Escaping the Climate Trap? Values, Technologies, and Politics*

Tim Besley[†] and Torsten Persson[‡]

November 2020

Abstract

It is widely acknowledged that reducing the emissions of greenhouse gases is almost impossible without radical changes in consumption and production patterns. This paper examines the interdependent roles of changing environmental values, changing technologies, and the politics of environmental policy, in creating sustainable societal change. Complementarities that emerge naturally in our framework may generate a "climate trap," where society does not transit towards lifestyles and technologies that are more friendly to the environment. We discuss a variety of forces that make the climate trap more or less avoidable, including lobbying by firms, private politics, motivated scientists, and (endogenous) subsidies to green innovation.

*We are grateful for perceptive comments by Philippe Aghion, David Baron, Xavier Jaravel, Bård Harstad, Elhanan Helpman, Gilat Levy, Linus Mattauch, and Jean Tirole, as well as participants in a Tsinghua University seminar, and LSE and Hong Kong University webinars. We also thank Azhar Hussain for research assistance. Financial support from the ERC and the Swedish Research Council is gratefully acknowledged.

[†]LSE, t.besley@lse.ac.uk.

[‡]IIES, Stockholm University, torsten.persson@iies.su.se

1 Introduction

What will it take to bring about the fourth industrial revolution that may be needed to save the planet? Such a revolution would require major structural changes in production as well as consumption patterns. Firms would have to invest on a large scale in technologies that generate lower greenhouse gas emissions, and households would have to consume goods that produce lower emissions.

Already these observations suggest that the required transformation can be reinforced by a key complementarity, akin to the one associated with so-called platform technologies (Rochet and Tirole 2003). If green technologies became more attractive, people would be more likely to evolve values that promote environmentally friendly lifestyles. And if more people were changing their lifestyle, firms would indeed be more likely to develop those green technologies. This paper formalizes how such complementarities may help drive two-way dynamics between values and technology.

Government intervention is also bound to have a major role in transformative change.[1] But the practical challenges are understated by studying exogenous policy paths, or paths chosen by benevolent social planners who can commit their society to future policy. To understand the conditions under which policy can feasibly lead a green transformation, considerations of incentive compatibility in the face of political objectives and commitment obstacles need to be taken on board. Changing values may also play a crucial role in decarbonization. However, as with policy, values are not exogenous and will evolve within the social system in a way that interacts with policy and technological change.

The paper is a first step towards exploring the economic and political conditions for a two-sided green transition. It identifies the preconditions for a "climate trap" rooted in the joint dynamics of environmentally-friendly values and technologies. One of the model's two building blocks is a population comprised of citizens who identify as environmentalists or materialists, and a socialization process which alters consumption patterns and policy preferences, as in Besley and Persson (2019a). The second is a model of endogenous technological change for green or brown technologies, akin to that in Acemoglu et al. (2012).[2] With these ingredients, we build a dynamic

[1] See OECD (2010) for a discussion of how fiscal instruments, i.e. taxes and subsidies are used.

[2] One key difference is that their model focuses on intermediate goods (energy), while

model where technologies, values, and policy coevolve.

Our baseline model is extremely simple, as the current generation can only express its value-based preferences via consumption choices and voting. However, we enrich this framework to study a broader range of political influence by activists or scientists (as members of civil society) and lobbying by firms. With these wider pathways, we gain better insights into how political change can boost or dampen economic change. This conveys additional insights into the kinds of enablers that help escape from the climate trap and the kind of frictions that deepen it.

The analysis delivers both positive and negative news for sustainability. Changing values can indeed support structural change towards predominantly green technologies, but this outcome is by no means guaranteed. Complementary technologies and values – as mediated by politics – create critical junctures that make the future virtuous or vicious. By studying these divergent dynamics, we formalize how a society can get stuck. In a climate trap, a long-run transition from business as usual to a low-pollution economy is *technologically* possible, but does not materialize due to interacting politics, technology, and values. The paper's punchline is a need to think harder about how a reorganization of politics, particularly enhancing the influence of environmentalists, can generate sustained change to deliver a "big push" to overcome the trap. It also suggests a fascinating research agenda on integrating endogenous policy, technology, and culture in an interdependent world.

The paper is organized as follows. Section 2 explains how our approach links to earlier research in different literatures. We then develop our framework step by step. Section 3 lays out a baseline economic model with static choices in consumption and production. Section 4 brings in dynamic choices in socialization and innovation and shows how two forms of corrective taxation are determined in two-party electoral competition. We then derive the full-equilibrium dynamics for values and technologies. These can converge to either a green or a brown steady state, where the latter may indeed be viewed as a climate trap with lower welfare. Finally, we discuss the comparative dynamics implied by the model.

The baseline model illustrates how the dynamic complementarity mech-

ours focuses on final goods. Another is that they work with a single representative consumer, while we allow for preference dynamics due to evolving consumer values. A third key difference is that they consider exogenous policies, whereas we consider endogenous policymaking.

anism mentioned at the outset might work. But it embodies restrictive assumptions about the actions available to households, firms, and political parties. Section 5 allows for richer behavior of these actors. The extensions all pick up salient points in recent discussions about the climate problem. Thus, we let citizens act upon their values, not just as consumers but also as innovators, and not just in public politics (as voters) but also in private politics (outside the electoral process). We allow firms to act not only in the economic but also in the political sphere (by lobbying politicians). Finally, we equip politicians with additional policy instruments such that they can influence firm-level decisions (and thus, indirectly, policies) not just in the present but also in the future. Section 6 takes stock of the results and concludes. Some analytical details and proofs are relegated to an (Online) Appendix.

2 Antecedents

This paper is related to many different lines of existing research in economics and other social sciences. In the interests of brevity, we sketch these links without attempting an exhaustive literature review.

First, theoretical models of endogenous technological change based on innovation can be classified into one of three main approaches. (1) *New* firms may innovate to produce *new* goods, as in Romer (1990). (2) *New* innovating firms may displace old firms in the production of *existing* goods, as in Aghion and Howitt (1992). (3) *Existing* firms may innovate in the production of *existing* goods, as in Krusell (1998). Our model follows the third approach. Garcia-Macia et al. (2019) discuss empirically these different forms of innovation and find that (3) is indeed the most important source of U.S. technological change.

Second, a growing literature studies the drivers of innovation in green versus brown technologies. These papers draw on theoretical findings by Acemoglu (2002) on directed and endogenous technical change, as well as empirical findings by Popp (2002) which link energy prices to energy-saving investments. Acemoglu et al. (2012) is an early theoretical contribution, with later work by Acemoglu et al. (2016), Aghion et al. (2016), and many others. Unlike that research, which focuses on green and brown technologies to produce intermediate goods (especially energy), we focus on green and brown technologies to produce consumption goods. Moreover, we allow substitution on the consumer side to occur via changing values, on top of the standard mechanism via incentives from relative prices, taxes, and qualities.

Third, our paper is related to research on policies to fight climate change. A classic approach builds on extensions of the neoclassical growth model like Nordhaus and Boyer (2000) and, more recently, Golosov et al. (2014). These papers add a simple carbon-cycle *cum* global-warming bloc, such that emission of greenhouse gases enduringly damages society. They study a social planner, who maximizes an exogenous objective function under full commitment, to derive a sequence of policies, typically for carbon taxes (see Hassler and Krusell 2018 for an overview).

Fourth, our model abstracts from mechanisms whereby greenhouse-gas emissions impose enduring effects on the climate. But it permits a richer analysis by studying optimal and credible policies in a simple model of electoral competition. While stylized, it emphasizes electioneering as does the political-economics literature. It is natural to assume that politicians cannot directly bind their successors – cf. how US President Trump pulled out of the Paris Accord signed by President Obama. Lacking commitment does not rule out strategic policymaking to affect future policy outcomes, as in the so-called "strategic debt" literature (e.g., Persson and Svensson 1989, Tabellini and Alesina 1990, Aghion and Bolton 1990). An extension of our baseline model in Section 5 studies such policymaking.

Fifth, standard approaches to the politics of environmentalism (see Oates and Portney 2003 for a review) are mostly static and thus treat underlying values and preferences as fixed. This approach has also studied how interest groups may engage in lobbying to move policy in their preferred direction. This is the subject in a further extension of our baseline model.

Sixth, yet another strand of research investigates "private politics," where activists pressure firms directly for change outside of the political system. A particularly important application concerns precisely actions against polluting firms and firms involved in fossil-fuel production (see Abito et al. 2019 for a review). Another extension in Section 5 shows how such activities can influence an economy's dynamic path.

Seventh, the paper encompasses ideas from the literature on endogenous cultural change by supposing that values and the concomitant preferences are responsive to developing policies and technologies. While the links between values and pollution taxes were explored already in Besley and Persson

(2019a), this paper combines them with directed technological change.[3] A related model with a second, nationalistic – rather than green – dimension of politics appears in Besley and Persson (2019b).

Eighth, and more generally, our approach to changing values is rooted in an earlier literature on cultural evolution beginning with Boyd and Richerson (1985) and Cavalli-Sforza and Feldman (1981) – Bisin and Verdier (2011) review economic applications. Our paper is related to Nyborg et al. (2006), who study how green consumers emerge in a model of pro-social motivation modeled as self-image.

Ninth, when values change, people alter their economic and policy preferences. Following Akerlof and Kranton (2000), we view this through the lens of identity formation, with environmentalists and materialists forming two identity groups. While this approach has a long history in sociology and social psychology, it has only recently gained currency in economics.[4]

3 The Economic Model

This section formulates our baseline model of consumption and production in each time period. This model has two sets of monopolistically competitive firms: one producing varieties of *brown* (polluting) goods, another producing varieties of *green* (non-polluting) goods. Consumers are of two types with different preference maps: *environmentalists* consume green goods and *materialists* brown goods.

Goods, consumers, and types Each citizen has an exogenous endowment ε of a numeraire good, the consumption of which is x. The numeraire can be transformed into two kinds of goods. A continuum of green goods is indexed by $i \in [0,1]$ with the quantity, quality, price, and the (producer) tax rate on green variety i being denoted by $\{y(i), q(i), p(i), t(i)\}$. Similarly, a continuum of brown goods is indexed by $j \in [0,1]$, with a corresponding quadruplet of variables $\{Y(j), Q(j), P(j), T(j)\}$.

[3] Mattauch et al. (2018) also consider policy implications with a form of endogenous values.

[4] See Bowles (1998) for a general discussion of preference change in economic models. Persson and Tabellini (2020) draw on lessons from several existing literatures when surveying research on the coevolution of values and institutions.

We focus on a case with symmetry within sectors where all brown and all green firms take the same actions and face the same tax rates, T and t respectively.[5]

One generation is economically active at each date and, in any period, the unit mass of citizens is divided into environmentalists, $\tau = e$, and materialists, $\tau = m$. Each type has the same income R (specified below). The two consumer types vary according to their consumption values, with environmentalists (materialists) only valuing green (brown) goods.[6] Let $\mu \in \left[\underline{\mu}, \overline{\mu}\right]$ denote the share of environmentalists, where $\underline{\mu}$ is a lower bound and $\overline{\mu}$ an upper bound with $\overline{\mu} > 1/2 > \underline{\mu}$. These bounds capture historical-cultural forces outside of the model. As we stress in the dynamic analysis, the environmentalist share changes over time, as new generations are socialized.

Demand patterns in society will change over time as values change rather than due to substitution between goods induced by price, tax and quality incentives for a fixed set of preferences. The discussion and data in Besley and Persson (2019b) suggest that environmental values do indeed vary across countries and age cohorts.

Materialists Materialists have the following preferences

$$U = x + \frac{1}{1-\sigma} \int_0^1 Q(j)^{\sigma} Y(j)^{1-\sigma} dj - \lambda \int_0^1 \bar{Y}(j)\, dj$$

with $\sigma < 1$. In this expression, $\lambda > 0$ represents pollution damage where the bar above the variable denotes the average population value. A single materialist consumer cannot affect the aggregate pollution level with his own behavior, and thus ignores the effect of his consumption on $\bar{Y}$. Pollution is therefore a classic externality. Materialists maximize these preferences subject to

$$R \geq x + \int_0^1 P(j)\, Y(j)\, dj + \int_0^1 p(i)\, y(i)\, di. \tag{1}$$

[5] Given the symmetry, readers may wonder why we do not simply study a monopoly firm in each sector. In that alternative model, however, these monopolists would be "large enough" to internalize the economy-wide consequnces of their production and investment decisions. By the monopolistic-competition assumption, we avoid this implausible property.

