Optimal fiscal policy in a second-best climate economy model with heterogeneous agents

Thomas Douenne *† Albert Jan Hummel *‡ Marcelo Pedroni *§

April 22, 2022

Disclaimer: This document is a work in progress. Please do not circulate or cite it without the consent of the authors.

Link to most recent version

Abstract

This paper studies optimal fiscal policy in a climate economy model with heterogeneous agents and distortionary taxes on labor and capital income. We derive optimal tax rules and show how they are modified relative to first-best and relative to the case without heterogeneity. We also explore quantitatively the role of heterogeneity for optimal carbon taxation in a version of the model that is calibrated to the U.S. economy.

JEL classification: E62, H21, H23, Q5

Keywords: Climate policy; Carbon taxes; Optimal taxation; Heterogeneous agents.

*University of Amsterdam, Amsterdam School of Economics. Present address: University of Amsterdam, Roetersstraat 11, 1018 WB Amsterdam, Netherlands.

†email: t.r.g.r.douenne@uva.nl

‡email: a.j.hummel@uva.nl

§email: m.pedroni@uva.nl

1 Introduction

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

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

We model this as a Ramsey problem in which the government chooses the level of linear taxes—in particular, taxes on labor and capital income, energy, and pollution—and a uniform lump-sum transfer to maximize aggregate welfare. Because agents are heterogeneous but tax instruments are anonymous, the government must rely on distortionary instruments to provide redistribution. We analytically characterize optimal tax formulas and study the implications of heterogeneity for optimal pollution taxation.

[To be included: results]

Our paper contributes to two strands of the literature. On the normative side, it contributes to the literature on the optimal taxation of pollution. Since the pioneering work of Pigou (1920), an extensive body of literature has studied pollution taxes in second best environments. Important papers in that literature include (among others) Sandmo (1975), Bovenberg and de Mooij (1994), Bovenberg and van der Ploeg (1994), Cremer et al. (1998), Kaplow (2012), Jacobs and de Mooij (2015), and Jacobs and van der Ploeg (2019). These papers usually focus on static settings and model the pollution externality in a stylized manner. By contrast, we develop a dynamic general equilibrium (DGE) model of the climate-economy, which enables us to study second-best environmental taxation in a richer setting. In doing so, our paper closely relates to Barrage (2019) who creates a critical bridge between the climate-economy literature and the dynamic public finance literature. Her framework integrates a climate-economy model in the spirit of Golosov et al. (2014) into the representative agent Ramsey model of Chari and Kehoe (1999). Our main innovation relative to Barrage (2019) is to introduce heterogeneous agents, which we see as critical for two reasons. First, this allows us to jointly study environmental and equity issues. In addition of the importance of equity in normative analysis, recent experience has shown that the distributional effects of environmental policies were also critical to ensure their public support.[1] Second, agents' heterogeneity provides a sound foundation for the study of second-best policies. In representative agent settings, the second-best environment arises because lump-sum transfers are assumed unfeasible: governments therefore *need to* rely on distortionary taxes to finance their expenditures. Yet, in practice lump-sum transfers are feasible as they simply correspond to the intercept on a tax scheme, and recent policy proposals such as the carbon tax and dividend advocated by the Climate Leadership Council even call for using such instruments to redistribute the carbon tax revenue.[2] With heterogeneous agents, lump-sum transfers are no longer excluded as long as they do not discriminate between agents. Although this non-distortionary source of public income is available, governments now *want to* use distortionary taxes to provide redistribution. Thus, the rationale behind distortionary taxation is entirely different in heterogeneous agents model, leading to reconsider the implications of the optimal tax results.

On the positive side, this paper contributes to the analysis of the distributional effects of environmental taxes in general equilibrium. An extensive literature has analyzed the distributional effects of environmental taxes through the consumption channel (for a recent survey, see Pizer and Sexton, 2019), generally pointing to regressive effects since the consumption share of polluting goods tends to decrease with income (Levinson and O'Brien, 2019). More recently, several authors have also analyzed the heterogeneous incidence of environmental taxes on households' income. While a number of papers found progressive effects due to the larger negative impact of the policy on capital income relative to labor income and transfers (see *e.g.* Rausch et al., 2011; Fullerton and Monti, 2013; Williams et al., 2015; Goulder et al., 2019), the recent work of Känzig (2021) shows—exploiting exogenous shocks to the EU-ETS price—that carbon taxation has a larger impact on poor households' income in the U.K. because these households are over-represented in pro-cyclical sectors that are more impacted by the tax. Känzig (2021) is also one of the few papers introducing heterogeneous agents in a climate-economy model. Based on the framework of Golosov et al. (2014) in which he introduces two types of agents (hand-to-mouth and savers) and nominal rigidities, he shows that redistributing the tax revenue through lump-sum transfers would reduce regressivity. In a different approach, Kotlikoff et al. (2021) introduce overlapping generations into a climate-economy model based on Golosov et al. (2014) and Nordhaus (2017) to quantitatively study the carbon price path that would lead to the largest uniform welfare

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

[2] See Wall Street Journal 2019's column signed by 3,354 American economists in support of carbon pricing with lump-sum rebates.

gain across current and future generations. In a related paper, Fried et al. (2018) study the effect of introducing a carbon tax with three alternative revenue-recycling schemes in a quantitative OLG model with heterogeneity within-generations. They focus on the non-environmental benefits of the policies, and investigate their effects both at the steady state and during transition. They show that while a uniform lump-sum rebate is more costly than reductions of the labor or capital tax rates in steady state, it is more favorable to the current generation and leads to less adverse distributional effects. In a working paper (Fried et al., 2021), the same authors focus exclusively on the steady-state and study the optimal recycling policy, but the carbon tax remains exogenous and the analysis abstracts from environmental effects.[3] Relative to this literature, our objective is to develop a framework to analytically and quantitatively study optimal carbon taxation in a dynamic general equilibrium setting with a rich representation of agents' heterogeneity. Because our model endogenizes the environmental externality, it also allows us to study the heterogeneous welfare impacts arising from the environmental effects of climate policies.

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

2 Model

The model builds on Barrage (2019): one sector of the economy produces a final good using capital, labor, and energy that is produced in the second sector. Energy production generates pollution that leads to environmental degradation, which in turn affects productivity and households' utility. The government finances an exogenous stream of expenditures using taxes on labor income, capital income, energy, and pollution, as well as a lump-sum tax. The key differences with Barrage (2019) are that in our model, agents are heterogeneous and the government has access to a (non-individualized) lump-sum 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 divided into types $i \in I$ of size π_i. Each agent of type $i \in I$ ranks streams of consumption of a final good $c_{i,t}$, labor supply $h_{i,t}$, and environmental degradation Z_t according to the preferences

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

[3]Other papers investigating quantitatively the double-dividend from carbon taxation with heterogeneous generations include Rausch (2013) and Rausch and Yonezawa (2018). Hassler and Krusell (2012) and Krusell and Smith (2015) abstract from modeling the fiscal system, but they study the heterogeneous impacts of climate change and carbon taxation across regions.

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

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

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

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. Ensuring no arbitrage opportunities requires $p_t = R_{t+1} p_{t+1}$, and defining $T \equiv \sum_{t=0}^{\infty} p_t T_t$, the budget constraint is equivalent to

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

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

$$\begin{aligned} \beta^t \frac{u_c\left(c_{i,t}, h_{i,t}, Z_t\right)}{u_c\left(c_{i,0}, h_{i,0}, Z_0\right)} &= p_t, \quad \forall\, t \geq 0, \\ \frac{u_h\left(c_{i,t}, h_{i,t}, Z_t\right)}{u_c\left(c_{i,t}, h_{i,t}, Z_t\right)} &= -\left(1-\tau_{H,t}\right) e_i w_t, \quad \forall\, t \geq 0, \end{aligned}$$

which holds across all agents.

2.2 Final-good sector

As in Barrage (2019), there are two production sectors. In the final-good sector, indexed by 1, 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 $K_{1,t}$, labor $H_{1,t}$, and energy E_t. The total factor productivity is given by $A_{1,t}$ and the function $D\left(Z_t\right)$ controls the damages to production implied by environmental degradation, with $D'(Z_t) > 0$. The output $Y_{1,t}$ is given by

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

The first order conditions for the firm problem are:

$$r_t = \left(1 - D\left(Z_t\right)\right) A_{1,t} F_K\left(K_{1,t}, H_{1,t}, E_t\right), \quad \forall\, t \geq 0, \tag{3}$$

$$w_t = \left(1 - D\left(Z_t\right)\right) A_{1,t} F_H\left(K_{1,t}, H_{1,t}, E_t\right), \quad \forall\, t \geq 0, \tag{4}$$

$$p_{E,t} = \left(1 - D\left(Z_t\right)\right) A_{1,t} F_E\left(K_{1,t}, H_{1,t}, E_t\right), \quad \forall\, t \geq 0. \tag{5}$$

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

2.3 Energy sector

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

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

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

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

where $\tau_{I,t}$ denotes the excise intermediate-goods tax on total energy, and $\tau_{E,t}$ denotes the excise tax on pollution emissions $E_t^M = \left(1-\mu_t\right) E_t$. Firms maximize profits subject to the technology constraint given by (6) by choosing the abatement term μ_t, capital $K_{2,t}$, and labor $H_{2,t}$.[4] The first order conditions are

$$r_t = \big(p_{E,t} - \tau_{I,t} - \tau_{E,t}(1-\mu_t) - \Theta_{E,t}(\mu_t, E_t)\big) A_{2,t} G_K\left(K_{2,t}, H_{2,t}\right), \quad \forall\, t \geq 0, \tag{7}$$

$$w_t = \big(p_{E,t} - \tau_{I,t} - \tau_{E,t}(1-\mu_t) - \Theta_{E,t}(\mu_t, E_t)\big) A_{2,t} G_H\left(K_{2,t}, H_{2,t}\right), \quad \forall\, t \geq 0, \tag{8}$$

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

Capital and labor are mobile across sectors, so the market clearing conditions give

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

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

2.4 Government

Each period the government finances the expenses G_t and lump sum transfers T_t with proportional income taxes on capital $\tau_{K,t}$ and labor $\tau_{H,t}$, total energy taxes $\tau_{I,t}$, emissions taxes $\tau_{E,t}$ and profit taxes

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

$\tau_{\pi,t}$. The government's intertemporal budget constraint is

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

2.5 Environmental degradation

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

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

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

2.6 Competitive equilibrium

Definition 1 *Given $\{a_{i,0}\}$, K_0, and 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 $\left\{\left(c_{i,t}, h_{i,t}\right)_i, Z_t, E_t, K_{1,t}, K_{2,t}, K_{t+1}, H_{1,t}, H_{2,t}, H_t\right\}_{t=0}^{\infty}$ such that: (i) agents choose $\left\{\left(c_{i,t}, h_{i,t}\right)_i\right\}_{t=0}^{\infty}$ to maximize utility subject to budget constraint (2) taking policies and prices (that satisfy $p_t = R_{t+1} p_{t+1}$) as given; (ii) firms maximize profits; (iii) the government's budget constraint (12) holds; (iv) markets clear: the resource constraints (6), (10), (11), and (13) hold, and*

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

3 Optimal tax rules

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

3.1 Ramsey problem

A simple characterization of equilibrium Because the government sets linear tax rates, all individuals face the same marginal rate of substitution between consumption and leisure. A direct implication is that the distribution of individual allocations (c_{it}, h_{it}) is efficient *given* aggregates (C_t, H_t, Z_t). Another way of stating this is that taxation is distortionary only to the extent it affects aggregates. Following Werning (2007), it is therefore possible to split up the optimal tax problem in two steps. The first is to determine individual allocations given aggregates, and the second is to determine the aggregates. Starting with the first step, denote by $\varphi \equiv \{\varphi_i\}$ a set of market weights with $\varphi_i \geq 0$. Using the property that individual allocations are efficient given aggregates, we can characterize these allocations by solving the following static sub-problem for each period t:

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

Here, $U\left(C_t, H_t, Z_t; \varphi\right)$ denotes the indirect aggregate utility function, computed using market weights and aggregates. In Section 3.4 below, we introduce a functional form for households' utility function in order to obtain expressions for $U\left(C_t, H_t, Z_t; \varphi\right)$, as well as for $c_{i,t}^m$ and $h_{i,t}^m$, solutions to problem (15). For now we chose to keep preferences unspecified to analyze optimal tax formulas with more generality. To reduce the notation burden and ease tractability, we simply make the assumption that utility is additively separable in Z, *i.e.* that we can write

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

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

$$\frac{U_H\big(C_t, H_t\big)}{U_C\big(C_t, H_t\big)} = \frac{u_h\left(c_{i,t}, h_{i,t}\right)}{u_c\left(c_{i,t}, h_{i,t}\right) e_i} = -w_t\left(1 - \tau_{H,t}\right),$$

and

$$\frac{U_C\big(C_t, H_t\big)}{U_C\big(C_0, H_0\big)} = \frac{u_c\left(c_{i,t}, h_{i,t}\right)}{u_c\left(c_{i,0}, h_{i,0}\right)} = \frac{p_t}{\beta^t},$$

where the variable Z has been omitted from the list of arguments in partial derivatives given the strong separability with consumption and labor. Using these relationships to substitute out for prices in agents' budget constraints, for any agent i we can arrive at the implementability condition that depends only on the aggregates C_t and H_t, and market weights φ

$$U_C\big(C_0, H_0\big)\big(R_0 a_{i,0} + T\big) \geq \sum_{t=0}^{\infty} \beta^t \Big(U_C\big(C_t, H_t\big) c_{i,t}^m\big(C_t, H_t; \varphi\big) + U_H\big(C_t, H_t\big) e_i h_{i,t}^m\big(C_t, H_t; \varphi\big)\Big), \quad \forall\, i. \tag{16}$$

The following Proposition follows immediately from the arguments above.

