Carbon taxation and precautionary savings*

Stefan Wöhrmüller

This draft: February 2023

Preliminary - Comments welcome
Please do not quote or redistribute without permission

Abstract

How does uninsurable idiosyncratic risk affect the optimal carbon tax? To answer this question, I augment a heterogenous-agent incomplete-markets model with a climate externality on total factor productivity and dirty energy demand of households and firms. A government sets a carbon tax on energy and redistributes its revenue lump-sum. I find that the optimal carbon tax is increasing in the level of uninsurable idiosyncratic risk, because the tax and transfer combination provides redistribution and insurance through higher transfers and by increasing wages and interest rates due to lower climate damages. This result depends on the availability of adjustable tax instruments.

Keywords: heterogenous agents, precautionary savings, carbon taxation
JEL codes: C23, D12, D31, Q50

*I am indebted to Christian Stoltenberg for his continuous guidance and support during my PhD. I am very grateful to Thomas Douenne, Sergio de Ferra, Kees Haasnoot, Alejandro Hirmas, Albert Jan Hummel, Bas Jacobs, Eva Janssens, Franc Klaassen, Kurt Mitman, Marcelo Pedroni, Rick van der Ploeg, Elisabeth Pröhl, Federica Romei, Timo Schenk, Randolph Sloof, and Konstantin Sommer for many helpful comments and conversations. All errors are my own. *Email:* s.h.p.woehrmueller@uva.nl

1 Introduction

How does uninsurable idiosyncratic earnings risk affect the optimal carbon tax? In this paper, I argue that an increase in the uninsurable idiosyncratic risk that households face gives rise to a higher optimal carbon tax in general equilibrium.

I begin by considering a standard heterogenous-agent incomplete-markets economy á la Aiyagari (1994) and enrich it along four dimensions. First, households supply labor and have Stone-Geary preferences over a clean and a dirty consumption good, where the dirty energy good is subject to a subsistence level. The subsistence level implies that carbon taxation would be regressive per se, as poorer households spend a larger fraction of their income on dirty goods. Second, the supply side features an energy producer and a final goods producer that uses capital, labor, and energy for production. Third, I introduce a climate externality. Energy production is pollution intensive and increases the stock of carbon in the atmosphere which in turn decreases total factor productivity of the final good firm. Fourth, the government has access to non-individualized transfers, capital taxes, labor income taxes, and carbon taxes.

In this model environment, households have a precautionary saving motive, because there is only one saving instrument available to self-insure against the uninsurable idiosyncratic risk and borrowing is limited by an exogenous constraint. This precautionary saving behavior implies a concave consumption function in current income and wealth with decreasing marginal propensities to consume clean and dirty goods. In other words, Engel curves over clean and dirty goods consumption are non-linear. I show that this result also holds for Stone-Geary preferences, which imply a linear expenditure system in a model environment without idiosyncratic risk. As a result, the optimal carbon tax set by the government might also take into account distributional and insurance concerns in addition to the climate externality.

To quantitatively study this interaction between carbon taxation and uninsurable idiosyncratic risk and precautionary savings, I calibrate and estimate the economic part of the climate-economy to match features of the U.S. economy. The calibration of the climate block is largely taken from the literature and represents climate impacts on a global scale. I then use my estimated model as a laboratory and optimize over the carbon tax in a stationary recursive equilibrium under an utilitarian welfare criterion. In particular, the government chooses the carbon tax and redistributes revenue lump-sum to households to maximize social welfare in the economy. Due to concavity of the utility function, the government has an implicit preference for redistribution and insurance.

The optimal carbon tax in the benchmark steady state is about 141 \$ per ton of CO_2 and emissions are reduced by almost half. To understand the effect of idiosyncratic risk and precautionary savings on the level of the carbon tax, I then repeat this optimization when I scale up and down the idiosyncratic labor productivity risk that agents face in the economy. As my main result, I find that the optimal carbon tax is increasing in the level of uninsurable idiosyncratic risk. This results holds for a comparative statics exercise, where model parameters are kept fixed, and for a case in which I recalibrate the model prior to optimization to match average hours worked and the capital-to-output ratio from the benchmark.

I identify two channels why a higher carbon tax is optimal in a high-risk environment. First, carbon tax revenue is redistributed lump-sum back to households, thus improving insurance and equality in the economy. Indeed, under optimal carbon taxes both consumption and net income inequality decrease. Second, as polluting is now more expensive, emissions and economic damages decrease, total factor productivity increases and thus also interest rates and wages. Both channels mirror state-independent insurance policies against the uninsurable idiosyncratic risk and especially higher wages are valued more by poorer households in a more uncertain environment.

To better understand what components gives rise to a non-zero carbon tax in this economy and the positive relation between optimal carbon taxation and uninsurable idiosyncratic risk, I conduct several exercises that change various features of the model. First, I introduce a skill-specific incidence rule for lump-sum transfers. Transfers are now proportional to ones labor productivity risk and are hence less redistributive. In response to this, the optimal carbon tax is reduced by half. The planner reduces the tax compared to the benchmark case, because transferring carbon tax revenue back to households now contributes less to redistribution and insurance.

Second, I optimize the carbon tax if I remove the subsistence level and damages, respectively. Eliminating the subsistence level increases the carbon tax only slightly. In this respect, the regressivity of the subsistence level does not seem to have a big quantitative impact. Eliminating climate damages, on the other hand, reduces the carbon tax to a measly 9$ per ton of CO_2. Naturally, this indicates that a carbon tax on its own is an ineffective way to raise revenue for redistribution and insurance. In this way, it predominantly increases the price of the dirty good without bringing the benefit of increasing wages and interest rates through lower climate damages on total factor productivity.

Note that the utilitarian criterion used so far conflates redistribution and efficiency motives of the planner. For this reason, I also use a criterion inspired by Bénabou (2002) that places no weight on interpersonal redistribution. The carbon tax under the efficient welfare criterion is about 6% lower than under the utilitarian welfare criterion. Furthermore, combining the efficient welfare criterion with no damages in the economy pushes the carbon tax to almost zero. These findings suggest that redistribution due to equality concerns of the planner is larger than for efficiency concerns. Under no damages, the carbon tax is reduced to a pure consumption tax on the dirty good and commodity tax differentiation does not yield large efficiency gains in this setup.

Finally, all of the above exercises keep other taxes in the economy, in particular labor income and capital income taxes, fixed. This exercise has major practical relevance, however, it is unsatisfactory from an economic theory perspective, as all inefficiencies and redistributive shortcomings of the economy are captured by the carbon tax alone. As a result, I repeat the main exercises from above, but now also let the government adjust labor income taxes. As before, lump-sum transfers adjust such that the government budget constraint holds. I find that average labor income taxes are increasing in uninsurable idiosyncratic risk and precautionary savings. This is intuitive, as the planner now engages in progressive redistribution by increasing labor taxes. At the same time, however, carbon taxes are now *decreasing*. Intuitively, marginal benefits with respect to redistribution from the carbon tax-transfer combination are now lower, and the price increases on the dirty good are hurting agents relatively more in a high-risk environment. Hence, if the government has additional tax instruments at its disposal, the relationship between uninsurable risk and carbon taxes reverses.

Related literature and contribution My paper contributes to several strands of the literature overarching optimal fiscal policy, consumption dynamics, and environmental economics.

My key contribution to this literature is the *joint analysis* of optimal carbon taxation in an environment with idiosyncratic risk which generates precautionary savings. The quantitative model combines a heterogeneous-agent incomplete market economy in the spirit of Bewley (1986); Huggett (1993); Aiyagari (1994) with a climate sector, which yields an endogenous distribution over income and wealth, and heterogeneity in marginal propensities to consume. Hence, my model allows to study the interaction of climate policies and economic inequalities in a unified framework. Thereby, I connect two lines of research.

The first line is a rapidly growing literature which analyzes optimal carbon taxation in quantitative macroeconomic models. Building on the seminal work by Nordhaus (1992, 1993), who developed the first integrated assessment model (IAM) to analyze climate damages within a centralized economic framework, several papers moved to decentralized market structure. For instance, Golosov, Hassler, Krusell and Tsyvinski (2014) derive a formula for the optimal carbon tax in a dynamic stochastic general-equilibrium model with an externality and resource scarcity. Building on their quantitative work, Barrage (2020) quantifies optimal carbon taxation in a model with tax distortions, and in turn, Douenne, Hummel and Pedroni (2023) quantify the additional impact of inequality.[1] None of these papers investigate settings with idiosyncratic risk, which is my main contribution compared to existing frameworks.[2]

The second line investigates the impact of idiosyncratic uncertainty and borrowing constraints on individual consumption demand. In particular, in the presence of idiosyncratic risk both prudence in preferences as well as borrowing constraints give rise to a precautionary saving motive which renders the consumption function concave in current income and wealth (Leland, 1968; Sandmo, 1970; Zeldes, 1989*b*,*a*; Kimball, 1990*a*,*b*; Carroll and Kimball, 1996; Huggett and Ospina, 2001; Carroll, Holm and Kimball, 2021).[3]

[1] Other quantitative examples study the optimal environmental policy in response to business cycles (Heutel, 2012) or nominal frictions and uncertainty (Annicchiarico and Di Dio, 2015), or in an overlapping generations framework (Kotlikoff, Kubler, Polbin and Scheidegger, 2021*a*; Kotlikoff, Kubler, Polbin, Sachs and Scheidegger, 2021*b*).

[2] An exception is Benmir and Roman (2022) who study the 2050 net zero emissions target for the U.S. in a HANK model. The main difference to the present paper is that I focus on the optimal carbon tax and model household consumption with quasi-homothetic preferences and two goods.

[3] Lugilde, Bande and Riveiro (2019) survey the empirical literature on precautionary savings. They conclude that papers which "test the effect of uncertainty about future income on consumption/saving decisions, especially [those] using micro data, tend to provide robust and convincing results as regards the existence of a precautionary motive for saving" (p.507). Examples of micro-panel studies in different countries include Carroll and Samwick (1997, 1998); Guariglia and Rossi (2002); Guariglia (2003); Lugilde, Bande and Riveiro (2018).

Second, my exercise builds on the theoretical literature on optimal carbon taxation. In particular, Jacobs and van der Ploeg (2019) show that the optimal carbon tax should be equal to the marginal external damage of pollution if Engel curves are linear and the social planner has access to a non-individual lump-sum transfer and linear income taxes.[4] In other words, the optimal carbon tax follows the Pigouvian rule (Pigou, 1920).[5] Intuitively, any demand change induced by the carbon tax can be undone by changing the lump-sum transfer and the income tax. The main difference in this paper is that I consider a quantitative model with idiosyncratic risk, CRRA utility, and borrowing constraints in which, as explained above, non-linear Engel curves are microfounded.[6]

Recent studies further extend these theoretical analyses under deterministic environments with tax distortions (Barrage, 2020) and inequality (Douenne et al., 2023). Compared to this theoretical literature I do not have analytical results concerning the optimal carbon tax, because a closed-form solution is not obtainable within the class of models I consider. Instead, I conduct counterfactual analyses to disentangle the main forces behind my results, as is common in this literature (see e.g. Conesa, Kitao and Krueger, 2009; Dyrda and Pedroni, 2023).

In addition, my paper relates to the literature on subsistence consumption of carbon-intensive goods and the incidence of taxation (Klenert and Mattauch, 2016; Klenert, Schwerhoff, Edenhofer and Mattauch, 2018). I contribute to this literature by proposing a novel strategy to estimate the structural parameters - including the subsistence level of dirty goods consumption - of the model via indirect inference (Guvenen and Smith, Jr., 2014; Stoltenberg and Uhlendorff, 2023).[7]

Lastly, my paper builds on the literature which studies how to optimally recycle carbon tax revenue (Fried, Novan and Peterman, 2018, 2021; Goulder, Hafstead, Kim and Long, 2019). This paper, on the other hand, examines the optimal level of the carbon tax, and the method of revenue recycling is, for now, set to lump-sum transfers.

The paper is organized as follows. Section 2 describes the quantitative model. Section 3 presents the data and outlines the estimation strategy. Section 4 briefly discusses the main quantitative exercise and presents the main results. Section 5 concludes.

2 An Economy with Idiosyncratic Risk, Two Goods, and a Climate Externality

This first two parts of this section describe the economic model used in the quantitative analyses to study the interaction between uninsurable idiosyncratic risk and optimal carbon taxes. Households face uninsurable idiosyncratic productivity risk and borrowing constraints, supply labor and

[4] This result is reminiscent of earlier studies by Angus Deaton (Deaton, 1979, 1981) in which he demonstrates that uniform commodity taxation is desirable under linear Engel curves and separability in consumption and leisure.

[5] This refers to Proposition 2 in Jacobs and van der Ploeg (2019).

[6] Jacobs and van der Ploeg (2019) is a specific application of a more general result that the optimal carbon tax equals the Pigouvian rate adjusted by the marginal cost of public funds (Sandmo, 1975; Bovenberg and de Mooij, 1994), which equals one under the optimal tax system (Jacobs and de Mooij, 2015; Jacobs, 2018).

[7] The references in the text refer to recent applications of indirect inference to dynamic macroeconomic models. A theoretical treatment can be found in Gourieroux, Monfort and Renault (1993) and Smith, Jr. (1993).

consume clean and dirty goods. The structure of production and the climate sector largely follows Barrage (2020) and Golosov et al. (2014), respectively. In the last part, I discuss household consumption and saving decisions in more detail. In particular, I discuss the emergence of non-linear Engel curves over consumption in this framework, even under quasi-homothetic preferences, which provides the theoretical rationale for carbon taxes to take into account distributional concerns (Jacobs and van der Ploeg, 2019).

2.1 Setup

Households Time is discrete, $t \in \{0, 1, \dots\}$, and there is no aggregate risk. The time period in the model is five years. The economy is populated by a continuum of infinitely-lived households of measure one. Households' preferences are represented by the utility function

$$\mathbb{E} \sum_{t=0}^{\infty} \beta^t \frac{\left(c_{it}^{\eta}(d_{it} - \underline{d})^{1-\eta}\right)^{1-\gamma}}{1-\gamma} + \chi \frac{(1-n_{it})^{1-\epsilon}}{1-\epsilon} \tag{1}$$

where c_{it} denotes the consumption flow of the clean good, d_{it} denotes the consumption flow of the dirty good, d denotes the subsistence level for the dirty consumption good, and n_{it} denotes labor supply of household i at time t. The time endowment of each household is normalized to 1. The future is discounted with factor β.

The first part of the preferences in Equation (1) nests a Stone-Geary utility in a CRRA specification. In particular, γ denotes the coefficient of relative risk aversion and $\underline{d}$ is the subsistence level for the dirty consumption goods. It is important to note that the elasticity of substitution between the clean and the dirty good is decreasing in the subsistence level (Baumgärtner, Drupp and Quaas, 2017).[8] η and $(1-\eta)$ are expenditure shares based on total income net subsistence consumption, as will become clear below. Regarding the second part, χ denotes the disutility of labor supply, and ϵ is related to the Frisch elasticity of labor supply, $\frac{1}{\epsilon}\frac{1-n}{n}$.

Households are subject to idiosyncratic productivity risk captured by a first-order Markov chain $\theta_t \in \Theta$ with $|\Theta| = S < \infty$ and transition matrix $\underset{S\times S}{\Gamma}$. An agents' pre-tax income is then determined by her productivity, the equilibrium wage per unit of productivity, w_t, and the amount of labor supply: $y_t^{pre} = w_t \theta_t n_t$. Pre-tax income is transformed into net (or after-tax) income using a net income function $\mathcal{T}(y) = y - T^y(y)$, where the tax function $T^y(\cdot)$ is to be specified below. Moreover, households have access to a one-period risk-free bond, a, as consumption insurance instrument. Capital income is taxed at rate τ^k and borrowing is restricted by an ad-hoc constraint $\underline{a}$. Lastly, share $(1-\mu)$ of energy produced is dirty and hence potentially subject to a carbon tax τ_d, which the energy producer passes-through at rate ϖ. The government pays lump-sum transfers g to the household.

[8] Under no subsistence consumption this elasticity is one (usual Cobb-Douglas case).

Hence, the household budget constraint is

$$c_t + (p_d + (1-\mu)\varpi\tau_d)d_t + a_{t+1} = \mathcal{T}(y^{pre}) + (1 + r(1-\tau^k))a_t + g,$$

where p_d denotes the price of the dirty good, respectively, r is the equilibrium interest rate. In the following, I define $\tilde{p} \equiv p_d + (1-\mu)\varpi\tau_d$.

Production I model two production sectors (Barrage, 2020; Douenne et al., 2023).

