *Journal of Environmental Economics and Management*, 95, 198–226.

Judd, K. L. (1985): “Redistributive taxation in a simple perfect foresight model,” *Journal of Public Economics*, 28, 59–83.

Känzig, D. R. (2023): “The unequal economic consequences of carbon pricing,” Tech. rep., National Bureau of Economic Research.

Kaplow, L. (2012): “Optimal Control of Externalities in the Presence of Income Taxation,” *International Economic Review*, 53, 487–509.

Klenert, D., G. Schwerhoff, O. Edenhofer, and L. Mattauch (2018): “Environmental Taxation, Inequality and Engels Law: The Double Dividend of Redistribution,” *Environmental & Resource Economics*, 71, 605–624.

Kotlikoff, L., F. Kubler, A. Polbin, J. Sachs, and S. Scheidegger (2021): "Making carbon taxation a generational win win," *International Economic Review*, 62, 3–46.

Kotlikoff, L., F. Kubler, A. Polbin, and S. Scheidegger (2024): "Can todays and tomorrows world uniformly gain from carbon taxation?" *European Economic Review*, 168, 104819.

Krusell, P. and A. A. Smith Jr (2022): "Climate change around the world," Mimeo.

Kuhn, M. and J.-V. 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, 2–73.

Kuhn, M. and L. Schlattmann (2024): *Distributional consequences of climate policies*, Centre for Economic Policy Research.

Le Grand, F., F. Oswald, X. Ragot, and A. Saussay (2022): "Fiscal policy for climate change," Mimeo.

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

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

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

Nordhaus, W. (2018): "Projections and uncertainties about climate change in an era of minimal climate policies," *American Economic Journal: Economic Policy*, 10, 333–60.

Nordhaus, W. D. (2017): "Revisiting the social cost of carbon," *Proceedings of the National Academy of Sciences*, 114, 1518–1523.

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

Piketty, T. and E. Saez (2013): "Optimal labor income taxation," in *Handbook of Public Economics*, Elsevier, vol. 5, 391–474.

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

Pirttilä, J. and M. Tuomala (1997): "Income tax, commodity tax and environmental policy," *International Tax and Public Finance*, 4, 379–393.

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

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

Ricke, K., L. Drouet, K. Caldeira, and M. Tavoni (2018): "Country-level social cost of carbon," *Nature Climate Change*, 8, 895–900.

Sachs, D., A. Tsyvinski, and N. Werquin (2020): "Nonlinear tax incidence and optimal taxation in general equilibrium," *Econometrica*, 88, 469–493.

Saez, E. (2002): "The desirability of commodity taxation under non-linear income taxation and heterogeneous tastes," *Journal of Public Economics*, 83, 217–230.

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

Sandmo, A. (1975): "Optimal Taxation in the Presence of Externalities," *Swedish Journal of Economics*, 77, 86–98.

Scheuer, F. and I. Werning (2017): "The taxation of superstars," *The Quarterly Journal of Economics*, 132, 211–270.

Storesletten, K., C. I. Telmer, and A. Yaron (2004): "Consumption and risk sharing over the life cycle," *Journal of monetary Economics*, 51, 609–633.

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

Trabandt, M. and H. Uhlig (2012): "How do Laffer curves differ across countries?" *Fiscal Policy After the Financial Crisis*, 211–249.

van der Ploeg, F., A. Rezai, and M. T. Reanos (2022): "Gathering support for green tax reform: Evidence from German household surveys," *European Economic Review*, 141, 103966.

Weitzman, M. L. (2009): "On modeling and interpreting the economics of catastrophic climate change," *Review of Economics and Statistics*, 91, 1–19.

Werning, I. (2007): "Optimal Fiscal Policy with Redistribution," *Quarterly Journal of Economics*, 122, 925–967.

Williams, R., H. Gordon, D. Burtraw, J. Carbone, and R. D. Morgenstern (2015): "The Initial Incidence of a Carbon Tax Across Income Groups," *National Tax Journal*, 68, 195–214.

Wöhrmüller, S. (2024): "Carbon taxation and precautionary savings," Mimeo.

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:

$$U(c_t, h_t, Z_t; \varphi) \equiv \max_{c_{i,t}, h_{i,t}} \sum_i \pi_i \varphi_i u(c_{i,t}, h_{i,t}, Z_t), \quad \text{s.t.} \sum_i \pi_i c_{i,t} = c_t, \quad \text{and} \quad \sum_i \pi_i e_i h_{i,t} = h_t. \tag{31}$$

The Lagrangian for this problem is

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

where θ_t^c and θ_t^h are Lagrange multipliers. Applying the envelope theorem to problem (31), 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 (31), 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

$$U_{c,t} = \varphi_i u_{c,i,t}, \tag{32}$$

$$U_{h,t} = \frac{\varphi_i u_{h,i,t}}{e_i}. \tag{33}$$

In any competitive equilibrium, these optimality conditions must hold for every agent i. Hence, using (32), (33), and agents' first-order conditions given by

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

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

we obtain

$$\frac{U_{h,t}}{U_{c,t}} = \frac{u_{h,i,t}}{u_{c,i,t} e_i} = -w_t (1 - \tau_{H,t}), \tag{34}$$

and

$$\frac{U_{c,t}}{U_{c,0}} = \frac{u_{c,i,t}}{u_{c,i,0}} = \frac{p_t}{\beta^t}. \tag{35}$$

Given the relationships above we can derive the implementability condition which relies only on the aggregates c_t, h_t, Z_t, and market weights φ. In this section we assume utility is additively separable in environmental damages, hence Z drops from the list of arguments. We turn to the non-separable case in Appendix F. Let $c_{i,t}^m(c_t, h_t; \varphi)$ and $h_{i,t}^m(c_t, h_t; \varphi)$ be the $\arg\max$ of problem (31). 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 (34) and (35) can be restated as

$$U_{c,0}\big(R_0 N_0 a_{i,0} + T\big) = \sum_{t=0}^{\infty} \beta^t N_t \Bigg(U_{c,t} c_{i,t}^m \big(c_t, h_t; \varphi\big) + U_{h,t} e_i h_{i,t}^m \big(c_t, h_t; \varphi\big) \Bigg), \quad \forall\ i. \tag{36}$$

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

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

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

$$\max_{\substack{\{C_t, H_{1,t}, H_{2,t}, K_{1,t}, K_{2,t}, \\ E_t, Z_t, \mu_t\}_{t=0}^{\infty}, T, \varphi, \tau_0^k}} \sum_t \beta^t N_t W\left(c_t, h_t, Z_t; \varphi, \theta, \lambda\right) - U_{c,0} \sum_i \pi_i \theta_i \left(R_0 N_0 a_{i,0} + T \right) \tag{37}$$

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 normalisation of market weights. Because this constraint is a simple normalisation, it has no impact on the resulting allocations.

In what follows, we assume that there is no initial wealth inequality, that is $a_{i,0} = a_0$ for every i. We relax this assumption in Appendix B.

A.2.2 First-order conditions

The first-order conditions are

$$[c_t] : W_{c,t} - \nu_{1,t} = 0, \quad \forall\, t \geq 0, \tag{38}$$

$$[H_{1,t}] : W_{h,t} + \nu_{1,t}\,(1 - D_t)\,A_{1,t}F_{H,t} = 0, \quad \forall\, t \geq 0, \tag{39}$$

$$[H_{2,t}] : W_{h,t} + \nu_{2,t}A_{2,t}G_{H,t} = 0, \quad \forall\, t \geq 0, \tag{40}$$

$$[K_{1,t+1}] : -\nu_{1,t} + [(1 - D_{t+1})\,A_{1,t+1}F_{K,t+1} + (1 - \delta)]\,\beta\nu_{1,t+1} = 0, \quad \forall\, t \geq 0, \tag{41}$$

$$[K_{2,t+1}] : -\nu_{1,t} + A_{2,t+1}G_{K,t+1}\beta\nu_{2,t+1} + (1 - \delta)\,\beta\nu_{1,t+1} = 0, \quad \forall\, t \geq 0, \tag{42}$$

$$[E_t] : -\nu_{1,t}\,(\Theta_{E,t} - (1 - D_t)\,A_{1,t}F_{E,t}) - \nu_{2,t} - \sum_{j=0}^{\infty}\beta^j\nu_{3,t+j}J_{E_t^M,t+j}\,(1 - \mu_t) = 0, \quad \forall\, t \geq 0, \tag{43}$$

$$[Z_t] : N_tW_{Z,t} - \nu_{1,t}D_t'A_{1,t}F_t + \nu_{3,t} = 0, \quad \forall\, t \geq 0, \tag{44}$$

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

$$[T] : \sum_i \pi_i\theta_i = 0. \tag{46}$$

For brevity, we have suppressed the terms associated with the derivatives of $U_{c,0}\sum_i \pi_i\theta_i\big(R_0N_0a_0 + T\big)$, as they vanish by equation (46) under no wealth inequality. Appendix B explains this result and its implications in detail, and Appendix B.2 presents the full first-order conditions including these terms. At $t = 0$, we also have that

$$[K_{1,0}] : [(1 - D_0)\,A_{1,0}F_{K,0} + (1 - \delta)]\,\nu_{1,0} - \kappa = 0, \tag{47}$$

$$[K_{2,0}] : A_{2,0}G_{K,0}\nu_{2,0} + (1 - \delta)\,\nu_{1,0} - \kappa = 0, \tag{48}$$

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 (39) and (40), 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 labour income taxes

From (38) and (39) we obtain

$$(1 - D(Z_t)) A_{1,t} F_{H,t} = -\frac{W_{h,t}}{W_{c,t}}, \quad \forall\, t \geq 0, \tag{49}$$

and using the intertemporal condition (41) we get

$$R^*_{t+1} \equiv 1 + r_{t+1} - \delta = \frac{1}{\beta}\frac{W_{c,t}}{W_{c,t+1}}, \quad \forall\, t \geq 0, \tag{50}$$

These two equations can be used to back out the optimal taxes on labour and capital income. Plugging (49) into (34) implies

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

which can be rearranged into

$$\tau_{H,t} = 1 - \frac{U_{h,t}}{U_{c,t}}\frac{W_{c,t}}{W_{h,t}}. \tag{51}$$

In any competitive equilibrium (35) 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 (50), it follows that

$$\frac{R_{t+1}}{R^*_{t+1}} = \frac{W_{c,t+1}}{W_{c,t}}\frac{U_{c,t}}{U_{c,t+1}}. \tag{52}$$

A.3.2 Excise taxes of energy and emissions

From the abatement first-order condition (45) and the energy firm abatement decision (9) we have that

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

From the climate variable first-order condition (44) 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

$$\tau_{E,t} = \frac{1}{\nu_{1,t}} \sum_{j=0}^{\infty} \beta^j \left(\nu_{1,t+j} D'_{t+j} A_{1,t+j} F_{t+j} - N_{t+j} W_{Z,t+j}\right) J_{E^M_t,t+j}. \tag{53}$$

From the energy first-order condition (43) we have that

$$-\nu_{1,t}\Big(\Theta_{E,t} + (1-\mu_t)\frac{\Theta_{\mu,t}}{E_t} - (1 - D(Z_t)) A_{1,t} F_{E,t}\Big) = \nu_{2,t}, \tag{54}$$

and combining the first-order conditions for sectoral labour supplies (39) and (40), 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 (54) 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

$$\tau_{I,t} = 0. \tag{55}$$

A.4 Explicit formulas

A.4.1 characterisation of equilibrium

Let us consider the following balanced-growth period utility function

$$u\left(c_i, h_i, Z\right) = \frac{\left(c_i (1 - \varsigma h_i)^{\gamma}\right)^{1-\sigma}}{1 - \sigma} + \hat{u}(Z). \tag{56}$$

To obtain explicit formulas, it is convenient to normalise 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 (56), the Lagrangian for the characterisation problem defined by (15) is