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

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

$$\max_{\substack{\{C_t, H_{1,t}, H_{2,t}, K_{1,t}, K_{2,t},\\ E_t, Z_t, \mu_t\}_{t=0}^{\infty}, T, \varphi, \tau_0^k \le 1}} \sum_{t,i} \beta^t \pi_i \lambda_i u\Big(c_{i,t}^m(C_t, H_t; \varphi), h_{i,t}^m(C_t, H_t; \varphi), Z_t\Big)$$

subject to

$$\begin{aligned}
U_C\big(C_0, H_0\big)\big(R_0 a_{i,0} + T\big) &\ge \sum_{t=0}^{\infty} \beta^t \left(U_C\big(C_t, H_t\big) c_{i,t}^m\Big(C_t, H_t; \varphi\Big) + U_H\big(C_t, H_t\big) e_i h_{i,t}^m\left(C_t, H_t; \varphi\right)\right), \quad \forall\ i,\\
\frac{F_K(K_{1,t} H_{1,t}, E_t)}{F_H(K_{1,t} H_{1,t}, E_t)} &= \frac{G_K(K_{2,t} H_{2,t})}{G_H(K_{2,t} H_{2,t})},\\
C_t + G_t + K_{t+1} + \Theta_t\left(\mu_t, E_t\right) &= \left(1 - D\left(Z_t\right)\right) A_{1,t} F\left(K_{1,t}, H_{1,t}, E_t\right) + \left(1 - \delta\right) K_t, \quad \forall\ t \ge 0,\\
E_t &= A_{2,t} G\left(K_{2,t}, H_{2,t}\right), \quad \forall\ t \ge 0,\\
Z_t &= J\left(S_0, E_0^M, ..., E_t^M, \eta_0, ..., \eta_t\right), \quad \forall\ t \ge 0,\\
K_{1,t} + K_{2,t} &= K_t, \quad \forall\ t \ge 0,\\
H_{1,t} + H_{2,t} &= H_t, \quad \forall\ t \ge 0.
\end{aligned}$$

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

3.2 General formulas

Initial capital taxes The first order condition with respect to τ_0^k is given by

$$U_C\left(C_0, H_0\right)\Big(\left(1 - D\left(Z_0\right)\right) A_{1,0} F_K\left(K_{1,0}, H_{1,0}, E_0\right) - \delta\Big) \sum_i \pi_i \theta_i a_{i,0} = 0$$

So, it is optimal to expropriate initial asset holdings until $\sum_i \pi_i \theta_i a_{i,0} = 0$. If this is not feasible, then the best the government can do is to raise the period-0 capital tax until $R_0 = 0$, which implies all wealth is appropriated. As explained by Werning (2007), tighter restrictions on initial capital taxation are difficult to justify because a wealth tax can be mimicked using consumption taxes. Hence, abstracting from consumption taxes, as we have done throughout, is only without loss of generality if we allow for wealth expropriation.

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

$$\tau_{H,t} = 1 - \frac{U_H\left(C_t, H_t\right)}{U_C\left(C_t, H_t\right)} \frac{W_C\left(C_t, H_t; \varphi, \theta, \lambda\right)}{W_H\left(C_t, H_t; \varphi, \theta, \lambda\right)},$$

and

$$\frac{R_{t+1}}{R^*_{t+1}} = \frac{W_C\left(C_{t+1}, H_{t+1}; \varphi, \theta, \lambda\right)}{W_C\left(C_t, H_t; \varphi, \theta, \lambda\right)} \frac{U_C\left(C_t, H_t\right)}{U_C\left(C_{t+1}, H_{t+1}\right)},$$

where the pseudo-utility function W is defined as

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

with

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

the aggregate utility based on the planner's weights,

$$IC_i(C_t, H_t) \equiv U_C\left(C_t, H_t\right) c^m_{i,t}\left(C_t, H_t; \varphi\right) + U_H\left(C_t, H_t\right) e_i h^m_{i,t}\left(C_t, H_t; \varphi\right), \tag{17}$$

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

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

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

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

$$\tau_{E,t} = \sum_{j=0}^{\infty} \beta^j \left(\frac{V_{C,t+j} + \sum_i \pi_i \theta_i MIC_{i,t+j}}{V_{C,t} + \sum_i \pi_i \theta_i MIC_{i,t}} D'\left(Z_{t+j}\right) A_{1,t+j} F\left(K_{1,t+j}, H_{1,t+j}, E_{t+j}\right) - \frac{V_{Z,t+j}}{V_{C,t} + \sum_i \pi_i \theta_i MIC_{i,t}} \right) J_{E^M_t, t+j}, \tag{18}$$

where $MIC_{i,t} \equiv \left(\partial IC_{i,t} / \partial C_t\right)$, and where the arguments to the functions $V_{C,t}$ and $MIC_{i,t}$ have been dropped to simplify notations. The term $V_{C,t} + \sum_i \pi_i \theta_i MIC_{i,t}$ appears from the substitution of $W_{C,t} = \nu_{1,t}$, where $\nu_{1,t}$ is the Lagrange multiplier on the planner's resource constraint. When the pollution tax increases, abatement increases, which increases the scarcity of the final good. The opportunity cost of increasing the pollution tax therefore corresponds to the marginal cost of increasing the final good's scarcity, which is equal to the marginal utility from consumption as computed using the planner's weights ($V_{C,t}$) minus a term which captures the marginal cost for the planner to implement its preferred allocation ($-\sum_i \pi_i \theta_i MIC_{i,t}$).

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

3.3 Comparison with first best

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

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

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

$$\tau_{E,t}^{FB} = \sum_{j=0}^{\infty} \beta^j \left(\left(\frac{V_{C,t+j}}{V_{C,t}} \right) D'\left(Z_{t+j}\right) A_{1,t+j} F\left(K_{1,t+j}, H_{1,t+j}, E_{t+j}\right) - \frac{V_{Z,t+j}}{V_{C,t}} \right) J_{E_t^M,t+j}.$$

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

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

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

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 ($-\sum_i \pi_i \theta_i MIC_{i,t}$) is not necessarily zero. In particular, we have

$$-\sum_i \pi_i \theta_i MIC_{i,t} = -\mathrm{cov}(\theta_i, MIC_{i,t}), \tag{19}$$

hence the marginal implementation cost is zero if and only if θ_i and $MIC_{i,t}$ are uncorrelated. As we show in appendix, the sign of this term also determines the sign of the marginal cost of public funds relative to 1, and whether the second best pollution tax should be set above or below the social cost of pollution.

Determinants of the marginal implementation cost Intuitively, $-\theta_i$ represents the shadow cost of implementing the desired allocation for agent i, which can be understood as the increase in implementation cost resulting from an additional unit of lump-sum transfer received by this agent. While this marginal cost is on average null, it may be positive for some agents and negative for others. Using functional forms below, we show that $-\theta_i$ is negative for households who are valued relatively more by the market than by the planner as compared to an average household.

The second term, $MIC_{i,t}$, represents how the difference between the agent current consumption and current labor income changes when more resources are available for consumption. At the optimum, agents' implementability conditions must be binding, hence with full expropriation of initial wealth—that, as we have shown, is optimal in this framework—we have

$$\sum_{t=0}^{\infty} \beta^t IC_i(C_t, H_t) = U_C\left(C_0, H_0\right) T, \tag{20}$$

from which we see that the discounted sum of $IC_{i,t}$ is invariant across types. Intuitively, this condition means that with initial wealth expropriation and a uniform lump-sum transfer, the discounted sum of expenditures minus the discounted sum of labor incomes must be the same for everyone. Still, $IC_{i,t}$ does not need to be constant across types in each period.

From equations (19), the marginal implementation cost will differ from zero only if individuals' expenditures minus labor income are responsive to contemporaneous changes in aggregate consumption C_t, and if these responses are heterogeneous. In particular, when preferences are such that individuals' expenditures minus labor income can be expressed as a constant fraction of aggregates, *i.e.* if we can write

$$IC_i(C_t, H_t) = m_i \tilde{IC}(C_t, H_t), \tag{21}$$

then from (20) we have that for any types i, j and any period t, $m_i = m_j$ and $MIC_{i,t} = MIC_{j,t}$. From (19), this implies that the marginal implementation cost is null in all periods and the second best tax is equal to the social cost of pollution. The reason is that increasing the pollution tax—and thereby leaving less resources available for consumption—affects the costs from satisfying (typically) poor agents' implementability constraint just as much as the benefits from satisfying (typically) rich agents' implementability constraint, so general fiscal motives do not affect the opportunity cost from corrective taxation in this case.

Timing of abatement and damages Going back to the pollution tax formula (18), the marginal implementation cost may imply deviations from the social cost of pollution for two reasons. First, a positive cost in period t means that that the opportunity cost of pollution abatement is lower in that period, which pushes the tax above the social cost of pollution. This effect is captured by the denominator of the formula. Second, a positive cost in period $t+j$ also means that having less production damages in that period is worth less, which pushes the tax below the social cost of pollution. This effect is captured by the numerator of the term multiplying production damages.

Focusing on production damages, we see that the marginal implementation cost operates as a form of discounting. If this cost increases over time, consumption is valued relatively more in the present than in the future, hence the pollution tax is set at a lower level. Conversely, a declining path for this term implies a higher tax. Turning to the utility part, the effect is again ambiguous and the tax is set to a higher (resp. lower) level to the extent that the marginal implementation cost is positive (resp. negative) in periods where the present value of utility damages are high.

Differences in allocations When the marginal implementation cost is null, the first and second best tax formulas coincide, and they are both equal to the social cost of pollution. Still, the actual tax levels may differ for two reasons.

The first reason is that when the tax system is different, aggregate variables generally take different values. When capital and labor are taxed, labor supply and investments are expected to be lower, hence output, consumption, and pollution are also expected to be lower along the optimal path. Since the pollution tax level is determined by the trade-off between the marginal utility of consumption and the marginal utility of pollution abatement, if both pollution and consumption are lower, the optimal tax will generally be set at a lower level since utility is concave in consumption and convex in pollution.[5]

The second reason is that the distribution of individual allocations also differs depending on the fiscal environment. Because individualized lump-sum transfers are not feasible in the second best, there are generally more consumption inequalities. The welfare gains from leaving more resources available for agents' consumption by decreasing the pollution tax may then be higher or lower compared to the first best depending on the curvature of agents' utility function. In the presence of inequalities, an increase in aggregate consumption is valued more to the extent that the average marginal utility is higher (by concavity of the utility function), but it is valued less to the extent that the inflow in consumption disproportionately goes in the hands of richer households with lower marginal utilities. Our analysis of functional form expressions below provides an illustration of these two opposite mechanisms.

3.4 Functional form expressions

Specification In the next section, we quantitatively analyze the optimal fiscal policies presented above. Before turning to these quantitative results, it is useful to investigate the theoretical predictions using the functional form for utility chosen in our quantitative analysis. Suppose agents have preferences over consumption, leisure and environmental degradation, with the following period utility function

$$u\left(c_i, h_i, Z\right) = \frac{\left(c_i(1-\varsigma h_i)^{\gamma}\right)^{1-\sigma}}{1-\sigma} + \frac{\left(1+\alpha_0 Z^2\right)^{-(1-\sigma)}}{1-\sigma}. \tag{22}$$

Capital and labor income taxes Without loss of generality, we can add a normalization constraint for the market weights to the Ramsey problem presented above (see appendix). We can then express the formulas for labor and capital income taxes as follows

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

and

$$\frac{R_{t+1}}{R^*_{t+1}} = \frac{\Phi - \Psi\varsigma\gamma\left(1-\sigma\right)\left(1-\varsigma H_{t+1}\right)^{-1}}{\Phi - \Psi\varsigma\gamma\left(1-\sigma\right)\left(1-\varsigma H_t\right)^{-1}},$$

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

with

$$\Phi = \sum_j \pi_j \frac{\lambda_j}{\varphi_j} + \big(1 - (1+\gamma)(1-\sigma)\big)\mathrm{cov}(\lambda_i/\varphi_i, \omega_i),$$

$$\Psi = -\frac{\mathrm{cov}(\lambda_i/\varphi_i, e_i)}{\varsigma},$$

where $\forall t$, $\omega_i = c_{i,t}/C_t$. We see that both the labor and the capital income tax rates are zero 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.

Marginal implementation cost Using our specification, we can also further examine the determinants of the marginal implementation cost that enters the pollution tax formula. First, using the first order condition of the Ramsey problem with respect to market weights, we can express the multipliers of the implementability constraints as

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

which means that the cost of implementing the planner's preferred allocation is positive (i.e. $-\theta_i$ is positive) for households who are valued relatively less by the market than by the planner as compared to an average household. Second, we can now write

$$-\sum_i \pi_i \theta_i MIC_{i,t} = (\sigma - 1)\frac{\mathrm{cov}(\theta_i, IC_{i,t})}{C_t}, \tag{23}$$

from which we see that the marginal implementation cost is always null if $\sigma = 1$. As shown in appendix, when more resources are available for agents' consumption, not only their consumption and real wage go up, but the price also goes down. When $\sigma = 1$, the price decline exactly offsets the increase in volume: expenditures and labor income remain unchanged, so that the planner has no need to adjust its transfer to ensure that the implementability constraints are satisfied. When $\sigma > 1$, the price effect dominates, so that an increase in aggregate consumption reduces the total amount of transfers needed to satisfy agents' implementability constraints. In periods when poor households (low θ_i) consume relatively more compared to what they earn (high $IC_{i,t}$), the aggregate implementation cost is positive and a reduction in transfers reduces the costs. Conversely, if the volume effect dominates ($\sigma < 1$), *or* if rich households temporarily consume relatively more compared to what they earn (low $\text{cov}(\theta_i, IC_{i,t})$), an increase in aggregate consumption increases implementation costs.

As shown in appendix, our utility function implies that individuals' consumption is a constant fraction of aggregate consumption, but individuals' labor supply is not proportional to aggregate labor, hence we cannot write $IC_i(C_t, H_t)$ as in (21). In particular, when transfers are positive (as they are in our quantitative analysis) less productive households work relatively less when aggregate labor supply is high, *i.e.*, for i, j such that $e_i > e_j$, $h_{i,t}/h_{j,t}$ is increasing in H_t. Under the assumption that higher productivity types also have a lower marginal utility of consumption, and thus a higher θ_i, the covariance term in equation (23) is positive (resp. negative) when the aggregate labor supply is relatively low (resp. high), and the marginal implementation cost is positive (resp. negative) if and only if increasing aggregate consumption decreases the amount of transfers necessary to satisfy agents' implementability conditions ($\sigma > 1$).

Marginal utility of consumption Under our functional form assumption, we can also sign the effect of inequalities on the marginal utility of consumption ($V_{C,t}$) as a function of the value of σ, which captures the utility curvature. In particular, when $\sigma = 1$, the increase in agents' average marginal utility exactly offsets the fact that higher marginal utility agents receive a lower share of aggregate consumption increases, hence an aggregate increase in consumption is valued identically in the first and second best. When the utility is more (resp. less) concave, consumption is valued more (resp. less) to the extent that there are more inequalities, shifting the second best pollution tax downward (resp. upward) compared to the first best where individualized transfers are used to reduce consumption inequalities (see Appendix A.4.3).