Final good sector In the final goods sector, indexed by 1, a final good Y is produced using a neoclassical aggregate production function

$$Y = (1 - \mathfrak{D}(S))\tilde{X}\tilde{F}_1(K_1, L_1, E^p) = X(S)\tilde{F}_1(K_1, L_1, E^p) = F_1(K_1, L_1, E^p; \tilde{X}, S) \tag{2}$$

with K_1 units of capital, L_1 efficiency units of labor, E^p units of energy as inputs, and total factor productivity $\tilde{X}$. The final good can either be consumed or invested. $\mathfrak{D}(S)$ represents climate damages to output as a function of the stock of atmospheric carbon S with $\mathfrak{D}'(S) > 0$. This modelling approach of climate damages follows the seminal work by Nordhaus (1991) and the more recent environmental macroeconomic literature.

Energy sector In the energy sector, indexed by 2, energy E is produced using a neoclassical aggregate production function

$$E = F_2(K_2, L_2) \tag{3}$$

with K_2 units of capital and L_2 efficiency units of labor. Energy is either consumed by households (dirty good) or used in production of the final good such that $E = E^p + D$. Following Barrage (2020), producers can provide a share μ from clean energy production, such that only $E^m = (1-\mu)E$ contributes to the stock of emissions. This clean technology is available at a cost of $\Psi(\mu)$ per unit of energy.

Lastly, capital and labor are fully mobile across sectors such that market clearing implies:

$$K = K_1 + K_2 \tag{4}$$

$$L = L_1 + L_2 \tag{5}$$

Government The government levies labor taxes on pre-tax income y^{pre} using the possibly non-linear labor tax function $T^y(y^{pre})$, a linear capital income tax τ^k as well as a carbon tax on dirty goods consumption τ_d. Moreover, it issues government debt B, and chooses lump-sum transfers g to balance its budget:

$$B_{t+1} + g_t = (1+r)B_t + \mathfrak{T}_t, \tag{6}$$

where $\mathfrak{T}_t$ denotes total tax revenue from labor, capital, and carbon taxes.

Climate sector

Carbon cycle The current level of atmospheric carbon concentration, S_t, depends on current and past emissions. In my case, emissions are related to energy produced net of the abated share:

$$S_t = \sum_{\tau=0}^{\infty}(1-\Phi_\tau)\left[(1-\mu_{t-\tau})E_{t-\tau}\right] = \sum_{\tau=0}^{\infty}(1-\Phi_\tau)E^m_{t-\tau}$$

where $1-\Phi_\tau = \varphi_L + (1-\varphi_L)\varphi_0(1-\varphi)^\tau$ with the following interpretation: φ_L is the share of carbon emitted which stays in the atmosphere forever; a share of $1-\varphi_0$ of the remaining $1-\varphi_L$ exits the atmosphere immediately; and a remaining share $(1-\varphi_L)\varphi_0$ that decays at geometric rate φ. To write it recursively, following Känzig (2023), I set and $\varphi_L = 0$ and write

$$S_t = (1-\varphi)S_{t-1} + \varphi_0 E^m_t \tag{7}$$

Recursive problem An agent is characterized by the by the pair $(a_{it} = a, \theta_{it} = \theta)$, the household state, and solves the following optimization problem

$$\begin{aligned} V(a,\theta) = \max_{c,d,n,a'} & \; u\big(c,d,n\big) + \beta\mathbb{E}_\theta V\big(a',\theta'\big) \\ & \text{subject to} \\ & c + \tilde{p}d + a' \leq (1 + r(1-\tau^k))a + \underbrace{w\theta n - T^y(w\theta n)}_{\mathcal{T}(w\theta n)} + g \\ & a' \geq \underline{a} \end{aligned} \tag{8}$$

2.2 Equilibrium

Let $A \equiv [\underline{a},\overline{a}]$ be the set of possible values for a_{it}. Define the state space by $\mathbb{S} \equiv A \times \Theta$ and let the σ-algebra $\Sigma_{\mathbb{S}}$ be defined as $B_A \otimes P(\Theta)$, where B_A is the Borel σ-algebra on A and $P(\Theta)$ is the power set of Θ. Finally, let $\mathcal{S} = (\mathcal{A} \times \Theta)$ denote a typical subset of $\Sigma_{\mathbb{S}}$. I define a stationary recursive equilibrium as follows.

Definition 1 (Stationary recursive equilibrium)**.** A stationary recursive equilibrium is a government policy $\{\tau_d, g\}$, a vector of aggregate quantities $\{Y, K_1, K_2, L_1, L_2, \mu, E, S\}$, a probability measure Λ defined over the measurable space $(\mathbb{S}, \Sigma_{\mathbb{S}})$, a set of policy functions $\{\boldsymbol{c}(a,\theta), \boldsymbol{d}(a,\theta), \boldsymbol{n}(a,\theta),$ $\boldsymbol{a}'(a,\theta)\}$, a set of prices $\{r, w, p_d\}$, and a set of policies $\{g, \tau_0, \tau_1, \tau_2, \tau^d, \tau_d\}$ such that: (i) given policies and prices, the decision rules solve the optimization problem Equation (8), (ii) the final goods firm chooses capital K_1, labor in efficiency units L_1, and energy E^p to maximize profits, (iii) the energy producer chooses capital K_2, labor in efficiency units L_2, and abatement μ to maximize profits, (iv) the government budget constraint

$$g + rB = \int_{(A\times\Theta)} T^y(w\theta\boldsymbol{n}(a,\theta))d\Lambda + \tau^k rA + \tau_d(1-\mu)E$$

holds, (v) the asset market clears

$$A \equiv \int_{(A\times\Theta)} \boldsymbol{a}'(a,\theta)d\Lambda = B + K$$

(vi) the goods market clears[9]

$$\int_{(A\times\Theta)} \boldsymbol{c}(a,\theta)d\Lambda + \delta K + \Psi(\mu)E = Y,$$

(vii) Λ is an invariant probability measure and satisfies for all $\mathcal{S} \in \Sigma_S$

$$\Lambda(\mathcal{S}) = \int_{(A\times\Theta)} Q((a,\theta),\mathcal{S})d\Lambda,$$

where Q is the associated Markov transition function induced by Γ and $\boldsymbol{a}'$, and (viii) the stock of emissions stays constant at $S = \frac{\varphi_0}{\varphi}(1-\mu)E$.

2.3 Quasi-homothetic preferences and concave consumption functions

Jacobs and van der Ploeg (2019) show, in a static setting where expenditure equals income, that under linear Engel curves, externality correcting taxes should be set at the Pigouvian rate. This subsection shows that the quantitative model outlined above implies *concave* consumption functions for clean and dirty goods over income and wealth due to the uninsurable idiosyncratic risk and precautionary saving behavior. Hence, there is a rationale for the carbon tax to deviate from the Pigouvian rate and to take distributional aspects into account.

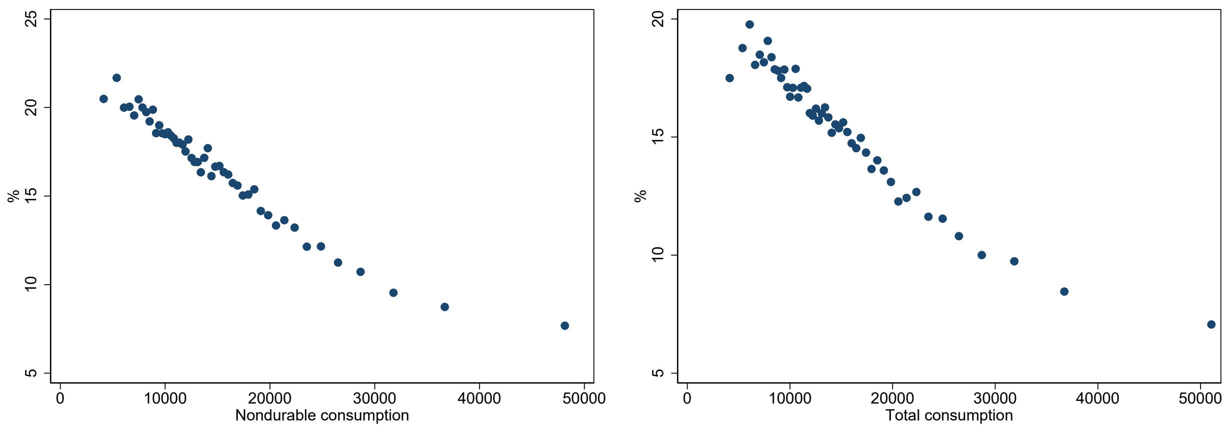
Stone-Geary preferences In the following, I will also argue that this concavity is present even for static Stone-Geary preferences as in Equation (1), which imply a linear Engel curves in settings with no uninsurable idiosyncratic risk and are often used in the environmental literature. The reason for choosing Stone-Geary type preferences is that they replicate the empirical fact of declining expenditure shares of carbon-intensive goods (dirty goods) such as energy.

Indeed, this declining relationship between energy expenditure share and total expenditure also holds for the PSID data that I later use for estimation. As Figure 1 shows, the expenditure share of US households on energy - defined as the sum of home fuel, heating, and electricity expenditure as a share of two different consumption measures in the PSID - decreases from 20% at the lower end of the expenditure distribution to around 8% at the upper end. Under Cobb-Dogulas utility, the expenditure share would be constant and independent of the expenditure level. However, the introduction of a subsistence level, $\underline{d}$, generates this pattern as households first have to cover the subsistence level before equating the (price-weighted) marginal utilities of the two goods (Equation (9)).

[9] This market clearing condition is actually redundant by Walras's law, but is nevertheless a useful check whether all equilibrium conditions are properly computed (Appendix A.3).

Figure 1: Energy expenditure relative to total expenditure

(a) Nondurable consumption expenditure (b) Total consumption expenditure



Note. This figure shows energy expenditure relative to consumption expenditure with and without durables for households in consumption expenditure 100 bins. Consumption includes expenses for food, gasoline, rent, utilities, communication, transportation, education, childcare, medical needs, vacations, clothing, and recreational activities. Durable components are car repair expenses, down-, loan-, and lease-payments for vehicle loans as well as other expenditure regarding vehicles. All variables have been adjusted using the OECD equivalence scale and are expressed in 2010-$.

Static household problem To further understand the role of the subsistence level, and to facilitate the discussion below, it is instructive to separate the household problem into a dynamic and a static one. In the dynamic problem, the household chooses how much to save for the next period, a_{it+1}, and how much to spend on consumption. Denote this latter total expenditure by e_{it}. In the static problem, the household allocates total expenditure between the clean and the dirty good, respectively. Formally, the household solves the following simple problem, in which e_{it} is predetermined:

$$u(e_{it}) = \max_{c_{it},d_{it}} c_{it}^{\eta}(d_{it} - \underline{d})^{1-\eta}$$

subject to:

$$c_{it} + \tilde{p}d_{it} = e_{it}$$
$$c_{it} \geq 0, \quad d_{it} \geq \underline{d}$$

The intra-temporal first-order condition of this problem is

$$u_d(c_{it}, d_{it}) = \tilde{p}u_c(c_{it}, d_{it}) \quad \Rightarrow \quad (1-\eta)c_{it} = \eta\tilde{p}(d_{it} - \underline{d}). \tag{9}$$

The solution to this problem is

$$\begin{aligned} c_{it} &= \eta\left(e_{it} - \tilde{p}\underline{d}\right), \\ d_{it} &= (1-\eta)\frac{e_{it}}{\tilde{p}} + \eta\underline{d}. \end{aligned} \tag{10}$$

Hence, in this simple setting, decision rules for clean and dirty consumption are linear in total expenditure. The subsistence level is merely a shifter of the expenditure expansion paths. This is an important feature of these particular preferences. In fact, the system of demand equations implied by them are referred to as the Linear Expenditure System (Stone, 1954).

Dynamic household problem Importantly, the main point of this subsection is then the following: Engel curves are *not* linear under the dynamic model described in Section 2.1, which features uninsurable risk and precautionary savings, even with Stone-Geary preferences (nested in a CRRA specification). The key to this observation lies in the concavity of the consumption function in heterogeneous-agent incomplete-markets models (Zeldes, 1989*b*; Carroll and Kimball, 1996), which is inextricably linked to the saving behavior of households (Huggett, 2004; Jensen, 2018). Due to uncertain future income or productivity, households accumulate precautionary savings and especially so when asset and/or income levels are low. Intuitively, the precautionary desire for households to self-insure against possible future negative income realizations increases with lower resources. Hence, poorer households with a relatively stronger precautionary motive have lower consumption and higher marginal propensities to consume (Jappelli and Pistaferri, 2017). In other words, the dirty good Engel curve is non-linear.[10]

Figure 2 illustrates these points and highlights the distinction between the static and dynamic framework. Panel 2a shows consumption functions for two productivity types as a function of assets; both are clearly concave and more so for lower levels of assets. Panel 2c shows expenditure on the dirty good as a function of total expenditure. We see that this relation is linear, relating to the static subproblem of the household (Equation (10)). Panel 2b shows the marginal propensity to consume the dirty good, what I term the *marginal propensity to consume (MPP)*, out of a windfall income gain of 1% of average income. We see that there is a distribution of MPPs, with higher marginal propensities for the lower productivity type. This heterogeneity is a clear indication of non-linear Engel curves. Overall, in the static stage, there is a linear mapping from total expenditures to expenditures for the dirty good. In the dynamic stage, however, there is a concave mapping from income or assets to total expenditure. Both taken together imply a concave mapping from income or assets to dirty goods expenditure or consumption.

The following proposition formalizes this discussion:

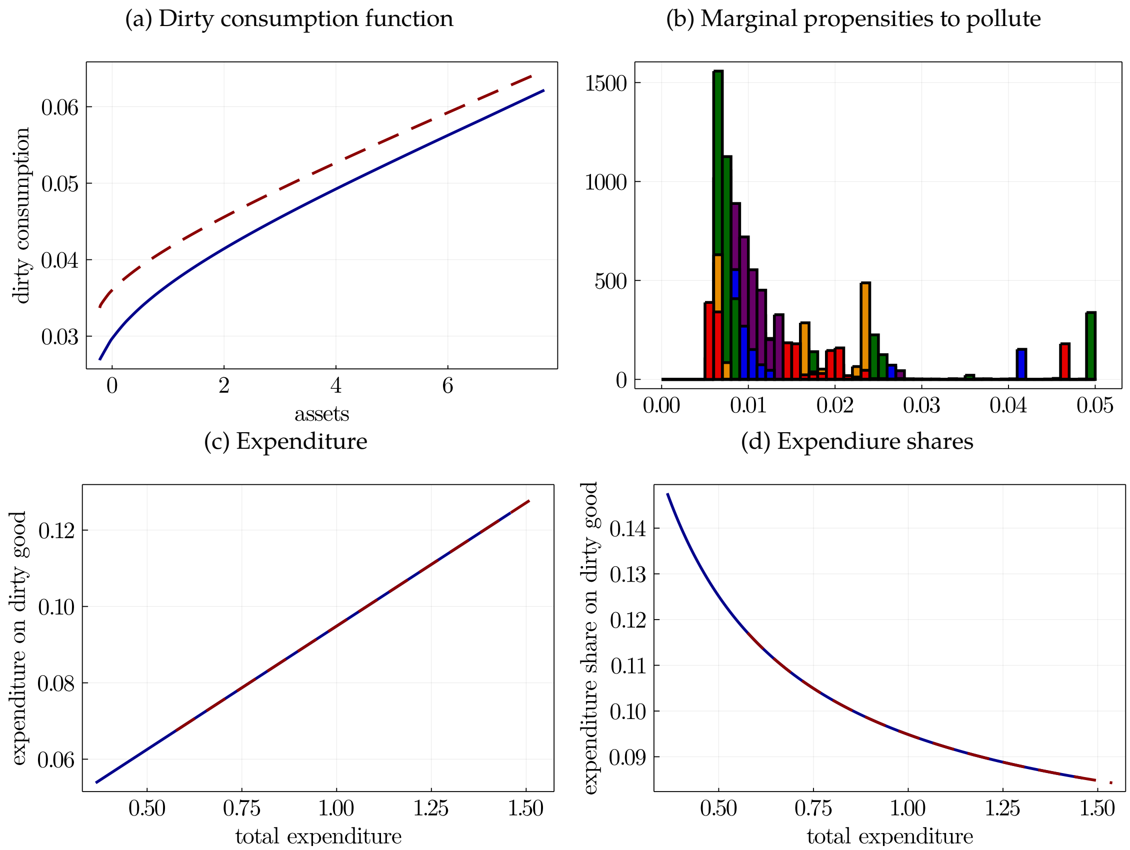
Proposition 1 (Non-linear Engel curves)**.** *Under (quasi-)homothetic preferences, inelastic labor supply, and for any labor-productivity Markov chain which induces non-negative consumption decisions, both the clean and dirty consumption good exhibits concave Engel curves w.r.t. to income and wealth:*

$$\boldsymbol{c}_{aa}(a,\theta) < 0,\ \boldsymbol{c}_{\theta\theta}(a,\theta) < 0 \quad \textit{and} \quad \boldsymbol{d}_{aa}(a,\theta) < 0,\ \boldsymbol{d}_{\theta\theta}(a,\theta) < 0$$

Proof. The proof of this proposition is a straightforward application of Theorem 1 in Carroll and Kimball (1996) for a finite horizon or Theorem 4 in Jensen (2018) for an extension to an infinite

[10] Carroll and Kimball (1996) discuss two cases under which the consumption function is linear. First, under isoelastic utility and interest rate uncertainty but no idiosyncratic uncertainty. Second, under CARA utility and labor income risk only.