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

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

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

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

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

$$c_{i,t}^m (c_t, h_t; \varphi) = \omega_i c_t, \tag{57}$$

$$1-\varsigma h_{i,t}^m (c_t, h_t; \varphi) = \frac{\omega_i}{e_i}(1-\varsigma h_t), \tag{58}$$

where

$$\omega_i = \frac{\left(\varphi_i (e_i)^{\gamma(\sigma-1)}\right)^{\frac{1}{\sigma-(1-\sigma)\gamma}}}{\sum_j \pi_j \left(\varphi_j e_j^{\gamma(\sigma-1)}\right)^{\frac{1}{\sigma-(1-\sigma)\gamma}}} = \left(\varphi_i (e_i)^{\gamma(\sigma-1)}\right)^{\frac{1}{\sigma-(1-\sigma)\gamma}}. \tag{59}$$

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

$$\begin{aligned} U(c_t, h_t, Z_t, \varphi) &= \sum_j \pi_j \varphi_j \left(\frac{\omega_j^{1+\gamma}}{e_j^{\gamma}}\right)^{1-\sigma} \frac{\left(c_t (1-\varsigma h_t)^{\gamma}\right)^{1-\sigma}}{1-\sigma} + \sum_i \pi_i \varphi_i \hat{u}(Z) \\ &= \frac{\left(c_t (1-\varsigma h_t)^{\gamma}\right)^{1-\sigma}}{1-\sigma} + \Gamma \hat{u}(Z), \end{aligned} \tag{60}$$

since from the normalisation 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 (36), substituting the derivatives of $U(c_t, h_t, Z_t; \varphi)$ into the definition of $W(c_t, h_t, Z_t; \varphi, \theta, \lambda)$ we get

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

Collecting terms and simplifying we obtain

$$W(c_t, h_t, Z_t; \varphi, \theta, \lambda) = \Phi \frac{\left(c_t (1-\varsigma h_t)^{\gamma}\right)^{1-\sigma}}{1-\sigma} + \hat{u}(Z) + \Psi U_{h,t}. \tag{61}$$

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 (51) we get

$$\tau_{H,t} = \frac{\Psi\varsigma (1-\varsigma h_t)^{-1}}{\Phi + \Psi\varsigma (1-\gamma(1-\sigma))(1-\varsigma h_t)^{-1}}, \tag{62}$$

substituting the derivatives into (52) yields

$$\frac{R_{t+1}}{R^*_{t+1}} = \frac{\Phi - \Psi\varsigma\gamma(1-\sigma)(1-\varsigma h_{t+1})^{-1}}{\Phi - \Psi\varsigma\gamma(1-\sigma)(1-\varsigma h_t)^{-1}}. \tag{63}$$

and substituting the derivatives into (53) we get

$$\tau_{E,t} = \frac{1}{\nu_{1,t}} \sum_{j=0}^{\infty} \beta^j \Big(\nu_{1,t+j} D'_{t+j} A_{1,t+j} F_{t+j} - N_{t+j} V_Z(Z_{t+j}) \Big) J_{E^M_t, t+j}, \tag{64}$$

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

$$\nu_{1,t} = W_{c,t} = V_{c,t} + \sum_i \pi_i \theta_i \mathcal{I}_{c,i,t}. \tag{65}$$

If we add—without loss of generality—the normalisation 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

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

and therefore

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

with

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

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

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

and, therefore

$$\theta_i = \sum_j \frac{\pi_j \lambda_j}{\varphi_j} - \frac{\lambda_i}{\varphi_i}. \tag{66}$$

This allows us to rewrite

$$\begin{aligned}
\Phi &= \sum_i \pi_i \omega_i \left(\frac{\lambda_i}{\varphi_i} + (1-\sigma)(1+\gamma) \left(\sum_j \frac{\pi_j \lambda_j}{\varphi_j} - \frac{\lambda_i}{\varphi_i} \right) \right) \\
&= \sum_j \pi_j \frac{\lambda_j}{\varphi_j} + \big(1 - (1+\gamma)(1-\sigma)\big) \mathrm{cov}(\lambda_i/\varphi_i, \omega_i), \qquad (67) \\
\Psi &= \frac{1}{\varsigma} \sum_j \pi_j \frac{\lambda_j}{\varphi_j} (1 - e_j) = -\frac{\mathrm{cov}(\lambda_i/\varphi_i, e_i)}{\varsigma}, \qquad (68)
\end{aligned}$$

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

Notice that labour and capital income taxes are zero whenever $\Psi = 0$, which, according to equation (68), 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

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

with

$$M \equiv \sum_{t=0}^{\infty} \beta^t N_t \gamma \left(c_t \left(1 - \varsigma h_t\right)^{\gamma} \right)^{1-\sigma} \left(1 - \varsigma h_t\right)^{-1}. \tag{69}$$

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

$$\omega_i = 1 + \frac{U_{c,0} R_0 N_0 (a_{i,0} - A_0) + M(e_i - 1)}{(1-\sigma)(1+\gamma) \sum_{t=0}^{\infty} \beta^t N_t \tilde{U}(c_t, h_t)}, \quad \forall\ i. \tag{70}$$

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

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

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 individualised lump-sum transfers. All first-order conditions remain the same except for the one with respect to T given by (46): since we now have individualised instruments T_i, we obtain

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

hence for all t, $\sum_i \pi_i \theta_i \mathcal{I}_{c,i,t} = 0$. From (66), 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 (53), 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 (64), 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 (65) to substitute in the definition of the MCF, we also obtain equation (23) stated in Proposition 2,

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

With balanced-growth preferences, using the expression of $U(c_t, h_t, Z_t, \varphi)$ given by (60) and the solutions for $c^m_{i,t}$ and $h^m_{i,t}$ given by (57) and (58), we can show that $\mathcal{I}_{i,t}$ defined by (19) can be expressed as

$$\mathcal{I}_{i,t} = \Big(c_t(1-\varsigma h_t)^{\gamma}\Big)^{(1-\sigma)} \left(\omega_i + \gamma\Big(\omega_i - \frac{e_i}{(1-\varsigma h_t)}\Big)\right), \tag{72}$$

from which we obtain

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

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

$$\sum_i \pi_i\theta_i\mathcal{I}_{c,i,t} = (1-\sigma)\frac{\mathrm{cov}(\theta_i, \mathcal{I}_{i,t})}{c_t}. \tag{73}$$

This term is equal to 0 when either σ tends to 1 or θ_i and $\mathcal{I}_{i,t}$ 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 $\mathcal{I}_{i,t}$,

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

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 $\mathcal{I}_{i,t}$ as a constant $\kappa_{\mathcal{I}}$ that does not depend on agents' type,

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

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

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

with $\sum_i \pi_i\theta_i = 0$. ■

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

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

hence using the explicit expression of $U(c_t, h_t, Z_t, \varphi)$ given by (60) and taking derivatives we have

$$V_{c,t} = \sum_i \pi_i \omega_i \frac{\lambda_i}{\varphi_i} U_{c,t}, \tag{74}$$

from which we can show that

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

From (52), and using the fact that

$$\mathrm{MCF}_t = \frac{W_{c,t}}{V_{c,t}}, \tag{75}$$

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 internalised if the capital tax is optimally set to zero for all future periods where current emissions generate production damages.

Price and volume effects To understand the role of the IES, it is useful to go back to the origin of the term $\mathcal{I}_{i,t}$. This term comes from households' budget constraint (2) in which we have substituted for the price and real wage using (34) and (35). From these equations, it appears that when making more resources available to households, the price goes down since, in equilibrium,

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

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.

labour 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 (72), we have

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

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

$$\frac{\partial \frac{\mathcal{I}_{i,t}}{\mathcal{I}_{j,t}}}{\partial h_t} = \frac{\varsigma\gamma(1+\gamma)\big(e_j\omega_i - e_i\omega_j\big)}{(1-\varsigma h_t)^2\Big(\omega_j(1+\gamma) - \frac{\gamma e_j}{(1-\varsigma h_t)}\Big)^2}. \tag{76}$$

From (70), 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 (76) is negative. This result means that when h_t is high relative to its average value, the relative labour 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 $\text{cov}(\theta_i, \mathcal{I}_{i,t}) < 0$. 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 labour supply is relatively high compared to its long-run value.

Proof of Proposition 3: From our characterisation 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 (74) 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

$$\begin{aligned} V_{c,t} &= \sum_i \pi_i \lambda_i \frac{u_{c,i,t} c_{i,t}}{c_t} \\ &= \sum_i \pi_i \lambda_i u_{c,i,t} + \text{cov}\Big(\lambda_i u_{c,i,t}, \frac{c_{i,t}}{c_t}\Big). \end{aligned} \tag{77}$$

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 (59), when σ tends to 1 we have $\omega_i = \varphi_i$, hence from (74) and using the normalisation 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 with initial wealth inequality

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

B.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 (37)) 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-optimise 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 was 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.

B.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

$$[c_0]: W_{c,0} - \nu_{1,0} - U_{cc,0} \sum_i \pi_i \theta_i R_0 a_{i,0} = 0, \tag{78}$$

$$[H_{1,0}]: W_{h,0} + \nu_{1,0} (1 - D_0) A_{1,0} F_{H,0} - U_{ch,0} \sum_i \pi_i \theta_i R_0 a_{i,0} \tag{79}$$

$$- N_0 U_{c,0} \sum_i \pi_i \theta_i a_{i,0} (1 - \tau_{K,0}) (1 - D_0) A_{1,0} F_{KH,0} + \Gamma_0 (F_{HH,0} G_{K,0} - F_{KH,0} G_{H,0}) = 0,$$

$$[H_{2,0}]: W_{H,0} + \nu_{2,0} A_{2,0} G_{H,0} - U_{ch,0} \sum_i \pi_i \theta_i R_0 a_{i,0} + \Gamma_0 (F_{H,0} G_{KH,0} - F_{K,0} G_{HH,0}) = 0, \tag{80}$$

$$[K_{1,0}]: ((1 - D_0) A_{1,0} F_{K,0} + (1 - \delta)) \nu_{1,0} - N_0 U_{c,0} \sum_i \pi_i \theta_i a_{i,0} (1 - \tau_{K,0}) (1 - D_0) A_{1,0} F_{KK,0} - \kappa$$

$$+ \Gamma_0 (F_{HK,0} G_{K,0} - F_{KK,0} G_{H,0}) = 0, \tag{81}$$

$$[K_{2,0}]: A_{2,0} G_{K,0} \nu_{2,0} + (1 - \delta) \nu_{1,0} - \kappa + \Gamma_0 (F_{H,0} G_{KK,0} - F_{K,0} G_{HK,0}) = 0, \tag{82}$$

$$[E_0]: - (\Theta_{E,0} - (1 - D_0) A_{1,0} F_{E,0}) \nu_{1,0} - \nu_{2,0} - \sum_{j=0}^{\infty} \beta^j \nu_{3,j} J_{E_0^M,j} (1 - \mu_0) \tag{83}$$

$$- N_0 U_{c,0} \sum_i \pi_i \theta_i a_{i,0} (1 - \tau_{K,0}) (1 - D_0) A_{1,0} F_{KE,0} + \Gamma_0 (F_{HE,0} G_{K,0} - F_{KE,0} G_{H,0}) = 0,$$

$$[Z_0]: N_0 W_{Z,0} - \nu_{1,0} D_0' A_{1,0} F_0 + \nu_{3,0} + N_0 U_{c,0} \sum_i \pi_i \theta_i a_{i,0} (1 - \tau_{K,0}) D_0' A_{1,0} F_{K,0} = 0, \tag{84}$$

$$[\mu_0]: -\nu_{1,0} \Theta_{\mu,0} + \sum_{j=0}^{\infty} \beta^j \nu_{3,j} J_{E_0^M,j} E_0 = 0. \tag{85}$$

