The Dynamics of
Environmental Politics and Values[*]

Timothy Besley
LSE

Torsten Persson
IIES, Stockholm University

June 5, 2019

Abstract

This paper develops a framework to study environmentalism as a cultural phenomenon, namely as reflecting a process of social identification with certain values. The model is used to explain how the shares of environmentalists and materialists in society can coevolve with taxes on emissions to protect society against damages caused by environmental degradation. These policies are determined by electoral competition. However, even though politicians internalize the welfare of those currently alive and pick Utilitarian optimal policies, the dynamic equilibrium paths of policies and evolving values may not converge to the steady state with the highest level of long-run welfare.

1   Introduction

Understanding the processes that shape policy is an important reason to study political economics. The models and approaches that have been developed in this field can help us identify a range of constraints on policy-making that explain the failures to adopt good policy outcomes. Depending on the application, these constraints can emanate from the policy preferences and hence the behavior of various actors, the technologies used in private and public production, the information and commitment abilities of policy-makers, and the institutional arrangements that allocate the use of political power.

The existing political-economics literature has mainly put the spotlight on the institutional framework. To the extent it considers dynamic factors, these typically relate to changes in institutions or (some form of) wealth, while policy preferences are treated as fixed. Given these preferences, policies shape payoffs over time by changing incentives rather than by systematically changing preferences or values.

Arguably, this is a restrictive model for many policy issues. However, endogenizing societal values as manifested in preferences remains a controversial topic and – as a result – economists rarely try to unpack the factors that drive preferences. In many cases, policies have to respond to social values and can also help shape those values. One such case is the topic of this paper: the determination of environmental values.

To study this issue requires breaking with the established tradition in much economics of focusing solely on materialistic preferences. Here, we will conceive of environmentalism as a fundamental value with consequences for consumption behavior – valuing a different life-style that limits pollution. Studying such values and their consequences is especially poignant given concerns about man-made climate change. There is overwhelming scientific evidence that, without substantial reductions in carbon emissions, we risk major disruption of human lifestyles or even human extinction. Hence, most people who see themselves as environmentalists try to reduce their carbon footprint. Moreover, they frequently do so in visible ways that signal their values to others, such as buying ecologically-sourced food, driving an electric vehicle, or riding a bicycle to work.

The standard economic prescription for combatting environmental damage, is to change policy, by imposing a Pigouvian tax on fossil fuels or by fine-tuning a system of tradable quotas for carbon emission. But it is moot how far such recommendations can be deployed successfully, unless we also

recognize that policy is the product of a political process – i.e., we need to treat policy is an equilibrium outcome, rather than as a primitive. The real obstacle to imposing desirable policies may thus lie with politics, even though government actions do reflect the views of their citizens, especially in democratic societies.

The (joint) actions of environmentalists can affect environmental damage directly by avoiding polluting activities. Their actions can also influence the political process and create stricter policies. If values are endogenous, environmentalists may also have an additional influence on future values. One channel of such influence would run through a large group of environmentalists creating an expectation of stricter environmental policies in the future.

Our paper explores these ideas by developing a framework to analyze the joint dynamics of values, politics, and environmental policy. A key feature of the model is to identify a range of complementarities between value adaptation and policy choice. We argue that such dynamics can have important implications for the long-run patterns of social change, where endogenous environmental values either grow or shrink over time. Our main focus is on macro-trends in preferences, when the key dynamics arise via a political externality due to the working of democratic society. This focus is rooted in fact – we show that societies vary in their core environmental attitudes and that these differences cannot be accounted for by individual characteristics.

The model we propose is very simple. It has two kinds of citizens: materialists and environmentalists. Policy is determined by electoral competition between two office-seeking parties that court citizens who are willing to switch their votes. As a result, the parties set environmental policies to cultivate the interests of the average swing voter – this results in a Pigouvian-like policy outcome. Relying on a very simple evolutionary process, we show how (expectations about) these policy choices can drive a society towards either environmentalism or materialism.

We also study the long-run welfare consequences of such changes in societal values. If environmental damages are large enough, a society's welfare is the highest when its population consists only of environmentalists, as this leads to the eradication of pollution. We argue that failure of a democratic political system to achieve this outcome reflects the system's inability to commit for the future. This inability gives a potential welfare-enhancing role for independent institutions insulated from politics (assuming that society can stand by its commitment to these institutions).

3

The remainder of the paper is organized as follows. The next section discusses some related literatures. Section 3 justifies our focus via some facts about individual-level and country-level environmental values from the World Values Survey. Section 4 lays out the economic and political sides of our theoretical framework, for a given share of environmentalists in society. Section 5 develops our dynamic model of changing values. Section 6 studies the welfare implications of our modeling. Section 7 mentions a few possible extensions of the analysis. Section 8 concludes.

2   Building Blocks

The analysis in this paper relates to different bodies of research. A key dimension of the analysis is indeed to cross-fertilize ideas from different parts of social science: economics, politics, sociology, psychology, and anthropology.

Policy responsiveness in static models   By now, a large and established theoretical and empirical literature shows how elections or pressure groups may shape policy within a given set of political institutions. Most of the models are static, studying how policies are chosen to affect the current generation of citizens with no implications for future policy. The main issue in such models is who gets what out of the political process. For instance, the classic model of Downs (1957) predicts that parties motivated mainly by winning an election adopt the preferred position of the median voter, if such a position exists.

Difficulties with equilibrium existence – except in very stylized policy-making environments – led to the development of models with shocks to voters, such that the winning probability varies smoothly with policy positions. Such probabilistic voting models have been used extensively (Coughlin 1992, Lindbeck and Weibull 1987, Persson and Tabellini 2000). Citizen-candidate models (Besley and Coate 1997, Osborne and Slivinski 1996) similarly develop a static framework where policy is chosen by elected officials, but where these now represent different groups in society. Although the exact mechanism varies, a key feature of all these models is that policy responds to citizens' policy preferences. Indeed, some would say that such responsiveness is the essence of democracy.

The standard normative benchmark in models of environmental economics is a Pigouvian tax which is set equal to the pollution externality associ-

ated with the marginal damage done to the environment. Following in the same tradition, the standard political-economics approach to environmentalism (reviewed by Oates and Portney 2003) supposes that underlying values and preferences are fixed. Instead, it examines a static setting, where interest groups lobby policy-makers to move policy in their preferred direction. These policy-makers have mixed motives over social welfare and money transfers, which are presumably related to the citizens' underlying preferences as expressed through elections.

Dynamic politico-economic models   A number of strategic models study how current policy choices can influence future elections or policy outcomes. A common theme in those models is the lack of policy-commitments. Our novel contribution is to draw out implications of these dynamic commitment issues for changes in societal values.

The first main application of dynamic political models was to explore dynamic interlinkages from debts and deficits. Thus, Persson and Svensson (1989) argue that debt policy can stray from the efficient path in political equilibrium, while Tabellini and Alesina (1990) show how political instability can give incumbent officeholders incentives to borrow from the future. In both these papers, current incumbents strategically alter future debt levels to manipulate the choices of future policymakers who may not share their own policy preferences.

Aghion and Bolton (1990) and Milesi-Ferretti and Spolaore (1994) develop models where policy is distorted because current policy choices affect which political party will win in the future. The use of policy in such models is deliberate and strategic and politicians understand that today's policies have future political implications. A similar effect has been used to explain other policy puzzles. For instance, Acemoglu and Robinson (2006) ask why some governments fail to invest in developing the economy and refer to a "political replacement effect," whereby today's policymaker fears that certain policies reduce her chances to survive in office.

Acemoglu (2003) and Besley and Coate (1998) point to a key feature of dynamic political models, namely that policies can be rendered inefficient by an inability to commit by political decision-makers – of course, this is related to the general time-consistency problem in policymaking discussed by Kydland and Prescott (1977). For example, in a two-group setting, a policy which could benefit both groups may not be implemented because one

group cannot commit to compensate another in the event of a transition in political power. Lacking commitment ability in the political process gives way to a potential role for non-elected independent institutions as sources of credible commitment. The classic example is to use an independent central bank in monetary policymaking when politicians are tempted to inflate the economy as in Rogoff (1985). Similarly, Acemoglu and Robinson (2000) point to the introduction of the franchise when incumbent elites are tempted to renege on promises of redistribution to the masses.

Yet another strand of dynamic modelling is developed in the literature on state capacity. Besley and Persson (2009, 2011) study models where the main dynamic force comes from the state investing in its own functions – the powers to tax, to adjudicate, and to deliver public goods. This analysis shows that finding institutional ways for more cohesive politics can serve as to spur such investments.

None of these dynamic models consider changing values and preferences as an explicit mechanism of economic change. These ideas are also developed by Besley and Persson (2019a), who build a multi-dimensional dynamic model to explore the rise of nationalist "identity politics". Besley (2019) looks at how compliance with taxation depends on preference types in the population with a dynamically evolving level of taxation, while Besley (2017) studies how redistributive policy changes endogenously with aspirational preferences in the population. Besley and Persson (2019b) ask how the design of democratic institutions interacts with democratic values over time. The general coevolution of institutions and culture is studied by Bowles, Choi, and Hopfensitz (2003) and Bisin and Verdier (2017).

Values, preferences, and identities   We will think of environmentalism as a particular form of pro-social preferences – i.e., as a commitment to a specific cause. This means that our model of environmentalists versus materialists relates to interconnected literatures on intrinsic motivations. It is well known that private intrinsic motives to do good can underpin pro-social actions. Andreoni (1989) notes that this can come from receiving status or acclaim, or simply from a "warm glow" associated with taking an action. We focus on the first of these motivations and model the utility from environmentalism as deriving from the signalling benefits associated with non-consumption of polluting goods. Even though an individual environmentalist cannot materially affect the level of pollution with her own actions, she con-

tributes to the environmental cause. Other ways to think about the same phenomenon would run via mission-driven preferences (Besley and Ghatak 2005), or via adoption of a particular social identity (Akerlof and Kranton 2005, 2010). Along the latter way, the identity of being an environmentalist includes not to consume goods that cause pollution.

We will adopt a specific micro-foundation based on Harbaugh (1998) and Benabou and Tirole (2006). Thus we suppose that the motive to become an environmentalist is to get social respect as a virtuous person. However, in our version of their social-signalling model, consumption of polluting goods is only imperfectly observed. This means that the value of social signals depends on the identity shares in the population. With a high fraction of environmentalists, it is more likely that someone not observed to consume polluting goods is in fact an environmentalist. By contrast, with a low share of environmentalists, there is little signalling value in abstaining from consuming polluting goods.