[6] Obviously, the assumption that each type only values a single type of goods is implausibly strong and made for convenience only. A systematic difference across identity types in their preferences for green and brown goods is all we need to make the same (qualitative) points.

These consumers buy no green goods and their demand for each brown variety is

$$Y = QP^{-\frac{1}{\sigma}}. \tag{2}$$

A higher price-quality ratio thus discourages consumption of the brown good.

Environmentalists Analogously, environmentalists have preferences

$$u = x + \frac{1}{1-\sigma}\int_0^1 q\left(i\right)^{\sigma} y\left(i\right)^{1-\sigma} di - \lambda \int_0^1 \bar{Y}\left(i\right) dj$$

and face the same budget constraint (1).[7] They do not buy brown goods and have demand for each green variety of

$$y = qp^{-\frac{1}{\sigma}}. \tag{3}$$

The demands for brown and green goods in (2) and (3) imply that pollution is $(1-\mu)Y = \bar{Y}$.

Firms To model innovation incentives, we suppose each green and brown variety is produced by a monopolist at a constant marginal cost χ. Firms care only about their own profit and are infinitely-lived, run by successive generations of managers who maximize long-run profits in each period. They are owned by consumers to whom profits are distributed.

Taking the demand function, marginal cost, and producer tax rate into account, profits for a typical brown-variety firm is

$$(1-\mu)\left[Q^{\sigma}Y^{1-\sigma} - (\chi + T)Y\right].$$

Choosing output – and thus the mark-up price – to maximize profit yields

$$Y = Q\left[\frac{\chi + T}{(1-\sigma)}\right]^{-\frac{1}{\sigma}}.$$

As in standard models, monopoly power makes profit-maximizing firms produce below the social optimum by charging a price above marginal cost (including taxes). Maximized profits per firm are given by

$$\Pi\left(Q, T, \mu\right) = \left(1-\mu\right)\sigma Q K(T), \tag{4}$$

[7] In the baseline model, environmentalists do not experience a higher disutility of pollution than materialists. One of the extensions in Section 5 allows their disutility to be higher.

where $K(T) = [(\chi + T) / (1 - \sigma)]^{1-\frac{1}{\sigma}}$. Profits are scaled by $1 - \mu$ due to a market-size effect: a larger share of materialists makes it more profitable to produce the goods they consume. Higher quality also enhances profits, while a higher brown-goods tax rate reduces them.

An analogous argument yields the green-firm profit function

$$\pi(q, t, \mu) = \mu\sigma q k(t), \tag{5}$$

where $k(t) = [(\chi + t) / (1 - \sigma)]^{1-\frac{1}{\sigma}}$.

Innovation Any existing brown (green) firm can improve the quality of its variety by hiring N (n) inventors/scientists as in Krusell (1998).[8] A fraction Ω of the population can train to become inventors/scientists at psychic cost ω, and contract to work for a firm at the time of training. In effect therefore, ω is the cost of hiring a scientist as long as some eligible non-scientists remain. The formal condition for this is $(1 - \mu) \int_0^1 N dj + \mu \int_0^1 n di < \Omega$, which we assume holds throughout.

By recruiting scientists, the firm raises its (next-period) product quality to

$$q\left[1 + \left(\frac{n}{q}\right)^{\varphi}\right] \text{ and } Q\left[1 + \left(\frac{N}{Q}\right)^{\varphi}\right].$$

Since $\varphi < 1$, inventive activity has decreasing returns. We study optimal innovation in Section 4.

Public finance In studying Pigouvian taxation, we abstract from redistributive issues and assume that all tax proceeds are paid back to consumers on a per-capita basis.[9] The government budget constraint is

$$T(1 - \mu) \int_0^1 Y dj + t\mu \int_0^1 y di = D,$$

where D is a per-capita "demogrant" which adds equally to each consumer's budget.

[8] In his model, unlike this one, inventors work on improving intermediate goods that serve as inputs to produce (a single form of) final goods.

[9] Although events such as the recent Gilet Jaunes demonstrations in France suggest that this conventional assumption may not be innoccous.

Economic payoffs The indirect utilities when firms maximize profits are:

$$V\left(T,t,q,Q\right)=R\left(T,t,\mu\right)+\frac{Q^{\sigma}Y^{1-\sigma}}{1-\sigma}-PY-\lambda\left(1-\mu\right)Y \tag{6}$$

for a materialist and

$$v\left(T,t,q,Q\right)=R\left(T,t,\mu\right)+\frac{q^{\sigma}y^{1-\sigma}}{1-\sigma}-py-\lambda\left(1-\mu\right)Y \tag{7}$$

for an environmentalist. Here,

$$R\left(T,t,\mu\right)\equiv\varepsilon+\Pi(T)+\pi(t)-\omega\left[\mu n+\left(1-\mu\right)N\right]+D \tag{8}$$

is income per capita: the exogenous endowment of the numeraire, per-capita profits from brown and green firms net of the cost of recruiting (training) scientists, and the government demogrant.

4 Dynamics

We now develop dynamics where the share of environmentalists and the qualities of both types of goods coevolve, as a result of households' socialization decisions and firms' innovation decisions. On top of this, we endogenize policy: two parties compete for power by setting producer taxes on green and brown goods in sequential elections.

Timing Time is infinite, discrete, and indexed by s. A time period in our framework simultaneously captures: the electoral cycle, the innovation horizon, and the generational gap in socialization. While a more realistic model would treat these alternative horizons in a more nuanced way, we carry on with a "sequential-generation model" without further apology, as it buys a great deal of analytical simplicity. When there is no risk of confusion, we use a short-hand notation for adjacent periods: z for z_s and z' for z_{s+1}.

Each period has five stages:

1. Society starts out with a stock of consumers/voters, and three state variables $\{q, Q, \mu\}$ – i.e., initial quality levels and values (the environmentalist share).

2. Parties announce electoral platforms. Tax rates $\{t, T\}$ are determined by the election outcome, which is subject to idiosyncratic and aggregate party shocks.

3. Current indirect utilities $\{v, V\}$ are determined by firms' outputs and prices and consumers' demand.

4. Next-period's qualities $\{Q', q'\}$ are determined by firms' innovation decisions.

5. Next-period's share of environmentalists, μ', is determined by a replacement socialization process, which is subject to family-specific shocks.

Equilibrium outcomes at stage 3 (for given taxes) were discussed in the previous section. We now embed this in the full model, working in reverse order. In the next two subsections, we thus study stages 5 and 4, taking (the sequence of) tax rates as given. In the following subsection, we close the model by exploring stage 2, where tax rates are determined in political equilibrium.

4.1 Socialization (stage 5)

The population turns over at the end of each period with the new generation of citizens being a driver of change.

New agents are socialized once and for all according to a "Darwinian" driver of values: the expected payoff from identifying with one type rather than the other. For this to be viable, those involved in socialization at s – physical parents, cultural parents, the young individuals themselves – must be able to assess the (hypothetical) well-being of adopting each identity. Specifically, denote the gain from being an environmentalist rather than a materialist at $s + 1$ by Δ'. The Darwinian property says that if $\Delta' > 0$, then the environmentalist share among the newly socialized, μ', increases relative to μ, more so when environmentalists are expected to thrive.

Another driver of socialization is social mixing. If parents have a unified view about environmentalism or materialism, this is more likely inherited by their children. Similarly, if cultural parents are involved, non-homogenous matching – and thus social mixing – of groups is important.

For concreteness, we base the rest of the analysis on a specific micro-founded sequential-generation model, where parents make consumption decisions and vote on behalf of their families before socializing their children at the end of period s.[10]

In that setting – see Section A.1 in the Appendix – the environmentalist share evolves according to:

$$\mu' = \mu + \varkappa 2\mu (1-\mu) \left[F(\beta \Delta') - \frac{1}{2} \right]. \tag{9}$$

Here, $F(\cdot)$ is the cumulative distribution function of a family-specific cultural-fitness shock, which is symmetric around a zero mean with density $f(\cdot)$. As a result, $F(\cdot)$ increases smoothly in Δ with $F(0) = 1/2$. On the right-hand side, $\varkappa$ reflects the extent of social mixing, exposing individuals to different ideas from their parents. This parameter helps govern the speed of cultural dynamics with a higher $\varkappa$ being associated with faster change.

But we can micro-found a similar equation in other ways. Suppose each new generation is not just socialized by their own parents, but by other mentors in society as well. Then, social-mixing parameter $\varkappa$ will also capture the rates at which different families meet across identities (see Besley and Persson 2020 for such an example in another context).

4.2 Investments in Innovation (stage 4)

To study innovation, let $\{\mathbf{t}, \mathbf{T}, \boldsymbol{\mu}\}$ denote future taxes and values from date s onwards. For the moment, we treat these as fixed. Firms invest in innovation to maximize the discounted sum of profits, using a discount factor denoted by β.

We can write the value functions associated with this problem as

$$\tilde{\pi}(q, \mathbf{t}, \boldsymbol{\mu}) = \begin{array}{c} \arg\max_{n \geq 0} \{\pi(q, t, \mu) - \omega n + \\ \beta \tilde{\pi} \left(q \left(1 + \left(\frac{n}{q} \right)^{\varphi} \right), \mathbf{t}', \boldsymbol{\mu}' \right) \} \end{array}$$

$$\tilde{\Pi}(Q, \mathbf{T}, \boldsymbol{\mu}) = \begin{array}{c} \arg\max_{N \geq 0} \{\Pi(Q, T, \mu)(1-\mu) - \omega N + \\ \beta \tilde{\pi} \left(Q \left(1 + \left(\frac{N}{Q} \right)^{\varphi} \right), \mathbf{T}', \boldsymbol{\mu}' \right) \}. \end{array}$$

[10] This kind of socialization model follows Bisin and Verdier (2001), Tabellini (2008), and Besley (2017).

All firms can bid to employ scientists at wage ω in a competitive market and do so to maximize their expected discounted profits, given (rationally) expected policies. The Euler equations associated with optimal innovation are

$$\left(\frac{N}{Q}\right)^{\varphi-1}\beta\varphi\sigma K(T')\left(1-\mu'\right)=\omega, \tag{10}$$

for a typical brown-variety firm, and

$$\left(\frac{n}{q}\right)^{\varphi-1}\beta\varphi\sigma k(t')\mu'=\omega, \tag{11}$$

for a typical green-variety firm. Firms thus hire scientists until their expected marginal gain in future profits equals their marginal cost. In our simple baseline model, only $\{t', T', \mu'\}$ and no current variables shape optimal investment decisions.[11] Moreover, each firm takes these expected future variables as given, since no individual firm is large enough to influence policy (or values) on its own.

Equilibrium structural change In view of (10) and (11), we can express the equilibrium growth rates of green and brown product qualities as follows:

$$\begin{aligned}\widetilde{g}\left(t',\mu'\right) &= \left[\frac{\beta\varphi\sigma k(t')\mu'}{\omega}\right]^{\frac{\varphi}{1-\varphi}} &(12)\\ \widetilde{G}\left(T',\mu'\right) &= \left[\frac{\beta\varphi\sigma K(T')\left(1-\mu'\right)}{\omega}\right]^{\frac{\varphi}{1-\varphi}}. &(13)\end{aligned}$$

These expressions incorporate the effects on innovation incentives of (expected) taxes, with each growth rate decreasing in its sectoral tax rate. Moreover, green (brown) quality growth is increasing (decreasing) in share of green consumers, due to the market-size effect.

In economic equilibrium, product quality in the two sectors therefore evolve according to

$$\begin{aligned}\widetilde{Q}\left(\mu',T',Q\right) &= Q(1+\widetilde{G}\left(T',\mu'\right))\\ \widetilde{q}\left(\mu',t',q\right) &= q(1+\widetilde{g}\left(t',\mu'\right)).\end{aligned}$$

[11] This will no longer be the case in the extension with innovation subsidies in Section 5 below.

4.3 Politics and Taxes (stage 2)

We know from Acemoglu et al. (2012), and subsequent work along the same lines, that brown-goods taxation may change the trajectory of economies with endogenous technological change. However, that research does not consider endogenous taxes. In this subsection, we explore the politics of environmental taxes in a simple model. Section 5 allows for a wider range of political influences.

As in Besley and Persson (2019a), our baseline model entails two-party competition with probabilistic voting (Lindbeck and Weibull 1987, Persson and Tabellini 2000).[12] We label the two (given) parties $P = A, B$ and assume these are solely motivated by winning elections. Each party proposes a tax platform: $\left\{T^P, t^P\right\}$.