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

From (81) and (82), it follows that

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

From (79) and (80), it follows that

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

Hence, putting these two equations together, we obtain

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

B.4 labour income taxes

From (79) and (78) we obtain

$$(1-D_0)\,A_{1,0}F_{H,0}=\frac{\left\{\begin{array}{c}-W_{h,0}+U_{ch,0}\sum_i\pi_i\theta_iR_0a_{i,0}\\ +N_0U_{c,0}\sum_i\pi_i\theta_ia_{i,0}\,(1-\tau_{K,0})\,(1-D_0)\,A_{1,0}F_{KH,0}-\Gamma_0\,(F_{HH,0}G_{K,0}-F_{KH,0}G_{H,0})\end{array}\right\}}{W_{c,0}-U_{cc,0}\sum_i\pi_i\theta_iR_0a_{i,0}} \tag{86}$$

Plugging (86) into (34) implies

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

which can be rearranged into

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

B.5 Capital income taxes

From (41) and (78) we obtain

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

In any competitive equilibrium (35) holds, which implies

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

Substituting this into (50), it follows that

$$\frac{R_1}{R_1^*}=\frac{W_{c,1}}{W_{c,0}-U_{cc,0}\sum_i\pi_i\theta_iR_0a_{i,0}}\frac{U_{c,0}}{U_{c,1}}. \tag{87}$$

B.6 Excise taxes of energy and emissions

From (9) and the abatement first-order condition (85) we have that

$$\tau_{E,0}=\frac{\Theta_{\mu,0}}{E_0}=\frac{1}{\nu_{1,0}}\sum_{j=0}^{\infty}\beta^j\nu_{3,j}J_{E_0^M,j}. \tag{88}$$

From the climate variable first-order condition (84) we have that

$$\nu_{3,0}=\nu_{1,0}D_0'A_{1,0}F_0-N_0W_{Z,0}-N_0U_{c,0}\sum_i\pi_i\theta_ia_{i,0}\,(1-\tau_{K,0})\,D_0'A_{1,0}F_{K,0}. \tag{89}$$

Substituting (89) into (88) 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 (83) 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 labour supplies (79) and (80), 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

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

C Optimal tax rules with energy consumption inequality

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.

C.1 characterisation of equilibrium

Let $\varphi \equiv \{\varphi_i\}$ be the market weights normalised 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:

$$\begin{aligned} U\left(c_t, d_t, h_t, Z_t; \varphi\right) &\equiv \max_{c_{i,t}, d_{i,t}, h_{i,t}} \sum_i \pi_i \varphi_i u_i\left(c_{i,t}, d_{i,t}, h_{i,t}, Z_t\right), \\ \text{s.t. } \sum_i \pi_i c_{i,t} = c_t, \quad &\text{and} \quad \sum_i \pi_i d_{i,t} = d_t, \quad \text{and} \quad \sum_i \pi_i e_i h_{i,t} = h_t. \end{aligned} \tag{90}$$

Using the utility function defined by equation (28) 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

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

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

C.2 Implementability condition

From the first-order conditions of problem (90) 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

$$\frac{U_{h,t}}{U_{c,t}} = \frac{u_{h,i,t}}{u_{c,i,t} e_{i,t}} = -w_t\left(1 - \tau_{H,t}\right), \tag{91}$$

$$\frac{U_{d,t}}{U_{c,t}} = \frac{u_{d,i,t}}{u_{c,i,t}} = p_{E,t} + \tau_{D,t}, \tag{92}$$

and

$$\frac{U_{c,t}}{U_{c,0}} = \frac{u_{c,i,t}}{u_{c,i,0}} = \frac{p_t}{\beta^t}. \tag{93}$$

Using (91), (92), and (93) to substitute in households' budget constraint (29), we obtain the implementability conditions

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

C.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

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

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.

C.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

$$\tau_{H,t} = 1 - \frac{U_{h,t}}{U_{c,t}} \frac{W_{c,t}}{W_{h,t}}, \tag{94}$$

$$\frac{R_{t+1}}{R^*_{t+1}} = \frac{W_{c,t+1}}{W_{c,t}} \frac{U_{c,t}}{U_{c,t+1}}, \tag{95}$$

$$\tau_{E,t} = \frac{1}{\nu_{1,t}} \sum_{j=0}^{\infty} \beta^j \left(\nu_{1,t+j} D'_{t+j} A_{1,t+j} F_{t+j} - N_{t+j} W_{Z,t+j} \right) J_{E_t^M,t+j}, \tag{96}$$

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

$$\tau_{D,t} = \frac{U_{d,t}}{U_{c,t}} - \frac{W_{d,t}}{W_{c,t}}. \tag{97}$$

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

$$\mathcal{I}_i(c_t, d_t, h_t, \varphi) \equiv U_{c,t} c^m_{i,t}\big(c_t, d_t, h_t; \varphi\big) + U_{d,t} d^m_{i,t}\big(c_t, d_t, h_t; \varphi\big) + U_{h,t} e_{i,t} h^m_{i,t}\big(c_t, d_t, h_t; \varphi\big), \tag{98}$$

we can express the MCF as

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

and we can re-write the optimal pollution tax given by (96) as

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

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

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

With balanced-growth preferences, substituting into (98) we obtain

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

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

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

where

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

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 (97), 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 (94) and (95), 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 ($\sum_i \pi_i \theta_i \mathcal{I}_{c,i,t}$) 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

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

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

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

Thus, the sign of the marginal implementation cost depends on a price effect through σ, and on an energy demand and labour supply effects from $\text{cov}(\theta_i, \mathcal{I}_{i,t})$. 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.

D Optimal tax rules with heterogeneous climate damages

Proof of Proposition 6: When utility is additively separable in consumption-leisure and environmental quality, the benchmark model presented in Appendix A can easily be generalised 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

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

The only difference with the benchmark model is the expression of $V_{Z,t}$, the marginal disutility 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.

E Optimal tax rules with third-best policies

In our baseline Ramsey problem, the government faces the constraint that 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 labour or capital income taxes. In this appendix, we consider the implications of introducing these constraints. We first summarise the theoretical results with fixed labour and then capital income taxes in Appendices E.1 and E.2. Appendix E.3 presents a quantitative assessment, and detailed derivations are presented in Appendices E.4 and E.5.

E.1 Fixed labour income tax

Let us assume that the planner cannot choose the labour income tax, which 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

$$\frac{U_{h,t}}{U_{c,t}} = -\left(1 - \bar{\tau}_H\right)\left(1 - D_t\right) A_{1,t} F_{H,t}, \tag{99}$$

which pins down the wedge between the marginal rate of substitution between consumption and leisure and the marginal product of labour. For a given value of $\bar{\tau}_H$, equation (99) puts a restriction on the implementable allocations that the planner must satisfy. Let $\beta^t \Lambda_t^H$ denote the multiplier on the constraint (99). 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 labour income tax is fixed at a sub-optimally high (resp. low) level. With the additional constraint (99) in each period t, the expression for the optimal pollution tax becomes (see Appendix E.4 for more details)[41]

$$\tau_{E,t} = \frac{1}{\nu_{1,t}} \sum_{j=0}^{\infty} \beta^j \left(\nu_{1,t+j} D'_{t+j} A_{1,t+j} F_{t+j} - N_{t+j} V_{Z,t+j} + \Lambda_{t+j}^H \left(1 - \bar{\tau}_H\right) D'_{t+j} A_{1,t+j} F_{H,t+j}\right) J_{E_t^M, t+j}, \tag{100}$$

where $\nu_{1,t}$ 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 term in parentheses, which Barrage (2020) refers to as the *fiscal interaction* term. The impact on optimal carbon taxes is in general ambiguous and reflects another reason for deviating from the Pigouvian tax rule. If τ_H is fixed at a level below the second-best labour tax, for instance, it is beneficial to reduce carbon taxes and allow climate damages to reduce labour productivity bringing the after-tax wage closer to the optimal level that would be reached without the constraint.

[41] Without constraint (99), it is optimal to equalize the marginal rate of technical substitution between capital and labour across both sectors: the government does not wish to distort production decisions. In the third best, with constraint (99), this is no longer the case, and it is optimal to deviate from zero excise energy taxes, $\tau_{I,t}$.

E.2 Fixed 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 $t \geq 0$,

$$\frac{U_{c,t}}{U_{c,t+1}} = \beta\Big(1 + (1 - \bar{\tau}_K)\left((1 - D_{t+1})\, A_{1,t+1} F_{K,t+1} - \delta\right)\Big), \tag{101}$$

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 labour income tax, equation (101) 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 if the capital income tax rate is fixed at a sub-optimally high level so that raising $\bar{\tau}_K$ in a particular period lowers welfare. With the additional constraint (101), the expression for the optimal pollution tax is modified to (see Appendix E.5 for more details)

$$\tau_{E,t} = \frac{1}{\nu_{1,t}} \sum_{j=0}^{\infty} \beta^j \left(\nu_{1,t+j} D'_{t+j} A_{1,t+j} F_{t+j} - N_{t+j} W_{Z,t+j} + \Lambda^K_{t+j}\,(1 - \bar{\tau}_K)\, D'_{t+j} A_{1,t+j} F_{K,t+j}\right) J_{E^M_t,t+j}, \tag{102}$$

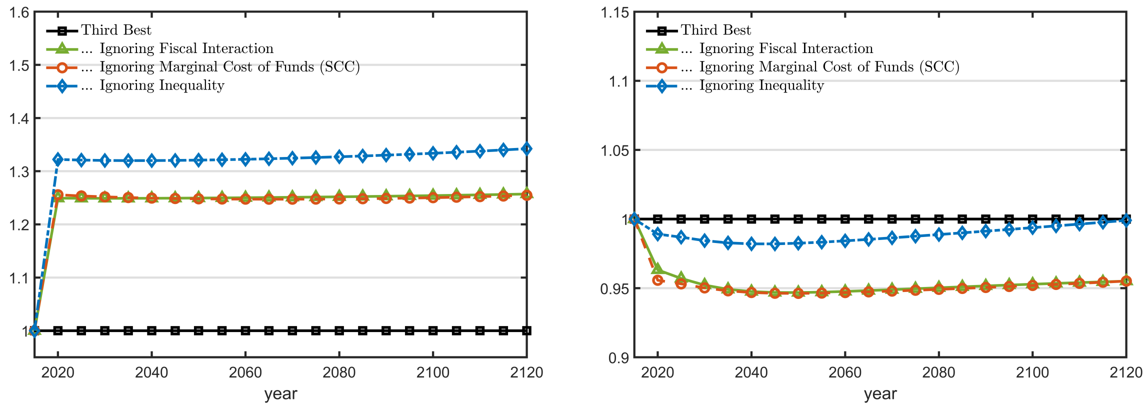
where again the last component captures the fiscal interaction term. The intuition is similar to before. If τ_K is fixed at a sub-optimally high level, for instance, it is beneficial to increase carbon taxes and further reduce climate damages to increase the marginal product of capital and bring the after-tax interest rate closer to the optimal level.

E.3 Quantitative analysis

Figure 8 below compares the third-best pollution tax normalised 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).[42] 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 labour 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, as can be seen from the small difference between the green and red lines in both panels, fiscal interactions now drive the carbon tax away from its Pigouvian level through the additional constraints that arise in the third-best environment. Interestingly, the fiscal interaction term lowers the optimal carbon tax when the labour 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 labour and the marginal product of capital and hence, the before-tax wage and interest rate. A higher before-tax wage, in turn, lowers

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



(a) Given labour income tax

(b) Given capital income tax

Figure 8: Third-Best Carbon Tax Decomposition.

Notes: The black line represents the second-best carbon tax normalised 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 (100) and for (b) the last term in equation (102). As in Figure 2b, 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.

welfare because the labour 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 labour 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 labour income tax is fixed.

Appendix H.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 labour income tax in our baseline scenario, this is not the case with additional constraints on instruments. In particular, when the government is forced to redistribute "too little" because labour income taxes are set below the optimum, the carbon tax revenue is mostly targeted toward redistribution, leading to more progressive effects.