In this sense, our paper is related to Nyborg et al (2006) who study the emergence of green consumers in a model of pro-social motivation modeled as self-image. They also emphasize the possibility of multiple steady states and discuss the role of tax policy in affecting long-run outcomes. Nyborg and Rege (2003) demonstrate how Norwegian smoking norms developed with changes in smoking regulation – specifically, they show how smokers' attention to their effects on others evolves over time. Our research is also related to the work by Lindbeck et al (1999), who propose a (static) voting model where policies and social norms are co-determined in equilibrium and show how this can give rise to multiple equilibria.

Cultural dynamics Generally speaking, our paper is part of a wider agenda aimed at studying the coevolution of values or cultures with other strategically designed outcomes. There is little doubt that drivers of preferences and values not only reflect inherited genetic endowments but are also shaped by cultural fitness. Cultural "memes" can be propagated by social influences transmitted by families, teachers, peer groups and other social networks. Such social factors surely act on us throughout the life cycle, but they may be particularly important during the formative years of childhood and adolescence. Social influence at that critical stage of life may thus leave a permanent mark on choices and behavior in adulthood.

Taking explicit cultural or value dynamics seriously is a recent develop-

ment in economics where culture had previously been thought of as an "error term". But economists increasingly appreciate that certain cross-sectional and time-series observations of cultural traits cry out for explanation, using economics tools and methods of empirical investigation.

Resistance to these ideas among economists create barriers to dialogue across disciplines. Thus, the idea that preferences are fluid and socially determined is readily accepted among sociologists. In a classic account, Bales and Parsons (1955) put it as follows:

> "If .. the essentials of human personality were determined biologically, independent of social systems, there would be no need for families ... It is because the human is not "born" but must be "made" through the socialization process that ... families are necessary. They are "factories" which produce human personalities.... We therefore suggest that the basic and irreducible functions of the family are two: first the primary socialization of children so that they can truly become members of the society ... ; second, the stabilization of the adult personalities of the society." (pages 16-17)

Among economists, Sam Bowles is a pioneer for this view (Bowles 1998):

> "the argument that economic institutions influence motivations and values is plausible, and the amount of evidence consistent with the hypothesis is impressive. Many ethnographic and historical studies, for example, recount the impact of modern economic institutions on traditional or indigenous cultures. The rapid rise of feminist values, the reduction in family size, and the transformation of sexual practices coincident with the extension of women's labor force participation likewise suggest that changes in economic organization may foster dramatic changes in value orientations." (page 76)

The ideas in this paper are also heavily influenced by the formal models of cultural change developed by evolutionary anthropologists (Boyd and Richerson 1985, Cavalli-Sforza and Feldman 1981). They borrow from the formal structures of population biology to model behavioral change as social learning that propagates behavior across populations.[1] This approach has

---

[1]Similar general ideas appear in the literature on evolutionary game theory, which is surveyed and developed in the books by Weibull (1995) and Sandholm (2010).

been influential in exploring the basis of unselfish behavior in kin groups or broader social groups. An important idea in this literature is the notion of cultural parents who influence the behavior of their offspring. Cultural parents are not confined to biological or foster parents, and can include a wide range of peers in education, social life, and education.

A canonical example in this research is the public-goods game, where rational self-interested individuals do not contribute because of their incentive to free-ride on the contributions of others. But this can be altered by three evolutionary mechanisms: mutations, genetic drift (relevant only in finite populations), and natural selection. Boyd and Richerson (1982) consider "conformist transmission" where individuals imitate the more common behavioral types among their cultural parents, which raises the frequency of these types in the population.

A growing literature in economics considers related ideas – see Bisin and Verdier (2011) for a useful review.[2] Bisin and Verdier (2001) develop a model where parents strategically socialize their children, by weighing future payoffs of children against the "social distance" between parents and children. Bisin and Verdier (2000) apply this approach to study the dynamics of religion and ethnicity. Kuran and Sandholm (2008) also develop a model of cultural integration. Bezin (2015) proposes a model of cultural evolution for environmental preferences based on private contributions to environmental protection.

The evolution of preferences we use in this paper builds on the indirect evolutionary approach introduced in Güth and Yaari (1992) and Güth (1995), who propose that preferences respond to payoffs in repeated games. In such models, whether or not preferences are observable is a key issue. In our setting, observability shapes short-run economic behavior, but is less important for the dynamics. This is because the key externality runs through an electoral process, where individuals (stochastically) vote in a sincere way. Ostrom (2000) emphasizes the indirect evolution approach in a context of collective action. By changing preferences, societies can become more or less cooperative and hence more or less able to solve collective-action problems. To date, however, applications of these ideas have focused on small-scale cooperation.

---

[2]See also Saez-Marti and Zillibotti (2008) for an overview of the issues.

3 Environmental Values

As already mentioned, those concerned about human impacts on the environment often suggest that changing values creates a route to more sustainable behavior and policy. Environmental values are most commonly expressed through spending patterns or recycling, but also through political activism. One outstanding issue is whether the underlying values are based on self-interest, humanistic altruism, or biospheric altruism. Dietz et al (2005) systematically review these issues across the social sciences. They discuss how environmental values can be thought about as a kind of "post-materialist" ethic associated with altruism. However, they lament that

"little can be said about the causes of value change and of the overall effects of value change on changes in behavior." (page 335).

Lorenzoni and Pidgeon (2006) discuss a range of polling data and note that environmental attitudes vary across populations, particularly between Americans and Europeans. In his overview of the growing social movements that put environmental values at the heart of campaigns to change policy, McAdam (2017) discusses why such values have spawned so little activism in the US.[3]

One of the key ideas in the model of the next section is that people have heterogeneous values regarding the need to protect the environment and that these attitudes shape their policy preferences. To shed some light on these attitudes and policy preferences, and their differences across individuals and countries, we turn to the World Values Survey (WVS).

Using the WVS data We use two questions from this survey. The first question appears in WVS waves 3, 4, 5 and 6 and is answered by about 250,000 people. It asks each respondent whether they would prioritize the environment over economic growth, or vice versa. We code the answer in a binary fashion, and set an indicator equal to one if s/he regards protecting the environment as the priority. We think of such individuals as self-identifying as environmentalists. This applies to 54 percent of the full, worldwide sample, although – as we will see below – this identification varies systematically across countries and across individual characteristics.

---

[3]Aghion et. al. (2019) explore the role of environmental attitudes in encouraging incentives to innovate. They show that such values interact with competition in promoting innovation.

The second question concerns policy and is posed in WVS waves 2, 3, 4 and 5 and is answered by about 190,000 people. It asks each respondent whether they favor an "increase in taxes if it is used to prevent environmental pollution." The four alternative answers are: "strongly agree", "agree", "disagree", or "strongly disagree". We code the response as favorable to environmental taxes if the respondent strongly agrees or agrees. This is true for about 44 percent of the sample. Once again, we will find stark individual and cross-country variation

Variation across individuals and countries We expect people to express different environmental attitudes depending on the period and circumstances in which they were socialized, and that this socialization predominantly occurs at the earlier, more formative, stages in life. To explore this, we construct a variable for birth cohort, for each ten-year period since the 1910s. As education is likely to influence an individual's attitudes to the environment, we also examine this variation using the WVS classification into three levels of education. Figure 1 includes two bar charts that show deviations from the country-specific mean of the answers to our two attitudinal questions. This isolates the idiosyncratic variation by cohort and education group. The left panel shows a clear variation: environmentalism, as well as the willingness to raise environmental taxes, are stronger among more recent cohorts. This suggests a shift in values across generations towards increasing environmentalism. The right panel shows a larger concern for the environment and a greater willingness to put up taxes among more highly educated groups.

Figure 2 displays four cross-country histograms for the share of environmentalists and the share of people willing to raise environmental taxes, defined by our binary classifications. To maximize the number of countries in the data, we average this across all WVS waves. The top two panels show the raw data, while the bottom two panels condition on a range of individual characteristics.[4] Responses are strikingly different across countries, whether we condition on individual characteristics or not. The share of environmentalists e.g., varies between 20 and 80 percent in the raw data and remains highly

---

[4]We estimate a linear probability model at the individual level with either environmental dummy variable on the left-hand side and right-hand side variables for gender, ten dummies for income groups, three for education groups, three age bands and wave dummies. To construct the bottom-row histograms, we average the residuals at the country level.

variable as we condition on individual traits. It is plausible to attribute these macro differences to different cultural values.

Figure 3 shows how our indicators for values and policy preferences covary across countries. It plots the average willingness to raise environmental taxes against the importance of protecting the environment, with the raw data in the left panel and the residuals used in Figure 2 in the right panel. Both graphs display a clear positive correlation, showing that values and policy preferences indeed go hand in hand. The correlation is strong, especially if we ignore the outliers on (both sides of) the horizontal axis. These graphs also suggest a country-specific component of social values, or culture, in the environmental domain.

Taken together, Figures 1-3 convey two salient patterns in the data. Environmental attitudes clearly vary across cohorts, as well as categories of education. They also display substantial macro, country-level, differences which cannot be explained by individual characteristics. These differences across generations and societies give some underpinnings to our modeling, which indeed implies society-specific environmental values that change systematically across generations.

4 Static Economics and Politics

We begin this modelling section by a short overview and commentary on the approach that we take.

4.1 Overview

There are three main elements in our modeling approach which are described in the next two sections. The first is an economic element, where we use a simple model with two types of citizens/consumers, who face an environmental policy (Subsection 4.1). The second is a political element, where we use a probabilistic-voting model of two parties that try to win elections with swing voters and loyal voters among the citizens (Subsection 4.2). The third is an evolutionary element, where values (citizen/consumer types) evolve over time in response to expected policy (Section 5). As discussed above, policymaking is a natural source of commitment problems in politics. What is novel here, as in Besley and Persson (2019a), is how the politics of policymaking shape the endogenous evolution of values.

Specific assumptions   We will make some specific assumptions which reflect a conscious effort to home in on the most some key aspects of environmental values. [5]

First, we characterize an environmentalist as someone who completely avoids consumption of (some) polluting goods. The virtue associated with this is earned from the social signalling that this permits.[6]

Second, we employ a specific political model where the equilibrium leads to parties converging on a policy that maximizes short-term utilitarian payoffs. This is not general, but it is a useful starting point since the welfare distortions that we study are then not attributable to the way that politics aggregates preferences.

Third, we adopt a specific model of intergenerational socialization where limited assortative matching creates scope for socialization based on the cultural-fitness advantage (or disadvantage) of environmentalists over materialists.

