UC Santa Barbara

UC Santa Barbara Electronic Theses and Dissertations

Title

Studies in Behavioral Environmental Economics

Permalink

https://escholarship.org/uc/item/6qz2k4qd

Author

Vora, Nirav

Publication Date

2024

Peer reviewed|Thesis/dissertation

eScholarship.org Powered by the California Digital Library
University of California

University of California
Santa Barbara

Studies in Behavioral Environmental Economics

A dissertation submitted in partial satisfaction
of the requirements for the degree

Doctor of Philosophy
in
Economics

by

N.R. Vora

Committee in charge:

Professor Ryan Oprea, Chair
Professor Kyle Meng
Professor Sevgi Yuksel

June 2024

The Dissertation of N.R. Vora is approved.

Professor Kyle Meng

Professor Sevgi Yuksel

Professor Ryan Oprea, Committee Chair

May 2024

Studies in Behavioral Environmental Economics

Copyright © 2024

by

N.R. Vora

Acknowledgements

This culmination of my time as a doctoral student would not have been possible without my family, friends, and advisors.

I benefited from the advice of many mentors at UCSB, foremost among them the members of my committee, Dr. Kyle Meng, Dr. Ryan Oprea, and Dr. Sevgi Yuksel. Each of them have been incredibly generous to me with their time, thoughts, and encouragement. Ryan, my chair, contributed so much to my growth as a researcher and also helped me to navigate the challenges of being a graduate student.

I also thank Dr. Ted Bergstrom, Dr. Javier Birchenall, Dr. Daniel Martin, and Mark Patterson for their generosity of time and advice, and many other members of my UCSB family, as well. This of course includes my fellow grad students. There are too many to name, but I particularly thank Jeff, Roberto, and Guglielmo for their feedback, collaboration and friendship.

Finally I cannot thank enough my family for their support over the years, and in particular my parents.

Curriculum Vitæ

N.R. Vora

Doctoral Studies

Ph.D. in Economics, University of California, Santa Barbara, 2024.

Fields of Interest

Environmental and Resource Economics, Microeconomic and Behavioral Theory, Public Policy, Experiments

Job Market Paper

Carbon Pricing and Bounded Rationality

Additional Research

Endogenous Green Preferences

Price, Demand, and Hypothetical Thinking

Preference for Sample Features and Belief Updating

Presentations

University of Melbourne, 2024

Western Economic Association International Conference, 2023

CESifo Venice Summer Institute, 2023

Occasional Workshop in Environmental and Resource Economics, 2023

Economic Science Association Annual Conference, 2022

Teaching and Research Assistant Experience

UC Santa Barbara 2018-present

· Median Mastery Score and Teaching Score: 1 - Excellent

· Head T.A. - Intr. to Macroeconomic Theory

· T.A. - Intr. to Microeconomic Theory, Intr. to Macroeconomic Theory, Intermediate Microeconomic Theory, Game Theory, Econometrics, Intr. to Financial Management

· Guest Lecture: Public Economics (Graduate level)

· R.A./G.S.R. - For Ryan Oprea, Christopher Costello & Kelsey Jack, and Javier Birchenall

· Laboratory for Aggregate Economics and Finance 2019-2024 ditor and contributor

· Environmental Markets Lab (emLab) 2022-2023 Graduate student researcher

Abstract

Studies in Behavioral Environmental Economics

by

N.R. Vora

My dissertation consists of three chapters that use a behavioral economic framework to enrich our understanding of the political economy of environmental policies.

Chapter 1 explores the role of cognitive frictions in the perceived distributional impacts of externality pricing. Despite their desirable properties, carbon prices have received very low public support. Leading explanations attribute the rejection of carbon pricing to preferences and ideology. I show that, instead, much of this rejection is due to mistakes voters make when reasoning about the consequences of such policies. Using laboratory experiments, I show that subjects have specific blind spots in reasoning about a key pricing policy structure, even in abstract tasks that remove scope for preferences and ideologies. Based on these findings, I introduce a new policy that is isomorphic to a leading carbon pricing proposal, but engineered to be robust to these cognitive frictions. The new policy (which I call "carbon penalty and reward") garners significantly more support than the original proposal (known as "carbon fee and dividend"). Using independent cognitive measures, I show that the increase in support is concentrated in people exhibiting systematic difficulty reasoning about a key policy structure.

Chapter 2, which is joint work with Guglielmo Zappalà, details how exposure to environmental policy in the past shapes preferences for them. Low public support has been an obstacle to the enactment of stronger environmental policies. Yet if policies are enacted, support for them may change. Using surveys covering 38 countries around the world, we study the dynamics of environmental policies and individual preferences over

time. Exploiting within-country, across birth-cohort variation in exposure to environmental policy stringency, we document that cohorts exposed to more stringent policies in the past are more supportive of environmental policies at the time of the survey, with the effect largely driven by exposure during a period of early adulthood known as the formative age window. This result even holds when evaluating exposure to a specific policy instrument, environmental taxes. Past exposure to these taxes improves support for them, but not for other environmental policies. This relationship suggests that a society's environmental policy attitudes evolve endogenously, with implications for normative frameworks used in welfare economics.

Chapter 3 evaluates how quick, inattentive decision-making systematically distorts perceptions of the effectiveness of pricing in reducing pollution externalities. When people are asked to make policy choices that affect a good's price – for example, a tax to reduce a negative externality – we traditionally assume that they accurately perceive quantities demanded conditional on price. But failures to consider substitutes and other adaptations to a permanent price shock can result, ex ante, in a perception of inelasticity and policy ineffectiveness. In past surveys, perceptions of ineffectiveness are among the main reasons voters do not support a carbon tax, believing it would fail to cut emissions unless the revenue were spent on offsetting projects. I implement an incentivized survey experiment to test whether increasing time and promoting hypothetical thinking increase subjective perception of responsiveness of demand to price, and I find meaningful, directional effects.

Carbon pricing and bounded reasoning

N. R. Vora*

Abstract

Despite their desirable properties, carbon prices have received very low public support. Leading explanations attribute the rejection of carbon pricing to preferences and ideology. I show that, instead, much of this rejection is due to mistakes voters make when reasoning about the consequences of such policies. Using laboratory experiments, I show that subjects have specific blind spots in reasoning about a key pricing policy structure, even in abstract tasks that remove scope for preferences and ideologies. Based on these findings, I introduce a new policy that is isomorphic to a leading carbon pricing proposal, but engineered to be robust to these cognitive frictions. The new policy (which I call "carbon penalty and reward") garners significantly more support than the original proposal (known as "carbon fee and dividend"). Using independent cognitive measures, I show that the increase in support is concentrated in people exhibiting systematic difficulty reasoning about a key policy structure.

Keywords: Externalities, carbon tax, political economy, cognitive frictions.

JEL Classification: H23, D91, Q52, Q58

*UC Santa Barbara, Department of Economics, 2127 North Hall University of California, Santa Barbara, CA 93106.

1 Introduction

Economists have long argued that directly pricing carbon with a tax is the ideal approach to mitigate climate change. Despite being efficient and effective, carbon taxes are not common and have low support in opinion polls. The leading carbon tax designs recycle the revenue collected from the tax into uniform payments. For many years, economists and policy experts expressed hope that carbon taxes, paired with payments from the revenue, would attract support from the many voters expected to financially benefit(Shultz and Halstead, 2018).[1] But public support for this Fee & Dividend policy, as it is commonly known, appears to be low in many countries, and is frequently rejected by voters most likely to benefit. The leading explanation is that ideology drives these choices, but these policies are not popular among the left or the right (Douenne and Fabre, 2022; Mildenberger et al., 2022).

I show that people make systematic mistakes when reasoning about the consequences of policy structures like Fee & Dividend by analyzing incentivized choices in laboratory experiments. These experiments use simple, abstract tasks that lack scope for preferences and ideologies. I identify the key behavioral blind spots and use them to design a new policy that is isomorphic to Fee & Dividend but robust to key cognitive frictions identified from the lab study. I test this new policy, Penalty & Reward, in a national survey, and independently measure the voters' policy reasoning skills, with incentives, on the abstract

[1]The carbon fee and dividend policy has the endorsement of thousands of economists, including 28 Nobel laureates. For detail, visit econstatement.org. It is the central policy proposal for two major climate policy advocacy organizations, the bipartisan Climate Leadership Council and the Citizen's Climate Lobby. The percentage of households expected to benefit, in consumption terms, varies by study ranging from 56%(Fremstad and Paul, 2019) to 70%(Horowitz et al., 2017) in the U.S., with benefiting households belonging predominantly to lower-income and middle-income quantiles.

tasks used in the lab. Despite being theoretically equivalent, Penalty & Reward is more than 30 percentage points more popular with voters than Fee & Dividend, and this gap in public support is concentrated in people who demonstrate systematic difficulties on the abstract tasks. These difficulties are not associated with political identity, occurring on both the left and the right.

The role of cognitive frictions in voter behavior, even when voters seem to make imprudent choices, cannot be measured from existing data, where individual tastes and information are potentially important confounds. My experimental design isolates the core reasoning required to evaluate real-world pricing policies, but in the absence of these confounds.[2] Despite a simplified, apolitical setting, subjects frequently make reasoning mistakes, and in a systematic way that resembles policy attitudes and outcomes documented throughout the world.

Choices in the lab reveal, among many possible cognitive mistakes that might bias policy selection, one mistake that stands out: Neglect of Distributed Effects. A voter deciding whether to support a carbon tax must consider multiple effects on her welfare, and the Distributed Effect is the proposed tax's welfare impact via distribution of tax revenues. When a voter belongs to a large population, Distributed Effects are a product of aggregate behavior and independent of a voter's own consumption decisions. Neglect of Distributed Effects, or NoDE, systematically biases cost-benefit analyses of policy proposals and drives frequent rejection of privately beneficial taxes and support of privately harmful subsidies. This finding motivates a new policy design that, in a test of support

[2]The setting is, by design, free of the following factors that cloud real-world analysis: variation in available information, risk preferences, time preferences, social preferences and ideological factors.

using a national sample, succeeds in winning public favor.

Distinguishing the role of cognitive frictions from preferences or beliefs is important, as it can tell us whether policies can be modified or clarified to be better received. If voters who reject Fee & Dividend despite being likely to benefit from it are revealing their true preferences – for instance, rejecting the policy on moral or political grounds – then how policy makers respond would be different than if the voters were making reasoning mistakes. Studies analyzing referendums and polls highlight the politicization of climate change. Yet it cannot be said from these studies whether preferences rooted in political ideology are the ultimate cause of puzzling policy rejections or, rather, that partisan voting is a symptom of high cognitive costs of assessing unfamiliar policies.[3]

In recent years, behavioral economics has increasingly pointed to errors and inattention, not preferences, as playing a dominant role in a wide range of choice anomalies.[4] Policy choices require hypothetical thinking and sufficient attention to various costs and benefits. An individual's incentive to expend cognitive effort on policy comprehension is low, but the aggregation of these small decisions is highly consequential (Downs, 1957). Addressing cognitive frictions can thus have clear welfare-improving policy implications. Identifying the troublesome mechanisms of voting errors in the lab can point to new policy designs, better capable of reflecting voters' preferences.

By removing the Distributed Effects that give voters difficulty, Penalty & Reward

[3]See Mildenberger et al. (2022); Douenne and Fabre (2022); Anderson et al. (2023). The studies find strong association of policy support with political identity, and also that inaccurate initial impressions can be difficult to correct with rich supplemental information.

[4]On taxes, health plan selection and other consumer decisions (Abeler and Jäger, 2015; Rees-Jones and Taubinsky, 2016, 2020; Bhargava et al., 2017; Chetty et al., 2009; Lacetera et al., 2012); and on risk, time, strategic, and other decision behavior that deviates from classic choice axioms (Oprea, 2022; Enke et al., 2023a; Banovetz and Oprea, 2023; Nielsen and Rehbeck, 2022).

is designed to be more transparent than Fee & Dividend. A Fee & Dividend can be understood to produce two groups: (i) net receivers, whose dividend exceeds their tax payments; and (ii) net payers, whose tax payments exceed their dividends. These two groups are separated by a threshold at the per capita level of carbon consumption, a fact that requires accurately estimating and aggregating multiple effects. Because this assessment breaks down under NoDE, I design Penalty & Reward to linearly reward below-mean consumers based on how far beneath the population average their emissions reach, and linearly penalize above-mean consumers for emissions beyond the average. Instead of everyone paying taxes and collecting dividends, Penalty & Reward conducts a single transaction. I demonstrate that when penalty and reward rates are set identical to the fee rate in Fee & Dividend, these policies are economically isometric. In a national survey that is representative on several key variables, I measure attitudes toward both policies and vulnerability to cognitive biases, including NoDE, that could explain differences in support for the two policies.