Figure 2: Decision rules, expenditure, and marginal propensities to pollute



Note. This left two panels illustrate the dynamic and static subproblem of the household. The top left panel shows the standard concave (dirty) consumption function in incomplete markets models due to precautionary savings from the dynamic problem. The bottom left panel shows the linear relationship between total expenditure and dirty goods expenditure arising from the static problem. The top right panel illustrates that the concavity from the dynamic problem is also with respect to income, as marginal propensities to pollute are on average larger for low productivity households. The bottom right panel shows the decreasing expenditure share on dirty goods induced by the subsistence level on dirty goods consumption.

horizon and borrowing constraints. The proofs of these theorems go through without any modification, but the period utility function is replaced by the indirect utility of the static subproblem, and households choose total expenditure and savings instead of consumption and savings. Intuitively, it relies on the fact that the composition of a linear (decision rule in the static problem) and a concave function (decision rule in the dynamic problem) yields a concave function. □

The takeaway of Proposition 1 and the preceding discussion is that the optimal carbon tax should take into account distributional concerns in quantitative heterogeneous-agent incomplete market models with precautionary savings.

3 Bringing the model to the data

The discussion in the last part of the previous section suggests that the presence of idiosyncratic risk and precautionary savings matters - in theory - qualitatively for optimal carbon taxation. In light of this, the rest of this paper asks whether these features are also quantitatively important? Hence, I calibrate the model from Section 2.1 and have to choose functional forms and parameter values. The latter are chosen in two steps. First, I set values according to the literature or to match aggregate data targets. Second, I estimate the remaining set of structural parameters, which only belong to the household problem, using indirect inference in partial equilibrium (PE). PE requires to take a stance on the determination of prices, since I do not impose market clearing, which I will discuss below. Table 1 summarizes the parameter values.

3.1 Calibration

Preferences I choose a standard value for relative risk aversion, $\gamma = 2$, and set ϵ to target an average Frisch elasticity of labor supply of one.

Labor productivity process I model the idiosyncratic productivity process as the sum of a persistent and a transitory shock (plus measurement error):

$$\log(\theta_{it}) = \kappa_{it} + \psi_{it} + \nu_{it}$$
$$\kappa_{it} = \rho\kappa_{it-1} + \varepsilon^{\kappa}_{it}.$$

In particular, the persistence process κ is modelled as an AR(1) with persistence ρ and variance of its innovation of $\sigma^2_{\varepsilon^\kappa}$; the transitory shocks ψ are independently and identically distributed with zero mean and variance σ^2_ψ; ν denotes (classical) measurement error with σ^2_ν.

I determine the (annual) variances of the parameters using pre-tax wage residuals estimated from PSID data between 2000 and 2006, following the strategy by Flodén and Lindé (2001), and translate them into the 5-year period unit of the model. Moreover, I follow Heathcote, Storesletten and Violante (2010) and Straub (2019) and set $\sigma^2_\nu = 0.02$ as estimated in French (2004) to identify the (annual) transitory shock. In Appendix B.2, I describe the estimation procedure in detail.

Final goods production The technology $\tilde{F}_1$ is assumed to be of the constant elasticity of substitution (CES) form

$$\left[(1-s)(K_1^\alpha L_1^{1-\alpha})^{\frac{\lambda-1}{\lambda}} + s(E^p)^{\frac{\lambda-1}{\lambda}}\right]^{\frac{\lambda}{\lambda-1}} \tag{11}$$

with λ as the elasticity of substitution between the capital-labor bundle and energy, and s a share parameter.[11] In equilibrium, the factors of production are rented at rates $r + \delta$, w, and $\tilde{p}$, such that by Euler's theorem: $Y = (r+\delta)K_1 + wL_1 + \tilde{p}E^p$, where δ denotes capital depreciation.

[11] van der Werf (2008) writes that "the (KL)E nesting structure, that is a nesting structure in which capital and labour are combined first, fits the data best, but we generally cannot reject that the production function has all inputs in one CES function". Another recent example where this particular nesting structure is used is Hassler, Krusell and Olovsson (2021).

I fix the gross capital share in production α at 0.36 based on standard estimates from the literature (Rognlie, 2016) and the elasticity of substitution between the capital-labor composite and energy λ at 0.547 as found in van der Werf (2008). I follow Straub (2019) and set δ to match a capital-to-output ratio of 3.05. The implied wealth to output ratio is 3.8, close to the most recent estimate of 4 in Piketty and Zucman (2014, Figure IV) for the US. The share parameter s is set to match an energy share of production of five percent.

I normalize output to unity using the technology parameter X. Moreover, recall that X is a product of net of climate damages and damages: $X = \tilde{X}(1 - \mathfrak{D}(S))$. Hence, during estimation, I ignore the stock of carbon in the atmosphere in the economy, for I could always update $\tilde{X}$ to cancel out any resulting damages.

Energy production Energy is produced using a Cobb-Douglas technology in capital and labor:

$$E = K_2^{\alpha_E} L_2^{1-\alpha_E}. \tag{12}$$

I set $\alpha_E = 0.597$ following Barrage (2020). Moreover, the abatement cost function is

$$\Psi(\mu) = c_1 \mu^{c_2}. \tag{13}$$

I follow DICE and set c_2 to 2.6. Hence, the cost function is convex in μ, implying that marginal costs are increasing in abatement. To pin down c_1, I use initial steady-state values from Douenne et al. (2023), who largely follow DICE 2016 in their calibration. In particular, the backstop price describes the price of emissions at which there is full abatement, $\mu = 1$, which implies for marginal abatement costs: $c_1 c_2 \mu^{c_2 - 1} E = c_1 c_2 E = P^{backstop} E$. The parameter c_1 is then chosen such that the backstop-price implied energy costs to GDP ratio in initial steady state, $\frac{P^{backstop} E}{Y}$, is equal to 0.27 as in Douenne et al. (2023).

Government I use the three parameter functional form by Gouveia and Strauss (1994) to model the labor income tax function:

$$T(y^{pre}) = \tau_0 \left(y^{pre} - \left((y^{pre})^{-\tau_1} + \tau_2 \right)^{-1/\tau_1} \right). \tag{14}$$

Gouveia and Strauss (1994) report $\tau_0 = 0.258$ and $\tau_1 = 0.768$ for the year 1989 - their most recent estimate. τ_2 is determined in estimation and is adjusted such that the government budget constraint is satisfied. I set the capital income tax τ^k to 0.36 as in Trabandt and Uhlig (2011). Lump-sum transfers g are set to 0.114 to match a transfer-to-GDP ratio of 11.4% (Dyrda and Pedroni, 2023).

Climate sector As discussed above, during estimation I ignore climate damages. For completeness, however, I also now describe how I model the climate sector of the economy, which - in the spirit of Nordhaus's DICE model (Nordhaus, 1992, 1993) - follows Golosov et al. (2014).

Carbon cycle To calibrate φ and φ_0 I follow Golosov et al. (2014).[12] φ is set to capture the fact that excess carbon has a mean-lifetime of about 300 years such that $(1 - \varphi)^{60} = 0.5$, while the

[12] Golosov and co-authors, in turn, cite Archer (2005) and the 2007 technial summary of the IPCC report (IPCC, 2007)

Table 1: Preset parameters for estimation

	Description	Value	Target/source
Preferences			
γ	Risk aversion	2.0	literature
ϵ	Curvature of utility from leisure	4.06	Average Frisch elasticity of unity
Productivities (annual)			
ρ	Productivty shock persistence	0.9327	PSID
$\sigma^2_{\epsilon^\kappa}$	Variance of innovations to persistent shock	0.0426	PSID
σ^2_ψ	Variance of transitory shocks	0.0507	PSID
σ^2_ν	Variance of measurement error	0.02	French (2004, p.608, Table 5)
Production			
Final goods production			
λ	Substitution elasticity	0.547	van der Werf (2008, p.2972, Table 3)
α	Capital share	0.36	literature
δ	Depreciation (annual)	0.140	annual $K/Y = 3.05$ (FRED)
X	Net total factor productivity	2.75	Normalize output to unity
s	Share parameter	0.0054	Energy share in production of 5%
ω	Pass-through coefficient	0.25	
Energy production			
α_2	Capital share	0.597	Barrage (2020)
Abatement			
c_1	Scale abatement cost function	1.64	Backstop price to GDP (see text)
c_2	Exponent abatement cost function	2.6	DICE 2016
Government			
τ_0	Average labor income tax	0.258	Gouveia and Strauss (1994, p.323, Table 1)
τ_1	Progressivity of labor income tax	0.768	Gouveia and Strauss (1994, p.323, Table 1)
τ_2	Scaling parameter	0.525	Government budget constraint
τ^k	Capital income tax	0.36	Trabandt and Uhlig (2011, p.311,Table 1)
B/Y	Public debt (annual) / GDP	0.73	FRED
g/Y	Transfers / GDP	0.114	Dyrda and Pedroni (2023)
Climate sector			
Damages			
ξ	Damage parameter	0.016	GDP loss of 5% under BAU
Carbon cycle			
φ	Emissions decay parameter	$1 - \exp(\log(0.5)/60)$	Golosov et al. (2014)
φ_0	Emissions share parameter	$0.5/((1-\varphi)^6)$	Golosov et al. (2014)

Note. This table shows preset and calibrated parameters of the quantitative model which is used to estimate the remaining parameters via indirect infernce. FRED datasources can be found in Appendix B.1.

calibration for φ_0 captures that half of the CO_2 emissions into the atmosphere are removed after 30 years: $\varphi_0 = \frac{0.5}{(1-\varphi)^6}$

Damage function The functional form for the damage function is taken from Golosov et al. (2014):

$$1 - \mathfrak{D}(S) = e^{-\xi S_t}, \tag{15}$$

where ξ governs the strength of output damages of a marginal increase in atmospheric carbon.[13] Later, I set the parameter ξ such that in the initial steady state, without carbon taxes, damages imply a total loss of 5% of GDP.

Prices and quantities Prices and quantities are pinned down using a system of equations implied by the supply side of the economy. I fix $\{r, L, D\}$ and solve for $\{w, K, K_1, K_2, L_1, L_2, p_d, E, E^p, Y\}$ using i) five first-order conditions of the firms ii) two technology definitions iii) energy market clearing and iv) factor market clearing. The annual interest rate fixed at 3%, slightly below the estimate of Jordà, Knoll, Kuvshinov, Schularick and Taylor (2019) for the post-1980 period. I fix $L = N \sum_S \theta_s f(\theta_s)$, where $f(\cdot)$ denotes the invariant productivity type distribution induced by the Markov chain and where I set $N = 0.363$ since I target this number in estimation. Given arbitrary model units, there is no observable counterpart to D and it is difficult to derive a theoretical starting point, as aggregate dirty goods consumption depends on the distribution of agents in the economy. Hence, I fix a starting point of $D = 0.041$ from based on test estimation runs on a coarser grid. Moreover, I later verify post estimation that the implied aggregate dirty goods consumption is in line with this initial guess.

The remaining structural parameters are (i) the utility elasticity η, (ii) the subsistence level $\underline{d}$, (iii) the disutility of labor χ, (iv) the borrowing limit $\underline{a}$, and (v) the discount factor β. I will estimate these parameters using indirect inference. In the following two subsections, I will describe the micro data and targets I use. Thereafter, I will describe the estimation procedure in more detail.

3.2 Data

I use data from the Panel Study of Income Dynamics between 2005-2018 to compute the micro moments which I target in estimation. The PSID is a widely used longitudinal survey containing information on household demographics, income, and wealth. In the waves of 1999 and 2005, respectively, the PSID extended its collection of consumption expenditure data. It now captures over 70 percent of all consumption items available in the Consumer Expenditure Survey (CEX) and around 70 percent of aggregate consumption in the national income and product accounts (NIPA) (Blundell, Pistaferri and Saporta-Eksten, 2016). The PSID was attested to be a high quality dataset in terms of general low sample attrition rates and high response rates (Andreski, Li, Samancioglu and Schoeni, 2014).

[13] As Golosov et al. (2014) explain, Equation (15) is an approximation that conflates the *concave* relationship between CO_2 concentrations and temperature, and a *convex* relationship between temperature and damages. In particular, it implies constant marginal damages - measured as a share of GDP: $\frac{\partial Y/\partial S}{Y} = -\xi$

Variables The following variables are all on the household level. For instance, income refers to income from both the head and the spouse in the household, if present. Moreover, all monetary variables in the analysis have been adjusted using the OECD equivalence scale and are expressed in 2010-$.

Income *Labor income* refers to all income from wages, salaries, commissions, bonuses, overtime and the labor part of business income *Total income* in addition includes transfers such as as well as social security income. Both income variables are net of taxes, which were computed using NBER's Taxsim program.

Consumption *Nondurable consumption* includes expenses for food, gasoline, rent, utilities, communication, transportation, education, childcare, medical needs, vacations, clothing, and recreational activities. *Total consumption* also includes durable components such as car repair expenses, down-, loan-, and lease-payments for vehicle loans as well as other expenditure regarding vehicles. Moreover, I define *energy expenditure* as expenditure on gasoline, electricity, and heating. All three categories are greenhouse gas intensive goods and are thus used as a data counterpart for the dirty consumption good in the model.

Wealth My wealth variable refers to financial wealth net of liabilities. In particular, I include the value of one's real estate assets net of remaining mortgages, checking and saving accounts, stocks, bonds, business assets, IRAs or other annuities, and cars. I subtract liabilities such as credit card debt, student debt, outstanding medical bills, legal debt, loans obtained from relatives, and business debt.

Sample My baseline sample includes all PSID waves from 2005-2019, and consists of households where the head is between 25 and 60 years. I exclude observations for which information on consumption, income, wealth, education, household size, and region is missing. Furthermore, I remove observations with labor income below half the state minimum wage as well as top and bottom 1% of the remaining observations on consumption, income, and wealth. This leaves me with a sample of 21,750 households, around 2700 observations per year.

Table D.1 in the appendix shows descriptive statistics about the data. The typical household head is 42 years old, male, and married with 3 family members in total.

3.3 Data targets in estimation

Using the PSID sample and variables as just described, I construct the following moments.

Hours worked The cross-sectional average of weekly working hours of household heads in my sample is 40.61. Given that a full week has 168 hours and assuming 8 hours per day for sleep and other personal care leaves 112 hours per week as time endowment (Guerrieri and Lorenzoni, 2017). Hence, average hours worked as a share of the total time endowment gives 36.3% which is targeted in estimation.

Table 2: Dirty good regression

	dirty goods expenditure			
	(1)	(2)	(3)	(4)
Total expenditure	0.0646	0.0643	0.0623	0.0486
	(0.0152)	(0.0152)	(0.0156)	(0.0113)
Constant	0.3571	0.3517	0.3682	0.4169
	(0.1004)	(0.0743)	(0.1016)	(0.0896)
Observations	22033	22033	22033	22027
R-squared	0.3059	0.2996	0.3077	0.3100

Note. This table shows second stage (IV) coefficients δ_1 and δ_0 of Equation (16) for different specifications. Column (1) is the baseline case as specified in the text. Column (2) omits the region dummies in the control vector. Column (3) additionally controls for liquid assets. Column (4) uses an alternative measure for consumption as the (instrumented) regressor. Standard errors are corrected for heteroskedasticity and clustered at the household level.

Wealth-to-income ratio To consider the distribution of endogenous variables in my model, I follow Stoltenberg and Uhlendorff (2023) and target two moments of the wealth-to-income distribution: the 10^{th} percentile as well as the median. A 10^{th} and 50^{th} quantile regression on a constant yields $\hat{\beta}_{10} = -0.253$ and $\hat{\beta}_{50} = 1.056$, respectively. Both values are precisely estimated.

Dirty good allocation rule In the data I only observe expenditures, that is, the product of price and quantity. Hence, the data counterpart to Equation (10) is

$$p_d d_{it} = \delta_0 + \delta_1 e_{it} + \mathbf{X}'_{it}\omega + \varepsilon_{it}, \tag{16}$$

where now e_{it} denotes *observed* total expenditure, $p_d d_{it}$ *observed* expenditure on dirty goods of household i at time t. $\mathbf{X}$ is a vector of controls including household-size dummies, household head's five-year age bracket, region, household and year dummies (Pedroni, Singh and Stoltenberg, 2022; Straub, 2019). Total expenditure will be instrumented by total income as in Blundell, Chen and Kristensen (2007). Lastly, to obtain a proper mapping between arbitrary units in the model and data variables in 2010-$ units, I scale all data variables with the average dirty goods expenditure, $D^{avg} = p_d \bar{d}$.[14]

Table 2 shows estimation results for the baseline specification (1) and various robustness exercises (2)-(4). Column (2) omits the region dummies in the control vector. Column (3) additionally controls for liquid assets. Column (4) uses an alternative measure for consumption as the (instrumented) regressor. In all specifications, the first stage F-statistic is well above 10.[15]

To interpret the coefficients, let us look at column (1). The coefficient on total expenditure is equal to 0.0646. According to Equation (10), this coefficient identifies $1 - \eta$, which suggests an η of 0.9354. Second, δ_0 identifies $\eta \frac{p_d d\eta}{p_d D^{avg}} = \eta \frac{d}{D^{avg}}$. This constant as well as δ_1 will be targeted in

[14] When presenting this paper at a conference, Alkis Blanz brought to my attention that Blanz and Kalkuhl (2022); Blanz, Eydam, Heinemann, Kalkuhl and Moretti (2023) also calibrate their model starting from such a dirty allocation rule under Stone-Geary preferences.