Voters Voters are of two kinds. Swing voters cast their ballots based on proposed policy platforms and loyal voters cast their ballots for one party independent of policy. This distinction follows a long-standing, political-science tradition based on the Michigan voting surveys. To simplify the algebra, the baseline model assumes the same proportion of swing voters among materialists and environmentalists (this is relaxed in Section 5.1).

Expected voter utility Now consider a particular voter and parent of identity type $\tau \in \{e, m\}$ when parties make their policy proposals at stage 2. Her expected utility, including the discounted payoff of her offspring, can be written as

$$\begin{aligned} W(\tau) &= I[e]v(t,T,q,Q) + (1-I[e])V(t,T,q,Q) + \\ &\quad \beta E\{I[e']v(t',T',q(1+\widetilde{g}(t',\mu'))+(1-I[e'])V(t',T',Q(1+\widetilde{G}(T',\mu')))\}, \end{aligned} \tag{14}$$

where $I[e]$, an environmentalism indicator, is 1 if $\tau = e$ and 0 if $\tau = m$ and $E\{\cdot\}$ is the expectations operator reflecting uncertainty (to be resolved at stage 5) about the type of offspring in mixed-identity marriages.[13]

Voters would also like parties to increase their children's' payoffs, on the second line of (14). However, parties have no way of affecting $\{\mu', t', T'\}$,

[12] We pick this particular formulation for pure convenience. As discussed in Besley and Persson (2019a), other political models would yield similar conclusions.

[13] In this calculus, parents consider the (hypothetical) payoffs of their children under alternative identities as either materialists or environmentalists. This is important when we consider welfare issues below.

which determine future payoffs via the second line of (14). They lack the power to commit to future policies. Moreover, current taxes $\{t, T\}$, which they do control, cannot be used to alter $\{\mu', t', T'\}$ indirectly. Thus, parties – as do individual firms and households – take the future path of policy $\{\mu', t', T'\}$ as given seeking to garner votes from swing voters by setting $\{t, T\}$ to increase the voters' current payoff, on the first line of (14).

Equilibrium tax rates Swing voters are subject to idiosyncratic and aggregate shocks and parties maximize their expected payoffs anticipating the distribution of these shocks. Since the analysis is standard, we relegate the details to the Appendix (Section A.2). To study equilibrium policy choices, we look for a Nash equilibrium in platforms and show that the model boils down to each party behaving "as if" it maximizes a short-term Utilitarian social-welfare function (the expression in the first line of (14). Equilibrium policies have the following property (see Section B.1 in the Appendix for the proof).

Proposition 1 *In political equilibrium, both parties choose the same tax rates*:

$$\hat{T} = (1-\sigma)\,\lambda - \sigma\chi \quad \textit{and} \quad \hat{t} = -\sigma\chi.$$

This result reflects the fact that equilibrium taxes play two roles. One is to offset the monopoly distortion from mark-up pricing due to imperfect competition. As a result, the green tax rate is negative and becomes a subsidy. The other is as a Pigouvian tax to correct the non-internalized damage caused by brown-sector pollution. Note that $\chi + \lambda$ is the brown-sector social marginal output cost, while χ is the green-sector social marginal cost. Whenever these costs are constant over time, as we assume here, so are the equilibrium tax rates.

Implications Equilibrium taxation affects firm pricing, equilibrium consumption, and emissions in each sector. Thus it also helps shape profits. In particular, the equilibrium value of variable $k(t)$ that enters (5) becomes

$$k(\hat{t}) = \left[\left(\chi + \hat{t}\right) / \left(1-\sigma\right)\right]^{1-\frac{1}{\sigma}} = \chi^{1-\frac{1}{\sigma}} > \left[\chi / \left(1-\sigma\right)\right]^{1-\frac{1}{\sigma}}. \tag{15}$$

Compared to the no-tax case, profits of green-sector firms are higher due to the subsidy on such goods.

Similarly, we can write variable $K(T)$ that enters (4) as

$$K(\hat{T}) = \left[\left(\chi + \hat{T}\right) / (1-\sigma)\right]^{1-\frac{1}{\sigma}} = [\chi + \lambda]^{1-\frac{1}{\sigma}} < [\chi / (1-\sigma)]^{1-\frac{1}{\sigma}}. \quad (16)$$

Compared to no taxes, taxation at $\hat{T} > \hat{t}$ lowers profits for brown-variety firms. Tax policy thus encourages green consumption and raises the profitability of green goods, but reduces consumption and profitability of brown goods.

4.4 Full Dynamic Equilibrium

We now put together the economic, social, and political dynamics to explore the path taken by society.

Payoffs and growth rates with optimal taxes Combining the result in Proposition 1 with indirect utilities (6) and (7) using (8), we obtain

$$U(Q) = r + \frac{\sigma}{1-\sigma} Q [\chi + \lambda]^{1-\frac{1}{\sigma}} \quad (17)$$

$$\text{and} \quad u(q) = r + \frac{\sigma}{1-\sigma} q \chi^{1-\frac{1}{\sigma}}, \quad (18)$$

as the welfare levels for materialists and environmentalists at equilibrium policy. In each of (17) and (18), $r = \varepsilon - \omega[\mu n + (1-\mu) N]$ reflects the quantity of consumption less the resource costs of innovation. The second term reflects the surplus from consumption of the favored set of varieties, given their equilibrium price.

With politically optimal taxes, the quality growth rates are

$$G(\mu', \lambda) = \left[\frac{\beta\varphi\sigma(\chi+\lambda)^{1-\frac{1}{\sigma}}(1-\mu')}{\omega}\right]^{\frac{\varphi}{1-\varphi}}$$

$$g(\mu') = \left[\frac{\beta\varphi\sigma(\chi)^{1-\frac{1}{\sigma}}\mu'}{\omega}\right]^{\frac{\varphi}{1-\varphi}},$$

implying that equilibrium qualities develop over time as

$$\hat{Q}(\mu', \lambda, Q) = Q[1 + G(\mu', \lambda)], \quad \hat{q}(\mu', q) = q[1 + g(\mu')]. \quad (19)$$

Since $\hat{Q}(\mu', \lambda, Q)$ and $\hat{q}(\mu', q)$ are monotonic in μ, the dynamic equilibrium is recursive in μ.

Value dynamics If the parameters satisfy (see the Appendix):

$$1-2\mu\left(1-\mu\right)\left(1-\nu\right)\beta f\left(\beta\Delta\left(\mu,q,Q\right)\right)\Delta_{\mu}\left(\mu,q,Q\right)>0 \text{ for all } \mu\in\left[\underline{\mu},\overline{\mu}\right], \tag{20}$$

then the value dynamics are well-behaved with convergence to a steady state. From (21), we know that the derivative of environmental fitness is positive, $\Delta_{\mu}>0$.This sign reflects a dynamic complementarity between values and technology based on the market-size effect. Specifically, a rising share of environmentalists boosts green-sector profits, which spurs innovation and quality in that sector. The complementarity is analogous to those that arise in the study of platform technologies.

Using (17), (18), and (19), we obtain the following expression for the equilibrium cultural fitness of environmentalism

$$\begin{aligned}\Delta' &= \Delta\left(\mu',q,Q\right)=u\left(\hat{Q}\left(\mu',\lambda,Q\right)\right)-U\left(\hat{q}\left(\mu',q\right)\right)\\ &= \frac{\sigma}{1-\sigma}\left[\hat{q}\left(\mu',q\right)\chi^{1-\frac{1}{\sigma}}-\hat{Q}\left(\mu',\lambda,Q\right)\left[\chi+\lambda\right]^{1-\frac{1}{\sigma}}\right].\end{aligned} \tag{21}$$

Putting this expression into (9), the equilibrium dynamics of the model become effectively one-dimensional. Combining (9) and (21), we can fully characterize these dynamics by

$$\text{sgn } \Delta'=\text{sgn } \delta(\mu',q/Q,\lambda),$$

where function δ is defined by

$$\delta(\mu',q/Q,\lambda)=\frac{q}{Q}\cdot\frac{1+g\left(\mu'\right)}{1+G\left(\mu',\lambda\right)}-\left[\frac{\chi}{\chi+\lambda}\right]^{\frac{1-\sigma}{\sigma}}. \tag{22}$$

The expected fitness of being an environmentalist rather than a materialist thus depends (i) positively on the current ratio of green-to-brown-goods quality, (ii) positively on the extent of environmentalism, via the relative green-to-brown quality growth rate, and (iii) negatively on the ratio of the (social) marginal costs of green goods and brown goods. The dynamics are fully characterized by the sign of $\delta\left(\mu',q/Q,\lambda\right)$, which determines the sign of Δ'. That, in turn, determines whether the share of people with green lifestyles – and the quality of green relative to brown goods – is growing or shrinking.

Steady states Under the mild assumption that $G(\underline{\mu}, \lambda) \geq g(\underline{\mu})$ (see Section B.2 in the Appendix), the equilibrium dynamics are divergent. They can be summarized as follows:

Proposition 2 *Given variables* $(\mu, q/Q)$ *and parameter* λ *– and condition* (20) *– values converge to an environmentalist steady state where* $\mu = \bar{\mu}$ *if* $\delta(\mu, q/Q, \lambda) > 0$ *and to a materialist steady state where* $\mu = \underline{\mu}$ *if* $\delta(\mu, q/Q, \lambda) < 0$.

These divergent dynamics reflect the aforementioned complementarity. Technical change becomes more and more directed towards whichever group of consumers that grows: environmentalists or materialists. This makes it more attractive to identify with the growing group. And so on.

Note that the divergent dynamics do not hinge on taxes being set in political equilibrium. That is, the basic complementarity between consumer values and producer technologies would apply in a model with exogenous taxes. As Proposition 2 shows, it is still present when parties set taxes (to maximize static welfare).

Proposition 2 also says that we only need to know the initial values of μ, q and Q, plus damage parameter λ, to know whether a society will converge to an environmentalist or materialist steady state. As is clear from the proof of Proposition 2, the sign of $\delta(\mu, q/Q, \lambda)$ implies the sign of $\delta(\mu', q/Q, \lambda)$ defined in (22). Moreover, if $\delta(\mu, q/Q, \lambda) < 0$, then society converges to the maximally materialist steady state with declining environmental values and faster quality improvements for brown than for green goods.

The model's dynamics are illustrated in Figure 1, which has three panels. In panel A, $\delta(\mu, q/Q, \lambda) > 0$ for all μ and society converges to $\bar{\mu}$ regardless of the starting value, while panel B portrays the opposite case where $\delta(\mu, q/Q, \lambda) < 0$ with convergence to $\underline{\mu}$. However, in panel C, there is an interior critical value of μ such that $\delta(\overline{\mu}, q/Q, \lambda) = 0$. Which steady state society converges to now depends on whether μ starts above or below this critical value. We find the last possibility the most interesting one.

Welfare comparisons We have already referred to the brown steady state – with maximal materialism – as a "climate trap". This could be viewed as a purely positive statement, i.e. a case where society makes maximal long-run investments in polluting goods. But the idea of a "trap" may also have a normative ring where welfare is lower in a brown steady state than a green steady state.[14] Investigating whether this may happen requires explicit welfare comparisons.

Since Bentham and Mill, Utilitarians have debated whether some preferences should be favored over others.[15] Some normative discussions led by environmentalists look upon materialist lifestyles as intrinsically less valuable and conclude that welfare must be lower according to such preferences as a form of paternalistic judgement. This observation highlights the difficulty of making standard welfare comparisons based on individual preferences when the composition of values (and hence of preferences) in the population differs across two steady states. Here, we do not take a stance on this issue and follow a standard Utilitarian approach. Thus, we ask whether a climate trap is possible in a normative sense, as evaluated by the sum of all citizens' utilities. Implicitly, this approach relies on neutral interpersonal comparisons between different types. That is, it rests on the same hypothetical comparison as the one that agents make in our model of cultural evolution (cf. Footnote 13).

On that basis, we compare steady-state welfare with $\mu = \bar{\mu}$ and $\mu = \underline{\mu}$, evaluated by the values held in those steady states and find

Proposition 3 *If λ is high enough, welfare in a green steady state at $\bar{\mu}$ is higher than in a brown steady state at $\underline{\mu}$, for any pair of initial qualities $\left\{\overline{q}_0, \underline{q}_0\right\}$ (provided that $\underline{\mu}$ is not too high, or ω not too low).*

The proof can be found in (Section B.3 of) the Appendix. But the intuition is simple enough. The green steady state has higher welfare when pollution damages captured by λ are large enough. Because the externality is internalized by Pigouvian taxes, the utility of materialists falls as λ gets larger while environmentalists become relatively better off. The bracketed qualification is sufficient to rule out a paradox, where the resources going into research in the green steady state are so large as to overturn the intuitive result.

