Climate, technology, family size; on the crossroad between two ultimate externalities

Reyer Gerlagh*

August 16, 2022

Abstract

We extend the Brock-Mirman model with endogenous growth through variety expansion, temperature rise affecting both the level *and* build up of TFP, human capital, and endogenous fertility. We derive closed-form solutions for capital investments, research, and the Social Cost of Carbon. Calibration of the fundamentals show that the social costs of carbon associated with reduced TFP growth are very large (median: 161 €/tCO_2) compared to previous SCC estimates based on climate change TFP-level effects. We also compare the contribution of population growth to welfare through its effects on knowledge creation versus its effect on accelerated climate change. We find in most cases a net positive externality of children on average welfare for other dynasties (median: 22 k€/child).

keywords: Climate change; R&D-based growth; Population level; Social Cost of Carbon
JEL codes: Q54; O44; Jll

*contact: r.gerlagh@uvt.nl; at the economics department of the Tilburg University. I'm grateful to Peter Kjaer Kruse-Andersen for spotting a few ambiguities, to Malik Curuk and Sjak Smulders for pointing out recent advances in empirical estimates of endogenous growth models, to Lucas Bretschger for nudging me to add some relevant discussions of assumption. Remaining errors are mine. Updated version here: https://surfdrive.surf.nl/files/index.php/s/EMDwsvMBv142hmZ.

1 Introduction

The DICE model (Nordhaus, 1993a,b) has become a benchmark in the literature that assesses the efficiency of climate policy through the lens of optimal growth theory. It embeds a reduced-form climate change description in the classic Ramsey-Cass-Koopmans framework, putting investments in emission reductions on the same footing as investment in capital. Both investments in capital and investments in climate protection trade off current against future consumption.[1] More recently, Golosov et al. (2014) reworked the climate-growth framework along the lines of a stochastic growth Brock-Mirman (1972) model. Careful choices of functional forms allow the authors to derive a closed-form solution for savings and the social cost of carbon.[2] The present paper further extends their model by adding endogenous growth via variety expansion (Jones, 1995), endogenous human capital formation, and endogenous family size (De La Croix and Gosseries, 2009), while keeping closed-form solutions for all decision variables. This is the first contribution of this paper.[3]

The extensions we provide help to connect the empirical literature, which identifies separate temporary effects of weather and climate change on output from more persistent effects on growth, with models for calculating the Social Costs of Carbon. Dell et al. (2012) and Burke et al. (2015) find significant statistical evidence that increased temperatures reduce economic growth, which accumulates to permanently reduced income, specifically in poor and warm countries.[4] Their macro-empirical findings are backed up by micro empirical studies. Masters and McMillan (2001) discuss empirical data suggesting that climatic conditions affect economic growth mechanisms. Graff Zivin et al. (2018) find effects

[1]Some may find the label 'optimal growth' pretentious for the DICE model, as it does not contain a description of innovation, and thus lacks the 'engine' of growth. I use the term 'optimal growth' as in the 1970s, where it refers to the inclusion of preferences over time, which marked an innovation of the Ramsey-Cass-Koopmans (RCK) model compared to the Solow model. In these models capital accumulation is an important contributor to growth and time preferences enabled the RCK model to define the concept of optimal capital accumulation. Nordhaus' innovation extends the reach of the RCK model to address optimal climate investments.

[2]A closed form solution for an important variables such as investments means it can be calculated as dependent on current parameters, state variables, and observable statistics in a finite number of standard operations. A typical example of a closed-form solution is the Brock Mirman model which gives savings rates that only depend on time preference and production parameters. Golosov et al. (2014) derive that the ratio of optimal carbon taxes to output only depends on time preference parameters, severity and persistence of climate change. Closed-form solutions contrast with open-form solutions that use infinite sums, integrals, and/or recurrent formulations with references to future variables.

[3]Note that our results are robust to stochastic productivity and damage shocks in e.g. Ω_t and Γ_t, as one can immediately see from the similarities with Brock and Mirman (1972) and Golosov et al. (2014). In this paper we do not focus on stochastic dynamics.

[4]Colacito et al. (2019) report similar evidence for temperate zone developed economies.

of temperatures on cognitive test results, and Park et al. (2020) show that these effects accumulate and become significant and persistent over time; heat reduces learning outcomes. Donadelli et al. (2021) associate temperature rise to reduced research investments. Such findings have far-reaching implications for climate policy; if global warming hinders the economic development of today's poorest countries, the social costs of carbon emissions are potentially very large (Dietz and Stern, 2015; Bretschger, 2017). Our closed-form analytical results enable a translation of these empirical findings on climate and growth in terms of the social costs of carbon. We quantify these, the second contribution of our paper.

The extended model also supports a structured discussion of endogenous population in a climate-macro-growth context, both from a congestion and innovation perspective, the third contribution. Population featured prominently in "the tragedy of the commons" where Hardin (1968) writes: "To couple the concept of freedom to breed with the belief that everyone born has an equal right to the commons is to lock the world into a tragic course of action." More recently, Harford (1998) and Schou (2002) show that when a public bad (pollution) is present in a model with endogenous fertility choice, two instruments are needed for Pareto efficiency: a Pigouvian tax on pollution and a tax per child. This tax equals the present discounted value of pollution taxes each descendant will pay (Gerlagh et al., 2018). Estimates of the carbon legacy associated with current reproduction decisions due to the additional emissions of children, grandchildren, and so on, suggests such a corrective birth tax must be very large. They exceed by magnitude the parent's direct emissions generated by day-to-day activities (O'Neill and Wexler, 2000; Murtaugh and Schlax, 2009; Bohn and Stuart, 2015; Wynes and Nicholas, 2017). To capture the mechanisms essential to this literature, we include endogenous family size decisions, in the spirit of Barro and Becker (1989); Peretto and Valente (2015); Brunnschweiler et al. (2020), in an optimal growth climate model, as in De La Croix and Gosseries (2009, 2012) and Gerlagh et al. (2018). We label this approach to the birth externality the congestion perspective.

Yet the population-congestion argument is not as clear-cut. Citizen not only do damage to their descendants through their emissions, they also contribute to the solution by increasing the speed of innovation. As Kuznets (1960) wrote: "We now face the question whether an increase in the absolute number of these contributors to new knowledge is likely to produce increasing, constant, or diminishing returns per head. [...] the argument stresses the importance of human beings not as producers of commodities and services, but as producers of new knowledge". The optimist perspective of humans as the ultimate resource of wealth has famously been advocated by Simon (1981). We label this the innovation perspective.

One needs a model with climate damages, endogenous growth through innovations, and endogenous fertility, to address jointly the 'congestion' and 'innovation' view. The nexus between growth, population, and climate was picked up quickly after endogenous growth models were developed. In an early paper, Gradus and Smulders (1993) find that pollution reduction increases growth in a model where pollution affects the health of workers and their ability to learn. While they refer to the effects of lead and air pollution on cognitive capacities, in the context of the above-mentioned empirical evidence, we conclude that similar mechanisms can be at play when applied to higher temperatures. Fankhauser and Tol (2005) present results from a model that resembles our description of innovation. They develop an integrated assessment model with physical and human capital in the spirit of Mankiw-Romer-Weil (1992). Peretto and Valente (2019) also add endogenous innovation to an integrated assessment model. Yet both analyses do not capture the recent empirical findings listed above; they do not separate climate change effects on output from those affecting the growth of TFP, but instead treat all damages as affecting output levels only. Fankhauser and Tol (2005) assume the same durability (depreciation) for human capital and physical capital. They also build a climate-plus-endogenous growth Romer (1990) type model, but again do not address the longer persistence of damages to knowledge. Our approach extends theirs, as we add and calibrate the standing-on-shoulders feature in our model so that it reproduces empirically plausible convergence rates; that is, our model features statistically observed persistent growth effects of climate damages. Golosov et al. (2014) also report that their model can be extended with endogenous growth. Yet they consider knowledge as an unintential by-effect of production; in this paper we treat innovation as the outcome of purposeful research, staying close to common assumptions in the endogenous growth literature. Bretschger et al. (2017) have endogenous growth with multiple sectors and regions, which we extend with endogenous population. Bretschger and Pattakou (2019) build an AK endogenous growth model with climate change. Our different description of endogenous growth is essential to separate level-damages from growth-damages. Kruse-Andersen (2019) builds a model with directed technical change, focusing on clean versus dirty innovations, while here we consider the interaction between climate change, population, and the overall rate of economic growth. Bretschger (2020) has a model with endogenous growth and population, and climate change combined with a Hotelling fossil fuel market. We add a description of climate damages that more closely follows patterns typical for natural science climate studies. Finally, we mention recent work by Peretto and Valente (2015) and Brunnschweiler et al. (2020) who connect economic growth and endogenous fertility, in the tradition of Galor (2011). Our model is conceptually very similar to those two papers, apart from the assumed interaction between fertility and innovations. In both models, the wage share in income increases with population size, increasing the opportunity costs of fertility, constructing a negative feedback loop. In our model set up, we do not construct a market-based negative feedback between population size and fertility. We will assume that some effects of population on utility do not operate through markets, which we label 'direct social and congestion effects'.[5]

An important caveat of our study is the omission of specific technology of emission reduction. We focus on aggregate growth and abstract from directed technological change that is essential for the transition from a fossil-fuel based economy to one driving on renewable energy. An early study of directed technological change in a climate-macro model is Gerlagh (2008). A recent state-of-the-art analysis is provided by Bretschger et al. (2017). Kruse-Andersen (2019) also develops a model with climate change and directed technical change.[6] For empirical evidence on directed green innovation, see Aghion et al. (2016); Calel and Dechezlepretre (2016). A second limitation is the absence of an international dimension. Climate change will affect warm and cold, and rich and poor countries, differently. We remain silent on the implications thereof; see Bretschger and Valente (2011); Bretschger and Suphaphiphat (2014) for a discussion.

2 Model set up

2.1 Households

Consider a Ramsey economy with dynasties $i \in [0, 1]$, each of population size $n_{i,t}$, with endogenous population growth. Each dynasty maximizes welfare (Brunnschweiler et al.,

[5]The indirect population-congestion effect works through resource markets, or indirectly and with a delay, through the state of the environment.

[6]Kruse-Andersen (2019) has special interest for this literature as it asks a different question from what is typical. He investigates whether increased innovation brought by increased population is sufficient to solve environmental problems *without* complementary environmental policy. The answer is negative. We will assume, throughout, that efficient climate policies are implemented, and ask whether, under those conditions, an increased population leads to higher or lower welfare.

2020):[7]

$$v_{i,t} = \sum_{j=0}^{\infty} \beta^j \left[\ln(c_{i,t+j}/n_{i,t+j}) + \gamma \ln(f_{i,t+j}) + u(N_{t+j}) \right]. \tag{1}$$

In addition to per capita consumption $c_{i,t}/n_{i,t}$ and fertility $f_{i,t}$, welfare also depends on population size (which is exogenous to the individual household) through concave $u(N_t)$. The extension captures two phenomena. First, humans are by nature a social species. When world population becomes small enough, an additional person could make a significant addition to the choice set, e.g. to find a suitable partner. The function $u(.)$ describes such social needs through $u'(0) = \infty$. At the other side of the spectrum, the earth cannot provide for the required space and material needs of a world with above 1000 billion humans, even with the most efficient recycling technology. Our description of production insufficiently captures this concern.[8] To ensure we do justice to the finite earth argument, we include these direct congestion effects through $u'(\infty) < 0$. For practical purposes[9] we set $u_t = \gamma_N \ln(N_t) - \gamma_m N_t$; $\gamma_N, \gamma_m \geq 0$.[10] The direct socializing and congestion effects balance if $u'(N_t) = 0 \Leftrightarrow N_t = \gamma_N/\gamma_m$. In our analysis, we will typically consider two cases separately, *without* direct population externalities $u(.) = 0$ and *with* direct population externalities $\gamma_N, \gamma_m > 0$. The main analysis considers full dynamic paths; we use one asterisk for variables that have steady state properties in the full dynamic model without direct population externalties, and we use two asterisks for the steady state of the economy with direct population externalities.[11]

Households maximize welfare subject to the budget constraint, labour supply constraint,

[7] It is notoriously difficult to define optimality with endogenous populations (Asheim and Zuber, 2014). We choose Millian welfare based on average utilities. An alternative is Benthamite welfare, which multiplies average utility by population size before aggregation. In the context of endogenous population size, Millian welfare tends to favor a very large population of almost starving individuals. To prevent such 'repugnant' outcomes, it has been proposed to subtract a critical utility level before aggregation (Blackorby et al., 1995). In our set up, welfare measures average utility of the first generation.

[8] In Section 2.2, we describe economic production as a process of creating value. Even though production has decreasing returns to scale due to fixed factors, the production specification does not detail the material side, and the model supports perpetual output growth through innovation as a feasible outcome.

[9] That is, the specification results in linear loci in the phase diagram Fig 1.

[10] To sketch intuition for the linear term describing congestion, let us consider $M = 1$ the maximal riches of nature available for rival consumption. Per capita rival consumption of nature becomes $m_t = M/N_t$. If we take $u_t = \gamma_N \ln(N_t) - \gamma_m/m_t$, we get the above parametric form for $u(.)$.

[11] There is a subtle difference between the paths for the two economies. In the model without direct population externalities, $u(.) = 0$, capital K_t is non-constant but the savings rate s_K^* and fertility f^* will be constant along the transitionary paths. In the model with direct population externalities, $\gamma_N, \gamma_m > 0$, fertility is non-constant but converges to a steady state: $f_t \to f^{**}$.

population dynamics and human capital dynamics:

$$c_{i,t} + s_{i,t+1} = w_t h_{i,t} l_{i,t} + r_t s_{i,t} + \tau_{n,t} n_{i,t}, \tag{2}$$

$$l_{i,t} = (1 - \phi f_{i,t} - x_{i,t} f_{i,t}) n_{i,t}, \tag{3}$$

$$n_{i,t+1} = (1 + f_{i,t} - \delta_N) n_{i,t}, \tag{4}$$

$$h_{i,t+1} = x_{i,t}^{\eta_s} h_t^{\eta_h}, \tag{5}$$

where $c_{i,t}$ is consumption, $s_{i,t}$ are savings, $l_{i,t}$ is labour supply, $f_{i,t}$ is the fertility rate, δ_N mortality, $h_{i,t}$ is human capital, w_t are wages, r_t are returns to investments, $\tau_{n,t}$ are per capita lump-sum government transfers, ϕ is the time for raising children, and $x_{i,t}$ is the time spent on schooling. Elasticity of human capital with respect to schooling and previous-period human capital satisfies $0 < \eta_s, \eta_h < 1$. Aggregate population and labour supply are $N_t = \int_i n_{i,t} = n_t$, $L_t = \int_i l_{i,t} = l_t$. As all households are identical, we refer to the representative household without subscript i and recycle the index i for intermediate producers, below.

2.2 Final Goods Production

The final good is produced by use of intermediates $i \in [0, A_t]$

$$Y_t = \Omega_t \left(\int_{i=0}^{A_t} (y_{i,t}^{\frac{\varepsilon-1}{\varepsilon}}) \right)^{\frac{\varepsilon}{\varepsilon-1}} \tag{6}$$

where Ω_t is climate-related productivity, $\varepsilon > 1$ is the elasticity of substitution between varieties, and intermediates are produced by monopolists who maximize profits

$$\max_{k_{i,t}, l_{i,t}, e_{i,t}} [p_{i,t} y_{i,t} - r_t k_{i,t} - w_t h_t l_{i,t} - q_{z,t} z_{i,t} - (q_{e,t} + \tau_{e,t}) e_{i,t} - \pi_{i,t}] \tag{7}$$

$$\text{s.t. } y_{i,t} = k_{i,t}^{\alpha} \left(f_t(z_{i,t}, e_{i,t}) \right)^{\kappa} (h_t l_{i,t})^{1-\alpha-\kappa} \tag{8}$$

with $\pi_{i,t}$ the royalties paid to the patent owner for the blueprint of variety i, $e_{i,t}$ is the use of fossil fuels as an exhaustible resource associated with greenhouse gas emissions with $q_{e,t}$ the Hotelling rent and $\tau_{e,t}$ a carbon tax, $z_{i,t}$ is the use of other natural resources in fixed supply, $\int_i z_{i,t} + \int_j z_{j,t} = 1$ (with j for innovators) with price $q_{z,t}$, and $f(.)$ is an emissions-resource composite (cf Gerlagh and Liski, 2018). We assume $f(.)$ is strictly concave, has constant-returns-to-scale; fossil fuels are not essential, $f_t(1, 0) > 0$, $\partial f_t(1, 0)/\partial e_t < \infty$, demand for fossil fuels is bounded, $\forall t : \exists \bar{e}_t : \partial f_t(1, \bar{e}_t)/\partial e_t = 0$. The fossil fuel supply is described in Section 2.4.

We have constant value share of capital α and labour share $1-\alpha-\kappa$. The share κ includes rent value for fossil fuels, minerals, land, but also the costs of emission allowances. By assumption, decarbonization of the economy does not change the natural resource value share κ, for example because renewable energy also depends on fixed factors in limited supply (e.g. good sites for hydro, wind, and solar).

Through free entry and exit on the intermediates market, royalties capture all monopoly rents, and FOCs give

$$p_{i,t}y_{i,t} = \frac{\varepsilon}{\varepsilon-1}\left(r_t k_{i,t} + w_t h_t l_{i,t} + q_{z,t}z_{i,t} + (q_{e,t}+\tau_{e,t})e_{i,t}\right) \tag{9}$$

$$\pi_{i,t} = \frac{1}{\varepsilon-1}\left(r_t k_{i,t} + w_t h_t l_{i,t} + q_{z,t}z_{i,t} + (q_{e,t}+\tau_{e,t})e_{i,t}\right) \tag{10}$$

Due to full symmetry, we can aggregate over intermediates and when we normalize prices for the final good to 1, we have:

$$Y_t = \Omega_t K_{Y,t}^{\alpha}\left(f_t(Z_{Y,t},E_{Y,t})\right)^{\kappa}\left(h_t L_{Y,t}\right)^{1-\alpha-\kappa} A_t^{\frac{1}{\varepsilon-1}} \tag{11}$$

$$r_t K_{Y,t} = \frac{\varepsilon-1}{\varepsilon}\alpha Y_t \tag{12}$$

$$q_{z,t}Z_{Y,t} + (q_{e,t}+\tau_{e,t})E_{Y,t} = \frac{\varepsilon-1}{\varepsilon}\kappa Y_t \tag{13}$$

$$w_t h_t L_{Y,t} = \frac{\varepsilon-1}{\varepsilon}(1-\alpha-\kappa)Y_t \tag{14}$$

$$\pi_t A_t = \frac{1}{\varepsilon}Y_t \tag{15}$$

where $K_{Y,t} = \int_i k_{i,t}$, $L_{Y,t} = \int_i l_{i,t}$, $Z_{Y,t} = \int_i z_{i,t}$, and $E_{Y,t} = \int_i e_{i,t}$.