[15] This rule of thumb is valid as I only consider one endogenous regressor and one instrument (Stock and Yogo, 2005).

Table 3: Targeted moments and estimated parameters

Moment	Estimated Parameters	Description
Wealth-to-income ratio - 10^{th} percentile	$\underline{a}$	Borrowing limit
Wealth-to-income ratio - Median	β	Discount factor
Expenditure regression (IV), δ_1	η	Clean good elasticity
Expenditure regression (IV), δ_0	$\underline{d}$	Subsistence level
Average hours worked	χ	Disutility of labor

Note. This table shows targeted moments an estimated parameters using indirect inference.

estimation.

3.4 Estimation with indirect inference

The remaining structural parameters $\{\eta, \underline{d}, \chi, \underline{a}, \beta\}$ of the model will be *jointly* estimated using indirect inference in a just identified system (same number of targets as unknown parameters) as shown in Table 3.

Objective function Denote the parameter vector by $\Theta = (\eta, \underline{d}, \chi, \underline{a}, \beta)$ and the vector of moments (as a function of the parameters) by $\mathcal{M} = \left(\delta_0(\Theta), \delta_1(\Theta), \frac{a}{y}\Big|_{50}(\Theta), \frac{a}{y}\Big|_{10}(\Theta), \overline{N}(\Theta)\right)$. Then define the vector of percentage deviations as

$$\mathcal{D} = \left(\mathcal{M} - \widehat{\mathcal{M}}\right) \oslash \widehat{\mathcal{M}}, \tag{17}$$

where $\widehat{\mathcal{M}}$ denotes the data counterpart of the respective moment conditions, and $\oslash$ denotes element-wise division. The objective function is then

$$\Theta^* = \arg\min_{\Theta} \mathcal{D}' W \mathcal{D}, \tag{18}$$

where W is a positive definite weighting matrix. I use the identity matrix as weighting matrix.

Given a parameter vector, I solve the model in partial equilibrium, where the tax parameter τ_2 is adjusted such that the government budget constraint holds. With a PE solution in hand, I construct the model-implied moments and use them in the objective function.

The objective function is minimized using a multistart global optimization algorithm ("TikTak"). For local minimization routines in this algorithm, I use the derivative-free BOBYQA routine from starting points that are determined using the Sobol sequence in a pre-testing phase (Guvenen, 2011; Arnoud, Guvenen and Kleineberg, 2019). See Appendix C for details.

Table 4: Estimation results

	$\underline{d}$	η	χ	β	$\underline{a}$
	Parameter estimates				
	0.016	0.9354	0.4651	0.8614	-0.2233
	(0.0045)	(0.0165)	(0.0272)	(0.0008)	(0.0206)
	Auxiliary model and other moments				
	δ_0	δ_1	$\overline{N}$	$\frac{a}{y}\Big\rvert_{50}$	$\frac{a}{y}\Big\rvert_{10}$
Data	0.3571	0.0646	0.3627	1.0551	-0.2527
Model	0.3571	0.0646	0.3626	1.0562	-0.2530

Note. Asymptotic standard errors in parentheses are computed using a non-parametric panel bootstrap with 200 repetitions; see Appendix C for details.

3.5 Estimation results

3.5.1 Parameter estimates

Table 4 shows the estimation results. I construct asymptotic standard errors using a non-parametric panel bootstrap with 200 repetitions (Appendix C).

Regarding the utility parameters, I estimate $\eta = 0.9354$ and $\underline{d} = 0.016$. Both parameters are precisely estimated and both coefficients of the auxiliary regression are close to their respective model counterpart. η is comparable with values in the literature, whereas $\underline{d}$ is relatively larger (Fried et al., 2018).[16]

Average hours worked, as a fraction of agents' time endowment, is pinned down well by the disutility of labor χ, which is estimated to equal 0.4651. Again, this parameter is estimated precisely.

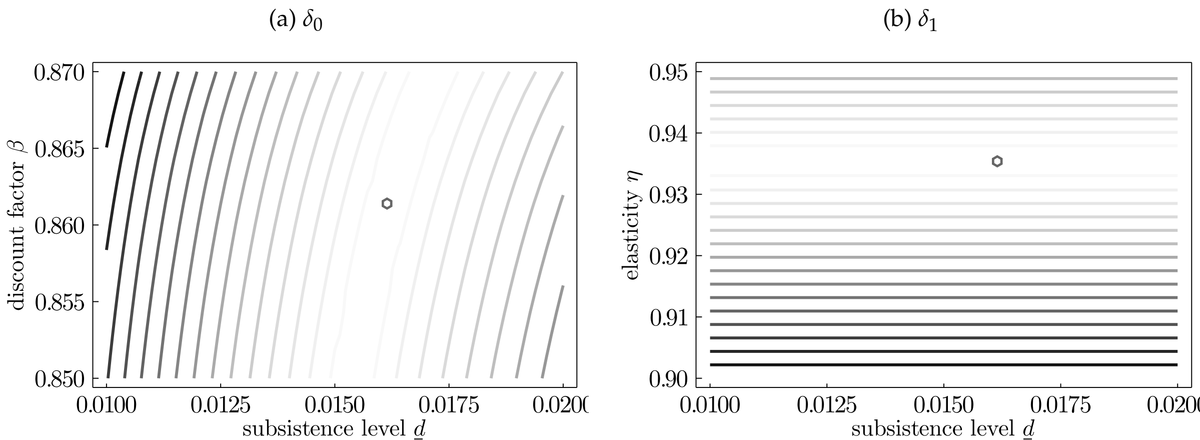
Given an intertemporal substitution elasticity of $1/\gamma = 1/2$, the discount factor β is relatively high reflecting the 5-year time period of the model. The borrowing limit is $\underline{a} = -0.2233$, which amounts to 33% of average gross income that can be borrowed every period. The model matches the two targeted moments of the wealth-to-income distribution well.

3.5.2 Identification

As common in these type of models, I have no proof of global identification of my parameters. However, I want to mention two points that indicate proper identification of the five parameters. First, I assess how different model-implied moments are affected when I change two of the five parameters and fix the remaining three at their best fit value. Figure 3 shows the results of this exercise and plots of moments implied by the structural model from their data counterparts when varying the parameters on the axes; lighter areas depict smaller deviations.

[16] Different calibrations, of course, give rise to different subsistence levels of dirty goods consumption. Hence, when comparing $\underline{d}$ to the literature, I compute expenditure on subsistence consumption, $p_d\underline{d}$.

Figure 3: Illustrating the identification of $\underline{d}$ and η



Note. Absolute deviation of moments implied by the structural model from their data counterparts when parameters on the respective axis are varying. Other parameters are fixed at their estimated value. Darker regions indicate higher deviation.

Contour plots indicate that the utility parameters are identified in two steps, as indicated by the decision rule in the static model. The relative prefence parameter is identified by the model-implied estimate for δ_1. Conditional on this value, δ_0 identifies the subsistence parameter of consumption. Similar arguments hold for the identification of the discount factor and the borrowing limit.

Second, after searching a grid of potential values in my optimization algorithm I choose the best 10% and start a local search step from these points. This search procedure converges to the same best fit values for various starting points,[17] which suggests that the model is globally identified.

4 Quantitative exercise

In this section, I use the estimated model as laboratory to answer the main question. To what extent do idiosyncratic risk and precautionary savings matter for the welfare-maximizing carbon tax? To this end, I first specifiy a social welfare function as the objective for the government (social planner). Next, I look for the welfare-maximizing carbon tax in general equilibrium, when the lump-sum transfer is adjusted to clear the government budget constraint, and other tax instruments are kept fixed. That is, the government collects carbon taxes and reimburses revenues via transfers. I then compare the level of this tax to various changes in the economic environment. Lastly, I repeat the main exercise if the government is also allowed to adjust the average labor income tax as well.

From now on, the economy is in general equilibrium (GE). In particular, I use the estimated and calibrated parameters, but let prices adjust to clear markets. Moreover, in the initial steady state without taxes, I back out the carbon cycle and damage function parameters as explained above. When optimizing over the carbon tax, the technology parameter $\tilde{X}$ is fixed, hence, any change in total factor productivity X is due to changes in climate damages.

[17] By construction, the algorithm with every new local search puts more weight on the current local minimum when setting the new starting value. However, it also converges to the same values for early starting values.

Social welfare function I assume that the social planner is utilitarian and maximizes social welfare defined over households' value functions in stationary equilibrium:

$$SW = \int_{(A\times\Theta)} V_{(\tau_d)}(a,\theta)\, d\Lambda_{(\tau_d)}. \tag{19}$$

In this specification, every household gets the same welfare weight. However, due to concavity in the utility function, the government has an implicit preference for redistribution, as the marginal utility of poorer households is higher than that of the rich.

The planner chooses τ_d to maximize Equation (19) while setting g to balance the government budget. Hence, the subscript (τ_d) stresses that the value function and invariant distribution are associated with this particular carbon tax. I also allow the government in another exercise to adjust both the carbon tax and the average labor income tax τ_0.

4.1 Welfare-maxizing carbon taxes

Optimal carbon taxes The optimal carbon tax equals $\tau_d^* = 1.096$, about 42% of the energy price, and a clean energy share of $\mu = 0.357$. Taking into account the pass-through parameter ω, this translates into a carbon price of about 141$ per ton of CO_2.[18]

Changes in aggregates and inequality Before discussing the impact of idiosyncratic risk and precautionary savings, it is instructive to first understand how the carbon tax affects aggregates and inequality. Thus, I compute and compare aggregate quantities and prices of the GE economy with no tax and the same economy with τ_d^*. Table 5 shows the results, where changes are displayed in percent between the two economies.

We see that households substitute away from dirty goods and increase their consumption in clean goods. Total dirty consumption decreases by 16.82% as the dirty good becomes more expensive. Moreover, average hours worked decrease. Note, however, that labor supply in efficiency units decreases by less, which suggests that labor supply shifts from less to more productive households. In fact, all three production inputs decrease. Nevertheless, overall output increases, because changes on the input side are dwarfed by reductions in emissions and thus environmental damages. Both interest rates and wages increase due to lower environmental damages which increase the marginal products of capital and labor, respectively.

Table 6 shows the results for household inequality statistics. In this case, I compare Gini coefficients for the different consumption goods and labor incomes before and after introducing the carbon tax. With respect to assets, which can attain negative values, I compare percentile differences. We see four sets of results. First, irrespective of the carbon tax, the gini on dirty goods consumption is smaller than the one on clean goods consumption. Intuitively, the subsistence level compresses the dirty consumption distribution from below.

[18] The first-order condition of the energy firm implies $\mu = ((\tau_d(1-\omega))/(c_1 c_2))^{\frac{1}{c_2-1}}$. Using the definition of the backstop price and its implied value of 550$ per ton of CO_2, I compute the implied \$/ton-CO_2 price as $(550/(1-\omega))\mu^{c_2-1}$.

Table 5: Changes in aggregate variables with optimal carbon tax

Variables	Percent change
Clean consumption, C	3.00
Dirty consumption, D	-16.82
Hours worked, N	-1.34
Assets, A	-2.12
Labor, L	-1.08
Capital, K	-2.57
Energy, E	-16.87
Output, Y	1.72
Wage, w	0.64
Interest rate, r	1.87
Transfers, g	41.54
Emissions, S	-46.54

Note. This table compares the changes in aggregate variables of the benchmark GE economy with and without the (optimized) carbon tax.

Second, the gini coefficient with the carbon tax is smaller for both the clean and dirty consumption than under no carbon tax, albeit for the clean good only slightly. A higher carbon tax increases the price of the dirty good, hence, households substitute away from dirty to clean. However, since the elasticity of substitution between the two goods is increasing in cash-on-hand due to the subsistence level (Baumgärtner et al., 2017), richer households decrease their dirty goods consumption by more and the dirty goods distribution gets compressed from above.

Third, gross labor income increases, whereas net labor income increases. Inequality in hours worked decreases slightly, but higher wages increase inequality in gross labor income as work shifts to richer households. Net labor income, on the other hand, takes into account progressive taxation and lump-sum transfers. Redistributing carbon tax revenue as lump-sum transfers, then, decreases net labor income inequality.

Fourth, asset inequality decreases. Instead of looking at gini coefficients I compute differences in the 5th and 50th percentiles that capture movements in the left part of the distribution as well as in the 50th and 95th percentiles that capture the right part of the distribution. Overall, the asset distribution shifts to the left with a higher share of agents at the borrowing constraint. The left part of the distribution moves by more than the right, as the P50-P5 distance decreases and the P95-P50 distance increases.

Welfare To understand where the welfare gains are coming from in the optimal carbon tax economy, I compute consumption equivalent variations (CEV). The CEV is the change in the consumption composite that makes an agent indifferent from switching from the pre-tax stationary equilibrium consumption-labor allocation to the one obtained under the optimal carbon tax. In particular,

Table 6: Inequality statistics

	Gini coefficient			
	Dirty consumption	Clean consumption	Gross labor income	Net labor income
Initial calibration	0.111	0.182	0.533	0.294
Optimal τ_d	0.095	0.178	0.550	0.280

	Percentile difference in assets	
	P50-P5	P95-P50
Initial calibration	0.685	2.767
Optimal τ_d	0.662	2.771

Note. This upper panel shows the gini coefficient for dirty and clean consumption, gross and net labor income for the initial calibration and after introducing the optimal carbon tax. The lower panel shows differences in percentiles in asset holdings for the initial calibration and after introducing the optimal carbon tax.

the CEV solves

$$\mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \tilde{u}(\tilde{c}_*, n_*) = \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \tilde{u}((1+CEV)\tilde{c}_0, n_0),$$

where the $\tilde{c}$ refers to the consumption composite $\tilde{c} = c^{\eta}(d - \underline{d})^{1-\eta}$. Denote the infinitely discounted sum of utilities as W such that the left-hand side is $W(\tilde{c}_*, n_*)$. Moreover, note that due to separability of the preferences, we can write $W(\tilde{c}_*, n_*) = W(\tilde{c}_*) + W(n_*)$. Then, the CEV is given by

$$CEV = \left(\frac{W(\tilde{c}_*, n_*) - W(n_0)}{W(\tilde{c}_0)} \right)^{\frac{1}{1-\gamma}} - 1.$$

Moreover, following Conesa et al. (2009), it is possible to split the contributions to the CEV which are stemming from changes in the consumption composite and changes in labor supply. The contribution of the change in consumption (or labor supply) can then further decomposed into a level effect and a distributional effect. The details of these decompositions are in Appendix D.4.

Table 7: Welfare gains from carbon taxation

Total CEV: 2.63					
	Consumption			Leisure	
Total	Level	Distribution	Total	Level	Distribution
1.48	0.69	0.78	1.13	0.93	0.21

Note. This table shows the welfare gains stemming from optimal carbon taxation expressed as consumption equivalent variation (CEV). The CEV is further decomposed in a component stemming from the change in consumption and labor supply (leisure), respectively. Each of these components can be further splitted in a level and a distributional effect.

Overall, the new stationary equilibrium features higher aggregate welfare, equivalent to an increase in the consumption composite of 2.63%. Table 7 shows that this increase stems virtually from all components of the CEV decomposition. The biggest contributor is the reduction in average labor supply, however, both the level effect and the distributional effect of consumption are quantitatively relevant.

The role of idiosyncratic risk To analyze the role of idiosyncratic risk and precautionary savings, I create a mean-preserving contraction/spread in the variance of the persistent part of the labor productivity of agents. In particular, I scale the variance of the persistent shock by a factor φ and re-normalize the mean productivity to one. Thus, $\varphi > 1$ represents an increase in the variance and so more risk. Moreover, I also compare different recalibrations: A comparative statics (CS) exercise in which all parameters are as in the benchmark, an exercise in which I adjust χ and β to match average hours worked and the capital-to-income ratio, respectively, as in the benchmark GE economy without the tax, and intermediate steps thereof.

The idea behind this exercise is that an increase in labor productivity risk increases the precautionary motive of households. Future labor productivity is more risky, hence, they respond by saving more or working more.

Figure 4 shows the results. I display the %-deviation to the benchmark carbon tax ($\varphi = 1$) under four different recalibrations and different levels of φ. Three things stand out. First, irrespective of recalibrating the model, an increase in idiosyncratic risk and precautionary savings increases the optimal carbon tax compared to the benchmark.