I find dramatically higher support for Penalty & Reward than for Fee & Dividend, showing that carbon pricing can be made much more popular by reducing the scope for cognitive frictions in policy decisions. Whereas 48% of voters express support for Fee & Dividend, 83% support Penalty & Reward. Other factors, like belief in climate science, play a role in my results, but Penalty & Reward's popularity advantage over Fee & Dividend holds across the political spectrum and is strongest among participants who discounted Distributed Effects in the abstract tasks.

This study contributes to our understanding of how support forms for important economic policies. The theoretical and practical advantages of a well-designed carbon tax

make the reasons for its rare adoption worth careful exploration.[5] Studies documenting low public support of carbon taxes date back to Baron and Jurney (1993). Findings in Dresner et al. (2006), Steg et al. (2006), Kallbekken and Aasen (2010)Kallbekken and Sælen (2011),Sælen and Kallbekken (2011), Baranzini and Carattini (2017), Carattini et al. (2017) Douenne and Fabre (2022), and Dechezleprêtre et al. (2022) together suggest that worry about impact to one's household and to low earners, perception of limited effectiveness, and concerns about how revenues would be spent have been consistent factors. Green subsidies tended to be more popular in comparison.

In recent years, Fee & Dividend has become the leading carbon pricing strategy because of its promising qualities. The fact that it is misunderstood is striking. After all, Fee & Dividend is designed to be among the simplest structured policies: a linear tax and a uniform lump sum offset. The perception that it is regressive (Douenne and Fabre, 2022; Mildenberger et al., 2022; Maestre-Andrés et al., 2019) contrasts with economic intuition, as well as with careful work from empirical public economists(Cronin et al., 2019; Metcalf, 2023, 2021; Fremstad and Paul, 2019; Budolfson et al., 2021; Boyce, 2018; Goulder et al., 2019). In France, Douenne and Fabre (2022) measure support for carbon Fee & Dividend at about 10%, despite the fact that 70% of households are expected to directly benefit: only 14% of households think that they will. Only 12% of expected winners who received an informational intervention positively updated their initial view of the policy. In my sample, treatments in which a Fee & Dividend proposal included supplemental sentences

[5]For appealing theoretical qualities, see Pigou (1924); Baumol and Oates (1971); Baumol (1972); Montgomery (1972); Tietenberg (1973); Nordhaus (2001); Baranzini et al. (2017); Cramton et al. (2017); Blanchard et al. (2023). For effectiveness reducing emissions at minimal cost in countries where it has been attempted, see Murray and Rivers (2015); Andersson (2019); Colmer et al. (2022); Metcalf (2021, 2023).

to assist in assessing the policy had similarly small effects, increasing the proportion supporting the policy by at most 3 percentage points.

By quantifying cognitive barriers to rational policy acceptance and designing a policy engineered to overcome these barriers, I contribute to an emerging literature that considers political economy through a behavioral lens. Other papers leverage controlled experiments to characterize how individuals understand and select policies. Among the many important findings from this literature are that consumers prefer taxes applied to a producer than to themselves, even though the tax is passed through(Sausgruber and Tyran, 2005, 2011); policy labels matter and revenue recycling can improve policy acceptability (Kallbekken et al., 2011); subsidies are generally preferred to taxes and smaller policy magnitudes are better tolerated(Cherry et al., 2012); repetition, feedback and learning can correct aversion to corrective taxes (Cherry et al., 2014); and when choosing among games, the equilibrium is not especially salient (Dal Bó et al., 2018). The experimental component of my project builds on this work by attempting to identify and quantify underlying behavioral mechanisms that drive preferences for tax/subsidy policies, and my results have broad implications for public policy and political economy. It is this quantification and comparison of mechanisms that informs the new climate policy design, the focus of this project.

While political orientation, beliefs, and tastes play a role in public support for policies, I find that these explanations are not necessary to generate the sometimes puzzling rejection of useful, beneficial proposals. Neglect of Distributed Effects can generate these decisions. Documenting strong support for Penalty & Reward suggests that people do not necessarily oppose carbon pricing, but that frames matter. A small amount of com-

plexity can play an outsize role in the presence of cognitive and attentional constraints. Addressing these bounds to rationality in policy design can reduce gaps between people's actual priorities and the policies they support.

The paper is organized as follows. Section 2 details possible mechanisms for biased policy evaluation. Section 3 details the design and results of the Diagnostic Experiment, the purpose of which is to generate and analyze a data set of abstract problem-solving behavior that can narrow down, among possible mechanisms, a primary driver of policy selection mistakes. That primary driver, Neglect of Distributed , or NoDE, motivates the a new policy design, Penalty & Reward, introduced in Section 4, that is isomorphic to Fee & Dividend and robust to NoDE. Section 5 details a survey experiment to elicit the popularity of Penalty & Reward and Fee & Dividend in a national sample, while gathering explanatory variables for policy support including cognitive factors. Section 6 presents the results of the survey; Section 7 elaborates on the implications of these results; and Section 8 concludes.

2 Cognitive biases in policy evaluation

In this section, I detail computational errors that can systematically bias evaluation of a policy like Fee & Dividend. I am careful to separate these from informational imperfections. A voter with perfect information about everyone's demand function can fail to use that information properly. For example, a voter may misjudge the size of what he loses in consumer surplus or what he gains from the dividend. A voter may fail to properly account for the private welfare impact of changes to consumption that follow a

policy introduction. Quantifying the relative role of such mistakes, among several others, has policy design implications. Ex ante, we don't know where reasoning may break down and therefore cast a broad net. Collecting data of policy selection decisions while carefully varying the scope for many potential mistakes could narrow down a list of possible errors to a mistake of highest importance in political outcomes.

To best distinguish cognitive factors from informational ones, I present a framework with perfect information: voters assess a proposal with full knowledge of the market price for carbon, their own carbon demand function, and the aggregate carbon demand function.[6] Yet assessment requires attending to the information and calculating net impact. A voter i has a demand function for carbon (i.e. fossil fuels): $q_i(p)$, and inverse demand function $p_i(q)$. Aggregate demand is $Q(p)$.[7] Let p_0 represent the price at present and p_τ represent the price after the introduction of a proposed tax τ. Voters belong to a very large population of size N. For this example, we can begin with the rational voter's estimation of the change to her own consumer surplus:

$$CS_{i,\tau} - CS_{i,0} = \int_0^{q(p_\tau)} p_i(q) - p_\tau \; dq \; - \int_0^{q(p_0)} p_i(q) - p_0 \; dq = - \int_{p_0}^{p_\tau} q_i(p) \; dp$$

Next we consider the welfare effect resulting from any change to the uniformly distributed effect D, which we can think of here as the introduction of a dividend from the

[6]For simplicity, I abstract from demand for various fossil fuels by simply referring to a market for carbon; I assume carbon is competitively produced at price p, and thus $\tau = p_\tau - p_0$; utility is quasilinear with respect to carbon. After voting on a proposal, it is either enacted or the status quo is maintained. Then, with full knowledge of the proposal's success or failure, voters make consumption choices.

[7]Throughout the paper I assume demand functions that are continuous and nonincreasing in p.

tax revenue:

$$D_{i,\tau} - D_{i,0} = \frac{\tau\ Q(p_\tau)}{N} \ - \ 0$$

This results in the near-term impact of the policy, which does not include a voter's private value for the policy's reduction to climate externalities:

$$V_i(FD) = -\int_{p_0}^{p_\tau} q_i(p)dp \ \ + \ \frac{\tau\ Q(p_\tau)}{N}$$

For now, I analyze the voter's decision problem considering only near-term consequences, which the literature points to as a key driver of negative carbon policy attitudes. For the large subset of voters unconcerned about climate change, for whatever reason, the near-term assessment represents their entire assessment problem. Later, I will discuss the inclusion of climate externalities to the assessment problem.

2.1 Assessing a Policy for the Near Term

The impact to consumer surplus of a tax (compensating variation) results in a negative **Consumer Effect**, or $-\int_{p_0}^{p_\tau} q_i(p)dp$. The policy's impact on the distribution of tax revenue, which I will call the **Distributed Effect**, is positive in this example, $+\frac{\tau Q(p_\tau)}{N}$. A voter might estimate these two effects with cognitive biases. For example, the Consumer Effect may seem more visible, familiar, or consequential to a voter than Distributed Effects. The fact that Distributed Effects accrue to a voter passively through the aggregate behavior of other consumers may cause it to be discounted and made peripheral to a voter's focus. Behavioral economics in recent years has documented many instances of

imperfect salience and difficulties with hypothetical thinking, both of which could play a role in voter assessment of the Distributed Effect. While a tax's Consumer Effect is salient and painful to all, its Distributed Effect could seem remote and difficult to quantify without experience, and a voter could misjudge the relative and absolute sizes of these two typically opposing private welfare effects.

Another key mistake that voters might make is to fail to account for how a policy introduction will change consumption. For example, a failure to reason hypothetically could cause a voter to evaluate a policy's impact as if the quantity demanded would be unaffected by the policy's introduction or account partially for the full magnitude of change.[8] I will refer to the calculation of a policy's welfare impacts, holding quantity demanded fixed, as the Direct Effect. I call the welfare adjustment that accounts for response of consumption to new prices as the Equilibrium Effect, borrowing the term from Dal Bó et al. (2018). The near-term value function for Fee & Dividend:

$$V_i(FD) = -\tau q_i(p_0) \; + \; \int_{p_0}^{p_\tau} q_i(p_0) - q_i(p) dp \quad + \; \frac{\tau \; Q(p_0)}{N} \; - \; \frac{\tau(Q(p_0) - Q(p_\tau))}{N}$$

The first term, $-\tau q_i(p_0)$ is the **Direct Effect**, the policy's impact to a voter's consumer surplus holding his consumption choices fixed. The next term, $\int_{p_0}^{p_\tau} q_i(p_0) - q_i(p) \; dp$ is the **Equilibrium Effect**, which corrects the Direct Effect to account for the voter's best consumption response. A tax raises prices, making some of the voter's carbon

[8] An elastic consumer who buys 20 gallons of gas a week, when asked to consider the minimum weekly payment they would be willing to accept in exchange for a permanent $1 increase in price might reflexively (and often mistakenly) think $20.

consumption unprofitable to continue. An Equilibrium Effect removes the losses incurred from unprofitable (and unlikely) purchases that would have occurred if the voter fixed consumption to the level prior to the policy's introduction. The Direct Effect and the Equilibrium Effect together compose the Consumer Effect. A failure of equilibrium thinking could also affect fair estimation of the policy's Distributed Effect, which results(almost entirely) from others' consumption behaviors. The term $\frac{\tau\ Q(p_0)}{N}$ is the **Distributed Direct Effect**, or the dividend that would result from holding aggregate quantity demanded fixed. The term $\frac{\tau(Q(p_0)-Q(p_\tau))}{N}$ represents the **Distributed Equilibrium Effect**, the adjustment to the Distributed Effect that reflects the shift in aggregate consumption attributable to the tax.

In tax (and subsidy) policies, the elements described above are additively separable. For a more detailed discussion of these four key effects, see Appendix A.1.

$$V(\Delta\tau \mid \tau_0) = \underbrace{\textit{Direct Effect} + \textit{Equil Effect}}_{\textit{Consumer Effect}} + \underbrace{\textit{Dist Direct Effect} + \textit{Dist Equil Effect}}_{\textit{Distributed Effect}}$$

Because of the errors that I've laid out, voters could have blind spots or simply misweight the four key effects. For a rational voter, ι and λ in Equation 1 and κ, η, ν, and θ in Equation 2 would equal 1; and biased voters would deviate from 1.

$$\hat{V}(\Delta\tau \mid \tau_0) \;=\; \iota\, \textit{Consumer Effect} \;+\; \lambda\, \textit{Distributed Effect} \tag{1}$$

$$\hat{V}(\Delta\tau \mid \tau_0) = \kappa\ \textit{Direct Eff} + \eta\ \textit{Equil Eff} + \nu\ \textit{Dist Direct Eff} + \theta\ \textit{Dist Equil Eff} \quad (2)$$

A data set of voters' choices across a sufficiently varied set of policy scenarios would allow us to identify the presence and relative importance of the biased weightings on the effects in Equations 1 and 2.