2.3 Innovation

Our description of innovation follows a standard semi-endogenous growth model, previously applied in a macro-climate context by Bretschger et al. (2017); Bretschger (2020). Ideas i are produced by innovators indexed j. The sector has open entry and exit. Each innovator produces a mass $a_{j,t+1}$ of new ideas, and the current stock of knowledge is

$$A_t = \int_j a_{j,t}. \tag{16}$$

We assume creative destruction; ideas are complementary to capital in the production process, which fully depreciates after each period. The innovation production function is

Cobb-Douglas

$$a_{j,t+1} = \Gamma_t k_{j,t}^{\alpha} \left(f_t(z_{j,t}, e_{j,t})\right)^{\kappa} (h_t l_{j,t})^{1-\alpha-\kappa} (X_t^A)^{-\psi} A_t^{\varphi}. \tag{17}$$

where Γ_t is a climate-dependent productivity factor, $X_t^A = \int_j k_{j,t}^{\alpha} f_{j,t}^{\kappa} (h_t l_{j,t})^{1-\alpha-\kappa}$ is a measure of the aggregate effort and $(X_t^A)^{-\psi}$ are decreasing returns due to standing on toes when inventors work on the same idea but only one receives the patent, while A_t^{φ} captures standing on shoulders with inventors building on existing knowledge.[12] For $\varphi < 1$, new ideas are getting harder to find when they accumulate, leading to semi-endogenous growth (Kruse-Andersen, 2017; Bloom et al., 2020).

Thus, the innovator maximizes

$$\max_{k_{j,t}, l_{j,t}} \left[\pi_{j,t+1} a_{j,t+1} / r_{t+1} - r_t k_{j,t} - q_{z,t} z_{j,t} - w_t h_t l_{j,t} - (q_{e,t} + \tau_{e,t}) e_{j,t}\right]. \tag{18}$$

As innovation production has CRS for firms j, we can aggregate over innovators:

$$A_{j,t+1} = \Gamma_t \left(K_{A,t}^{\alpha} \left(f(Z_{A,t}, E_{A,t})\right)^{\kappa} (h_t L_{A,t})^{1-\alpha-\kappa}\right)^{1-\psi} A_t^{\varphi}, \tag{19}$$

where $K_{A,t} = \int_j k_{j,t}$, $L_{A,t} = \int_j l_{j,t}$, $Z_{A,t} = \int_j z_{j,t}$ and $E_{A,t} = \int_j e_{j,t}$, and we have the zero-profit condition

$$\pi_{t+1} A_{t+1} / r_{t+1} = r_t K_{A,t} + w_t h_t L_{A,t} + q_{z,t} Z_{A,t} + (q_{e,t} + \tau_{e,t}) E_{A,t}. \tag{20}$$

2.4 Fossil Fuels and Climate Change

Carbon dioxide is the main greenhouse gas. Various climate-economy models thus connect emissions to fossil fuel supply, which they describe as an exhaustible resource in fixed cumulative supply (Sinn, 2008; Bretschger, 2020). These studies cast climate policy in a frame of optimal timing of extraction. Another perspective considers coal reserves, a major fossil fuel, as so large that it is an abundant resource that carries (almost) no scarcity rent (Golosov et al., 2014).

Our model treats the fossil fuel market as a simple exhaustible resource depletion problem. We abstract from extraction costs, so that prices follow the Hotelling rule. One interpretation is that the resource owner collects royalties $q_t E_t$ while extraction costs and costs of production of fossil fuels is implicit in $f(.)$. Careful reading of the proofs in the

[12] To keep our closed-form solutions, we impose the same capital-innovation elasticity as for capital-output α.

appnendix, and of Golosov et al. (2014) shows that results are robust to a more extensive description with both exhaustible (oil) and plentiful (coal) fossil fuels. The model also covers the case when a global constraint on cumulative emissions keeps some fossil fuels unexploited, rendering the Hotelling rent for those fossil fuels (e.g. coal) zero as the tightening of climate policies marks the end of the fossil fuel era (Gerlagh, 2011; Welsby et al., 2021).

We will find that the implementation of the optimum requires no taxes on fossil fuel reserves or extraction, thus fossil fuel prices increase with the effective returns on savings for households, which equal the marginal productivity of capital r_t multiplied by a factor $\sigma_{k,t}$ that corrects for a possible capital income tax: $q_{t+1} \leq \sigma_{t+1} r_{t+1} q_t$ with equality when $R_{t+1} > 0$. The resource owner maximizes the resource value:

$$\max\ q_t E_t + \frac{q_{t+1}}{\sigma_{k,t+1} r_{t+1}} R_{t+1} \tag{21}$$

$$\text{s.t. } R_{t+1} = R_t - E_t \tag{22}$$

with $R_t, E_t \geq 0$.

Carbon dioxide emissions increase atmospheric concentrations, but there is an exchange between the atmosphere, plants, soil, and water bodies including the oceans. The exchange between reservoirs leads to a dynamic system of differential equations that can be proxied through a set of 'boxes', each receiving a share of emissions, with its own depreciation rate. In turn atmospheric concentrations increase forcing, which slowly heats up the atmosphere and oceans. The coupled system can be captured through an emissions-response sequence, θ_i, which describes the increase in global average temperatures i periods after releasing carbon dioxide into the atmosphere (Gerlagh and Liski, 2018).

$$T_t = \sum_{i=1}^{\infty} \theta_i E_{t-i}. \tag{23}$$

Temperature rise reduces output, a level-effect, but also hampers growth (Masters and McMillan, 2001; Dell et al., 2012; Burke et al., 2015), which we capture through its effect on learning by innovation (Graff Zivin et al., 2018; Park et al., 2020; Donadelli et al., 2021). We borrow the functional form from Golosov et al. (2014):[13]

$$\Omega(T_t) = e^{-\delta_Y T_t}, \tag{24}$$

$$\Gamma(T_t) = e^{-\delta_A(\varepsilon-1)T_t}. \tag{25}$$

where the term $\varepsilon - 1$ is used for normalization; for equal $\delta_A = \delta_Y$, both functions will describe the same reduction of output in the next period, given a one-time temperature rise.[14] We have left out damages to capital, see van der Ploeg and de Zeeuw (2018) for an analysis.[15]

The interpretation of the damage coefficients δ_Y, δ_A warrant some annotation, especially the damages for growth. In his paper, Tol (2009) collects the most recent impact estimates of climate change on GDP. Depending on temperature increase, he finds a loss in GDP between 1%-5%. Hsiang et al. (2017) estimate costs of about 1.2 per cent of US GDP per Kelvin temperature rise. In their review of climate damage estimates, Howard and Sterner (2017) indicate higher percentage GDP losses from increases in temperature.[16] Based on that evidence we assume that each $TtCO_2$ reduces world output by about 1%, as a level effect, with an interval $[0.005, 0.015]$ for sensitivity analysis.

When temperature rise affects economic growth, the calibration of δ_A requires a more subtle reasoning. For example, Dell et al. (2012) estimate that poor countries see a 1.5 per cent point decline in the growth rate per Kelvin temperature rise. When one period in the model represents ten years, a 1K temperature rise thus lowers output in the next period by about 15 per cent, that is, $\delta_A = 0.15$. But growth effects are unevenly distributed, geographically. Burke et al. (2015) find no effect for countries with annual average temperatures of about 13 degrees Celsius, and a decline of about 1 per cent in the growth rate per Kelvin temperature rise for countries with annual average temperature of about 23 degrees Celsius. Observing that poor countries have warmer climates and rich countries mostly sit in temperature climate zones, the two references for calibrating δ_A fit consistently.[17] While high-income countries still make up more than half of the

[13]Going through the proofs and comparing with Golosov et al. (2014), it is clear from the similarities in model that we can add stochastic dynamics for both parameters δ_Y and δ_A, and also add stochastics for productivity levels. All results will be maintained based on expected parameter values.

[14]For comparison with other models that have endogenous population and technology: Schou (2002) assumes a flow pollutant, while Bretschger (2020) assumes that damages per ton CO_2 decrease inversely proportional with cumulative emissions.

[15]In our model, in which capital depreciates fully within each period, capital and output damages are equivalent. Capital damages are more important to account for in a model with incomplete capital depreciation.

[16]A recent study, Miller et al. (2021), also finds relatively high agricultural losses from temperature rise.

[17]Colacito et al. (2019) finds significant effects of quarterly temperature anomalies for the US, but

world economy, the formerly developing countries witness a rapidly rising population and economic share of in the world economy, including a rising share in frontier technologies. For our quantitative assessment, we consider a conservative 0.3 per cent point decline in the growth rate per Kelvin temperature rise as world average, that is accumulated over a period of ten years, our central parameter choice is $\delta_A = 0.03$ with sensitivity range $[0.01, 0.05]$.

2.5 Taxes and savings

Household aggregate savings are invested in capital, patents and fuel resources:

$$\int_i s_{i,t+1} = S_{t+1} = S_{K,t+1} + S_{A,t+1} + S_{R,t+1}. \tag{26}$$

For capital investments, households or investors can directly buy final goods, $S_{K,t+1} = K_{t+1}$, and receive returns on their owned capital r_{t+1}, with a tax or subsidy $\sigma_{k,t+1}$, so that their income from capital investments equal $\sigma_{k,t+1} r_{t+1} S_{K,t+1}$.

For patents, investors pay the inventors for their costs of research at time t,

$$S_{A,t+1} = r_t K_{A,t} + w_t h_t L_{A,t} + q_{z,t} Z_{A,t} + (q_{e,t} + \tau_{e,t}) E_{A,t}. \tag{27}$$

Inventors collect patent revenues $\pi_{t+1} A_{t+1}$ at time $t+1$, pay taxes or receive subsidies $\sigma_{a,t+1}$, and return net revenues to the investors. Household income from investments in patents thus amounts to $\sigma_{a,t+1} \pi_{t+1} A_{t+1}$. Arbitrage between the two investment opportunities imply that net returns on patents investments must equal net returns on capital investments $\sigma_{k,t+1} r_{t+1}$; (20) is replaced by:

$$\frac{\sigma_{a,t+1}}{\sigma_{k,t+1} r_{t+1}} \pi_{t+1} A_{t+1} = r_t K_{A,t} + w_t h_t L_{A,t} + q_{z,t} Z_{A,t} + (q_{e,t} + \tau_{e,t}) E_{A,t}. \tag{28}$$

The value of exhaustible fuel resources equal the price times the stock. Given arbitrage in savings and a return of $\sigma_{k,t+1} r_{t+1}$ on capital investments, the future value for exhaustible resources at which households are willing to buy at time t and hold until $t + 1$ the exhaustible resource results in

$$S_{R,t+1} = q_{t+1} R_{t+1} / \sigma_{k,t+1} r_{t+1}. \tag{29}$$

not when aggregated to annual variations. Yet Kim et al. (2021) present US evidence for significant macro-economic costs of climate-related extreme events for more recent years.

In addition to capital income taxes, the government levies labour income taxes $\sigma_{l,t}$. For population policies, we also introduce fiscal policies based on family size. These are interpreted broadly as family planning policies, but for convenience of notation are formally described similarly to a birth tax $\tau_{f,t}$. Finally, the government returns all tax revenues equally over all citizen, that is, proportional with household size through lump-sum transfers $\tau_{n,t}$. The household budget constraint (2) becomes:

$$c_{i,t} + s_{i,t+1} = \sigma_{l,t} w_t h_{i,t} l_{i,t} + \sigma_{k,t} r_t s_{i,t} - \tau_{f,t} f_t n_t + \tau_{n,t} n_{i,t} \tag{30}$$

To close, the government receives revenues from emission allowances, $\tau_{e,t} E_t$ and we assume that she also receives revenues from other scarce natural resources $q_{z,t} Z_t$. We will find in Proposition 2 that she needs no taxes for schooling and fuel resource depletion. Assuming closed per-period government budgets, lump-sum transfers equal net tax revenues[18]

$$\tau_{n,t} N_t = q_{z,t} Z_t + \tau_{e,t} E_t + \tau_{f,t} f_t N_t + (1 - \sigma_{l,t}) w_t h_t L_t + (1 - \sigma_{k,t}) r_t K_t + (1 - \sigma_{a,t+1}) \pi_t A_t \tag{31}$$

2.6 Aggregation and Equilibrium

Let $s_{K,t}$ be the savings share of output Y_t invested in new capital K_{t+1}, and $s_{A,t}$ the share of capital and labour used for innovation.[19] Define the capital-labour aggregate $X_t(.) \equiv K_t^{\alpha}(f_t(1, E_t))^{\kappa}(h_t L_t)^{1-\alpha-\kappa}$. The aggregate economy is described through the representative household welfare (1), labour supply (3), population growth (4), human capital accumulation (5), the commodity balance using (11), innovation (19), resource

[18]The assumption is not without consequences, as Ricardian equivalence does not hold with endogenous fertility decisions (Lapan and Enders, 1990). We discuss this further in Section 3.2 on taxes that decentralize the social optimum.

[19]Note that the combined investment rate in capital and knowledge per value of all production factors equals $s_{A,t} + s_{K,t}(1 - s_{A,t})$.

depletion (22) and the temperature response (23). In aggregate these equations become:

$$V_t = \sum_{i=0}^{\infty} \beta^i \left[\ln(C_{t+i}/N_{t+i}) + \gamma ln(f_{t+i}) + u(N_t)\right] \tag{32}$$

$$C_t + K_{t+1} = \Omega_t(.) A_t^{\frac{1}{\varepsilon-1}} (1 - s_{A,t}) X_t(.) \tag{33}$$

$$A_{t+1} = \Gamma_t(.)(s_{A,t} X_t(.))^{1-\psi} A_t^{\varphi} \tag{34}$$

$$N_{t+1} = (1 + f_t - \delta_N) N_t, \tag{35}$$

$$h_{t+1} = x_t^{\eta_s} h_t^{\eta_h} \tag{36}$$

$$R_{t+1} = R_t - E_t \tag{37}$$

$$L_t = (1 - \phi f_t - x_t f_t) N_t \tag{38}$$

$$T_t = \sum_i \theta_i E_{t-i} \tag{39}$$

where we use brackets in $\Omega_t(.)$, $\Gamma_t(.)$, $X_t(.)$ for shorthand, and to emphasize the dependence on other variables, e.g. that output depends on temperatures through $\Omega_t(.)$,which in turn depends on past emissions.

Given the description of household behavior, final good producers, intermediate good producers, innovators, the dynamics of climate change, and fiscal policies, we can define equilibrium.[20]

Definition 1 (Equilibrium)**.** *An equilibrium is an allocation $N_t, f_t, x_t, h_t, L_t, K_t, Y_t, C_t, A_t, E_t, R_t$, supported by prices $w_t, r_t, \pi_t, q_{z,t}, q_{e,t}$, and policies $\sigma_{l,t}, \sigma_{k,t}, \sigma_{a,t}, \tau_{f,t}, \tau_{e,t}, \tau_{n,t}$, such that households maximize utility, and firms (final good, intermediate goods and innovators, fossil fuel owners) maximize profits.*

As an alternative to fiscal policy variables $\sigma_{l,t}, \sigma_{k,t}, \sigma_{a,t}, \tau_{f,t}, \tau_{e,t}, \tau_{n,t}$, we can identify an equilibrium, or the supporting policies, through intensive control variables. We consider the savings rates $s_{K,t}, s_{A,t}$, the intensity of carbon pricing relative to output $g_t \equiv (\partial Y_t/\partial E_t - q_t)(1/Y_t)$, and the fertility and education choices f_t, x_t. The control-variables are particularly useful to characterize an equilibrium as there are many instances in which they remain constant over an equilibrium, both on the balanced growth path as well as on the transitionary paths.

Definition 2 (history-independent policies 1)**.** *A policy (or the allocation produced by the policy), is said to be in the class $\mathcal{P}(s_K)$, $\mathcal{P}(s_A)$, $\mathcal{P}(g)$, $\mathcal{P}(f)$, $\mathcal{P}(x)$, when the corresponding*

[20] Note that $q_{z,t}$ is considered a price rather than a policy variable, as its level is determined by supply and demand, $Z_t = 1$.

policy choice variable $s_{K,t}, s_{A,t}, g_t, f_t, x_t$ *is a sequence (over time) independent of the (current) state of world* (K_{t_0}, A_t, $(E_{t-i})_{i=1}^{\infty}$, N_t, h_t, R_t)*.*

We speak interchangeably of policies, allocations, or equilibria in the class $\mathcal{P}(.)$, whichever is convenient. We can meaningful intersect classes, e.g. $\mathcal{P}(s_K, s_A) = \mathcal{P}(s_K) \cap \mathcal{P}(s_A)$ is the set of policies with both capital savings rates and relative innovation efforts independent of history. We do not define the property $\mathcal{P}(E)$ as we do not have an intensive control variable for fossil fuel use, or emissions E_t, which is independent of the fossil fuel stock R_t. That is, the carbon tax $\tau_{e,t}$ may satisfy property $\mathcal{P}(g)$, but the market price q_t varies with R_t in a way for which we have no closed-form solutions.[21]

The above definition characterizes classes of dynamic equilibria including paths out-of-balanced growth. It defines classes of economies where *some* intensive variables remain constant along the transitionary dynamics. The basic Brock-Mirman and Solow models, as examples, are used as basis for empirical studies to describe transitionary dynamics such as (conditional) convergence of low-income countries that catch up with high-income countries. Yet these models have a constant savings rate, and thus fall in the class $\mathcal{P}(s_K)$. Our model extends the BM and Solow model property $\mathcal{P}(s_K)$ to include constant innovation expenditure shares through the property $\mathcal{P}(s_A)$. The model then also provides closed-form solutions for the convergence rate, which helps to discipline the model parameters (Appendix B, see also C.1 for some transition dynamics).[22]

The economy without direct population externalities $u(.) = 0$ supports long-run balanced growth for population with constant fertility and ever-rising (or declining) population size, and convergent human capital, $h_t \to h_\infty$. Appendix B.1 derives balanced growth properties. Long-run economic growth depends on two mechanisms associated with innovations: standing on shoulders (φ) and standing on toes (ψ). If the former is sufficiently strong, innovations can build on each other without diminishing returns. Varieties continue to accumulate and economic growth continues forever. Even if population as the ultimate source of ideas is constant, each cohort adds to the stock of ideas. When standing on shoulders is weaker, or a larger share of innovations are duplicates (standing on toes), innovations still build on each other but to a lesser extend, and the growth of the number of varieties diminishes in the long run if population stabilizes. In that case, long-run economic growth through an ever-increasing expansion of varieties remains feasible only if supported

[21] See comment in the proof of Lemma 1.