Time and types   The model has infinite time, which is labelled by s. For most purposes, we will think of s as labeling a sequence of generations. A single generation of adults is alive at each date. The population includes two types of citizens denoted by $\tau \in \{m, e\}$ where m stands for materialists and e for environmentalists. Let μ_s be the proportion of environmentalists in the population at date s.

4.2   Economics

Everybody in society has the same level of income y. Citizens choose a life style which determines how much they pollute the environment with their consumption. We associate pollution – such as carbon emissions – with a specific component of consumption, a good denoted by c which can be taxed at rate t. There is also another type of non-polluting (ecologically friendly) consumption, n, with price $p \geq 1$. If consuming non-polluting goods is more expensive, then the inequality is strict. Consumers are of two types: materialists and environmentalists.

---

[5]Readers who would like to see a treatment with fewer specific assumptions can find a more general version of the model in the Web Appendix.

[6]Many of the results in the paper would continue to hold if we assumed a fixed utility from the "warm glow" of being an environmentalist, as in the philantropy model of Andreoni (1989).

Materialists   Materialists have preferences

$$u^m = \log(Ac) + n - \lambda C,$$

which are linear in the numeraire, n, and where the final term is the disutility from environmental damage, which we assume is proportional λ to the average per-capita consumption of the polluting good denoted by C. It is easier to pursue the analysis in terms of $\alpha = \log(A) - 1$, a parameter that additively shifts materialist preferences and therefore plays a role in the utility comparisons between types and steady states, which we conduct below.

The budget constraint for consumption is $y + r = c(1 + t) + pn$, where r is a lump-sum government transfer. Optimal polluting consumption is given by

$$\widehat{c}(t, p) = \arg\max_c \left\{ \alpha + 1 + \log(c) + \frac{y + r}{p} - \frac{(1 + t)c}{p} \right\} = \frac{p}{(1 + t)}.$$

As each consumer is small, she cannot affect C by her own actions and hence ignores the effect of consumption on overall pollution. We assume that $0 < \widehat{c}(t, p) < \frac{y}{p}$ and let

$$v(t, p) = \alpha + 1 + \log(\widehat{c}(t, p)) - \frac{(1 + t)\widehat{c}(t, p)}{p} = \alpha + \log\left(\frac{p}{(1 + t)}\right) \quad (1)$$

be the indirect utility function from good c.

Environmentalists   Environmentalists get utility from social signalling as well as from consuming the numeraire. Specifically, they have preferences

$$u^e = n - (\lambda + \theta)C + V(\mu)$$

and, as do materialists, face a budget budget constraint $y + r = c(1 + t) + pn$. Since environmentalists get no utility from the polluting good, it is optimal for them to set $c = 0$.

Apart from non-polluting consumption, the environmentalist utility function has two additional components. Like materialists, environmentalists suffer from pollution, but they put an additional (relative) weight on it – specifically, they perceive an extra cost θ, such that their total cost of pollution is $(\lambda + \theta)C$. The second component is a measure of private virtue which,

following Harbaugh (1998) and Benabou and Tirole (2006), we model as the social-signalling benefit from being an environmentalist. To motivate this, we assume that types are not directly observed but inferred from imperfectly observed behavior. Specifically, if a consumer sets $c = 0$, this can be observed while if $c > 0$ it is observed with probability $\rho \in (0, 1)$.

By Bayes rule, we can write the probability that an individual observed setting $c = 0$, is an environmentalist:

$$\varphi(\mu) = \frac{\mu}{(1 - \mu)\rho + \mu}.$$

Note that $\varphi(\mu)$ is increasing in μ with $\varphi(1) = 1$ and $\varphi(0) = 0$.

We suppose that the virtue utility from being an environmentalist is $V(\mu) = \chi\varphi(\mu)$ where $\chi > 0$ is the gain in social recognition from being perceived as an environmentalist rather than a materialist. Assuming that χ is positive is a simple way to micro-found why it is attractive to be an environmentalist: although you forgo some private consumption, you gain social respect. This corresponds to what is often called "virtue signalling". Our specific micro-foundation generates a *positive* link between utility from virtue and μ – in a society with many (few) environmentalists, it is more (less) likely that someone not observed choosing $c > 0$ is indeed an environmentalist. This link will mean that virtue signalling has social effects by generating benefits to being an environmentalist which can promote cultural change.

Policy preferences   To close the model, we first assume that tax revenue is rebated back to both groups of consumers on a uniform basis through per-capita (lump-sum) grant:

$$r = Ct.$$

We then substitute from the consumer and government budget constraints and use the equilibrium condition that $C = (1 - \mu)\widehat{c}(t, p)$. Finally, we set p to 1 so that we can eliminate it as an argument from the functions above.[7]

---

[7]Setting $p = 1$ is for notational convenience, but it is *not* a pure normalization. In a fuller (closed-economy) model, p would reflect the supply side and the relative costs of producing goods favored by environmentalists and materalists. As such, it would affects the utility comparison between the two types. In this way, changing technologies for producing the two goods (which would be reflected in p) could also affect the evolution of values. This makes sense: if substitutes for non-polluting goods become cheaper, it becomes more attractive to become an environmentalist, everything else equal.

This allows us to write the policy preferences of the two types

$$u^\tau(t,\mu) = \begin{cases} \chi\varphi(\mu) - (\lambda + \theta - t)(1 - \mu)\widehat{c}(t) + y & \text{if } \tau = e \\ v(t) - (\lambda - t)(1 - \mu)\widehat{c}(t) + y & \text{if } \tau = m. \end{cases} \quad (2)$$

As $\theta > 0$ and $v(t)$ is decreasing in t, the tax rate preferred by environmentalists is always strictly higher than that preferred by materialists.

We assume throughout that

$$\alpha + \log\left(\frac{1}{1 + \lambda + \theta}\right) < 0 < \alpha + \log\left(\frac{1}{1 + \lambda}\right), \quad (3)$$

which requires A and λ to be small enough relative to θ. Roughly speaking, the salience of the (relative) weight attached to pollution damages thus has to be sufficiently different across environmentalists and materialists.

4.3   Politics

As in Besley and Persson (2019a), we could think about a model with underlying cleavages in two dimensions: party politics and identity politics. Party politics would reflect conflicts in a fixed dimension, like class or religion, which creates loyalty to a particular party for some groups of voters. (Such preferences have been documented extensively in survey data.) Identity politics – represented by the values held by types τ – would be imperfectly correlated with party loyalty. These values would be associated with conflicting preferences over environmental taxes t, which spill over to preferences over party platforms. The model developed below has only the second dimension, however.

Parties   Consider a model of two-party competition with probabilistic voting. We label the (given) parties A and B and assume that they are solely motivated by winning elections. Each of the two parties chooses a party platform for its proposed environmental tax rate: $\{t^A, t^B\}$.

We pick this particular formulation for pure convenience. It will clearly illustrate how our framework departs from standard models by allowing the population types to evolve over time. But the effects of changing types would extend to any kind of model where a higher population share of a certain type

moves policy towards the one preferred by that type.[8] In the probabilistic-voting approach, this happens smoothly (see, e.g., Lindbeck and Weibull 1987, or Persson and Tabellini 2000).

Voters There are two kinds of voters. Swing voters cast their ballots based on proposed policy platforms. Loyal voters always cast their ballots for one of the parties. This distinction follows a long-standing, political-science tradition based on the Michigan voting surveys. To simplify the algebra, we assume that each party has the same fraction of loyal voters and that the same fraction of each type are swing voters.[9]

Following the probabilistic-voting approach, the party choice by swing voters is also subject to idiosyncratic (voter specific) and aggregate (affecting all voters) shocks. A swing voter of type τ supports party A if

$$u^\tau(t^A, \mu) + \varepsilon + \zeta \geq u^\tau(t^B, \mu),$$

where ε is the idiosyncratic shock and ζ the aggregate shock. Both shocks are assumed to be uniformly distributed: ε on $[-1/\epsilon, 1/\epsilon]$ and ζ on $[-1/\psi, 1/\psi]$ where in each case a positive shock favours party A. This simple formulation – together with our specific assumptions about individual utilities – gives a closed-form solution for policy.

Using these assumptions and integrating over ε, we find the proportion of type τ swing voters who vote for party A:

$$\frac{1}{2} + \epsilon \left[u^\tau(t^A, \mu) - u^\tau(t^B, \mu) + \zeta \right]. \tag{4}$$

We assume an interior solution – i.e., that (4) lies between zero and one.

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

$$\zeta + \Omega\left(t^A, t^B, \mu\right) \geq 0, \tag{5}$$

where

$$\Omega\left(t^A, t^B, \mu\right) = \begin{array}{l} \mu \left[u^e(t^A, \mu) - u^e(t^B, \mu) \right] \\ + (1 - \mu) \left[u^m(t^A, \mu) - u^m(t^B, \mu) \right]. \end{array}$$

---

[8]As discussed in the Web Appendix, the main insights would also hold with policy-motivated parties, or with citizen candidates, as long as these would have, direct or indirect, motives to court non-environmentalist swing voters as part of their electoral strategy.

[9]As shown in the Web Appendix, this assumption is easy to generalize.

The first term in (5) just depends on whether the realized aggregate shock ζ favors party A, while the second depends on whether the policies on offer allow the party to court swing voters.

Integrating over ζ, gives us the probability that party A wins the election as:

$$q^A = \frac{1}{2} + \psi\Omega\left(t^A, t^B, \mu\right),\tag{6}$$

assuming an interior solution.[10] Party B wins with the complementary probability $q^B = 1 - q^A = \frac{1}{2} - \psi\Omega\left(t^A, t^B, \mu\right)$. It follows that the probability of winning for each party is given by the same function of its own tax rate. Given the expression for $\Omega\left(t^A, t^B, \mu\right)$, the parties are thus effectively maximizing the same Utilitarian social-welfare function.

Equilibrium tax rates To study equilibrium policy choices, we look for a Nash equilibrium where each party optimizes its policy platform given the decision of the other. In view of the comments above, the political equilibrium maximizes a Utilitarian objective similar to what would emerge from a standard Pigouvian model. Specifically, we have:

Proposition 1 *Both parties pick the same tax rate to maximize their winning probability*

$$t^A = t^B = \widehat{t}\left(\mu\right) = \lambda + \mu\theta.$$

Proof. To prove this, note that (2) implies

$$t = \arg\max \ (1 - \mu)\left\{\frac{(t - \lambda - \mu\theta)}{1 + t} + v\left(t\right)\right\}.$$

So the first-order condition is

$$\frac{1}{1 + t} - \frac{t - \lambda - \mu\theta}{(1 + t)^2} - \frac{(1 + t)}{(1 + t)^2} = 0.$$

Solving this expression yields the result. ∎

---

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

Observations We end this section with three observations about the solution in Proposition 1.