E.4 Optimal tax rules with fixed labour income taxes

Suppose that labour 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,

$$\frac{U_{h,t}}{U_{c,t}} = -\left(1-\bar{\tau}_H\right)\left(1-D_t\right)A_{1,t}F_{H,t}, \tag{103}$$

$$F_{H,t}G_{K,t} = F_{K,t}G_{H,t}. \tag{104}$$

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

$$[c_t] : W_{c,t} - \nu_{1,t} + \Lambda_t^H \vartheta_{c,t} = 0, \quad \forall\, t \geq 0, \tag{105}$$

$$[H_{1,t}] : W_{h,t} + \nu_{1,t} \left(1 - D_t\right) A_{1,t} F_{H,t} + \Lambda_t^H \Big(\vartheta_{h,t} + \left(1 - \bar{\tau}_H\right) \left(1 - D_t\right) A_{1,t} F_{HH,t} \Big)$$
$$+ \Gamma_t^H \Big(F_{HH,t} G_{K,t} - F_{KH,t} G_{H,t} \Big) = 0, \quad \forall\, t \geq 0, \tag{106}$$

$$[H_{2,t}] : W_{h,t} + \nu_{2,t} A_{2,t} G_{H,t} + \Lambda_t^H \vartheta_{h,t} + \Gamma_t^H \left(F_{H,t} G_{KH,t} - F_{K,t} G_{HH,t} \right) = 0, \quad \forall\, t \geq 0, \tag{107}$$

$$[K_{1,t+1}] : -\nu_{1,t} + \Big(\left(1 - D_{t+1}\right) A_{1,t+1} F_{K,t+1} + \left(1 - \delta\right) \Big) \beta \nu_{1,t+1} + \beta \Lambda_{t+1}^H \Big(\left(1 - \bar{\tau}_H\right) \left(1 - D_{t+1}\right) A_{1,t+1} F_{HK,t+1} \Big)$$
$$+ \beta \Gamma_{t+1}^H \Big(F_{HK,t+1} G_{K,t+1} - F_{KK,t+1} G_{H,t+1} \Big) = 0, \quad \forall\, t \geq 0, \tag{108}$$

$$[K_{2,t+1}] : -\nu_{1,t} + A_{2,t+1} G_{K,t+1} \beta \nu_{2,t+1} + \left(1 - \delta\right) \beta \nu_{1,t+1}$$
$$+ \beta \Gamma_{t+1}^H \left(F_{H,t+1} G_{KK,t+1} - F_{K,t+1} G_{HK,t+1} \right) = 0, \quad \forall\, t \geq 0, \tag{109}$$

$$[E_t] : -\nu_{1,t} \left(\Theta_{E,t} - \left(1 - D_t\right) A_{1,t} F_{E,t} \right) - \nu_{2,t} - \sum_{j=0}^{\infty} \beta^j \nu_{3,t+j} J_{E_t^M,t+j} \left(1 - \mu_t\right)$$
$$+ \Lambda_t^H \left(\left(1 - \bar{\tau}_H\right) \left(1 - D_t\right) A_{1,t} F_{HE,t} \right) + \Gamma_t^H \left(F_{HE,t} G_{K,t} - F_{KE,t} G_{H,t} \right) = 0, \quad \forall\, t \geq 0, \tag{110}$$

$$[Z_t] : N_t W_{Z,t} - \nu_{1,t} D_t' A_{1,t} F_t + \nu_{3,t} - \Lambda_t^H \left(1 - \bar{\tau}_H\right) D_t' A_{1,t} F_{H,t} = 0, \quad \forall\, t \geq 0, \tag{111}$$

$$[\mu_t] : -\nu_{1,t} \Theta_{\mu,t} + \sum_{j=0}^{\infty} \beta^j \nu_{3,t+j} J_{E_t^M,t+j} E_t = 0, \quad \forall\, t \geq 0, \tag{112}$$

$$[T] : \sum_i \pi_i \theta_i = 0, \tag{113}$$

$$[\varphi_i] : \sum_t \beta^t N_t W_{\varphi_i,t} - \frac{\zeta}{\sigma - \left(1 - \sigma\right) \gamma} \frac{\pi_i \omega_i}{\varphi_i} = 0, \tag{114}$$

where

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

E.4.1 Capital income taxes and multipliers on new constraints

From (106) and (105) we obtain

$$(1 - D_t) A_{1,t} F_{H,t} = - \frac{W_{h,t} + \Lambda_t^H \Big(\vartheta_{h,t} + \left(1 - \bar{\tau}_H\right) \left(1 - D_t\right) A_{1,t} F_{HH,t} \Big) + \Gamma_t^H \Big(G_{K,t} F_{HH,t} - G_{H,t} F_{KH,t} \Big)}{W_{c,t} + \Lambda_t^H \vartheta_{c,t}}, \quad \forall\, t \geq 0, \tag{115}$$

and using the intertemporal condition (108) we get

$$R^*_{t+1} \equiv 1 + r_{t+1} - \delta$$
$$= \frac{1}{\beta} \frac{\left\{ \begin{array}{c} W_{c,t} + \Lambda_t \vartheta_{c,t} - \beta \Lambda^H_{t+1} \Big((1-\bar{\tau}_H)(1-D_{t+1}) A_{1,t+1} F_{HK,t+1} \Big) \\ -\beta \Gamma_{t+1} \left(F_{HK,t+1} G_{K,t+1} - F_{KK,t+1} G_{H,t+1} \right) \end{array} \right\}}{W_{c,t+1} + \Lambda^H_{t+1} \vartheta_{c,t+1}}. \quad \forall\, t \geq 0. \tag{116}$$

Solving (106) and (107), and (108) and (109) for $\nu_{2,t}/\nu_{1,t}$, and equating both equations, using (104), yields

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

where

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

Combining this equation with (115) we can then solve for

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

In any competitive equilibrium (35) 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 (116), it follows that

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

E.4.2 Excise taxes of energy and emissions

From (9) and the abatement first-order condition (112) we have that

$$\tau_{E,t} = \frac{\Theta_{\mu,t}}{E_t} = \frac{1}{\nu_{1,t}} \sum_{j=0}^{\infty} \beta^j \nu_{3,t+j} J_{E^M_t,t+j}. \tag{117}$$

From the climate variable first-order condition (111) we have that

$$\nu_{3,t} = \nu_{1,t} D'_t A_{1,t} F_t - W_{Z,t} + \Lambda^H_t (1-\bar{\tau}_H) D'_t A_{1,t} F_{H,t}. \tag{118}$$

Substituting (118) into (117) we obtain the optimal pollution tax

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

From the energy first-order condition (110) 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 labour supplies (106) and (107), 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

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

E.5 Optimal tax rules with fixed capital income taxes

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

$$\frac{U_{c,t}}{U_{c,t+1}}=\beta\Big(1+(1-\bar{\tau}_K)\,((1-D_{t+1})\,A_{1,t+1}F_{K,t+1}-\delta)\Big), \tag{119}$$

$$F_{H,t}G_{K,t}=F_{K,t}G_{H,t}. \tag{120}$$

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

$$[c_t] : W_{c,t} - \nu_{1,t} - \frac{\Lambda_{t+1}^K}{N_t} \frac{U_{cc,t}}{U_{c,t+1}} + \frac{\Lambda_t^K}{N_t} \frac{U_{c,t-1} U_{cc,t}}{\beta U_{c,t}^2} = 0, \quad \forall\ t \geq 0, \tag{121}$$

$$\begin{aligned}[H_{1,t}] : W_{h,t} &+ \nu_{1,t} \left(1 - D_t\right) A_{1,t} F_{H,t} - \frac{\Lambda_{t+1}^K}{N_t} \frac{U_{ch,t}}{U_{c,t+1}} + \frac{\Lambda_t^K}{N_t} \frac{U_{c,t-1} U_{ch,t}}{\beta U_{c,t}^2} \\ &+ \Lambda_t^K \left(1 - \bar{\tau}_K\right) \left(1 - D_t\right) A_{1,t} F_{KH,t} + \Gamma_t^K \Big(F_{HH,t} G_{K,t} - F_{KH,t} G_{H,t} \Big) = 0, \quad \forall\ t \geq 0,\end{aligned} \tag{122}$$

$$\begin{aligned}[H_{2,t}] : W_{h,t} &+ \nu_{2,t} A_{2,t} G_{H,t} - \frac{\Lambda_{t+1}^K}{N_t} \frac{U_{ch,t}}{U_{c,t+1}} + \frac{\Lambda_t^K}{N_t} \frac{U_{c,t-1} U_{ch,t}}{\beta U_{c,t}^2} \\ &+ \Gamma_t^K \Big(F_{H,t} G_{KH,t} - F_{K,t} G_{HH,t} \Big) = 0, \quad \forall\ t \geq 0,\end{aligned} \tag{123}$$

$$\begin{aligned}[K_{1,t+1}] : -\nu_{1,t} &+ \Big(\left(1 - D_{t+1}\right) A_{1,t+1} F_{K,t+1} + (1 - \delta) \Big) \beta \nu_{1,t+1} + \beta \Lambda_{t+1}^K \Big(\left(1 - \bar{\tau}_K\right) \left(1 - D_{t+1}\right) A_{1,t+1} F_{KK,t+1} \Big) \\ &+ \beta \Gamma_{t+1}^K \Big(F_{HK,t+1} G_{K,t+1} - F_{KK,t+1} G_{H,t+1} \Big) = 0, \quad \forall\ t \geq 0,\end{aligned} \tag{124}$$

$$\begin{aligned}[K_{2,t+1}] : -\nu_{1,t} &+ A_{2,t+1} G_{K,t+1} \beta \nu_{2,t+1} + (1 - \delta) \beta \nu_{1,t+1} \\ &+ \beta \Gamma_{t+1}^K \Big(F_{H,t+1} G_{KK,t+1} - F_{K,t+1} G_{HK,t+1} \Big) = 0, \quad \forall\ t \geq 0,\end{aligned} \tag{125}$$

$$\begin{aligned}[E_t] : -\nu_{1,t} &\left(\Theta_{E,t} - \left(1 - D_t\right) A_{1,t} F_{E,t} \right) - \nu_{2,t} - \sum_{j=0}^{\infty} \beta^j \nu_{3,t+j} J_{E_t^M, t+j} \left(1 - \mu_t\right) \\ &+ \Lambda_t^K \Big(\left(1 - \bar{\tau}_K\right) \left(1 - D_t\right) A_{1,t} F_{KE,t} \Big) + \Gamma_t^K \Big(F_{HE,t} G_{K,t} - F_{KE,t} G_{H,t} \Big) = 0, \quad \forall\ t \geq 0,\end{aligned} \tag{126}$$

$$[Z_t] : N_t W_{Z,t} - \nu_{1,t} D_t' A_{1,t} F_t + \nu_{3,t} - \Lambda_t^K \Big(\left(1 - \bar{\tau}_K\right) D_t' A_{1,t} F_{K,t} \Big) = 0, \quad \forall\ t \geq 0, \tag{127}$$

$$[\mu_t] : -\nu_{1,t} \Theta_{\mu,t} + \sum_{j=0}^{\infty} \beta^j \nu_{3,t+j} J_{E_t^M, t+j} E_t = 0, \quad \forall\ t \geq 0, \tag{128}$$

$$[T] : \sum_i \pi_i \theta_i = 0, \tag{129}$$

$$[\varphi_i] : \sum_t \beta^t W_{\varphi_i,t} - \frac{\zeta}{\sigma - (1 - \sigma) \gamma} \frac{\pi_i \omega_i}{\varphi_i} = 0. \tag{130}$$