[22] When a model includes endogenous population but also property $\mathcal{P}(f)$, we can treat population as a variable independent of other state dimensions. That is, property $\mathcal{P}(f)$ supports the use of population growth as independent factor in an empirical growth model. Similarly including education as independent variable in an empirical growth estimate assumes the property $\mathcal{P}(x)$.

by an ever-increasing population; the model is of the semi-endogenous growth type.[23] For our quantitative analysis below, we follow Jones (2002) and calibrate a semi-endogenous growth model.

The economy with population socializing, and congestion, $\gamma_N, \gamma_m > 0$ naturally defines a relaxation of the class of policies:

Definition 3 (History-independent policies 2)**.** *We define the classes $\mathcal{P}'(f), \mathcal{P}'(x)$ if f_t, x_t do not depend on capital, technology and past emissions, but only depend on (N_t, h_t). Note that $\mathcal{P}(f) \subset \mathcal{P}'(f)$, $\mathcal{P}(x) \subset \mathcal{P}'(x)$.*

For $u(.) = 0$ we will define the class of allocations $\mathcal{P}^* \equiv \mathcal{P}(s_K, s_A, g, f, x)$, and for $\gamma_N, \gamma_m > 0$ use the same label slightly differently: $\mathcal{P}^* \equiv \mathcal{P}(s_K, s_A, g) \cap \mathcal{P}'(f, x)$.

3 Social Optimum

3.1 Optimal Control

In the original Brock-Mirman (1972) model, full capital depreciation and logarithmic utility can be chosen as parametric forms supporting optimal savings as a constant fraction of output, in a context of stochastic total factor productivity. Krusell et al. (2002) use the same set up to establish closed-form solutions for optimal capital taxes, with property $\mathcal{P}(s_K)$, in an economy with time-inconsistent consumers. Golosov et al. (2014) showed how climate change dynamics and damages can be included, adding a closed-form solution for carbon taxes that are proportional to output, $\partial Y_t/\partial E_t = q_t + gY_t$. That is, the ratio between carbon taxes and output, for which we use the variable g, is a constant that only depends on the description of climate change dynamics, damages, and time preferences; optimal climate policies are in the class $\mathcal{P}(g)$. Gerlagh and Liski (2018) and Iverson and Karp (2021) merge the previous two studies and find that optimal climate policy with time-inconsistent preferences remain in the class $\mathcal{P}(s_K, g)$. Here we further extend the framework by including endogenous variety expansion, human capital, and fertility, while keeping constant ratios to characterize the decision variables that implement the equilibrium.

[23] We can see this in the appendix, eq (135). There is a continuum of combinations for standing on shoulders φ and standing on toes ψ that will produce a zero at the denominator, which means the model supports full endogenous growth with constant population. For any lower value of φ, the denominator will be positive and the model supports semi-endogenous growth. To guarantee finite growth, we impose $(1-\alpha)(\varepsilon-1)(1-\varphi) > \alpha(1-\psi)$

The proposition below shows that our model integrates all domains apart from the exhaustible resource use within the tradition of Brock-Mirman.[24] For our model, the class of allocations $\mathcal{P}^*$ envelopes the social optimum, $SO \in \mathcal{P}^*$, along the transitionary equilibrium path.[25]

Proposition 1 (Social optimum characterization)**.** *For any initial state, the Social Optimum is characterized through a constant capital investments share $s_{K,t} = s_K^*$ and constant innovation efforts $s_{A,t} = s_A^*$:*

$$s_K^* = \alpha\beta \left[1 + \frac{\beta(1-\psi)}{(\varepsilon-1)(1-\beta\varphi)}\right] \tag{40}$$

$$s_A^* = \frac{\beta(1-\psi)}{(\varepsilon-1)(1-\beta\varphi)+\beta(1-\psi)} \tag{41}$$

while the social costs of carbon equals the marginal productivity of fossil fuels, and is proportional to output, $\partial Y_t/\partial E_t = g_t Y_t$. The factor $g_t = g^*$ is constant and describes the discounted cumulative sum of the climate emissions-response:[26]

$$g^* = \left[\delta_Y + \frac{\beta\delta_A}{1-\beta\varphi}\right]\sum_{i=1}^{\infty}\beta^i\theta_i. \tag{42}$$

Though the carbon tax is set by g^*, the end-user price for fossil fuels also depends on its market price q_t and emissions E_t^* vary with the available resources R_t.

If $u(.) = 0$, time spent per child on education, $x_t = x^*$, and fertility $f_t = f^*$ are constant, jointly determined through the equations

$$\frac{\phi f^* + x_t f^*}{1-\phi f^* - x^* f^*} = \frac{\gamma + \beta\zeta_N\widetilde{f}}{(1-\alpha-\kappa)\zeta_X}, \tag{43}$$

$$\frac{x^* f^*}{1-\phi f^* - x^* f^*} = \frac{\eta_s\beta}{1-\beta\eta_h}, \tag{44}$$

where $\widetilde{f} = fN_t/N_{t+1} = f^*/(1+f^*-\delta_N)$ is the (constant) share of new born in the population, ζ_X is the (constant) value of all output in a period relative to the value of consumption (89)*, and ζ_N is the (constant) value of the population size relative to the*

[24]Hassler et al. (2018) presents a closed form for resource extraction if one agent owns the exhaustible resource without any other income.

[25]The absence of time subscripts below only applies to these intensive variables. For all other variables, time subscripts are required to describe dynamics.

[26]Fig 1 of Dietz and Venmans (2019) suggests that θ_i is almost constant, so that the last term in the equation below becomes $\theta/(1-\beta)$.

value of consumption (91). *Fertility f^* increases (decreases) while schooling x^* decreases (increases) with preferences for children γ (elasticity of human capital with respect to schoolng η_s, η_h).*

If $\gamma_N, \gamma_m > 0$, population and fertility (N_t, f_t) converge monotonically to a steady state with constant population N^{**} and reproductive fertility $f^{**} = \delta_N$. Fertility and population in balanced growth (steady state) in the two models relate as:

$$f^* > \delta_N \Leftrightarrow N^{**} > \gamma_N/\gamma_m. \tag{45}$$

Upon closer inspection, the control variable values of the Social Optimum are intuitive. For interpreting the capital investment share (40), recall that in the Brock-Mirman model the savings share is $\alpha\beta$. In our model, capital investments have an extra return, as they also contribute to technology, in addition to their direct contribution to (next-period) output. The technology contribution is one-period delayed, which explains the β in the second term. The term $1/(\varepsilon - 1)$ measures the value of technology relative to output, while the $(1 - \psi)$ measures the returns to scale and $1/(1 - \beta\varphi)$ the persistence of contributions to innovation.

For interpreting the innovation investment share (41), it is essential to recognize that $s_A^*/(1 - s_A^*)$ measures the investments in innovation relative to those in final goods.

$$\frac{s_A^*}{1 - s_A^*} = \frac{\beta(1 - \psi)}{(\varepsilon - 1)(1 - \beta\varphi)} \tag{41'}$$

Now we see that this ratio exactly repeats the same arguments as for the second term in s_K^*; the above intuition can be copied.

The Social Costs of Carbon rule (42) extends the closed-form solution by Golosov et al. (2014) as it attributes a substantial part of its value to effects of climate change on innovation mechanisms (Masters and McMillan, 2001; Graff Zivin et al., 2018; Park et al., 2020). It extends previous studies with climate damages and innovation as these do not provide closed-form solutions for the SCC and thus need to rely on simulations for its calculation. Furthermore, we note that though various climate-economy models feature innovation, Fankhauser and Tol (2005) attributes all damages to output reduction, Bretschger and Pattakou (2019) to capital depreciation, and (Bretschger, 2020) does not have a carbon tax. The details of the SCC rule are readily understood. Damages to output enter the formula similar to Golosov et al. (2014). Damages to innovation affect output with one period delay, the β in the nominator, but these damages persist over next periods by rate φ, the $1-\beta\varphi$ in the denominator.[27]

The fertility condition (43) states that time spent on raising children relative to labour equals the ratio of the value of offspring relative to the value of output. The time spent on schooling relative to labour (44) calculates the relative increase in future labour revenues gained by spending one unit of time on education. In the context of other climate-fertility models, our results add a positive innovation effect of increased population compared to (Schou, 2002; Bohn and Stuart, 2015; Gerlagh et al., 2018) while it adds crowding out of scarce resources through population growth compared to Bretschger (2020). A further discussion of innovation and resource scarcity effects on the birth externality is provided when we derive fertility policies below (52). As a result, we will find both positive and negative birth-externalities as a possible outcome, dependent on parameter values (Table 2).

We can also consider the role of individual parameters in the social optimum characterization. When we interpret higher carbon taxes as an investment in the future, we see that all investments increase with the weight given to future generations β, and decrease with standing on toes ψ, substitutability between intermediates ε, and increase with durability of technology φ. These properties are intuitive and do not require much explanation.

The class of history-independent policies spans a large set of allocations. If we think of a 'business as usual' that is characterized by market failure for climate change, $g=0$, while other policies implement the social optimum, we will also find $BAU \in \mathcal{P}^*$. More generally, the class admits for many policy distortions that leave some wedge in the economy, for example between the social and private value of capital investments, or the social and private value of innovations. Such distortions do not necessarily arise from imperfect policy making; they may also result from incomplete information about the fundamental parameter values. If we are uncertain about the degree of crowding out between innovations ψ, or the spillovers between subsequent generations of patents φ, fiscal policies such as tax exemptions for research cannot be optimally set. The savings rate, research investments, fertility rate or schooling expenditures may be set too high or too low, but as long as the control variables defined in the above definition remain independent of the state variables, the equilibrium still falls in class $\mathcal{P}^*$. This class of policies has rather useful welfare properties, summarized in the following lemma.

[27] Intuitively, we can use the result to also gauge the effect of damages for capital. Consider the case that capital would not depreciate fully, but has persistence of φ_K, and that climate damages to the capital stock are evaluated as equivalent to δ_K share of output. The aggregated sum of damages would then result in a carbon-pricing term between the square brackets of $\delta_K/(1-\beta\varphi_K)$.

Lemma 1 (separable log-linear welfare)**.** *If $u(.) = 0$, within the class of equilibria $\mathcal{P}(s_K, s_A, g, f, x)$, welfare satisfies*

$$V_t = \zeta_K \ln(K_t) + \zeta_A \ln(A_t) + \zeta_N \ln(N_t) + \zeta_h \ln(h_t) - \sum_{i=1}^{\infty} \Theta_i E_{t-i} + \overline{V}_t(R_t). \tag{46}$$

If $\gamma_N, \gamma_m > 0$, within the class of equilibria $\mathcal{P}(s_K, s_A, g) \cap \mathcal{P}'(f, x)$, welfare satisfies

$$V_t = \zeta_K \ln(K_t) + \zeta_A \ln(A_t) - \sum_{i=1}^{\infty} \Theta_i E_{t-i} + \overline{V}_t(R_t, N_t, h_t). \tag{47}$$

The weights ζ_K, ζ_A and parameters Θ_i are the same in both cases. Weights $\zeta_K, \zeta_A, \zeta_N, \zeta_h$ and parameters describing the social costs of past emissions Θ_i are constant over time, and do not depend on the the level of (past, present and future) savings rates $s_{K,t}$, innovation shares $s_{A,t}$, or climate policies g_t. Such policy choices enter the function sequence $(\overline{V}_t(.))_t$.

As with Proposition 1, the lemma applies to the dynamic equilibrium also off steady state; the parameter ζ_K measures how current welfare depends on the current capital stock, given the future adjustment dynamics for investments, education, carbon taxes, etc. The parameters in the lemma can be interpreted as elasticities of permanent (real) income with respect to assets. Consider the case that consumption increases by 1 per cent, perpetually. Welfare will then increase by $1/(1-\beta)$ per cent. Thus, When capital increases by 1 per cent, welfare increases by as much as if permanent income would rise by $(1-\beta)\zeta_K$ per cent. The same argument holds for the other stock variables. That is, the parameters describe the elasticities of permanent income, scaled by $1/(1-\beta)$.[28]

The lemma has appeal beyond its use in this study; it is instrumental for solving Markov Perfect equilibria in which agents behave strategically, e.g. because of time-inconsistent preferences (Krusell et al., 2002; Iverson and Karp, 2017; Gerlagh and Liski, 2018).

Lemma 1 is a typical feature of the Brock-Mirman model set up, that secures validity of results for individual policy domains when we deviate from the model, or first-best policies, in other policy domains. It importantly extends the reach of the first proposition. It informs us that the policy choices for investments in capital, innovation, carbon taxes,

[28]The permanent income referred to in this paragraph differs from the long-run permanent effect of a change ìn some state variable. In the calibration, we shall see that the model is of the semi-endogenous growth type, which means that the effect of a change in capital, technology or human capital on income vanishes over time. A shock in population size leads to a permanent change in the level of per capita output and consumption. A shock in climate conditions also has a permanent income effect if climate change is permanent, that is, if θ_i is bounded away from zero. In Appendix C.2 we derive those permanent income effects of current emissions.

human capital, and population are separable. It specifically confirms that for second-best equilibria the efficient carbon pricing still follows the same rule (42).

Corollary 1 (Carbon taxes in second-best)**.** *Given any sequence of history-independent policies for capital, innovation, fertility, education, $s_{K,t}, s_{A,t}, f_t, x_t$ in the class $\mathcal{P}(s_K, s_A) \cap \mathcal{P}'(f, x)$, not necessarily first-best, climate policies maximize welfare if carbon taxes are set by the Social Costs of Carbon* (42)*.*

The corollary also speaks to the debate between positive and descriptive discounting. Suppose that time preferences β have been estimated empirically, e.g. based on household savings decisions. Furthermore, assume that macro-economic observations for capital and innovation investments are inconsistent with their calculated social optimal levels; there are often constraints in imperfect economies that create a wedge between the private and social returns to investment (Stern, 2008). In that case, (42) can still be used to guide optimal carbon pricing. There is no need to first improve the efficiency of capital or innovation markets for climate policy to be effective in improving welfare.

The corollary is also useful for those who consider climate change a fundamentally ethical problem, which requires a trade-off between future and present generations that cannot be left to markets. Within the domain of prescriptive discounting, the β is based on ethical considerations, independently of macro-economic descriptive capital returns statistics. The ethical proposition shared by many illustrious economists is that β should be close to unity (Stern, 2008). In this context, the above proposition and corollary ensure that, also if macro-economic investments are too low given the weight attributed to future generations, the SCC formula (42) can still be used for optimal climate policies.[29]

Lemma 1 is also informative for the second major theme of this study, the contribution of population size to average utility:

Corollary 2 (Population and welfare)**.** *If $u(.) = 0$, in all equilibria in the class $\mathcal{P}(s_K, s_A, g, f, x)$, including the social optimum, welfare W_1 increases with the initial population size N_1 iff $\zeta_N > 0$, i.e. iff*

$$(\alpha(1-\beta)+\kappa)(\varepsilon-1)(1-\beta\varphi) < \beta(1-\alpha(1-\beta)-\kappa)(1-\psi) \tag{48}$$

The corollary provides sharp parametric conditions for welfare to increase (decrease) with population size, and though the inequality looks complicated, its derivation in the appendix, specifically (91), informs us about its interpretation.

[29] Scovronik et al. (2016) provides an example of calculations for optimal carbon pricing with an exogenous capital savings rate.

The various parameters have the expected signs in the inequality, e.g. the role of technology. A positive population effect is more likely when varieties are more complementary (lower ε), and technology has more positive spillovers to future innovations φ. Negative population effects appear when efforts for innovations have decreasing returns to scale (ψ), and when technology is exogenous, $\psi = 1$, a negative population effect always prevails. A negative effect is more likely when capital and natural resources are an important factor, reducing the returns to scale in production. Without capital and resources, that is if labour is the only production factors, $\alpha = \kappa = 0$, we always have positive returns to population since production is proportional to population size, and innovation adds to that, leading to increasing returns. To summarize, the formula contains both the congestion and innovation perspectives described in the introduction. Decreasing returns to scale associated with capital and natural resources cause a negative effect of population size on per capita welfare, whereas endogenous growth yields a positive effect of population size on per capita welfare.

3.2 Decentralization

We now present the implementation of the social optimum through fiscal policies, that is, we determine optimal policies in equilibrium. Whereas in this section we assume optimal policies over all policy domains; these policies are independent between domains. Corollary 1 establishes the social cost of carbon as an efficient climate policy rule in second best. We can read the proposition below as four independent policy prescriptions.