There are other cognitive phenomena that may bias policy selection. A voter may find the presence of a tax or subsidy complex, wish to avoid the cognitive cost of evaluating them, and thus reject them: I call this **Policy Aversion** and its opposite, Policy Bias. A voter may be averse to deviations from the status quo, resulting in **Default Bias** (or its opposite, default aversion).[9] Policy Aversion and Default Bias can result from avoidance of cognitive costs. Under another cognitive phenomenon, **Loss Aversion With Mental Accounts**, or LAMA, voters perceive the total effect of a policy with exaggeration to the magnitudes of a policy's negative parts. This biases cost-benefit analysis by always overweighting a policy's true costs. For a tax proposal (like Fee & Dividend), the Consumer Effect is typically negative and the Distributed Effect positive, resulting in a weight larger than 1 to the Consumer Effect, in that instance. For a subsidy proposal (or a proposal to lower a tax), the Distributed Effect is negative and would get a subjective weight larger than 1. These misweightings can lead to mistakes. I provide a deeper discussion of LAMA and its potential role in policy selection in Appendix section A.3.

[9] In Section 3, controlling for Policy Bias and Default Bias motivates the balanced inclusion of subsidy proposals and default policies where the subject must evaluate a proposal to remove it.

The literature highlights pessimism on near-term household impact as a driver of Fee & Dividend rejection, and associates pessimism with ideology. Between biases on Consumer Effects, Distributed Effects, or Equilibrium Effects, and Policy Aversion, Default Bias and LAMA, the results of a diagnostic experiment will allow us to narrow our attention on a key mistake to motivate an improved policy design.

2.2 Climate Externalities

Carbon taxes also produce private welfare impacts in the long run through the reduction in climate externalities. Unlike the calculation of immediate financial impact, it is difficult to determine whether a voter's circumstances, beliefs, or risk, time, and social preferences could rationalize their estimate of a proposal's "Climate Effect" on their welfare. To best isolate cognitive costs from information and taste-based explanations, I seek to detect dominated financial decisions that cannot be rationalized, and thus focus on mistakes in reasoning through policies' near-term effects.

I now add a term, $\phi_i(Q(p_0)-Q(p_\tau)$, to the value function, where ϕ_i is a voter's private benefit from reduction to climate externalities, for each unit carbon reduced.[10]

$$V_i(FD) = -\int_{p_0}^{p_\tau} q_i(p)dp \;\; + \;\; \frac{\tau \; Q(p_\tau)}{N} \;\; + \;\; \phi_i(Q(p_0) - Q(p_\tau))$$

Climate Effects share many attributes of Distributed Effects, and may be subject to similar discounting or neglect. To a voter, they both result (almost entirely) from the consumption choices of others and may be less salient or familiar. Regardless, for any

[10]We can think of as a certainty equivalent that incorporates beliefs, and risk and time preferences.

$\tau \geq 0$ and $\phi_i \geq 0$, the term $\phi_i(Q(p_0) - Q(p_\tau) \geq 0$. Then, when a voter rejects a financially beneficial carbon tax, the private total cost of their selection mistake is at least as large as the cost of their near-term financial mistake, and probably larger.

3 Diagnostic Experiment

A voter's choice can appear to conflict with the voter's best interests. Could these by driven by mistakes in reasoning about policy? The Diagnostic Experiment is designed to: (1) study whether people make mistakes in abstract choice problems in the lab that test the reasoning required for policy selection in the world; and (2) test which of the potential mechanisms discussed in Section 2 drive these policy selection mistakes.

3.1 Design

My goal is to create a simple decision environment capable of generating a rich data set in which potential cognitive errors of Section 2 can be identified and quantified. This requires sufficient variation in economic environments and policy options. Each subject encounters 30 of these variations. For details, see Appendix B.4 and **??**.

In the decision environment, there is a fictional good. The design uses real resale values for these fictional objects to induce preferences, shown as a demand function on the left of the screen in Figure 1. The economy's participants include **one human subject** and **1,000 robots**. Each robot has its own preferences, shown on the right of the screen in Figure 1. The 1,000 robots have identical preferences. An experimental

currency unit, "francs," are used to scale up the numbers.[11]

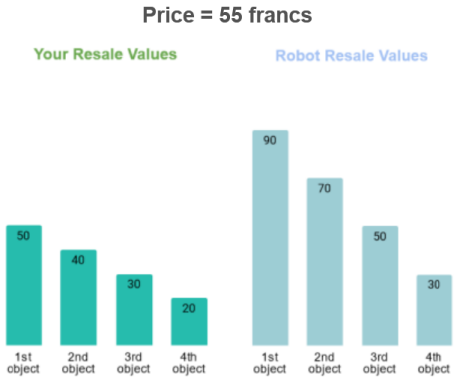


Figure 1: An example of a status-quo environment with prices and resale values.

The supply of objects is inexhaustible with an exogenous base price. Any purchaser, human or robot, can buy up to four. Purchaser budgets are always set sufficiently high to never constrain a human or robot's capacity to buy objects, up to four. Subjects must complete comprehension checks to ensure they understand how to determine privately optimal purchasing behavior for purchasers in this environment.

Subjects are then informed that all purchasing decisions will be made automatically for them and for the robots to maximize each purchaser's profit from purchasing objects, taking into account price and any other relevant facts in the environment. Hence, no privately harmful purchasing mistake can possibly be made by any purchaser, human or robot. The only choice the subject (human) must make is a Policy Choice (which, in the lab, is always referred to as a "Feature Choice" to maintain abstraction and avoid evoking any public policy connotations). The human subject is the sole voter in this environment. The Policy Choice outcome will be in place and taken into account when automatic purchasing choices are made on behalf of each purchaser. A Policy will mirror a

[11] Twenty francs exhange for $1. Resale is described as happening automatically after purchase.

tax or subsidy, with a lump-sum offset that preserves the proposal's revenue neutrality.[12]



Figure 2: Consider this simple example involving a tax-and-dividend proposal. It results in no loss of consumer surplus, and a dividend gain via recycled aggregate tax revenue of 20 francs. Therefore, the optimal choice is to "Accept."

The Policy Choice page always comes after subjects have viewed details of the status-quo environment, and those status-quo details are visible while the subject makes the Policy Choice. Importantly, the Policy Choice varies by round, and does not always involve a choice of introducing a Policy. The Policy Choice can be either: (1) to accept or reject a new proposed Policy, or (2) to accept or reject a proposed reversal of an existing Policy. The value to the subject of the Policy Choice is contingent on the automatic,

[12]Taxes are offset with positive dividends, subsidies with negative dividends. Uncertainty, temporal factors, trust and impact on other humans are absent. This simplifies policy selection into a perfect-information computation task. Verbiage of policy, politics, or pollution, too, are absent.

optimal purchases that are made subsequently on behalf of each purchaser. Remember, no purchaser, human or robot, can buy an object if that purchase is harmful to the purchaser. Similarly, no purchaser, human or robot, can fail to buy an object if that purchase is beneficial to the purchaser.

Alongside the binary Policy Choice decision, I elicit a subjective valuation for the Policy Choice decision, relative to the rejected choice, using a slider bar task. I ask subjects what it is worth to them, in francs, to get their choice, if the certain alternative is to implement the rejected choice. Subjects are told that an accurate report is in their best financial interest, and if they wish are able to click to read details of incentive compatibility, based on the Becker-DeGroot-Marshak mechanism.[13]

The primary mistake that a subject can make is to choose a dominated option in the binary Policy Choice. A secondary mistake is to misvalue a policy. All subjects who err in the binary Policy Choice automatically err in subjective policy valuation. However, a subject who makes a wrong Policy Choice and values their choice highly is "more wrong" than a subject who makes a wrong Policy Choice and values their choice close to zero.

For the 30 randomly assigned rounds in which subjects participate, I vary the scope for Consumer Effects, Distributed Effects, Equilibrium Effects, Distributed Equilibrium Effects, LAMA Bias Effects, Policy Bias Effects, and Default Bias Effects. This is accomplished by varying parameters by round, including policy type (tax and subsidy), policy

[13]In these details, subjects are told that their vote will be implemented with 95% probability, but that with 5% probability the rejected option will be implemented and the subject will be given a 200 franc supplement to their budget. A number between 0 and 200 will be drawn, and if that number exceeds their valuation, they keep the 200 francs and the rejected Policy Choice is implemented. If the drawn number is beneath their valuation, they value of the drawn number is deducted from their 200 francs supplement, and the subject's Policy Choice is implemented.

magnitudes, status-quo conditions, prices and demand functions.[14] See Appendix section ??, for full details about the experimental parameters, by round. For implementation details see Appendix section B.2.

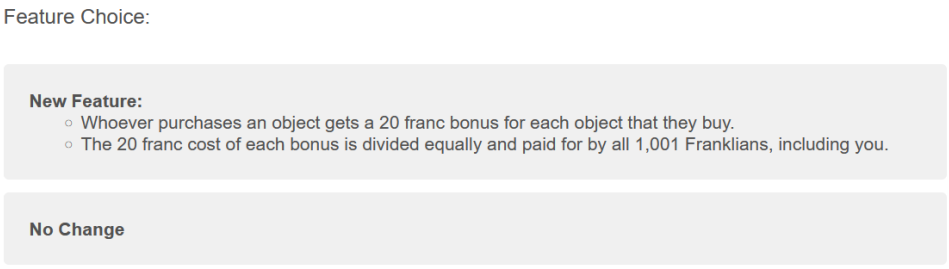


Figure 3: An example of a subsidy policy proposal.

Including subsidy proposals is valuable for a number of reasons. Discounting the Distributed Effect can result in the rejection of privately optimal taxes. However, Policy Aversion and LAMA could each produce rejection of taxes, as well. To remind you, policy-averse voters tend to reject policies because they do not wish to assess their welfare implications while LAMA voters tend to reject policies because they exaggerate the costs of any policy relative to the benefits. Including subsidy proposals in many rounds allows for instances that separate the behavioral responses of subjects who discount Distributed Effects from those who are policy averse or have LAMA.

3.2 Measuring behavioral biases

The willingness to pay responses allow us to measure how subjects utilize the experiment round's key parameters to form a subjective valuation for the policy proposal. Below, I assign behavioral weights to evaluation elements detailed in Section 2. Due to additive

[14]Subjects are aware that one round will be randomly chosen to determine their pay.

separability in the two-parameter case (Consumer Effect and Distributed Effect) and the four-parameter case (Direct Effect, Equilibrium Effect, Distributed Direct Effect, and Distributed Equilibrium Effect), expressed in Equations 1 and 2 of the previous section, the marginal impact of any of these four effects on a voter's subjective value is the behavioral weight. Therefore a linear regression of participant i's value for proposal j on the four objective effect parameters of proposal j will return an unbiased estimator of the subjective weights, given a sufficiently large set of proposals J that are linearly independent. Deviation from the rational benchmark of 1 on the estimated coefficient for the parameters represents a misweighting.[15]

Among the cognitive mechanisms that must be assessed is Loss Aversion with Mental Accounts, or LAMA, when a voter aggregates a perceived total effect of a policy move with he exaggeration to components that have negative welfare effects. It might be that the Consumer Effect of a proposal is negative, as is commonly the case with proposals to introduce a tax or remove a subsidy. It might be that the Distributed Effect is negative, as is commonly the case when the proposal introduces a subsidy or removes a tax. To exaggerate costs when summing the costs and benefits of a proposal is distinct from the consistent misweighting of the Consumer Effect or Distributed Effect.[16] (Remember NoDE is a consistent misweighting of Distributed Effects). The regression equation below can measure and control for LAMA, absorbing the systematic exaggeration of costs, by including as an explanatory parameter the magnitude of either of the two effects that represent costs in that round. This variable is called "Negative Term Magnitude," and

[15]For the population-wide estimation, I cluster observations at the subject level.

[16]LAMA is conditional overweighting on either the Consumer Effect and the Distributed Effect if its sign is negative.

estimating a nonzero coefficient on it represents deviation from rationality.