Second, the change in the carbon tax due to re-scaling labor productivity risk is fairly symmetric. For instance, the optimal carbon tax under CS decreases by approximately 6% when the variance of the persistent shock is reduced by 25% and increases by approximately 6% when the variance is raised by 25%.

Third, while matching average hours worked only slightly increases the change in the optimal carbon tax, compared to a CS exercise, matching the capital-to-output ratio decreases the impact of a change in labor productivity risk by about 1.2 percentage points. When matching both, an increase of the persistent shock variance by 25% raises the carbon tax by about 3.8%.

Contributors to the carbon tax In order to better understand what components gives rise to a non-zero carbon tax in this economy and to the positive gradient with respect to the degree labor productivity risk, I now conduct several exercises that change various features of the model. First, I introduce a skill-specific incidence rule, $\bar{g}(\theta)$, for lump-sum transfers. In the benchmark economy, every households gets the same level of transfers; g is non-individualized. With an incidence rule, I make the transfer proportional to a household's labor productivity type. In particular, the functional form for the rule is

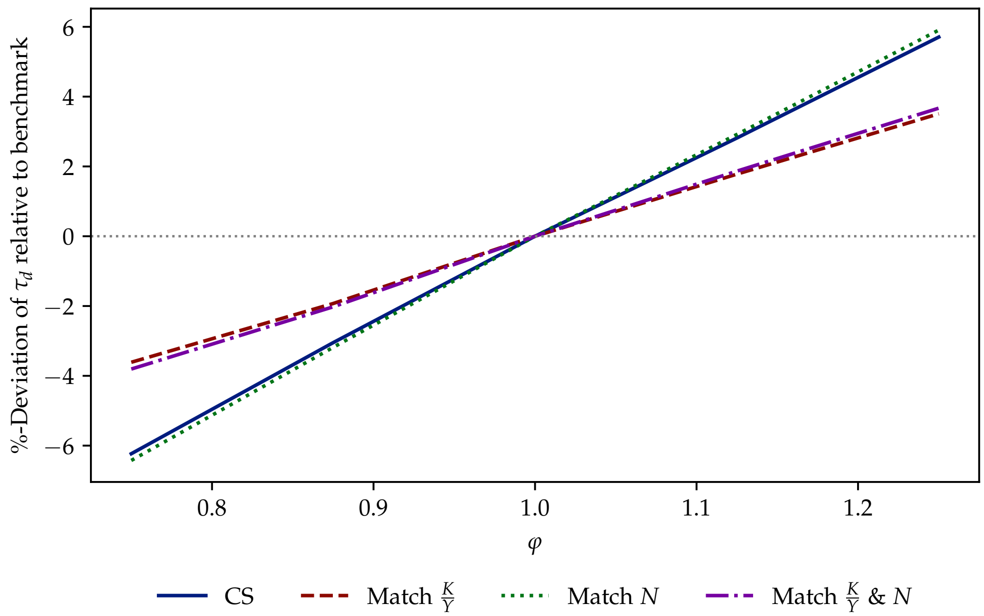
$$\bar{g}(\theta) = \frac{\theta}{\sum_S \theta_s f(\theta_s)},$$

and the budget constraint is now specified as

$$c + \tilde{p}d + a' \leq (1 + r(1 - \tau^k))a + \mathcal{T}(w\theta n) + g\bar{g}(\theta).$$

A proportional incidence rule reduces the progressivity of the carbon tax-transfer combination. Moreover, and specific to the economy at hand, it also reduces the insurance and redistributive

Figure 4: Optimal carbon tax for different levels of idiosyncratic productivity risk relative to benchmark



Note. This figure shows the percentage change of the optimal carbon tax relative to the benchmark for different levels of idiosyncratic risk, which is scaled with the factor φ. The different curves depict different choices regarding recalibration of the economy prior to optimizing the carbon tax: The blue solid curve is a compartive statics (CS) exercise that keeps all parameters as in benchmark and the purple dashed-dotted line matches both average hours worked, N, as well as the capital-to-output ratio, K/Y. Other curves show the intermediate cases.

capabilities the government provides to households. When a negative labor productivity shock occurs for a household, the transfer received is also smaller than previous period.

The first column of Table 8 shows the optimal carbon taxes under this incidence rule. Compared to the benchmark economy tax is approximately halved. Forgoing a uniform distribution of transfers increases the carbon tax burden on poorer households, who now receive lower transfers while still spending relatively more for dirty goods.

The last three columns compute the optimal carbon tax under no damages, no subsistence level, and both, respectively. Under no damages, I set $\xi = 0$. The optimal carbon tax is then around 0.066, considerably lower than in the benchmark case. In this setup, a carbon tax increases the price of the dirty good and its revenues can be redistributed, but it does not provide indirect economic benefits by increasing total factor productivity through lower climate damages. The carbon tax has been reduced to a pure consumption tax on the dirty good.

Under no subsistence level, I set $\underline{d} = 0$. In this case, the optimal carbon tax is slightly higher compared to the benchmark. The presence of subsistence needs makes the carbon tax, in and of itself, regressive. Hence, poorer households are affected relatively more by the tax. As such, removing the subsistence level gives room for the government to set higher carbon taxes. From a quantitative point of view, however, the change in carbon taxes is small.

Table 8: Optimal carbon taxes under different model specifications

Benchmark: 1.096				
	Transfer incidence	No damages	No subsistence level	No damages and no subsistence level
Comparative statics	0.574	0.066	1.099	0.067
Match K/Y & N	0.518	0.067	1.164	0.092

Note. This table compares the optimal carbon tax under different model specifications. The first column implements the optimal carbon tax under a proportional skill-specific transfer rule instead of lump-sum transfers. The second column removes economic damages. The third column removes the subsistence level. The fourth column computes the optimal carbon tax under no subsistence level and no economic damages.

Table 9: Optimal carbon taxes under the efficient welfare criterion

Utilitarian welfare	Efficient welfare	Efficient welfare & No damages
1.096	1.036	0.0047

Note. This table compares the optimal carbon tax under the utilitarian welfare criterion and an efficient welfare criterion inspired by Bénabou (2002) (middle column). The last column, in addition, removes economic damages.

Equity vs efficiency The utilitarian welfare criterion used so far conflates redistributive and efficiency/insurance concerns of the economy. Even in an economy without idiosyncratic labor productivity shocks, redistributing from rich to poor increases welfare due to the concavity of the utility function (Bénabou, 2002). Hence, to isolate equity and efficiency considerations, I follow Bénabou (2002) and Bakış, Kaymak and Poschke (2015) and construct a different welfare measure which aggregates certainty-equivalent levels of consumption and hours worked, instead of utility levels. To be precise, the effiecient welfare is constructed as follows:

$$SW^{Eff} = \frac{1}{1-\gamma}\left(\int \bar{\tilde{c}}(a,\theta)d\Lambda\right)^{1-\gamma} + \frac{\chi}{1-\epsilon}\left(\int (1-\bar{n}(a,\theta))d\Lambda\right)^{1-\epsilon},$$

where $\bar{\tilde{c}}$ and $\bar{n}$ denote certainty-equivalent levels of the consumption-composite and hours-worked, respectively. The certainty-equivalent level of the consumption-composite solves the equation

$$\mathbb{E}\sum_t \beta^t \frac{\tilde{c}(a,\theta)^{1-\gamma}}{1-\gamma} = \frac{1}{1-\beta}\frac{\bar{\tilde{c}}(a,\theta)^{1-\gamma}}{1-\gamma};$$

similarly for hours worked.

Table 9 shows the optimal carbon tax under this new welfare criterion and compares it to the utilitarian benchmark. The carbon tax under the efficient welfare criterion is about 6% lower than under the utilitarian welfare criterion. This suggests that equity concerns have a positive impact on the carbon tax. As discussed above, this is due to the lump-sum redistribution of transfers, without which the carbon tax would be regressive in the presence of a subsistence level.

To further understand the contribution of efficiency considerations, the last column then shows the optimal carbon tax under the efficient welfare criterion when there are no damages ($\xi = 0$); in this case, the model has been recalibrated to match average hours worked and the capital-to-output ratio. The resulting optimal carbon tax is very small with an optimized value of 0.005. Comparing this to the no-damage value from Table 8 suggests that redistribution due to equality concerns is considerably larger than for efficiency concerns.

Lastly, note that while Table 5 suggests that labor supply shifts from low productive to high productive households and thereby increasing efficiency, when looking at average labor productivity this effect seems to be quite small. Average labor productivity in the economy, L/N, changes from 1.012 in the initial calibration to 1.015 in the economy under optimal carbon taxes (Table D.4).

Taking stock The optimal level of the carbon tax seems to be driven by the size of climate damages, but also inequality and insurance concerns of the government. Households, especially income-poor ones, benefit from higher lump-sum transfers in form of recycled carbon tax revenue as well as higher wages and interest rates due to lower climate damages. Both direct and indirect components contribute to more welfare in a high-risk environment, yielding a positive relation between carbon taxes and uninsurable idiosyncratic risk and precautionary savings.

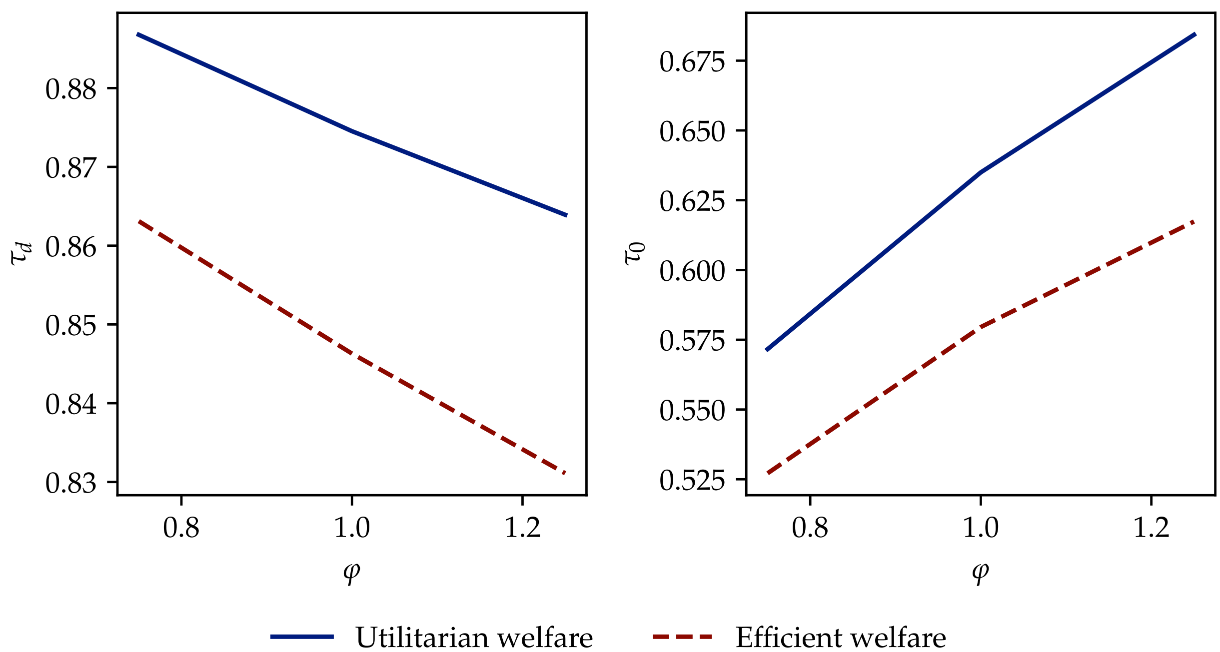
4.2 Adjusting labor income taxes

The exercises conducted so far are essentially third-best policy exercises. I keep labor income and capital income taxes fixed and only adjust carbon taxes. While this is of major practical relevance, as governments do not always adjust other taxes when implementing corrective carbon taxation, it is unsatisfactory from a economic theory perspective. Changing the carbon tax in this economy means that it not only adjusts to internalize the climate externality, but it also raises revenue for redistribution and insurance. When the government has more instruments at its disposal, however, there are more suitable ways to achieve the latter.

In the following, I thus allow the government to adjust the average labor tax parameter τ_0, in addition to τ_d, to maximize welfare in the stationary equilibrium. Recall that even a flat tax with a non-zero lump-sum transfer constitutes a progressive tax system, because average tax rates increase with income. Since the government in the present model does have access to non-zero lump-sum transfers, increasing the average labor income parameter enables the government to create a more progressive tax system.

Figure 5 shows the results. The left panel shows the optimal carbon tax for different labor productivity risk for both utilitarian and efficient welfare. The right panel shows the corresponding average labor tax parameter. The results suggest that the income tax does take over redistributive and insurance purposes that fell on the carbon tax before. First, the carbon tax is lower and the average labor income tax parameter is higher than under the benchmark steady state calibration. Second, the effect of idiosyncratic risk and the precautionary motive is now reversed! The carbon tax is decreasing in the scale parameter φ, whereas the average income tax parameter is increasing. These changes also hold when using the efficient welfare criterion. Intuitively, in a high-risk environment, higher carbon taxes hurt agents relatively more, because the marginal benefits of re-

Figure 5: Optimal carbon taxes and optimal (average) labor income taxes



Note. This figure shows optimal carbon taxes and optimal average labor income taxes for different levels of idiosyncratic risk, scaled by the factor φ, if both taxes are allowed to be adjusted by the social planner. The solid line depicts results under utilitarian welfare. The dashed line depicts results under efficient welfare.

distribution and insurance are lower due to a better structure of the labor income tax system.

Indeed, Table D.3 shows that inequality in all components is lower when, in addition to carbon taxes, average labor income taxes also get optimized. The key redistribution channel are the lump-sum transfers, which are now 125% higher compared to the initial equilibrium (Table D.2). Moreover, the main welfare gains stem from a lower level of average hours worked. The CEV contribution of consumption is slightly negative, where the negative level effect of consumption slightly surpasses the distributional effect (Table D.5).

As a robustness exercise, I also let the government optimize over τ_d, τ_0 as well as τ_1; the government can also affect the tax-rate progression in the tax schedule. Figure D.1 shows that carbon taxes and average labor income taxes are still decreasing and increasing in the degree of idiosyncratic risk, respectively. Interestingly, the tax-rate progression drops from 0.768 in the initial calibration to 0.207 in the optimized economy and is decreasing in idiosyncratic risk. The main difference in this optimization exercise is that the effect of τ_1 mainly stems from improvements in efficiency. There is a strong increase in average labor productivity, and welfare gains are largest due to a level decrease in average hours worked. At the same time, inequality in consumption and net labor income increases.

5 Conclusion

In this paper, I studied the optimal carbon tax in a climate-economy model with idiosyncratic risk and borrowing constraints in general equilibrium. I first calibrated and estimated the model on U.S. household panel data. In a next step, I used the model as a laboratory and optimize over the carbon tax in a general equilibrium steady state under an utilitarian welfare criterion.

To analyze the role of idiosyncratic risk and precautionary savings, I create a mean-preserving contraction/spread in the variance of the persistent part of the labor productivity of agents. I find that the optimal carbon tax is increasing in the desire for precautionary savings. When recycling the revenue lump-sum, the optimal carbon tax carbon tax also functions as a means of redistribution and serves as an insurance device for the uninsurable idiosyncratic productivity shocks through i) transfers and ii) increasing wages and interest rates due to lower climate damages on production and thus higher factor productivity. This result depends on the availability of tax instruments of the social planner.

Lastly, I find two more avenues of future research of interest. First, the novelty in this paper was the introduction of idiosyncratic risk. Another interesting direction would be to study the distributional consequences of *aggregate* climate uncertainty. This is in particular the case if there is heterogeneous incidence of pollution damages, for instance, due to different abilities in adaptation.

Second, political support for carbon taxation or other forms of corrective pricing has so far been weak.[19] This is also partly due to distributional concerns, as the yellow-vest movements in France or Canada demonstrate (Douenne and Fabre, 2022). Thus, an exploration of household heterogeneity in a political climate-economy would be worthy of future research.

References

Aiyagari, S. Rao. 1994. "Uninsured Idiosyncratic Risk and Aggregate Saving." *Quarterly Journal of Economics*, 109(3): 659–684.

Altonji, Joseph G., and Lewis M. Segal. 1996. "Small-Sample Bias in GMM Estimation of Covariance Structures." *Journal of Business & Economic Statistics*, 14(3): 353–366.

Andreski, Patricia, Geng Li, Mehmet Zahid Samancioglu, and Robert Schoeni. 2014. "Estimates of Annual Consumption Expenditures and Its Major Components in the PSID in Comparison to the CE." *American Economic Review*, 104(5): 132–135.

Annicchiarico, Barbara, and Fabio Di Dio. 2015. "Environmental policy and macroeconomic dynamics in a new Keynesian model." *Journal of Environmental Economics and Management*, 69: 1–21.