[14] We do not consider the more anlaytically demanding problem of comparing the entire dynamic paths, including the transitions to these steady states.

[15] Mill argued that some activities are "higher pleasures" and should count for more in welfare comparisons. This led him to state that "It is better to be a human being dissatisfied than a pig satisfied; better to be Socrates dissatisfied than a fool satisfied. And if the fool, or the pig, are of a different opinion, it is because they only know their own side of the question. The other party to the comparison knows both sides" Mill (1863, page 9).

Proposition 3 states conditions under which convergence to the brown steady-state is indeed associated with a climate trap (in a normative sense).[16] Moreover, (22) implies that a starting point with q/Q close enough to zero, will always yield convergence to a welfare-inferior steady state even when λ is large. Thus the risk of a climate trap is greater in a society that starts out with a substantial quality advantage for brown goods and a large externality. Because the world has only recently woken up to the climate emergency, this seems like the relevant case to consider. The case for highlighting a possible climate trap would only be reinforced if one (i) added concerns about a stock externality,[17] (ii) introduced uncertainty about future damages,[18] (iii) adopted environmental paternalism, or (iv) took the view that the true value of λ is somehow underestimated.

Implications We end this section by discussing some implications of Proposition 2 (and, for the normative points, Proposition 3) and our baseline model, including how parameter shifts might alter society's dynamic path.

For the production side of the economy, the model predicts a changing pattern on the equilibrium path along with a changing μ. Where environmental values are growing, quality growth more and more favors green goods. Even if initially $q/Q < 1$, green product quality will catch up with – and eventually overtake – brown product quality. The physical consumption of green goods is thus increasing faster than that of brown goods. In the end, the economy converges to a maximal rate of green-goods quality growth $g(\bar{\mu})$ and a minimal growth rate of brown-goods quality $G(\bar{\mu}, \lambda)$.

Pollution costs in any period are given by $\lambda (1 - \mu) Y$. In the baseline model, the pollution path need not be monotonic over time, as Y can rise (by brown-goods quality growth encouraging consumption) even though μ is rising. In the steady state, pollution keeps growing even in the green steady state as long as $\bar{\mu} < 1$ – i.e., not all consumers become environmentalists.

[16]Note that our model does *not* say that convergence to any brown steady state leads to lower welfare. But it does say that this will be the case for high enough λ.

[17]Our framework could be extended to have a stock externality where λ rises in pace with the build up of carbon emissions, as in Acemoglu et al (2012). That extension would only reinforce the finding in Proposition 3 if the starting point has a low value of q/Q. This would be particularly relevant in a multi-country world, where λ depended on global cumulated emissions that each country could only marginally influence.

[18]Weitzman (2009) argues forcefully that that uncertainty about future damages λ should make us focus on cases that can potentially entail catastropic outcomes.

However, we could easily add other features to the model to make brown goods decline or disappear. For example, competition for scarce production factors between green- and brown-goods producers would eventually make brown goods unprofitable or inviable and drive emissions to zero.

Comparative dynamics Propositions 2 and 3 show how values (culture) may sustain a climate trap. The same properties of the model may also bring about path dependence. Consider two societies with identical economic opportunities and technologies. These may diverge just because initial values are different, say, $\mu^H > \mu^L$ with $\delta(\mu^H, \frac{q}{Q}, \lambda) > 0 > \delta(\mu^L, \frac{q}{Q}, \lambda)$. Under this condition, green values and technology would decline from μ^L, but rise from μ^H. Initial values alone can thus make a crucial difference, with one society becoming greener and cleaner, and another – with identical economic fundamentals – becoming browner and dirtier. A shock to environmentalism, e.g., following a natural disaster, could have a similar effect.

Long-term effects of shocks Finally, we discuss the impact of two "MIT-shocks," (i.e., unanticipated shocks to parameters or variables at some time period s, with perfect foresight about these parameter or variables from then on). The first is a hypothetical rise of perceived pollution damages from λ^L to $\lambda^H > \lambda^L$. This rise could reflect science or salience. The former could be stronger (and commonly believed) warnings by climate scientists about the destructive effects of carbon emissions. The latter could be "declarations of a climate emergency" by fellow citizens – through social media, physical demonstrations, or other channels – that the problem of climate change is worse than previously thought.

Our framework predicts that politics would respond to such a shift by a higher tax on brown goods. This, in turn, would reorient technological change from brown to green technologies. If the shift $\lambda^H - \lambda^L$ is large enough, or society close enough to the critical juncture where $\delta(\mu, \frac{q}{Q}, \lambda) = 0$, it could be that $\delta(\mu, \frac{q}{Q}, \lambda^H) > 0 > \delta(\mu, \frac{q}{Q}, \lambda^L)$. This could put the economy on a different dynamic path, where values now evolve to escape the climate trap.

The second MIT-shock is a pro-green technology shift, raising q from q^L to q^H. This could be due to opening up imports of new green technologies developed in other countries (think e.g., cheap solar cells from China). By making it cheaper to identify as an environmentalist, the shift $q^H - q^L$ might take the economy out of the climate trap if $\delta(\mu, \frac{q^H}{Q}, \lambda) > 0 > \delta(\mu, \frac{q^L}{Q}, \lambda)$.

We summarize these observations as

Corollary *Large enough positive exogenous shocks to parameter λ or variables q and μ can make society escape a climate trap and converge to $\overline{\mu}$ rather than $\underline{\mu}$.*

In all cases, the complementarities in our model make the new trajectory sustainable via mutually reinforcing technology, politics and values.

We note that the shocks do *not* have to be permanent for this to happen, as long as values develop far enough that the condition $\delta(\mu, \frac{q}{Q}, \lambda) > 0$ holds, once the shift subsides. Another way to say this is that the model can exhibit hysteresis: temporary shocks can have permanent effects.

5 Enriching the Model

Our baseline model described in Section 3 and 4 makes restrictive assumptions about the actions available to households, firms, and political parties. In this section, we enrich these actions.

A key part is expanding the nature of political influence. First, we allow environmentalists to influence outcomes beyond their effect on consumer demand. Following Besley and Persson (2019a), we make the natural assumption that the salience of environmental outcomes is stronger for members of this group. This will enhance their influence as swing voters. Second, we allow environmentalism to affect the behavior of scientists or engage in influence activities, often referred to as "private politics" (Baron 2003). Third, we consider political activity by firms in the form of lobbying. Considering these extensions, we gain new insights into how political factors can perpetuate or end a climate trap.

Finally, we enrich the policy space by introducing a subsidy to green innovation, such that politicians can affect the welfare not only of current voters but also of their children. This can speed up a transition out of a climate trap, although politics remain central in making this happen.

5.1 Environmental Salience

Part and parcel of being an environmentalist is to view pollution as a salient policy issue. We now extend our baseline model so that environmentalists have a higher weight on pollution in their preference function. This is realistic if environmentalism responds to media coverage and to increasing climate awareness. In standard static models, this would only influence policy outcomes by changing political priorities. In our dynamic model, such responses also give additional impetus to the (political) feedback mechanism as environmentalists become more likely to act as swing voters. This makes a climate trap less likely.

Differential salience across types To capture a different salience of pollution, suppose (see Besley and Persson 2019a) that the period-utility function of environmentalists has an additional weight θ on pollution damages

$$x + \int_0^1 \frac{qy^{1-\sigma}}{1-\sigma} di - (\lambda + \theta) \int_0^1 \bar{Y}(j)\, dj.$$

As we discuss and formally analyze in Besley and Persson (2019b), higher salience in an identity group (like environmentalists) can also reflect a stronger – collective rather than individual – identity, due to a social movement among its members.

With this change, we may go through the same steps as in Sections 3 and 4. Doing so, modifies Proposition 1 to

Proposition 1′ *In political equilibrium, where environmentalists have additional salience of pollution θ and make up a fraction μ of the population, both parties choose the same tax rates*:

$$\hat{T} = (1-\sigma)(\lambda + \mu\theta) - \sigma\chi \quad \textit{and } \hat{t} = -\sigma\chi.$$

The corrective tax on polluting goods now reflects a weighted group average of the damages from pollution. Crucially, $\hat{T}$ depends on μ_s when $\theta > 0$. As a result, the brown-goods tax is no longer constant over time but rises (falls) as more (less) people identify as environmentalists. This strengthens the complementarity driving the dynamics: more environmentalists makes politicians set higher brown-goods taxes, which makes it even more attractive to become an environmentalist.

Formally, Proposition 2 still governs the dynamics, but the functional form of $\delta(\cdot)$ becomes

$$\delta(\mu, q/Q, \lambda) = \frac{q}{Q} \cdot \frac{1 + g(\mu)}{1 + G(\mu, \lambda + \mu\theta)} - \left[\frac{\chi}{\chi + (\lambda + \mu\theta)}\right]^{\frac{1-\sigma}{\sigma}}.$$

Earlier, the positive effect of μ on δ – and hence on the sign of Δ – only reflected an "economic complementarity" via the green vs. brown growth rate $g(\mu)/G(\mu,\lambda)$. But present extension adds a "political complementarity" via the brown tax rate. As a result, society is less likely to end up in a climate trap.

It would be interesting to extend the model to encompass partisan (citizen) candidates, who are either materialists or environmentalists. We would then expect that, when a society is close to close to a critical juncture, a popularity shock that led to a power shift from a materialist to an environmentalist could spell the end of a climate trap.

Differential composition of swing voters Another mechanism through which a stronger environmentalist engagement for greener policies will show up in politics would arise if more environmentalists would switch their vote based on the pollution-tax platform in the electoral campaign. Our baseline model abstracts from this, as environmentalists and materialists have the same share of swing voters.

Assume instead (see Besley and Persson 2019a, Online Appendix) that a share $\epsilon > 1/2$ of environmentalists are swing voters, while only a share $(1-\epsilon)$ of materialists are swing voters. Then, the objective function maximized by politicians no longer has pure population weights, as it becomes biased towards swing voters.[19] As a result, the common brown-goods tax rate becomes $\hat{T} = (1-\sigma)(\lambda + 2\epsilon\mu\theta) - \sigma\chi$. As environmentalists dominate among swing voters, policy responds even more strongly to the environmentalist share than in the previous extension.

Formally, Proposition 2 still governs the dynamics, but we have to change the functional form to

$$\delta(\mu, q/Q, \lambda) = \frac{q}{Q} \cdot \frac{1+g(\mu)}{1+G(\mu, \lambda + 2\epsilon\mu\theta)} - \left[\frac{\chi}{\chi + (\lambda + 2\epsilon\mu\theta)}\right]^{\frac{1-\sigma}{\sigma}} .$$

[19] Formally, the baseline model had a function $\Omega\left(T^A, t^A, T^B, t^B, \mu\right)$ maximized by politicians as derived in Appendix B.1. This function is now replaced by

$$\Gamma\left(t^A, T^A, t^B, T^B, \mu\right) =$$

$$\epsilon\mu\left[v(t^A, T^A, q, Q) - v\left(t^B, T^B, q, Q\right)\right] + (1-\epsilon)(1-\mu)\left[V(t^A, T^A, q, Q) - V\left(t^B, T^B, q, Q\right)\right],$$

The upshot is similar to that with an increased weight on pollution, except that the political complementarity is now stronger and more so the larger is the $\epsilon - 1/2$ wedge.

Implications of environmental salience Growing rhetoric among environmentalists can persuade them to vote more for parties that offer stringent environmental policies. Even as a minority, environmentalists can get disproportionate attention and push up taxes on brown goods, which can help break the climate trap. Of course, things can go the other way if materialists get upset – as when Gilets-Jaunes protests made French President Macron back off from a proposed hike of gasoline taxes. Our model highlights the long-run effects of such phenomena and shows how preference intensity can help shape policy dynamics.

5.2 Motivated Citizens

We now allow environmentalist citizens to act on their identity not only as consumers, but also as scientists/inventors, or as activists in private politics. These actions change investment or production costs for green or brown firms. This channel does not reflect that environmentalists have greater preference intensity, but that their actions alter the interplay between technological and value change.

Motivated scientists Scientists form an important part of civil society. Organizations like the National Academy of Sciences, the Royal Society, or the Royal Swedish Academy of Sciences project views and values that sometimes clash with political authority. For example, scientists like Rachel Carson were the first to alert the world to ugly pollution and climate dynamics.