Proposition 2 (Decentralization of SO)**.** *Given the social optimum $s_K^*, s_A^*, g^*, f^*, x^*$, there is a unique policy vector $\sigma_l^*, \sigma_k^*, \sigma_a^*, \tau_{f,t}^*, \tau_{e,t}^*$ that implements the SO as equilibrium. Optimal schooling and fuel resource depletion does not require fiscal intervention.*

$$\sigma_l^* = \sigma_k^* = \frac{\varepsilon}{\varepsilon - 1} \tag{49}$$

$$\sigma_a^* = \frac{1-\psi}{1-\beta\varphi} \tag{50}$$

$$\tau_{e,t}^* = \frac{\varepsilon}{\varepsilon - 1} g^* Y_t \tag{51}$$

$$\tau_{f,t}^* = -\beta(\zeta_K + (1-\psi)\zeta_A + \zeta_{R,t+1} + \zeta_{N,t+1})\frac{C_t}{N_{t+1}} \tag{52}$$

A fertility above reproduction in the model without direct population externalities implies *higher birth taxes in the model with direct externalities*

$$f^* > \delta_N \Leftrightarrow \zeta_N^{**} < \zeta_N^* \Leftrightarrow \tau_f^{**} > \tau_f^* \tag{53}$$

where birth taxes are evaluated for identical consumption and population levels.

In equilibrium the intermediate sector pays production factors below marginal productivity due to monopolistic supply of intermediates; the income tax σ_l^* and σ_k^* correct for that (49). Innovations have positive externalities due to standing on shoulders (φ); the innovator does not appropriate the full value of future use of its innovations. There are also negative externalities of innovation efforts due to standing on toes (ψ). The innovation subsidy (tax) σ_a^* weighs the relative importance of both externalities (50). Yet the model does not describe the need for government funding, while assuming lump-sum transfers as feasible redistribution mechanism. We thus do not consider the income and innovation taxes comparable with those observed in data and will not report these in our quantitative analysis.[30] Emissions cause future damages which the carbon tax g^* corrects for, yet the carbon tax is also reduced in response to monopoly pricing. While carbon taxes also need correction for the costs of public funds, in line with previous climate-macro studies we will report the values that result from the first order conditions (51).

Equation (52) describes the birth tax; it measures the externality of fertility decisions, that is, the change in welfare for all other dynasties $j \neq i$ caused by one additional child in dynasty i. It is almost simple, and at first sight somewhat surprising. The first observation is that the fertility tax or subsidy does not only depend on the pure population externality as captured by ζ_N, but it depends on the overall returns to scale for all aggregate stocks together, capital, technology, exhaustible resources, and population. To better understand the formula, we proceed in two steps. First, we abstract from standing on toes, assuming $\psi = 0$. Then we explain its role in the birth externality.

Abstracting from standing on toes, $\psi = 0$, the sign for the birth externality equals the sign of overall returns to scale.

$$\tau_{f,t}^* > 0 \Leftrightarrow \zeta_K + \zeta_A + \zeta_{R,t+1} + \zeta_{N,t+1} < 0$$

This is intuitive. Consider the case that the economy has constant returns to scale. If both capital and population increase by the same relative amount, average utility will

[30]For calibrated parameters presented in Table 1, we find for most cases no need for financial support for private R&D, $\sigma_A^* < 1$. This may come as a surprise but is consistent with recent empirical estimates (Montmartin and Massard, 2015).

not change. The constant returns to scale at the aggregate economy match the constant returns to scale holding at the dynasty's level. If a dynasty increases its savings and number of children proportionally, then average welfare for its offspring will not change. That is, individual trade-offs with respect to savings and family size reflect the aggregate trade-offs, and no birth tax is required.[31] If the aggregate economy has decreasing returns, then more people crowd out each other in pursuit of a fixed factor, signaling a negative birth externality.

Then, what is the role of standing on toes, ψ, in the birth externality? We best understand this by viewing the extreme case with $\psi = 1$. In that case, technology is exogenous and the optimal innovation effort becomes zero. Yet, from the outside, if an external force could increase A_t for one period, then output and welfare would increase. Thus, $\zeta_A > 0$, but as technology is not (effectively) produced within the economy, it must not count in the returns to scale measure used for the birth externality.

4 Quantitative Assessment

The closed-form solutions enable us to calibrate parameters and calculate the social costs of carbon without the need for dynamic simulations. Validity of resulting numbers is constrained by the assumptions required to derive solutions. Two remarks are in order. First, regarding the restrictions we impose on parametric forms. Van den Bijgaart et al. (2016) and Rezai and Van der Ploeg (2016) present extensive calculations comparing a simple closed-form carbon pricing rule derived from a model with strong assumptions with those that come from more elaborate dynamic models in which those rules cannot be formally derived. They find that the strong assumptions on parametric forms cause biases in quantitative results that are very small compared to the range originating from parameter uncertainty.[32] Our second remark considers the effect of sub-optimal policies in other policy domains on optimal climate policy choices. As a key interest is the social cost of carbon, we follow the relevant literature building on Nordhaus (1993a) assuming that optimal policies are in place for all other policy domains.[33]

We present results for a series of 1000 alternative parameter values. The table below lists the targeted macro-economic moments, and the lower 5 percentile, median, and upper

[31] For this line of reasoning, it is important that the measure of returns to scale at the aggregate level includes all stocks that are held as private property.

[32] As case in point, in Appendix C.3, we find a wide dispersion of SCC levels associated with level and growth damages. All these calculated results use the same parametric form. Thus, basic assumptions on targeted macro-moments such as the convergence rate have substantial effects on calculated SCC levels.

[33] Corollorary 1 suggests that results are not too dependent on this assumption.

95 percentile of resulting parameter values. The endogenous growth model allocates a value share $1/\varepsilon$ of output to technology. For the elasticity of substitution, the literature provides a wide range of estimates, with high elasticities for high-digit level of aggregation, and much lower values for low-digit estimates. Broda and Weinstein (2006) find a low value of 2.2 at 3-digit level post-1990, but also values above 10 at more disaggregate level. We choose a uniform distribution for all parameters and macro targets, $\varepsilon \in [3, 7]$. The technology value share is statistically counted for as gross profit in the labour-vs-profits distribution of value added, so that we add it to investments and calibrate the elasticity of output with respect to capital α to have total investments in the range of [20,40] per cent, which gives a median $\alpha = 0.26$ in the 5-95 percentile interval $\alpha \in [0.12, 0.39]$. We choose the total share of natural resources in value of output as $\kappa \in [0.05, 0.15]$, representing among other things minerals and land for agriculture and renewables. We calculate $\zeta_{R,t}$ based on the market capitalization. In the model the market value of exhaustible resources relative to capital and technology is given by $\zeta_{R,t}/(\zeta_K + (1-\psi)\zeta_A)$ (122). The energy companies jointly make up a few percentages of the world market capitalization; we set this share to cover the range $[0.02, 0.08]$.[34]

Based on Golosov et al. (2014), we consider a pure rate of time preference of 2 per cent per year ranging between 1 and 3 per cent, using a decade as time unit in our model we have $\beta \in [0.74, 0.90]$. We calibrate the standing on shoulders φ and standing on toes ψ jointly, requiring a convergence speed between 1 and 3 per cent per year, and per capita income growth between 20 and 60 per cent of population growth (Jones, 2002; Fernald and Jones, 2014). Largest part of US post-WWII growth is estimated to come from increasing R&D shares, that is, transitionary dynamics and increased education; only a small part comes from growing number of innovations associated to labor or population growth. See Appendix B for details. We find substantial standing on shoulders, but also a large standing on toes. Climate sensitivity is set at a median 0.7 Kelvin per Teraton CO2 with range $[0.4, 1.0]$, immediate and constant (Dietz and Venmans, 2019; Van der Ploeg et al., 2020). Climate damages on the output level is set between a half and 1.5 per cent per Kelvin, while the decrease in TFP is set to decrease growth between 0.1 and 0.5 per cent per Kelvin accumulating to 1-5 per cent per decade, see Section 2.4 for further discussion thereof.

[34]Stranded fossil fuel assets under optimal climate policy would be reason for a downwards adjustment of the market capitalization (Welsby et al., 2021).

Table 1: Parameters and Macro Targets

Parameter	**Description**	**Value**	**Source / Targeted Moment**
α	Capital-output elasticity	(0.12,0.26,0.39)	Savings share
β	Pure discount	(0.74,0.82,0.90)	Return on capital
δ_Y	Climate damage for output [/K]	(0.005,0.01,0.015)	Hsiang et al. 2017
δ_A	Climate damage for growth [/K]	(0.01,0.03,0.05)	Dell et al. 2012; Burke et al. 2015
ε	Elasticity of demand	(3,5,7)	Industry mark up
φ	Standing on shoulders	(0.71,0.79,0.88)	Convergence of 1-3% p.y.
κ	Natural resource share in output	(0.05,0.1,0.15)	Resource shares
ψ	Standing on toes	(0.51,0.80,0.93)	Income growth, $g_Y/g_L = 1.2 - 1.6$
θ_i	Climate sensitivity [K/TtCO2]	(0.4,0.7,1.0)	Climate literature

The triples for $\beta, \delta_Y, \delta_A, \varepsilon, \kappa, \theta$ present the lower bound, median, and upper bound for chosen uniform distributions, while the triples for α, φ, ψ present 5,50,95 percentiles that come out of the calibration process.

Putting these parameters into Proposition 1, we can derive optimal capital investment and innovation shares, and the social costs of carbon conditional on income, presented in the next table. We consider world output Y_t of 700 trillion euros per decade and a population of 7.9 billion. The social costs of carbon are split in one part associated with level damages, and the other part associated with reduced growth. The result on the SCC associated with level damages is consistent with the existing literature. Based on output-level damages, we find a social costs of carbon between 10 and 38 euro per ton CO2. This is in the same range of values presented in (Golosov et al., 2014). Strikingly, the climate-damages affecting growth are valued a magnitude larger, between 54 and 304 €/tCO2. Thus, we confirm the expectations by (Dietz and Stern, 2015; Bretschger, 2017) of very high social costs of carbon if climate change hinders the economic development, as empirically established by Dell et al. (2012) and Burke et al. (2015).[35][36]

[35]The empirical evidence to support the calibration of δ_A is limited. If innovation takes place mostly in temperate climate regions, where innovation does not suffer from global warming, it is corrected downwards.

[36]See Appendix C.3 for a figure with SCC results level versus growth effects.

Table 2: Outcomes for calibrated model

Variable	Description	Value
$s_A + s_K(1-s_A)$	Aggregate investment rate	$(0.25, 0.33, 0.40)$
s_K	Capital Investment share	$(0.12, 0.24, 0.34)$
s_A	Research share	$(0.06, 0.11, 0.18)$
$\tau_{E,t}$	SCC [€/tCO2]	$(71, 161, 332)$
	level effect [€/tCO2]	$(10, 21, 38)$
	growth effect [€/tCO2]	$(54, 142, 304)$
$(1-\beta)\zeta_K$	capital-permanent income elasticity	$(0.03, 0.07, 0.12)$
$(1-\beta)(1-\psi)\zeta_A$	technology-permanent income elasticity	$(0.02, 0.04, 0.06)$
$(1-\beta)\zeta_{R,t}$	resource-permanent income elasticity	$(0.002, 0.005, 0.010)$
$(1-\beta)\zeta_N$	population-permanent income elasticity	$(-0.18, -0.05, 0.09)$
$\tau_{f,t}$	birth tax [k€/cap]	$(-69, -22, 21)$

The triples for aggregate investments presents the lower bound, median, and upper bound for the targeted macro variable. The other triples present 5,50,95 percentiles. The Social Cost of Carbon is partitioned in its two components. The last two lines with population elasticity and birth tax require $u(.) = 0$

The table also presents permanent income elasticities. The elasticities of annualized real income with respect to capital $(1-\beta)\zeta_K$ and technology $(1-\beta)\zeta_A$ are positive through all parameter choices. A one-time increase in capital (technology) of one per cent increases welfare by as much as a permanent consumption growth of 0.07 (0.04) percent (median values). Welfare increases with exhaustible resources, $(1-\beta)\zeta_{R,t} > 0$, but the elasticities are small.

The elasticity for population $(1-\beta)\zeta_N$ requires $u(.) = 0$; we cannot calibrate γ_N, γ_m on observable data. Its value is often negative, but moderate. A one-time increase in population of one per cent reduces welfare as much as a permanent per-capita consumption drop of 0.05 per cent (median value). It looks as if the negative crowding of a common pool (Hardin, 1968) dominates the positive contribution to ideas (Kuznets, 1960), but the model does not rule out the inverse.

Yet, while ζ_N suggests a negative effect of population on welfare, there are other observations that temper this interpretation. The second observation is that, in a semi-endogenous growth model, long run income growth is proportional to population growth. The model is calibrated on the interpretation of historic data that a one percent population growth sustains an 0.4 [0.2,0.6] per cent per capita income growth (Jones, 2002). The same relation holds in levels: long-run TFP increases proportionally to long run population. Through that mechanism, more people will contribute to more welfare by endogenous adjustment of capital and innovation.

The third observation is based on the last row of the table. Parents take into account that children in a larger family will share the inheritance. What matters for correcting policies are those welfare effects not internalized within households. The birth externality as expressed in (52) is positive in most cases, i.e. the birth tax negative. Larger families receive a subsidy of 22 thousand euro per child (median result), though the calibration interval does not rule out a birth tax.

Combining the first three observations, starting at the social optimum, we find that a policy directed towards smaller families will increase welfare of the immediate next generation due to an increase in per capita wealth, it will decrease welfare in the long run because of lower TFP levels, and it will decrease welfare of the (altruist) parents distorting fertility away from the optimum. Yet there is a fourth observation further complicating the interpretation.

Neither the negative population elasticity, nor the positive birth externality, predict whether, in the model with $u(.) = 0$, the optimal allocation will show an increasing or declining population. The model describes an economy in which the socially optimal allocation exhibits constant family size; population either increases exponentially without bound, or completely collapses. The property is not unique to our model (cf Bretschger, 2020; Jones, 2020). What matters most for long term population dynamics (in our model with $u(.) = 0$) is the family size preferences γ.

Assuming $\gamma_N, \gamma_m > 0$, Proposition 1 shows how population socializing naturally avoids collapse, while congestion limits the growth. The proposition qualifies the base model with $u(.) = 0$. If innovations are sufficiently valuable, (52) informs us about the birth subsidy associated with economies of scale on the production side. The calibration is based on past trends; with some caution its quantitative results from Table 2 may be interpreted as applicable to the calibration period. Assuming $\gamma_N, \gamma_m > 0$ could be more appropriate for a long-run perspective. It informs us that even if there are economies of scale that, when considered on their own, warrant a subsidy for children, if such would lead to long-term fertility levels beyond reproduction the social cost of children will rise so much that the birth subsidy must be reduced or become a birth tax to bend population growth into a stable population. The same argument goes the other way. Even if scarce natural resources warrant a birth tax, it needs correction if the policy leads to a population collapse. In the long run, the social value of a human community and the carrying capacity of the earth dominate returns to scale in production.

Finally we compare our findings with literature. Schou (2002) finds a negative birth externality associated with scarce resources as a common pool, and a positive birth externality associated with knowledge spillovers for abatement technology. As his model has no endogenous TFP, the former tends to dominate, explaining a more pessimist outcome as we find. Gerlagh et al. (2018) has no endogenous technology, and thus only identifies the negative birth externality. Kruse-Andersen (2017) does not consider optimal climate policy. In our model, carbon taxes rise proportionally with aggregate outcome, so that an increased population directly implies an increased carbon tax. Without that adjustment, it is understandable that Kruse-Andersen (2017) finds population growth to negatively affect welfare. Bretschger (2020) considers the benefit of increased population through endogenous technology, while marginal climate change damages are assumed to decrease with cumulative emissions. At the same time, fossil fuels are assumed an exhaustible resource in a competitive market. Thus, while he keeps the positive population externalities there is no open pool resource crowding out problem. This explains his finding of no negative population externality.

5 Conclusions

We developed and analyzed a model with three aims in mind. First to integrate the literature on endogenous technological change and family planning into the recent climate economics models that provide closed-form solutions. Second, to use the framework to incorporate recent empirical findings on climate change and economic growth in calculations on the social costs of carbon. Third, to use the model to better understand the possible role of family size as part of optimal climate policy. Can the integrated framework help us to weigh the finite earth argument, where more people implies lower per capita consumption, versus a perspective where more people lead to an increasing pool of productive ideas?

Against these questions, let us collect the findings. A key message from Nordhaus' (1993a) seminal work has been that climate policies can be understood as investments in future welfare, to be treated on equal terms with investments in capital, infrastructure and knowledge. Optimal climate policy is an investment portfolio decision. Golosov et al. (2014) subsequently showed conditions under which the net present value of climate change damages brought by one unit of carbon dioxide can be calculated relatively straightforwardly, as the net present value of the emissions-damage response function discounted at the pure rate of time preference. We extended those results for global warming damaging the engine of growth, as supported by recent empirical literature. The approach we followed separates four decision domains; investments in capital, climate policy, innovation, and human capital jointly with fertility. By construction, each of these domains is decided independently of the others. The approach supports simple rules that describe the first-order tradeoffs for each domain; admittedly, abstracting from interactions between domains. The closed form supports a relatively convenient calculation of the social cost of carbon based on empirical climate-damage estimates.

As to the third aim, the integrated model presents both negative and positive externalities of a larger population, through congestion and innovation, in one framework. We can derive conditions for a positive or negative effect of population on average welfare. Calibrating the parameters suggests a somewhat surprising mixed combination of findings. Per capita welfare decreases with the number of people, all other things equal. The social optimum requires a birth *subsidy* for decentralization (for most parameter values). And optimal population may either increase without bound, or collapse, dependent on the (calibrated) fertility preferences.

The above applies if we abstract from direct population externalities $u(.) = 0$. While our analysis supports the idea that climate concerns do not stand in the way of a positive contribution of people to the engine of growth, we note that the production side of our model only describes man-made production of goods. Scarcity of nature, the pressure on natural reserves, worldwide biodiversity loss, but also positive social interactions, are all not captured in the description if $u(.) = 0$. The expected population growth to ten billion people and beyond is probably good for man-made material welfare. Yet, whether it also adds to a broad measure of welfare, probably depends mostly on our preferences for less tangible features such as biodiversity, nature, and space to live. These preferences are not only hard to estimate, they also likely differ so much between people, that the concept of an optimal population is unobtainable.