Archer, David. 2005. "Fate of fossil fuel CO2 in geologic time." *Journal of Geophysical Research*, 110(C9): C09S05.

Arnoud, Antoine, Fatih Guvenen, and Tatjana Kleineberg. 2019. "Benchmarking Global Optimizers." National Bureau of Economic Research w26340, Cambridge, MA.

[19] As of 2020, only 18% of global CO_2 emissions are internalized (Ritchie and Rosado, 2022).

Bakış, Ozan, Barış Kaymak, and Markus Poschke. 2015. "Transitional dynamics and the optimal progressivity of income redistribution." *Review of Economic Dynamics*, 18(3): 679–693.

Barrage, Lint. 2020. "Optimal Dynamic Carbon Taxes in a Climate–Economy Model with Distortionary Fiscal Policy." *The Review of Economic Studies*, 87: 1–39.

Baumgärtner, Stefan, Moritz A. Drupp, and Martin F. Quaas. 2017. "Subsistence, Substitutability and Sustainability in Consumption." *Environmental and Resource Economics*, 67(1): 47–66.

Benmir, Ghassane, and Josselin Roman. 2022. "The Distributional Costs of Net-Zero: A Heterogeneous Agent Perspective." *Manuscript*, 100.

Bewley, Truman F. 1986. "Stationary monetary equilibrium with a continuum of independently fluctuating consumers." In *Contributions to Mathematical Economics in Honor of Gerard Debreu*. . Werner Hildenbrand and Andreu Mas-Colell (eds.) ed., 79–102. North-Holland.

Blanz, Alkis, and Matthias Kalkuhl. 2022. "Climate Change and long-run Inequality: The Case of Food prices in the Developing World." *Manuscript*.

Blanz, Alkis, Ulrich Eydam, Maik Heinemann, Matthias Kalkuhl, and Nikolaj Moretti. 2023. "Fiscal Policy and Energy Price Shocks." *CEPA Discussion Papers*, 70.

Blundell, Richard, Luigi Pistaferri, and Itay Saporta-Eksten. 2016. "Consumption Inequality and Family Labor Supply." *American Economic Review*, 106(2): 387–435.

Blundell, Richard, Xiaohong Chen, and Dennis Kristensen. 2007. "Semi-Nonparametric IV Estimation of Shape-Invariant Engel Curves." *Econometrica*, 75(6): 1613–1669.

Bénabou, Roland. 2002. "Tax and Education Policy in a Heterogeneous-Agent Economy: What Levels of Redistribution Maximize Growth and Efficiency?" *Econometrica*, 70(2): 481–517.

Bovenberg, A. Lans, and Ruud A. de Mooij. 1994. "Environmental Levies and Distortionary Taxation." *The American Economic Review*, 84(4): 1085–1089.

Cameron, Colin A., and Pravin K. Trivdei. 2005. *Microeconometrics: methods and applications.* Cambridge University Press.

Carroll, Christopher D. 2006. "The method of endogenous gridpoints for solving dynamic stochastic optimization problems." *Economics Letters*, 91(3): 312–320.

Carroll, Christopher D., and Andrew A. Samwick. 1997. "The nature of precautionary wealth." *Journal of Monetary Economics*, 40(1): 41–71.

Carroll, Christopher D., and Andrew A. Samwick. 1998. "How Important Is Precautionary Saving?" *The Review of Economics and Statistics*, 80(3): 410–419.

Carroll, Christopher D., and Miles S. Kimball. 1996. "On the Concavity of the Consumption Function." *Econometrica*, 64(4): 981–992.

Carroll, Christopher D., Martin B. Holm, and Miles S. Kimball. 2021. "Liquidity constraints and precautionary saving." *Journal of Economic Theory*, 195: 105276.

Conesa, Juan Carlos, Sagiri Kitao, and Dirk Krueger. 2009. "Taxing Capital? Not a Bad Idea After All!" *American Economic Review*, 99(1): 25–48.

Deaton, Angus. 1979. "Optimally uniform commodity taxes." *Economics Letters*, 2(4): 357–361.

Deaton, Angus. 1981. "Optimal Taxes and the Structure of Preferences." *Econometrica*, 49(5): 1245–1260.

Douenne, Thomas, Albert Jan Hummel, and Marcelo Pedroni. 2023. "Optimal Fiscal Policy in a Second-Best Climate Economy Model with Heterogeneous Agents." *Manuscript*.

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

Dyrda, Sebastian, and Marcelo Pedroni. 2023. "Optimal Fiscal Policy in a Model with Uninsurable Idiosyncratic Income Risk." *The Review of Economic Studies*, 90(2): 744–780.

Flodén, Martin, and Jesper Lindé. 2001. "Idiosyncratic Risk in the United States and Sweden: Is There a Role for Government Insurance?" *Review of Economic Dynamics*, 4(2): 406–437.

French, Eric. 2004. "The Labor Supply Response to (Mismeasured but) Predictable Wage Changes." *The Review of Economics and Statistics*, 86(2): 602–613.

Fried, Stephie, Kevin Novan, and William B. Peterman. 2018. "The distributional effects of a carbon tax on current and future generations." *Review of Economic Dynamics*, 30: 30–46.

Fried, Stephie, Kevin Novan, and William B. Peterman. 2021. "Recycling Carbon Tax Revenue to Maximize Welfare." *Finance and Economics Discussion Series*, 2021(023): 1–30.

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

Goulder, Lawrence H., Marc A.C. 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.

Gourieroux, C., A. Monfort, and E. Renault. 1993. "Indirect Inference." *Journal of Applied Econometrics*, 8: S85–S118.

Gouveia, Miguel, and Robert P. Strauss. 1994. "Effective federal individual income tax functions: an exploratory empirical analysis." *National Tax Journal*, 47(2): 317–339.

Guariglia, Alessandra. 2003. "Saving behaviour and earnings uncertainty: Evidence from the British Household Panel Survey." In *Family, Household and Work*. , ed. Klaus F. Zimmermann and Michael Vogler, 135–150. Berlin, Heidelberg:Springer Berlin Heidelberg.

Guariglia, Alessandra, and Mariacristina Rossi. 2002. "Consumption, habit formation, and precautionary saving: evidence from the British Household Panel Survey." *Oxford Economic Papers*, 54(1): 1–19.

Guerrieri, Veronica, and Guido Lorenzoni. 2017. "Credit Crises, Precautionary Savings, and the Liquidity Trap." *Quarterly Journal of Economics*, 132(3): 1427–1467.

Guvenen, Fatih. 2011. "Macroeconomics with Heterogeneity: A Practical Guide." 73.

Guvenen, Fatih, and Anthony A. Smith, Jr. 2014. "Inferring Labor Income Risk and Partial Insurance From Economic Choices." *Econometrica*, 82(6): 2085–2129.

Hassler, John, Per Krusell, and Conny Olovsson. 2021. "Directed Technical Change as a Response to Natural Resource Scarcity." *Journal of Political Economy*, 129(11): 3039–3072.

Heathcote, Jonathan, Kjetil Storesletten, and Giovanni L. Violante. 2010. "The Macroeconomic Implications of Rising Wage Inequality in the United States." *Journal of Political Economy*, 118(4): 681–722.

Heutel, Garth. 2012. "How should environmental policy respond to business cycles? Optimal policy under persistent productivity shocks." *Review of Economic Dynamics*, 15(2): 244–264.

Huggett, Mark. 1993. "The risk-free rate in heterogeneous-agent incomplete-insurance economies." *Journal of Economic Dynamics and Control*, 17(5-6): 953–969.

Huggett, Mark. 2004. "Precautionary Wealth Accumulation." *The Review of Economic Studies*, 71(3): 769–781.

Huggett, Mark, and Sandra Ospina. 2001. "Aggregate precautionary savings: when is the third derivative irrelevant?" *Journal of Monetary Economics*, 48(2): 373–396.

IPCC. 2007. "Climate change 2007 the physical science basis." In: Solomon, S., Qin, D., Manning, M., Marquis, M., Averyt, K., Tignor, M.M.B., Miller Jr., H.L.R., Chen, Z. (Eds.), Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press.

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

Jacobs, Bas, and Frederick van der Ploeg. 2019. "Redistribution and pollution taxes with non-linear Engel curves." *Journal of Environmental Economics and Management*, 95: 198–226.

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.

Jappelli, Tullio, and Luigi Pistaferri. 2017. *The Economics of Consumption: Theory and Evidence.* Oxford University Press.

Jensen, Martin Kaae. 2018. "Distributional Comparative Statics." *The Review of Economic Studies*, 85(1): 581–610.

Jordà, Òscar, Katharina Knoll, Dmitry Kuvshinov, Moritz Schularick, and Alan M. Taylor. 2019. "The Rate of Return on Everything, 1870–2015." *The Quarterly Journal of Economics*, 134(3): 1225–1298.

Kimball, Miles S. 1990*a*. "Precautionary Saving and the Marginal Propensity to Consume." National Bureau of Economic Research Working Paper 3403.

Kimball, Miles S. 1990*b*. "Precautionary Saving in the Small and in the Large." *Econometrica*, 58(1): 53–73.

Klenert, David, and Linus Mattauch. 2016. "How to make a carbon tax reform progressive: The role of subsistence consumption." *Economics Letters*, 138: 100–103.

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

Känzig, Diego R. 2023. "The unequal economic consequences of carbon pricing." *Manuscript*.

Kotlikoff, Laurence, Felix Kubler, Andrey Polbin, and Simon Scheidegger. 2021*a*. "Pareto-improving carbon-risk taxation." *Economic Policy*, 36(107): 551–589.

Kotlikoff, Laurence, Felix Kubler, Andrey Polbin, Jeffrey Sachs, and Simon Scheidegger. 2021*b*. "Making Carbon Taxation a Generational Win Win." *International Economic Review*, 62(1): 3–46.

Krueger, D., K. Mitman, and F. Perri. 2016. "Macroeconomics and Household Heterogeneity." In *Handbook of Macroeconomics*. Vol. 2, 843–921. Elsevier.

Leland, Hayne E. 1968. "Saving and Uncertainty: The Precautionary Demand for Saving." *The Quarterly Journal of Economics*, 82(3): 465–473.

Lugilde, Alba, Roberto Bande, and Dolores Riveiro. 2018. "Precautionary saving in Spain during the great recession: evidence from a panel of uncertainty indicators." *Review of Economics of the Household*, 16(4): 1151–1179.

Lugilde, Alba, Roberto Bande, and Dolores Riveiro. 2019. "Precautionary Saving: A Review of the Empirical Literature." *Journal of Economic Surveys*, 33(2): 481–515.

Nordhaus, William D. 1991. "To Slow or Not to Slow: The Economics of The Greenhouse Effect." *The Economic Journal*, 101(407): 920–937.

Nordhaus, William D. 1992. "An Optimal Transition Path for Controlling Greenhouse Gases." *Science*.

Nordhaus, William D. 1993. "Rolling the 'DICE': an optimal transition path for controlling greenhouse gases." *Resource and Energy Economics*, 15(1): 27–50.

Pedroni, Marcelo, Swapnil Singh, and Christian A Stoltenberg. 2022. "Advance Information and Consumption Insurance: Evidence and Structural Estimation." *Tinbergen Discussion Paper 22-032/VI*.

Pigou, Arthur. 1920. *The Economics of Welfare.* London:Macmillan.

Piketty, Thomas, and Gabriel Zucman. 2014. "Capital is Back: Wealth-Income Ratios in Rich Countries 1700–2010." *The Quarterly Journal of Economics*, 129(3): 1255–1310.

Powell, M J D. 2009. "The BOBYQA algorithm for bound constrained optimization without derivatives." *Cambridge NA Report NA2009/06*, 26.

Ritchie, Hanna, and Pablo Rosado. 2022. "Which countries have put a price on carbon?" *Our World in Data*.

Rognlie, Matthew. 2016. "Deciphering the Fall and Rise in the Net Capital Share: Accumulation or Scarcity?" *Brookings Papers on Economic Activity*, 2015(1): 1–69.

Rouwenhorst, Geert K., ed. 1995. *Asset Pricing Implications of Equilibrium Business Cycle Models. Frontiers of Business Cycle Research*. Thomas F. Cooley, ed. Princeton ed., New Jersey: Princeton University Press.

Sandmo, A. 1970. "The Effect of Uncertainty on Saving Decisions." *The Review of Economic Studies*, 37(3): 353–360.

Sandmo, Agnar. 1975. "Optimal Taxation in the Presence of Externalities." *The Swedish Journal of Economics*, 77(1): 86–98.

Smith, Jr., Anthony A. 1993. "Estimating Nonlinear Time-Series Models Using Simulated Vector Autoregressions." *Journal of Applied Econometrics*, 8: 63–84.

Stock, James, and Motohiro Yogo. 2005. "Asymptotic distributions of instrumental variables statistics with many instruments." *Identification and inference for econometric models: Essays in honor of Thomas Rothenberg*, 6: 109–120.

Stoltenberg, Christian A, and Arne Uhlendorff. 2023. "Consumption choices and earnings expectations: empirical evidence and structural estimation." 52.

Stone, Richard. 1954. "Linear Expenditure Systems and Demand Analysis: An Application to the Pattern of British Demand." *The Economic Journal*, 64(255): 511–527.

Straub, Ludwig. 2019. "Consumption, Savings, and the Distribution of Permanent Income." *Manuscript*, 85.

Trabandt, Mathias, and Harald Uhlig. 2011. "The Laffer curve revisited." *Journal of Monetary Economics*, 58(4): 305–327.

van der Werf, Edwin. 2008. "Production functions for climate policy modeling: An empirical analysis." *Energy Economics*, 30(6): 2964–2979.

Young, Eric R. 2010. "Solving the incomplete markets model with aggregate uncertainty using the Krusell–Smith algorithm and non-stochastic simulations." *Journal of Economic Dynamics and Control*, 34(1): 36–41.

Zeldes, Stephen P. 1989*a*. "Consumption and Liquidity Constraints: An Empirical Investigation." *Journal of Political Economy*, 97(2): 305–346.

Zeldes, Stephen P. 1989*b*. "Optimal Consumption with Stochastic Income: Deviations from Certainty Equivalence." *The Quarterly Journal of Economics*, 104(2): 275.

Appendix

A Model - Details

A.1 Households

I repeat the recursive household problem for ease of exposition.

$$
\begin{aligned}
V(a,\theta) = \max_{c,d,n,a'} \; & u(c,d,n) + \beta \mathbb{E}_\theta V(a',\theta') \\
& \text{subject to} \\
& c + (p_d + (1-\mu)\varpi\tau_d)d + a' \leq (1 + r(1-\tau^k))a + \underbrace{w\theta n - T^y(w\theta n)}_{\mathcal{T}(w\theta n)} + g \\
& a' \geq \underline{a} \quad n \geq 0
\end{aligned}
$$

Defining $r(1-\tau^k) \equiv \tilde{r}$ and $p_d + (1-\mu)\varpi\tau_d \equiv \tilde{p}_d$, the Bellman equation is

$$
\begin{aligned}
V(a,\theta) = \max_{c,d,n,a'} u(c,d,n) + \beta \mathbb{E}_\theta V(a',\theta') - \pi^1 \Big(c + \tilde{p}_d d + a' - (1+\tilde{r})a - \mathcal{T}(w\theta n) - g \Big) \\
- \pi^2 \Big(\underline{a} - a' \Big) + \pi^3 n,
\end{aligned}
$$

where π^1, π^2, π^3 denote the Lagrange multipliers on the budget, borrowing, and non-negativity constraint, respectively.

In the following, I use the common notation that $\frac{\partial u(c,d,n)}{\partial c} \equiv u_c$. The first-order conditions of the household are

$$
\begin{aligned}
[c]: &\quad u_c = \pi^1 \\
[d]: &\quad u_d = \tilde{p}_d \pi^1 \\
[n]: &\quad u_n = -\pi^1 \mathcal{T}_n(w\theta n) w\theta - \pi^3 \\
[a']: &\quad \beta \mathbb{E}_\theta \left[V_a(a',\theta') \right] = \pi^1 - \pi^2
\end{aligned}
$$

Substituting out the multiplier π^1, assuming an interior solution for labor ($\pi^3 = 0$), and using the Envelope condition $V_a(a,\theta) = \pi^1(1+\tilde{r})$ we get

$$u_c \tilde{p}_d = u_d \tag{A.20}$$

$$u_n = -u_c \mathcal{T}_n(w\theta n) w\theta \tag{A.21}$$

$$u_c \geq \beta(1+\tilde{r}) \mathbb{E}_\theta \left[u_{c'} \right] \tag{A.22}$$

Equation (A.20) is the *intra-temporal* first-order condition between the two consumption goods. Re-arranging yields that in the optimum, the marginal rate of substitution (MRS) between the clean and the dirty good, u_c/u_d, equals the relative price of the clean good in terms of the dirty good,

$1/\tilde{p}_d$. Note that the marginal rate of transformation (MRT) between the two goods is $1/p_d$, hence, a positive carbon tax distorts the social optimal goods allocation.