We now show that the collective action of scientists can work via market incentives if scientists who care about pollution may be more attracted to green sectors.[20] Formally, let scientists be "motivated agents" in the language of Besley and Ghatak (2005). Specifically, let the share of inventors with environmental values coincide with the population share μ. Moreover,

[20] https://www.bloomberg.com/news/articles/2019-08-01/the-oil-industry-s-talent-pipeline-slows-to-a-trickle reports that fossil fuel companies are now having increasing difficulties in attracting new graduates.

environmentalist inventors accept a lower wage $(1-\gamma)\omega$ – with compensating differential $\gamma < 1$ – if they work on green rather than brown innovations.

Firms now offer different training contracts for scientists: green-sector firms are able to offer $w = (1-\gamma)\,\omega$, while the brown-sector wage is $W = \omega$. This cost advantage will show up in the number of scientists hired. While the brown quality growth rate is the same as in the baseline model, the green rate becomes $g\left(\mu/\left(1-\gamma\right)\right) > g\left(\mu\right)$.

Implications of motivated inventors The fitness of environmentalism is now given by

$$\delta\left(\mu, q/Q, \lambda, \gamma\right) = \frac{q}{Q} \cdot \frac{1 + g\left(\mu/\left(1-\gamma\right)\right)}{1 + G\left(\mu, \lambda\right)} - \left[\frac{\chi}{\chi + \lambda}\right]^{\frac{1-\sigma}{\sigma}},$$

higher than in the baseline model. All else equal, the climate trap is less likely. As before, close to a critical juncture, a boost to green growth may push society across this threshold. Scientists exercise political power, not through collective action but indirectly through the market system. However, if scientists created a collective identity as environmentalists, such salience could be a way of increasing γ.

This extension emphasizes how market-based mechanisms can help avoid a climate trap. Environmentalism now occurs on the supply side of firms and reinforces any market-size effect on the demand side.

Private politics Environmentalists frequently engage in private politics (Baron 2003, Abito et al. 2019). Resource companies fear movements like the Rainforest Action Network, who pressure brown firms outside the standard political process.[21] This can involve sit-ins, public product boycotts, or publicity campaigns, where environmentalists threaten brown firms to lower their emissions.[22] Such threats raise costs, by driving firms to invest in additional security measures or in PR-activities to offset negative reputational consequences.

In the simplest possible model, activists push up the marginal cost of all brown firms by $\mu\lambda d$, where $d > 0$ denotes the expected damages imposed per firm. We take such activity as exogenous (but could easily endogenize

[21] See https://www.ran.org/.

[22] Bezin (2015) proposes a model of cultural evolution for environmental preferences based on private contributions to environmental protection.

the incentives to engage in protest). Current protest costs cut current profits and production, and expected protest costs affect investments in innovation via future expected profits.

The equilibrium tax rates are set as in Section 4 and now become

$$\hat{T} = (1-\sigma)\,\lambda - \sigma\,(\chi + \mu\lambda d)).$$

This means a *lower* brown-goods tax than in the baseline model where it was $\hat{T} = (1-\sigma)\,\lambda - \sigma\chi$. The reason is that optimal policy undoes a portion of the higher marginal production cost to prevent it from being passed on to consumers. However, the *after-tax* brown-goods marginal production cost does go up, to $(1-\sigma)(\chi + \lambda(1+\mu d)) > (1-\sigma)(\chi+\lambda)$.

Implications of private politics Private politics reshapes the technology and value dynamics both directly and indirectly. Since brown goods are more expensive to produce, innovation incentives in such goods are weaker. The indirect effect alters the value dynamics, as we now have

$$\delta\,(\mu, q/Q, \lambda(1+\mu d)) = \frac{q}{Q}\frac{1+g\,(\mu)}{1+G\,(\mu, \lambda\,(1+\mu d))} - \left[\frac{\chi}{\chi+\lambda\,(1+\mu d)}\right]^{\frac{1-\sigma}{\sigma}}.$$

On a path where μ is rising (falling), the marginal damage cost is increasing (decreasing) over time. A larger group of environmentalists increases the pressure on firms that produce brown goods and that magnifies the feedback effects that we have studied. But the qualitative picture is unchanged, meaning that the sign of $\delta\,(\mu, q/Q, \lambda\,(1+\mu d))$ fully shapes the dynamics.

We have not allowed activism to directly enter the payoffs of environmentalists. If activism is "warm-glow," or negative-reciprocity, it will further boost payoffs of environmentalists and hence further raise $\delta\,(\mu, q/Q, \lambda\,(1+\mu d))$. If activism is costly, the fitness of environmentalism will still rise, as long as the pressure costs do not outweigh the lower quality and higher costs of brown goods.

A more general point Actions of climate activists are often dismissed as social signalling. Our analysis shows that such actions can have auxiliary static and dynamic effects. It gives a different gloss on the welfare effects of private politics, which are often decried as "distortions." Think about this as an application of the theory of the second best. If the political process does not deliver an optimum – here, due to a lack of commitment – private politics may enhance welfare.

5.3 Lobbying by Firms

Governments are lobbied in many domains, and climate politics is no exception. A vast literature points out how lobbying favors organized firms at the expense of consumers. In our framework, lobbying may also have dynamic implications. To illustrate these points, we extend the baseline model whereby firms in both sectors can pay campaign contributions to political parties with policies favorable to their profits. We follow the approach of Baron (1994), where opportunistic parties choose policy platforms partly to please prospective contributors who can help them win elections. This captures the intuitive policy biases from one group of firms (green or brown) being better organized than another.

Basics As before, we study sector-wide taxes $\{t, T\}$. A closed interval $[0, \phi]$ of green-sector firms belong to a coalition that lobby political parties. Specifically, each participating firm pays a campaign contribution c^P to party P at cost $\frac{1}{2}(c^P)^2$. In the same way, an interval $[0, \Phi]$ of brown-sector firms make contributions C^P at cost $\frac{1}{2}(C^P)^2$. The contributions raise party P's probability of winning, in proportion to parameter ξ. Firm coalitions decide on contributions after parties have designed their policy platforms, but before the election.

Equilibrium policy In one intermediate step, we derive the optimal contributions of each firm. In another, we derive the common maximand of the two political parties, which augments the earlier Utilitarian objective by a weighted average of profits in the two sectors. After these steps (see Section A.4 in the Appendix), we can state the main result in this subsection

Proposition 4 *In political equilibrium, with lobbying by organized firms, both parties choose the same tax rates*

$$\hat{T} = \frac{(1-\sigma)\lambda - \sigma\chi(1+\Phi\xi(1-\sigma))}{1+\Phi\xi(1-\sigma)\sigma} \text{ and } \hat{t} = -\sigma\chi\frac{(1+\xi\phi(1-\sigma))}{(1+\xi\phi(1-\sigma)\sigma)}.$$

The two expressions may look complex at a first glance, but they squarely encompass the distortions of lobbying. Specifically, the taxes coincide with those in Proposition 1, when either $\xi = 0$ – money is ineffective in politics – or $\Phi = \phi = 0$ – no firms are organized to lobby. As ξ increases, the subsidy on green goods rises and the tax on brown goods falls. However, this strikes differently across green and brown sectors if Φ and ϕ differ – i.e., lobbying organization is asymmetric. If brown firms are more established and organized ($\Phi > \phi$), then this lowers T relative to t.

Implications of lobbying To see the impact of lobbying, note that the ratio of social marginal costs $\chi / (\chi + \lambda)$ in the equilibrium policies gets adjusted to

$$\frac{1 + \Phi \xi (1 - \sigma) \sigma}{1 + \phi \xi (1 - \sigma) \sigma} \cdot \frac{\chi}{\chi + \lambda}.$$

This adjusted marginal-cost ratio reveals that a larger coalition Φ makes policy more advantageous to brown firms and thus encourages their output and innovation. On the contrary, a higher ϕ encourages greater green-firm output and innovation. If $\Phi > \phi$, then the condition for escaping the climate trap, $\delta (\mu, q/Q, \lambda) > 0$, is *less* likely to hold. Hence, organized lobbying by brown-sector firms raises the likelihood of a climate trap, while lobbying by green-sector firms has the opposite effect.

The net effect of lobbying thus depends on which sector is better organized. If brown firms have an edge, this produces policy inertia and makes a climate trap more likely. Similar reasoning invokes cynicism about policymaking in the US, or the EU, where established brown firms have huge lobbying operations in Washington DC, and in Brussels. In sum, our dynamic model shows how lobbying may contribute to a long-run climate trap, something static models of interest groups do not pick up.

5.4 Innovation Subsidies

In the baseline model, equilibrium taxes maximize the current payoff for parents, but do not take payoffs for children into account. This reflects lack of commitment in the political process. In this subsection, we allow the government to subsidize innovation in green goods. This opens the door to "strategic" policy making where current policy affects future outcomes.[23]

Many governments pursue such policies (see OECD, 2010, for an overview). We already know from Acemoglu et al. (2012) that a temporary innovation subsidy can move the economy to a new trajectory by crossing a tipping

[23] Similar issues would arrise if we considered direct public investments in R&D.

point. But we do not know whether offering such a subsidy is consistent with political equilibrium.

Policy objective Let b denote an ad valorem subsidy to hiring green-sector scientists. The per-scientist cost thus becomes $\omega(1-b)$ and the total cost $\omega b \mu n$. Given separable production across periods (and quasi-linear utility), we can write the intertemporal policy objective as

$$\sum_{s=0}^{\infty} \beta^s \widehat{W}(q_s, Q_s, \mu_s, b_s),$$

where

$$\widehat{W}(q_s, Q_s, \mu_s, b_s) \quad = \mu_s u(q_s, Q_s) + (1-\mu_s) U(q_s, Q_s) - \omega b \mu_s n_s$$

and $U(q_s, Q_s)$ and $u(q_s, Q_s)$ are defined in (17) and (18). This is Utilitarian welfare net of subsidy costs when both tax rates are optimally chosen in each period (taxes do not vary with the investment subsidy).

As the subsidy will lower the per-scientist cost for green firms from ω to $\omega(1-b)$, green-good quality growth becomes

$$g(\mu'/(1-b)) = \left[\frac{\varphi \sigma \chi^{1-\frac{1}{\sigma}} \mu'}{\omega(1-b)}\right]^{\frac{\varphi}{1-\varphi}},$$

which is increasing in b. The cost of a subsidy to generate this growth is $b\mu\omega q g(\mu'/1-b)^{\frac{1}{\varphi}}$.[24]

To study the optimal innovation subsidy, write the value function associated with this choice

$$\begin{gathered} w(\mu, Q, q) \equiv \max_{b \geq 0} \{\hat{W}(\mu, Q, q) - b\mu\omega q g(\mu'(b), b)^{\frac{1}{\varphi}} + \\ \beta w(\mu'(b), Q[1 + (G(\mu'(b), \lambda))], q[1 + g(\mu'(b)/(1-b))])\}, \end{gathered} \tag{23}$$

where μ is the evolving state variable. Equation (23) writes μ' as dependent on b through (9). Product qualities also depend on b – through (19). Political parties can thus influence future outcomes via their choice of innovation subsidy. But as their electoral platforms maximize Utilitarian welfare at any date, both parties pursue a consistent objective.

[24] Recall that $qg^{\frac{1}{\varphi}}$ green-sector scientists are needed to generate a growth rate of g.

The equilibrium innovation subsidy To derive the optimal subsidy, we maximize $w(\mu, Q, q)$ defined in (23) with respect to b, restricting our attention to non-negative solutions.[25] Three elements determine the optimal subsidy. The first is how innovation responds to the subsidy, which depends on $1/(1-\varphi)$. The second is how the subsidy "corrects" firms' focus on profits rather than total welfare, which depends on the markup rate $1/(1-\sigma)$. The third element is more novel: green firms that invest in quality raise future environmental values, which encourages green-goods consumption and profits. Individual firms do not internalize this value externality, but political parties do when they decide on subsidies.

Putting the three elements together gives (see section B.5 in the Appendix)

Proposition 5 *The optimal green R&D subsidy is*

$$b = \max\left\{\frac{\rho - \mu' + \kappa'}{\rho}, 0\right\},$$

where $\rho = \frac{1}{1-\sigma} + \mu\varphi$ *and "cultural multiplier"* $\kappa' = \frac{\beta}{\omega}\frac{dw}{d\mu'}\frac{d\mu'}{dn}$.

To understand the result, first ignore the cultural multiplier (set $\kappa' = 0$). Then, a subsidy is justified only when $\rho > \mu'$, which is always true for small enough μ'. With firm markups $1/1-\sigma$ (and $\sigma < 1$), a subsidy offsets firms' underestimation of the societal R&D benefits (they maximize profit rather than welfare). But as environmental values become more wide-spread (μ' goes up), such subsidies become more costly to the current generation.