First, the proposition gives the lowest tax rate as $t(0) = \lambda$. This is the conventional Pigouvian tax that exactly corrects for the environmental externality in a population where everybody is a materialist and causes pollution. For positive shares of environmentalists, the tax rate is higher. Note, however, that the political equilibrium produces a Utilitarian optimal policy for any given μ. Hence, there is no political failure according to that conventional criterion. This makes the political equilibrium tax a useful and important benchmark for the analysis to follow in Section 6. There, we will show that with endogenous values the traditional Pigouvian optimum does not necessarily yield the appropriate optimum, even if we stick to utility-based criteria.

Second, when $\mu = 1$, we have $t(1) = \lambda + \theta$. This tax rate is higher than when $\mu = 0$, but still finite. However, there is no consumption of the polluting good, as all the materialists have disappeared.

Third, define

$$u^\tau(\mu) = u^\tau(\widehat{t}(\mu), \mu), \tag{7}$$

the equilibrium utility of type τ when the population has a fraction μ of environmentalists. It is clear that $u^e(\mu)$ is increasing, but $u^m(\mu)$ decreasing, in μ. This is because a higher fraction of environmentalists makes politicians put more weight on their preferences relative to those of materialists.

5   Dynamics of Environmentalism

The dynamics of preferences is the least standard element of our analysis. This section specifies the evolutionary model we rely on, analyzes the resulting dynamics, and describes the model's steady state(s). Given the binary types and our formulation of the evolution process, that analysis turns out to be simple.

The evolution of values We posit a class of dynamics in which the main driver depends only on the relative expected payoff from being one type rater than another.[11] The proportion of a type in the population that "thrives"

---

[11]This is only one possibility out of several: fitness could also be purely material – e.g., based on real income levels – or purely social – e.g., based on dominant behavior in a relevant peer group.

in a particular cultural environment will tend to increase depending on the size of this payoff difference. For this to be viable, there needs to be a capacity for intrapersonal comparisons of utility between materialists and environmentalists, i.e. parents need to assess the gain in psychological well-being from socializing their children to be different types.[12]

Hence, the key magnitude that drives the socialization process is the utility difference between environmentalists and materialists. For an anticipated fraction of environmentalists μ_{s+1}, this difference is

$$\Delta\left(\mu_{s+1}\right) = u^e(\mu_{s+1}) - u^m(\mu_{s+1}). \tag{8}$$

Using the properties (7) at the end of Section 4, we can show that (8) is increasing in μ (see further below).

Specific micro-foundation   We work with a specific micro-founded model where, as in Besley (2017), cultural transmission takes place across successive generations. We suppose that there are two generations alive at each date, "parents" and "children". Only parents are allowed to vote (and also make consumption decisions on behalf of their children). To keep the population balanced, every family has two parents and two children. Reproduction follows a matching process in which a fraction β of mating is assortative – i.e., parents have the same type. The remaining fraction $1 - \beta$ of parents are randomly matched and hence some couples will have different types.

Children are socialized by their parents. To simplify the analysis, we assume that two parents of the same type guarantees that their common type

---

[12]This is only one specific socialization protocol where cultural fitness is based on payoff differences. We could also have followed the formulation in Sandholm (2010), where individual types evolve sporadically (with inertia), and where switches depend on current behavior and opportunities (myopia). This approach is underpinned by a revision protocol $\varsigma_s^{i,j} \in [0,1]$ for $i,j \in \{e,m\}$ that specifies a time-varying conditional switch rate from type i to j given the payoffs and proportion of types in the population. In our forward-looking model, this would yield:

$$\mu_{s+1} - \mu_s = (1 - \mu_s)\,\varsigma_s^{e,m} - \mu_s\varsigma_s^{m,e},$$

where

$$\varsigma_s^{m,e} > 0 \iff \Delta(\mu_{s+1}) > 0 \ \text{ and } \ \varsigma_s^{e,m} > 0 \iff \Delta(\mu_{s+1}) < 0.$$

is passed along to their children.[13] However, whether a child with mixed parents becomes an environmentalist depends on any cultural fitness advantage $\Delta\left(\mu_{s+1}\right)$ – i.e., the utility difference for the two types in the next period – when the child has become an adult. The child's type also depends on a family-specific shock ν that has infinite support and distribution function $G\left(\cdot\right)$, which is symmetric around a zero mean with density $g\left(\cdot\right)$. The condition for becoming an environmentalist is $\Delta\left(\mu_{s+1}\right)\geq\nu$ so that the probability that an individual with mixed parents becomes an environmentalist is $G\left(\Delta\left(\mu_{s+1}\right)\right)$. Given a continuum of families, this will also be the proportion of environmentalists among those with mixed parents. Note that $G\left(\cdot\right)$ increases smoothly in Δ with $G\left(0\right)=1/2$.

Value dynamics   Given this framework, the proportion of the population who are environmentalists at date $s+1$ given that μ_s are environmentalists at s evolves according to:

$$\mu_{s+1} = \mu_s + 2\mu_s\left(1-\mu_s\right)\left(1-\beta\right)\left[G\left(\Delta\left(\mu_{s+1}\right)\right)-\frac{1}{2}\right]. \qquad (9)$$

To interpret this expression, note that assortative matching preserves the proportion of environmentalists. However, among the randomly matched, a fraction $\left(\mu_s\right)^2$ are matched with other environmentalists. The fraction of mixed-parent households is therefore $2\mu_s\left(1-\mu_s\right)$.[14]

Timing   The timing of the dynamic model is as follows

1. There is an initial share of environmentalists in the population represented by μ_s.

2. Parties choose policy and compete for office leading to a tax rate t_s (as described in Section 4.2)

---

[13]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.

[14]Note that the population fraction that matches assortatively does not affect the steady state of the model, only its speed of convergence as long as $\beta < 1$ – i.e., some matching is random. Parameter β can be thought of as crudely measuring the openness of social structures, as assortative matching will tend to entrench existing values while lower β implies more rapidly changing values.

3. Payoffs of citizens are realized (as described in Section 4.1).

4. Citizens match, a new generation is born and children are socialized leading to μ_{s+1}.

A political complementarity We have seen that the key driver of the dynamics is $\Delta(\mu)$. Using (2), we have

$$
\begin{aligned}
\Delta\left(\mu_{s+1}\right) &= u^e(\mu_{s+1}) - u^m(\mu_{s+1}) \qquad\qquad (10)\\
&= \chi\varphi\left(\mu_{s+1}\right) - \theta\left(1 - \mu_{s+1}\right)\widehat{c}\left(\widehat{t}\left(\mu_{s+1}\right)\right) - v\left(\widehat{t}\left(\mu_{s+1}\right)\right).
\end{aligned}
$$

Note that this expression depends on the future equilibrium tax rate which depends on μ_{s+1}. Since $\widehat{t}\left(\mu_{s+1}\right)$ is determined by next period's political equilibrium, it is taken as given by parents who cannot unilaterally influence μ_{s+1}. It is straightforward to see that a higher value of μ_{s+1} favors environmentalists since:

$$
\Delta_\mu\left(\mu\right) = \chi\varphi_\mu\left(\mu\right) - v_t\widehat{t}\left(\mu\right)\widehat{t}_\mu - \theta(1 - \mu)\widehat{c}_t\widehat{t}_\mu + \theta\widehat{c} > 0.
$$

The first term is positive as the social-signalling mechanism is more effective for a higher value of μ. The second term is positive because a higher μ favors environmentalists in politics and this gives lower indirect utility for materialists via higher taxation. Moreover, the third and fourth terms are positive, as the higher taxes cuts pollution and makes environmentalists better off, directly and indirectly. This implies that there is a positive complementarity between policy and preference evolution, created by the responsiveness of electoral politics to the proportion of environmentalists in the population.

Dynamics and steady states for values To study the steady states, we begin with two key observations.

First, when there are very few environmentalists around, the cultural fitness advantage lies with the materialists. To see this, note that as $\mu \to 0$, then $\varphi(\mu) \to 0$ which implies that $\Delta(0) < 0$.[15] Since there is no signalling value from not observing $c > 0$ when there are no environmentalists in the population, it is always best to be a materialist in a world where more or less everyone else is. Moreover, this is true even if being an environmentalist is a respected thing, i.e. $\chi > 0$.

---

[15] Note that (3) implies that $v\left(\widehat{t}\left(0\right)\right) = \alpha + \log\left(\frac{1}{1+\lambda}\right) > 0$.

Second, when there are very few materialists in the population, the cultural fitness advantage lies with the environmentalists. To see this, note that as $\mu \to 1$, then $v\left(\widehat{t}\left(1\right)\right) \to \alpha + \log\left(\frac{1}{1+\lambda+\theta}\right)$ implying that $\Delta\left(1\right) > 0$ (by (3)). In this limit, creating pollution is expensive, so the signalling rents from being thought of as an environmentalist are sufficient to give a fitness advantage to environmentalists given that we have assumed high enough value of θ to make the tax rate faced by materialists high enough.

These two observations, along with the fact that $\Delta\left(\mu\right)$ is continuous and increasing, mean that there must exist a critical value $\hat{\mu} \in [0,1]$ such that

$$\Delta\left(\hat{\mu}\right) = \chi\varphi\left(\hat{\mu}\right) - \theta\left(1-\hat{\mu}\right)\widehat{c}\left(\widehat{t}\left(\hat{\mu}\right)\right) - v\left(\widehat{t}\left(\hat{\mu}\right)\right) = 0. \tag{11}$$

Moreover $\Delta\left(\mu\right) > 0$ if and only $\mu > \hat{\mu}$.