Equation (A.21) is the *intra-temporal* first-order condition between clean consumption and labor. Similar arguments regarding the MRS and MRT as above apply. One could have stated the condition in terms of the dirty good. Again, a positive carbon tax (in addition to the labor tax) distorts the labor supply decision of the household, as it makes leisure cheaper relative to the dirty good.

Equation (A.22) is the *inter-temporal* first-order condition and is the familiar Euler equation. When the borrowing constraint is not binding, π^2 is zero and the equation holds with equality. Here, the capital income tax drives a wedge between MRS ($\frac{u_c}{\beta u_{c'}}$) and MRT $(1+r)$.

A.2 Firms

Energy producer The energy producer maximizes profits by choosing capital, K_2, labor L_2, and the fraction of abatement μ under perfect competition using a constant returns to scale technology. It takes prices (r, w, p_d), pass-through opportunities ϖ, as well as policy variables as given and obtains zero profits in equilibrium:

$$\max_{K_2, L_2, \mu} \quad \{p_d F_2(K_2, L_2) - (1-\mu)(1-\varpi)\tau_d F_2(K_2, L_2) - (r+\delta)K_2 - wL_2 - \Psi(\mu)F_2(K_2, L_2)\},$$

where I already substituted in the technology constraint $E = F_2(K_2, L_2)$.

The first-order conditions are

$$\begin{aligned}
[p_d - (1-\mu)(1-\varpi)\tau_d - \Psi(\mu)]F_{2,K_2} &= r + \delta \\
[p_d - (1-\mu)(1-\varpi)\tau_d - \Psi(\mu)]F_{2,L_2} &= w \\
\tau_d(1-\varpi) &= \Psi'(\mu)
\end{aligned}$$

Final good producer The final good firm maximizes its profits by choosing capital, K_1, labor L_1, and energy E^p under perfect competition using a constant returns to scale technology. Hence, it takes prices (r, w, p_d), policy variables as well as carbon-tax pass-through of the energy producer as given and obtains zero profits in equilibrium:

$$\max_{K,L,E} \quad \{F(K, E, L; X, S) - (r+\delta)K_1 - wL_1 - (p_d + (1-\mu)\varpi\tau_d)E^p\}$$

Moreover, the firm compensates households for depreciation, δ.

The first-order condition of the firm implied by profit maximization - substituting in the CES production function from the main text - are

$$
\begin{aligned}
p_d + (1-\mu)\varpi\tau_d = \frac{\partial F_1}{\partial E^p} &= X\left[(1-s)(K_1^\alpha L_1^{1-\alpha})^{\frac{\lambda-1}{\lambda}} + s(E^p)^{\frac{\lambda-1}{\lambda}}\right]^{\frac{1}{\lambda-1}} s(E^p)^{-\frac{1}{\lambda}} \\
w = \frac{\partial F_1}{\partial L_1} &= X\left[(1-s)(K_1^\alpha L_1^{1-\alpha})^{\frac{\lambda-1}{\lambda}} + s(E^p)^{\frac{\lambda-1}{\lambda}}\right]^{\frac{1}{\lambda-1}} (1-s)(K_1^\alpha L_1^{1-\alpha})^{\frac{-1}{\lambda}}(1-\alpha)K_1^\alpha L_1^{-\alpha} \\
r + \delta = \frac{\partial F_1}{\partial K_1} &= X\left[(1-s)(K_1^\alpha L_1^{1-\alpha})^{\frac{\lambda-1}{\lambda}} + s(E^p)^{\frac{\lambda-1}{\lambda}}\right]^{\frac{1}{\lambda-1}} (1-s)(K_1^\alpha L_1^{1-\alpha})^{\frac{-1}{\lambda}}\alpha K_1^{\alpha-1} L_1^{1-\alpha}
\end{aligned}
$$

As usual, the prices of the inputs are equal to their marginal products.

A.3 Goods market clearing

Aggregate the household budget constraint over household and impose asset market clearing $A = B + K$ in the steady-state:

$$
C + (p_d + (1-\mu)\varpi\tau_d)D + (B+K) = (1 + (1-\tau^k)r)(B+K) + wL - \int T^y d\Lambda + g.
$$

Rewrite the government budget constraint as $g - \int T^y d\Lambda = \tau^k r(K+B) + \tau_d(1-\mu)E - rB$[20], plug it in the aggregated household constraint above and collect terms:

$$
C + (p_d + (1-\mu)\varpi\tau_d)D = rK + wL + \tau_d(1-\mu)E.
$$

Extend with δK and $(p_d + (1-\mu)\varpi\tau_d)E^p$, recall that $E^p + D$, $K_1 + K_2 = K$, $L_1 + L_2 = L$, and use Euler's theorem to obtain:

$$
\begin{aligned}
C + \delta K = &\underbrace{(r+\delta)K_1 + wL_1 + (p_d + (1-\mu)\varpi\tau_d)E^p}_{Y} \\
&+ \tau_d(1-\mu)E - (p_d + (1-\mu)\varpi\tau_d)(E^p + D) + (r+\delta)K_2 + wL_2
\end{aligned}
$$

Finally, use the first-order conditions of the energy producer and again use Euler's theorem to write

$$
\begin{aligned}
C + \delta K = Y + [p_d - (1-\mu)\tau_d(1-\varpi) - \Psi(\mu)](\underbrace{F_{2,K_2}K_2 + F_{2,L_2}L_2}_{E}) + \tau_d(1-\mu)E - (p_d + (1-\mu)\varpi\tau_d)E \\
C + \delta K + \Psi(\mu)E = Y + \underbrace{[p_d - (1-\mu)\tau_d(1-\varpi)]E + \tau_d(1-\mu)E - (p_d + (1-\mu)\varpi\tau_d)E}_{0} \\
C + \delta K + \Psi(\mu)E = Y
\end{aligned}
$$

Hence, the final good can be used for consumption, investment, and abatement.[21]

[20] To be precise, the government receives carbon tax revenue from the household, $\tau_d(1-\mu)\varpi D$, from the final goods firm, $\tau_d(1-\mu)\varpi E^p$, and from the energy producer $\tau_d(1-\mu)(1-\varpi)E$. These three terms sum up to $\tau_d(1-\mu)E$.

[21] The second term on the left-hand-side equals investment I, since in the steady-state the law of motion of capital $K_{t+1} = (1-\delta)K_t + I_t$ collapses to $\delta K = I$.

B Calibration - Details

B.1 Macroeconomic variables

Capital-output-ratio (K/Y) Current-Cost Net Stock of Fixed Assets (2014) in current dollars divided by Gross Domestic Product (2014) in current dollars; both series from the FRED database with series tags K1TTOTL1ES000 and GDPA, respectively.

Bond-output-ratio (B/Y) Federal Debt Held by the Public (2014) in current dollars divided by Gross Domestic Product (2014) in current dollars; both series from the FRED database with series tags FYGFDPUN and GDPA, respectively.

B.2 Estimation of the productivity process

In the following, I explain how I estimate the labor productivity process which I use in my quantitative model.[22] I follow Flodén and Lindé (2001) and measure productivity as agent's "hourly [pre-tax] wage rate relative to all other agents" (p. 416). The data is taken from the PSID and refers to labor income as described in the main text. The productivity process is estimated on yearly data. To avoid notational clutter, I denote yearly time-steps by τ, compared with a model period (5 years) denoted by t.

As my productivity process, I take the following standard persistent-transitory specification for log wages at year τ

$$\log \hat{\theta}_{i\tau} = f(\mathbf{X}_i, \beta) + \kappa_{i\tau} + \psi_{i\tau} + \nu_{i\tau} \tag{B.23}$$

$$\kappa_{i\tau} = \rho \kappa_{i\tau-1} + \varepsilon^{\kappa}_{i\tau}, \tag{B.24}$$

where $f(\mathbf{X}_i, \beta)$ denotes a set of individual-specific controls, $\kappa_{i\tau}$ is an AR(1) process with persistence ρ and innovation-variance $\sigma^2_{\varepsilon^\kappa}$, $\psi_{i\tau}$ is a transitory component with variance σ^2_{ψ}, and $\nu_{i\tau}$ is measurement error.

Given the short time horizon when estimating the productivity process, instead of estimating the household fixed effect directly, I model the permanent component by controlling for individual-specific characteristics, $\mathbf{X}_i$.

Moreover, the measurement error term cannot be identified from the transitory term. Hence, I follow the literature and set the variance of the measurement error term to 0.02 (French, 2004; Heathcote et al., 2010; Straub, 2019).

The estimation then proceeds along the following steps. First, I residualize log wages using

$$\log \theta_{i\tau} = \log \hat{\theta}_{i\tau} - f(\mathbf{X}_i, \hat{\beta}).$$

Second, I compute empirical variances and covariances from these residuals and stack them in the

[22] The exposition here follows the one in Straub (2019) who uses a similar strategy to estimate a process for log income and from whose description I learned a lot.

Table B.10: Estimated parameters - Productivity process

ρ	$\sigma^2_{\varepsilon^\kappa}$	σ^2_ψ
0.9327	0.0426	0.0507
(0.0092)	(0.0060)	(0.0049)

Note. This table shows the estimated parameters of the productivity process. Standard errors are bootstrapped using a non-parametric block bootstrap at the household level with 500 iterations.

vector $\vec{\mathfrak{M}}$. The theoretical variances and covariances, on the other hand, can be computed using

$$\begin{aligned} var(\log\theta_{i\tau}) &= \frac{\sigma^2_{\varepsilon^\kappa}}{1-\rho^2} + \sigma^2_\psi + \sigma^2_\nu \\ cov(\log\theta_{i\tau}, \log\theta_{i\tau-h}) &= \rho^h \frac{\sigma^2_{\varepsilon^\kappa}}{1-\rho^2}. \end{aligned}$$

I denote the stacked theoretical (co)variances by $\vec{\mathfrak{m}}(\rho, \sigma^2_{\varepsilon^\kappa}, \psi_{it})$. This formulation stresses that $\vec{\mathfrak{m}}$ is a function of the parameters that we seek to estimate.

Lastly, I apply a minimum distance estimation (MDE) to minimize the weighted distance, $\mathfrak{U}(\rho, \sigma^2_{\varepsilon^\kappa}, \sigma^2_\psi) = \vec{\mathfrak{M}} - \vec{\mathfrak{m}}$, between theoretical and empirical moments/covariances:

$$\min_{\rho, \sigma^2_{\varepsilon^\kappa}, \sigma^2_\psi} \mathfrak{U}(\rho, \sigma^2_{\varepsilon^\kappa}, \sigma^2_\psi)' \, \mathfrak{W} \, \mathfrak{U}(\rho, \sigma^2_{\varepsilon^\kappa}, \sigma^2_\psi)$$

As is standard in this procedure, I use the identity matrix as weighting matrix $\mathfrak{W}$ which was shown to be more robust to small sample bias (Altonji and Segal, 1996).

Table B.10 shows the result of the MDE. Standard errors are obtained by using a non-parametric block bootstrap (Cameron and Trivdei, 2005, p.362/p.377).

5-year time period To translate these values that were estimated on annual data to their 5-year model counterparts, I proceed in two steps: First, I iterate the persistent component backward such that

$$\kappa_{i\tau} = \rho^5 \kappa_{i\tau-5} + \underbrace{\sum_{s=0}^{4} \rho^s \varepsilon^\kappa_{i\tau-s}}_{\tilde{\varepsilon}^\kappa_{i\tau}}. \tag{B.25}$$

I can compute the variance of $\tilde{\varepsilon}^\kappa_{i\tau}$ given the annual estimate:

$$\sigma^2_{\tilde{\varepsilon}^\kappa} = var(\tilde{\varepsilon}^\kappa_{i\tau}) = \sum_{s=0}^{4} \rho^{2s} var(\varepsilon^\kappa_{i\tau-s}) = \sum_{s=0}^{4} \rho^{2s} \sigma^2_{\varepsilon^\kappa} \tag{B.26}$$

This gives $\tilde{\rho} = 0.7058$ and $\sigma^2_{\tilde{\varepsilon}^\kappa} = 0.1647$.

Second, I set $\sigma^2_{\tilde{\psi}} = \sigma^2_\psi$.

The approach I take here yields similar results as the one proposed by Krueger, Mitman and

Perri (2016), who convert annual to quarterly estimates. They also take $\tilde{\rho} = \rho^5$ and $\sigma^2_{\tilde{\psi}} = \sigma^2_{\psi}$. However, they then choose $\sigma^2_{\tilde{\varepsilon}^\kappa}$ such that

$$\frac{\sigma^2_{\varepsilon^\kappa}}{1-\rho^2} = \frac{\sigma^2_{\tilde{\varepsilon}^\kappa}}{1-\tilde{\rho}^2}. \tag{B.27}$$

This would give $\sigma^2_{\varepsilon^\kappa} = 0.1643$.

C Estimation - Details

Pre-testing and local step The estimation procedure follows the TikTak algorithm in Arnoud et al. (2019). The goal is to find the parameter vector which minimizes the objective function from the main text.

In the pre-test step, I first draw $N = 1000$ quasirandom Sobol starting points in the five dimensional parameter space. Thereby, I specify bounds for all parameters and later make sure that the algorithm finds an interior solution is not constrained by these choices. For each of these starting points I solve the partial equilibrium economy, simulate an artificial panel and construct the moment conditions and evaluate the objective function.

In the next step, the local step, I pick the ten ($= 0.01N$) points which gave rise to the lowest objective value from the pre-test step. From each of these points, I use the BOBYQA algorithm by Powell (2009) to find a local minimum. The algorithm terminates if either the absolute or the fractional tolerance on the parameter vector is smaller than $1e^{-5}$. The global minimum is then the parameter vector which attains the lowest value of all these ten runs. In my application, the final objective value is $1.97e^{-7}$.

Simulation of the artifical panel I simulate an artificial panel with 25000 agents and 250 time periods. All agents start with zero assets. I then simulate the economy forward using the stationary policy functions and a simulated Markov chain of the labor productivity process. To construct the simulated moments and to run the IV regressions from the main text, I only use the last 5 time periods of the simulated panel to limit dependence on initial conditions. Moreover, the idiosyncratic shocks are always drawn using the same seed to ensure comparability between runs.

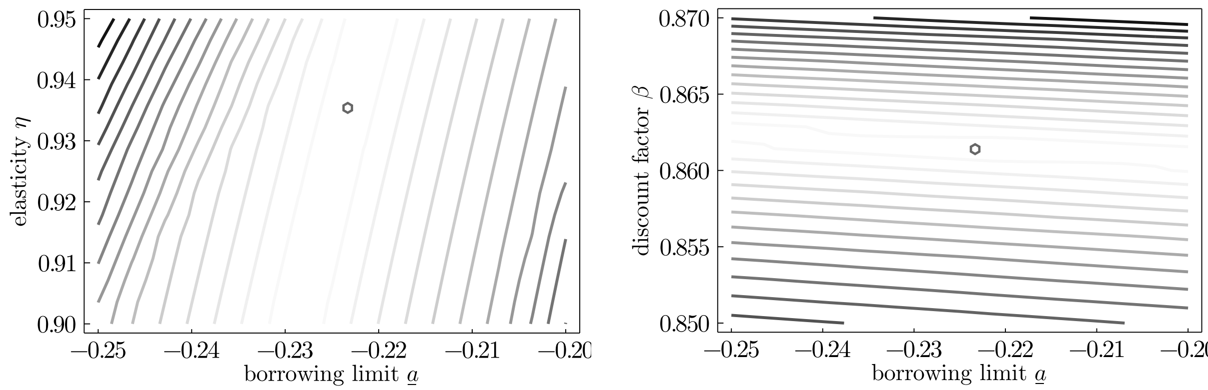
Standard errors Gourieroux et al. (1993) show that show that under no observable exogenous variables that enter the moments, the asymptotic variance-covariance matrix is

$$\mathbb{COV} = \left(1+\frac{1}{B}\right)\left[\frac{\partial\mathcal{M}^*}{\partial\Theta}^\top W\ \frac{\partial\mathcal{M}^*}{\partial\Theta}\right]^{-1}\frac{\partial\mathcal{M}^*}{\partial\Theta}^\top W\ cov(\mathcal{M}^B)\ W\ \frac{\partial\mathcal{M}^*}{\partial\Theta}\left[\frac{\partial\mathcal{M}^*}{\partial\Theta}^\top W\ \frac{\partial\mathcal{M}^*}{\partial\Theta}\right]^{-1}$$

where B denotes the number of bootstrap repetitions, $\mathcal{M}^* = \mathcal{M}(\Theta^*)$ and $cov(\mathcal{M}^B)$ is the covariance matrix of the bootstrapped moments.

Figure C.6: Illustrating the identification of $\underline{a}$ and β

(a) Wealth-to-income ratio, 10th (b) Wealth-to-income ratio, Median



Note. Absolute deviation of moments implied by the structural model from their data counterparts when parameters on the respective axis are varying. Other parameters are fixed at their estimated value. Darker regions indicate higher deviation.