The cultural multiplier The term in κ' shows how the innovation subsidy can influence values. This happens in two ways, both of which can be seen in (23). The first way runs through the effect of environmental values on welfare, $dw/d\mu$, which has the same sign as Δ'. If the subsidy makes more people green and this group has a cultural-fitness advantage, $\Delta' > 0$, voters collectively would like to subsidize green investment. But in a climate trap, $\Delta' < 0$, this effect makes a subsidy less valuable. The second way the subsidy shapes the cultural multiplier runs via expression $w_Q \partial Q'/\partial\mu' + w_q \partial q'/\partial\mu'$

[25] In principle, a government that wanted to promote a brown future could tax green innovation but we ignore this possibility here.

which is ambiguous in sign. This expression captures how a higher μ reduces brown-goods innovation and raises green-goods innovation.

Due to these potentially countervailing effects, we have no clear-cut result for the sign of κ'. The cultural multiplier can thus boost or dampen other effects of the subsidy. But it is a "multiplier" affecting the *pace* of change, not a "driver" affecting the *direction* of change. To guarantee positive green innovation subsidies would require policy to fall into the hands of those who care more than voters about a green future.

Implications of endogenous R&D subsidies An innovation subsidy affects the dynamic path via the growth of green-good quality. The new expression governing the value dynamics becomes

$$\delta\left(\mu, q/Q, \lambda\right) = \frac{q}{Q} \cdot \frac{1 + g\left(\mu/\left(1-b\right)\right)}{1 + G\left(\mu, \lambda\right)} - \left[\frac{\chi}{\chi + \lambda}\right]^{\frac{1-\sigma}{\sigma}}.$$

The larger is $b > 0$, the higher is green growth, which lowers the chance that society is stuck in a climate trap. However, a larger b is more likely when $\Delta' > 0$. Hence, endogenously chosen innovation subsidies may decrease the likelihood of a climate trap with a large and positive cultural multiplier.

If $\Delta' < 0$, the bulk of citizens value brown over green lifestyles and politics follow that preference. That said, an interesting possibility is a polity close to a tipping point where an innovation subsidy implies $\Delta' > 0$ even though $\Delta < 0$. This instrument thus allows policy-makers to reverse the direction of an economy on its way to a brown steady state.

The strategic policy assumes that voters understand that political decisions can internalize the externality from green investment on values. This requires a certain degree of farsightedness and rationality. If voters were more behavioral, green-innovation subsidies may not emerge in electoral politics.

All in all, this subsection illustrates how a policy instrument that is capable of affecting values may not be used to tilt them in a green direction. This is yet another reminder to take incentive compatibility in policymaking into account.

6 Conclusions

The main motivation for this paper is the climate emergency, where certain kinds of production and consumption contribute disproportionately to carbon emissions. We have proposed a model where the coevolution of values, technology, and politics shapes society's dynamic path and long-run outcome. Government policies are endogenous to politics, but subject to a credibility problem: current generations may care about the future and politicians may internalize their caring, but they cannot commit future policy. This limits the capacity of current generations to internalize the effect of future policy on the evolution of values and technologies.

Our baseline model pivots around a dynamic complementarity, which generates a prospective "climate trap". Technologies and values interact non-linearly, producing two alternative dynamic paths – one with ever greener values and economic transformation, another with ever browner values and growing emissions. Around tipping points (critical junctures), small changes can have non-marginal, long-run consequences. The analysis highlights how different features of economics and politics shape the dynamics, including a case when strategically set policies can influence future welfare.

By creating a tractable tool for studying the interplay of politics, technologies, and values, the paper opens up a range of issues for future work. A natural extension would be to add intrinsically motivated environmental entrepreneurs running some green-goods firms.[26] This would promote green structural change in a similar way as the motivated inventors in Section 5.2. In a richer model with private savings and portfolio investments, environmental citizens may also "boycott" investments into brown-goods firms, thus driving apart innovation costs of brown and green firms.[27]

Our modeling constitutes but a first, preliminary step. Three natural extensions, beyond the scope of this paper, would further raise its relevance to climate-change debates. First, as mentioned in Section 4, it would be interesting to include stock-related and time-related aspects of the climate externality. An economy may structurally change towards green consumption, but the transition may not be fast enough to avoid a climate disaster (see Aghion et al., 2012 with exogenous policy). We could also consider investments and actions that operate directly on pollution damages (λ in our model). Second, global climate externalities among interacting policymakers also merit further analysis. Spillovers in both technology and values may affect sustainable paths. For example, one country's promotion of green

[26] See Hart and Zingales (2017) for a discussion of different corporate objectives.

[27] See Gollier and Pouget (2014) and Broccardo et al (2020) for the development of models along these lines.

technologies could generate positive global spillovers. Third, it would be interesting to calibrate the model and hence get a better quantitative sense of the magnitude of the qualitative effects we have identified.

A central contribution of the paper is to bring politics more squarely into the formal study of environmental dynamics and policy. While green R&D subsidies and brown-goods taxes could hypothetically alter a society's trajectory, this will not happen without political implementation. However, the takeaway is *not* that paying attention to equilibrium policymaking is only a further constraint. Indeed, our model suggests that politics itself trigger transformative change via empowered climate activists and scientists, reduced influence of brown lobbies, or a higher weight on environmentalist views in the policy process.

The overall message of the paper is that economists should pay closer attention to changing values, as well as to political incentives, when studying the dynamics of sustainable technologies.

References

[1] Abito, José Miguel, David Besanko, and Daniel Diermeier [2019], *Corporate Reputation and Social Activism: Strategic Interaction, Firm Behavior, and Social Welfare*, New York, NY: Oxford University Press.

[2] Acemoglu, Daron [2002], "Directed Technical Change," *Review of Economic Studies* 69(4), 781-809.

[3] Acemoglu, Daron, Philippe Aghion, Leonardo Bursztyn, and David Hémous [2012], "The Environment and Directed Technical Change," *American Economic Review* 102(1), 131–166.

[4] Acemoglu, Daron, Ufuk Akcigit, Douglas Hanley, and William Kerr [2016], "The Transition to Clean Technology," *Journal of Political Economy* 124(1), 52–104.

[5] Acemoglu Daron, Philippe Aghion, Lint Barrage, and David Hémous [2019], "Climate Change, Directed Innovation, and Energy Transition: The Long-Run Consequences of the Shale-Gas Revolution," unpublished typescript.

[6] Aghion, Philippe and Patrick Bolton [1990], "Government Domestic Debt and the Risk of Default: A Political-Economic Model of the Strategic Role of Debt," in Dornbusch, Rudiger and Mario Draghi (eds.), *Public Debt Management: Theory and Practice*, Cambridge, UK: Cambridge University Press.

[7] Aghion, Philippe, Antoine Dechezleprêtre, David Hémous, Ralf Martin, and John Van Reenen [2016], "Carbon Taxes, Path Dependency and Directed Technical Change: Evidence from the Auto Industry," *Journal of Political Economy* 124(1), 1–51.

[8] Aghion, Philippe, Roland Benabou, Ralf Martin, and Alexandra Roulet [2019], "Environmental Preferences and Technological Choices: Is Market Competition Clean or Dirty?" unpublished manuscript.

[9] Aghion, Philippe and Peter Howitt [1992], "A Model of Growth Through Creative Destruction," *Econometrica* 60(2), 323–351.

[10] Akerlof, George and Rachel Kranton [2000], "Economics and Identity," *Quarterly Journal of Economics* 115(3), 715–753.

[11] Baron, David P. [1994], "Electoral Competition with Informed and Uniformed Voters," *American Political Science Review* 88(1), 33-47.

[12] Baron, David P. [2003], "Private Politics," *Journal of Economics & Management Strategy* 12(1), 31–66.

[13] Besley, Timothy [2017], "Aspirations and the Political Economy of Inequality," The 2016 Hicks Lecture in Economics, *Oxford Economic Papers* 69(1), 1–35.

[14] Besley, Timothy and Maitreesh Ghatak [2005], "Competition and Incentives with Motivated Agents," *American Economic Review* 95(3), 616-636.

[15] Besley, Timothy and Torsten Persson [2019a], "The Dynamics of Environmental Politics and Values," *Journal of the European Economic Association* 17(4), 993–1024.

[16] Besley Timothy and Torsten Persson [2019b], "The Rise of Identity Politics," unpublished manuscript.

[17] Besley Timothy and Torsten Persson [2020], "Organizational Dynamics: Culture, Design and Performance," unpublished manuscript.

[18] Bezin, Emeline [2015], "A Cultural Model of Private Provision and the Environment," *Journal of Environmental Economics and Management* 71(3), 109–124.

[19] Bisin, Alberto and Thierry Verdier [2001], "The Economics of Cultural Transmission and the Dynamics of Preferences," *Journal of Economic Theory* 97(2), 298-319.

[20] Bisin, Alberto and Thierry Verdier [2011], "The Economics of Cultural Transmission and Socialization," Chapter 9 in Benhabib, Jess (ed.), *Handbook of Social Economics*, Volume 1A, Amsterdam, NL and Oxford, UK: North-Holland, Elsevier.

[21] Bowles, Samuel [1998], "Endogenous Preferences: The Cultural Consequences of Markets and Other Economic Institutions," *Journal of Economic Literature* 36, 75-111.

[22] Boyd, Robert and Peter Richerson [1985], *Culture and the Evolutionary Process*, Chicago, IL: University of Chicago Press.

[23] Broccardo, Eleonora, Oliver Hart, and Luigi Zingales [2020], "Exit vs. Voice," unpublished manuscript.

[24] Cavalli-Sforza, Luigi L. and Marcus Feldman [1981], *Cultural Transmission and Evolution*, Princeton, NJ: Princeton University Press.

[25] Dasgupta, Partha [2008], "Discounting Climate Change," *Journal of Risk and Uncertainty* 37(2/3), Special Issue on Discounting Dilemmas, 141-169.

[26] Garcia-Macia, Daniel, Chang-tai Hsieh, and Peter Klenow [2019], "How Destructive is Innovation?" *Econometrica* 87(5), 1507-1541.

[27] Gollier, Christian and Sébastien Pouget [2014], "The "Washing Machine": Investment Strategies and Corporate Behavior with Socially Responsible Investors" TSE Working Paper, No. 457.

[28] Golosov, Mikhail, John A. Hassler, Per Krusell, and Aleh Tsyvinsky [2014], "Optimal Taxes on Fossile Fuel in General Equlibrium," *Econometrica* 82(1), 41–88.

[29] Greaker, Mads and Kristoffer Midttømme [2016], "Network Effects and Environmental Externalities: Do Clean Technologies Suffer from Excess Inertia?" *Journal of Public Economics* 143(3), 27-38.

[30] Hart, Oliver and Luigi Zingales [2017], "Companies Should Maximize Shareholder Welfare Not Market Value," *Journal of Law, Finance, and Accounting,* 2(2): 247–75.

[31] Harstad, Bård [2019], "Technology and Time Inconsistency," forthcoming in *Journal of Political Economy.*

[32] Hassler, John, A. and Per Krusell [2018], "Environmental Macroeconomics: The Case of Climate Change," in Dasgupta, Partha, Subherendu K. Pattanyak, and V. Kerry Smith (eds.), *Handbook of Environmental Economics,* Volume 4, Amsterdam, NL and Oxford, UK: North-Holland, Elsevier.

[33] Krusell, Per [1998], "How is R&D Distributed Across Firms? A Theoretical Analysis," unpublished typescript.

[34] Lindbeck, Assar and Jörgen Weibull [1987], "Balanced Budget Redistribution as the Outcome of Political Competition," *Public Choice* 52(3), 273-297.

[35] Mattauch, Linus, Cameron Hepburn, and Nicholas Stern [2018], "Pigou Pushes Preferences: Decarbonisation and Endogenous Values," Grantham Research Institute on Climate Change and the Environment, Working Paper No. 314.

[36] Mayer, Colin [2018], *Prosperity: Better Business Makes the Greater Good*. Oxford University Press.

[37] Mill, John Stuart [1863], *Utilitarianism*, 1910 Everyman edition, London: J.M. Dent and Sons Ltd.

[38] Nordhaus, William [2008], *A Question of Balance: Weighing the Options on Global Warming Policies*, New Haven and London: Yale University Press.

[39] Nordhaus, William and Joseph Boyer [2000], *Warming the World. Economic Models of Global Warming*, Cambridge, MA: MIT Press.