References

Aghion, P., Dechezleprêtre, A., Hemous, D., Martin, R., and Van Reenen, J. (2016). Carbon taxes, path dependency, and directed technical change: Evidence from the auto industry. *Journal of Political Economy*, 124(1):1–51.

Asheim, G. B. and Zuber, S. (2014). Escaping the repugnant conclusion: Rank-discounted utilitarianism with variable population. *Theoretical Economics*, 9(3):629–650.

Barro, R. J. and Becker, G. S. (1989). Fertility choice in a model of economic growth. *Econometrica: journal of the Econometric Society*, pages 481–501.

Blackorby, C., Bossert, W., and Donaldson, D. (1995). Intertemporal population ethics: critical-level utilitarian principles. *Econometrica: journal of the Econometric Society*, pages 1303–1320.

Bloom, N., Jones, C. I., Van Reenen, J., and Webb, M. (2020). Are ideas getting harder to find? *American Economic Review*, 110(4):1104–44.

Bohn, H. and Stuart, C. (2015). Calculation of a population externality. *American Economic Journal: Economic Policy*, 7(2):61 – 87.

Bretschger, L. (2017). Climate policy and economic growth. *Resource and Energy Economics*, 49:1–15.

Bretschger, L. (2020). Malthus in the light of climate change. *European Economic Review*, page 103477.

Bretschger, L., Lechthaler, F., Rausch, S., and Zhang, L. (2017). Knowledge diffusion, endogenous growth, and the costs of global climate policy. *European Economic Review*, 93:47–72.

Bretschger, L. and Pattakou, A. (2019). As bad as it gets: how climate damage functions affect growth and the social cost of carbon. *Environmental and resource economics*, 72(1):5–26.

Bretschger, L. and Suphaphiphat, N. (2014). Effective climate policies in a dynamic north–south model. *European Economic Review*, 69:59–77.

Bretschger, L. and Valente, S. (2011). Climate change and uneven development. *The Scandinavian Journal of Economics*, 113(4):825–845.

Brock, W. and Mirman, L. (1972). Optimal growth under uncertainty. *Journal of Economic Theory*, 4:479–513.

Broda, C. and Weinstein, D. E. (2006). Globalization and the gains from variety. *The Quarterly journal of economics*, 121(2):541–585.

Brunnschweiler, C. N., Peretto, P. F., and Valente, S. (2020). Wealth creation, wealth dilution and demography. *Journal of Monetary Economics*.

Burke, M., Hsiang, S. M., and Miguel, E. (2015). Global non-linear effect of temperature on economic production. *Nature*, 527(7577):235–239.

Calel, R. and Dechezlepretre, A. (2016). Environmental policy and directed technological change: evidence from the european carbon market. *Review of economics and statistics*, 98(1):173–191.

Colacito, R., Hoffmann, B., and Phan, T. (2019). Temperature and growth: A panel analysis of the united states. *Journal of Money, Credit and Banking*, 51(2-3):313–368.

De La Croix, D. and Gosseries, A. (2009). Population policy through tradable procreation entitlements. *International Economic Review*, 50(2):507–542.

De La Croix, D. and Gosseries, A. (2012). The natalist bias of pollution control. *Journal of Environmental Economics and Management*, 63(2):271–287.

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

Dietz, S. and Stern, N. (2015). Endogenous growth, convexity of damage and climate risk: how nordhaus' framework supports deep cuts in carbon emissions. *The Economic Journal*, 125(583):574–620.

Dietz, S. and Venmans, F. (2019). Cumulative carbon emissions and economic policy: in search of general principles. *Journal of Environmental Economics and Management*, 96:108–129.

Donadelli, M., Grüning, P., Jüppner, M., and Kizys, R. (2021). Global temperature, r&d expenditure, and growth. *Energy Economics*, page 105608.

Fankhauser, S. and Tol, R. S. (2005). On climate change and economic growth. *Resource and Energy Economics*, 27(1):1–17.

Fernald, J. G. and Jones, C. I. (2014). The future of us economic growth. *American economic review*, 104(5):44–49.

Galor, O. (2011). *Unified growth theory*. Princeton University Press.

Gerlagh, R. (2008). A climate-change policy induced shift from innovations in carbon-energy production to carbon-energy savings. *Energy Economics*, 30(2):425–448.

Gerlagh, R. (2011). Too much oil. *CESifo Economic Studies*, 57(1):79–102.

Gerlagh, R. and Liski, M. (2018). Consistent climate policies. *Journal of the European Economic Association*, 16(1):1–44.

Gerlagh, R., Lupi, V., and Galeotti, M. (2018). Family planning and climate change. *CESifo Working Paper*, 7421:Accepted for Scandinavian Journal of Economics.

Golosov, M., Hassler, J., Krusell, P., and Tsyvinski, A. (2014). Optimal taxes on fossil fuel in general equilibrium. *Econometrica*, 82(1):41–88.

Gradus, R. and Smulders, S. (1993). The trade-off between environmental care and long-term growth—pollution in three prototype growth models. *Journal of Economics*, 58(1):25–51.

Graff Zivin, J., Hsiang, S. M., and Neidell, M. (2018). Temperature and human capital in the short and long run. *Journal of the Association of Environmental and Resource Economists*, 5(1):77–105.

Hardin, G. (1968). The tragedy of the commons. *Science*, 162(3859):1243–1248.

Harford, J. D. (1998). The ultimate externality. *The American Economic Review*, 88(1):260–265.

Hassler, J., Krusell, P., Olovsson, C., and Reiter, M. (2018). Integrated assessment in a multi-region world with multiple energy sources and endogenous technical change. *Unpublished paper*.

Howard, P. H. and Sterner, T. (2017). Few and not so far between: a meta-analysis of climate damage estimates. *Environmental and Resource Economics*, 68(1):197–225.

Hsiang, S., Kopp, R., Jina, A., Rising, J., Delgado, M., Mohan, S., Rasmussen, D., Muir-Wood, R., Wilson, P., Oppenheimer, M., et al. (2017). Estimating economic damage from climate change in the united states. *Science*, 356(6345):1362–1369.

Iverson, T. and Karp, L. (2017). Carbon taxes and climate commitment with non-constant time preference. *CUDARE WPs*.

Iverson, T. and Karp, L. (2021). Carbon taxes and climate commitment with non-constant time preference. *The Review of economic studies*, 88(2):764–799.

Jones, C. I. (1995). R & d-based models of economic growth. *Journal of political Economy*, 103(4):759–784.

Jones, C. I. (2002). Sources of us economic growth in a world of ideas. *American Economic Review*, 92(1):220–239.

Jones, C. I. (2020). The end of economic growth? unintended consequences of a declining population. Technical report, National Bureau of Economic Research.

Kim, H. S., Matthes, C., and Phan, T. (2021). Extreme weather and the macroeconomy. *Available at SSRN 3918533*.

Kruse-Andersen, P. (2017). Testing r&d-based endogenous growth models. *Available at SSRN 2947528*.

Kruse-Andersen, P. (2019). Directed technical change, environmental sustainability, and population growth. *Univ. of Copenhagen Dept. of Economics Discussion Paper*, (19-12).

Krusell, P., Kurusçu, B., and Smith, A. A. (2002). Equilibrium welfare and government policy with quasi-geometric discounting. *Journal of Economic Theory Journal of Economic Theory*, 105:42–72.

Kuznets, S. (1960). Population change and aggregate output. In *Demographic and economic change in developed countries*, pages 324–351. Columbia University Press.

Lapan, H. E. and Enders, W. (1990). Endogenous fertility, ricardian equivalence, and debt management policy. *Journal of Public Economics*, 41(2):227–248.

Mankiw, N. G., Romer, D., and Weil, D. N. (1992). A contribution to the empirics of economic growth. *The quarterly journal of economics*, 107(2):407–437.

Masters, W. A. and McMillan, M. S. (2001). Climate and scale in economic growth. *Journal of Economic Growth*, 6(3):167–186.

Miller, S., Chua, K., Coggins, J., and Mohtadi, H. (2021). Heat waves, climate change, and economic output. *Journal of the European Economic Association*.

Montmartin, B. and Massard, N. (2015). Is financial support for private r&d always justified? a discussion based on the literature on growth. *Journal of Economic Surveys*, 29(3):479–505.

Murtaugh, P. A. and Schlax, M. G. (2009). Reproduction and the carbon legacies of individuals. *Global Environmental Change*, 19:14 – 20.

Nordhaus, W. D. (1993a). Optimal greenhouse-gas reductions and tax policy in the" dice" model. *The American Economic Review*, 83(2):313–317.

Nordhaus, W. D. (1993b). Rolling the "dice": An optimal transition path for controlling greenhouse. *Resource and Energy Economics*, 15:27–50.

O'Neill, B. C. and Wexler, L. (2000). The greenhouse externality to childbearing: a sensitivity analysis. *Climatic Change*, 47:283 – 324.

Park, R. J., Goodman, J., Hurwitz, M., and Smith, J. (2020). Heat and learning. *American Economic Journal: Economic Policy*, 12(2):306–39.

Peretto, P. F. and Valente, S. (2015). Growth on a finite planet: resources, technology and population in the long run. *Journal of Economic Growth*, 20(3):305–331.

Peretto, P. F. and Valente, S. (2019). Growth with deadly spillovers. *CESifo Working Paper*, 7761.

Rezai, A. and Van der Ploeg, F. (2016). Intergenerational inequality aversion, growth, and the role of damages: Occam's rule for the global carbon tax. *Journal of the Association of Environmental and Resource Economists*, 3(2):493–522.

Romer, P. M. (1990). Endogenous technological change. *Journal of political Economy*, 98(5, Part 2):S71–S102.

Schou, P. (2002). Pollution externalities in a model of endogenous fertility and growth. *International Tax and Public Finance*, 9:709 – 725.

Scovronik, N., Budolfdon, M. B., Denning, F., Fleurbaey, M., Siebert, A., Socolow, R. H., Spears, D., and Wagner, F. (2016). Impact of population growth and population ethics on climate change mitigation policy. *Proceedings of the National Academy of Sciences*, 114(46):12338 – 12343.

Simon, J. (1981). The ultimate resource.

Sinn, H.-W. (2008). Public policies against global warming: a supply side approach. *International tax and public finance*, 15(4):360–394.

Stern, N. (2008). The economics of climate change. *American Economic Review*, 98(2):1–37.

Tol, R. S. (2009). The economic effects of climate change. *Journal of economic perspectives*, 23(2):29–51.

Van den Bijgaart, I., Gerlagh, R., and Liski, M. (2016). A simple formula for the social cost of carbon. *Journal of Environmental Economics and Management*, 77:75–94.

van der Ploeg, F. and de Zeeuw, A. (2018). Climate tipping and economic growth: Precautionary capital and the price of carbon. *Journal of the European Economic Association*, 16(5):1577–1617.

Van der Ploeg, R., Dietz, S., Rezai, A., Venmans, F., et al. (2020). Are economists getting climate dynamics right and does it matter? *CESifo Working Paper*, 8122.

Weitzman, M. L. (1973). Duality theory for infinite horizon convex models. *Management Science*, 19(7):783–789.

Welsby, D., Price, J., Pye, S., and Ekins, P. (2021). Unextractable fossil fuels in a 1.5 c world. *Nature*, 597(7875):230–234.

Wynes, S. and Nicholas, K. A. (2017). The climate mitigation gap: education and government recommendations miss the most effective individual actions. *Environmental Research Letters*, 12:1–9.

A Appendix

A.1 Proofs

Preliminaries. Elasticities of the value function.
Consider the social optimum welfare function $V(K_t, A_t, N_t, h_t, (E_{t-i})_i, R_t)$, and we define the semi-elasticities $\zeta_{K,t} = (\partial V_t/\partial K_t)K_t$, $\zeta_{A,t} = (\partial V_t/\partial A_t)A_t$, $\zeta_{N,t} = (\partial V_t/\partial N_t)N_t$, $\zeta_{h,t} = (\partial V_t/\partial h_t)h_t$, $\zeta_{R,t} = (\partial V_t/\partial R_t)R_t$. These semi-elasticities have the following meaning: if the capital stock increases (exogenously) by 1 per cent at the start of time t, than welfare at that time increases by the same amount as if consumption would increase by $(1-\beta)\zeta_{K,t}$ per cent, forever.

Write down the Lagrangean for the planner's recursive optimization:

$$\mathcal{L} = V_t = \ln\left(\frac{C_t}{N_t}\right) + \gamma\ln(f_t) + \beta V(K_{t+1}, A_{t+1}, N_{t+1}, h_{t+1}, (E_{t+1-i})_i, R_{t+1}) \tag{54}$$

$$+\beta\lambda_{K,t+1}\left[\Omega_t(.)A_t^{\frac{1}{\varepsilon-1}}(1-s_{A,t})X_t(.) - C_t - K_{t+1}\right] \tag{55}$$

$$+\beta\lambda_{A,t+1}\left[\Gamma_t(.)(s_{A,t}X_t(.))^{1-\psi}A_t^{\varphi} - A_{t+1}\right] \tag{56}$$

$$+\beta\lambda_{N,t+1}\left[(1-\delta_N+f_t)N_t - N_{t+1}\right] \tag{57}$$

$$+\beta\lambda_{h,t+1}\left[x_t^{\eta_s}h_t^{\eta_h} - h_{t+1}\right] \tag{58}$$

$$+\beta\lambda_{R,t+1}\left[R_t - E_t - R_{t+1}\right] \tag{59}$$

where we substituted $X_t(.) = (K_t)^{\alpha}(f_t(1,E_t))^{\kappa}[(h_t(1-\phi f_t - x_t f_t)N_t)]^{1-\alpha-\kappa}$. Taking the FOCs for $K_{t+1}, A_{t+1}, N_{t+1}$ immediately expresses the welfare elasticities in terms of Lagrangean dual variables. Here we use subscript t not having established yet that the parameters ζ are constant and independent of state variables:

$$\zeta_{K,t} = \lambda_{K,t}K_t \ ; \ \zeta_{A,t} = \lambda_{A,t}A_t \ ; \ \zeta_{N,t} = \lambda_{N,t}N_t \ ; \ \zeta_{h,t} = \lambda_{h,t}h_t \ ; \ \zeta_{R,t} = \lambda_{R,t}R_t. \tag{60}$$

For interpretation, note that the variable $\zeta_{K,t}$ measures the marginal effect of a relative increase in capital on welfare, measured in welfare units ('utils'). It also measures the value of K_t (shadow price times level), in utils. We therefore refer to $\zeta_{K,t}$ (and other $\zeta_{.,t}$) interchangeable as the 'semi-elasticities' and the 'auxiliary value variables'.

Proof of Proposition 1 (Social optimum characterization with $u(.) = 0$).

Proof. We will use equivalence to the above welfare program, but for the proof itself, find a slightly different formulation more convenient. Instead of having K_{t+1} as control/state variable, we'd like to work with the investment share out of output, $s_{K,t}$. We write down the Lagrangean for the planner's optimization:

$$\mathcal{L} = \sum_{t=0}^{\infty} \beta^t \left[\ln\left(C_t/N_t\right) + \gamma \ln(f_t)\right] \tag{61}$$

$$+ \sum_{t=0}^{\infty} \beta^t \lambda_{Y,t} \left[\Omega_t(.) A_t^{\frac{1}{\varepsilon-1}} (1 - s_{A,t}) X_t(.) - Y_t\right] \tag{55'}$$

$$+ \sum_{t=0}^{\infty} \beta^t \lambda_{C,t} \left[(1 - s_{K,t}) Y_t - C_t\right] \tag{55''}$$

$$+ \sum_{t=0}^{\infty} \beta^{t+1} \lambda_{K,t+1} \left[s_{K,t} Y_t - K_{t+1}\right] \tag{55'''}$$

$$+ \sum_{t=0}^{\infty} \beta^{t+1} \lambda_{A,t+1} \left[\Gamma_t(.) (s_{A,t} X_t(.))^{1-\psi} A_t^{\varphi} - A_{t+1}\right] \tag{56'}$$

$$+ \sum_{t=0}^{\infty} \beta^{t+1} \lambda_{N,t+1} \left[(1 - \delta_N + f_t) N_t - N_{t+1}\right] \tag{57'}$$

$$+ \sum_{t=0}^{\infty} \beta^{t+1} \lambda_{h,t+1} \left[x_t^{\eta_s} h_t^{\eta_h} - h_{t+1}\right] \tag{58'}$$

$$+ \sum_{t=0}^{\infty} \beta^{t+1} \lambda_{R,t+1} \left[R_t - E_t - R_{t+1}\right] \tag{59'}$$

where we substituted $X_t(.) = (s_{K,t-1} Y_{t-1})^{\alpha} (f_t(1, E_t))^{\kappa} [(h_t(1 - \phi f_t - x_t f_t) N_t)]^{1-\alpha-\kappa}$ and we use *current* value dual variables λ. Comparing the two programs, we immediately see that the two programs are equivalent in the sense that the dual variables with the same label have the same value, and the dual variables for C_t, K_{t+1} and Y_t, respectively, are connected through $\lambda_{C,t} = \beta \lambda_{K,t+1}$ (FOC for $s_{K,t}$), $\lambda_{Y,t} = (1 - s_{K,t}) \lambda_{C,t} + s_{K,t} \beta \lambda_{K,t+1} = \beta \lambda_{K,t+1}$ (FOC for Y_t), $\lambda_{C,t} = 1/C_t$ (FOC for C_t). Furthermore, the value of final good production measured in utils ($1/C_t$) is given by the auxiliary value variable $\zeta_{Y,t} = \lambda_{Y,t} Y_t$, the value of capital $\zeta_{K,t} = \lambda_{K,t} K_t$, the value of technology $\zeta_{A,t} = \lambda_{A,t} A_t$, we also use $\zeta_{N,t} = \lambda_{N,t} N_t$ and for the value of human capital $\zeta_{h,t} = \lambda_{h,t} h_t$. For future reference, it is also useful to define the auxiliary value variable of X_t: $\zeta_{X,t} \equiv \zeta_{Y,t} + \beta(1 - \psi) \zeta_{A,t+1}$ and for the exhaustible resource $\zeta_{R,t} = \lambda_{R,t} R_t$, though this dual variable will not be constant.