4 Quantitative analysis

This section explores quantitatively the implications of heterogeneity in productivity for the optimal taxation of carbon, capital income, and labor income. As in Barrage (2019), we consider a climate-economy model based on Nordhaus' DICE model. While Barrage (2019) considers a planner setting taxes for the global economy, we adopt a slightly different approach: we consider a global economy with the economic features of the U.S. economy, *i.e.* we parametrize the income per capita, the productivity distribution, and the fiscal system to match U.S. data, but we scale our economy so that output and emissions match global data. The objective is to determine how an economy with important inequalities and responsible for a significant share of global emissions like the U.S. should design his fiscal system if it were to internalize the global impact of its emissions, assuming that the rest of the world would behave identically.

4.1 Calibration

Below we describe our calibration choices. Because in our baseline model we also allow for an expropriatory wealth tax, we first demonstrate results without heterogeneity in asset holdings. A potential micro-foundation is that wealth and productivity are positively correlated, which makes it optimal for the planner to tax all initial wealth at a confiscatory rate. In Section 5, we consider further constrained environments where the planner cannot optimize over all fiscal instruments, as well as additional sources of heterogeneity. When the capital income tax is exogenously fixed, the planner is no longer able to fully expropriate wealth in the first period and the initial distribution of assets matters.

4.1.1 Climate model

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

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

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

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

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

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

where χ_t^{ex} is exogenous forcing. A positive radiative forcing means that the Earth receives more energy from the sun than it emits back to space, hence the climate warms.

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

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

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

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

4.1.2 Damages

We also model production damages as in DICE 2016, with

$$D_t = a_1 Z_t^{At} + a_2 (Z_t^{At})^{a_3}. \tag{24}$$

Since DICE does not distinguish between production and utility damages, we follow Barrage (2019) to decompose the damages from DICE into a production and a utility component. We apply her decomposition and assign 74% of damages at 2.5°C warming to output, and 26% to utility. This provides an adjusted value for the parameter a_2 in equation (24), and enables us to determine the preference parameter α_0.

[To be included: alternative damage scenario]

4.1.3 Households

The inter-temporal population-augmented aggregate utility as computed using market weights is

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

where $c_t \equiv C_t/N_t$ and $h_t \equiv H_t/N_t$ are the per capita average consumption and labor supply, and $Z_t \equiv Z_t^{At}$ is the atmospheric temperature. To ensure that aggregate emissions remain consistent with DICE, we calibrate the growth rate of population accordingly. Because we also want to match the GDP per capita of the U.S., we set the population levels as U.S. population multiplied by the ratio of world to U.S. GDP in 2011-2015.

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

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, we use the SCF 2013. We divide the sample of working households into ten groups of hourly wage deciles (*i.e.*, $I = 10$, and for all i, $\pi_i = 0.1$), with an hourly wage of \$6.44 for the bottom productivity group and \$101.35 for the top productivity group, and normalize productivity levels such that $\sum_i \pi_i e_i = 1$. The full procedure is described in appendix.

4.1.4 Production

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

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

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

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

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

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

where $c_{1,t}c_2 = P_t^{\text{backstop}}$ represents the backstop price, *i.e.* the price at which it becomes economical to abate 100% of emissions. As in DICE 2016, we assume that this price is \$550/tCO_2 in the initial period, and declines at a rate of 0.5% per year. We also calibrate the exponent $c_2 = 2.6$ as in DICE.

4.1.5 Government

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

To calibrate fiscal policy, we first require the empirical counterparts of taxes. In the model, there are four taxes: a tax $\tau_{K,t}$ on capital income, a tax $\tau_{H,t}$ on labor income, an excise (intermediate-goods) tax $\tau_{I,t}$ on total energy and a tax $\tau_{E,t}$ on pollution emissions. We set the tax rates on capital and labor income in line with Trabandt and Uhlig (2012), who conduct a detailed analysis of fiscal policies in the U.S. and a number of European countries. Using a comprehensive measure of taxes on capital income, they find that on average, capital income in the U.S. is taxed at a rate of 41,4%, hence we set a time-invariant $\tau_K = 0.411$ in our baseline.[6] They find that labor income in turn, is taxed at a rate of

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

22.1%. Combined with a tax rate on consumption of 4.6%, this translates into an effective tax rate on labor income of 25.5%, or $\tau_H = 1 - (1 - 0.221)/(1 + 0.046) = 0.255$.

Turning to energy taxes, from the final-goods firm's problem it is clear that τ_I only affects decisions on total energy usage. We follow Barrage (2019) and set the intermediate-goods tax at $\tau_I = 0$. Regarding the tax on pollution emissions τ_E, we do not set it to zero as in Barrage (2019). Instead, we set it at a level so that, in our calibrated economy, 3% of total energy is obtained from clean technologies (Nordhaus, 2017). This requires $\tau_E = 2.01\$/tCO_2$ in 2015. The reason for deviating from Barrage (2019) is that setting $\tau_E = 0$ in our baseline implies a business-as-usual (BAU) scenario where no policies are implemented to mitigate climate change, which seems unrealistic.

To calibrate initial, outstanding debt B_0 at the start of the economy, we calculate the difference between total liabilities and financial assets from the U.S. government's balance sheet, both as a percentage of GDP.[7] Following Barrage (2019) and in order to facilitate reproducing results for other countries, these data are obtained from the IMF Government Finance Statistics. This gives an average debt-to-GDP ratio of approximately 111% over the period 2011–2015. Because in our model a period corresponds to five years, we set $B_0/Y_{1,0} = 1.11/5 = 0.222$, or 22.2%.

Lastly, we require an empirical counterpart of government spending. In our model, G_t denotes government consumption of the final good, while T captures the present value of all lump-sum transfers households receive from the government. To better align the model with the data and to analyze business-as-usual scenarios, we follow Barrage (2019) and split up total government spending into final good spending G_t^C and *exogenous* transfers G_t^T that are provided to households. The total transfers households receive thus consist of this exogenous component G_t^T and the endogenous component T.[8] To obtain the empirical counterparts of G_t^C and G_t^T, we proceed as in Barrage (2019) and collect data on U.S. government expenses from the IMF Government Finance Statistics. Averaging over the years 2011–2015, government consumption is $G_0^C/Y_{1,0} = 0.158$, or 15.8%, while government transfers are $G_0^T/Y_{1,0} = 0,145$, or 14.5%.[9] 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.

4.2 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.

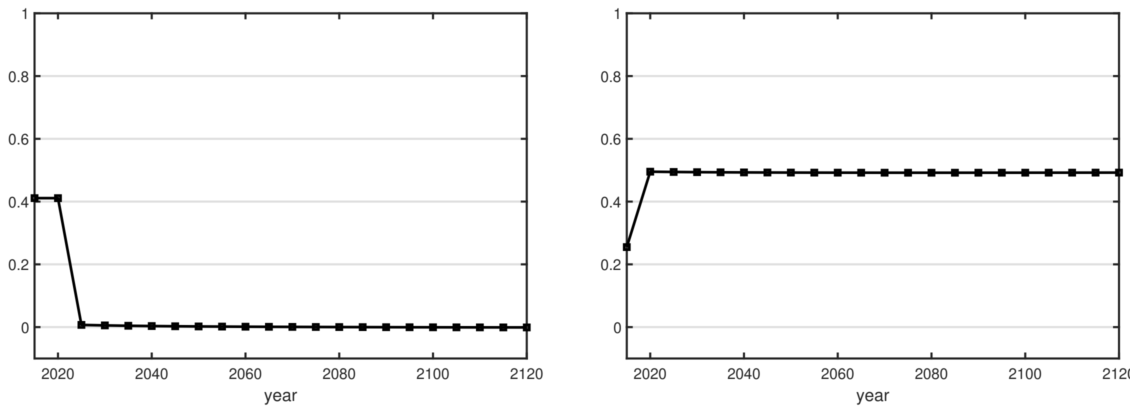
[7] The numbers are calculated at the "General Government" level.

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

[9] As in Barrage (2019), 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 expense. For transfers G_t^T, we include social benefits.

4.2.1 Optimal policy

Figure 1 shows the path of optimal taxes on capital and labor income. Because all wealth is expropriated initially, the capital income tax converges to zero in the second period and remains at this level thereafter. The next section examines scenarios with further constraints on policy instruments leading to deviations from this result. The labor income tax significantly goes up in the first period, at close to 50%, and stabilizes at this level. Figure 2 shows the optimal path of carbon taxes: in the baseline scenario, the tax starts at 21.5\$/tCO_2 in 2020 and goes up to reach 227.3\$/tCO_2 a century later.



(a) Optimal Capital-Income Taxes

(b) Optimal Labor-Income Taxes

Figure 1: Optimal Income Taxes

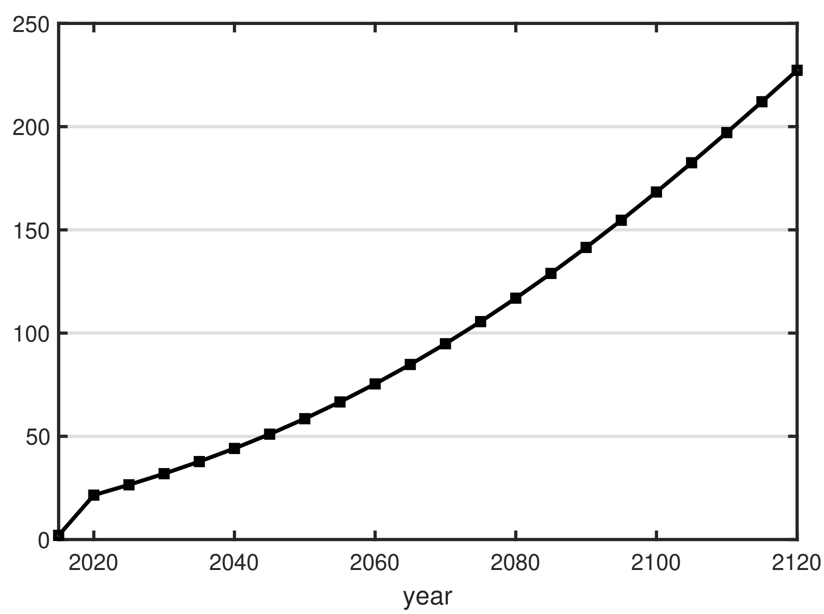


Figure 2: Optimal Carbon Taxes (\$/tCO_2)

Figure 3 compares the second best pollution tax normalized to 1 (black line) to what it would be if the marginal implementation cost was null in all periods (red line)—which also corresponds to the social cost of carbon (SCC) evaluated at the second best allocation—and to what it would be if all agents had equal marginal utilities of consumption (blue line). The marginal implementation cost appears to play an insignificant role: the social cost of carbon is only 0.5% above the second best carbon tax in the initial period, a difference that becomes even smaller in subsequent periods. Thus, even in the presence of distortionary taxation, it is optimal to set the carbon tax "almost" at the social cost of carbon. However, the discrepancy between the blue and red lines indicates that the social cost of carbon itself is significantly affected by the presence of inequalities. The reason is that the social cost of carbon represents the monetary value of climate damages, and is determined by the arbitrage between reducing damages and increasing aggregate consumption. As explained in Section 3, a marginal unit of aggregate consumption is valued more in the presence of inequalities if the marginal utility is sufficiently declining in consumption. Intuitively, an increase in aggregate consumption is valued less to the extent that it disproportionately goes in the hand of richer households, but it is valued more to the extent that the average marginal utility becomes higher if some people have relatively low consumption levels. In particular, with our specification, consumption inequalities call for lower carbon taxes for $\sigma > 1$. With $\sigma = 1.45$, we see that ignoring consumption inequalities would lead to a social cost of carbon higher by on average 4.2% over the next century.

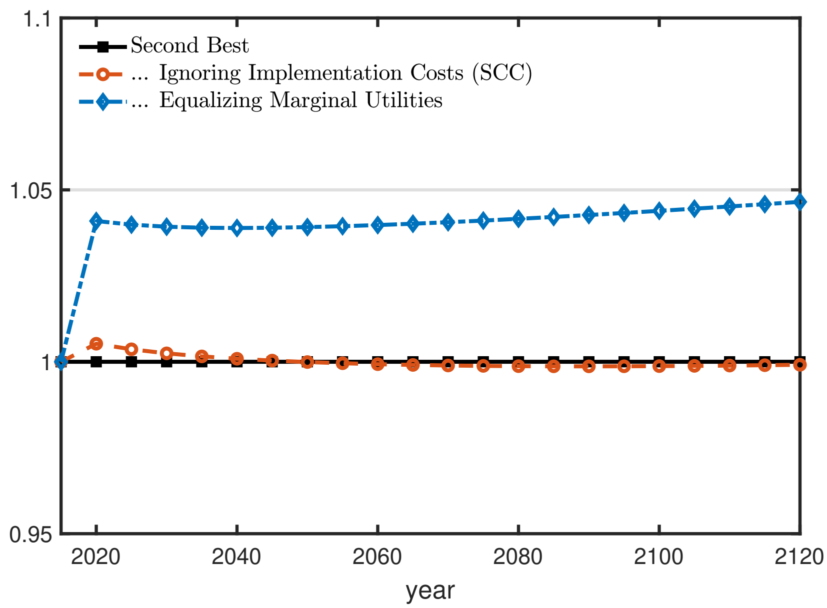


Figure 3: Carbon Tax Decomposition

Table I below reports the adjustments made to the government budget between our baseline second best scenario and a "climate skeptic scenario" in which the planner ignores the anthropogenic origin of climate change. Specifically, this climate skeptic planner sets all taxes optimally but behave as if the climate variable was exogenous and not driven by human-made emissions. The objective of this experiment is to see how the planner should adjust the fiscal system once it acknowledges the necessity to address climate change. [To be completed]

Table I: Government Budget Adjustment

	Revenue Source			Revenue Use		
	Labor-Inc. Tax	Capital-Inc. Tax	Carbon Tax	Gov. Cons.	Transfers	Interest
No Carbon Tax	0.335	0.006	0.000	0.172	0.146	0.023
Optimal Carbon Tax	0.329	0.006	0.010	0.171	0.151	0.023
Change	-1.8%	-0.6%		-0.8%	3.5%	1.2%