E.5.1 Capital income taxes and multipliers on new constraints

From(122) and (121) we obtain

$$\left(1 - D_t\right) A_{1,t} F_{H,t} = \frac{- \left\{ \begin{array}{c} W_{h,t} - \frac{\Lambda_{t+1}^K}{N_t} \frac{U_{ch,t}}{U_{c,t+1}} + \frac{\Lambda_t^K}{N_t} \frac{U_{c,t-1} U_{ch,t}}{\beta U_{c,t}^2} + \Lambda_t^K \left(1 - \bar{\tau}_K\right) \left(1 - D_t\right) A_{1,t} F_{KH,t} \\ + \Gamma_t^K \left(F_{HH,t} G_{K,t} - F_{KH,t} G_{H,t} \right) \end{array} \right\}}{W_{c,t} - \frac{\Lambda_{t+1}^K}{N_t} \frac{U_{cc,t}}{U_{c,t+1}} + \frac{\Lambda_t^K}{N_t} \frac{U_{c,t-1} U_{cc,t}}{\beta U_{c,t}^2}}, \tag{131}$$

and using the intertemporal condition (124) we get

$$
\begin{aligned}
R^*_{t+1} \equiv & 1 + r_{t+1} - \delta \\
= & \frac{1}{\beta} \frac{\left\{ \begin{array}{c} W_{c,t} - \frac{\Lambda^K_{t+1}}{N_t} \frac{U_{cc,t}}{U_{c,t+1}} + \frac{\Lambda^K_t}{N_t} \frac{U_{c,t-1} U_{cc,t}}{\beta U^2_{c,t}} - \beta \Lambda^K_{t+1} \left(1 - \bar{\tau}_K\right) \left(1 - D_{t+1}\right) A_{1,t+1} F_{KK,t+1} \\ -\beta \Gamma^K_{t+1} \left(F_{HK,t+1} G_{K,t+1} - F_{KK,t+1} G_{H,t+1}\right) \end{array} \right\}}{W_{c,t+1} - \frac{\Lambda^K_{t+2}}{N_{t+1}} \frac{U_{cc,t+1}}{U_{c,t+2}} + \frac{\Lambda^K_{t+1}}{N_{t+1}} \frac{U_{c,t} U_{cc,t+1}}{\beta U^2_{c,t+1}}}. \qquad (132)
\end{aligned}
$$

Solving (106) and (107), and (108) and (109) for $\nu_{2,t}/\nu_{1,t}$, and equating both equations, using (120), yields

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

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 (35) 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 (132), it follows that

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

Plugging (131) into (34) implies

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

which can be rearranged into

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

E.5.2 Excise taxes of energy and emissions

From (9) and the abatement first-order condition (128) we have that

$$\tau_{E,t} = \frac{\Theta_{\mu,t}}{E_t} = \frac{1}{\nu_{1,t}} \sum_{j=0}^{\infty} \beta^j \nu_{3,t+j} J_{E_t^M,t+j}. \tag{133}$$

From the climate variable first-order condition (127) we have that

$$\nu_{3,t} = \nu_{1,t} D_t' A_{1,t} F_t - N_t W_{Z,t} + \Lambda_t^K \left(1 - \bar{\tau}_K\right) D_t' A_{1,t} F_{K,t}. \tag{134}$$

Substituting (134) into (133) 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 (126) 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 labour supplies (122) and (123), 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

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

F Optimal tax rules with non-separable utility

In our benchmark economy, we have assumed that households' utility is additively separable between consumption and leisure on the one hand, and climate damages on the other, as in Barrage (2020). To explore the role of this assumption, we now turn to a more general utility function where environmental damages are not separable from consumption and leisure. The derivation of the new optimal tax rules closely follows the method presented in Appendix A. This appendix highlights the differences with this benchmark.

F.1 Implementability conditions

Following the same steps as in Appendix A.1, we can show that conditions (34) and (35) still hold since they can be obtained without making any specific assumption about separability. With non-separable utility, however, the solutions for individual consumption and labour supply also depend on the environmental variable Z. As a result, the implementability constraints can be stated as

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

F.2 Ramsey problem

Following the same steps as in Appendix A.2, we can define the following pseudo-utility function

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

The Ramsey problem can then be formulated exactly like in the additively separable case: the only difference between the two problems is the expression of the pseudo-utility function W. It follows that the tax formulas expressed as functions of the derivatives of W also remain unchanged, *i.e.* we still have, $\forall t \geq 0$,

$$\begin{aligned} \tau_{H,t} &= 1 - \frac{W_{c,t}}{W_{h,t}} \frac{U_{h,t}}{U_{c,t}}, \\ R_{t+1} &= R_{t+1}^* \frac{U_{c,t}}{W_{c,t}} \frac{W_{c,t+1}}{U_{c,t+1}}, \\ \tau_{E,t} &= \frac{1}{W_{c,t}} \sum_{j=0}^{\infty} \beta^j \bigg[\Big(W_{c,t+j} D'_{t+j} A_{1,t+j} F_{t+j} - N_{t+j} W_{Z,t+j} \Big) J_{E_t^M, t+j} \bigg], \\ \tau_{I,t} &= 0. \end{aligned}$$

F.3 characterisation of equilibrium

To understand the implications of the separability assumption, let us consider the following period utility function,

$$u\big(c_i, h_i, Z\big) = \frac{\big(c_i(1-\varsigma h_i)^\gamma \chi(Z)\big)^{1-\sigma}}{1-\sigma} + \hat{u}(Z).$$

This specification generalises our baseline additively separable utility by allowing environmental damages to also affect households' marginal utility from consumption and leisure through the function $\chi(Z)$, with $\chi'(Z) < 0$, $\lim_{Z\to\infty} \chi(Z) \geq 0$, and normalised such that $\chi(0) = 1$. Using this functional form, the new Lagrangian of the characterisation problem is, $\forall t \geq 0$,

$$\mathcal{L}_t = \sum_i \pi_i \varphi_i \Bigg[\frac{\big(c_{i,t}(1-\varsigma h_{i,t})^\gamma \chi(Z_t)\big)^{1-\sigma}}{1-\sigma} + \hat{u}(Z_t)\Bigg] + \theta_t^c\big(c_t - \sum_i \pi_i c_{i,t}\big) - \theta_t^h\big(h_t - \sum_i \pi_i e_i h_{i,t}\big).$$

Following the same steps as in Appendix A.4.1, we can show that individual consumption and labour supply decisions can be expressed with the same functions of aggregates as in the additively separable case,

$$\begin{aligned} c_{i,t}^m\big(c_t, h_t, Z_t; \varphi\big) &= \omega_i c_t, \\ 1 - \varsigma h_{i,t}^m\big(c_t, h_t, Z_t; \varphi\big) &= \frac{\omega_i}{e_i}\big(1 - \varsigma h_t\big), \end{aligned}$$

where, making the same normalisation assumption as in the additively separable case,

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

From the above, it follows that the environmental variable Z affects individual consumption only through aggregate consumption, and affects individual labour supply only through aggregate labour supply. We can now write the aggregate indirect utility $U\big(c_t, h_t, Z_t; \varphi\big)$ in terms of the aggregates c_t, h_t, Z_t, and market weights φ,

$$\begin{aligned} U\big(c_t, h_t, Z_t, \varphi\big) =& \sum_i \pi_i \varphi_i \left(\frac{\omega_i^{1+\gamma}}{e_i^\gamma}\right)^{1-\sigma} \frac{\big(c_t\big(1-\varsigma h_t\big)^\gamma \chi(Z_t)\big)^{1-\sigma}}{1-\sigma} + \sum_i \pi_i \varphi_i \hat{u}(Z_t) \\ =& \frac{\big(c_t\big(1-\varsigma h_t\big)^\gamma \chi(Z_t)\big)^{1-\sigma}}{1-\sigma} + \Gamma \hat{u}(Z_t), \end{aligned}$$

where we have again used the same normalisation of market weights as in the additively separable case, and with $\Gamma \equiv \sum_i \pi_i \varphi_i$.

F.4 Explicit tax formulas

Using our functional form, we can express the pseudo-utility function as

$$
\begin{aligned}
W(c_t, h_t, Z_t; \varphi, \theta, \lambda) = & \sum_i \pi_i \frac{\lambda_i \omega_i}{\varphi_i} \frac{\big(c_t(1-\varsigma h_t)^\gamma \chi(Z_t)\big)^{1-\sigma}}{1-\sigma} + \sum_i \pi_i \lambda_i \hat{u}(Z_t) \\
& + \sum_i \pi_i \theta_i \Big[\omega_i (1+\gamma)\big(c_t\big(1-\varsigma h_t\big)^\gamma \chi(Z_t)\big)^{1-\sigma} - \gamma e_i \big(1-\varsigma h_t\big)^{-1}\big(c_t\big(1-\varsigma h_t\big)^\gamma \chi(Z_t)\big)^{1-\sigma}\Big], \\
= & \Phi \frac{\big(c_t\big(1-\varsigma h_t\big)^\gamma \chi(Z_t)\big)^{1-\sigma}}{1-\sigma} + \hat{u}(Z_t) + \Psi U_{h,t},
\end{aligned}
$$

with again

$$
\begin{aligned}
\Phi &\equiv \sum_i \pi_i \omega_i \Big(\frac{\lambda_i}{\varphi_i} + (1-\sigma)(1+\gamma)\theta_i\Big), \\
\Psi &\equiv \sum_i \frac{\pi_i \theta_i e_i}{\varsigma}.
\end{aligned}
$$

With our new functional form, U_{hh}/U_h and U_{hc}/U_c remain unchanged relative to the additively separable case. It follows that the explicit formulas for labour and capital income taxes given by (62) and (63) remain exactly unchanged. Turning to the pollution tax, we now have

$$
\begin{aligned}
\tau_{E,t} = \sum_{j=0}^{\infty} \beta^j \Bigg(& \frac{V_{c,t+j} + \sum_i \pi_i \theta_i \mathcal{I}_{c,i,t+j}}{V_{c,t} + \sum_i \pi_i \theta_i \mathcal{I}_{c,i,t}} D'_{t+j} A_{1,t+j} F_{t+j} \\
& - N_{t+j} \frac{\tilde{V}_{Z,t+j} + \hat{V}_{Z,t+j} + \sum_i \pi_i \theta_i \mathcal{I}_{Z,i,t+j}}{V_{c,t} + \sum_i \pi_i \theta_i \mathcal{I}_{c,i,t}} \Bigg) J_{E_t^M, t+j}, \quad \forall t \geq 0,
\end{aligned} \tag{135}
$$

with

$$
\begin{aligned}
\tilde{V}(c_t, h_t, Z_t, \varphi, \lambda) &\equiv \sum_i \pi_i \frac{\lambda_i \omega_i}{\varphi_i} \frac{\big(c_t(1-\varsigma h_t)^\gamma \chi(Z_t)\big)^{1-\sigma}}{1-\sigma}, \\
\hat{V}(Z_t) &\equiv \hat{u}(Z_t),
\end{aligned}
$$

and

$$
\mathcal{I}_{i,t} \equiv U_{c,t} c_{i,t}^m + U_{h,t} e_i h_{i,t}^m = \Big[\omega_i(1+\gamma) - \gamma e_i \big(1-\varsigma h_t\big)^{-1}\Big] \Big(c_t \big(1-\varsigma h_t\big)^\gamma \chi(Z_t)\Big)^{1-\sigma}. \tag{136}
$$

Notice that equation (135) generally applies to any utility function where damages are modelled both additively and non-additively, *i.e.* it is not specific to the balanced growth preference assumption. The novel feature relative to the baseline formula is the additional term, $\sum_i \pi_i \theta_i \mathcal{I}_{Z,i,t+j}$. This term, which Barrage (2020) labels the *offer curve impact* of environmental damages, represents how much environmental degradation affects the planner's ability to implement its preferred allocation. As can be seen from equation (136), this effect matters only to the extent that the environmental variable affects the marginal utility from consumption and leisure. Using our functional form, we have

$$
\mathcal{I}_{Z,i,t} = \mathcal{I}_{c,i,t} \frac{\tilde{V}_{Z,t}}{V_{c,t}},
$$

which depends only on non-separable damages $\tilde{V}_{Z,t}$, but not on separable damages $\hat{V}_{Z,t}$. Following the definition of the MCF (see Section 3), we have

$$V_{c,t} + \sum_i \pi_i \theta_i \mathcal{I}_{c,i,t} = V_{c,t} \times \text{MCF}_t, \quad \forall t \geq 0,$$

hence we can re-write the optimal pollution tax as

$$\tau_{E,t} = \sum_{j=0}^{\infty} \beta^j \left[\frac{\text{MCF}_{t+j}}{\text{MCF}_t} \left(\frac{V_{c,t+j}}{V_{c,t}} D'_{t+j} A_{1,t+j} F_{t+j} - N_{t+j} \frac{\tilde{V}_{Z,t+j}}{V_{c,t}} \right) - N_{t+j} \frac{\hat{V}_{Z,t+j}}{\text{MCF}_t \times V_{c,t}} \right] J_{E_t^M, t+j}, \quad \forall t \geq 0. \tag{137}$$

The offer curve impact term therefore leads to adjust the non-additive utility damages with their contemporaneous MCF. Intuitively, these non-additive damages are akin to consumption damages as they affect the marginal utility of consumption. Hence, these are internalised in the same way as production damages.