To capture Default Bias or Default Aversion (where subjects systematically avoid or prefer change from the status quo) the regression includes a control term that I call "Deviation Magnitude,"($\mid \tau_1 - \tau_0 \mid$) that measures in absolute terms the amount of change that is being proposed in the Policy Choice. Similarly, a term called "Policy Magnitude"($\mid \tau_1 \mid - \mid \tau_0 \mid$) is included to control for Policy Bias or Policy Aversion.[17]

Using subjective valuations of policy proposals, we can estimate the weights in Equation 1, along with the other cognitive factors:

$$\begin{aligned} V_{ij} = \alpha \ + \ \beta_1 \ \textit{Consumer Effect}_j \ + \ \beta_2 \ \textit{Distributed Effect}_j \\ + \gamma_1 \ \textit{Neg. Term Magn}_j + \gamma_2 \ \textit{Policy Magn}_j + \gamma_3 \ \textit{Dev. Magn}_j + \epsilon_{ij} \end{aligned} \tag{3}$$

Similarly, we can perform a regression to assess the subjective weightings on subeffects shown in Equation 2:

$$\begin{aligned} V_{ij} = \alpha + \beta_1 \textit{Dir Eff}_j + \beta_2 \textit{Equil Eff}_j + \beta_3 \textit{Dist Dir Eff}_j + \beta_4 \textit{Dist Equil Eff}_j \\ + \gamma_1 \ \textit{Neg. Term Magn}_j + \gamma_2 \ \textit{Policy Magn}_j + \gamma_3 \ \textit{Dev. Magn}_j + \epsilon_{ij} \end{aligned} \tag{4}$$

In Equations 3 and 4, I compare the β coefficients to the rational benchmark of 1 and the γ coefficients (Negative Term Magnitude, Policy Magnitude, and Deviation Magnitude) to a rational benchmark of 0. Estimating the cognitive factors together and comparing them to their rational benchmarks will will reveal which factors are more

[17] Remember that subsidies are represented as taxes(τ) with negative values.

consequential. The degree of deviation from the rational benchmark corresponds proportionally to money that subjects failed to capture, in expectation.[18]

Neglect of Distributed Effects can be measured nonparametrically by evaluating the binary feature choices made by subjects. Correct choices will be harder to make for subjects with NoDE when the magnitude of the Distributed Effect (DE) exceeds the magnitude of the Consumer Effect (CE). Comparing correctness rates when $\mid DE \mid > \mid CE \mid$ to when $\mid CE \mid \geq \mid DE \mid$ or the overall correctness rate can illustrate whether NoDE plays a role in policy selection in these simplified, abstract experiments.

3.3 Results and Discussion

Subjects frequently make mistakes in policy decisions, even in this simplified and incentivized perfect-information setting. Of binary choices, 26% were dominated. For rounds in which the magnitude of Distributed Effects exceeded the magnitude of Consumer Effects, 39.7% of binary choice responses were dominated; for rounds in which Distributed Effects were smaller than Consumer Effects, 20.7% were dominated.[19] Thus, rounds with high scope for selection mistakes due to discounting of Distributed Effects had twice the error rate of rounds with low scope for such mistakes.

In contrast, among rounds with high scope for errors driven by discounting of Equi-

[18] Across the rounds that subjects can receive, some rounds lack the scope for any scope for Equilibrium Effects. In other words, the proposal does not change any consumer's best consumption response. Estimating Specification Equation 3 solely on these rounds offers another strategy for isolating weightings on Consumer and Distributed Effects without confounding Equilibrium Effects.

[19] Random guessing would produce an expected 50% correctness rate. A simple heuristic would work well for the 87% of rounds in which there are distinct high demand and low demand types for the human subject and robot counterparts. Among these questions, a heuristic in which a low-demand type rejects subsidies and accepts taxes and a high-demand type accepts subsidies and rejects taxes produces dominant binary choices 94% of the time, or 81% of the total available rounds.

librium Effects, mistake rates were only slightly higher than the overall rate. When Equilibrium Effects exceed Direct Effects, the binary choice error rate was 28%, versus 26% over all rounds. Equilibrium Effects appear to play secondary role in generating policy selection and valuation mistakes, which is confirmed in the regression results.

Table 1 distills the extent of deviation from rational benchmarks, by behavioral factor, in policy valuation responses, and is drawn from the regression results in Appendix Tables 4 and 5. It clearly demonstrates the outsize role of Distributed Effects in generating valuation errors. Each dollar of Equilibrium Effects (when separate from Distributed Effects) is undervalued by 22%, on average (column 3). Loss Aversion with Mental Accounts (LAMA) mistakes generate similar financial costs as Equilibrium Effects for subjects (column 3). More than twice those losses are generated by errors undervaluing Distributed Effects, which are discounted by 59% when separate from Equilibrium Effects. An alternative way of calculating the weight on Distributed Effects, independent of Equilibrium Effects, is shown in column 1, in which the weighting on Distributed Effects is calculated in the subset of problems that lack scope for Equilibrium Effects. There, we see a similar magnitude of discounting by half (-0.500) relative to the rational benchmark.[20] Note that, in contrast, Consumer Effects and Direct Effects are valued consistently at near the rational benchmark.

On an individual subject level, the distribution of the coefficient on the Distributed Effect reflects wide variation in Appendix Figure 8, with most subject responses reflecting

[20]In column 3, we can see full neglect, on average, for Distributed Equilibrium Effects. This discount is moderately larger than the effect we might expect if the discounting of Distributed Effects and Equilibrium Effects were independent of one another ($\beta_4 < \beta_2 * \beta_3$). The welfare implications of the full neglect of this term are ambiguous in the presence of the other discounting, but would costly when voters encounter subsidy proposals. It combines two effects we are interested in independently, and therefore I focus on measures of the effects when independent of each other.

a weight of between 0 and the rational benchmark of 1.

Table 1: Deviation from rational benchmarks for subjective valuations

	Deviation from rational benchmark		
	Equil Eff Absent	All Rounds	All Rounds
Consumer Effect	0.113	−0.134	
Direct Effect			0.049
Equilibrium Effect			−0.221*
Distributed Effect	−0.500***	−0.790***	
Distributed Direct Effect			−0.590***
Distributed Equil Effect			−1.011***
LAMA	−0.043	−0.258**	−0.225**
Policy Bias	0.187	−0.200**	0.042
Default Bias	−0.427	−0.227	−0.210

Notes: Columns 1 and 2 above result from estimating Equation 3 (first on the subset of rounds where Equilibrium Effects are absent; second on the full set of rounds) and are drawn from Columns 4 and 2 of Appendix Table 4. Column 3 result from Estimating Equation 4 across all rounds and are from Appendix Table 5 Column 2. rounds. Given the regression specifications, the rational benchmark for Consumer Effect and Distributed Effect and their subeffects are 1. Rational benchmark for LAMA, Policy Bias, and Default Bias are 0. *p<0.1; **p<0.05; ***p<0.01.

Across all specifications, the fact that Distributed Effects are weighted meaningfully lower than Consumer Effects indicates that the dominated binary choices characterized earlier in this section are to be expected. These mistakes resemble the stylized policy attitudes documented in the policy survey literature: taxes are under-favored and subsidies are over-favored (see Section 1).[21] Having removed preference confounds and controlled for alternative behavioral factors, the results suggest attentional and computational explanations have a role to play in puzzling voting behavior out in the world, most significantly via Neglect of Distributed Effects, or NoDE. For the remainder of the

[21]The implications of this finding likely extend beyond preference for the special case studied here, taxes and subsidies with uniformly distributed offsets as in Fee & Dividend. Neglecting Distributed Effects could be a cognitive element that underlies fiscal illusion (Puviani, 1903; Buchanan and Wagner, 2013; Dell'Anno and Mourao, 2012); consumption responses to tax changes that deviate from Ricardian equivalence (Bernheim, 1987); and preference for predistributive policies over redistributive policies (Kuziemko et al., 2023).

paper, NoDE will be the primary cognitive factor of interest.

4 Penalty & Reward: Theoretical Properties

I design a policy, Penalty & Reward, that removes the scope for neglect of Distributed Effects, or NoDE. First, I present a behavioral model for NoDE, that can explain the dominant mistake in the Diagnostic Experiment and motivates the design of Penalty & Reward. Next, I present the Penalty & Reward policy and explain why it is outcome equivalent, in theory, to Fee & Dividend. Lastly, I detail how a voter with NoDE would perceive Penalty & Reward and Fee & Dividend differently.

4.1 Neglect of Distributed Effects (NoDE)

A voter i evaluates a tax (or subsidy) proposal as follows:

$$\hat{V}_i(\Delta\tau \mid \tau_0) \;=\; \textit{Consumer Effect} \;+\; \lambda_i \; \textit{Distributed Effect}$$

where $\lambda_i = 1$ for a rational voter, and $0 \leq \lambda_i < 1$ for a NoDE voter.

A NoDE voter would partially or completely discount the value of a dividend when evaluating a Fee & Dividend proposal.

4.2 Penalty & Reward: Definition

The Penalty & Reward policy imposes either (i) a linear tax on units consumed above a threshold in a period, for any consumer whose total consumption exceeds the threshold,

or (ii) a linear reward for on the difference between the threshold and the quantity consumed for any consumer whose total consumption is beneath the threshold. The level of the threshold can change from one time period to another.

4.3 Demand and welfare equivalence to Fee & Dividend

In this section I will show that, when the penalty and reward rates are identical to tax rates of Fee & Dividend, and when the Penalty & Reward threshold is set at the population mean level of consumption for the period, these two policies theoretically result in equivalent demand and welfare.

A consumer's utility function under a general policy that taxes carbon at rate τ and provides a lump-sum payment D:

$$V_i^{FD}(q_{carb}, q_{clean}) = U_i(q_{carb}, q_{clean}) + m_i - p_{carb}q_{carb} - p_{clean}q_{clean}$$

$$\underbrace{- \tau q_{carb} + D}$$

Tax & Payment

The consumer maximizes utility with respect to two goods, q_{carb} and q_{clean}, taking endowment m_i, prices p_{carb} and p_{clean}, tax rate τ, and payment D as given. Fee & Dividend is an instance of such a policy, where D is an equal per capita share of the aggregate revenue.

Under Penalty & Reward, a consumer faces a similar maximization problem, taking m_i, p_{carb}, p_{clean}, a Penalty & Reward threshold, A, and penalty rate γ and reward rate γ

as given:

$$V_i^{PR}(q_{carb}, q_{clean}) = U_i(q_{carb}, q_{clean}) + m_i - p_{carb}q_{carb} - p_{clean}q_{clean}$$

$$\underbrace{-\ \gamma\ (\mathbf{1}_{q_{carb}>A})(q_{carb} - A)\ +\ \gamma\ (1 - \mathbf{1}_{q_{carb}>A})(A - q_{carb})}$$

Penalty & Reward

$$= U_i(q_{carb}, q_{clean}) + m_i - p_{carb}q_{carb} - p_{clean}q_{clean}$$

$$-\ \gamma q_{carb} +\ \gamma A$$

For any tax rate τ and lump-sum payment D in a Fee & Dividend regime, when Penalty & Reward rate $\gamma = \tau$ and Penalty & Reward threshold $A = \frac{D}{\tau}$, consumer utility functions under Penalty & Reward and Fee & Dividend are equivalent. Thus demand $(q^*_{carb}, q^*_{clean})_i = \underset{q_{carb}, q_{clean}}{argmax}\ V_i(q_{carb}, q_{clean})$ and indirect utility $V_i(m_i, p_{carb}, p_{clean})$ are equivalent under each policy.[22] [23]

When $D = \tau\ \bar{q}_{carb}$, where $\bar{q}_{carb} \equiv \frac{\sum_i q_{carb,i}}{N}$, the mean level of consumption across the population, we have Carbon Fee & Dividend policy, with its property of revenue

[22] The equivalence of demand for each consumer implies equivalence of aggregate demand and aggregate damage from externalities. I've demonstrated this as a partial equilibrium result, taking prices as given under each policy, but it extends to general equilibrium. The policy, as framed here, does not directly affect supply. By showing that demand is equivalent under each policy, prices, too, must be equivalent for any aggregate supply function, assuming a unique equilibrium. Also, a variation of Fee & Dividend that taxes producers instead of on consumers should have no impact in equilibrium prices, demand or welfare, due to equivalence of incidence.