An important feature of the economy is that we can solve for the optimal path block-recursively, in four blocks. As a first block, we solve for capital and innovation investments and the associated auxiliary value variables: $s_{K,t}, s_{A,t}, \zeta_{K,t}, \zeta_{A,t}$. We do so for an arbitrary path for the other control and state variables. We will see that savings rates and the auxiliary value variables $s_K, s_A, \zeta_K, \zeta_A$ are constant along the optimal path, independent of the level of state variables, and independent of the other control variables g_t, f_t, x_t: their values do not change when e.g. schooling or fertility deviates from the social optimum. Then as second block we derive the remaining auxiliary value variables for population, human capital, and the capital-labour composite: $\zeta_{X,t}, \zeta_{N,t}, \zeta_{h,t}$, which we also find to be constant along the optimal path. Then, as third block, we can solve for the social costs of carbon and optimal emissions as described through policy g_t. As the fourth block, we solve for optimal fertility f_t and schooling x_t.

We start with the first block, and solve for the optimal $(s_{K,t}, s_{A,t}, \zeta_{K,t}, \zeta_{A,t})$ for an arbitrary path for the other control and state variables. Consider the welfare program (32)-(39). Recall that we have 5 control variables $(s_{K,t}, s_{A,t}, f_t, x_t, E_t)$. In this first block, we show that for any given path (not necessarily optimal) for control variables (f_t, x_t, E_t) with (consistent) state variables (N_t, h_t, R_t, T_t), there is a *unique* optimal path for $(s^*_{K,t}, s^*_{A,t})$ and associated (K^*_t, A^*_t) that is independent of the other control and state variables. The policies and auxiliary value variables $s^*_K, s^*_A, \zeta^*_K, \zeta^*_A$ are constant.

We assume fertility, population, schooling, human capital, emissions and temperature as given. We rewrite the welfare program (32)-(39) using savings shares $s_{K,t}$ instead of investments as control variable, and we use auxiliary variables $\widetilde{x} = \ln(X)$ and inequalities instead of equalities to cast the welfare program in a strict convex feasibility space. That is, the only endogenous variables we consider here are $\{s_{K,t},\ s_{A,t},\ \widetilde{k}_t,\ \widetilde{a}_t,\ \widetilde{y}_t,\ \widetilde{c}_t\}$. The auxiliary program reads

$$\max V_0 = \sum_{t=0}^{\infty} \beta^t \left[\widetilde{c}_t - \widetilde{n}_t + \gamma ln(f_t)\right] \tag{62}$$

$$\widetilde{y}_t \leq \frac{1}{\varepsilon - 1}\widetilde{a}_t + \widetilde{x}_t(.) + \ln(1 - s_{A,t}) - \delta_Y T_t \tag{63}$$

$$\widetilde{c}_t \leq \ln(1 - s_{K,t}) + \widetilde{y}_t \tag{64}$$

$$\widetilde{k}_{t+1} \leq \ln(s_{K,t}) + \widetilde{y}_t \tag{65}$$

$$\widetilde{a}_{t+1} \leq \varphi\widetilde{a}_t + (1 - \psi)\left(\widetilde{x}_t(.) + \ln(s_{A,t})\right) - \frac{\delta_A}{\varepsilon - 1}T_t \tag{66}$$

where we substituted $\widetilde{x}_t(.) = \alpha\widetilde{k}_t + \kappa\ln(f_t(1, E_t)) + (1 - \alpha - \kappa)\ln[(h_t(1 - \phi f_t - x_t f_t)N_t)]$. We see all equations are linear in variables $\widetilde{k}_t, \widetilde{a}_t, \widetilde{y}_t, \widetilde{c}_t$ and the right-hand sides are strictly concave in the variables $s_{K,t}, s_{A,t}$; thus feasibility space is strictly convex. It follows that the maximization solution is unique and identification of an allocation that satisfies the first order conditions is sufficient for optimality. We note that uniqueness of the optimal solution and sufficiency of FOCs carries over to the original welfare program (the one before taking logs with variables Y_t, A_t, etc).

The Lagrangean for this welfare program reads

$$\mathcal{L}_t = \sum_{t=0}^{\infty} \beta^t \left[\widetilde{c}_t - \widetilde{n}_t + \gamma ln(f_{t+i})\right] \tag{67}$$

$$+ \sum_{t=0}^{\infty} \beta^t \zeta_{Y,t} \left[\frac{1}{\varepsilon - 1}\widetilde{a}_t + \widetilde{x}_t(.) + \ln(1 - s_{A,t}) - \delta_Y T_t - \widetilde{y}_t\right] \tag{68}$$

$$+ \sum_{t=0}^{\infty} \beta^t \zeta_{C,t} \left[\ln(1 - s_{K,t}) + \widetilde{y}_t - \widetilde{c}_t\right] \tag{69}$$

$$+ \sum_{t=0}^{\infty} \beta^{t+1} \zeta_{K,t+1} \left[\ln(s_{K,t}) + \widetilde{y}_t - \widetilde{k}_{t+1}\right] \tag{70}$$

$$+ \sum_{t=0}^{\infty} \beta^{t+1} \zeta_{A,t+1} \left[\varphi\widetilde{a}_t + (1 - \psi)\left(\widetilde{x}_t(.) + \ln(s_{A,t})\right) - \frac{\delta_A}{\varepsilon - 1} T_t - \widetilde{a}_{t+1}\right] \tag{71}$$

with current auxiliary value variables $\zeta_{Y,t}, \zeta_{C,t}, \zeta_{K,t}, \zeta_{A,t}$. Note that these ζ's have exactly the same value and meaning as in the original welfare program without logs. We immediately see from the FOC for $\widetilde{c}_t$ that $\zeta_{C,t} = 1$ The FOCs for $\widetilde{y}_t$, $\widetilde{k}_t$, $\widetilde{a}_t$, $s_{K,t}$, $s_{A,t}$ are given by

$$\zeta_{Y,t} = 1 + \beta\zeta_{K,t+1} \tag{72}$$

$$\begin{aligned} \zeta_{K,t} &= \alpha\zeta_{Y,t} + \alpha\beta(1 - \psi)\zeta_{A,t+1} \\ &= \alpha + \alpha\beta\zeta_{K,t+1} + \alpha\beta(1 - \psi)\zeta_{A,t+1} \end{aligned} \tag{73}$$

$$\zeta_{A,t} = \frac{1 + \beta\zeta_{K,t+1}}{\varepsilon - 1} + \beta\varphi\zeta_{A,t+1} \tag{74}$$

$$\frac{s_{K,t}}{1 - s_{K,t}} = \beta\zeta_{K,t+1} \tag{75}$$

$$\frac{s_{A,t}}{1 - s_{A,t}} = \frac{(1 - \psi)\zeta_{A,t+1}}{\zeta_{K,t+1}} \tag{76}$$

The first equation states that, as marginal utility equals inverse consumption, the util value of output equals the value of consumption plus the contribution of capital investments to next-period production. The value of capital adds to that the subsequent period value of knowledge. The third equation describes the value of technology as a share of output value (dependent on the elasticity ε) plus the next-period value. The fourth equation states that the value of savings relative to consumption is given by the next-period capital value. The last equation states that innovation versus consumption good expenditures are proportional to their value shares.

A closer look at the FOCs reveals a remarkable feature, typical for the Brock-Mirman model. In most other dynamic models, the dynamic equations for dual variables include references to state variables. Imposing stationarity then implies imposing a balanced growth path for those state variables. In the Brock Mirman model, and our extension, the state variables are substituted out of the FOCs that describe the dynamics for ζ. We see that the above *two* equations (73) and (74) have *two* variables $\zeta_{K,t}, \zeta_{A,t}$, which implies that (unless singular) they contain a stationary solution $\zeta_{Y,t} = \zeta_Y$ and $\zeta_{A,t} = \zeta_A$. That is, we can find a stationary solution for $\zeta_{K,t}, \zeta_{A,t}$. Before we do so, let us rule out non-stationary solutions.

We rewrite the difference system for the zetas in vector-matrix notation: $\zeta_t = (\zeta_{K,t}, \zeta_{A,t})'$ is the dual variable vector and M is the 2x2 matrix

$$M = \begin{pmatrix} \alpha\beta & \alpha\beta(1-\psi) \\ \frac{\beta}{\varepsilon-1} & \beta\varphi \end{pmatrix} \tag{77}$$

so that the dual value variables satisfy the linear difference equation

$$\zeta_t = (\alpha, \frac{1}{\varepsilon-1})' + M\zeta_{t+1} \tag{78}$$

Let $\zeta^* = (\zeta_K^*, \zeta_A^*)$ be the stationary (vector) solution, $\zeta^* = (1/\beta, 1/(\varepsilon-1))' + M\zeta^*$ and $\Delta_t = \zeta_t - \zeta^*$ the deviation from the stationary solution, so that $\Delta_t = M\Delta_{t+1}$. Let λ^* be an eigenvalue of M and Δ^* the associated eigenvector, which satisfies $M\Delta^* = \lambda^*\Delta^*$. A deviation from the stationary solution that is proportional to an eigenvector, say $\Delta_0 = \Delta^*$, increases exponentially over time (decreases going backwards in time) by factor $1/\lambda^*$: $\Delta_t = \Delta_0/(\lambda^*)^t$. We will now establish that all eigenvalues are in the interval $(-\beta, \beta)$, which implies that any deviation increases in size at a factor larger than $1/\beta$ and thus only the stationary solution can satisfy the transversality condition $\lim_{t\to\infty} \beta^t\zeta_t \to 0$.

To establish that all $\lambda^* \in (-\beta, \beta)$, we consider the (quadratic) eigenfunction $p(\lambda) = det(M - \lambda I)$:

$$p(\lambda) = (\lambda - \alpha\beta)(\lambda - \beta\varphi) - \frac{\alpha\beta^2(1-\psi)}{\varepsilon - 1} \tag{79}$$

The eigenvalues for M equal the roots of $p(.)$. We find that $p(\lambda)$ takes its minimum value for $\lambda = \beta(\alpha + \varphi)/2$, which is between 0 and β, and returns a negative value. Thus $p(.)$ has two real-valued roots. For $\lambda = \beta$ we find $p(\beta) = \beta^2[(1-\alpha)(1-\varphi) - \alpha(1-\psi)/(\varepsilon-1)] > 0$ since we required $(1-\alpha)(\varepsilon-1)(1-\varphi) > \alpha(1-\psi)$ to guarantee finite growth, see also Footnote 23 and eq (135). For $\lambda = -\beta$ we find $p(-\beta) = \beta^2[(1+\alpha)(1+\varphi) - \alpha(1-\psi)/(\varepsilon-1)] > \beta^2[(1-\alpha)(1-\varphi) - \alpha(1-\psi)/(\varepsilon-1)] > 0$. We conclude that both eigenvalues are in the interval $(-\beta, \beta)$; the dual variables $\zeta_{K,t}, \zeta_{A,t}$ immediately take their stationary values ζ_K, ζ_A; they cannot follow another non-stationary path.[37] As we note in the main text, the constant values for the dual variables and savings rates are compatible with rich transitionary dynamics for state and flow dynamics. We provide some simulation results in Appendix C.1.

We find the equations that define the stationary solution as:

$$\zeta_K = \alpha + \alpha\beta\zeta_K + \alpha\beta(1-\psi)\zeta_A \tag{80}$$

$$\zeta_A = \frac{1+\beta\zeta_K}{\varepsilon-1} + \beta\varphi\zeta_A \tag{81}$$

We can now derive the auxiliary value variables for output, knowledge, and the capital-labor composite:

$$\zeta_Y = \frac{(\varepsilon-1)(1-\beta\varphi)}{(1-\alpha\beta)(\varepsilon-1)(1-\beta\varphi) - \alpha\beta^2(1-\psi)} \tag{82}$$

$$\zeta_K = \frac{\alpha(\varepsilon-1)(1-\beta\varphi) + \alpha\beta(1-\psi)}{(1-\alpha\beta)(\varepsilon-1)(1-\beta\varphi) - \alpha\beta^2(1-\psi)} \tag{83}$$

$$\zeta_A = \frac{1}{(1-\alpha\beta)(\varepsilon-1)(1-\beta\varphi) - \alpha\beta^2(1-\psi)} \tag{84}$$

[37] We find that instability of deviations from the stationary path and finite growth require the same conditions; this is no coincidence. From dual theory, we know that for a convex infinite horizon finite sum optimization problem (in our model: finite sum is equivalent to finite growth), it is sufficient to find *a* path for control, state and co-state variables that satisfies the FOCs. There can exist no other paths that satisfy the transversality conditions (Weitzman, 1973). From a technical perspective, our zetas define the separating plane, and local optimization (e.g. per period) is a sufficient condition for a global optimum. One can also understand sufficiency without separating planes. The recursive per-period optimization defines a contracting mapping from value-function space to itself, $V \to \widehat{V} : \widehat{V}_t = \max u_t + \beta V_{t+1}$, but only if growth is finite. Contraction means that the distance between two value functions strictly decreases through the mapping. It implies that there is a unique fixed point: any value function that is per-period consistent, as the one we construct with the stationary zetas, must be the unique infinite-horizon value function.

Taking the FOCs for $s_{K,t}$ and $s_{A,t}$ (75), (76) then gives constant s_K, s_A:

$$\frac{s_K}{1-s_K} = \beta\zeta_K \tag{85}$$

$$\frac{s_A}{1-s_A} = \frac{(1-\psi)\zeta_A}{\zeta_K} \tag{86}$$

which gives the main text optimal investments:

$$s_K^* = \alpha\beta\left[1+\frac{\beta(1-\psi)}{(\varepsilon-1)(1-\beta\varphi)}\right] \tag{40}$$

$$s_A^* = \frac{\beta(1-\psi)}{(\varepsilon-1)(1-\beta\varphi)+\beta(1-\psi)} \tag{41}$$

This is the unique stable, and stationary solution $s_K^*, s_A^*, \zeta_Y^*, \zeta_A^*$ for the equations (73)-(76). Importantly, though savings and auxiliary value variables are constant, the state and flow variables $\widetilde{y}_t$, $\widetilde{k}_t$, $\widetilde{a}_t$ (or Y_t, K_t, A_t) follow their dynamic equations (63)-(66) (with equalities) and are non-stationary (see Appendix C.1). For the original economy, given K_0, A_0 we determine X_0, which determines Y_0, A_1 through s_A^*, and Y_0 determines K_1 through s_K^*, etc. (abstracting from climate change, population and education).

At a deeper level, we note from the derivation above that the value of output and innovation ζ_K^*, ζ_A^* comes out of the FOCs for $\widetilde{k}_t$, $\widetilde{a}_t$, and is thus *independent* of the investment share $s_{K,t}$ and innovation share $s_{A,t}$, that is, also when $s_{K,t}, s_{A,t}$ follow some exogenous non-optimal path.

For future reference (when identifying capital and research policies), it is useful to note that capital used in final goods production, $K_{Y,t} = (1-s_A)K_t$, so that (12) becomes $(1-s_A)r_{t+1}K_{t+1} = \alpha[(\varepsilon-1)/\varepsilon]Y_{t+1}$, into $(1-s_A)s_KY_t = \alpha[(\varepsilon-1)/\varepsilon]Y_{t+1}/r_{t+1}$. Thus,

$$\frac{Y_{t+1}}{r_{t+1}} = \beta\frac{\varepsilon}{\varepsilon-1}Y_t \tag{87}$$

The value of next-period output, valued at the market interest rate that firms pay, equals the current value of output multiplied by the time preference factor β and the term $\varepsilon/(\varepsilon-1)$, caused by capital receiving a too low return due to monopolistic pricing. The utility rate of substitution for final goods equals

$$\frac{C_{t+1}}{\beta C_t} = \frac{\varepsilon}{\varepsilon-1}r_{t+1}. \tag{88}$$

This immediately pinpoints $\sigma_k^* = \varepsilon/(\varepsilon-1)$ for implementation of optimal savings (see proof for Prop 2). This ends the first block.

We now move to the second block, where we characterize the dual variables $\zeta_{N,t}, \zeta_{h,t}$ through the Lagrangean of the full welfare program (61)-(58)$'$. First we establish the value for X_t, $\zeta_X \equiv \zeta_Y + \beta(1-\psi)\zeta_A$:

$$\zeta_X = \frac{(\varepsilon-1)(1-\beta\varphi)+\beta(1-\psi)}{(1-\alpha\beta)(\varepsilon-1)(1-\beta\varphi)-\alpha\beta^2(1-\psi)} \tag{89}$$

We read from the above that if technology has a negligible value share in output, $\varepsilon \to \infty$, we have $\zeta_X \to 1/(1-\alpha\beta)$ as in the Brock Mirman model. In that case, due to decreasing returns to scale, the value of labour falls short of the value of consumption: $(1-\alpha-\kappa)\zeta_X < 1$. On the other hand, if technology has a large value share in output, $\varepsilon \searrow 1+\alpha\beta(1-\psi)/(1-\alpha\beta)(1-\beta\varphi)$, returns to innovation are so large that we find $\zeta_X \to \infty$ and the value of labour exceeds the value of consumption.

The FOC for N_t is

$$\lambda_{N,t} = \beta\lambda_{N,t+1}(1-\delta_N+f_t)+(1-\alpha-\kappa)\frac{\zeta_X}{N_t}-\frac{1}{N_t} \tag{90}$$

We multiply by N_t/β^t and rewrite it as

$$\zeta_{N,t} = (1-\alpha-\kappa)\zeta_X - 1 + \beta\zeta_{N,t+1} \Rightarrow \zeta_N = \frac{(1-\alpha-\kappa)\zeta_X - 1}{1-\beta} \tag{91}$$

The equation informs us that if technology has a negligible value share in output, $\varepsilon \to \infty$, we have $\zeta_N < 0$: an increasing population reduces welfare due to decreasing returns to scale in production. On the other hand, if technology has a large value share in output, $\varepsilon \searrow 1+\alpha\beta(1-\psi)/(1-\alpha\beta)(1-\beta\varphi)$, we find $\zeta_N > 0$: an increasing population increases welfare due to increasing returns to scale in production.