Following the same steps as in Appendix A.5, one can show that the welfare-weighted average MCF from period 0 onward is again equal to 1, that the MCF is equal to 1 in all periods when the IES is equal to unity, and that the ratio $\text{MCF}_{t+j}/\text{MCF}_t$ is 1 as long as going forward the optimal capital tax is null. Hence, the implications of tax distortions for optimal carbon taxation are the same as in the benchmark where environmental damages enter the utility function only in an additively separable way.

Turning to the effect of inequality, the main difference with the additively separable case is that non-separable damages, as captured by $\tilde{V}_Z/V_c$, are *not* affected by consumption inequality. In particular, we have

$$\frac{\tilde{V}_{Z,t+j}}{V_{c,t}} = \frac{\chi'(Z_t)}{\chi(Z_t)} c_t,$$

which depends on aggregates only. Thus, the optimal pollution tax is affected by inequality only through the term $\hat{V}_{Z,t+j}/V_{c,t}$, *i.e.* only through the additively separable damages. Thus, if part of environmental utility damages enter in a non-additively separable way, the effect of inequality on the optimal pollution tax could become even smaller than in the baseline.

G Calibration

G.1 Government

Regarding taxes on capital and labour income, we follow Trabandt and Uhlig (2012), who conduct a detailed analysis of fiscal policies in the US and a number of European countries. Using a comprehensive measure of taxes on capital income, they find that, on average, capital income in the US is taxed at a rate of 41.4%. Hence, we set a time-invariant $\tau_K = 0.411$ in our baseline.[43] They find that labour

[43] 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 labour income taxes, personal income taxes are adjusted to account for payroll taxes and social security contributions.

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-labour wedge of 25.5%, or $\tau_H = 1 - (1 - 0.221)/(1 + 0.046) = 0.255$.

To obtain the empirical counterparts of the two components of government spending, government consumption G_t^C and exogenous transfers G_t^T, we collect data on US government expenses from the IMF Government Finance Statistics. As in Barrage (2020), we include the following categories from the expense breakdown in G_t^C: compensation of employees, use of goods and services, subsidies, grants, and other expenses. For transfers G_t^T, we include social benefits. Averaging over the years 2011–2015, government consumption is $G_0^C/Y_{1,0} = 0.158$, while government transfers are $G_0^T/Y_{1,0} = 0.145$. To keep the sizes comparable to GDP going forward, both government consumption and exogenous transfers grow at the sum of technological progress and population growth.[44]

G.2 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. The annual labour 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 labour supply (*i.e.*, the sum of the respondent and their partner's labour supply). This number reflects how much household 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 into 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.

G.3 Wealth distribution

For each of the ten productivity groups, we again divide households into ten deciles of net worth. For each sub-group, we compute the average net worth. This provides a table in which households are split into 100 groups of equal size, with each of these groups having different average hourly wages and net

[44] 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. This does not have important implications for the results, except that the average level of lump-sum transfers is affected.

worth.[45]

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 into 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.

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 numbers 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.

G.4 Baseline hours worked

We also use the SCF 2013 to compute the initial labour 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 labour supply, we do not eliminate outliers based on their hourly wage or labour supply. In particular, we keep unemployed households for whom the hourly wage is not observed, as dropping them would lead to an overestimation of the average labour supply. For all households in the sample, we divide the annual labour supply by the number of working-age individuals (individuals between 18 and 65). This yields an average of 1440 hours annually. Assuming a maximum labour supply capacity of 52 weeks per year and 100 hours per week per individual, this yields an average labour supply of 0.277 of the maximum capacity.

[45] 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.

G.5 Other parameters

All parameters used in the calibration are summarised in Tables VI and VII below.

G.6 Stone-Geary calibration

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

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

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,

$$d_{h,t}(p_{E,t} + \tau_{D,t}) = \bar{d}_{h,t}(p_{E,t} + \tau_{D,t}) + \epsilon c_{h,t}. \tag{138}$$

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 (138),

$$d_h(p_E + \tau_D) = \beta_d + \beta_\epsilon c_h + \mu_h, \tag{139}$$

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 Survey (CEX). We estimate regression (139) 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 (138)—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,

$$\bar{d}_h(p_E + \tau_D) = \sum_{j \in J} \beta_j \mathcal{I}_h^j + \eta_h, \tag{140}$$

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

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.

expenditures. In order to characterise 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 labour 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 household members were paid on average for each hour of work they supplied in the past year.[46]

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 into five groups of hourly wage quintiles. For each of the five groups, we again divide households into 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.[47] 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

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

[47]We divide households into only 15 necessity groups to mitigate the potential overestimation of consumption heterogeneity due to measurement error at the household level in the CEX and to avoid negative values for the necessity levels.

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.

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 (139) and (140) 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 normalised 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 into 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.

Additional parameter adjustments We target the share of emissions coming from households direct energy consumption in the residential and transportation sectors based on the US EPA. In 2013 (our initial period), 17% of US 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 US aggregate emissions. To target this number, 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 does not affect our results in the benchmark model. To remain consistent with our previous targets for the initial labour supply, Frisch elasticity, initial emissions, share of utility damages, and capital share relative to labour, we also adjust the values of ς, γ, $A_{2,2015}$, α_0, and α. Table VIII reports the value of the adjusted parameters.

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.

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		
Z_{2015}^{At}	Initial atmospheric temperature change (C since 1900)	0.85
Z_{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
Preferences			
β	Utility discount rate (per year)	1/(1.015)	DICE-2016
σ	Inverse of IES	1.45	DICE-2016
η^F	Frisch elasticity of labour supply	0.75	Chetty et al (2011)
ς	Labour disutility coefficient	1.875	To target η^F and h_{2015}
γ	Labour disutility exponent	0.753	To target η^F and h_{2015}
α_0	Relative preference for the environment	7.88e-05	Adapted from Barrage (2019)
Production damages			
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
Production first sector			
α	Return to scale on labour 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 labour 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 labour supply per capita	0.277	Computed from SCF
$A_{1,2015}$	Initial TFP sector 1	141.9	To target Y_{2015}
Production second sector			
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}
Abatement costs			
$P_{2015}^{\text{backstop}}$	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
Government			
G_t/Y_t	Government spending to GDP ratio	0.3030	IMF-GFS
B_{2015}	Initial net public debt to GDP ratio	0.2220	IMF-GFS
$\tau_{H,2015}$	Initial tax rate on labour 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

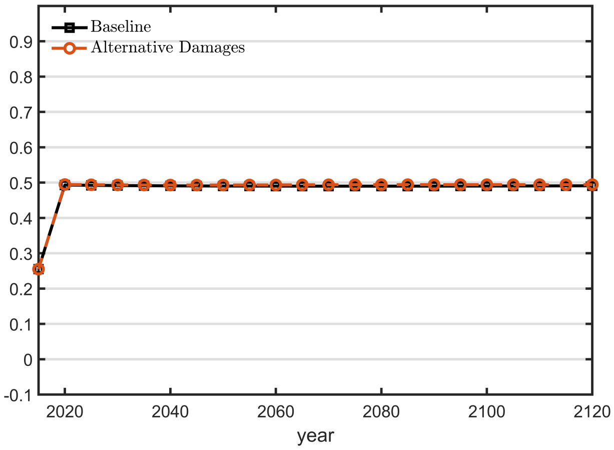
Table VIII: Calibration Summary: Economic Parameters with Stone-Geary Preferences.

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 labour sector 1	0.259
ς	Labour disutility coefficient	1.881
γ	Labour disutility exponent	0.728
$A_{2,2015}$	Initial TFP sector 2	20.4
α_0	Relative preference for the environment	7.92e−05

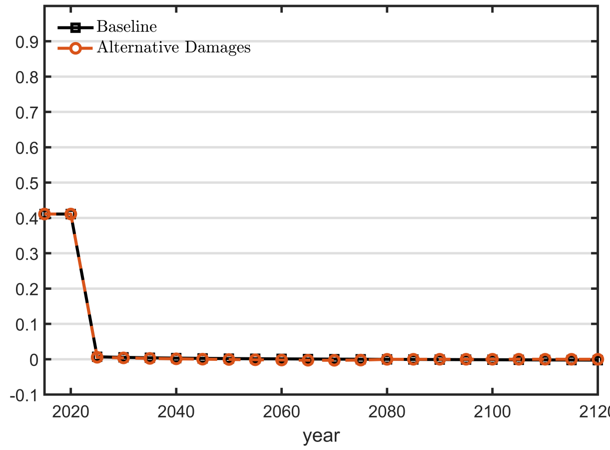
Notes: The table reports the values of the parameters used in the calibration of the extended version of the model with two goods and Stone-Geary utility (see Section 6.2). The parameters are selected to obtain energy expenditure shares consistent with the CEX and a share of aggregate emissions coming from households' energy consumption consistent with EPA's estimates.

H Additional quantitative results

H.1 Alternative damages



(a) Optimal labour income taxes



(b) Optimal capital income taxes

Figure 9: Optimal Income Taxes, Alternative Damages.

Notes: Figures show the path of second-best labour 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).

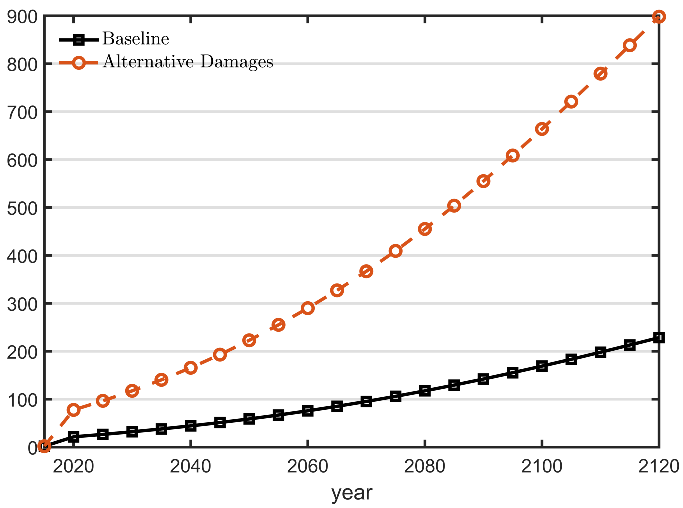
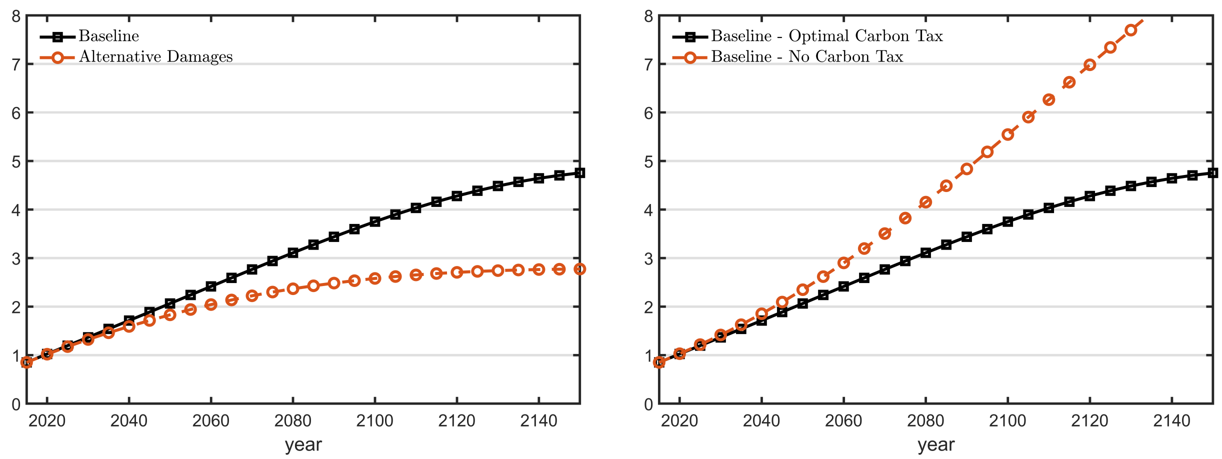


Figure 10: Optimal Carbon Taxes ($\$/tCO_2$), Alternative Damages.