Finally, we make the weak assumption that[16]

$$1 - 2\mu\left(1-\mu\right)\left(1-\beta\right)g\left(\Delta\left(\mu\right)\right)\Delta_\mu\left(\mu\right) > 0 \tag{12}$$

for all $\mu \in [0,1]$. Since $\Delta\left(\mu\right)$ is increasing, (12) implies that under any reasonable definition of stability, an interior steady state at $\hat{\mu}$ is unstable. Hence, the only stable steady states are at the extremes: $\mu = 0$ and $\mu = 1$. Using these observations and letting $\mu_0 \in [0,1]$ denote the initial value of μ, we have:

Proposition 2 *If $\mu_0 > \hat{\mu}$, the polity monotonically approaches the steady state $\mu = 1$. Otherwise, it monotonically approaches the steady state $\mu = 0$.*

Proof. Use the first-order approximation

$$G\left(\Delta\left(\mu_{s+1}\right)\right) \simeq G\left(\Delta\left(\mu_s\right)\right) + g\left(\Delta\left(\mu_s\right)\right)\left(\mu_{s+1} - \mu_s\right)\Delta_\mu$$

of (9) around μ_s to obtain

$$\mu_{s+1} - \mu_s \simeq 2\mu_s\left(1-\mu_s\right)\left(1-\beta\right)\left[G\left(\Delta\left(\mu_s\right)\right) - \frac{1}{2} + g\left(\Delta\left(\mu_s\right)\right)\left(\mu_{s+1} - \mu_s\right)\Delta_\mu\right].$$

We can rewrite this expression as

$$\mu_{s+1} - \mu_s \simeq \frac{2\mu_s\left(1-\mu_s\right)\left(1-\beta\right)}{1 - 2\mu_s\left(1-\mu_s\right)\left(1-\beta\right)g\left(\Delta\left(\mu_s\right)\right)\Delta_\mu}\left[G\left(\Delta\left(\mu_s\right)\right) - \frac{1}{2}\right].$$

---

[16]In the second term of the condition, $\mu\left(1-\mu\right)$ is maximized at 0.25, while β and g are both smaller than 1. Thus, (12) holds unless Δ_μ is very large.

The denominator on the right-hand side is positive by (12). Since $G(\cdot)$ is increasing with $G(0) = 1/2$, it follows from the term in square brackets that $\mu_{s+1} - \mu_s > 0$ iff $\Delta(\mu_s) > 0$, which requires $\mu_s > \hat{\mu}$. Applying the expression for $\mu_{s+1} - \mu_s$ to $s = 0, 1, 2, ...$ and noting that $\Delta(\mu)$ is monotonically increasing gives the result. ∎

Thus, the model predicts convergence to either $\mu = 1$ or $\mu = 0$ depending on the starting value relative to $\hat{\mu}$. Societies with different starting values can end up on divergent paths towards different steady states, even if their fundamental parameters are the same. Of course, this raises the question where μ_0 comes from. But many societal and cultural factors outside of the model could affect the initial share of environmentalists.

To see the logic behind the proposition, note that the environmental tax goes up as the share of environmentalists increases. This, together with our formulation of the utility benefit from social signalling, $\chi\varphi(\mu)$, creates a complementarity between the fraction of environmentalists, μ, and the payoff difference between environmentalists and materialists, $\Delta(\mu)$. The sign of $\Delta(\mu)$, which governs whether μ is increasing or decreasing, switches from negative to positive as μ increases. Which steady state the economy converges to depends on the starting value μ_0 relative to the critical value $\hat{\mu}$, at which $\Delta(\mu)$ switches sign.

Implied policy dynamics   The model predicts that changing environmental values drive changes in the pollution tax:

$$t_{s+1} - t_s = \theta \left[\mu_{s+1} - \mu_s \right].$$

These policy dynamics have two underpinnings. On the one hand, current tax policy responds to policy preferences, which depend on current values. On the other hand, the evolution of values depends on the expected fitness of environmentalists compared to materialists, which is influenced by expected tax policy.

According to the model, the shares of environmentalists and the strictness of environmental policy should be positively correlated. However, since both policy and values are endogenously dependent on each other, that correlation is not causal in a simple, unidirectional sense.

Comparative dynamics   The analysis allows us to think about how key parameters affect the evolution of values. Observe first that λ does not

impact cultural evolution directly, as it is a common component of $u^\tau(t, \mu)$ for both types τ. The direct effect of θ is to reduce the prospects for cultural change as it lowers the utility of environmentalists, all else equal. But higher λ and θ both lead to higher taxes t, which indirectly raises the prospect for environmentalism.

The effect of χ, social recognition for environmentalists, is unambiguous. A permanent hike in the respect environmentalists perceive to get from others, increase the prospect for an environmentalist culture. This reinforces our earlier claim that virtue signalling can have dynamic effects.

6   Welfare Implications

Undertaking welfare analysis with changing preferences is well-known to be challenging. However, it also raises some interesting issues. Can we really say that a society comprising more or less of one type in the population is better off in a well-defined sense? This section explores that question.

A welfarist approach   Following the conventional (welfarist) approach, we could define a social welfare function as a function of the utilities of both types. We therefore work with a class of additive social welfare functions where

$$W\left(u^e, u^m, \mu\right) = \mu\omega(u^e(\mu)) + (1 - \mu)\,\omega\left(u^m(\mu)\right), \tag{13}$$

and $\omega\left(\cdot\right)$ is an increasing, concave function. (If $\omega\left(\cdot\right)$ is linear, then we have a Utilitarian welfare function.) To work with this welfare objective, we need to assume that payoffs are both measurable and comparable across individuals with different preferences. However, we have already implicitly done so in our dynamics based on $\Delta\left(\mu\right),$ as these suppose that parents are capable of comparing the payoffs of different types when they socialize their children.

The welfare analysis in our model is interesting, in that the standard approach to environmental policy would simply say that the optimal policy is to tax pollution at the Pigouvian level $(t = \lambda + \mu\theta)$ and that welfare cannot feasibly be any higher than at that tax rate. But once we allow for the possibility that values (fractions of types) are endogenous, this is no longer correct. We have to ask whether a society of environmentalists is happier if the environmental externality is completely eliminated in the long run, and not just mitigated via taxation. To answer this question, we need to compare welfare at the two steady states.

From (2) and (13), steady-state welfare in the two cases is

$$W\left(u^{e}, u^{m}, 1\right) = \omega\left(\chi + y\right)$$

and (using $\lambda = \hat{t}\left(0\right)$)

$$
\begin{aligned}
W\left(u^{e}, u^{m}, 0\right) &= \omega(v(\hat{t}\left(0\right)) + y - \left[\lambda - \hat{t}\left(0\right)\right]\hat{c}\left(\hat{t}\left(0\right)\right) \\
&= \omega(v(\hat{t}\left(0\right)) + y).
\end{aligned}
$$

These expressions imply:

Proposition 3 *Welfare in the two steady states depend on parameter values as follows:*

1. *If $\alpha < \chi$, welfare is always higher with $\mu = 1$*

2. *If $\alpha \geq \chi$, there exists a threshold value of λ such that welfare is higher with $\mu = 1$ – i.e., $\chi > v\left(\hat{t}\left(0\right)\right)$ for all λ above this threshold.*

Proof. If $\alpha < \chi$, steady-state utility with $\mu = 1$ is always higher than that in the all-materialists steady state since

$$
\begin{aligned}
\chi &> v(\hat{t}\left(0\right)) \\
&= \alpha + \log\left(\frac{1}{1+\lambda}\right)
\end{aligned}
$$

for all all $\lambda \geq 0$. Then, $\omega\left(\chi + y\right) > \omega(v(\hat{t}\left(0\right)) + y - \lambda\hat{c}\left(\hat{t}\left(0\right)\right))$. Now consider $\alpha > \chi$. Suppose that $\lambda = 0$. Then, (1) implies that $v\left(\hat{t}\left(0\right)\right) = \alpha > \chi$. What if $\lambda > 0$? Because $v\left(\hat{t}\left(0\right)\right) = \alpha + \log(\frac{1}{1+\lambda})$, there exists λ such that $0 < v\left(\hat{t}\left(0\right)\right) < \chi$ so that (3) holds. Then, for large enough λ, the consumption utility of materialists is always lower than the social-signalling utility of environmentalists . Hence for all $\alpha > \chi$, there exists a value of λ for which being an environmentalist yields higher long-run utility than being a materialist. ■

Discussion Proposition 3 makes intuitive sense. In its first case, the social-signalling benefit of environmentalism is so strong that welfare is higher in a population consisting only of environmentalists. The second case is perhaps more interesting. It says that when $\lambda = 0$, materialism yields higher utility,

26

and there is no need for a corrective tax But when λ is higher – and reaches a certain level – high taxation is needed even in a population of materialists. This means that their welfare is lower compared to the welfare level in an all-environmentalist population, which does not consume polluting good c even if the social-signalling benefit from environmentalism is very small.

The combination of Propositions 2 and 3 says that, whichever steady state is long-run optimal society may not converge to it. If $\alpha < \chi$, then beginning with $\mu_0 < \hat{\mu}$ a society will converge to $\mu = 0$, the sub-optimal steady state. The same is true when $\alpha > \chi$, provided that λ is large enough. Convergence to environmentalism in our model requires a mass of environmentalists above a critical tipping point – without other forces supporting environmentalism, this will not happen.

The reasons that (3) plays a key role in this result is that we require that θ is large enough to reduce the payoffs of materialists sufficiently when the proportion of environmentalists grows.   To see this, recall the tax rate increases in proportion to θ.

That preference parameters play a role in these results is a natural feature of utility-based comparison of welfare across types. An underlying feature of the model is that the welfare comparisons are based on steady-state *long-run* utility differences, but cultural evolution depends on expected *short-run* utility differences, between environmentalists and materialists. As we have shown, this can result in a failure to develop welfare-improving environmentalist values if a society begins with a low share μ of such types in the population.

The divergence between long-run welfare and the path driven by the value dynamics could be reinforced if citizens also have biased beliefs, or do not value pollution for other reasons. For example, suppose that citizens underestimate λ by failing to internalize expert opinions on the consequences of pollution. This would be an additional reason why convergence to $\mu = 0$ might be suboptimal. Of course, the direction of the welfare distortion is contingent on the nature of biases and so *a priori* it could go either way.

Failing democratic politics? It is well-known that elections need not deliver welfare-maximizing outcomes. The classic example if the tyranny of the majority and the consequent need to protect minority rights through courts and constitutional provisions. However, in this case the possibility of a suboptimal outcome is due to a lack of commitment.

The issue of the optimal time path for pollution taxes, when cultural change matters, lies beyond the scope of this paper. But we can illustrate the importance of commitment by taking the perspective of very patient policymaker who cares only about long-run welfare. Suppose that a policymaker is unconstrained by politics and can commit to a *constant* tax $\tilde{t} \gtrless \hat{t}(1)$ for all future time-periods. By committing to this tax rate, the policymaker can influence the direction of the value dynamics – particularly when the polity might otherwise converge to $\mu = 0$.