The FOC for h_t is

$$\lambda_{h,t} = \frac{1-\alpha-\kappa}{h_t}\left(\lambda_{Y,t}Y_t + \beta\lambda_{A,t+1}A_{t+1}\right) + \eta_h\beta\lambda_{h,t+1}\frac{h_{t+1}}{h_t} \tag{92}$$

We rewrite these using $\zeta_{h,t} = \lambda_{h,t}h_t$ for the relative value of human capital. We then see that the value of human capital equals the discounted cumulative value of output:

$$\zeta_{h,t} = (1-\alpha-\kappa)\zeta_X + \beta\eta_h\zeta_{h,t+1} \Rightarrow \zeta_h = \frac{(1-\alpha-\kappa)}{1-\beta\eta_h}\zeta_X \tag{93}$$

This ends the second block.

We can now move on to the third block for characterizing emissions and the SCC from the FOCs for R_t, E_t:

$$\lambda_{R,t} \geq \beta\lambda_{R,t+1} \tag{94}$$

$$\lambda_{Y,t}\frac{\partial Y_t}{\partial E_t} = \beta\lambda_{R,t+1} + \sum_{i=1}^{\infty}\theta_i\beta^t\left[\delta_Y\zeta_Y + \delta_A\beta\zeta_A\right] \tag{95}$$

with the first weak inequality an equality iff $R_{t+1} > 0$. We immediately see we can set the Hotelling rent to $q_t = \lambda_{R,t}/\lambda_{Y,t}$ so that we can rewrite the FOC for E_t as a Social Costs of Carbon formula with $\tau_{e,t} = g^*Y_t$:

$$g^* = \left[\delta_Y + \frac{\beta\delta_A}{1-\beta\varphi}\right]\sum_{i=1}^{\infty}\beta^i\theta_i. \tag{42}$$

We emphasize that the above FOCs do not depend on optimality of $s_{K,t}, s_{A,t}$, nor on the choices for fertility and human capital studied below. The SCC rule remains valid also for exogenously given $s_{K,t}, s_{A,t}, f_t, h_t$.

Finally, the fourth block solves for fertility and education. Consider the FOCs in (61)-(58)$'$ for f_t, x_t:

$$\frac{\gamma}{f_t} + \beta\lambda_{N,t+1}N_t = (\phi + x_t)\frac{1-\alpha-\kappa}{l_t}\left(\lambda_{Y,t}Y_t + \beta\lambda_{A,t+1}A_{t+1}\right) \tag{96}$$

$$\eta_s\frac{\beta\lambda_{h,t+1}h_{t+1}}{x_t} = f_t\frac{1-\alpha-\kappa}{l_t}\left(\lambda_{Y,t}Y_t + \beta\lambda_{A,t+1}A_{t+1}\right) \tag{97}$$

Rewriting the FOC for fertility f_t tells that the ratio between total time spent on raising children $\phi f_t + x_t f_t$ versus labour supply, equals the ratio between direct utility derived from children plus the next-period value of a larger population, versus the value of adding to labour supply.

$$\frac{\phi f^* + x^* f^*}{1 - \phi f^* - x^* f^*} = \frac{\gamma + \beta\zeta_N\widetilde{f}}{(1-\alpha-\kappa)\zeta_X}. \tag{43}$$

where $\widetilde{f} = fN_t/N_{t+1} = f^*/(1 + f^* - \delta_N)$ is the (constant) share of new born in the population. Rewriting the FOC for schooling x_t tells that the ratio between total time spent on schooling $x_t f_t$ versus labour supply, equals the ratio between direct utility derived from children plus the next-period value of a larger population, versus the value of adding to labour supply. The last FOC informs us that the time share of educating children increases with the value share of human capital η, and decreases with the time costs of raising children:

$$\frac{x^* f^*}{1-\phi f^* - x^* f^*} = \frac{\eta_s \beta \zeta_h}{(1-\alpha-\kappa)\zeta_X} = \frac{\eta_s \beta}{1-\beta\eta_h} \tag{44}$$

$$\Rightarrow x^* f^* = \frac{\beta\eta_s}{1-\beta+\beta\eta_n+\beta\eta_s}(1-\phi f^*) \tag{98}$$

Note that the fraction on the RHS is always smaller than half. Thus, we can picture the last equation as a line in $(\phi f^*, x^* f^*)$-space, from the right-lower corner $(1,0)$ to the left-upper corner $(0, \frac{\beta\eta_s}{1-\beta+2\beta\eta})$. Running over this line, we can evaluate (44)', and see that the LHS increases monotonically from a value below one to infinity. The RHS, meanwhile increases or decrease, dependent on the sign of ζ_N, starting at $\frac{\gamma}{(1-\alpha-\kappa)\zeta_X}$. Thus, fertility f^* increases (decreases) while schooling x^* decreases (increases) with γ (η). Q.E.D.

Sanity check.
As the proof above is rather complex, it is useful to check some results against a simple benchmark, without resources, no climate change, and exogenous technology, $\kappa = 0$, $\psi = 1$. We then expect that a proportional increase in capital and population does not affect welfare, which we confirm:

$$\zeta_K + \zeta_N = \frac{\alpha}{1-\alpha\beta} + \frac{1-\alpha}{(1-\alpha\beta)(1-\beta)} - \frac{1}{1-\beta} = \frac{\alpha-\alpha\beta+1-\alpha-1+\alpha\beta}{(1-\alpha\beta)(1-\beta)} = 0. \tag{99}$$

Proof of Proposition 1 (Social optimum characterization with $\gamma_N, \gamma_m > 0$).

Proof. We establish monotonicity of population dynamics. The equivalence in the last line of the proposition follows immediately. Write down the Lagrangean, which exactly copies (61)-(59)', apart from (61), which becomes

$$\mathcal{L} = \sum_{t=0}^{\infty} \beta^t \left[\ln(C_t/N_t) + \gamma_f \ln(f_t) + \gamma_N \ln(N_t) - \gamma_m N_t\right] ... \tag{100}$$

Fig 1 draws the phase diagram in state-co-state space for $(N_t, \zeta_{N,t})$.[38] The FOC for N_t

[38] The lines do not need to cross above the x-axis! $\zeta_N^{**} < 0$ is a possible outcome

(91) provides the dynamics for $\zeta_{N,t}$ and the locus for $\Delta\zeta_{N,t} = 0$:

$$\zeta_{N,t} = (1-\alpha-\kappa)\zeta_X - 1 + \gamma_N - \gamma_m N_t + \beta\zeta_{N,t+1} \tag{101}$$

$$\Delta\zeta_{N,t} = 0 \Rightarrow \zeta_{N,t} = \frac{(1-\alpha-\kappa)\zeta_X - 1 + \gamma_N - \gamma_m N_t}{1-\beta} \tag{102}$$

The locus is linear with negative slope $-\gamma_m/(1-\beta)$. For larger population size (right of the locus), the top equation tells that $\zeta_{N,t} < \zeta_{N,t+1}$, that is, the optimal path moves up.

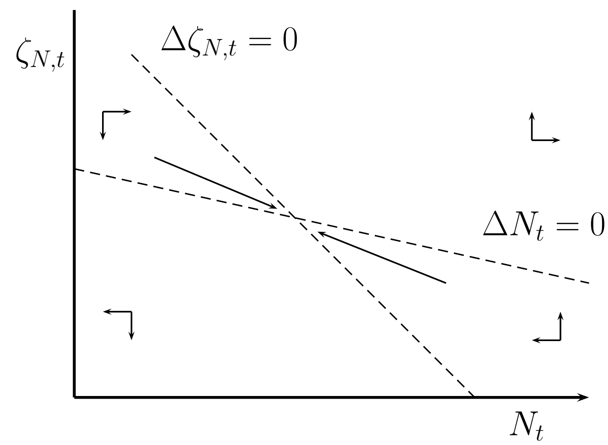

Figure 1: Phase diagram for population

Now consider the FOC for x_t (97) which gives

$$x_t f_t = \frac{\eta_s \beta \zeta_h}{(1-\alpha-\kappa)\lambda_X}(1 - \phi f_t - x_t f_t) \tag{103}$$

$$\Rightarrow x_t f_t = \frac{\eta_s \beta \zeta_h}{(1-\alpha-\kappa)\lambda_X + \eta_s\beta\zeta_h}(1-\phi f_t) \tag{104}$$

$$\Rightarrow \frac{\phi f_t + x_t f_t}{1-\phi f_t - x_t f_t} = \frac{(1-\alpha-\kappa)\lambda_X + \eta_s\beta\zeta_h}{(1-\alpha-\kappa)\lambda_X}\frac{\phi f_t}{1-\phi f_t} + \frac{\eta_s\beta\zeta_h}{(1-\alpha-\kappa)\lambda_X} \tag{105}$$

The point to take from the above is that, if we have a constant value for fertility over some domain, e.g. reproduction $f_t = \delta_N$, we must have expenditures on education $x_t f_t$ to also be a constant on that domain. We use that for identification of the locus for $\Delta N_t = 0$, by substituting $\zeta_{N,t+1}$ out of (101), through the FOC for f_t (96), rewritten as:

$$(\phi f_t + x_t f_t)(1-\alpha-\kappa)\lambda_X = \gamma + \beta\frac{f_t}{1+f_t-\delta_N}\zeta_{N,t+1}(1-\phi f_t - x_t f_t) \tag{106}$$

That is, for the locus with $\Delta N_t = 0$, we must have $f_t = \delta_N$, so that on the locus $x_t f_t$ is a constant (from the FOC for x_t), and thus $\zeta_{N,t+1}$ is a constant. It then follows from (101) that the locus is linear with (negative) slope $-\gamma_m$. That is, the locus $\Delta N_t = 0$ cuts the locus $\Delta\zeta_{N,t} = 0$ from below, as in the diagram. Above the locus $\Delta N_t = 0$, population increases.

Combining the dynamics of the four quadrants, we find one stable saddle path, as depicted in the figure, with population below steady state characterized by higher dual variable value $\zeta_{N,t}$, and higher fertility rates.

Note that the steady state (using two asterisks for the case $\gamma_N, \gamma_m > 0$), has $f^{**} = \delta_N$. This fixes x^{**} through (98), and h^{**} through (5). Then through (43) we have ζ_N^{**}, and then through the above we have N^{**}.

Now consider the locus for $\Delta\zeta_{N,t} = 0$ for $u(.) = 0$, which is a horizontal line (102). Assume it lies above the steady state (N^{**}, ζ_N^{**}), which means that $\zeta_N^* > \zeta_N^{**}$ so that $f^* > f^{**} = \delta_N$. From (102) we also see that $\gamma_N < \gamma_m N^{**}$, thus $N^{**} > \gamma_N/\gamma_m$. That is, $f^* > \delta_N \Leftrightarrow N^{**} > \gamma_N/\gamma_m$. Q.E.D.

Proof of Lemma 1 (Separable log-linear welfare).

Proof. Recall that $\partial V_t/\partial K_t = \zeta_K/K_t$, as defined at the start of the proof for Prop 1. The constancy of ζ_K independent of the level of K_t implies that V_t is separable and logarithmic in K_t: $V_t = \zeta_K \ln(K_t)$ plus something for other stocks. Thus the logarithmic part of valuation follows from the observation of constant ζ_K (83), ζ_A (84), and for $u(.) = 0$, constant ζ_N (91), and ζ_h (93). Note that the lemma also implies that optimal emissions E_t^* do not depend on capital and knowledge, and for $u(.) = 0$ also not on population and human capital stock. Emissions thus depend on the fossil fuel resource that is, for $u(.) = 0$: $E_t^* = E_t^*(R_t)$, while for $\gamma_N, \gamma_m > 0$, $E_t^* = E_t^*(N_t, h_t, R_t)$. Indeed, a similar feature has been identified in Gerlagh and Liski (2018) at the end of their proof of Lemma A.1.

Finally, we can construct the Θ_i through the marginal effects of E_{t-i} on welfare, directly through output Y_t, indirectly through its effect on A_{t+1}:

$$\Theta_i = [(\delta_Y \zeta_Y + \delta_A \beta \zeta_A)/(1-\beta)]\,\theta_i \tag{107}$$

Q.E.D.

Proof of Corollary 1 (Carbon taxes in second-best).

Proof. First, we note that the corollary can be derived from carefully re-reading the proof of Prop 1. That proof is divided in 4 blocks. As discussed in that proof, when we derive

(95) and subsequently (42), the SCC formula does not depend on the optimality of the other policy choice variables.

Yet, we can also see the corollary emanating from lemma 1. When we solve for the social optimum recursively,

$$\max V_t = \ln(C_t/N_t) + \gamma \ln(f_t) + \beta V_{t+1}, \tag{108}$$

the first order conditions for emissions E_t, when maximizing $u_t + \beta V_{t+1}$ are independent of capital savings and innovation share, fertility and schooling. Thus the social optimum SCC conditions hold independently. Q.E.D.

Proof of Corollary 2 (Population and welfare).

Proof. From the Lagrangean, set up at the start of the proof of Prop 1 above, it is obvious that $dW_t/dN_t = \zeta_N/N_t$. Substituting (89) in (91) we get (48):

$$(1-\beta)\zeta_N = \frac{(1-\alpha-\kappa)(\varepsilon-1)(1-\beta\varphi) + \beta(1-\psi)(1-\alpha-\kappa)}{(1-\alpha\beta)(\varepsilon-1)(1-\beta\varphi) - \alpha\beta^2(1-\psi)} - 1 \tag{109}$$

Q.E.D.

Proof of Proposition 2 (Decentralization of SO).

Proof. From $q_t = \lambda_{R,t}/\lambda_{Y,t}$, (88) and (94), we get $q_{t+1} = \sigma_{k,t+1} r_{t+1} q_t$ so that there is no need for taxes on exhaustible resource holdings.

Consider the Lagrangean for the representative household:

$$\mathcal{L}^H = \sum_{t=0}^{\infty} \beta^t \left[\ln(c_t/n_t) + \gamma \ln(f_t)\right] \tag{110}$$

$$+ \sum_{t=0}^{\infty} \beta^t \lambda_{c,t}^H [\sigma_{l,t} w_t h_t (1 - \phi f_t - x_t f_t) n_t + \sigma_{k,t} r_t s_t - \tau_{f,t} f_t n_t + \tau_{n,t} n_t - c_t - s_{t+1}] \tag{111}$$

$$+ \sum_{t=0}^{\infty} \beta^{t+1} \lambda_{n,t+1}^H \left[(1 - \delta_N + f_t) n_t - n_{t+1}\right] \tag{112}$$

$$+ \sum_{t=0}^{\infty} \beta^{t+1} \lambda_{h,t+1}^H \left[x_t^{\eta_s} h_t^{\eta_h} - h_{t+1}\right] \tag{113}$$

where the last two equations are the same as (57′), (58′), with lower case variables n_t to denote that we consider household's own variables. In equilibrium, all lower-case and capital variables are identical, $C_t = c_t, N_t = n_t$, etc. The dual variables have superscript H, to separate those from the dual variables in the aggregate Lagrangean (61)-(58′).

For later reference, as for the planner's program, here we also use the recursive formulation, $v_t(s_t, n_t, h_t) = u(c_t/n_t, f_t) + \beta v_{t+1}(s_{t+1}, n_{t+1}, h_{t+1})$, and the semi-elasticities $\zeta^H_{s,t} = (\partial v_t/\partial s_t)s_t$, $\zeta^H_{n,t} = (\partial v_t/\partial n_t)n_t$, $\zeta^H_{h,t} = (\partial v_t/\partial h_t)h_t$. An important feature is that, due to the linear budget constraint, for an individual dynasty, welfare is linear in the number of people and savings:

$$\zeta^H_{s,t} + \zeta^H_{n,t} = 0. \tag{114}$$

We compare the above Lagrangean with the SO Lagrangean (61)-(58). We first take the FOC for c_t

$$\lambda^H_{c,t} = 1/c_t, \tag{115}$$

evaluated at the SO, so that we immediately see that the two optimization problems have the same dual variables for the final good: $\lambda^H_{c,t} = \lambda_{Y,t} = 1/C_t$.

We then consider the FOC for s_{t+1},

$$\lambda^H_{c,t} = \beta\sigma_k\lambda^H_{c,t+1}r_{t+1}, \tag{116}$$

evaluated at the SO. Substituting the SO properties $\lambda_{c,t}Y_t = \lambda_{c,t+1}Y_{t+1}$, $r_{t+1}K_{t+1} = \alpha((\varepsilon - 1)\zeta_X/\varepsilon\zeta_Y)Y_{t+1}$, and $s_K = \alpha\beta\zeta_X/\zeta_Y$ in the above equation, we conclude that $\sigma^*_k = \varepsilon/(\varepsilon - 1)$ (49). As labour is supplied inelastically, its tax rate has no effects on the allocation. Yet for consistency, to value labour supply by the household at its social marginal value, we must implement $\sigma^*_l = \sigma^*_k = \varepsilon/(\varepsilon - 1)$

Innovators maximize the post-tax value of patents minus costs (15). We use (87) to express both left and RHS in terms of current output:

$$\frac{\sigma_{a,t+1}}{\sigma_{k,t+1}r_{t+1}}\pi_{t+1}A_{t+1} = \sigma_{a,t+1}\frac{Y_{t+1}}{\sigma_{k,t+1}r_{t+1}\varepsilon} = \frac{\beta\sigma_a}{\varepsilon}Y_t \tag{117}$$

$$r_tK_{A,t} + w_th_tL_{A,t} + q_{z,t}Z_{A,t} + (q_{e,t} + \tau_{e,t})E_{A,t} = \frac{s_A}{1 - s_A}\frac{\varepsilon - 1}{\varepsilon}Y_t = \frac{\beta(1 - \psi)}{\varepsilon(1 - \beta\varphi)}Y_t \tag{118}$$

which immediately gives $\sigma_a = (1 - \psi)/(1 - \beta\varphi)$ (50).