4.2.2 Welfare effects

Figure 4 displays the increase in consumption that would be necessary in the climate skeptic scenario to make households as well-off as in the optimal scenario in each period and for each productivity group. While the average long run gains are positive for all productivity groups (the average discounted gain is 4.1%), the "myopic" period welfare gains are heterogeneous over time and between groups. While the increase in the lump-sum transfer initially benefits poor households relatively more, in the long-run the decrease in the labor income tax benefits rich households relatively more. Overall, welfare gains are progressive but mostly negative in the 21$^{\text{st}}$ century and positive but regressive after 2100. [To be completed]

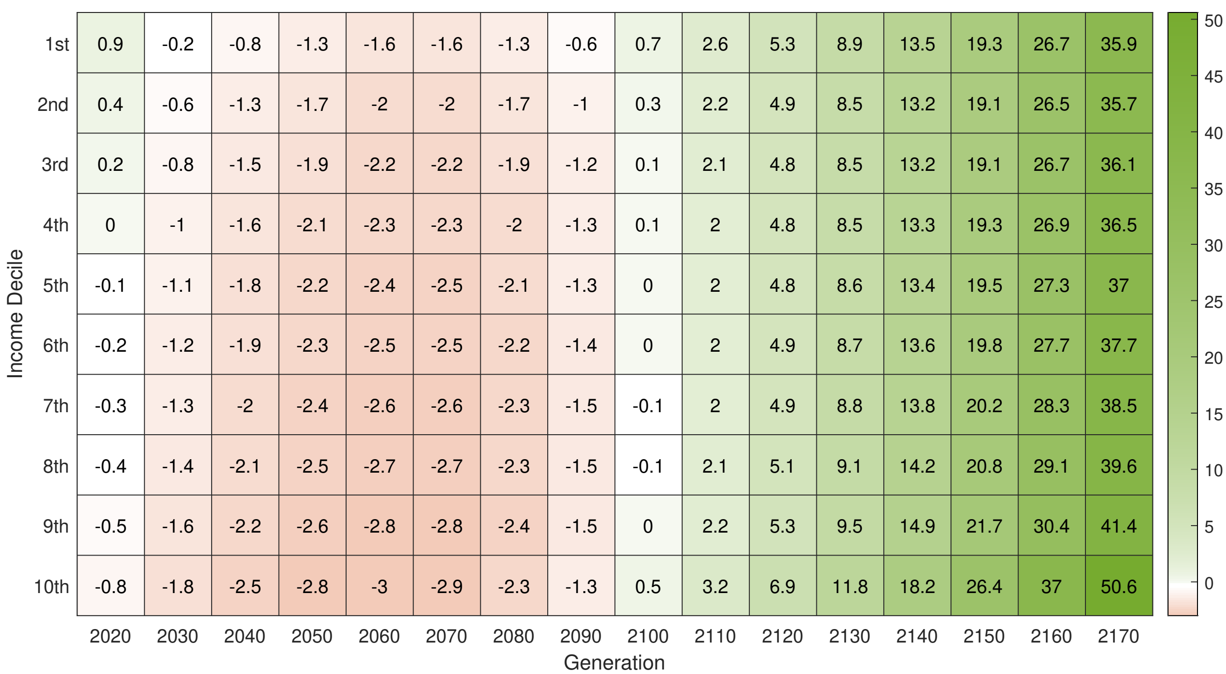


Figure 4: Myopic Welfare Gains (%)

5 Extensions

5.1 Business-as-usual scenarios

We have considered a Ramsey problem in which the government faces two key constraints: only linear and anonymous instruments can be used. Still, this set of fiscal instruments confers a lot of power to the government, arguably more than what most governments have. When introducing a carbon tax policy, a government may not have complete freedom to adjust labor or capital income taxes. In particular, the full expropriation of asset holdings in the initial period that is optimal in our benchmark is not a realistic policy option. To explore these issues, we now turn to fiscal environments with additional constraints on the set of available instruments.

5.1.1 Exogenous labor income tax

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

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

Let $\beta^t \Lambda_t^H$ denote the multiplier on the constraint (25). In addition, let $\beta^t \Omega_t^H$ be the multiplier on the constraint that the marginal rate of technical substitution between capital and labor is the same across both sectors.[10] As shown in appendix, the constrained second-best optimal capital income tax in this scenario is given by

[to be included]

and the constrained second-best optimal pollution tax becomes

[to be included]

5.1.2 Exogenous capital income tax

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

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

[10] Without condition (25), one can show that the multiplier associated with this constraint is optimally zero: the government does not wish to distort production decisions, which also explains why $\tau_{I,t} = 0$. With additional restrictions on the tax system, this is no longer generally true.

Let $\beta^t\Lambda^K_{t+1}$ be the multiplier on this constraint and $\beta^t\Omega^K_t$ the multiplier on the constraint that the marginal rate of technical substitution between capital and labor is the same across sectors. As shown in appendix, the constrained second-best optimal labor income tax in this scenario becomes

[to be included]

and the constrained second-best optimal pollution tax becomes

[to be included]

5.2 Additional sources of heterogeneity

5.2.1 Optimal tax rules

Model Our benchmark model considers heterogeneous agents who differ in productivity and initial asset holdings. To further explore the role of agents heterogeneity on optimal fiscal policy, we now introduce two additional ingredients to our benchmark model: a second consumption good, and heterogeneous preferences. We assume that a household of type i derives utility from the consumption of a final good $c_{i,t}$, labor supply $h_{i,t}$, environmental degradation Z_t, and the consumption of a "dirty" good $d_{i,t}$ according to a utility function

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

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

$$E_t = E_{1,t} + D_t,$$

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

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

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

$\bar{d}_i$, this specification also allows us to consider non-linear *aggregate* Engel curves as well as horizontal heterogeneity.[11,12] In addition, we assume that agents' relative sensitivity to the environmental variable Z is also type specific and denoted χ_i, normalized such that $\sum_i \pi_i \chi_i = 1$.

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

$$\sum_{t=0}^{\infty} p_t \Big(c_{i,t} + d_{i,t}(p_{E,t} + \tau_{D,t}) - (1 - \tau_{H,t})\, w_t e_i h_{i,t} \Big) \leq R_0 a_{i,0} + T. \tag{26}$$

Solution method We apply the same solution method as in our benchmark model. Using the method of Werning (2007), we can express individual allocations as a function of aggregate variables and market weights. These expressions allow us to write the aggregate utility function $U(C_t, D_t, H_t, Z_t, \varphi)$ and individual implementability conditions necessary to solve the Ramsey problem based on aggregate variables and market weights only.

Optimal taxes As shown in appendix, the second best labor income tax in this extended framework is

$$\tau_{H,t} = \frac{\Psi \varsigma (1 - \varsigma H_t)^{-1}}{\Phi + \Psi \frac{\varsigma \left(1 - \gamma(1-\sigma)\right)}{(1 - \varsigma H_t)} - \Lambda_t \frac{\epsilon(\sigma - 1)}{(D_t - \bar{D}_t)}},$$

the capital income can be obtained from

$$\frac{R_{t+1}}{R^*_{t+1}} = \frac{\Phi + \Psi \frac{\varsigma\gamma(\sigma-1)}{(1-\varsigma H_{t+1})} - \Lambda_{t+1} \frac{\epsilon(\sigma-1)}{(D_{t+1} - \bar{D}_{t+1})}}{\Phi + \Psi \frac{\varsigma\gamma(\sigma-1)}{(1-\varsigma H_t)} - \Lambda_t \frac{\epsilon(\sigma-1)}{(D_t - \bar{D}_t)}},$$

the excise tax on energy remains unchanged at $\tau_{I,t} = 0$, and the households energy consumption excise tax is

$$\tau_{D,t} = \frac{\Lambda_t \epsilon C_t}{\Phi + \frac{\Psi\varsigma\gamma(\sigma-1)}{(1-\varsigma H_t)} - \frac{\Lambda_t \epsilon(\sigma-1)}{(D_t - \bar{D}_t)}},$$

with

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

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

[12] Horizontal heterogeneity arises when individuals 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 (Cronin et al., 2019; Pizer and Sexton, 2019), and their implications for the design of tax reforms (Sallee, 2019).

Turning to the pollution tax, we obtain the same general formula as in our benchmark model

$$\tau_{E,t} = \sum_{j=0}^{\infty} \beta^j \left(\frac{V_{C,t+j} + \sum_i \pi_i \theta_i MIC_{i,t+j}}{V_{C,t} + \sum_i \pi_i \theta_i MIC_{i,t}} D'\left(Z_{t+j}\right) A_{1,t+j} F\left(K_{1,t+j}, H_{1,t+j}, E_{t+j}\right) \right.$$
$$\left. - \frac{V_{Z,t+j}}{V_{C,t} + \sum_i \pi_i \theta_i MIC_{i,t}} \right) J_{E_i^M, t+j}.$$

Comparison with the benchmark model Besides the differences in the path of allocations, the key differences with our benchmark lie in the expressions of the marginal implementation cost and the marginal utilities.

As shown in appendix, we can again express the marginal implementation cost as the covariance between θ_i and $MIC_{i,t}$, but this last term takes a different form. While in our benchmark the covariance was positive when increasing aggregate consumption led richer households (higher θ_i) to consume relatively more or work relatively less compared to poorer households, now its value is also higher when the energy needs of richer agents increase relative to poorer households. Using our functional form for utility, we again see that this additional energy demand effect would disappear if the relative energy consumption of two agents was constant over time, or simply unaffected by changes in aggregate consumption (which is the case when $\sigma = 1$).

The expression of the marginal utility of consumption ($V_{C,t}$) is also affected by the presence of a second consumption good, since utility is not strongly separable in C and D. Regarding the marginal dis-utility from pollution ($V_{Z,t}$), the energy consumption good has no direct impact, but heterogeneity in the relative sensitivity to the environmental variable captured by the distribution of χ_i may play a role. Indeed, we now have

$$V_{Z,t} = -(1 + \mathrm{cov}(\lambda_i, \chi_i)) 2\alpha_0 Z_t (1 + \alpha_0 Z_t^2)^{\sigma-2},$$

so that the distribution of χ_i matters in the marginal valuation of pollution to the extent that it is correlated with the planner's weights. When the agents most valued by the planner are more sensitive to pollution, the tax is set at a higher level. If we assume that the planner has utilitarian preferences however, then for all i, λ_i is constant and the distribution of χ_i has no impact on the aggregate marginal dis-utility from pollution.

The role of preferences heterogeneity In the special case where preferences are homogeneous (*i.e.* assuming that for $t \geq 0$ and for all i, $\bar{d}_{i,t} = \bar{d}_t$ and $\chi_i = 1$), we have $\Lambda_t = 0$, and all tax formulas remain unchanged relative to our benchmark. In particular, although poor households spend a larger share of their budget in energy, the pollution tax formula remains the same and the excise tax on energy consumption is null. 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 distribution of market weights is affected by the consumption of a second good: having a second good modeled as a necessity generates a fixed-cost to households welfare, which affects the whole distribution of welfare. Hence, even though the optimal tax formulas are preserved, the level of the tax rates will be affected by this additional source of heterogeneity as the formulas will be evaluated at a different allocation.

With heterogeneous preferences for energy consumption, Λ_t is not generally equal to zero anymore. When the consumption threshold ($\bar{d}_i$) varies positively with the relative planner's weight (λ_i/φ_i)—*i.e.* when individuals who are relatively more valued by the planner are also the ones with higher energy needs—then Λ_t is negative. This has two effects. First, it affects the implementation cost. The sign of this effect depends on the value of σ, which once again captures the price effect discussed above. For $\sigma > 1$, a negative Λ_t will lower the labor income tax, the pollution tax, and the absolute value of the excise tax on energy consumption. The second effect is captured by the numerator of the excise tax on energy consumption: when Λ_t is negative, this tax is negative. The logic behind this result is that the *aggregate* Engel curve being non-linear with heterogeneous preferences, commodity taxes offer an additional levy for redistribution. When the agents who are valued relatively more by the planner also have higher energy needs, the planner can target these agents by subsidizing the energy good.

The sign and magnitude of the previous mechanisms therefore depend on the distribution of $\{\bar{d}_i\}_{i\in I}$, both between and within productivity types. First, as less productive types tend to have higher marginal utilities of consumption, the relative planner's weights are generally higher for these agents. Λ will therefore be lower (resp. higher) to the extent that less (resp. more) productive agents have on average higher energy needs. Second, for a given productivity level, agents with higher energy needs will also tend to have higher marginal utilities of consumption because of the higher fixed cost that they incur. This horizontal heterogeneity will therefore drive the value of Λ downward. Our quantitative analysis below uses data on U.S. households energy consumption to illustrate the impact of these two sources of heterogeneity.

5.2.2 Quantitative analysis of the extended model

For each of the ten productivity groups described above, we compute the initial distribution of net worth from the SCF, and the distribution of energy needs from the Consumer Expenditure Surveys (CEX). The full procedures are described in appendix.

[To be included: quantitative analysis.]

6 Conclusion

What are the implications of heterogeneity in productivity and asset holdings for optimal carbon pricing? This paper attempts to shed light on this question in a climate-economy model, where environmental degradation generates both production and utility externalities. We extend the analysis from Barrage (2019) by including agent heterogeneity, which provides a micro-foundation for the use of distortionary taxes on labor and capital income. We study both theoretically and quantitatively how different sources of heterogeneity and a concern for redistribution affects the optimal carbon tax relative to the first-best (Pigouvian) scenario and relative to the case with a representative agent. [To be included: results.]

References

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

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

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

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

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

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

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

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

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

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

Fried, Stephie, Kevin Michael Novan, and William Peterman (2021) "Recycling Carbon Tax Revenue to Maximize Welfare."

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

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

Golosov, Mikhail, John Hassler, Per Krusell, and Aleh Tsyvinski (2014) "OPTIMAL TAXES ON FOSSIL FUEL IN GENERAL EQUILIBRIUM," *Econometrica*, 82 (1), 41–88, http://www.jstor.org/stable/24029171.