We now show that such commitment can always yield a steady state with $\mu = 1$:

Proposition 4 *If society can commit to a constant tax rate, there exists $\tilde{t}$ such that it will converge $\mu = 1$ for $\mu_0 \in [0, 1]$*

Proof. By (10) the evolutionary dynamics be governed by:

$$
\begin{aligned}
\Delta\left(\mu_{s+1}\right) &= \chi\varphi\left(\mu_{s+1}\right) - \theta\left(1 - \mu_{s+1}\right)\hat{c}\left(\tilde{t}\right) - v\left(\tilde{t}\right) \\
&= \chi\varphi\left(\mu_{s+1}\right) - \theta\frac{\left(1 - \mu_{s+1}\right)}{1 + \tilde{t}} - \alpha - \log\left(\frac{1}{1 + \tilde{t}}\right).
\end{aligned}
$$

Note that for large enough $\tilde{t}$, $\Delta\left(\mu_{s+1}\right) > 0$ for all μ_{s+1} since $\log\left(\frac{1}{1+\tilde{t}}\right) \to -\infty$ as $\tilde{t} \to \infty$. ∎

This result says that commitment can always attain a long-run welfare optimum, in the case where $\mu = 1$ is optimal. This is because with $\tilde{t}$ high enough, it becomes very unattractive to be a materialist. Putting together Propositions 3 and 4, it follows that for λ high enough, commitment gives access to the long-run optimal steady state. This means that the value of commitment in environmental policy can go up, if pollution problems – represented by λ – turn out to be more serious than previously thought.

The mechanism behind this result is that (expected) policy not only influences current payoffs, but also the value dynamics and hence, indirectly, future welfare levels. Whether commitment matters, depends on a comparison of $\tilde{t}$ and λ. When $\tilde{t} < \lambda$, commitment will not affect the future trajectory of a society. However, for small enough λ, we have already seen that the long-run outcome will be $\mu = 0$. The interesting case is thus when $\tilde{t} > \lambda$. With θ large enough, committing to the preferred tax rate of environmentalists, $\hat{t}(1) = \lambda + \theta$, will be sufficient for convergence to $\mu = 1$. With θ smaller, $\tilde{t} > \lambda + \theta$ may be needed.

Commitments vs. horizons In cases where $\tilde{t} > \hat{t}(\mu)$ is needed to secure convergence to the welfare optimum, we can think of this as a failure of the political Coase theorem as outlined in Acemoglu (2003). That the political system is constrained to offer $\hat{t}(\mu)$ may imply that the long-run welfare optimum may not emerge. Of course, this analysis is also related to the classic discussion of time-inconsistency issues in Kydland and Prescott (1977) and Fischer (1980). Politics imposes a certain kind of time-consistency (incentive-compatibility) condition, which may contradict the optimal policy with commitment.

As discussed in a similar context – with endogenous manager types rather than endogenous consumer types – in Besley and Persson (2018), the problem is *not* that the decision-makers have a short horizon. Our earlier result on equilibrium policies would result even if politicians did internalize the future. Without the ability to commit to future policies, politicians as well as everybody else in society at s must take next period's equilibrium policy – which will depend on μ_{s+1} – as given. And, as we have seen, μ_{s+1} is itself based on future expected policy and hence not affected by current policy. However, this may no longer be true if the model had another state variable beyond values, such as the state of the environment.

These observations have more general resonance for thinking about optimal environmental policy with changing values (culture). A forward-looking policymaker, who anticipates changing values, may want to commit to a more draconian policy than would be justified by current preferences. One interesting problem for future work would be to study the optimal sequence of taxes to trade off the welfare of current and future generations, and to compare this sequence with the political-equilibrium sequence. Unlike the constant tax rate in Proposition 4, this sequence would involve trading off near-term losses against long-term gains. Even in cases where values will converge to $\mu = 1$ in the long-run, a policymaker could desire a faster or slower pace of cultural change than would occur when policies are chosen in political equilibrium in each period.

Remedies and enforcement Many (e.g., Stern 2015) have argued that, if a society believes that welfare is higher with environmentalism in the long run, it needs to implement stricter environmental policies than those consistent with the current Pigouvian optimum and the current political equi-

librium.[17] Our analysis gives a normative and non-Paternalistic justification for non-majoritarian policy in an environmental context. The democratic policy process responds to short-term preferences, but this is no guarantee that society will converge to a long-term welfare optimum.[18] This links our analysis to the research on directed technical change that has stressed the value of policies different from the static Pigouvian optimum (Greaker and Midttømme 2016, Harstad 2012).

Let us finally speculate about possible remedies in view of the results in Propositions 3 and 4. One way to think about them would be that they may justify a role for international organizations such as the EU (see Harstad 2016). Such an organization could encourage policies which are not political equilibria for every country, as member countries with high levels of environmentalism would create a positive externality by pushing up environmental taxation.

Another way to think about implementation of a better long-run equilibrium would be to invoke a role for environmental lobby groups. These would be pushing policy away from the conventional Pigouvian optimum. Via a kind of second-best logic, however, this may move the political equilibrium in a desirable direction, once the impact of changing values is taken into account.

Yet another avenue to implementation might run through the judicial process.[19] Courts could adjudicate in favor stricter environmental policies, e.g., if future generations were given rights over current policies. But this would require that politicians who legislate those rights understood that environmentalism is endogenous, and that this legislation could not be repealed for short-run political gain.

---

[17]Mattauch et al (2018) propose a model of endogenous preferences in the context of Pigouvian taxation to support this conclusion.

[18]It would be interesting to extend the model to environmental damages based on the cumulated stocks rather than the current flows of pollution – the right assumption in the case of fossil-fuel emissions and climate change.

[19]Interestingly, one of the goals of the Climate Leadership Council (CLC) – a club of private companies, including the oil giants – is to introduce a $40-a-ton fee on carbon-dioxide emissions. This fee would be a quid pro quo for removing current climate change regulations and protecting companies from federal and state tort liability for historic emissions. (Guardian 2019)

7 Extensions and Open Issues

Our paper is just a first pass at these issues and, at best, contributes to the beginning of a research programme. This section outlines a few possible extensions and issues for future research.

In our framework, an individual is an environmentalist mainly to convey to others (through her consumption decisions) that she cares about a certain cause, rather than because she realistically expects to make a difference by herself. We have not allowed environmentalists to influence social activism or political behavior outside of voting. Social movements and pressure groups – such as the UK "Extinction Rebellion" and the protests among young European people started by Greta Thunberg's school strike – may enhance the collective voice of environmentalists. In a standard setting, this might move policy (increase $\hat{t}(\mu)$ in our simple model). But our approach has also stressed how altered policies might increase the cultural fitness of the environmental movement and change its numbers (raise future values of μ). In Besley and Persson (2019a), we look at social movements among nationalists which enter endogenously and enhance the salience of nationalism. The insights from that paper could be married with those from this paper.

We have also maintained a fixed party structure. However, the emergence of Green parties that seek direct policy influence, particularly by exploiting the coalition structures of proportional representation, may enhance the power of environmentalists and give them further power over policy (increase $\hat{t}(\mu)$ further). This will also have dynamic consequences if it affects the attractiveness of becoming an environmentalist. Green party entry could be modeled using the same approach as Besley and Persson (2019a), who look at endogenous nationalist party entry.

Another interesting extension would be via direct socialization through the education system. We have already seen a link between education and environmental values in the WVS data. This link might reflect a general human-capital effect of reading more about the adverse consequences of human life styles for the planet. But, of course, governments may aim publicly-funded education towards changing values. This is something that could be exploited in both directions (e.g., raising or lowering μ). In this context, there could be a role for forward-looking strategic policymaking by government. Normative analysis would be politically controversial, but a positive analysis would point to the same kind of political constraints as the choice of environmental regulation. In a similar vein, free and independent media, as

well as government-controlled media could influence values directly through their reporting.

Finally, we have worked with a static model of society. But many aspects of the environment – not the least when it comes to climate change – are inherently dynamic. Modeling the interaction of a changing environment and evolving values is a challenging but important task. As more information becomes available and events make the consequences of climate change more salient, we would expect politics, policies and values to respond.

8   Concluding Comments

At the root of this paper is the obvious point that any kind of environmental policy in a democratic society is constrained by what the citizens want. This has been vividly illustrated by recent real-world events. When, in 2018, French President Emmanuel Macron tried to raise the tax on motor fuels – a move that many would describe as environmentally sound – the *Gilets Jaunes* took to the streets of Paris to protest. U.S. President Donald Trump's recent decision to withdraw from the Paris climate-change agreement was very popular among his supporters. Many well-meaning people who are environmentalists would advocate bans on polluting emissions, but ignore the fact that those are far from being a political equilibrium. These political constraints have both static and dynamic consequences.

By building a role for changing values, the environmental politics in our framework moves beyond the standard Pigouvian models of policymaking that dominate the literature. This is timely given the current dynamics of social movements that – as mentioned in the previous section – aim at creating behavioral and policy change. We have illustrated the interplay between forces that affect behavior in traditional ways (via economic and political incentives) and those that influence values.

Social movements stress the importance of declaring a "climate emergency," which is often dismissed as an empty gesture of virtue signalling. But our analysis shows why virtue signalling can indeed be a driver of cultural change. If environmentalists were just miserable about pollution and climate damages and took actions which reduced their own material living standards, environmentalism as a social movement may not catch on. In our model, the positive message of environmentalism as a virtue (a higher χ in our model) is one of the forces that drives environmental values. Declaring

a common cause can increase the perceived virtue of private actions. Exploring such issues further opens up a rich potential agenda on the political economics of policy.

We have used our model to explore a new issue in political economics, namely the interplay between democratic politics and endogenous environmental values. In our setting, policy not only shapes current welfare outcomes, but expected policy also influences future policies via evolving values. By responding to citizens' preferences, politics can create a kind of momentum that may drive multiple steady states. There is no reason to believe that society will converge to the long-run outcome with the highest welfare level. Moreover, this is true even though the political equilibrium picks out a Pigouvian-like optimum based on current preferences (i.e., $\lambda + \mu\theta$). In our model, political preferences are not distorted away from standard welfare objectives – in fact, the probabilistic-voting framework produces a Utilitarian policy outcome.

Although our application is specific, we believe it delivers some takeaway messages of wider significance.

First, policy choices can affect the socialization of types by affecting their cultural fitness. Even though environmentalism is a natural application, we believe that this insight fits many other contexts.