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. The initial level (for 2015) is set exogenously to its current level obtained from Nordhaus (2017).



(a) Alternative Damages (b) No Carbon Tax

Figure 11: Temperature change: Optimal Path, Alternative Damages, and No Carbon Tax.

Notes: Panels show the path of atmospheric temperature change (Z_t^{At}) for the baseline calibration (black) and (a) the alternative-damages calibration (red) or (b) the business-as-usual no-carbon-tax scenario (red), expressed in degrees Celsius.

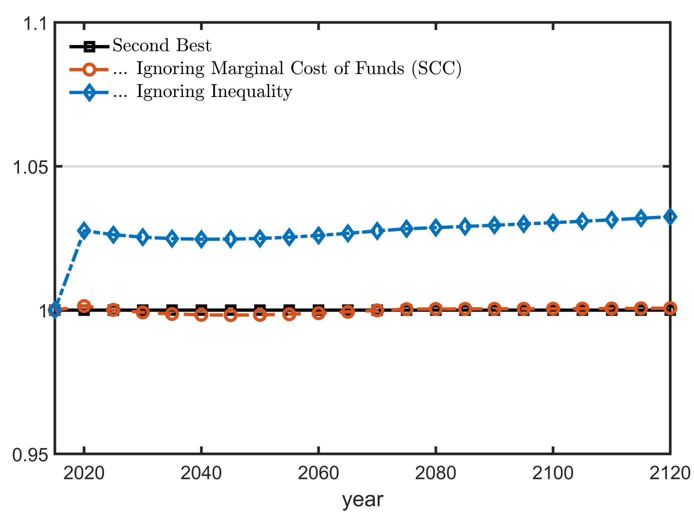
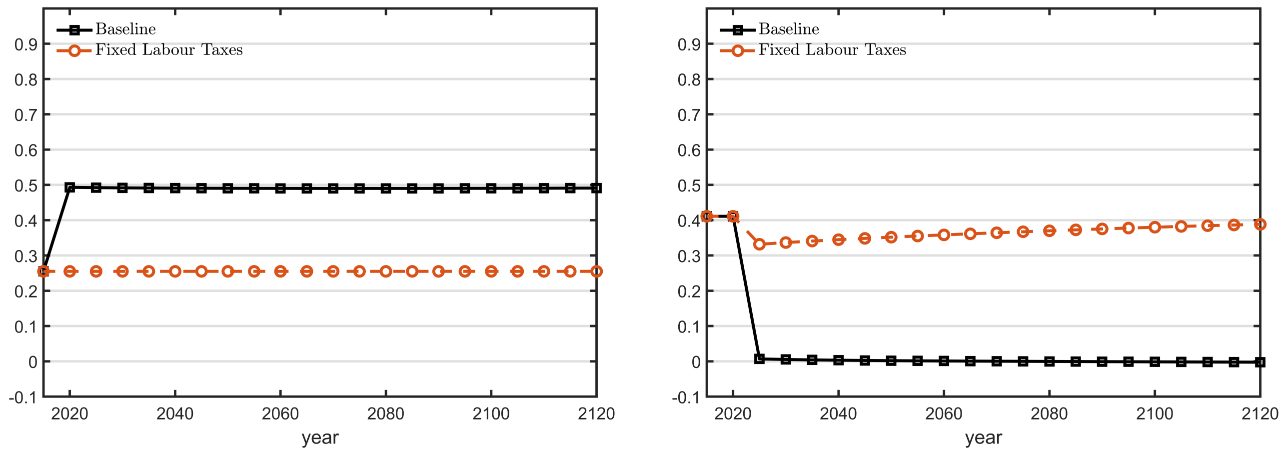


Figure 12: Carbon Tax Decomposition, Alternative Damages.

Notes: The black line represents the second-best carbon tax normalised 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.

H.2 Third-best policies

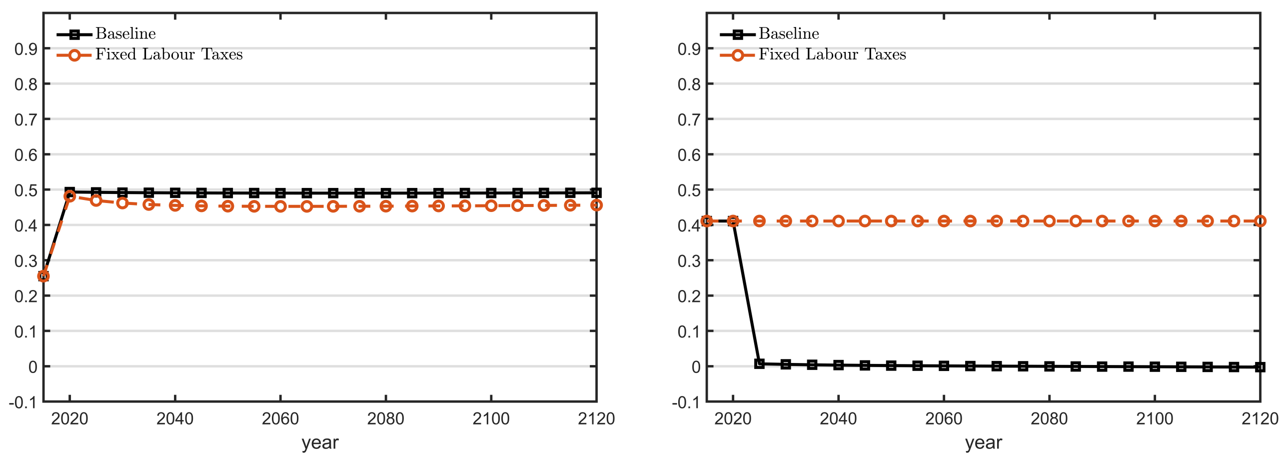


(a) Optimal labour income taxes

(b) Optimal capital income taxes

Figure 13: Optimal Income Taxes, Given Labour Tax.

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

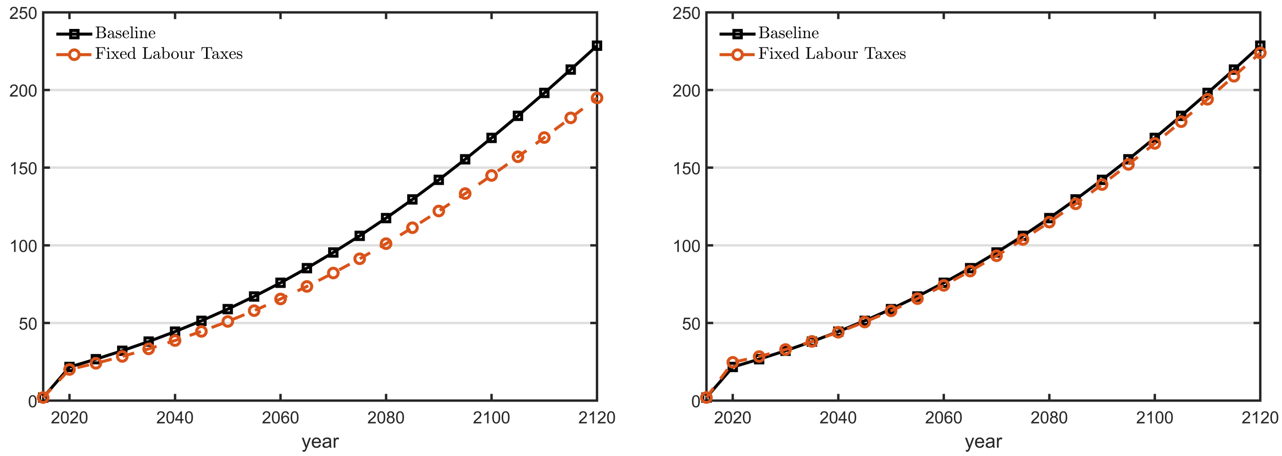


(a) Optimal labour income taxes

(b) Optimal capital income taxes

Figure 14: Optimal Income Taxes, Given Capital Tax.

Notes: Figures show the path of second-best labour 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 labour income tax

(b) Given capital income tax

Figure 15: Optimal Carbon Taxes ($/tCO$_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 labour and capital income taxes (red), expressed in dollars per ton of CO_2. The 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 Labour Income Taxes.

	Revenue Source			Revenue Use		
	Labour	Capital	Carbon	Gov. Cons.	Transfer	Interest
No Carbon Tax	17.2%	5.5%	0.0%	16.0%	5.0%	2.1%
Optimal Carbon Tax	17.0%	5.3%	1.0%	15.8%	6.0%	2.1%
Change	−0.1%	−0.2%	1.0%	−0.2%	0.9%	0.0%

Notes: For given labour 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		
	Labour	Capital	Carbon	Gov. Cons.	Transfer	Interest
No Carbon Tax	31.5%	6.2%	0.0%	18.6%	16.5%	2.1%
Optimal Carbon Tax	30.8%	6.1%	1.3%	18.2%	17.3%	2.2%
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.

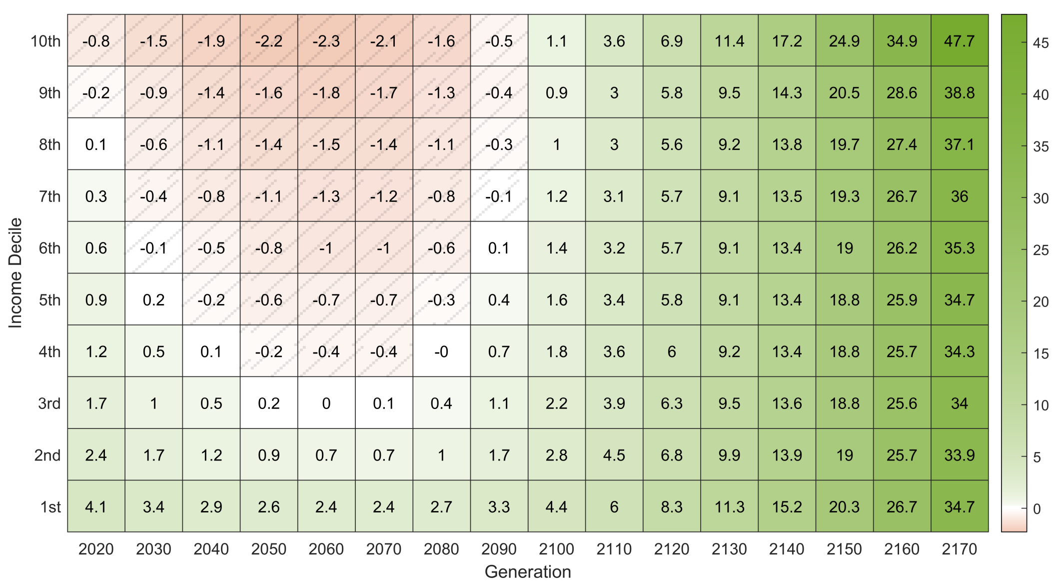


Figure 16: Period Welfare Gains (%), Given Labour Income Taxes.