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

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

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

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

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

Kaplow, Louis (2012) "OPTIMAL CONTROL OF EXTERNALITIES IN THE PRESENCE OF INCOME TAXATION," *International Economic Review*, 53 (2), 487–509, http://www.jstor.org/stable/23251596.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Appendices

A Optimal tax rules in the benchmark model

A.1 Implementability conditions

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

$$U(C_t, H_t, Z_t; \varphi) \equiv \max_{c_{i,t}, h_{i,t}} \sum_i \pi_i \varphi_i u(c_{i,t}, h_{i,t}, Z_t), \quad \text{s.t.} \sum_i \pi_i c_{i,t} = C_t, \quad \text{and} \quad \sum_i \pi_i e_i h_{i,t} = H_t. \tag{27}$$

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. Assuming additive separability in Z and applying the envelope theorem to problem (27), we get

$$U_C(C_t, H_t) = \theta_t^c, \quad \text{and} \quad U_H(C_t, H_t) = -\theta_t^h.$$

From the first order conditions of problem (27), we also have

$$\varphi_i u_c(c_{i,t}, h_{i,t}) = \theta_t^c, \quad \text{and} \quad \varphi_i u_h(c_{i,t}, h_{i,t}) = -e_i \theta_t^h.$$

It follows that

$$U_C(C_t, H_t) = \varphi_i u_c(c_{i,t}, h_{i,t}), \tag{28}$$

$$U_H(C_t, H_t) = \frac{\varphi_i u_h(c_{i,t}, h_{i,t})}{e_i}. \tag{29}$$

In any competitive equilibrium these optimality conditions must hold for every agent i. Hence, using (28), (29), and agents' first order conditions given by

$$\beta^t \frac{u_c(c_{i,t}, h_{i,t})}{u_c(c_{i,0}, h_{i,0})} = p_t, \quad \forall\, t \geq 0, \tag{30}$$

$$\frac{u_h(c_{i,t}, h_{i,t})}{u_c(c_{i,t}, h_{i,t})} = -(1 - \tau_{H,t}) e_i w_t, \quad \forall\, t \geq 0, \tag{31}$$

we obtain

$$\frac{U_H(C_t, H_t)}{U_C(C_t, H_t)} = \frac{u_h(c_{i,t}, h_{i,t})}{u_c(c_{i,t}, h_{i,t}) e_i} = -w_t(1 - \tau_{H,t}), \tag{32}$$

and

$$\frac{U_C(C_t, H_t)}{U_C(C_0, H_0)} = \frac{u_c(c_{i,t}, h_{i,t})}{u_c(c_{i,0}, h_{i,0})} = \frac{p_t}{\beta^t}. \tag{33}$$

Given the relationships above we can derive the implementation condition which relies only on the aggregates C_t, H_t, and market weights φ. Let $c^m_{i,t}(C_t, H_t; \varphi)$ and $h^m_{i,t}(C_t, H_t; \varphi)$ be the arg max of problem (27). The budget constraint of agent i implies

$$\sum_{t=0}^{\infty} p_t \left(c^m_{i,t}(C_t, H_t; \varphi) - (1 - \tau_{H,t}) w_t e_i h^m_{i,t}(C_t, H_t; \varphi) \right) \leq R_0 a_{i,0} + T,$$

which using (32) and (33) can be restated as

$$U_C(C_0, H_0)\big(R_0 a_{i,0} + T\big) \geq \sum_{t=0}^{\infty} \beta^t \Big(U_C\big(C_t, H_t\big) c^m_{i,t}\big(C_t, H_t; \varphi\big) + U_H\big(C_t, H_t\big) e_i h^m_{i,t}\big(C_t, H_t; \varphi\big) \Big), \quad \forall\, i. \tag{34}$$

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

$$\begin{aligned} W(C_t, H_t, Z_t; \varphi, \theta, \lambda) \equiv & \sum_i \pi_i \lambda_i u\left(c^m_{i,t}(C_t, H_t; \varphi), h^m_{i,t}(C_t, H_t; \varphi), Z_t\right) \\ & + \sum_i \pi_i \theta_i \left[U_C(C_t, H_t) c^m_{i,t}(C_t, H_t; \varphi) + U_H(C_t, H_t) e_i h^m_{i,t}(C_t, H_t; \varphi) \right] \end{aligned}$$

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

$$\max_{\substack{\{C_t, H_{1,t}, H_{2,t}, K_{1,t}, K_{2,t}, \\ E_t, Z_t, \mu_t\}_{t=0}^{\infty}, T, \varphi, \tau_0^k \leq 1}} \sum_{t,i} \beta^t W(C_t, H_t, Z_t; \varphi, \theta, \lambda) - U_C(C_0, H_0) \sum_i \pi_i \theta_i (R_0 a_{i,0} + T)$$

subject to

$$\begin{aligned} C_t + G_t + K_{t+1} + \Theta_t(\mu_t, E_t) &= (1 - D(Z_t)) A_{1,t} F(K_{1,t}, H_{1,t}, E_t) + (1 - \delta) K_t, \quad \forall\, t \geq 0, \\ E_t &= A_{2,t} G(K_{2,t}, H_{2,t}), \quad \forall\, t \geq 0, \\ Z_t &= J\left(S_0, E^M_0, ..., E^M_t, \eta_0, ..., \eta_t\right), \quad \forall\, t \geq 0, \\ \frac{F_K(K_{1,t} H_{1,t}, E_t)}{F_H(K_{1,t} H_{1,t}, E_t)} &= \frac{G_K(K_{2,t} H_{2,t})}{G_H(K_{2,t} H_{2,t})}, \\ K_{1,t} + K_{2,t} &= K_t, \quad \forall\, t \geq 0, \\ H_{1,t} + H_{2,t} &= H_t, \quad \forall\, t \geq 0, \end{aligned}$$

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

A.2.2 First order conditions

The first order conditions are

$$[C_t]: W_C(C_t, H_t; \varphi, \theta, \lambda) - \nu_{1,t} = 0, \quad \forall\, t \geq 0, \tag{35}$$

$$[H_{1,t}]: W_H(C_t, H_t; \varphi, \theta, \lambda) + \nu_{1,t}(1 - D(Z_t))\, A_{1,t} F_H(K_{1,t}, H_{1,t}, E_t) = 0, \quad \forall\, t \geq 0, \tag{36}$$

$$[H_{2,t}]: W_H(C_t, H_t; \varphi, \theta, \lambda) + \nu_{2,t} A_{2,t} G_H(K_{2,t}, H_{2,t}) = 0, \quad \forall\, t \geq 0, \tag{37}$$

$$[K_{1,t+1}]: -\nu_{1,t} + \left[(1 - D(Z_{t+1}))\, A_{1,t+1} F_K(K_{1,t+1}, H_{1,t+1}, E_{t+1}) + (1-\delta)\right] \beta \nu_{1,t+1} = 0, \quad \forall\, t \geq 0, \tag{38}$$

$$[K_{2,t+1}]: -\nu_{1,t} + A_{2,t+1} G_K(K_{2,t+1}, H_{2,t+1})\, \beta \nu_{2,t+1} + (1-\delta)\, \beta \nu_{1,t+1} = 0, \quad \forall\, t \geq 0, \tag{39}$$

$$\begin{aligned}[E_t]: & -\nu_{1,t}\left(\Theta_{E,t}(\mu_t, E_t) - (1 - D(Z_t))\, A_{1,t} F_E(K_{1,t}, H_{1,t}, E_t)\right) - \nu_{2,t} \\ & - \sum_{j=0}^{\infty} \beta^j \nu_{3,t+j} J_{E_t^M, t+j}(1 - \mu_t) = 0, \quad \forall\, t \geq 0,\end{aligned} \tag{40}$$

$$[Z_t]: W_Z(C_t, H_t, Z_t; \varphi, \theta, \lambda) - \nu_{1,t} D'(Z_t)\, A_{1,t} F(K_{1,t}, H_{1,t}, E_t) + \nu_{3,t} = 0, \quad \forall\, t \geq 0, \tag{41}$$

$$[\mu_t]: -\nu_{1,t} \Theta_{\mu,t}(\mu_t, E_t) + \sum_{j=0}^{\infty} \beta^j \nu_{3,t+j} J_{E_t^M, t+j} E_t = 0, \quad \forall\, t \geq 0, \tag{42}$$

$$[T]: \sum_i \pi_i \theta_i = 0, \tag{43}$$

and at $t = 0$,

$$\left[\tau_0^k\right]: U_C(C_0, H_0)\left((1 - D(Z_0))\, A_{1,0} F_K(K_{1,0}, H_{1,0}, E_0) - \delta\right) \sum_i \pi_i \theta_i a_{i,0} = 0 \tag{44}$$

$$[K_{1,0}]: \left[(1 - D(Z_0))\, A_{1,0} F_K(K_{1,0}, H_{1,0}, E_0) + (1-\delta)\right] \nu_{1,0} - \kappa = 0 \tag{45}$$

$$[K_{2,0}]: A_{2,0} G_K(K_{2,0}, H_{2,0})\, \nu_{2,0} + (1-\delta)\, \nu_{1,0} - \kappa = 0 \tag{46}$$

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

$$(1 - D(Z_0))\, A_{1,0} F_K(K_{1,0}, H_{1,0}, E_0)\, \nu_{1,0} = A_{2,0} G_K(K_{2,0}, H_{2,0})\, \nu_{2,0},$$

which together with (36) and (37), implies that

$$\frac{F_K(K_{1,0}, H_{1,0}, E_0)}{F_H(K_{1,0}, H_{1,0}, E_0)} = \frac{G_K(K_{2,0}, H_{2,0})}{G_H(K_{2,0}, H_{2,0})}. \tag{47}$$

A.3 Optimal taxes

A.3.1 Capital and Labor income taxes

From (35) and (36) we obtain

$$(1 - D(Z_t))\, A_{1,t} F_H(K_{1,t}, H_{1,t}, E_t) = -\frac{W_H(C_t, H_t; \varphi, \theta, \lambda)}{W_C(C_t, H_t; \varphi, \theta, \lambda)}, \quad \forall\, t \geq 0, \tag{48}$$

and using the intertemporal condition (38) we get

$$R_{t+1}^{*} \equiv 1 + r_{t+1} - \delta = \frac{1}{\beta}\frac{W_C\left(C_t, H_t; \varphi, \theta, \lambda\right)}{W_C\left(C_{t+1}, H_{t+1}; \varphi, \theta, \lambda\right)}, \quad \forall\, t \geq 0, \tag{49}$$

These two equations can be used to back out the optimal taxes on labor and capital income. Plugging (48) into (32) implies

$$\frac{U_H\left(C_t, H_t\right)}{U_C\left(C_t, H_t\right)} = \frac{W_H\left(C_t, H_t; \varphi, \theta, \lambda\right)}{W_C\left(C_t, H_t; \varphi, \theta, \lambda\right)}\left(1 - \tau_{H,t}\right),$$

which can be rearranged into

$$\tau_{H,t} = 1 - \frac{U_H\left(C_t, H_t\right)}{U_C\left(C_t, H_t\right)}\frac{W_C\left(C_t, H_t; \varphi, \theta, \lambda\right)}{W_H\left(C_t, H_t; \varphi, \theta, \lambda\right)}. \tag{50}$$

In any competitive equilibrium (33) holds, which together with $p_t = R_{t+1}p_{t+1}$ implies

$$\frac{U_C\left(C_{t+1}, H_{t+1}\right)}{U_C\left(C_t, H_t\right)}\beta R_{t+1} = 1.$$

Substituting this into (49), it follows that

$$\frac{R_{t+1}}{R_{t+1}^{*}} = \frac{W_C\left(C_{t+1}, H_{t+1}; \varphi, \theta, \lambda\right)}{W_C\left(C_t, H_t; \varphi, \theta, \lambda\right)}\frac{U_C\left(C_t, H_t\right)}{U_C\left(C_{t+1}, H_{t+1}\right)}. \tag{51}$$

A.3.2 Excise taxes of energy and emissions

From the abatement first-order condition (42) and the energy firm abatement decision (9) we have that

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

From the climate variable first-order condition (41) we have that

$$\nu_{3,t} = \nu_{1,t} D'\left(Z_t\right) A_{1,t} F\left(K_{1,t}, H_{1,t}, E_t\right) - W_Z\left(C_t, H_t, Z_t; \varphi, \theta, \lambda\right),$$

hence the pollution tax is given by

$$\tau_{E,t} = \frac{1}{\nu_{1,t}}\sum_{j=0}^{\infty}\beta^j\left(\nu_{1,t+j} D'\left(Z_{t+j}\right) A_{1,t+j} F\left(K_{1,t+j}, H_{1,t+j}, E_{t+j}\right) - W_Z\left(C_{t+j}, H_{t+j}, Z_{t+j}; \varphi, \theta, \lambda\right)\right) J_{E_t^M,t+j}. \tag{52}$$

From the energy first-order condition (40) we have that

$$-\nu_{1,t}\Big(\Theta_{E,t}(\mu_t, E_t) + (1-\mu_t)\frac{\Theta_{\mu,t}(\mu_t, E_t)}{E_t} - (1 - D(Z_t))A_{1,t}F_E\left(K_{1,t}, H_{1,t}, E_t\right)\Big) = \nu_{2,t}, \tag{53}$$

and combining the first-order conditions for sectoral labor supplies (36) and (37), it follows that

$$\frac{\nu_{2,t}}{\nu_{1,t}} = \frac{\left(1 - D\left(Z_t\right)\right) A_{1,t} F_H\left(K_{1,t}, H_{1,t}, E_t\right)}{A_{2,t} G_H\left(K_{2,t}, H_{2,t}\right)}.$$

From (4) and (8) we also have

$$\frac{(1-D(Z_t))\,A_{1,t}F_H(K_{1,t},H_{1,t},E_t)}{A_{2,t}G_H(K_{2,t},H_{2,t})} = p_{E,t}-\tau_{I,t}-\tau_{E,t}(1-\mu_t)-\Theta_{E,t}(\mu_t,E_t)$$

hence using (5), (9), and (53) we have

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

and therefore

$$\tau_{I,t}=0. \tag{54}$$

A.4 Explicit formulas

A.4.1 Characterization of equilibrium

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

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

Using the period utility function defined in (22), the Lagrangian for the characterization problem defined by (15) is

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

The first order conditions are

$$[c_{i,t}]:\varphi_i\left(c_{i,t}\left(1-\varsigma h_{i,t}\right)^{\gamma}\right)^{1-\sigma}c_{i,t}^{-1}=\theta_t^c,\quad \forall\, t\geq 0, \tag{55}$$

$$[h_{i,t}]:\varphi_i\left(c_{i,t}\left(1-\varsigma h_{i,t}\right)^{\gamma}\right)^{1-\sigma}\gamma\varsigma\left(1-\varsigma h_{i,t}\right)^{-1}=e_i\theta_t^h,\quad \forall\, t\geq 0, \tag{56}$$