[40] Nyborg, Karine, Richard B. Howarth, and Kjell Arne Brekke [2006], "Green Consumers and Public Policy: On Socially Contingent Moral Motivation," *Resource and Energy Economics* 28(4), 351–366.

[41] Oates, Wallace and Robert Portney [2003], "The Political Economy of Environmental Policy," in Mäler, Karl-Göran and Jeffrey R. Vincent (eds.), *Handbook of Environmental Economics*, Volume 1, Amsterdam, NL and Oxford, UK: North-Holland, Elsevier.

[42] OECD [2010], *Taxation, Innovation and the Environment*, available at http://www.oecd.org/israel/ taxationinnovationandtheenvironment.htm

[43] Persson, Torsten and Lars E.O. Svensson [1989], "Why a Stubborn Conservative Would Run A Deficit: Policy with Time-Inconsistent Preferences," *Quarterly Journal of Economics* 104(2), 325-346.

[44] Persson, Torsten and Guido Tabellini [2000], *Political Economics: Explaining Economic Policy*, Cambridge, MA: MIT Press.

[45] Persson, Torsten and Guido Tabellini [2020], "Culture, Institutions and Policy," forthcoming in Bisin, Alberto and Giovanni Federico (eds.), *Handbook of Historical Economics*.

[46] Rochet, Jean-Charles and Jean Tirole [2003], "Platform Competition with Two-Sided Markets," *Journal of the European Economic Association* 1(4), 990–1029.

[47] Romer, Paul [1990], "Endogenous Technological Change," *Journal of Political Economy*, 98(5), S71–S102.

[48] Sandholm, William H. [2010], *Population Games and Evolutionary Dynamics*, Cambridge, MA: MIT Press.

[49] Stern, Nicholas [2006], *The Stern Review of the Economics of Climate Change*, Cambridge, UK: Cambridge University Press.

[50] Tabellini, Guido [2008], "The Scope of Cooperation: Values and Incentives," *Quarterly Journal of Economics* 123(3), 905–950.

[51] Tabellini, Guido and Alberto Alesina [1990], "Voting on the Budget Deficit," *American Economic Review* 80(1), 37-49.

[52] Weitzman, Martin [2009], "On Modeling and Interpreting the Economics of Catastrophic Climate Change," *Review of Economics and Statistics*, 91(1), 1–19.

Appendix

A Additional Material

A.1 Socialization

One way to derive (9) in Section 4.1 is to assume a process of family-based socialization, as in Bisin and Verdier (2001) and Tabellini (2008). Here, we follow the same family-based approach as in Besley (2017). All children have two parents and parents have two children. Reproduction follows a matching process, where a fraction υ of matching is assortative – i.e., parents have the same identity. The remaining fraction $1 - \upsilon$ are randomly matched, which results in some mixed-identity couples. To simplify, we assume two parents of the same type to pass this type on to their children.[28] However, a child with mixed parents may identify as an environmentalist depending on fitness advantage $\beta\Delta'$ – next period's expected-utility difference, discounted by factor β – when the child is adult. The child's identity also depends on a family-specific shock ψ with infinite support and distribution function $F(\cdot)$,which is symmetric around a zero mean with density $f(\cdot)$. A mixed-parent child becomes an environmentalist if $\beta\Delta' \geq \psi$, so the probability of this event is $F(\beta\Delta')$. With a continuum of families, this is the proportion of environmentalist children of mixed parents. Note that $F(\cdot)$ increases smoothly in Δ with $F(0) = 1/2$. This yields

$$\mu' = \mu + (1 - \upsilon) 2\mu (1 - \mu) \left[F(\beta\Delta') - \frac{1}{2} \right]. \tag{24}$$

To interpret this expression, note that assortatively matched couples preserve the proportion of environmentalists. Among the randomly matched, a fraction μ^2 involve two environmentalists. The fraction of mixed-parent households is therefore $2\mu(1-\mu)$. Defining $\varkappa = (1 - \upsilon)$ gives equation (9).

Although we have motivated the model by socialization by parents, a similar story would hold in a wider setting. We could think about peer-group formation at a critical stage of life where people could sort into either

[28] This is clearly a strong assumption, adopted here to make the analysis sharper and simpler. One could consider alternatives, such as a fixed "mutation" rate in homogenous groups.

homogenous groups or mixed groups. If the mixed groups were more open to change, social mixing would again drive the dynamics.

A.2 Deriving the probability of electoral victory

A materialist swing voter supports party A if

$$V(t^A, T^A, q, Q) + \eta + \zeta \geq V(t^B, T^B, q, Q),$$

where η is the idiosyncratic shock and ζ the aggregate shock. Both shocks are assumed to be uniformly distributed: η on $[-1/K, 1/K]$ and ζ on $[-1/X, 1/X]$. This simple formulation – and our specific assumptions about individual utilities – gives a simple solution for policy.

Integrating over η, we can now find the share of materialist swing voters who vote for party A:

$$\frac{1}{2} + E\left[V(t^A, T^A, q, Q) - V(t^B, T^B, q, Q) + \zeta\right]. \tag{25}$$

We assume an interior solution – i.e., (25) lies strictly in the unit interval. A parallel expression holds for environmentalist swing voters.

Party A wins the election if it gets more than half of the votes. This will happen if

$$\zeta + \Gamma\left(t^A, T^A, t^B, T^B, \mu\right) \geq 0, \tag{26}$$

where

$$\Gamma\left(t^A, T^A, t^B, T^B, \mu\right) = \begin{array}{c} \mu\left[v(t^A, T^A, q, Q) - v\left(t^B, T^B, q, Q\right)\right] \\ +(1-\mu)\left[V(t^A, T^A, q, Q) - V\left(t^B, T^B, q, Q\right)\right]. \end{array}$$

The first term in (26) is positive if the realized aggregate shock ζ favors party A, while the second is positive if the party's policy platform allows it to court swing voters.

Integrating over ζ (and exploiting the uniform density), gives us the probability that party A wins the election:

$$z^A = \frac{1}{2} + X\Gamma\left(t^A, T^A, t^B, T^B, \mu\right), \tag{27}$$

assuming an interior solution.[29] Party B wins with the complementary probability $z^B = 1 - z^A = \frac{1}{2} - X\Gamma\left(t^A, T^A, t^B, T^B, \mu\right)$. Each party's probability

[29] This will always be the case if X is small enough – i.e., there is a wide enough support for aggregate shock ζ.

of winning is thus given by the same function. Given the expression for $\Gamma\left(t^A, T^A, t^B, T^B, \mu\right)$, this common objective function is concave. Moreover, it is "as if" each party is maximizing a Utilitarian social-welfare function defined over the short-run parental payoffs. This is a useful benchmark, as the political equilibrium maximizes static welfare, as do classic Pigouvian taxes.

A.3 Alternative dynamics

We now show that there are two possibilities when looking at long-run steady states and the dynamic path towards them given initial conditions $(\mu, q/Q)$. The first case is where an economy always always ends up in a green steady state. The condition for this is given in:

Proposition A1 *If $g\left(\underline{\mu}\right) > G\left(\underline{\mu}, \lambda\right)$, then societal values always converge to $\mu = \bar{\mu}$ for all $\mu_0 \in \left[\underline{\mu}, \bar{\mu}\right]$. Depending on initial conditions, values may not evolve monotonically along the equilibrium path.*

Proof. The result is obvious if $\delta\left(\mu, q/Q, \lambda\right) > 0$ since $\mu_s > \mu$ for all $s > 0$ using (9) and (20). So now consider the case where $\delta\left(\mu, q/Q, \lambda\right) < 0$. Observe first that if $g\left(\underline{\mu}\right) > G\left(\underline{\mu}, \lambda\right)$, then $g\left(\mu\right) > G\left(\mu, \lambda\right)$ for all $\mu \geq \underline{\mu}$. Thus q/Q is an increasing sequence in s. However, μ need not be. But $\delta_s = \delta\left(\underline{\mu}, q_s/Q_s, \lambda\right)$ is increasing in s and there exists $\widetilde{q/Q}$ such that $\delta\left(\underline{\mu}, \widetilde{q/Q}, \lambda\right) = 0$. Hence there exists $\tilde{s}$ such that $\delta_{\tilde{s}} > 0$ for all $q_{\tilde{s}}/Q_{\tilde{s}} > \widetilde{q/Q}$. Then (9) implies that μ_s is an increasing sequence for all $s > \tilde{s}$. The result holds a fortiori, if there exists $\mu_s > \underline{\mu}$ such that $\delta\left(\mu_s, q_s/Q_s, \lambda\right) > 0$. ■

This case is interesting as it can be thought of as a case where economics and politics together allow a country to escape the climate trap. It will hold if the Pigouvian tax on brown goods is large enough, so that even with a small market for green goods, there are stronger incentives to invest in green rather than brown goods. In this case, we could begin with $\delta\left(\mu, q/Q, \lambda\right) < 0$ which means that cultural dynamics are unfavorable to combatting climate change. However, things will eventually turn around since q/Q will increase over time to a point where $\delta\left(\mu, q/Q, \lambda\right) > 0$ and now values will change in the green direction, driven by technological progress.

A.4 Political objectives with lobbying

Given our assumptions in Section 6.3, we assume that the coalition of green firms agrees on contributions that maximize the expected profits of a typical green-variety firm

$$E(\pi) = z^A \pi(t^A) + (1 - z^A)\pi(t^B) - \frac{1}{2}[(c^A)^2 + (c^B)^2].$$

The total contributions collected by each party are

$$\int_0^\phi c^P di + \int_0^\Phi C^P dj = \phi c^P + \Phi C^P.$$

These monies (or other resources) allow parties to monotonically raise their probability of winning elections. To simplify, we use a reduced-form parametric formulation (see Persson and Tabellini 2000, ch. 7), where total campaign contributions of the two parties modify (27), the probability of winning, as

$$z^A = \frac{1}{2} + X\{\Gamma\left(T^A, t^A, T^B, t^B, \mu\right) + \sqrt{\xi}[\phi c^A + \Phi C^A - (\phi c^B + \Phi C^B)]\}. \quad (28)$$

Parameter $\xi > 0$ measures how effectively money influences electoral outcomes.

Using (28), we obtain optimal green-firm contributions as

$$c^A = \max\{0, X\sqrt{\xi}[\pi(t^A) - \pi(t^B)]\} \quad \text{and} \quad c^B = \max\{0, X\sqrt{\xi}[\pi(t^B) - \pi(t^A)]\}. \quad (29)$$

In words, a firm only pays to the one party whose policy yields higher profits. By a similar argument, optimal brown-variety contributions are

$$C^A = \max\{0, X\sqrt{\xi}[\Pi(T^A) - \Pi(T^B)]\} \quad \text{and} \quad C^B = \max\{0, X\sqrt{\xi}[\Pi(T^B) - \Pi(T^A)]\}. \quad (30)$$

Next, we substitute the optimal contributions in (29) and (30), integrate these up over all firms, and substitute the result into (28) to get

$$\begin{aligned} z^A &= \frac{1}{2} + X\{\Gamma\left(T^A, t^A, T^B, t^B, \mu\right) + \\ &\quad X\xi[\phi\mu(\pi(t^A) - \pi(t^B)) + \Phi\left(1 - \mu\right)\left(\Pi(T^B) - \Pi(T^A)\right)]\}. \end{aligned} \quad (31)$$

Compared to (27), the third term now adds a weighted average of profits in the two sectors. Hence the optimal strategy will no longer be Utilitarian, as in the baseline model. This reflects the (rational) expectation that a policy boosting profits in a sector will generate contributions from its lobbying coalition, which – in turn – will help the party win the election. As party B maximizes $1 - z^A$, it once again faces a symmetric problem to A.

Finally, we derive the implications of the model. First, substitute from equations (4) and (5) for equilibrium profits into the political objective (31). Maximizing the resulting expression with respect to the two policy instruments, we can carry out the steps in the proof of Proposition 4 (see below).