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 labour income taxes.

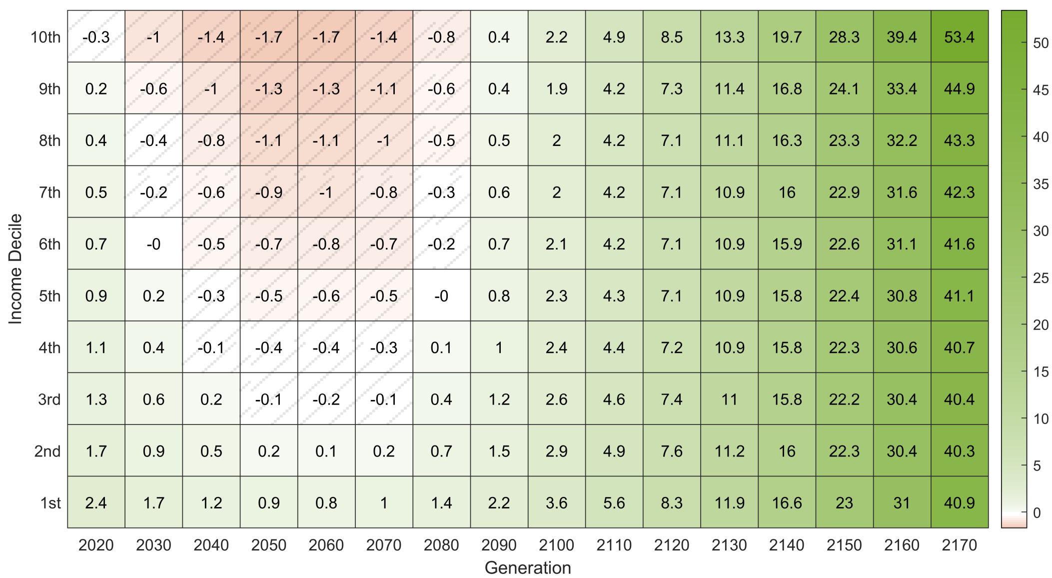
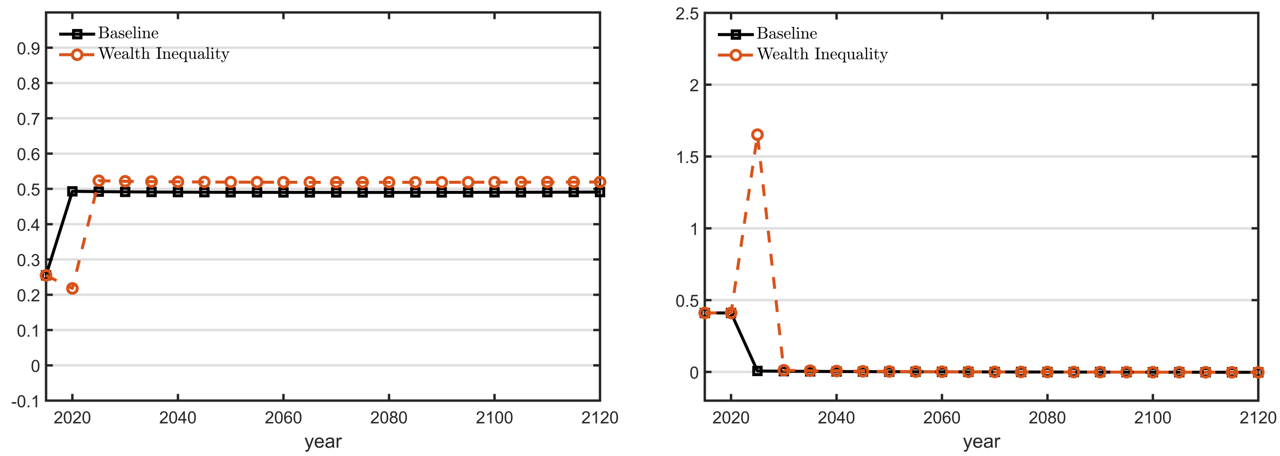


Figure 17: Period Welfare Gains (%), Given Capital Income Taxes.

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.

H.3 Initial wealth inequality



(a) Optimal labour 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 labour 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).

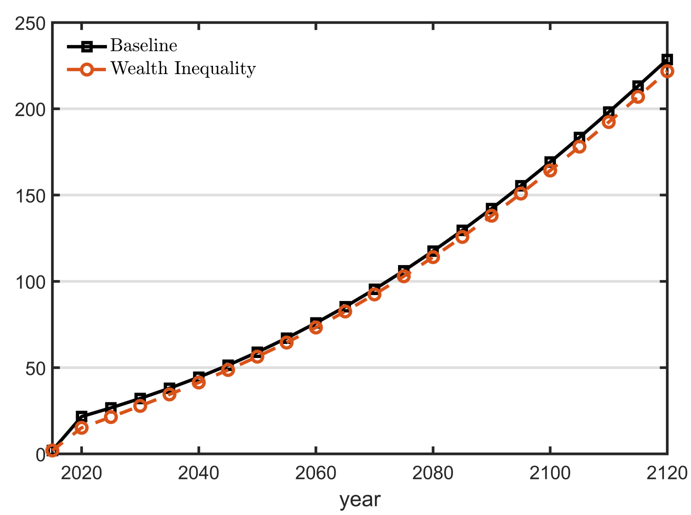


Figure 19: Optimal Carbon Taxes (\$/tCO_2), Initial Wealth Heterogeneity and Exogenous Initial Capital Tax.

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. The 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		
	Labour	Capital	Carbon	Gov. Cons.	Transfer	Interest
No Carbon Tax	34.2%	3.2%	0.0%	17.9%	18.1%	1.5%
Optimal Carbon Tax	33.5%	3.2%	1.1%	17.6%	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.

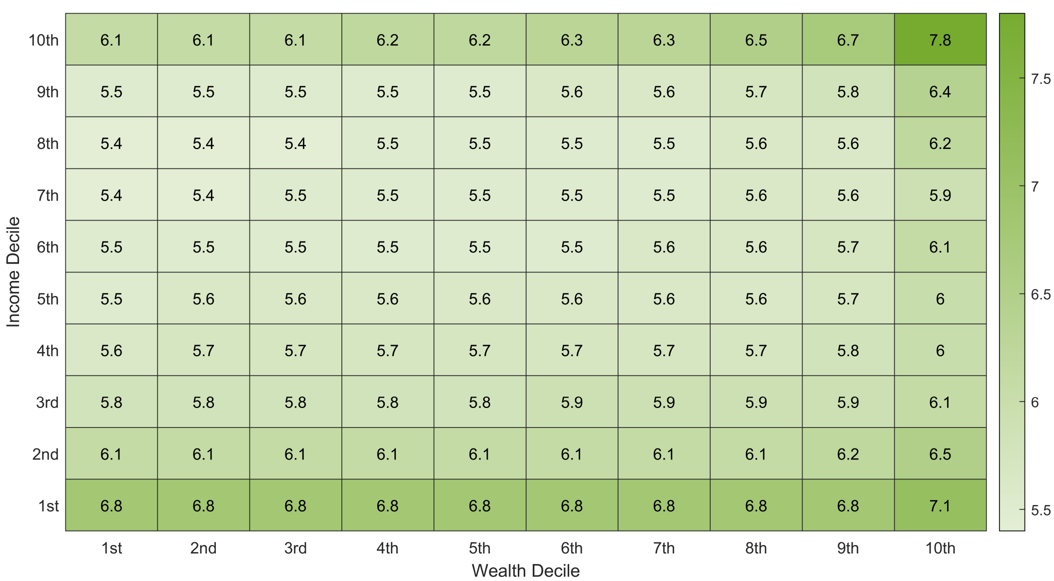
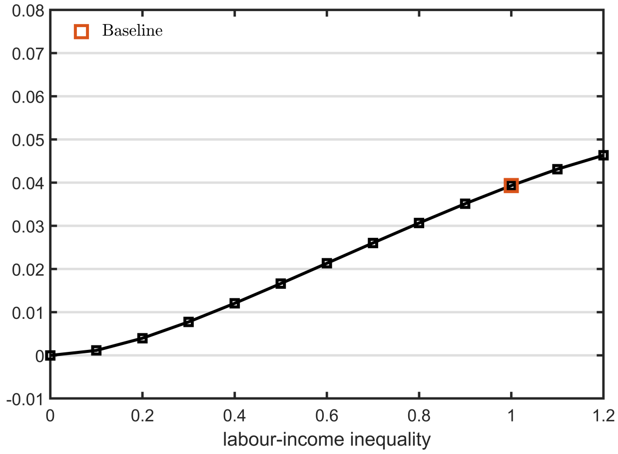


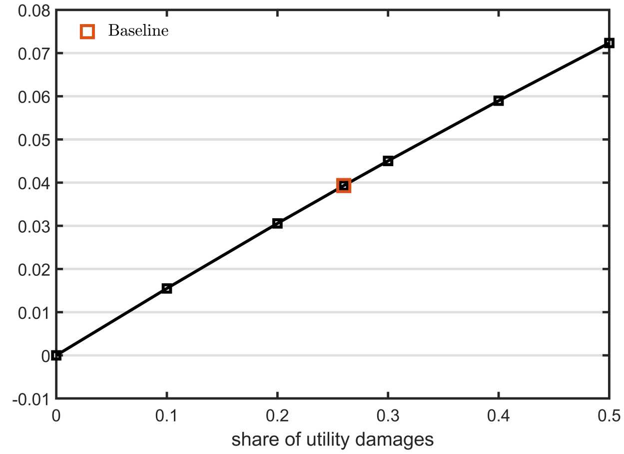
Figure 20: Welfare Gains (%), Initial Wealth Heterogeneity and Exogenous Initial Capital Tax.

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.

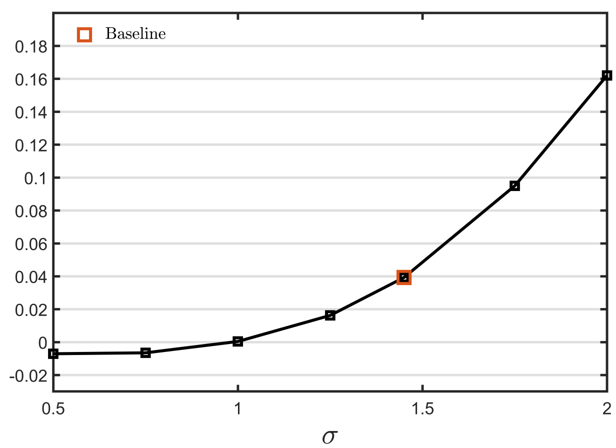
H.4 Sensitivity of inequality effects to calibration choices



(a) Effect of Labour Income Inequality



(b) Effect of Share of Utility Damages



(c) Effect of σ

Figure 21: How Inequality Effects Change with Different Levels of Labour 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 labour income inequality. (a) To obtain different levels of labour 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 labour 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.

I 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 aggregates. 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 elabourate further on this part.

3. *Compute terms that appear in the optimality conditions of the Ramsey planner:* Compute M using equation (69), then obtain ω_i and φ_i, for all i, from equations (70) and (71)—equations (57) and (58) can be used to obtain individual allocations and welfare. Next, obtain Φ and Ψ using equations (67) and (68). Equation (61) then gives $W_{c,t}$, $W_{h,t}$, and $W_{Z,t}$, for all t.

4. *Update policy:* Use equations (51), (52), (55), and (64) 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.

The initial balanced growth path is calibrated to 2015. One period in the model corresponds to 5 years. We allow policy, climate, and exogenous variables to evolve over 45 periods, i.e., from 2015 to 2240. After that, we give the model an additional 25 periods (2240 to 2365) to converge to the final balanced growth path.