In particular, I have 5770 unique households in my sample. I draw $B = 200$ random samples, with replacement, of these households to construct $cov(\mathcal{M}^B)$ from the data. The draws are panel draws (block draws), that is, when I draw a specific household, I keep all household-year observations.

The gradient $\frac{\partial \mathcal{M}^*}{\partial \Theta}$ is a Jacobian, where the elements give partial derivatives from the structural parameters to the model moments:

$$\frac{\partial \mathcal{M}^*}{\partial \Theta} = \begin{bmatrix} \frac{\partial M_1^*}{\partial \eta} & \frac{\partial M_1^*}{\partial \underline{d}} & \frac{\partial M_1^*}{\partial \chi} & \frac{\partial M_1^*}{\partial \beta} & \frac{\partial M_1^*}{\partial \underline{a}} \\ \frac{\partial M_2^*}{\partial \eta} & \frac{\partial M_2^*}{\partial \underline{d}} & \frac{\partial M_2^*}{\partial \chi} & \frac{\partial M_2^*}{\partial \beta} & \frac{\partial M_2^*}{\partial \underline{a}} \\ \frac{\partial M_3^*}{\partial \eta} & \frac{\partial M_3^*}{\partial \underline{d}} & \frac{\partial M_3^*}{\partial \chi} & \frac{\partial M_3^*}{\partial \beta} & \frac{\partial M_3^*}{\partial \underline{a}} \\ \frac{\partial M_4^*}{\partial \eta} & \frac{\partial M_4^*}{\partial \underline{d}} & \frac{\partial M_3^*}{\partial \chi} & \frac{\partial M_4^*}{\partial \beta} & \frac{\partial M_4^*}{\partial \underline{a}} \\ \frac{\partial M_5^*}{\partial \eta} & \frac{\partial M_5^*}{\partial \underline{d}} & \frac{\partial M_3^*}{\partial \chi} & \frac{\partial M_5^*}{\partial \beta} & \frac{\partial M_5^*}{\partial \underline{a}} \end{bmatrix}$$

The partial derivatives are approximated with a numerical two-sided difference. Finally, the weighting matrix W is an identity matrix. The standard errors are then on the diagonal of $\mathbb{COV}$.

Identification II Figure C.6 shows another set of numerical identification plots concerning the two moments of the wealth-to-income distribution. In panel (a), we see that the borrowing limit identifies the 10th percentile of the wealth-to-income distribution. Importantly, this results holds for a given value of β, as otherwise, the discount factor would also affect this particular moment. However, in panel (b), we see that β is pinned down by the median wealth-to-income ratio, even for different values of the borrowing limit.

D Additional tables & figures

D.1 Descriptive statistics

Table D.1: Summary statistics

Variable	Mean	Std. dev.	Median
Demographics			
Age	41.52	10.12	41
Sex	0.821	0.383	1
Household size	2.919	1.448	3
Number of children	0.990	1.203	1
Married	0.663	0.473	1
Education			
Elementary or middle school	0.0721	0.259	0
Finished high school	0.244	0.430	0
Some college	0.281	0.449	0
Finished college	0.213	0.410	0
Postgrad. qualification	0.189	0.392	0
Region			
Northeast	0.166	0.372	0
Midwest	0.302	0.459	0
South	0.337	0.473	0
West	0.195	0.397	0
Net income	32,080	19,127	28,132
Net labor income	41,119	29,023	34,637
Wealth	93,804	179,069	30,237
Nondurable consumption	14,919	7,714	13,119
Total consumption	17,136	9,502	14,798

Note. This table shows summary statistics regarding demographic and economic variables of the data which is used in estimation.

D.2 Aggregate and distributional statistics compared to benchmark

Table Table D.2 shows the change in aggregates for the quantitative exercises in Section 4.2 where the government can also adjust averages labor income taxes by changing τ_0 or, in addition, also adjust tax progressivity by changing τ_1.

Table D.2: Percent changes in aggregate variables for different optimizations compared to initial calibration

	Optimal τ_d	Optimal τ_d, τ_0	Optimal τ_d, τ_0 & τ_1
Clean consumption, C	3.00	-4.82	-4.89
Dirty consumption, D	-16.82	-18.71	-18.78
Hours worked, N	-1.34	-7.68	-11.03
Assets, A	-2.12	-13.86	-13.91
Labor, L	-1.08	-6.68	-7.27
Capital, K	-2.57	-16.76	-16.84
Energy, E	-16.87	-20.69	-20.80
Output, Y	1.72	-8.10	-8.52
Wage, w	0.64	-3.03	-2.83
Interest rate, r	1.87	39.17	37.37
Transfers, g	41.54	125.05	199.42
Emissions, S	-46.54	-45.28	-45.08

Note. This table compares the changes in aggregate variables of the economy i) under the optimized carbon tax ii) under optimized carbon taxes and average labor income taxes iii) under optimized carbon taxes, average labor income taxes, and optimal tax progressivity with the economy under no carbon taxes and income tax parameters as in the initial calibration.

Table D.3: Inequality statistics

	Gini coefficient			
	Dirty consumption	Clean consumption	Gross labor income	Net labor income
Initial calibration	0.111	0.182	0.533	0.294
Optimal τ_d	0.095	0.178	0.550	0.280
Optimal τ_d, τ_0	0.078	0.149	0.506	0.217
Optimal τ_d, τ_0 & τ_1	0.081	0.155	0.565	0.234

	Percentile difference in assets	
	P50-P5	P95-P50
Initial calibration	0.685	2.767
Optimal τ_d	0.662	2.771
Optimal τ_d, τ_0	0.640	2.298
Optimal τ_d, τ_0 & τ_1	0.590	2.507

Note. This table compares the inequality statistics of the economy i) under the optimized carbon tax ii) under optimized carbon taxes and average labor income taxes iii) under optimized carbon taxes, optimized average labor income taxes, and optimal tax progressivity with the economy under no carbon taxes and income tax parameters as in the initial calibration.

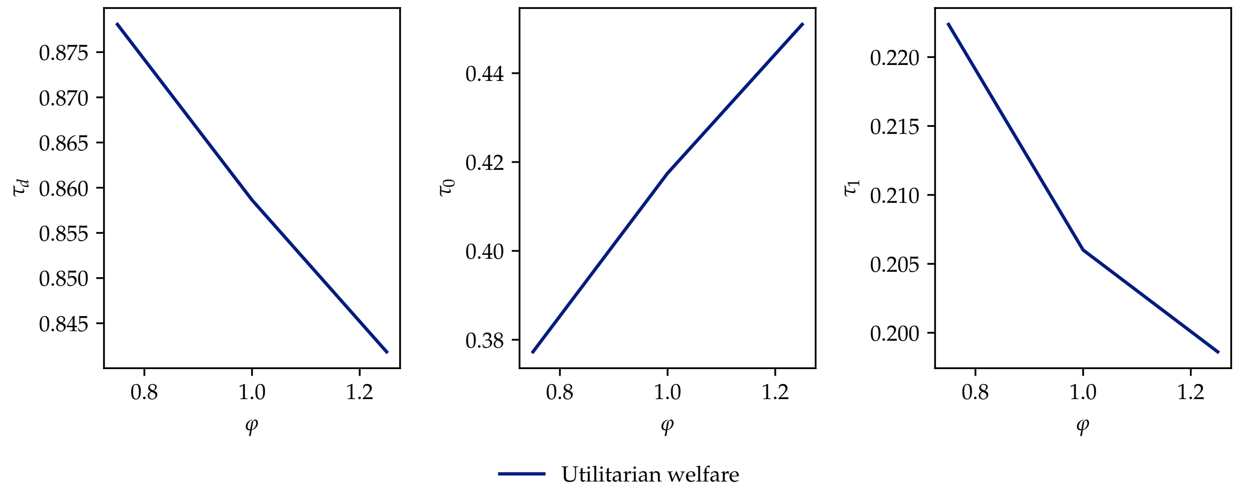
Table D.4: Labor productivity for different optimization exercises

	Initial calibration	Optimal τ_d	Optimal τ_d, τ_0	Optimal τ_d, τ_0 & τ_1
L/N	1.012	1.015	1.023	1.055

Note. This table shows the aggregate labor productivity of the economy for the initial calibration, the economy i) under the initial calibration ii) under the optimized carbon tax iii) under optimized carbon taxes and average labor income taxes and iv) under optimized carbon taxes, optimized average labor income taxes, and optimal tax progressivity. Aggregate labor productivity is measured as the ratio between total labor supply (in efficiency units) and total labor supply (hours worked): $\frac{\int n(a,\theta)\theta d\Lambda}{\int n(a,\theta) d\Lambda}$

D.3 Optimal labor income taxes II

Figure D.1: Optimal carbon taxes and optimal labor income taxes for different degrees of idiosyncratic risk



Note. This figure shows optimal carbon taxes, optimal average labor income taxes, and the optimal level of progressivity for different levels of idiosyncratic risk, scaled by the factor φ, if the taxes τ_d, τ_0, and τ_1 can be adjusted jointly by the social planner. The solid line depicts results under utilitarian welfare.

D.4 Consumption equivalent variations

Recall that CEV solves

$$\mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \tilde{u}(\tilde{c}_*, n_*) = \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \tilde{u}((1+CEV)\tilde{c}_0, n_0).$$

Similar to Conesa et al. (2009), we can define CEV_C and CEV_N that solve

$$\mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \tilde{u}(\tilde{c}_*, n_0) = \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \tilde{u}((1+CEV_C)\tilde{c}_0, n_0)$$
$$\mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \tilde{u}(\tilde{c}_*, n_*) = \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \tilde{u}((1+CEV_N)\tilde{c}_*, n_0),$$

such that

$$(1+CEV_C) = \left(\frac{W(\tilde{c}_*)}{W(\tilde{c}_0)}\right)^{\frac{1}{1-\gamma}} \quad \text{and} \quad (1+CEV_N) = \left(\frac{W(\tilde{c}_*, n_*) - W(n_0)}{W(\tilde{c}_*)}\right)^{\frac{1}{1-\gamma}}.$$

It follows that $(1+CEV) = (1+CEV_C)(1+CEV_N)$ and approximately $CEV \approx CEV_C + CEV_N$. Moreover, following again Conesa et al. (2009), it is possible to split CEV_C into a level and a distributional component defined by

$$\mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \tilde{u}(\hat{\tilde{c}}_0, n_0) = \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \tilde{u}((1 + CEV_{CL})\tilde{c}_0, n_0)$$
$$\mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \tilde{u}(\tilde{c}_*, n_0) = \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \tilde{u}((1 + CEV_{CD})\hat{\tilde{c}}_0, n_0),$$

where $\hat{\tilde{c}}_0 = \frac{\tilde{C}_*}{\tilde{C}_0}\tilde{c}_0$ is the consumption composite scaled by the ratio of aggregate consumption composite in the respective stationary equilibria. One can then show that $CEV_{CL} = \frac{\tilde{C}_*}{\tilde{C}_0} - 1$ and $CEV_{CD} = \left(\frac{W(\tilde{c}_*)}{W(\tilde{c}_0)}\right)^{\frac{1}{1-\gamma}} \frac{\tilde{C}_*}{\tilde{C}_0} - 1$ such that $(1 + CEV_C) = (1 + CEV_{CL})(1 + CEV_{CD})$.

The split between CEV_{NL} and CEV_{ND} is a different in that it is not a true decomposition. I define CEV_{NL} similar to the consumption case

$$\mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \tilde{u}(\tilde{c}_*, \hat{n}_0) = \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \tilde{u}((1 + CEV_{NL})\tilde{c}_*, n_0),$$

where $\hat{n} = \frac{1-N_*}{1-N_0}(1 - n_0)$ effectively refers to scaling leisure in the utility function. Then

$$CEV_{NL} = \left(\frac{W(\tilde{c}_*) + \left(\frac{1-N_*}{1-N_0}\right)^{1-\epsilon} W(n_0) - W(n_0)}{W(\tilde{c}_*)} \right)^{\frac{1}{1-\gamma}}$$

and CEV_{ND} is defined residually such that $(1 + CEV_N) = (1 + CEV_{NL})(1 + CEV_{ND})$.

Table D.5: Consumption equivalent variation decomposition for different optimization exercises

Optimal τ_d& τ_0 - Total CEV: 5.66					
	Consumption			Leisure	
Total	Level	Distribution	Total	Level	Distribution
-0.52	-6.24	6.11	6.20	4.98	1.15

Optimal τ_d, τ_0 & τ_1 - Total CEV: 6.54					
	Consumption			Leisure	
Total	Level	Distribution	Total	Level	Distribution
-1.15	-6.81	6.08	7.74	6.98	0.72

Note. This upper panel shows the welfare gains when both carbon taxes and average income taxes are optimized, compared to the initial calibration. The lower panel shows the welfare gains when all three carbon taxes, average income taxes, and tax progressivity are optimized, compared to the initial calibration.

E Computational appendix

E.1 Computing the household's optimal decision rules and invariant distribution

I use a variant of the endogenous gridpoint method (EGM) to solve the household's decision problem. Compared to the basic version developed by Carroll (2006), my version accommodates two goods and endogenous labor supply with possibly non-linear taxation.

Grids I represent asset positions by discrete points on a exponentially-spaced grid $\mathcal{A} \subset [\underline{a}, \overline{a}]$, where $\overline{a}$ is chosen large enough such that the upper bound is never binding. I discretize the productivty Markov process with a finite-state Markov chain using Rouwenhorst (1995)'s method. The inputs for this method, such as the persistence parameter ρ, are obtained in Appendix B.2.

Endogenous gridpoint method

Step 1 I start with a guess of the clean consumption policy function defined on the *future* asset and productivity grid, $\boldsymbol{c}(a', \theta')$. Using the intra-temporal first-order condition between clean and dirty consumption, I can express the dirty consumption policy function $\boldsymbol{d}(a', \theta')$ as a function of $\boldsymbol{c}(a', \theta')$:

$$\boldsymbol{d}(a', \theta') = \frac{1-\eta}{\tilde{p}_d \eta} \boldsymbol{c}(a', \theta') + \underline{d}. \tag{E.28}$$

Step 2 Hence, for each pair (a', θ) where the household is not constrained and the Euler equation (EE) holds with equality, I can solve *analytically* for the value $\boldsymbol{c}(a', \theta)$.[23] $\boldsymbol{c}(a', \theta)$ is essentially on the left-hand side of the EE and represents the value of consumption today, which is consistent with having a' assets tomorrow if the productivity shock today is θ:

$$u_c(\boldsymbol{c}(a', \theta)) = \beta(1+\tilde{r})\mathbb{E}_\theta \left[u_c(\boldsymbol{c}(a', \theta')) \right] \tag{E.29}$$

Note that I write u_c explicitly as a function of $\boldsymbol{c}$ only, as the utility is separable in consumption and labor, and $\boldsymbol{d}$ is implied by Equation (E.28).

Step 3 With $\boldsymbol{c}(a', \theta)$ in hand, I can solve for $\boldsymbol{n}(a', \theta)$ using the intra-temporal FOC between clean consumption and labor. In the following, I assume an interior solution:

$$-u_n(\boldsymbol{n}(a', \theta)) = u_c(\boldsymbol{c}(a', \theta))\mathcal{T}_n, \tag{E.30}$$

where $\mathcal{T}_n$ denotes $\frac{\partial \mathcal{T}}{\partial n}$. Under linearity of $\mathcal{T}$, Equation (E.30) can also be solved analytically for $\boldsymbol{n}(a', \theta)$. Otherwise, a root-finding step has to be implemented at every point in the state space. In the benchmark case, I use a version of Brent's method, modified to take into account the corner solution if $\boldsymbol{n}(a', \theta) = 0$.

[23] Of course, this step depends on the invertability of the utility function. Other functional forms for the consumption composite might not make this feasible.

Step 4 I can then invert the budget constraint to solve for the value of assets today, $a^*(a', \theta)$, which are consistent with the future assets (on grid) and the choices made above.

$$a^* = \frac{1}{1+\tilde{r}} \left(\boldsymbol{c}(a', \theta) + (p_d + \tau_d)\boldsymbol{d}(a', \theta) + a' - \mathcal{T}\left(w\theta \boldsymbol{n}(a', \theta)\right) \right), \tag{E.31}$$

implying $\tilde{c}(a^*, \theta) = \boldsymbol{c}(a', \theta)$. Note that these a^* are not on the grid (whence the name) and change each iteration. To obtain a new guess for the clean consumption policy function which is defined on the grid, I linearly interpolate on $(a^*, \tilde{c}(a^*, \theta))$ and apply this mapping to the exogenous grid a'. Use the new guess as a starting point in Step 2 above.

I repeat the above iteration procedure until convergence between two successive clean consumption policy functions is achieved: $||\boldsymbol{c}^{n+1} - \boldsymbol{c}^n|| < 10^{-7}$, where $|| \cdot ||$ denotes the supnorm and n is the iteration counter.

Density discretization With the policy functions in hand, I discretize the invariant density and iterate on it using Young (2010)'s lottery method.