[23] Doing this welfare comparison using a static model requires the assumption that time lags either either do not exist, do not matter, or do not differ between these polices. If dividends can be paid arbitrarily frequently, for instance, we have no need to consider how a lag between first-period tax payments and first-period dividend could result in a welfare difference than if tax payments and dividends occurred simultaneously. In the survey, I address this by offering an initial prepaid expected dividend, which should remove any potential advantage to Penalty & Reward.

neutrality.[24] The threshold for the equivalent Penalty & Reward policy is $A = \bar{q}_{carb}$, which is therefore also revenue neutral. The sum of transfers for each policy can be expressed as follows:

$$\sum_i (-\tau q_{carb,i} + D) = \tau \sum_i (-q_{carb,i} + \bar{q}_{carb}) = 0$$

$$\sum_i (-\gamma q_{carb,i} + \gamma A) = \gamma \sum_i (-q_{carb,i} + \bar{q}_{carb}) = 0$$

For either of the two proposals, a policy choice might be made with uncertainty about the consumption choices of others. This uncertainty would then apply directly to the size of the dividend in Fee & Dividend. Under Penalty & Reward, this uncertainty would apply to the consumption threshold required to maintain revenue neutrality. (The threshold separates fee payers from reward receivers.) In both Fee & Dividend and Penalty & Reward, the uncertainty pertains only to the level of $\bar{q}_{carb}$, and is thus identical uncertainty under both regimes. Therefore, if aggregate consumption information is imperfect, risk aversion will not result in any differences between this two policies in incentives, quantities demanded, or welfare.

4.4 Behavioral Distinctions in How Policies Are Perceived

A voter must evaluate how a proposed policy will change his or her welfare, to determine whether and how intensely to support the proposal. Proposals can have Consumer Effects, which are more direct and transparent to voters, and Distributed Effects, which

[24] As population size N becomes arbitrarily large, the correlation between relationship between $\bar{q}_{carb}$ and $q_{carb,i}$ approaches 0. Therefore, a consumer chooses a bundle effectively takes $\bar{q}_{carb}$ as an exogenous constant.

are less direct and opaque to NoDE voters. Below I revisit the framework from Section 2 for policy valuation.[25]

As shown earlier in this section, the unbiased value for both the Penalty & Reward and Fee & Dividend policies are equivalent. However, Penalty & Reward consists of only transparent effects, lacking a mechanism that produces any Distributed Effect. Therefore, a NoDE voter evaluating a Penalty/Reward policy would come to the same value as an equivalent non-NoDE counterpart. I denote the subjective value function of a NoDE voter i with $\tilde{V}_i$.

$$V_i(FD) = V_i(PR) = \tilde{V}_i(PR) = -\left(\int_{p_0}^{p_\tau} q_i(p)dp\right) \; + \; \frac{\tau \; Q(p_\tau)}{N} \; + \; \phi_i(Q(p_0) - Q(p_\tau))$$

The first term represents the change in consumer surplus, with $q_i(p)$ as the individual demand function for carbon. The middle term is the dividend, with $Q(p)$ representing the aggregate demand function for carbon. The final term is the private benefit from the change in the externality, with ϕ_i representing an individual's private damage per unit damage. Prices p_0 and p_τ correspond to effective prices of carbon without taxes and with tax rate τ, respectively.

In comparison to the transparency of Penalty & Reward, Fee & Divdend consisting of a direct, familiar Consumer Effect and a relatively opaque Distributed Effect. As a result, aggregating these two effects when evaluating Fee & Dividend, a NoDE voter discounts the Distributed Effect.

[25] Again for simplicity, I abstract from demand for various fossil fuels by simply referring to a market for carbon; I assume carbon is competitively produced at price p; and I assume quasilinear utility.

$$\tilde{V}_i(FD) = -\left(\int_{p_0}^{p_\tau} q_i(p)dp\right) \; + \; \lambda_i \frac{\tau \; Q(p_\tau)}{N} \; + \; \phi_i(Q(p_0) - Q(p_\tau))$$

While the difference in the value of these two policies is 0, a NoDE voter, with $0 \leq \lambda_i < 1$, assuming that both the tax rate and resulting aggregate consumption quantity is greater than zero.

$$V_i(PR) - V_i(FD) = 0$$

$$\tilde{V}_i(PR) - \tilde{V}_i(FD) = \tau(1 - \lambda_i)(\bar{Q}(p_\tau)) > 0$$

An alternative way to describe this is that under both policies there is an actual and perceived consumption threshold above which you lose from the policy, and below which you gain from the policy. In reality, the threshold is the same under both policies. This threshold is transparent under Penalty & Reward. But under Fee & Dividend, it is opaque; for subjects with NoDE, the subjective threshold separating fee payers from reward receivers is lower, suggesting that more consumers will lose from the policy than is actually the case.

There are two alternative or additional cognitive mechanisms, False Correlation and Mental Budgeting that could result in greater support for Penalty & Reward than for Fee & Dividend that I test in the survey and am able to control for in the evaluation of my primary NoDE hypothesis. I provide background on these mistakes and detail the theoretical implications in Appendix A.4.

5 Survey-Experiment Design

The survey-experiment has two main purposes: (1) to assess whether people indeed more strongly support Penalty & Reward, a carbon pricing strategy designed to overcome cognitive frictions identified in the Diagnostic Experiment, than Fee & Dividend; and (2) to understand whether the cognitive mechanism of NoDE can indeed help to explain any difference of popularity between the two policies.

Accomplishing this requires survey participants both to assess policies and to participate in incentivized tasks akin to those shown in the Diagnostic Experiment. I designed the survey to also include questions that test for alternative or additional mechanisms and to measure other voter characteristics that might drive policy attitudes.

I will detail the three pieces of the survey: **Policy Elicitation**, in which participants rate the two carbon policies; **Abstract Policy Task**, which tests for NoDE, a cognitive factor that biases policy perceptions; and a **Questions** section that includes markers for two additional cognitive biases and other relevant factors. Afterward, I detail supplemental treatments, along with delivery and sample characteristics.

5.1 Policy Elicitation

Half of subjects will see the policy section first. The order in which the two policies are displayed is randomized by individual, and that order remains constant throughout the survey, to avoid any confusion.

Please read the following policies completely. You will answer several questions about both of them.

Carbon fees and rebate payments: The government would tax greenhouse gas emissions by $100 per ton. This is roughly equivalent to a $1 tax per gallon of gasoline. The government gathers the money collected from the new emissions tax and uses all of the money to send every living resident an equal rebate payment. Before the first week of the tax, one estimated rebate payment is paid to all people in advance. Then, at the end of each week, the government sends a rebate payment that evenly divides the week's tax collection.
ss

Carbon penalties and rewards: The government would tax greenhouse gas emissions by $100 per ton only on a person's emissions that exceed the average individual level for the week. If your emissions are above average, then the higher you emit, the more tax you pay. Also, the government rewards $100 per ton for how much below average a person emits. If your emissions are below average, the less you emit, the bigger your reward. Taxes and rewards are roughly equivalent to $1 per gallon of gasoline and paid weekly.

Participants then provide four ratings for each policy: **Overall Support**; **Emissions**, or how successful a policy will be in curbing climate emissions; **Your Finances**, or how the policy will affect your household financially; and **Low Income**, or how the policy

will affect low income individuals. (For full text of each rating prompt, see Appendix Section C.1.1.) The order of these questions is randomized (as shown in Appendix Figure 10). For each rating question, subjects respond using a slider bar with an integer from -100 to +100, and the text of the policy is always visible. (For an image, see Appendix Figure 11.)

5.2 Abstract Policy Task

Using decision environments from the Diagnostic Experiment, I measure NoDE for each survey participant. Half of subjects will encounter this section first. In each of 15 rounds, participants are offered a choice: to adopt a proposal or to maintain the status quo. A proposal can take one of two forms: (1) accept or reject a new proposed policy, or (2) to accept or reject a proposed reversal of an existing policy.[26]

Every subject encounters the same 15 rounds, but in an order randomized by subject. The 15 rounds vary in policy type (tax or subsidy), status quo (policy or no policy), and optimal choice (accept or reject). Most importantly, the policies vary in the magnitude of the change in Consumer Effect and the Distributed Effect. This is accomplished by varying parameters across the 15 rounds, such as prices, policy proposals, demand functions assigned to the participant, and demand functions assigned to her robot counterparts.[27]

To minimize the survey length and difficulty, I do not elicit willingness to pay as I did in the lab experiment. To measure neglect of Distributed Effects, I first calculate the

[26]Participants earn a $ 1 bonus if a they made the better of the policy choices in a randomly drawn round. Consistent with the Diagnostic Experiment, the Abstract Policy Task never uses the word "policy," and always uses the word "feature." Features mimic taxes and subsidies without evoking the connotations of public policies.

[27]For a detailed look at all 15 of the rounds, please see Appendix C.3.

coefficients on the Consumer Effect and the Distributed Effect when they are used jointly as regressors, with controls, on each participant's 15 binary outcome choices.[28] Then for each participant a calculate the difference between these coefficients.[29]

$$\mathbf{1}_{ij}^{Accept} = \beta_{0,i} + \beta_{1,i}\ Consumer\ Effect_j + \beta_{2,i}\ Distributed\ Effect_j + X_j'\gamma_i + \epsilon_{ij}$$

$$NoDE\ Score_i = (\beta_{1,i} - \beta_{2,i}) * \kappa\ (a\ scaling\ constant)$$

In essence, the NoDE Score is the difference between the marginal effect of the Consumer Effect and the Distributed Effect in a linear probability model of binary choices, by individual.

5.3 Questions section

Finally, subjects complete a questionnaire. All subjects encounter the questionnaire after the Policy Elicitation and Abstract Policy Task sections. The questionnaire includes state of residence, a demographic trait that Prolific does not provide.

There are six questions to assess factors that might predict support for redistributive climate change policies broadly: political orientation score, belief in climate science, faith in how government spends money, impact of a carbon tax on own job income, belief in appropriateness of redistribution from high earners to low earners, and personal emissions

[28] Controls include terms for LAMA, Policy Aversion and Default Bias, shown in Section 3.2

[29] I multiply this difference in coefficients for each subject by a fixed scaling constant $\kappa = 40$ to allow my regression coefficients to not be large.

relative to the population average. See Appendix section C.1.2 for more details.

Included are two questions to test alternative or additional mechanisms that would produce differences in support between Fee & Dividend and Penalty & Reward: Mental Budgeting and False Correlation mistakes. For theoretical background, see Appendix Section A.4. With the Fee & Dividend policy displayed in both: (1) Mental budgeting asks: "*Do you agree with the following statement? Because carbon taxes in the "Carbon fees and rebate payments" policy make fossil fuels more expensive than before, the payments that people get back from the policy will be spent on their fossil fuel budget.*" with Agree as 1 and Disgaree as 0; (2) False Correlation asks: "*True or False? Under the policy, I should not change my behavior because the more fossil fuel tax I pay, the more I'll get back in payments.*" with True as 1 and False as 0.

5.4 Supplemental treatments

The elements discussed thus far comprise the main study. But I test variations in the Fee & Dividend policy text that add information meant to improve receptivity and have been proposed or tested in the literature.[30] The goal is to compare support for Penalty & Reward to support of these "assisted" versions of Fee & Dividend. [31]

The first treatment, which I refer to as COMPUTE, tries to clarify the likely financial impact of Fee & Dividend by appending this statement to the end of the policy description: *Individuals who emit below the average level will receive a higher rebate payment*

[30]Informational interventions like these have had modest or disappointing results in the past (Mildenberger et al., 2022; Douenne and Fabre, 2022).

[31]Each of the two treatments enroll 50 participants, were designed after the collection of the baseline study results, and were not pre-registered.

than what they paid in emission taxes. Individuals who emit above the average level will spend more in taxes than they will receive in the rebate payment.