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 $H_t=\sum_i \pi_i e_i h_{i,t}$)

$$C_t=\left(\theta_t^c\left(\frac{\theta_t^h}{\theta_t^c}\frac{1}{\gamma\varsigma}\right)^{\gamma(1-\sigma)}\right)^{-\frac{1}{\sigma-(1-\sigma)\gamma}}\sum_i \pi_i\left(\frac{e_i^{\gamma(1-\sigma)}}{\varphi_i}\right)^{-\frac{1}{\sigma-(1-\sigma)\gamma}} \tag{57}$$

$$1-\varsigma H_t=\frac{\theta_t^c}{\theta_t^h}\gamma\varsigma\left(\theta_t^c\left(\frac{\theta_t^h}{\theta_t^c}\frac{1}{\gamma\varsigma}\right)^{\gamma(1-\sigma)}\right)^{-\frac{1}{\sigma-(1-\sigma)\gamma}}\sum_i \pi_i\left(\frac{e_i}{\varphi_i}^{\gamma(1-\sigma)}\right)^{-\frac{1}{\sigma-(1-\sigma)\gamma}} \tag{58}$$

It follows that

$$c_{i,t}^{m}\left(C_t,H_t;\varphi\right)=\omega_i C_t, \tag{59}$$

$$1-\varsigma h_{i,t}^{m}\left(C_t,H_t;\varphi\right)=\frac{\omega_i}{e_i}\left(1-\varsigma H_t\right), \tag{60}$$

where

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

Hence, we can write aggregate indirect utility $U\left(C_t,H_t,Z_t;\varphi\right)$ in terms of the aggregates C_t, H_t, and Z_t

$$U\left(C_t,H_t,Z_t,\varphi\right)=\sum_j \pi_j\varphi_j\left(\frac{\omega_j^{1+\gamma}}{e_j^{\gamma}}\right)^{1-\sigma}\frac{\left(C_t\left(1-\varsigma H_t\right)^{\gamma}\right)^{1-\sigma}}{1-\sigma}+\sum_i \pi_i\varphi_i\frac{\left(1+\alpha_0 Z_t^2\right)^{-(1-\sigma)}}{1-\sigma}, \tag{61}$$

$$=\frac{\left(C_t\left(1-\varsigma H_t\right)^{\gamma}\right)^{1-\sigma}}{1-\sigma}+\Gamma\frac{\left(1+\alpha_0 Z_t^2\right)^{-(1-\sigma)}}{1-\sigma}, \tag{62}$$

since from the normalization of market weights we have

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

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

A.4.2 Explicit tax formulas

From (34), substituting the derivatives of U into the definition of $W\left(C_t,H_t,Z_t;\varphi,\theta,\lambda\right)$ we get

$$\begin{aligned}W\left(C_t,H_t,Z_t;\varphi,\theta,\lambda\right)=&\sum_i \pi_i\lambda_i\left(\frac{\omega_i}{\varphi_i}\frac{\left(C_t\left(1-\varsigma H_t\right)^{\gamma}\right)^{1-\sigma}}{1-\sigma}+\frac{\left(1+\alpha_0 Z_t^2\right)^{-(1-\sigma)}}{1-\sigma}\right)\\&+\sum_i \pi_i\theta_i\left[\left(C_t\left(1-\varsigma H_t\right)^{\gamma}\right)^{1-\sigma}\omega_i-\gamma\left(C_t\left(1-\varsigma H_t\right)^{\gamma}\right)^{1-\sigma}\left(1-\varsigma H_t\right)^{-1}\left(e_i-\omega_i\left(1-\varsigma H_t\right)\right)\right]\end{aligned} \tag{63}$$

Collecting terms and simplifying we obtain

$$W\left(C_t,H_t,Z_t;\varphi,\theta,\lambda\right)=\Phi\frac{\left(C_t\left(1-\varsigma H_t\right)^{\gamma}\right)^{1-\sigma}}{1-\sigma}+\frac{\left(1+\alpha_0 Z_t^2\right)^{-(1-\sigma)}}{1-\sigma}+\Psi U_H\left(C_t,H_t\right). \tag{64}$$

where

$$\Phi\equiv\sum_i \pi_i\omega_i\left(\frac{\lambda_i}{\varphi_i}+\left(1-\sigma\right)\left(1+\gamma\right)\theta_i\right), \tag{65}$$

$$\Psi\equiv\sum_i \frac{\pi_i\theta_i e_i}{\varsigma}. \tag{66}$$

Substituting the derivatives into equation (50) we get

$$\tau_{H,t} = \frac{\Psi\varsigma\left(1-\varsigma H_t\right)^{-1}}{\Phi + \Psi\varsigma\left(1-\gamma\left(1-\sigma\right)\right)\left(1-\varsigma H_t\right)^{-1}}, \tag{67}$$

substituting the derivatives into (51) yields

$$\frac{R_{t+1}}{R^*_{t+1}} = \frac{\Phi - \Psi\varsigma\gamma\left(1-\sigma\right)\left(1-\varsigma H_{t+1}\right)^{-1}}{\Phi - \Psi\varsigma\gamma\left(1-\sigma\right)\left(1-\varsigma H_t\right)^{-1}}, \tag{68}$$

and substituting the derivatives into (52) we get

$$\tau_{E,t} = \frac{1}{\nu_{1,t}}\sum_{j=0}^{\infty}\beta^j\Big(\nu_{1,t+j}D'\left(Z_{t+j}\right)A_{1,t+j}F\left(K_{1,t+j},H_{1,t+j},E_{t+j}\right) - V_Z(Z_{t+j})\Big)J_{E_t^M,t+j}, \tag{69}$$

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

$$\nu_{1,t} = V_{C,t} + \sum_i \pi_i\theta_i MIC_{i,t}. \tag{70}$$

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

$$\sum_t \beta^t W_{\varphi_i}\left(C_t,H_t,Z_t;\varphi,\theta,\lambda\right) - \frac{\zeta}{\sigma-\left(1-\sigma\right)\gamma}\frac{\pi_i\omega_i}{\varphi_i} = 0, \quad \forall\ i.$$

From this equation we have that

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

and therefore

$$\frac{\lambda_i}{\varphi_i} + \theta_i = \frac{\zeta}{\left(1-\sigma\right)\left(1+\gamma\right)\sum_{t=0}^{\infty}\beta^t U\left(C_t,H_t\right)}, \quad \forall\ i.$$

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

$$\sum_j \frac{\pi_j\lambda_j}{\varphi_j} = \frac{\zeta}{\left(1-\sigma\right)\left(1+\gamma\right)\sum_{t=0}^{\infty}\beta^t U\left(C_t,H_t\right)},$$

and, therefore

$$\theta_i = \sum_j \frac{\pi_j\lambda_j}{\varphi_j} - \frac{\lambda_i}{\varphi_i}. \tag{71}$$

This allows us to rewrite

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

where the last result is obtained using the normalization of productivity levels, $\sum_i \pi_i e_i = 1$. The implementability conditions can be rewritten as

$$\omega_i = \frac{U_C(C_0, H_0)(R_0 a_{i,0} + T) + M e_i}{(1-\sigma)(1+\gamma)\sum_{t=0}^{\infty} \beta^t U(C_t, H_t)}, \quad \forall\ i, \tag{72}$$

with

$$M \equiv \sum_{t=0}^{\infty} \beta^t \gamma \left(C_t (1 - \varsigma H_t)^{\gamma} \right)^{1-\sigma} (1 - \varsigma H_t)^{-1}.$$

Since

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

we can express market weights as

$$\varphi_i = \frac{\omega_i^{\sigma-(1-\sigma)\gamma}}{e_i^{\gamma(\sigma-1)}} = \frac{1}{e_i^{\gamma(\sigma-1)}} \left(\frac{U_C(C_0, H_0)(R_0 a_{i,0} + T) + M e_i}{(1-\sigma)(1+\gamma)\sum_{t=0}^{\infty} \beta^t U(C_t, H_t)} \right)^{\sigma-(1-\sigma)\gamma}$$

A.4.3 Comparison with first best

First best pollution tax To compare our second-best results with the first best, we solve the same Ramsey problem except that we now allow for individualized lump-sum transfers. All first order conditions remain the same except for the one with respect to T given by (43): we now have

$$\theta_i = 0, \quad \forall i, \tag{73}$$

hence for all t, $\sum_i \pi_i \theta_i MIC_{i,t} = 0$. From (71), this also implies that

$$\frac{\lambda_i}{\varphi_i} = \sum_j \frac{\pi_j \lambda_i}{\varphi_i}, \quad \forall i, \tag{74}$$

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 (52), we can express the first best Pigouvian tax as

$$\tau_{E,t}^{FB} = \sum_{j=0}^{\infty} \beta^j \left(\frac{V_C(C_{t+j}, H_{t+j})}{V_C(C_t, H_t)} D'(Z_{t+j}) A_{1,t+j} F(K_{1,t+j}, H_{1,t+j}, E_{t+j}) - \frac{V_Z(Z_{t+j})}{V_C(C_t, H_t)} \right) J_{E_t^M, t+j}. \tag{75}$$

The marginal cost of funds Let us now decompose the first best tax rule into a production damage component and a utility damage component:

$$\tau_{E,t}^{FB,Y} = \sum_{j=0}^{\infty} \beta^j \frac{V_C(C_{t+j}, H_{t+j})}{V_C(C_t, H_t)} \Big(D'\left(Z_{t+j}\right) A_{1,t+j} F\left(K_{1,t+j}, H_{1,t+j}, E_{t+j}\right) \Big) J_{E_t^M,t+j} \ ,$$

$$\tau_{E,t}^{FB,U} = (-1) \sum_{j=0}^{\infty} \beta^j \frac{V_Z(Z_{t+j})}{V_C(C_t, H_t)} J_{E_t^M,t+j} \ .$$

If we define the marginal cost of funds as

$$\mathrm{MCF}_t \equiv \frac{\nu_{1,t}}{V_C\left(C_t, H_t; \varphi\right)},$$

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

$$\Delta_{t+s} \equiv \frac{\beta^j D'\left(Z_{t+s}\right) A_{1,t+s} F\left(K_{1,t+s}, H_{1,t+s}, E_{t+s}\right) J_{E_t^M,t+s}}{\sum_{j=0}^{\infty} \beta^j \Big(D'\left(Z_{t+j}\right) A_{1,t+j} F\left(K_{1,t+j}, H_{1,t+j}, E_{t+j}\right) J_{E_t^M,t+j} \Big)} \ ,$$

then the second best tax given by (69) can be re-written as a function of the marginal cost of funds and the first best tax rule evaluated at the second best allocation

$$\tau_{E,t} = \sum_{j=0}^{\infty} \frac{\mathrm{MCF}_{t+j}}{\mathrm{MCF}_t} \Delta_{t+j} \, \tau_{E,t}^{FB,Y}\Big|_{SB} + \frac{\tau_{E,t}^{FB,U}\Big|_{SB}}{\mathrm{MCF}_t} \ .$$

From (51), and using the fact that

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

we can also write the ratio of MCFs as

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

from which we see that the ratio is equal to 1 if the capital tax is null for all future periods where current emissions generate production damages. Thus, as in Barrage (2019), the optimal tax on production damage is not distorted as long as, going forward, the capital income tax is optimally set to zero. Substituting for $\nu_{1,t}$ in the definition of the MCF, we also see that

$$\mathrm{MCF}_t = 1 + \frac{\sum_i \pi_i \theta_i MIC_{i,t}}{V_C\left(C_t, H_t; \varphi\right)},$$

from which it appears that the MCF is equal to one when the implementation cost $-\sum_i \pi_i \theta_i MIC_{i,t}$ is null. In this situation, the second best pollution tax corresponds to the first best tax formula evaluated at the second best allocation.

The marginal implementation cost Using our functional form for U, we can show that

$$IC_{i,t} = \Big(C_t(1-\varsigma H_t)^{\gamma}\Big)^{(1-\sigma)} \left(\omega_i + \gamma\Big(\omega_i - \frac{e_i}{(1-\varsigma H_t)}\Big)\right), \tag{76}$$

from which we can write

$$MIC_{i,t} = (1-\sigma)\frac{IC_{i,t}}{C_t}. \tag{77}$$

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

$$-\sum_i \pi_i\theta_i MIC_{i,t} = (\sigma-1)\frac{\text{cov}(\theta_i, IC_{i,t})}{C_t}. \tag{78}$$

This term is equal to 0 when either $\sigma = 1$, or θ_i and $IC_{i,t}$ are uncorrelated.

Price effect To understand the role of σ, it is useful to go back to the origin of the the term $IC_i(C_t, H_t)$. This term comes from households' budget constraint (2) in which we have substituted for the price and real wage using (32) and (33). From these equations, it appears that when making more resources available to households, the price goes down since

$$p_t = \beta^t\Big(\frac{C_t}{C_0}\Big)^{-\sigma}\Big(\frac{1-\varsigma H_t}{1-\varsigma H_0}\Big)^{\gamma(1-\sigma)}.$$

When $\sigma = 1$, the price effect exactly offsets the volume effect so that households' expenditures and nominal income remain unchanged, hence the planner does not need to change the value of the lump-sum transfer and the implementation cost remains constant.

Labor supply effect To determine the sign of the covariance term, we can examine the ratio of the period implemantation cost for two agents i and j such that $e_i > e_j$. From (76), we have

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

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

$$\frac{\partial \frac{IC_{i,t}}{IC_{j,t}}}{\partial H_t} = \frac{\varsigma\gamma(1+\gamma)\big(e_j\omega_i - e_i\omega_j\big)}{(1-\varsigma H_t)^2\big(\omega_j(1+\gamma) - \frac{\gamma e_j}{(1-\varsigma H_t)}\big)^2}. \tag{79}$$

From (72), we can also show that with full wealth expropriation, when transfers 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 (79) is negative. This result means that when H_t is high relative to its average value, the relative labor supply of highly productive households compared to less productive households is higher, hence more productive households need lower transfers to satisfy the planners' allocation at that period. If the more productive also have a lower marginal utility of consumption (hence a higher θ_i), then the covariance term in equation (78) is negative when aggregate labor supply is relatively high. Conversely, when transfers are negative, the derivative in (79) is positive and the covariance term is negative when the aggregate labor supply is low.