B Proof of Propositions

B.1 Proposition 1

Proof. To prove this, we first solve for the optimal tax rate on brown-good varieties. The key observation is that – substituting from (4), (5), and (6)-(8) – we can write each party's problem as maximizing

$$\begin{aligned} \mu v(t,T,q,Q,\mu) + (1-\mu) V(t,T,q,Q,\mu) &= \varepsilon + \mu \left[\frac{q^\sigma y^{1-\sigma}}{1-\sigma} - \chi y\right] + \\ &\quad (1-\mu)\left[\frac{Q^\sigma Y^{1-\sigma}}{1-\sigma} - \chi Y\right] - \lambda (1-\mu) Y. \end{aligned} \tag{32}$$

To get this note that

$$\begin{aligned} V(T,t,q,Q) &= \varepsilon + (1-\mu)[PY - (\chi+T)Y] \\ &\quad + \mu[py - (\chi+t)y] + (1-\mu)TY + \mu t y + \frac{Q^\sigma Y^{1-\sigma}}{1-\sigma} - PY - \lambda(1-\mu)Y \end{aligned} \tag{33}$$

for a materialist and

$$\begin{aligned} v(T,t,q,Q) &= \varepsilon + (1-\mu)[PY - (\chi+T)Y] \\ &\quad + \mu[py - (\chi+t)y] + (1-\mu)TY + \mu t y + \frac{q^\sigma y^{1-\sigma}}{1-\sigma} - py - \lambda(1-\mu)Y \end{aligned} \tag{34}$$

for an environmentalist. The optimum with respect to T satisfies

$$(1-\mu)\left[\frac{\partial\left[\frac{Q^\sigma Y^{1-\sigma}}{1-\sigma} - \chi Y\right]}{\partial T} - \lambda\frac{\partial Y}{\partial T}\right] = (1-\mu)\left[\frac{\sigma\chi + T}{1-\sigma} - \lambda\right]\frac{\partial Y}{\partial T} = 0. \tag{35}$$

Solving for T yields the result. Analogously, for t, the optimum condition is

$$\mu\left[\frac{\partial\left[\frac{q^{\sigma}y^{1-\sigma}}{1-\sigma}-\chi y\right]}{\partial t}\right]=\mu\left[\frac{\sigma\chi+t}{1-\sigma}\right]\frac{\partial y}{\partial t}=0. \tag{36}$$

Solving for t gives the result. ■

B.2 Proposition 2

Proof. To see this, note that if $\delta\left(\mu,\frac{q}{Q},\lambda\right)>0$, we have $\mu'\geq\mu$, from (9) and (20), with strict inequality if $\mu>\underline{\mu}$. Moreover, from (19) $\frac{q'}{Q'}\geq\frac{q}{Q}$ with strict inequality if $\mu>\underline{\mu}$. By induction, this implies that $\delta\left(\mu',\frac{q'}{Q'},\lambda\right)>\delta\left(\mu,\frac{q}{Q},\lambda\right)$, and thus $\mu\geq\mu'$ for all future time periods. A parallel argument establishes that μ is a decreasing sequence if $\delta\left(\mu,\frac{q}{Q},\lambda\right)<0$. ■

B.3 Proposition 3

Proof. To carry out the welfare comparison between the steady states, we write period-s utilitarian welfare as a weighted average of the equilibrium indirect utility of materialists and environmentalists in (17) and (18), also using the equilibrium expressions for n and N in (10) and (11). This yields:

$$\begin{aligned}W(\mu,\lambda)&=\varepsilon+\frac{\sigma}{1-\sigma}\left[\mu q\chi^{1-\frac{1}{\sigma}}+(1-\mu)Q\left[\chi+\lambda\right]^{1-\frac{1}{\sigma}}\right] \\ &-\omega\mu q\left[\frac{\omega}{\beta\varphi\sigma\chi^{1-\frac{1}{\sigma}}\mu}\right]^{\frac{1}{\varphi-1}}-\omega(1-\mu)Q\left[\frac{\omega}{\beta\varphi\sigma\left[\chi+\lambda\right]^{1-\frac{1}{\sigma}}(1-\mu)}\right]^{\frac{1}{\varphi-1}}.\end{aligned} \tag{37}$$

At $\bar{\mu}$ growth rates are $\{\bar{g},\bar{G}\}$ and at $\underline{\mu}$ they are $\{\underline{g},\underline{G}\}$ with $\bar{G}<\underline{G}$, and $\underline{g}<\bar{g}$. Note also that we can write $n=qg^{\frac{1}{\varphi}}$ and $N=QG^{\frac{1}{\varphi}}$.

Let us indicate the two steady states by and . With arbitrary initial conditions, we can write welfare in the green steady state as

$$\sum_{s=0}^{\infty}\beta^{s}\left[\begin{array}{c}\varepsilon+\frac{\sigma}{1-\sigma}\left[\bar{\mu}\overline{q}_0(1+\bar{g})^{s}\chi^{1-\frac{1}{\sigma}}+(1-\bar{\mu})\overline{Q}_0\left(1+\bar{G}\right)^{s}\left[\chi+\lambda\right]^{1-\frac{1}{\sigma}}\right] \\ -\bar{\mu}\omega\overline{q}_0(1+\bar{g})^{s}\bar{g}^{\frac{1}{\varphi}}-(1-\bar{\mu})\overline{Q}_0\omega\left(1+\bar{G}\right)\bar{G}^{\frac{1}{\varphi}}\end{array}\right],$$

and welfare in the brown steady state as

$$\sum_{s=0}^{\infty}\beta^s\left[\begin{array}{c}\varepsilon+\frac{\sigma}{1-\sigma}\left[\underline{\mu q}_0\left(1+\underline{g}\right)^s\chi^{1-\frac{1}{\sigma}}+\left(1-\underline{\mu}\right)\underline{Q}_0\left(1+\underline{G}\right)^s\left[\chi+\lambda\right]^{1-\frac{1}{\sigma}}\right]\\-\underline{\mu}\omega\underline{q}_0\left(1+\underline{g}\right)^s\underline{g}^{\frac{1}{\varphi}}-\left(1-\underline{\mu}\right)\underline{Q}_0\omega\left(1+\underline{G}\right)^s\underline{G}^{\frac{1}{\varphi}}\end{array}\right].$$

As λ gets large, $\bar{G}\to\underline{G}\to 0$ and $[\chi+\lambda]^{1-\frac{1}{\sigma}}\to 0$. Thus we are left with a comparison between

$$\sum_{s=0}^{\infty}\beta^s\left[\varepsilon+\frac{\sigma}{1-\sigma}\left[\bar{\mu}\overline{q}_0\left(1+\bar{g}\right)^s\chi^{1-\frac{1}{\sigma}}\right]-\bar{\mu}\omega\overline{q}_0\left(1+\bar{g}\right)^s\bar{g}^{\frac{1}{\varphi}}\right]$$

and

$$\sum_{s=0}^{\infty}\beta^s\left[\varepsilon+\frac{\sigma}{1-\sigma}\left[\underline{\mu q}_0\left(1+\underline{g}\right)^s\chi^{1-\frac{1}{\sigma}}\right]-\underline{\mu}\omega\underline{q}_0\left(1+\underline{g}\right)^s\bar{g}^{\frac{1}{\varphi}}\right]$$

We need the first expression to dominate the second for all qualities $\left\{\overline{q}_0,\underline{q}_0\right\}$. There are two ways to ensure that the $\bar{\mu}$ steady state has higher welfare. One is that $\underline{\mu}$ is low enough. The other way is that ω is high enough, since then $\bar{g}\to\underline{g}\to 0$. ■

B.4 Proposition 4

Proof. To show this, observe that (using the envelope condition)

$$\pi_t(t)=-\mu y(t)$$

and

$$\Pi_T(T)=-(1-\mu)Y(T).$$

Carrying out the same steps as in the proof of Proposition 1, we get the first-order condition

$$\mu\left[\frac{\sigma\chi+t}{1-\sigma}\right]\frac{\partial y}{\partial t}-\mu\xi\phi y(t)=0.$$

Now, observe that

$$\frac{y}{\partial y/\partial t}=-\sigma(\chi+t).$$

Then, the optimal tax/subsidy solves

$$\left[\frac{\sigma\chi+t}{1-\sigma}\right]+\sigma\xi\phi(\chi+t)=0.$$

Similarly, the optimum with respect to T satisfies

$$(1-\mu)\left[\frac{\sigma\chi+T}{1-\sigma}-\lambda\right]\frac{\partial Y}{\partial T}-(1-\mu)\,\xi\Phi Y\,(t)\,.$$

Noting that

$$\frac{Y}{\partial Y/\partial T}=-\sigma\,(\chi+T)$$

yields

$$\left[\frac{\sigma\chi+T}{1-\sigma}-\lambda\right]+\sigma\,(\chi+T)\,\xi\Phi=0.$$

■

B.5 Proposition 5

Proof. The formula for the optimal subsidy is given by the first-order condition when maximizing (23) with respect to b_s

$$-\mu\omega n-b\mu\omega\frac{\partial n}{\partial b}+\beta\frac{dw}{d\mu'}\frac{d\mu'}{dn}\frac{\partial n}{\partial b}+\beta\frac{\partial w}{\partial q'}\frac{\partial q'}{\partial n}\frac{\partial n}{\partial b}=0,$$

which we can rewrite as

$$-\frac{\mu\omega n}{\partial n/\partial b}-b\mu\omega+\beta\frac{dw}{d\mu'}\frac{d\mu'}{dn}+\beta\frac{\partial w}{\partial q'}\frac{\partial q'}{\partial n}=0.$$

We observe that

$$\frac{n}{\partial n/\partial b}=(1-\varphi)\,(1-b)\,,$$

and that an optimal solution will set b such that

$$\beta\frac{\partial w}{\partial q'}\frac{\partial q'}{\partial n}=\beta\mu'\frac{\partial q'}{\partial n}\frac{(y')^{1-\sigma}}{1-\sigma}.$$

The profit-maximizing condition for investment in quality is

$$\beta\mu'\frac{\partial q'}{\partial n}\,(y')^{1-\sigma}=\omega\,(1-b)\,,$$

which implies

$$\beta\frac{\partial w}{\partial q'}\frac{\partial q'}{\partial n}=\frac{\omega\,(1-b)}{1-\sigma}.$$

If we define

$$\kappa' = \frac{\beta}{\omega}\frac{dw}{d\mu'}\frac{d\mu'}{dn},$$

the optimal subsidy formula becomes

$$-\mu\left[(1-b)(1-\varphi)+b\right]+\kappa'+\frac{(1-b)}{1-\sigma}=0.$$

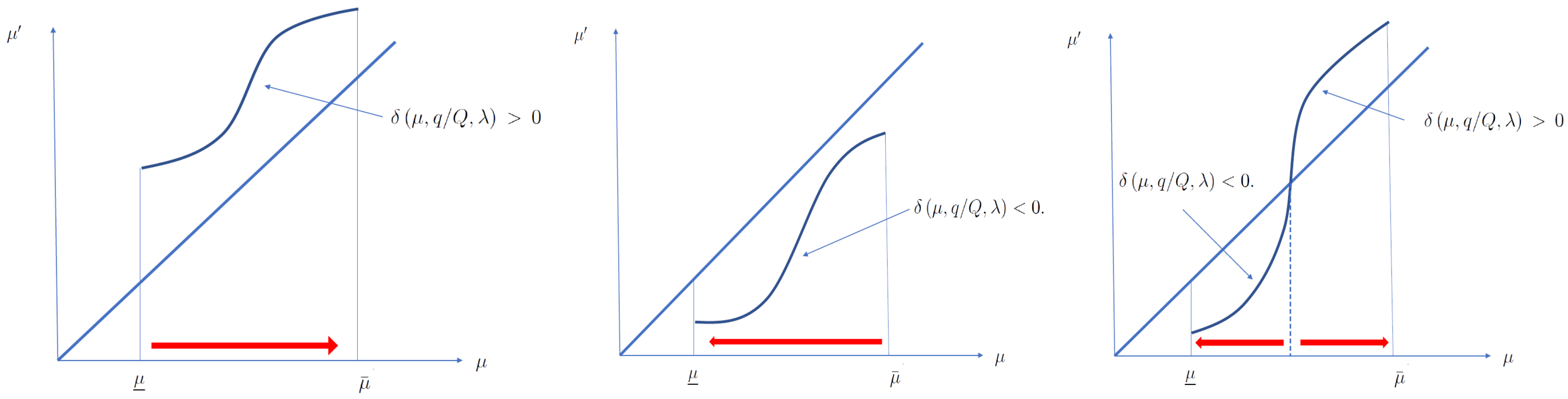
Solving for b yields

$$b\left[\frac{1}{1-\sigma}+\mu\varphi\right]=\kappa'+\frac{1}{1-\sigma}+\mu(1-\varphi),$$

such that

$$b=\max\left\{\frac{\kappa'+\rho-\mu}{\rho},0\right\}$$

where $\rho=\frac{1}{1-\sigma}+\mu\varphi>0$. ■



Panel A　　Panel B　　Panel C

Figure 1