The second treatment, which I refer to as EXPERT, tries to clarify the impact of Fee & Dividend by instead adding this sentence to the standard Fee & Dividend description: *Experts say that most below-average emitters will benefit by gaining money from this arrangement and that most above-average emitters will lose money.*

5.5 Robustness to single/sequential policy presentation

Providing two policies simultaneously offers participants an opportunity to convey relative preference or indifference between Fee & Dividend and Penalty & Reward. Policy proposals can also be presented alone, without alternative options, and for that reason I conduct a robustness study in which support is measured one policy at a time. Participants convey their support for a randomly assigned policy from the pair, either Fee & Dividend or Penalty & Reward. (While reading and answering rating questions about it, participants are unaware of an alternative policy they will be asked to rate subsequently.) A between subjects comparison of ratings for first-shown policies would be equivalent to a study in which each subject was shown only one policy.

5.6 Delivery and sample characteristics

The surveys are delivered online to 678 United States participants, 18 or older, via Prolific. Participants know nothing about the survey's topic prior to beginning it. Because political representativeness is potentially important to a fair answering of the re-

search questions, subjects are enrolled in a manner that delivers 40% Republicans, 40% Democrats, and 20% independents for the entire study and its subsets, including each treatment and robustness sample. The final sample is representative of national age, gender, race and geographic characteristics.[32]

As measured in their Abstract Policy Task performance, participants were much likelier to show NoDE than not, and there was substantial heterogeneity. The extent of bias varies widely by individual as shown in the distribution of NoDE Scores in Appendix Figure 12. Notably, there is no meaningful correlation between NoDE Score and political orientation.[33] Overall, participants answered 56% of the binary choice questions correctly, with a 46% correctness rate when the Distributed Effect's magnitude exceeded the Consumer Effect's magnitude and a 68% correctness rate when the Consumer Effect's magnitude exceeded the Distributed Effect's magnitude. The structure of these outcomes resembles results of the Diagnostic Experiment. Approximately 13 percent of participants answered the False Correlation question incorrectly, and 60 percent answered the mental budgeting question incorrectly.

[32]Prolific provides age, sex, race and party affiliation, though some participants decline to provide each detail. I supplement these details with my state-of-residence question. As shown in Appendix Figure 12, there is broad distribution of ages, with mean 42.7. Survey distribution occurred at times of day that enabled daytime participation in both Eastern, Central and Pacific time zones. There was broad geographic participation, with the 43 most populous states represented. 47.0% of subjects identify as male, and of the subjects who provided race, 63.1% identified as White, compared to 59.3% at the population level. In addition 13.8% described themselves as students, and 11.1% as unemployed. Of the 347 enrolled in the main study and two supplemental treatments the base pay rate was $4. Upon entering the study did they learned of the possibility of earning a $1 bonus through the Abstract Policy Task. Participants took on average 22 minutes and earned $4.54, on average. For the shorter robustness study, 331 subjects spent about 5 minutes and earned $1.

[33]The Pearson correlation coefficient (-0.07) between NoDE Score and Left-Right Score is not statistically meaningful (p-value of 0.167) and indicates slightly lower NoDE bias on the right. This suggests that even if voters on the right are averse to real taxes, they are very slightly likelier to support an abstract-setting tax when it is optimal than those on the left.

6 Survey-Experiment Results

Participants overwhelmingly favored Penalty & Reward, with 71 percent of participants rating it higher than Fee & Dividend and 17 percent rating it lower in the Baseline study, which excludes the supplemental information treatments. Users were asked to rate their opposition or support for a policy on a -100 to +100 scale. Their numerical response is referred to throughout as "support intensity."

Figures 4 and 5 offer a deeper look at support for the policies across participants and across subgroups, like low NoDE Score, high NoDE Score, right of center politically, and left of center. In Figure 4, individuals across the board tend to favor Penalty & Reward over Fee & Dividend, with an average Penalty & Reward Premium – or the difference in support intensity for the two isomorphic policies – of 49 points.[34] The effect is greater among individuals showing more pronounced NoDE in the Abstract Policy Task. The Penalty & Reward Premium is 33 points larger for individuals with a NoDE score above the median (0.17) than for individuals with scores below it.

A high-NoDE individual's rating for Fee & Dividend is -13.4, or 19.3 points below the rating given by a low-NoDE individual. Meanwhile, the average support for Fee & Dividend among those on the left is 28.2 points higher than average support on the right. For Penalty & Reward the left-right gap is 24 points. This 19.3 point gap in support for Fee & Dividend between high-NoDE and low-NoDE individuals is more than two-third the size of the gap in support for Fee & Dividend between right- and left-

[34]I conduct a robustness exercise on Policy Elictation, in which I presented 331 participants with policies sequentially. I compared the ratings of the policy shown first, the results of which are shown in Table 10. Broadly speaking, the gap in support for the two policies is robust but moderated in size. I elaborate further in a discussion in the Appendix with Table 10.

leaning individuals. This is just one piece of evidence that cognitive factors may deserve nearly as much attention as political identity in driving support for these policies. On average, participants on the political right support Penalty & Reward to a greater extent than participants on the left support Fee & Dividend, by a margin of 16.5 points. The difference is highly significant (Wilcoxon rank sum p-value: 0.0044).

For Figure 5, I transform baseline support intensities into a binary variable of support or non-support.[35] Of the subgroups presented, Fee & Dividend enjoys majority support only among the left-leaning and low-NoDE groups, while Penalty & Reward is supported by a solid majority of every subgroup. The patterns in the proportion measure are reflective of the mean intensity scores in Figure 4.[36]

To test the role of the cognitive factors on the difference of support between Penalty & Reward and Fee & Dividend, I employ the following regression, the results of which are shown in Table 2:

$$\begin{aligned} \textit{Penalty Reward Premium}_i = \beta_0 \; + \; & \beta_1 \textit{NoDE Score}_i \\ & + \beta_2 \mathbf{1}(\textit{Correl Error})_i + \beta_3 \mathbf{1}(\textit{Mental Budget})_i + \beta_4 \mathbf{1}(\textit{Exp First})_i \\ & + \beta_5 \mathbf{1}(\textit{Compute Prompt})_i + \beta_6 \mathbf{1}(\textit{Expert Prompt})_i + X_i'\gamma + \epsilon_i \end{aligned} \tag{5}$$

[35] I define support as providing a rating of 5 or more on the -100 to +100 scale.

[36] To formally test hypotheses, I conduct Wilcoxon rank sum tests shown in Table 9. The first column tests my pre-registered hypothesis that Penalty & Reward would be supported higher than Fee & Dividend, and the results are strongly significant. The second column is one of several tests for my second pre-registered hypothesis, that the Penalty & Reward Premium (the difference in support compared to Fee & Dividend) was higher for above-median NoDE-score types than for for below-median types. The outcome of this test, too, is strongly significant. As mentioned earlier, I also find that the higher distribution of support for Penalty & Reward among right-leaning participants than the support of Fee & Dividend among left-leaning participants is statistically significant.

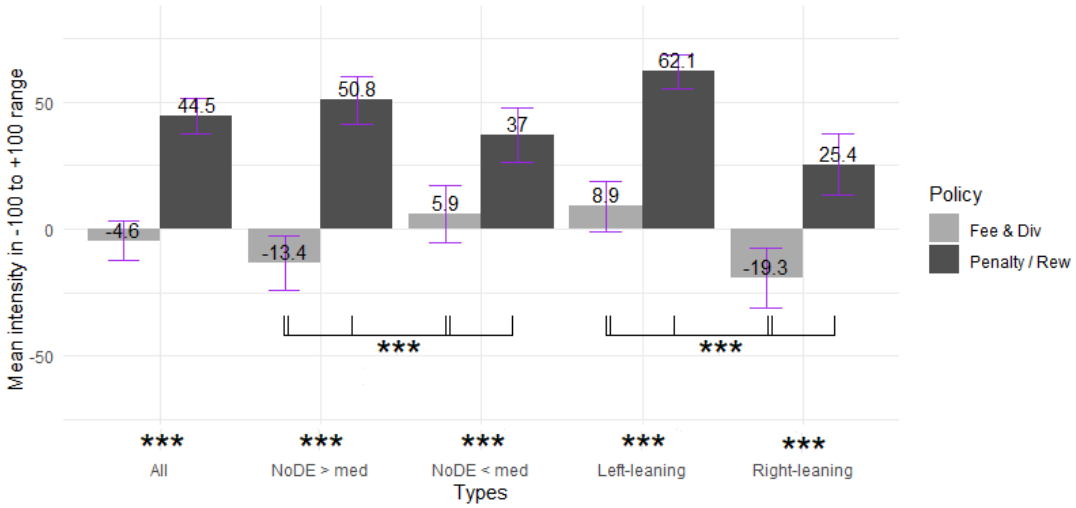


Figure 4: Mean support for policies by group in the baseline study, shown with 95 percent confidence intervals. Note that "bias" refers to NoDE score.

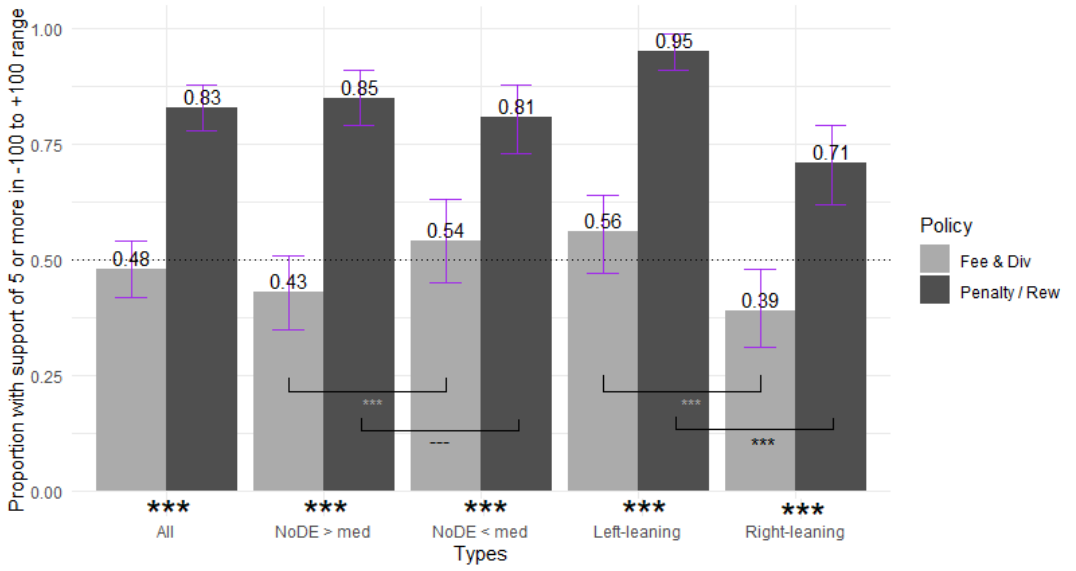


Figure 5: Proportion in support of policies by group in the baseline study, shown with 95 percent confidence intervals. Note that "bias" refers to NoDE score.

The results shown in Table 2 suggest that cognitive factors play an important role in the Penalty & Reward Premium. Neglect of Distributed Effects (NoDE), as well as the Mental Budgeting mistakes play a consistent predictive role in the hypothesized direction. (Remember that a subject making a mental budgeting mistake believes that dividends will allocated to carbon-intensive consumption.) A full point of the bias score measuring NoDE would be expected to reduce the Penalty & Reward Premium by an estimated 30 to 50 points, depending on the specification.[37] One standard deviation of NoDE score (0.3 points) predicts 0.2 standard deviations in the Penalty & Reward Premium (14.4 points).

The study included information treatments intended to help participants reason about the private welfare impact of a policy. The regression in Table 2 suggests that the information treatments (which added words to the policy text) do not appear to play a significant role in reducing (or increasing) the Penalty & Reward Premium. Interestingly, doing the Abstract Policy Task prior to expressing one's views on the two real-world policies reduces the Penalty & Reward Premium, although with weak statistical significance across specifications. However, the magnitude of the effect is comparable to the better of the two information treatments. Experience thinking about the policy selection abstractly was helpful in reducing an irrational gap in support between the policies. This could suggest that the formation of intuition from experience can improve performance better than instruction via information treatments, in line with recent findings (Esponda et al., 2020). It is important to distinguish that participants in this study did not receive any feedback about their performance on the Abstract Policy Task until the entire survey

[37] The range for NoDE scores observed in the sample is 1.5 points.