Differences from individual allocations We can express the aggregate utility defined using the planner's weights as follows

$$V(C_t, H_t, Z_t; \varphi, \lambda) = \frac{\sum_i \pi_i \lambda_i u(c_{i,t}, h_{i,t})}{\sum_i \pi_i \varphi_i u(c_{i,t}, h_{i,t})} \frac{\left(C_t \left(1 - \varsigma H_t\right)^{\gamma}\right)^{1-\sigma}}{1-\sigma} + \frac{\left(1 + \alpha_0 Z_t^2\right)^{-(1-\sigma)}}{1-\sigma},$$

hence the marginal utility of consumption from the planner's perspective is

$$V_{C,t} = \frac{\sum_i \pi_i \lambda_i u(c_{i,t}, h_{i,t})}{\sum_i \pi_i \varphi_i u(c_{i,t}, h_{i,t})} U_{C,t}.$$

From our characterization problem, we know that market weights are determined by the following expression

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

from which we can rewrite

$$V_{C,t} = \sum_i \pi_i \lambda_i \frac{u_{c,i,t} c_{i,t}}{C_t} \tag{80}$$

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 $u(c, h)$ is concave in c, 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. Substituting $u_{c,i,t}$ by its functional expression in (80), we have

$$V_{C,t} = \sum_i \pi_i \lambda_i \frac{\left(c_{i,t}(1 - \varsigma h_{i,t})^{\gamma}\right)^{1-\sigma}}{C_t}$$

and we see that when $\sigma = 1$, 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 Stone-Geary utility and heterogeneous preferences

B.1 Characterization of equilibrium

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).

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

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

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

$$\begin{aligned} U\left(C_t, D_t, H_t, Z_t; \varphi\right) \equiv & \max_{c_{i,t}, d_{i,t}, h_{i,t}} \sum_i \pi_i \varphi_i u_i\left(c_{i,t}, d_{i,t}, h_{i,t}, Z_t\right), \\ \text{s.t.} \quad \sum_i \pi_i c_{i,t} = C_t, \quad \text{and} \quad & \sum_i \pi_i d_{i,t} = D_t, \quad \text{and} \quad \sum_i \pi_i e_i h_{i,t} = H_t. \end{aligned} \tag{81}$$

Following the same steps as in Appendix A, we obtain the following solutions for this problem

$$c^m_{i,t}\left(C_t, D_t, H_t; \varphi\right) = \omega_i C_t, \tag{82}$$

$$d^m_{i,t}\left(C_t, D_t, H_t; \varphi\right) = \bar{d}_{i,t} + \omega_i\big(D_t - \bar{D}_t\big), \tag{83}$$

$$1 - \varsigma h^m_{i,t}\left(C_t, D_t, H_t; \varphi\right) = \frac{\omega_i}{e_i}(1 - \varsigma H_t), \tag{84}$$

where

$$\omega_i = \left(\varphi_i e_i^{\gamma(\sigma-1)}\right)^{\frac{1}{1-(1+\epsilon+\gamma)(1-\sigma)}} \tag{85}$$

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

$$U\left(C_t, D_t, H_t, Z_t\right) = \frac{\left(C_t(D_t - \bar{D}_t)^{\epsilon}\left(1 - \varsigma H_t\right)^{\gamma}\right)^{1-\sigma}}{1-\sigma} + \Gamma_\chi \frac{\left(1 + \alpha_0 Z_t^2\right)^{-(1-\sigma)}}{1-\sigma}, \tag{86}$$

with $\Gamma_\chi \equiv \sum_i \pi_i \varphi_i \chi_i$.

B.2 Implementability condition

From the first order conditions of problem (81) and applying the envelope theorem we have

$$U_C\left(C_t, D_t, H_t\right) = \varphi_i u_{c,i}\left(c_{i,t}, d_{i,t}, h_{i,t}\right), \tag{87}$$

$$U_D\left(C_t, D_t, H_t\right) = \varphi_i u_{d,i}\left(c_{i,t}, d_{i,t}, h_{i,t}\right), \tag{88}$$

$$U_H\left(C_t, D_t, H_t\right) = \frac{\varphi_i u_{h,i}\left(c_{i,t}, d_{i,t}, h_{i,t}\right)}{e_i}, \tag{89}$$

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

$$\frac{U_H\big(C_t, D_t, H_t\big)}{U_C\big(C_t, D_t, H_t\big)} = \frac{u_{h,i}\left(c_{i,t}, d_{i,t}, h_{i,t}\right)}{u_{c,i}\left(c_{i,t}, d_{i,t}, h_{i,t}\right) e_{i,t}} = -w_t\left(1 - \tau_{H,t}\right), \tag{90}$$

$$\frac{U_D\big(C_t, D_t, H_t\big)}{U_C\big(C_t, D_t, H_t\big)} = \frac{u_{d,i}\left(c_{i,t}, d_{i,t}, h_{i,t}\right)}{u_{c,i}\left(c_{i,t}, d_{i,t}, h_{i,t}\right)} = p_{E,t} + \tau_{D,t}, \tag{91}$$

and

$$\frac{U_C\big(C_t, D_t, H_t\big)}{U_C\big(C_0, D_0, H_0\big)} = \frac{u_{c,i}\left(c_{i,t}, d_{i,t}, h_{i,t}\right)}{u_{c,i}\left(c_{i,0}, e_{i,0}, h_{i,0}\right)} = \frac{p_t}{\beta^t}, \tag{92}$$

where $u_{c,i}$ (resp. $u_{d,i}$, $u_{h,i}$) denotes the derivative of u_i with respect to c_i (resp. d_i, h_i). Using (90), (91), and (92) to substitute in households' budget constraint (26), we obtain the implementability conditions

$$\begin{aligned} U_C\big(C_0, D_0, H_0\big)\,(R_0 a_{i,0} + T) \geq \sum_{t=0}^{\infty} \beta^t \Big(& U_C\big(C_t, D_t, H_t\big) c_{i,t}^m\big(C_t, D_t, H_t; \varphi\big) \\ & + U_D\big(C_t, D_t, H_t\big) d_{i,t}^m\,(C_t, D_t, H_t; \varphi) \\ & + U_H\big(C_t, D_t, H_t\big) e_i h_{i,t}^m\,(C_t, D_t, H_t; \varphi) \Big), \quad \forall\ i. \end{aligned} \tag{93}$$

B.3 Ramsey problem

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

$$\begin{aligned} W\big(C_t, D_t, H_t, Z_t; \varphi, \theta, \lambda\big) \equiv & \sum_i \pi_i \lambda_i u_i\big(c_{i,t}^m\big(C_t, D_t, H_t; \varphi\big), d_{i,t}^m\big(C_t, D_t, H_t; \varphi\big), h_{i,t}^m\big(C_t, D_t, H_t; \varphi\big), Z_t\big) \\ & + \sum_i \pi_i \theta_i \Big[U_C\big(C_t, D_t, H_t\big) c_{i,t}^m\big(C_t, D_t, H_t; \varphi\big) \\ & + U_D\big(C_t, D_t, H_t\big) d_{i,t}^m\big(C_t, D_t, H_t; \varphi\big) + U_H\big(C_t, D_t, H_t\big) e_{i,t} h_{i,t}^m\big(C_t, D_t, H_t; \varphi\big) \Big], \end{aligned}$$

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

$$\max_{\substack{\{C_t, H_{1,t}, H_{2,t}, K_{1,t}, K_{2,t}, \\ D_t, E_{1,t}, Z_t, \mu_t\}_{t=0}^{\infty}, T, \varphi, \tau_0^k \leq 1}} \sum_{t,i} \beta^t W\left(C_t, D_t, H_t, Z_t; \varphi, \theta, \lambda\right) - U_C\left(C_0, D_0, H_0, Z_0\right) \sum_i \pi_i \theta_i \left(R_0 a_{i,0} + T\right),$$

subject to

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

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

B.4 Optimal taxes

Tax formulas From the first order conditions of the Ramsey problem, we can show that

$$\tau_{H,t} = 1 - \frac{U_H\left(C_t, D_t, H_t\right)}{U_C\left(C_t, D_t, H_t\right)} \frac{W_C\left(C_t, D_t, H_t; \varphi, \theta, \lambda\right)}{W_H\left(C_t, D_t, H_t; \varphi, \theta, \lambda\right)}, \tag{94}$$

$$\frac{R_{t+1}}{R^*_{t+1}} = \frac{W_C\left(C_{t+1}, D_{t+1}, H_{t+1}; \varphi, \theta, \lambda\right)}{W_C\left(C_t, D_t, H_t; \varphi, \theta, \lambda\right)} \frac{U_C\left(C_t, D_t, H_t\right)}{U_C\left(C_{t+1}, D_{t+1}, H_{t+1}\right)}, \tag{95}$$

$$\tau_{E,t} = \frac{1}{\nu_{1,t}} \sum_{j=0}^{\infty} \beta^j \left(\nu_{1,t+j} D'\left(Z_{t+j}\right) A_{1,t+j} F\left(K_{1,t+j}, H_{1,t+j}, E_{t+j}\right) - W_Z\left(C_{t+j}, H_{t+j}, Z_{t+j}; \varphi, \theta, \lambda\right)\right) J_{E^M_t, t+j}, \tag{96}$$

and

$$\tau_{I,t} = 0. \tag{97}$$

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

$$W_D\left(C_t, D_t, H_t, Z_t; \varphi, \theta, \lambda\right) = W_C\left(C_t, D_t, H_t, Z_t; \varphi, \theta, \lambda\right)(1 - D(Z_t)) A_{1,t} F_E(K_{1,t}, H_{1,t}, E_{1,t}),$$

which together with (91) and the final good firm's first order condition with respect to $E_{1,t}$ (given by (5) in the benchmark model) gives

$$\tau_{D,t} = \frac{U_D(C_t, D_t, H_t)}{U_C(C_t, D_t, H_t)} - \frac{W_D(C_t, D_t, H_t, Z_t; \varphi, \theta, \lambda)}{W_C(C_t, D_t, H_t, Z_t; \varphi, \theta, \lambda)}. \tag{98}$$

Using our functional form assumption, we can rewrite the pseudo-utility function as follows

$$W\left(C_t, D_t, H_t, Z_t; \varphi, \theta, \lambda\right) = \Phi U(C_t, D_t, H_t) + \frac{\sum_i \pi_i \lambda_i \chi_i}{\sum_i \pi_i \varphi_i \chi_i} U(Z_t) + \Psi U_H(C_t, D_t, H_t) + \Lambda_t U_D(C_t, D_t, H_t), \tag{99}$$

with

$$U(C_t, D_t, H_t) = \frac{(C_t (D_t - \bar{D}_t)^{\epsilon} (1 - \varsigma H_t)^{\gamma})^{1-\sigma}}{1-\sigma},$$
$$U(Z_t) = \Gamma_{\chi} \frac{(1 + \alpha_0 Z_t^2)^{-(1-\sigma)}}{1-\sigma},$$

where

$$\Phi \equiv \sum_i \pi_i \omega_i \Big(\frac{\lambda_i}{\varphi_i} + \big(1-\sigma\big)\big(1+\epsilon+\gamma\big)\theta_i\Big), \tag{100}$$

$$\Psi \equiv \frac{1}{\varsigma} \sum_i \pi_i \theta_i e_i, \tag{101}$$

$$\Lambda_t \equiv \sum_i \pi_i \theta_i \bar{d}_{i,t}. \tag{102}$$

Substituting the derivatives into equations (94), (95), and (98), we get

$$\tau_{H,t} = 1 - \frac{\Phi + \Psi\frac{U_{CH}}{U_C} + \Lambda_t\frac{U_{CD}}{U_C}}{\Phi + \Psi\frac{U_{HH}}{U_H} + \Lambda_t\frac{U_{DH}}{U_H}} = \frac{\Psi\varsigma(1-\varsigma H_t)^{-1}}{\Phi + \Psi\frac{\varsigma\left(1-\gamma(1-\sigma)\right)}{(1-\varsigma H_t)} + \Lambda_t\frac{\epsilon(1-\sigma)}{(D_t-\bar{D}_t)}}, \tag{103}$$

$$\frac{R_{t+1}}{R^*_{t+1}} = \frac{\Phi + \Lambda_{t+1}\frac{U_{CD_{t+1}}}{U_{C_{t+1}}} + \Psi\frac{U_{CH_{t+1}}}{U_{C_{t+1}}}}{\Phi + \Lambda_t\frac{U_{CD_t}}{U_{C_t}} + \Psi\frac{U_{CH_t}}{U_{C_t}}} = \frac{\Phi + \Lambda_{t+1}\frac{\epsilon(1-\sigma)}{(D_{t+1}-\bar{D}_{t+1})} - \Psi\frac{\varsigma\gamma(1-\sigma)}{(1-\varsigma H_{t+1})}}{\Phi + \Lambda_t\frac{\epsilon(1-\sigma)}{(D_t-\bar{D}_t)} - \Psi\frac{\varsigma\gamma(1-\sigma)}{(1-\varsigma H_t)}}, \tag{104}$$

and

$$\tau_{D,t} = \frac{\Lambda_t(D_t-\bar{D}_t)^{-1}U_D}{\Phi U_C + \Psi U_{HC} + \Lambda_t U_{DC}} = \frac{\Lambda_t\epsilon C_t}{\Phi + \frac{\Psi\varsigma\gamma(\sigma-1)}{(1-\varsigma H_t)} - \frac{\Lambda_t\epsilon(\sigma-1)}{(D_t-\bar{D}_t)}}. \tag{105}$$

If we define

$$V\left(C_t, D_t, H_t, Z_t; \varphi, \theta, \lambda\right) \equiv \sum_i \pi_i\lambda_i u_i\left(c^m_{i,t}\left(C_t, D_t, H_t; \varphi\right), d^m_{i,t}\left(C_t, D_t, H_t; \varphi\right), h^m_{i,t}\left(C_t, D_t, H_t; \varphi\right), Z_t\right)$$

and

$$\begin{aligned} IC_i(C_t, D_t, H_t) \equiv U_C\left(C_t, D_t, H_t\right)c^m_{i,t}\left(C_t, D_t, H_t; \varphi\right) + U_D\left(C_t, D_t, H_t\right)d^m_{i,t}\left(C_t, D_t, H_t; \varphi\right) \\ + U_H\left(C_t, D_t, H_t\right)e_{i,t}h^m_{i,t}\left(C_t, D_t, H_t; \varphi\right) \end{aligned}$$

we can also express the optimal pollution tax as

$$\begin{aligned} \tau_{E,t} = \sum_{j=0}^{\infty}\beta^j\Bigg(\frac{V_{C,t+j} + \sum_i \pi_i\theta_i MIC_{i,t+j}}{V_{C,t} + \sum_i \pi_i\theta_i MIC_{i,t}} D'\left(Z_{t+j}\right) A_{1,t+j} F\left(K_{1,t+j}, H_{1,t+j}, E_{t+j}\right) \\ - \frac{V_{Z,t+j}}{V_{C,t} + \sum_i \pi_i\theta_i MIC_{i,t}}\Bigg) J_{E^M_t, t+j}. \end{aligned}$$

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 implementation cost ($-\sum_i \pi_i\theta_i MIC_{i,t}$) 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

$$-\sum_i \pi_i\theta_i MIC_{i,t} = (\sigma-1)\frac{\mathrm{cov}(\theta_i, IC_{i,t})}{C_t},$$

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

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

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

The value of the optimal tax also depends on the marginal dis-utility from pollution ($V_{Z,t}$) which now accounts for the weights χ_i. In particular, we now have

$$\begin{aligned} V_{Z,t} =& -\sum_i \pi_i \lambda_i \chi_i 2\alpha_0 Z_t (1+\alpha_0 Z_t^2)^{\sigma-2} \\ =& -(1+\mathrm{cov}(\lambda_i,\chi_i))2\alpha_0 Z_t(1+\alpha_0 Z_t^2)^{\sigma-2} \end{aligned}$$

where the last result is obtained using the normalization of the sums of λ_i and χ_i. Thus, when the planner has utilitarian preferences, for all i, $\lambda_i = 1$ and the distribution of χ_i has no impact on the aggregate marginal dis-utility form pollution. When the planner values more (resp. less) agents with higher marginal dis-utility from pollution, then the tax is set at a higher (resp. lower) level.