Second, with endogenous values one must consider how social welfare depends upon the composition of population types. But then one has to grapple with the thorny issue whether citizens in some societies have "better values" than others. Our paper has suggested a new way of looking at the welfare economics of environmental taxation. In particular, we show why one may not want to succumb to the usual Pigouvian logic that optimal policy should reflect only current preferences. If society's preferences are themselves endogenous, then long-run desirable policies may be a lot more draconian. However, democratic societies would find it very hard to bring such draconian policies about. That some of today's citizens ignore environmental degradation does not make the problem go away, and the experienced utility of living in a damaged environment may eventually come home to roost in a variety of ways. A similar logic may apply in other policy spheres.

Third, there is no reason to believe that a evolutionary process for values will converge to a long-run social optimum. A system where *relative,* rather than absolute, payoffs drive cultural dynamics will almost always deliver such a conclusion. This is further compounded when politics ensures that current preferences drive policy choices.

Fourth, our framework has highlighted how a political process, where policy is made by current majorities not only affects current outcomes but also emerging values. If we assume – as did Acemoglu and Robinson (2000) – that it is easier to commit to future institutions for policymaking than to future policies, our results suggest that it may be desirable to find institutional frameworks which reduce the responsiveness of environmental policy to current preferences. This may seem to run against one of the assumed virtues of democracy, to deliver policy outcomes that respond to the wishes of the current majority. Yet, many societies routinely delegate policy choices to more far-sighted institutions in other domains, such as central banking.

Finally, economists have been reluctant to embrace cultural dynamics in their analyses. However, our modeling approach suggests that such reluctance could neglect an important aspect of policymaking. Some may find it unpalatable to say that we have to change people's values to fundamentally change the world. But as we have shown, thinking about values is a complement to the conventional approach to optimal policy choices. More generally, our analysis suggests that failing to consider how social and cultural values change in response to policy may give an incomplete account of human progress.

References

[1] Acemoglu, Daron, [2003], "Why Not a Political Coase Theorem? Social Conflict, Commitment, and Politics," *Journal of Comparative Economics*, 31, 620-652.

[2] Acemoglu, Daron and James Robinson, [2000], "Why Did the West Increase the Franchise?", *Quarterly Journal of Economics*, 115, 1167-1199.

[3] Acemoglu, Daron and James Robinson, [2006], "Economic Backwardness in Political Perspective," *American Political Science Review*, 100, 115-131.

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

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

[6] Akerlof, George and Rachel Kranton, [2000], "Economics and Identity," *Quarterly Journal of Economics*, 115, 715-753.

[7] Akerlof, George and Rachel Kranton, [2010], *Identity Economics: How Our Identities Shape Our Work, Wages, and Well-Being*, Princeton: Princeton University Press.

[8] Almond, Gabriel and Sidney Verba, [1963], *The Civic Culture: Political Attitudes and Democracy in Five Nations*, New York, NY: Sage Publications.

[9] Andreoni, James, [1989], "Giving with Impure Altruism: Applications to Charity and Ricardian Equivalence," *Journal of Political Economy*, 97, 1447-1458.

[10] Benabou, Roland and Jean Tirole, [2006], "Incentives and Prosocial Behavior," *American Economic Review*, 96, 1652-1678.

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

[12] Besley, Timothy, [2019], "State Capacity, Reciprocity and the Social Contract," forthcoming in *Econometrica*.

[13] Besley, Timothy and Stephen Coate, [1997], "An Economic Model of Representative Democracy," *Quarterly Journal of Economics*, 112, 85-114.

[14] Besley, Timothy and Stephen Coate, [1998], "Sources of Inefficiency in a Representative Democracy: A Dynamic Analysis," *American Economic Review*, 88, 139-156.

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

[16] Besley Timothy and Torsten Persson, [2009], "The Origins of State Capacity: Property Rights, Taxation and Politics," *American Economic Review*, 99, 1218-1244.

[17] Besley Timothy and Torsten Persson, [2011], *Pillars of Prosperity: The Political Economics of Development Clusters*, Princeton, NJ: Princeton University Press.

[18] Besley Timothy and Torsten Persson, [2018] "Organizational Dynamics: Culture, Design, and Performance," mimeographed manuscript.

[19] Besley Timothy and Torsten Persson, [2019a], "The Rise of Identity Politics," mimeographed manuscript.

[20] Besley Timothy and Torsten Persson, [2019b], "Democratic Values and Institutions," *American Economic Review: Insights*, 1, 59-76.

[21] Bezin, Emeline, [2015], "A Cultural Model of Private Provision and the Environment," *Journal of Environmental Economics and Management*, 71, 109-124.

[22] 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, Elsevier.

[23] Bisin, Alberto and Thierry Verdier, [2017], "On the Joint Evolution of Culture and Institutions," NBER Working Paper No. 23375

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

[25] Bowles, Samuel, Jung-Kyoo Choi, and Astrid Hopfensitz, [2003], "The Co-evolution of Individual Behaviors and Social Institutions," *Journal of Theoretical Biology*, 223, 135-147.

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

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

[28] Coughlin, Peter, [1992], *Probabilistic Voting Theory*, Cambridge: Cambridge University Press.

[29] Dekel, Eddie, Jeffrey Ely, and Okan Yilankaya, [2007], "Evolution of Preferences," *Review of Economic Studies*, 74, 685-704.

[30] Dietz, Thomas, Amy Fitzgerald, and Rachael Shwom, [2005], "Environmental Values," *Annual Review of Environment and Resources*, 30, 335-372.

[31] Downs, Anthony, [1957], *An Economic Theory of Democracy*, New York, NY: Harper Collins.

[32] Druckman, James N. and Arthur Lupia, [2016], "Preference Change in Competitive Political Environments," *Annual Review of Political Science*, 19, 13-31.

[33] Fischer, Stanley, [1980], "Dynamic Inconsistency, Cooperation and the Benevolent Dissembling Government," *Journal of Economic Dynamics and Control*, 2, 93-107.

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

[35] Guardian, [2019], "Microsoft Joins Group Seeking to Avoid Historic Climate Change Lawsuit," https://www.theguardian.com/technology/2019/may/01/microsoft-joins-group-seeking-to-avoid-climate-change-lawsuit

[36] Güth, Werner and Menaham E. Yaari, [1992], "Explaining Reciprocal Behavior in Simple Strategic Games: An Evolutionary Approach," in Witt, Ulrich (ed.) *Explaining Process and Change: Approaches to Evolutionary Economics*, Ann Arbor, MI: University of Michigan Press.

[37] Harbaugh, William, [1998], "What Do Donations Buy? A Model of Philanthropy Based on Prestige and Warm Glow," *Journal of Public Economics*, 67, 269-284.

[38] Harstad, Bård, [2012], "Climate Contracts: A Game of Emissions, Investments, Negotiations, and Renegotiations," *Review of Economic Studies*, 79, 1527-57.

[39] Harstad, Bård, [2016], "The Dynamics of Climate Agreements," *Journal of the European Economic Association*, 14, 719-752.

[40] Inglehart, Ronald, Miguel Basanez, Jaime Diez-Medrano, Loek Halman, and Ruud Luijkx (eds.), [2004], *Human Beliefs and Values: A Cross-cultural Sourcebook Based on the 1999–2002 Values Surveys*, Mexico City: Siglo XXI

[41] Krusell, Per and Jose-Victor Rios-Rull, [1996], "Vested Interests in a Theory of Growth and Stagnation," *Review of Economic Studies*, 63, 301–29.

[42] Krusell, Per, Vincenzo Quadrini, and Jose-Victor Rios-Rull, [1997], "Politico-Economic Equilibrium and Economic Growth," *Journal of Economic Dynamics and Control,* 21, 243-272.

[43] Kydland, Finn E. and Edward C. Prescott, [1977], "Rules Rather than Discretion: The Inconsistency of Optimal Plans," *Journal of Political Economy*, 85, 473-492.

[44] Lindbeck, Assar, Sten Nyberg, and Jörgen Weibull, [1999], "Social Norms and Economic Incentives in the Welfare State," *Quarterly Journal of Economics*, 114, 1-35.

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

[46] Lorenzoni, Irene and Nick Pidgeon, [2006], "Public Views on Climate Change: European and USA Perspectives," *Climatic Change*, 77, 73-95.

[47] McAdam, Doug, [2017], "Social Movement Theory and the Prospects for Climate Change Activism in the United States," *Annual Review of Political Science*, 20, 189-208.

[48] 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.

[49] Meltzler, Alan and Scott Richards, [1981], "A Rational Theory of the Size of Government," *Journal of Political Economy*, 89, 914-927.

[50] Milesi-Ferretti, Gian-Maria and Enrico Spolaore, [1994], "How Cynical Can an Incumbent Be? Strategic Policy in a Model of Government Spending, " *Journal of Public Economics*, 55, 121-40.

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

[52] 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, 351–366.

[53] Nyborg, Karine and Mari Rege, [2003], "On Social Norms: The Evolution of Considerate Smoking Behavior," *Journal of Economic Behavior & Organization*, 52, 323-340.

[54] Oates, Wallace and Robert Portney, [2003], "The Political Economy of Environmental Policy," in *Handbook of Environmental Economics*,

Volume 1, Mäler, Karl-Göran and Jeffrey R. Vincent (eds.), Amsterdam: Elsevier.

[55] Osborne, Martin and Al Slivinski, [1996], "A Model of Political Competition with Citizen-candidates," *Quarterly Journal of Economics*, 111, 65-96.

[56] Persson, Torsten, Gérard Roland, and Guido Tabellini, [1997], "Separation of Powers and Political Accountability," *Quarterly Journal of Economics*, 112, 1163-1202.

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

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

[59] Rogoff, Kenneth, [1985], "The Optimal Degree of Commitment to an Intermediate Target," *Quarterly Journal of Economics,* 100, 1169-1190.

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

[61] Stern, Nicholas H., [2015], *Why Are We Waiting? The Logic, Urgency, and Promise of Tackling Climate Change*, Cambridge MA: MIT Press.

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

[63] Tsebelis, George, [2002], *Veto Players: How Political Institutions Work*, New York and Princeton, NJ: Russell Sage Foundation and Princeton University Press.

[64] Weibull, Jörgen, [1995], *Evolutionary Game Theory*, Cambridge, MA: MIT Press.