Table 2: Effect of Behavioral Factors on Differences in Policy Support

	Dependent variable: Penalty & Reward Premium		
	Support for P&R minus support for F&D		
	(1)	(2)	(3)
Constant	35.767***	4.109	−12.877
	(8.245)	(21.035)	(35.353)
NoDE Score	35.310**	29.689**	48.185***
	(14.664)	(15.028)	(17.279)
False Correlation error	14.812	16.090	19.612
	(12.267)	(11.896)	(14.010)
Mental budgeting error	16.179**	15.609**	17.880**
	(7.462)	(7.487)	(8.798)
Survey order: experiment first	−13.988*	−14.349*	−18.597*
	(7.966)	(8.092)	(9.740)
"Compute" prompt	−13.092	−16.164*	−2.544
	(9.823)	(9.793)	(11.489)
"Expert" prompt	−4.498	−0.869	1.624
	(10.248)	(11.038)	(13.486)
Controls:			
Support factors		X	X
Demographic/employment			X
Observations	347	347	253
Mean outcome	45.176	45.176	48.091
R^2	0.067	0.097	0.154
Adjusted R^2	0.050	0.062	0.089

Notes: The table above results from estimating Equation 5, which predicts Penalty & Reward premium, i.e. each subject's support for Penalty & Reward minus their support for Fee & Dividend. If Fee & Dividend is favored, the premium is negative. Column 1 estimates over the full sample with the key cognitive explanatory factors: NoDE Score, False Correlation error, and Mental Budgeting error, and Survey order: Experiment First. "Compute" prompt and "Expert" prompt represent the information treatments meant to enhance Fee & Dividend that a subset of participants received. Column 2 includes a set of controls derived from answers to survey questions that would predict support for policies in general (Relative emissions, Belief in climate change, Left-right score, Belief job benefits from carbon tax, Belief job hurt from carbon tax, Trust government use of revenue, and Support for redistribution). Column 3 includes a set of demographic controls (Age, Male, Black, Student, and Unemployed) on the large subset of participants for whom all demographic details were available.Extended table: Table 6. *p<0.1; **p<0.05; ***p<0.01.

was complete.

Additional support of hypothesized mechanisms are shown in Table 3, which creates Factor Premiums, or differences in the ratings of policies across the three factors discussed earlier in Section 5.1 (Emissions, Your Finances, and Low Income). NoDE is associated with a higher factor premium in the relative rating for Penalty & Reward in the two factors we would expect to be most affected, (1) the effect on your own finances and (2) the effect on low income individuals. Based on the theoretical analysis in Section 4.4, these two channels concerning who wins and loses financially appear to be clarified by Penalty & Reward's design. Most survey participants believe themselves below-average emitters of carbon, shown in Figure 12, and therefore we would expect them to respond positively to Penalty & Reward concerning the financial impact to their own household.[38]

Analyzing the effect of controls on the outcome variable, which is visible in the extended version of Table 2, Appendix Table 6, we see no strong effects of control variables on Penalty & Reward Premium, apart from belief in climate science, which is strongly and positively predictive of the premium. This may result from the possibility that the expected predictive effect of climate science for Fee & Dividend support is masked by a distaste for the policy, indicated in Appendix Table 7. Because participants believing in climate science overwhelmingly expressed support for Penalty/Reward, climate science beliefs are predictive of the gap in support between these two policies.

In Appendix Table 7, we can see the predictive power of the behavioral and control variables on support for each policy. NoDE is strongly and negatively associated with

[38]In analyzing the predictiveness of rating factors on overall policy support, we see in Appendix Table 8 that, in both policies, the effectiveness in reducing emissions receives the highest weighting. Impact on one's own finances follows next, with impact on low-income persons weighed least of the three.

Table 3: Baseline Study: Factor Premium for Penalty & Reward

	Dependent variable: Factor premium of P&R over F&D		
	Emissions	Your Finances	Low Income
	(1)	(2)	(3)
Constant	28.713	1.516	−48.417*
	(25.267)	(20.742)	(27.440)
NoDE Score	10.222	29.958**	32.225**
	(14.008)	(14.470)	(14.341)
False Correlation error	9.146	16.616	33.415**
	(12.069)	(11.896)	(12.904)
Mental budgeting error	17.085**	3.786	0.767
	(8.514)	(8.757)	(10.296)
Survey order: experiment first	−2.848	1.602	−2.642
	(8.871)	(8.357)	(9.718)
Controls:			
Support factors	X	X	X
Demographic/employment	X	X	X
Observations	248	248	248
R^2	0.049	0.043	0.080
Adjusted R^2	0.008	0.002	0.041

Notes: The table above estimates the effect of cognitive bias on the Factor Premium for Penalty Reward across three factors. The first of the factors, "Emissions" stems from a rating given to each of the two policies based on their effectiveness at reducing emissions. The Emissions factor premium is Emissions rating given to Penalty & Reward minus the Emissions rating given to Fee & Dividend. The second factor, "Your Finances," stems from ratings to each policy based on its impact to your personal household. The third factor, "Low Income" stems from each policy's impact on low earners. See Appendix Section C.1.1 for more detail on the factor questions. The controls include the Support factors (Relative emissions, Belief in climate change, Left-right score, Belief job hurt from carbon tax, Trust government use of revenue, and Support for redistribution) and Demographic/employment factors (Age, Male, Black, Student, and Unemployed) *p<0.1; **p<0.05; ***p<0.01.

support for Fee & Dividend policy, as hypothesized, with little significant effect on support for Penalty & Reward. In line with theoretical predictions, one's own emissions relative to the average are strongly and negatively associated with support for either policy, as is the belief that one's job will be negatively affected by a carbon tax. As noted earlier, belief in climate science is strongly associated with support for Penalty & Reward but not with Fee & Dividend.

Political orientation is not meaningfully predictive of support for either policy in the presence of the other regressors. Leaning right has a slight positive association with support for either policy in Appendix Table 7, when controlling for belief in climate science, cognitive predictors, demographics and the other questions included in the survey. The belief in climate science variable appears to be capturing the portion of variation in policy support that is commonly attributed to political orientation. When belief in climate science is removed as a variable from the regression, political orientation is predictive of policy support. Belief in climate science has become highly politicized and is strongly correlated with left-right score in the sample.[39] Still, the results imply that a conservative who believes in climate science is no less likely to support carbon pricing policies than would a liberal who believes in climate science, when controlling for cognitive and other factors.

It appears that accounting for cognitive bias reduces the scope for politicization of carbon pricing, but reaching a firm conclusion on this is difficult. In Figure 6, we can see that among participants on the right, for whom support for Fee & Dividend was weak,

[39]The Pearson correlation coefficient is -0.65 and it is extremely statistically significant, with (-0.58,-0.71) as the 95% confidence interval.

support was weakest among high NoDE subset. This same group saw a far sharper increase in support for Penalty & Reward than their low-NoDE counterparts. While this is indicative of the role of partisanship diminishing when addressing cognitive frictions, we see a very similar effect among left-leaning participants.

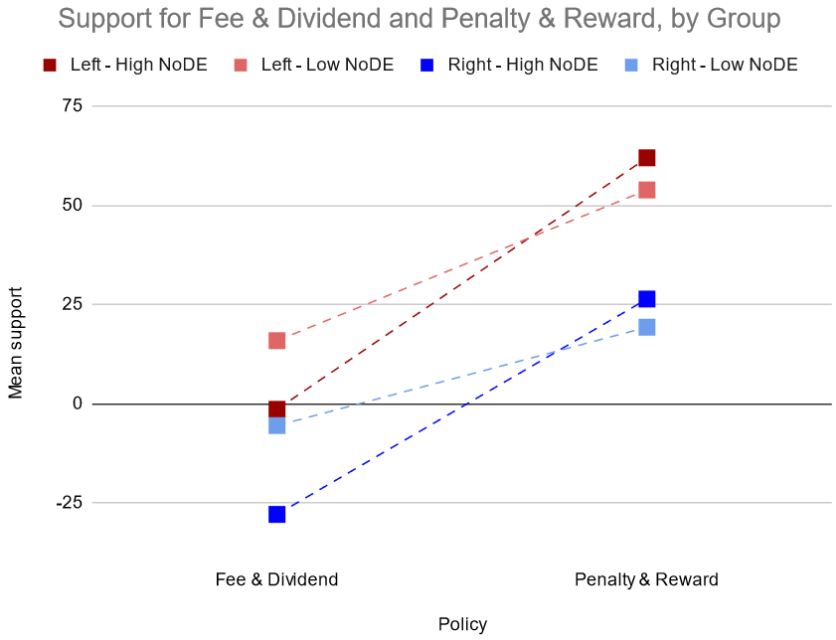


Figure 6: Comparison of mean policy support by left-leaning voters and right-leaning voters, each separated into high- and low-NoDE subsets.

Considering the proportions in support of the policy, the political gap narrows dramatically , with all political affiliations supporting Penalty & Reward, at the mean. Only the left-leaning subsample supported Fee & Dividend. But remarkably, there is a sharp increase in the intensity of support among left-leaning individuals for Penalty & Reward. While the literature notes that support for carbon pricing typically splits along party lines, the results here reveal very tepid support for carbon taxes on the left. Lowering the cognitive costs, for both political affiliations, appears to be a more important effect than the moderate gap between the two political groups, which largely remains in the

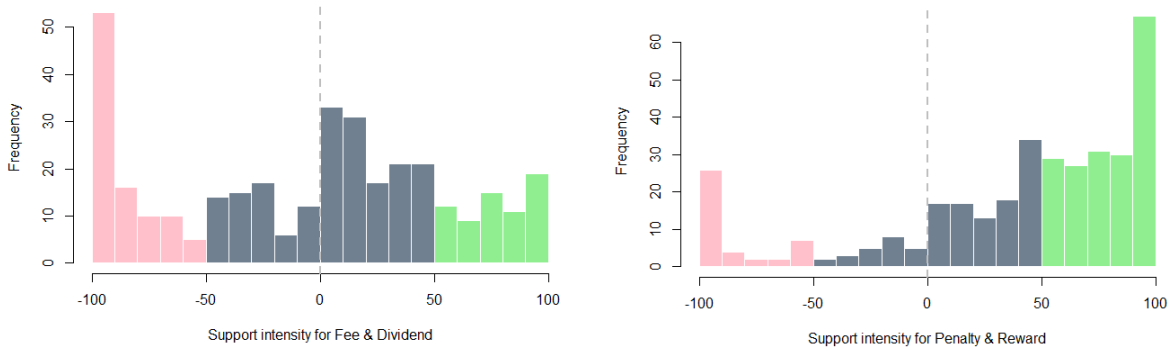


Figure 7: Distribution of Overall Rating for Fee & Dividend and Penalty & Reward.

support of the new policy, as well.

Lastly, by looking at the distribution of support intensity among subjects, we see in Figure 7 that Fee & Dividend has higher density of ratings in the vicinity of uncertainty or indifference than does Penalty & Reward. A policy with support that is less certain due to its complexity may be more vulnerable to misleading characterizations from vested interests.

7 Discussion

This study approached the role of cognitive frictions on carbon pricing support from multiple angles: a diagnostic approach on lab subjects engaging with an abstract policy; a test of a new policy designed to sidestep a key bias; and a test of the association of problem-solving ability on policy attitudes. The combination conveys a clear message: Neglect of Distributed Effects is a cognitive/attentional phenomenon and it affects how people choose real policies. For voters with constraints on attention or computation, Fee & Dividend's opacity puts it at a disadvantage. This disadvantage may be larger in the

public discourse, where a noisy information environment, influence from vested interests, and the lack of extrinsic incentives to evaluate costs and benefits can result in outcomes like the 10% support in France for a carbon tax and rebate, shown in Douenne and Fabre (2022). This study shows that policy design, informed by behavioral theory and evidence, can result in policies that produce much better intuitions about policy efficacy. Accurate initial impressions and intuitions may be crucial, as subsequent, supplemental information in this study, as with others, is often ignored. When we lessen the burden of choosing well, public preferences appear to respond.

In much of the literature about the public acceptability of carbon pricing, the conclusion that political orientation drives opinion is not what I observe in the analysis of the survey data. Many on the left reject Fee & Dividend, perhaps because they perceive it to be regressive. And even though for both policies, support is stronger on the left than on the right, the explanatory power of political orientation disappears when controlling for cognitive factors, belief in climate change, and other covariates. For studies that assess the role of partisanship in support for environmental policy, it is worth considering presenting policies (or creating policies), so that all of the important information can be easily and quickly integrated. My results suggest that the ability to form a clear intuition about how a beneficial policy works gives rise to strongly favorable opinions, which could be more resistant to misleading messaging.