Conveniently, the tax rates imply that the representative households' labour income, capital rent income, and patent income equal the value in the Social Optimal analysis, :

$$\lambda^H_{c,t}\sigma_{l,t}w_t h_t L_t = (1-\alpha-\kappa)\zeta_X \tag{119}$$

$$\lambda^H_{c,t}\sigma_{k,t}r_t K_t = \zeta_K = \alpha\zeta_X \tag{120}$$

$$\lambda^H_{c,t}\sigma_{a,t}\pi_t A_t = (1-\psi)\zeta_A \tag{121}$$

Recall that $S_{K,t+1} = K_{t+1}$, and from (27) and (28), we get $\sigma_{k,t+1}r_{t+1}S_{A,t+1} = \sigma_{a,t+1}\pi_{t+1}A_{t+1}$. Together with the above, we have that the value of savings for households, equals the value of capital plus the share of technology produced, plus the value of exhaustible resources:

$$\lambda_{c,t+1}\sigma_{k,t+1}r_{t+1}S_{t+1} = \zeta_K + (1-\psi)\zeta_A + \zeta_{R,t+1}. \tag{122}$$

Considering the FOCs for schooling and human capital we find that the households FOC are equal to those for the planner. There is no need for fiscal policies for education.

To derive the fertility tax, we need to look at the FOC for fertility. Now, the FOC for fertility f_t also looks the same for the household as for the planner (96), apart from the fertility tax term $\tau_{f,t}$:

$$\frac{\gamma}{f_t} + \beta\lambda^H_{n,t+1}n_t = \lambda^H_{c,t}((\phi + x_t)\sigma_l w_t h_t - \tau_{f,t})n_t \tag{123}$$

Along the optimal path, $N_t = n_{i,t}$, and the labour value in (96) equals that in (123), so that

$$\tau^*_{f,t} = \frac{\beta\lambda_{N,t+1} - \beta\lambda^H_{n,t+1}}{\lambda^H_{c,t}} \Rightarrow \tag{124}$$

$$\tau^*_{f,t}N_{t+1} = \frac{\beta\zeta_N - \beta\zeta^H_{n,t+1}}{\lambda^H_{c,t}} \tag{125}$$

$$= \frac{\beta\zeta_{N,t+1} + \beta\zeta^H_{s,t+1}}{\lambda^H_{c,t}} \tag{126}$$

$$= \frac{\beta\zeta_{N,t+1} + \beta\lambda^H_{c,t+1}\sigma_k r_{t+1}S_{t+1}}{\lambda^H_{c,t}} \tag{127}$$

$$= \beta(\zeta_K + (1-\psi)\zeta_A + \zeta_{R,t+1} + \zeta_{N,t+1})C_t \tag{128}$$

where the third line follows from (114), the fourth line is the FOC for s_{t+1}, and the last line is from (122). The above derivation reveals that if part of the produced value of the assets $\zeta_K + (1-\psi)\zeta_A$ is saved publicly rather than privately, the fertility tax is adjusted accordingly.

See end of proof for Prop 1 for $f^* > \delta_N \Leftrightarrow \zeta_N^{**} < \zeta_N^*$. Q.E.D.

B Appendix. Calibration

B.1 Balanced Growth

When calibrating the model on historic growth, we can make two distinct assumptions. One is that energy use is a fixed factor, so that the balanced growth equations for (33),(34), abstracting from climate damages, become

$$(1-\alpha)g_Y = \frac{g_A}{\varepsilon - 1} + (1-\alpha-\kappa)g_L \tag{129}$$

$$(1-\varphi)g_A = (1-\psi)(\alpha g_Y + (1-\alpha-\kappa)g_L) \tag{130}$$

At the other end, we can consider fossil fuel energy use as growing at approximately proportional to total output, so that the share of production factors that grow at the same rate as output becomes $\alpha + \kappa$. This gives a set of balanced growth equations for (33),(34), abstracting from climate damages, to become

$$(1-\alpha-\kappa)g_Y = \frac{g_A}{\varepsilon - 1} + (1-\alpha-\kappa)g_L \tag{131}$$

$$(1-\varphi)g_A = (1-\psi)((\alpha+\kappa)g_Y + (1-\alpha-\kappa)g_L) \tag{132}$$

The first choice gives for income growth as dependent on population growth:

$$\begin{aligned} (1-\alpha)g_Y &= \frac{(1-\psi)(\alpha g_Y + (1-\alpha-\kappa)g_L)}{(\varepsilon-1)(1-\varphi)} + (1-\alpha-\kappa)g_L \\ \Rightarrow g_Y &= \frac{(\varepsilon-1)(1-\varphi)(1-\alpha-\kappa) + (1-\psi)(1-\alpha-\kappa)}{(\varepsilon-1)(1-\varphi)(1-\alpha) - \alpha(1-\psi)} g_L \end{aligned} \tag{133}$$

The second choice gives for income growth as dependent on population growth:

$$\begin{aligned} (1-\alpha-\kappa)g_Y &= \frac{(1-\psi)((\alpha + -\kappa)g_Y + (1-\alpha-\kappa)g_L)}{(\varepsilon-1)(1-\varphi)} + (1-\alpha-\kappa)g_L \\ \Rightarrow g_Y &= \frac{(\varepsilon-1)(1-\varphi)(1-\alpha-\kappa) + (1-\psi)(1-\alpha-\kappa)}{(\varepsilon-1)(1-\varphi)(1-\alpha-\kappa) - (\alpha+\kappa)(1-\psi)} g_L \end{aligned} \tag{134}$$

It is clear that, for given parameters, the second calibration choice results in higher output per capita g_Y/g_L. Our preferred choice is to calibrate parameters on a relation in-between:

$$\begin{aligned} g_Y &= \frac{(\varepsilon-1)(1-\varphi)(1-\alpha)+(1-\psi)(1-\alpha}{(\varepsilon-1)(1-\varphi)(1-\alpha)-\alpha(1-\psi)} g_L \\ \Rightarrow g_Y &= \frac{(\varepsilon-1)(1-\varphi)+(1-\psi)}{(\varepsilon-1)(1-\varphi)-\frac{\alpha}{1-\alpha}(1-\psi)} g_L \\ \Rightarrow g_Y - g_L &= \frac{(1-\psi)}{(1-\alpha)(\varepsilon-1)(1-\varphi)-\alpha(1-\psi)} g_L \end{aligned} \tag{135}$$

Thus, growth increases with standing on shoulders (φ) and also with returns to scale in innovation production ψ.

B.2 Rate of convergence

Log-deviations for output $\breve{y}_t = \ln(Y_t) - \ln(Y_t^*)$ innovations $\breve{a}_t$ and capital $\breve{k}_t$, when we consider (33),(34), abstracting from climate damages, become

$$\breve{k}_{t+1} = \breve{y}_t = \frac{\breve{a}_t}{\varepsilon-1} + \alpha\breve{k}_t \tag{136}$$

$$\breve{a}_{t+1} = \varphi\breve{a}_t + \alpha(1-\psi)\breve{k}_t \tag{137}$$

We solve the two equations for possible eigenvalues λ, by imposing that $\breve{k}_t = \lambda^t$ and $\breve{a}_t = a \cdot \lambda^t$:

$$\lambda = \frac{a}{\varepsilon-1} + \alpha, \tag{138}$$

$$a\lambda = \varphi a + \alpha(1-\psi). \tag{139}$$

Substitution of the first in the second gives a quadratic polynomial in λ:

$$\lambda^2 - (\alpha+\varphi)\lambda + \alpha(\varphi - \frac{1-\psi}{\varepsilon-1}) = 0. \tag{140}$$

We search for the dominant eigenvalue, that is the one with largest absolute value. As the parabole has negative slope at $\lambda = 0$, it follows that the largest eigenvalue is dominant, thus we have

$$\lambda^* = \frac{1}{2}\left(\alpha+\varphi+\sqrt{(\alpha+\varphi)^2 - 4\alpha\left(\varphi - \frac{1-\psi}{\varepsilon-1}\right)}\right) \tag{141}$$

For our calibration procedure, we target an annual convergence rate of 2 per cent, that is, $\lambda^* = 0.98^{10} = 0.817$.

C Further results

C.1 Dynamics for the core capital-innovation model (62)-(66)

We focus here on model (62)-(66), where the only endogenous variables are capital, technology, output, consumption, and investments in capital and innovation. This model composes the capital-innovation core of the broader model that has added schooling, human capital, fertility, population, emissions and climate change. Thus the graphs below present the core dynamics of the 2-state capital-innovation model without the schooling-fertility-climate response, which are kept constant, unaffected by the initial state and dynamics of capital and technology.

We established for the core model constant dual variables $\zeta_{\cdot,t}$ and savings rate s_K, s_A, which was sufficient for the theory in main text. The transitionary dynamics for capital K_t and technology A_t display the dynamic behavior of the model. We construct two simulation series based on central parameters $\alpha = 0.219$, $\beta = 0.817$, $\varphi = 0.795$, $\psi = 0.765$, consistent with the central macro targets.

The first simulation series focuses on transitionary dynamics as dependent on the initial state of capital and technology. Figure 2 presents the capital and technology stocks, relative to the steady state, where we let initial values vary by plus-minus 10 per cent, keeping all other state variables the same. Each arrow presents one period of about 10 years. The figure shows that capital adjusts relatively fast, while technology is more persistent. Starting with a high level of technology but low level of capital (left upper corner), capital quickly increases and even exceeds the steady state value from the second period onward, after which both state variable converge to their steady state. The attracting (steep) paths from north-east and south-west to the steady state define the long-term convergence rate for this economy, are associated with the largest eigenvalue of the system (141), and are used to calibrate ψ, φ in Appendix B.2.

Figure 3 explicitly considers the dynamic effects of a drop in output associated with costly emissions reductions. Its purpose is to present the long-run amplification of costs and benefits of climate policies, as immediate cost feed into lower investments in capital and innovation, further reducing output beyond the direct costs. The same argument applies to benefits from prevented climate damages, which exceed the direct prevented

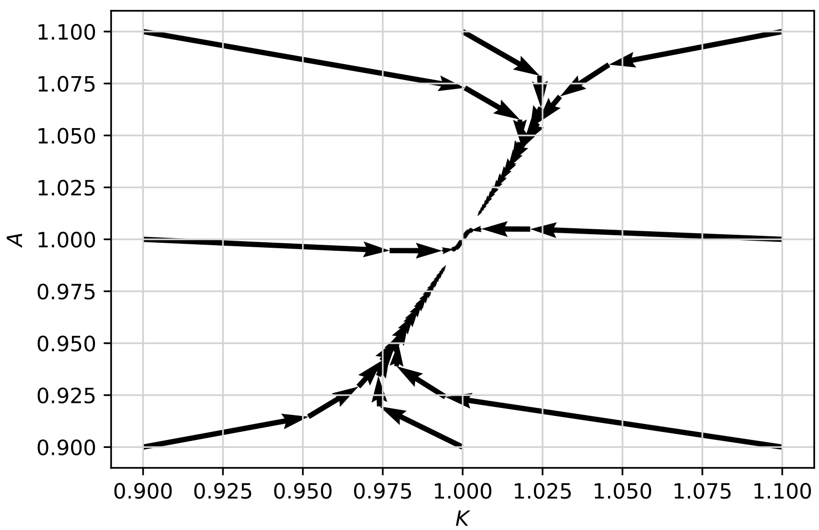

Figure 2: Convergence trajectories for capital and technology

damages substantially. As thought experiment, we start the economy along some reference path, and then assume a decrease in emissions such that output directly decreases by 10 per cent, forever. Technically we reduce $f_t(1, E_t)^\kappa$ by 10 per cent, while abstracting from benefits of reduced climate damages through Ω_t and Γ_t. The graph shows that after one period, capital build up has decreased by 10 per cent, technology has decreased by 5 per cent. Subsequently, reduced capital and technology further depress output the next period. The graph shows that in the long run capital scales with output, and the capital output ratio quickly converges to its steady state value. Technology takes a much longer time to converge. After 10 periods of about 10 years each, the initial drop in output of 10 percent has amplified to a decrease by 16.6 per cent. The level of technology also has decreased by more than 10 per cent, though technology has still not fully converged to its steady state level.

C.2 Permanent climate-change effects

We re-use notation from the section above to study the permanent effects of emissions on output, dropping subscript t. Permanent log-deviations for output and capital $\breve{y} = \breve{k}$ and innovations $\breve{a}$, when we consider (33),(34), comparing two scenarios with different

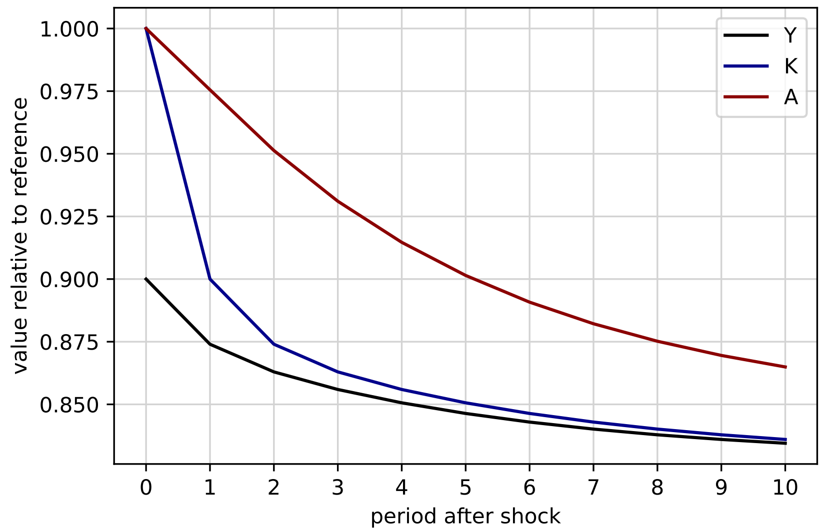

Figure 3: Dynamics for output, capital, and technology, over 10 periods, after a reduced emissions policy shock that directly and permanently reduces output by 10 per cent.

temperature levels $\breve{t} = \ln(T_\infty) - \ln(T_\infty^*)$, become:

$$\breve{k} = \breve{y} = \delta_Y \breve{t} + \frac{\breve{a}}{\varepsilon - 1} + \alpha \breve{k} \tag{142}$$

$$\breve{a} = \delta_A \breve{t} + \varphi \breve{a} + \alpha(1-\psi)\breve{k} \tag{143}$$

which gives

$$(1-\alpha)\breve{y} = -\delta_Y \breve{t} + \frac{\breve{a}}{\varepsilon - 1} \tag{144}$$

$$(1-\varphi)\breve{a} = -(\varepsilon-1)\delta_A \breve{t} + \alpha(1-\psi)\breve{y} \tag{145}$$

and then

$$\breve{y} = -\frac{\delta_Y + \delta_A}{(1-\alpha) - \frac{\alpha(1-\psi)}{(\varepsilon-1)(1-\varphi)}} \breve{t} \tag{146}$$

For $\lim_{i\to\infty} \theta_i = \theta_\infty > 0$, we have that $d\breve{t}_\infty / dE_t = \theta_\infty$, so that the long-term income effects of current emissions become

$$\breve{y} = -\frac{\delta_Y + \delta_A}{(1-\alpha) - \frac{\alpha(1-\psi)}{(\varepsilon-1)(1-\varphi)}} \theta_\infty \tag{147}$$

C.3 Variation in SCC estimates

Here we present a figure underlying the SCC results presented in Table 2. The vertical and horizontal black lines indicate the 5,50,95 percentiles.

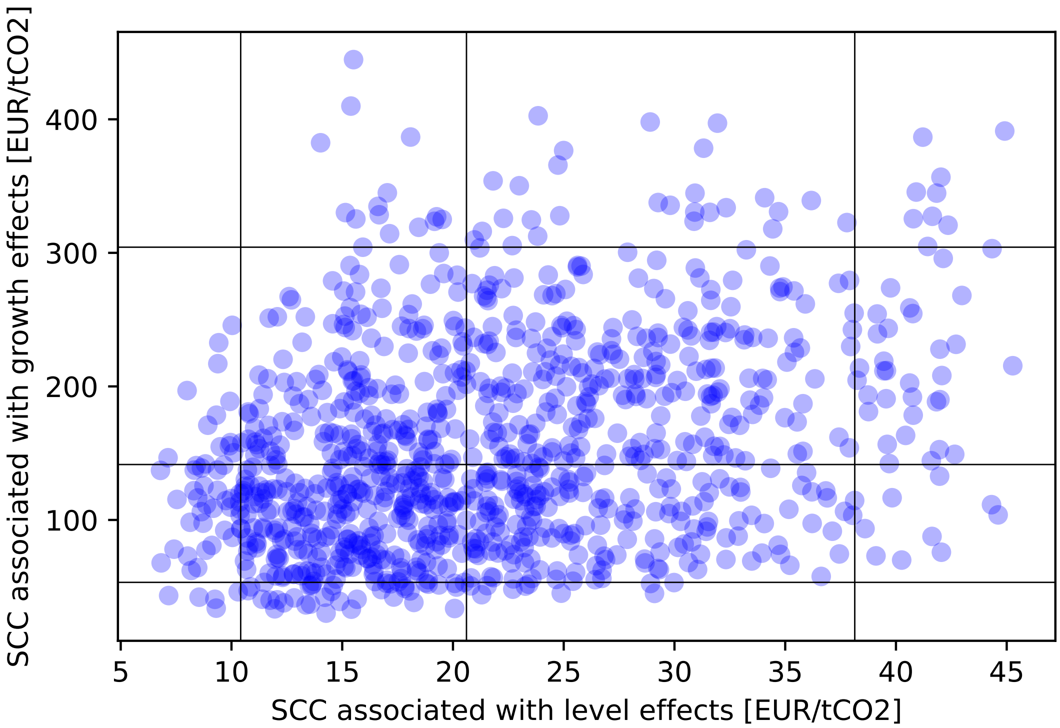

Figure 4: SCC estimates, decomposed in level and growth effects, for Monte Carlo set of parameters. (sample size=1000)