**Figure 1: Environmental Attitudes by Birth-decade and Education**
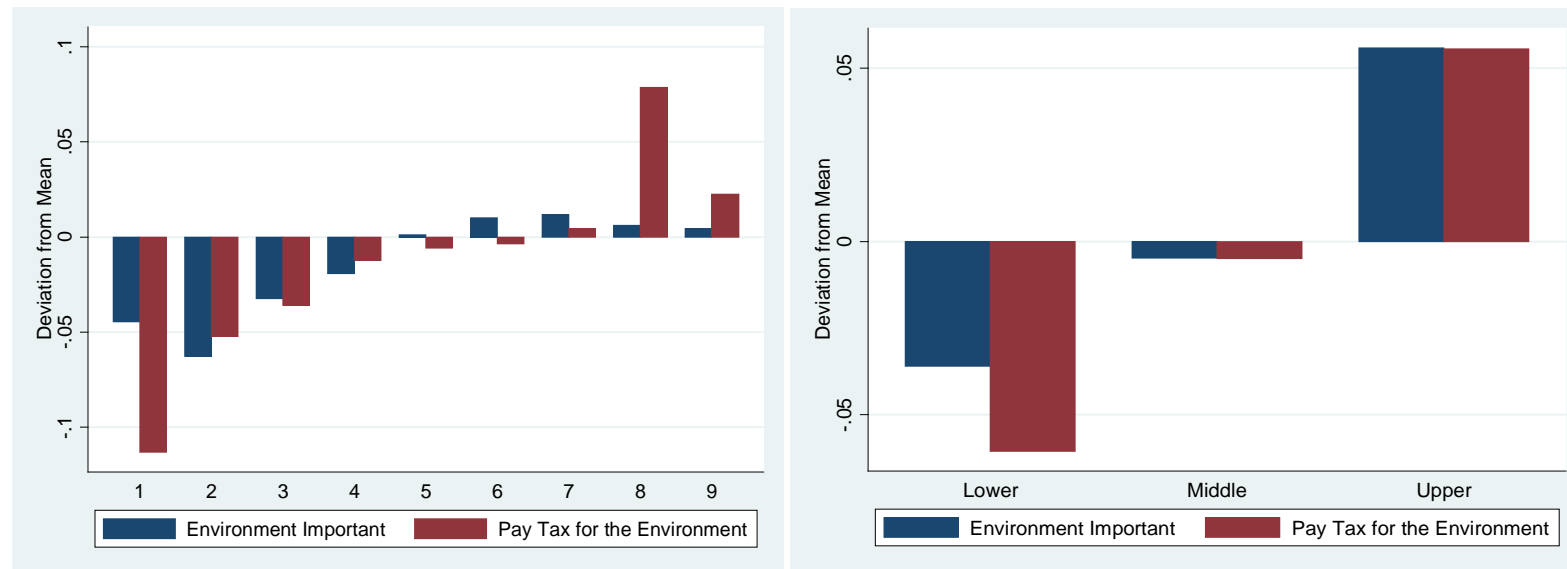
Notes: The graphs show deviations from means for different groups using data from answers to two WVS questions, one where respondents are asked whether the environment is an important policy priority (question B002 in the last WVS wave), and the second whether they "strongly agree" or "agree" with an "increase in taxes if used to prevent environmental pollution" (question B008). The left graph shows average deviations from overall country means among respondents who belong to 10-year cohorts born in the 1910s and onwards, while the right graph shows average deviations from overall country means among respondents in three groups according to their level of education.

Figure 2: Cross-country Variation in Environmental Attitudes
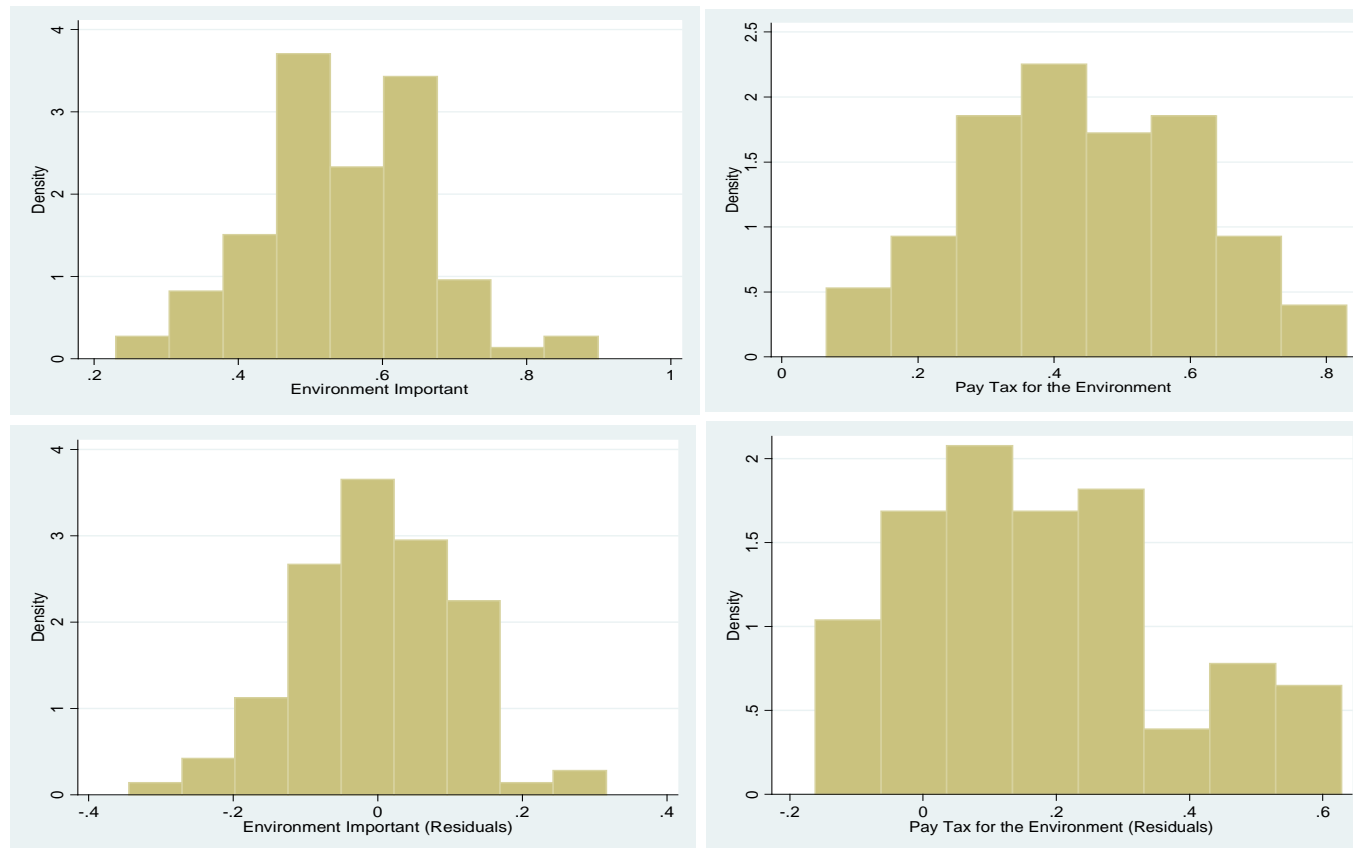
Notes: Each histogram shows the variation across countries in the share of respondents who think the environment is a priority (left two graphs), and whether they support taxes to help the environment (right two graphs) – see the text and Note to Figure 1. The top row shows the average raw share, while the bottom row shows the average share adjusted for individual characteristics. The latter is based on a linear regression at the WVS individual level with an individual dummy on the LHS, and a dummy for gender, ten dummies for income groups, three for education groups, three age bands, and WVS wave dummies on the RHS.

Figure 3: Country-level Correlation Between Environmentalism and Support for Environmental Taxes
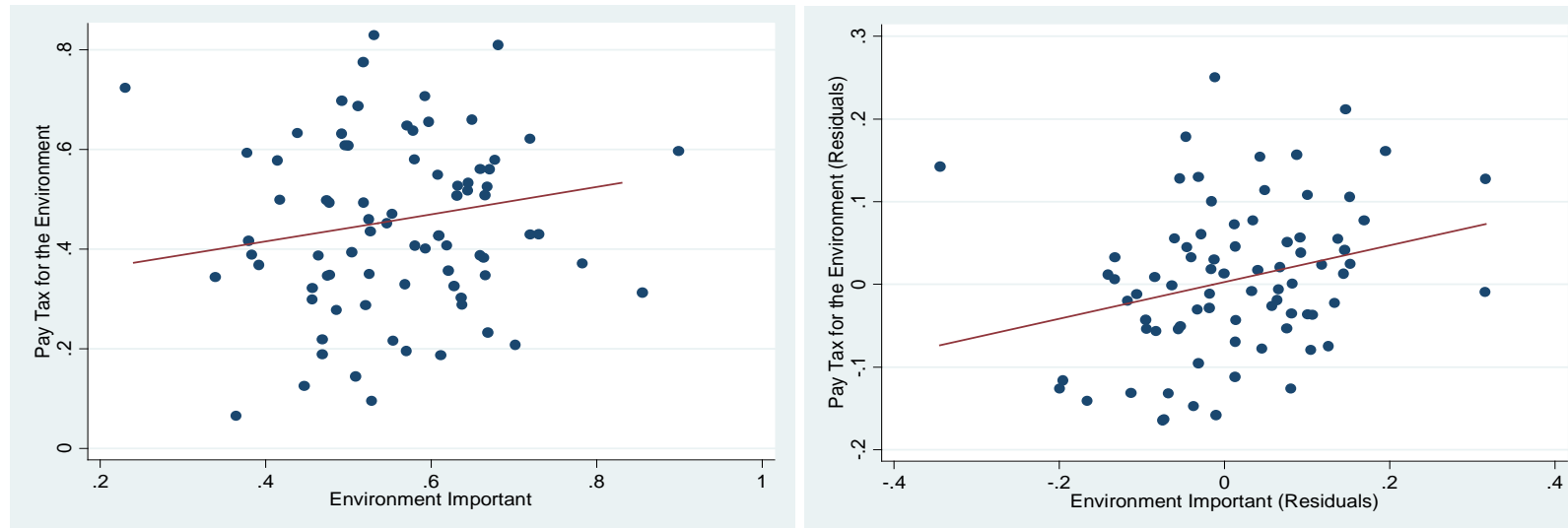
Notes: Both graphs show the correlation in the WVS between country means for holding environmentalist values and for supporting taxes to help the environment. The left graph shows the raw data, while the right graphs show the mean country residuals adjusted for individual characteristics. To define these residuals, we run a linear regression at the WVS individual level with an individual dummy on the LHS, and a dummy for gender, ten dummies for income groups, three for education groups, three age bands, and WVS wave dummies on the RHS.