Note that we can again use the first order conditions with respect to market weights to obtain

$$\theta_i = \sum_j \frac{\pi_j \lambda_j}{\varphi_j} - \frac{\lambda_i}{\varphi_i}, \tag{106}$$

from which we can rewrite

$$\begin{aligned} \Phi =& \sum_i \pi_i \omega_i \Big(\frac{\lambda_i}{\varphi_i} + \big(1-\sigma\big)\big(1+\epsilon+\gamma\big)\Big(\sum_j \frac{\pi_j\lambda_j}{\varphi_j} - \frac{\lambda_i}{\varphi_i}\Big)\Big) && (107) \\ =& \sum_j \pi_j \frac{\lambda_j}{\varphi_j} + \Big(1-(1+\epsilon+\gamma)(1-\sigma)\Big)\mathrm{cov}(\lambda_i/\varphi_i,\omega_i), && (108) \\ \Psi =& \frac{1}{\varsigma}\sum_i \pi_i \Big(\sum_j \frac{\pi_j\lambda_j}{\varphi_j} - \frac{\lambda_i}{\varphi_i}\Big) e_i && (109) \\ =& -\frac{\mathrm{cov}(\lambda_i/\varphi_i, e_i)}{\varsigma}, && (110) \\ \Lambda_t =& \sum_i \pi_i \Big(\sum_j \frac{\pi_j\lambda_j}{\varphi_j} - \frac{\lambda_i}{\varphi_i}\Big)\bar{d}_{i,t} && (111) \\ =& -\mathrm{cov}(\lambda_i/\varphi_i, \bar{d}_{i,t}), && (112) \end{aligned}$$

and obtain an expression for market weights

$$\varphi_i = \frac{1}{e_i^{\gamma(\sigma-1)}}\left(\frac{U_C(C_0,D_0,H_0)(R_0 a_{i,0}+T) + \sum_t \beta^t\Big(U_H(C_t,D_t,H_t)\frac{e_i}{\varsigma} - U_D(C_t,D_t,H_t)\bar{d}_{i,t}\Big)}{(1-\sigma)(1+\epsilon+\gamma)\sum_t \beta^t U(C_t,D_t,H_t)}\right)^{1-(1+\epsilon+\gamma)(1-\sigma)}$$

C Calibration

C.1 Household heterogeneity

Productivity distribution We calibrate the ability distribution on the basis of hourly wage data that we obtain from the Survey of Consumer Finances (SCF). For each of the 6,015 households in the

2013 wave of the survey, we sum the hours worked on their main job and potential additional job(s) in a normal week. Annual labor supply of the respondent and their partner is then calculated by multiplying weekly hours worked by 52 minus the number of weeks they have spent unemployed during the past 12 months minus the number of weeks spend on holidays (which we assume equals 3 for each worker). The household hourly wage is then obtained as the household annual income from wages and salaries before taxes, divided by the household total annual labor supply (i.e., the sum of the respondent and their partner's labor supply). This number reflects how much households members were paid on average for each hour of work they supplied in the past year.

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

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

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

Distribution of energy consumption Let us denote $X_{i,t}$ the expenditure share of energy for an household of type i at time t,

$$X_{i,t} \equiv \frac{d_{i,t}\big(p_{E,t} + \tau_{D,t}\big)}{d_{i,t}\big(p_{E,t} + \tau_{D,t}\big) + c_{i,t}}. \tag{113}$$

From the households' first order conditions we have

$$\frac{u_{d,i,t}}{u_{c,i,t}} = \frac{\epsilon c_{i,t}}{d_{i,t} - \bar{d_{i,t}}} = p_{E,t} + \tau_{D,t}, \tag{114}$$

with $u_{d,i,t}$, $u_{c,i,t}$ the marginal utility of energy and final good consumption of agent i at time t. Substituting the previous expression into (113) we get

$$X_{i,t} = \frac{d_{i,t}\big(\frac{\epsilon c_{i,t}}{d_{i,t} - \bar{d_{i,t}}}\big)}{d_{i,t}\big(\frac{\epsilon c_{i,t}}{d_{i,t} - \bar{d_{i,t}}}\big) + c_{i,t}}.$$

Table II: Distribution of households hourly wages and net worth by productivity deciles (rows) and net worth deciles (columns), controlling for generational differences.

	Net worth deciles										Hourly wage
Productivity deciles	1st	2nd	3rd	4th	5th	6th	7th	8th	9th	10th	
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 for each productivity group. Columns 1 to 10 correspond to net worth decile groups within productivity groups. The number reported in these columns are the average net worth for each group. 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.

Rearranging the previous equation, we can express the necessity parameter of agent i in period t ($\bar{d}_{i,t}$) as a function of its observed consumption level ($d_{i,t}$), its observed energy consumption share ($X_{i,t}$), and the parameter of relative preference for energy (ϵ) common to all households

$$\bar{d}_{i,t} = d_{i,t}\left(1 - \epsilon\frac{(1 - X_{i,t})}{X_{i,t}}\right). \tag{115}$$

We obtain the initial distribution of households' energy expenditures and energy consumption shares from the Consumer Expenditure Surveys (CEX). 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_i) are obtained by summing expenditures on gasoline and motor oil, electricity, natural gas, fuel oil, and other fuels. The energy expenditure shares (X_i) are obtained by dividing energy expenditures by total expenditures. To determine hourly wages, we apply the same procedure as with the SCF. We first compute the household annual wage by summing the income received from salary or wages before taxes. We then compute the annual labor supply of the respondent and its partner: we multiply the number of hours usually worked per week by the number of weeks worked in the past twelve months, minus 3 weeks of imputed holidays. The household hourly wage is then the ratio of the household annual wage over annual hours. Just like with the SCF data, this number reflects how much households members were paid on average for each hour of work they supplied in the past year.[13]

In order to characterize the joint distribution of hourly wages and energy expenditure shares, we restrict our sample to working households, following the same sample definition as with the SCF. Using

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

this sub-sample, we divide households in ten groups of hourly wage deciles. For each group, we compute the average hourly wage. For each of the ten groups, we divide again households in five weighted quintiles of energy expenditure share, and compute the average energy expenditure share. This provides a table in which households are split in 50 groups of equal size, with for each of these groups the average hourly wage and the energy expenditure share.[14] Since energy consumption shares do not appear to be strongly determined by age among working households, we do not control for generational differences. However, we control for seasonality and yearly variations that could lead to overestimate consumption heterogeneity. We proceed in the same way as with generational controls: we group individuals based on their ranking relative to the people interviewed in the same month and same year. We then compute the average for each group over all time periods. The resulting distribution of initial energy shares $\{X_i\}_{i \in I}$ is presented in Table III.

Finally, in order to obtain the distribution of $\bar{d}_i$, we also need to determine ϵ. Relative to the data, the model gives us a degree of freedom, hence we assume ϵ is such that the group i with the lowest consumption share has $\bar{d}_i = 0$, which gives $\epsilon \simeq 0.0263$. [To be added: calibration going forward.]

Table III: Distribution of households energy expenditure shares by productivity deciles (rows) and expenditure share quintiles (columns), controlling for seasonality and time trend.

		Expenditure share quintiles					Average
		1st	2nd	3rd	4th	5th	
Productivity deciles	1st	2.69%	7.59%	11.42%	15.88%	24.39%	12.70%
	2nd	3.50%	8.07%	11.48%	15.26%	22.83%	12.51%
	3rd	4.13%	8.29%	11.31%	14.78%	21.79%	12.33%
	4th	4.09%	7.99%	10.86%	14.00%	20.46%	11.84%
	5th	4.09%	7.63%	10.33%	13.39%	19.45%	11.20%
	6th	3.93%	7.23%	9.75%	12.78%	18.86%	10.74%
	7th	3.83%	6.90%	9.25%	12.03%	17.89%	10.19%
	8th	3.47%	6.22%	8.44%	11.17%	16.96%	9.45%
	9th	3.04%	5.63%	7.76%	10.29%	16.05%	8.76%
	10th	2.56%	4.95%	7.01%	9.65%	15.60%	8.16%

Note: The rows correspond to productivity (*i.e.* hourly wage) decile groups. The column "Average" corresponds to the average energy expenditure share for each productivity group. Columns 1 to 5 correspond to energy expenditure share quintile groups within productivity decile 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 weighted average across all periods. Example: 2.69% in the 1st row, 1st column, means that among the people that belong to the bottom 10% of the hourly wage distribution at the month × year they were interviewed, the 20% with lowest energy shares spend on average 2.69% of their total expenditures in energy. Sample: CEX from 2011 to 2015, only workers included.

[14] We chose to divide each decile group in quintiles instead of deciles in order to mitigate the impact of potential outliers.

C.2 Parameters choice

Baseline hours worked We also use the SCF 2013 to compute the initial labor supply that we impute to the model. To do so, we again restrict the sample to individuals between 18 and 65 years old. However, because our aim is not to compute hourly wages but to look at the average labor supply, we do not eliminate outliers based on their hourly wage or labor supply. In particular, we keep unemployed households for whom the hourly wage is not observed, as dropping them would lead to overestimate the average labor supply. For all households in the sample, we divide the annual labor supply by the number of working age individuals (individuals between 18 and 65). This yields an average of 1440 hours annually. Assuming a maximum labor supply capacity of 52 weeks per year and 100 hours per week per individual, this yields an average labor supply of 0.277 of the maximum capacity.

Table IV: Calibration summary: climate parameters.

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
χ_{2015}^{Ex}	Initial forcings of non-CO2 GHG (Wm-2)	0.5
χ_{2100}^{Ex}	2100 forcings of non-CO2 GHG (Wm-2)	1
$g_{\chi^{\text{Ex}}}$	Rate of convergence of χ	1/17
Temperature		
T_{2015}	Initial atmospheric temperature change (C since 1900)	0.85
T_{2015}^{Lo}	Initial lower stratum temperature change (C since 1900)	0.0068
ζ_1	Climate model coefficient	0.1005
ζ_2	Climate model coefficient	1.1875
ζ_3	Climate model coefficient	0.088
ζ_4	Climate model coefficient	0.025

Note: All parameters are taken from DICE (2016).

Table V: Calibration summary: economic parameters.

Parameter	Description	Value	Source
<u>Preferences</u>			
β	Utility discount rate (per year)	1/(1.015)	DICE 2016
σ	Inverse of IES	1.45	DICE 2016
η^F	Frisch elasticity of labor supply	0.75	Chetty et al (2011)
ς	Labor dis-utility coefficient	1.885	To target η^F and h_{2015}
γ	Labor dis-utility exponent	0.709	To target η^F and h_{2015}
α_0	Relative preference for the environment	7.61e-05	Adapted from Barrage (2019)
<u>Production damages</u>			
a_1	Damage intercept	0	DICE 2016
a_2	Damage coefficient quadratic term	0.00175	DICE 2016 adjusted
a_3	Damage exponent	2	DICE 2016
<u>Production first sector</u>			
α	Return to scale on labor sector 1	0.3	DICE 2016
ν	Return to scale on energy sector 1	0.04	Golosov et al (2014)
δ	Depreciation rate on capital (per year)	0.1	DICE 2016
r_{2015}	Initial net rate of return on capital	0.023	To target steady state
Y_{2015}	Initial output (in trillions 2015 USD)	70.807	World Bank (2011-2015)
$hh_{1,2015}$	Initial share of labor in sector 1	0.976	To equate MPL across sectors
$kk_{1,2015}$	Initial share of capital in sector 2	0.926	To equate MPL across sectors
E_{2015}	Initial industrial emissions ($GtCO_2$ per year)	35.85	DICE 2016
h_{2015}	Initial labor supply per capita	0.277	Computed from SCF
$A_{1,2015}$	Initial TFP sector 1	141.9	To target Y_{2015}
<u>Production second sector</u>			
α_E	Return to scale on capital sector 2	0.403	Barrage (2019)
$A_{2,2015}$	Initial TFP sector 2	87.1	To target E_{2015}
<u>Abatement costs</u>			
$P_{2015}^{\text{backstop}}$	Backstop price in 2015 (in \$/tCO_2)	550	DICE 2016
$g_{P^{\text{backstop}}}$	Decline rate backstop price (per period)	2.5%	DICE 2016
c_2	Exponent abatement cost function	2.6	DICE 2016
μ_{2015}	Initial abatement share	0.03	DICE 2016
<u>Government</u>			
G_t/Y_t	Government spending to GDP ratio	0.3030	IMF-GFS
B_{2015}	Initial public debt to GDP ratio	0.2220	IMF-GFS
$\tau_{H,2015}$	Initial tax rate on labor income	0.255	Trabandt & Uhlig (2012)
$\tau_{K,2015}$	Initial tax rate on capital income	0.411	Trabandt & Uhlig (2012)

Calibration summary: economic parameters (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

Note: The adjustments relative to DICE 2016 for population and damages are described in the calibration section.