Separating the explanatory power of political identity from cognitive factors is important. A recent paper, Kotchen (2022), advocates labeling corrective tax proposals with a ratio of the value of the externality reduction to the revenue collected by the policy. This is in part to address weak support on the right, who may not view government use of

tax revenues as desirable. It suggests considering corrective taxes only with large ratios (greater than 1). A consequence of proposing only taxes with ratios greater than one would be to tax the use of coal, but not of other fossil fuels. But this ratio lacks a clear, objective welfare interpretation. And if revenues are returned as dividends, then it even more difficult to rationalize poor public support in this framework.

I show that voter discounting of the Distributed Effects of corrective tax proposals make it more likely that policies with smaller revenue collection would be better supported, but also that this is rooted in cognitive factors and not solely by preferences or beliefs about the purpose, prudence or efficiency of government spending. In Abstract Policy Tasks, I can rule out concerns about prudence, efficiency, or trustworthiness of institutions. If these policy judgments were rooted in cognitive constraints, we would not necessarily want to enshrine such mistaken notions of policy effects into the kinds of policies we prioritize. It is possible that we can do better than that with thoughtful policy design.

By avoiding taxing and rebating everyone, Penalty & Reward shrinks the perceived scale of the policy and the concerns about tax revenue raised in Kotchen (2022). As corrective tax designs like Fee & Dividend are presently framed, a major portion of the tax money collected makes an unnecessary round trip to government and back, acquiring a subjective discount along the way, based on my findings. By removing the round trip, Penalty & Reward slashes the scale of collections and mechanically increases the ratio that Kotchen (2022) describes.

Going forward, analysis of real world climate policy support could be extended to non-carbon-pricing approaches, like subsidies or command and control policies. The

Diagnostic Experiment results suggest a bias favoring subsidies among participants, and it could be useful to measure if NoDE is predictive of support for environmental subsidies over taxes. Subsidies, despite being less efficient, appear to be widely popular in the world. My results and behavioral model suggest that opacity and complexity have the ability to improve support for certain policies, just as it can diminish support for certain policies. It would be interesting if the Distributed Effect-related opacity about how subsidies are paid for were important to its popularity and could somehow be reversed. Understanding what features of policies can be adjusted to make their true effects more intuitive to the public is an important, underexplored area of research.

Among the arguments emphasized by many advocates for Fee & Dividend policy is that it progressively redistributes money, independent of its positive effects of mitigating climate change.[40] A majority of households, mainly with lower incomes, are expected to be able to increase consumption due to Fee & Dividend's immediate financial benefits.[41] Yet, Fee & Dividend proposals offer its beneficiaries a salient but small cost, and an opaque but large benefit, and I have shown how this fosters bad policy selection and can aggregate into harms for most and benefits for a few. If easier to understand, the public may consider the use of a simplified policy like Penalty & Reward more ethical.[42] The

[40]This is highlighted in academic papers (Klenert et al., 2018; Budolfson et al., 2021); public statements from economists and public policy experts Akerlof et al. (2019); Center (2008); and statements of advocacy groups, the Climate Leadership Council and the Citizen's Climate Lobby.

[41]The top 10 percent of earners in the U.S. contribute 40 percent of the total emissions, and the average person among households in the top 0.1% of emitters in the U.S. contributed 57 times the emissions of the average person in the lowest decile (Starr et al., 2023a,b).

[42]Ambuehl et al. (2022) explores how incentives disproportionately attract people whose learning costs are higher in employment and consumer choice settings. For example, a physically risky job with a high wage has a transparent incentive with an opaque cost. Voters constrained in attention or computation would optimally use less information before making a decision, but are more likely to be disappointed in the end than the more discerning cognitively advantaged types. Most people find this dynamic unethical (Ambuehl and Ockenfels, 2017).

cost of understanding a difficult policy may cause an expected beneficiary to misperceive it, to avoid engaging with it and take cues from their party, or to abstain from voting because of confusion.[43] As a result, popular measures of support may not reflect the true preferences of the public. The voter's welfare and aggregate welfare can be hampered by ignoring the cognitive costs of policy evaluation.

Finally, the primary purpose of Penalty & Reward in this study is as a tool to better understand voter preferences. But in Appendix section D, I discuss strategies, costs, and benefits of literal and approximate implementations of Penalty & Reward to amenably mitigating negative externalities.

8 Conclusion

The fact that voters' reasoning is bounded requires us to move away from the assumption that voters always express their true preferences. The material incentives to think carefully about policy tend to be quite low for many people, and so it is important to understand the cognitively demanding elements of evaluation when designing policies. Environmental policies can be especially complex. Corrective taxes involve many moving parts, including changes to consumer surplus, usage of revenue, and reductions in externalities. This study identifies and measures mechanisms that suggests voters do not perfectly attend to each of these parts. Yet this does not necessarily mean societies must abandon useful tools like carbon taxes. The seeming failure to adopt carbon prices has

[43]Enke et al. (2023b) demonstrates that a lack of confidence resulting from cognitive difficulty can often cause individuals to select out of a participatory democratic process. An equitable climate policy that benefits all but a few can fail to realize if its beneficiaries happen to be more constrained, less confident, and thus more inclined to select out of participation.

shifted the tone of environmental economics in recent years, with many taking as given that the efficient policies at the center of the discipline have little hope of manifesting, due to public disapproval and shifting attention to "second-best" approaches.[44] Much attention has been given to polarization of beliefs or attitudes, for which little can be done by intervention.

The findings of this study suggest that some of this pessimism may be misplaced. Appealing to voters' financial self-interest is indeed a way to potentially bridge a divide between worldviews; however, to be successful, it must take into account the high sensitivity of the public to complexity. Cases in which simplification can be accomplished at little cost could be low-hanging fruit as societies try to build consensus around challenging and urgent priorities like the need to dramatically reduce greenhouse gas emissions without excessive damage to economic well-being.

[44]Many papers assessing "second-best" approaches include, as motivation, the observation that carbon taxes have failed to gain political momentum; economists have been scolded by academics of other disciplines for their attachment to carbon pricing (Kopp, 2022); and economists themselves have blogged or tweeted about moving on from their ideal tool of carbon taxes (Ho, 2021).

References

Abeler, J. and Jäger, S. (2015). Complex tax incentives. *American Economic Journal: Economic Policy*, 7(3):1–28.

Akerlof, G., Aumann, R., Deaton, A., Diamond, P., Engle, R., Fama, E., Hansen, L., Hart, O., Holmström, B., Kahneman, D., et al. (2019). Economists' statement on carbon dividends. *Wall Street Journal*, 16.

Ambuehl, S. and Ockenfels, A. (2017). The ethics of incentivizing the uninformed: A vignette study. *American Economic Review*, 107(5):91–95.

Ambuehl, S., Ockenfels, A., and Stewart, C. (2022). Who opts in? composition effects and disappointment from participation payments. *Review of Economics and Statistics*, pages 1–45.

Anderson, S., Marinescu, I., and Shor, B. (2023). Can pigou at the polls stop us melting the poles? *Journal of the Association of Environmental and Resource Economists*, 10(4):903–945.

Andersson, J. J. (2019). Carbon taxes and co2 emissions: Sweden as a case study. *American Economic Journal: Economic Policy*, 11(4):1–30.

Banovetz, J. and Oprea, R. (2023). Complexity and procedural choice. *American Economic Journal: Microeconomics*, 15(2):384–413.

Baranzini, A. and Carattini, S. (2017). Effectiveness, earmarking and labeling: testing

the acceptability of carbon taxes with survey data. *Environmental Economics and Policy Studies*, 19:197–227.

Baranzini, A., Van den Bergh, J. C., Carattini, S., Howarth, R. B., Padilla, E., and Roca, J. (2017). Carbon pricing in climate policy: seven reasons, complementary instruments, and political economy considerations. *Wiley Interdisciplinary Reviews: Climate Change*, 8(4):e462.

Baron, J. and Jurney, J. (1993). Norms against voting for coerced reform. *Journal of Personality and Social Psychology*, 64(3):347.

Baumol, W. J. (1972). On taxation and the control of externalities. *The American Economic Review*, 62(3):307–322.

Baumol, W. J. and Oates, W. E. (1971). The use of standards and prices for protection of the environment. In *The Economics of Environment: Papers from Four Nations*, pages 53–65. Springer.

Bernheim, B. D. (1987). Ricardian equivalence: An evaluation of theory and evidence. *NBER macroeconomics annual*, 2:263–304.

Bhargava, S., Loewenstein, G., and Sydnor, J. (2017). Choose to lose: Health plan choices from a menu with dominated option. *The Quarterly Journal of Economics*, 132(3):1319–1372.

Blanchard, O., Gollier, C., and Tirole, J. (2023). The portfolio of economic policies needed to fight climate change. *Annual Review of Economics*, 15(1):689–722.

Boyce, J. K. (2018). Carbon pricing: effectiveness and equity. *Ecological Economics*, 150:52–61.

Buchanan, J. M. and Wagner, R. E. (2013). *Fiscal responsibility in constitutional democracy*, volume 1. Springer Science & Business Media.

Budolfson, M., Dennig, F., Errickson, F., Feindt, S., Ferranna, M., Fleurbaey, M., Klenert, D., Kornek, U., Kuruc, K., Méjean, A., et al. (2021). Protecting the poor with a carbon tax and equal per capita dividend. *Nature Climate Change*, 11(12):1025–1026.

Carattini, S., Baranzini, A., Thalmann, P., Varone, F., and Vöhringer, F. (2017). Green taxes in a post-paris world: are millions of nays inevitable? *Environmental and Resource Economics*, 68:97–128.

Center, C. T. (2008). Pricing carbon efficiently and equitably. *Carbon Tax Center, May*, 23.

Cherry, T. L., Kallbekken, S., and Kroll, S. (2012). The acceptability of efficiency-enhancing environmental taxes, subsidies and regulation: An experimental investigation. *Environmental Science & Policy*, 16:90–96.

Cherry, T. L., Kallbekken, S., and Kroll, S. (2014). The impact of trial runs on the acceptability of environmental taxes: Experimental evidence. *Resource and Energy Economics*, 38:84–95.

Chetty, R., Looney, A., and Kroft, K. (2009). Salience and taxation: Theory and evidence. *American economic review*, 99(4):1145–1177.

Colmer, J., Martin, R., Muûls, M., and Wagner, U. J. (2022). Does pricing carbon mitigate climate change? firm-level evidence from the european union emissions trading scheme.

Cramton, P., MacKay, D. J., Ockenfels, A., and Stoft, S. (2017). *Global carbon pricing: the path to climate cooperation*. The MIT Press.

Cronin, J. A., Fullerton, D., and Sexton, S. (2019). Vertical and horizontal redistributions from a carbon tax and rebate. *Journal of the Association of Environmental and Resource Economists*, 6(S1):S169–S208.

Dal Bó, E., Dal Bó, P., and Eyster, E. (2018). The demand for bad policy when voters underappreciate equilibrium effects. *The Review of Economic Studies*, 85(2):964–998.

Dechezleprêtre, A., Fabre, A., Kruse, T., Planterose, B., Chico, A. S., and Stantcheva, S. (2022). Fighting climate change: International attitudes toward climate policies. Technical report, National Bureau of Economic Research.

Dell'Anno, R. and Mourao, P. (2012). Fiscal illusion around the world: an analysis using the structural equation approach. *Public finance review*, 40(2):270–299.

Douenne, T. and Fabre, A. (2022). Yellow vests, pessimistic beliefs, and carbon tax aversion. *American Economic Journal: Economic Policy*, 14(1):81–110.

Downs, A. (1957). An economic theory of political action in a democracy. *Journal of political economy*, 65(2):135–150.

Dresner, S., Dunne, L., Clinch, P., and Beuermann, C. (2006). Social and political

responses to ecological tax reform in europe: an introduction to the special issue. *Energy policy*, 34(8):895–904.

Enke, B., Graeber, T., and Oprea, R. (2023a). Complexity and time. Technical report, National Bureau of Economic Research.

Enke, B., Graeber, T., and Oprea, R. (2023b). Confidence, self-selection, and bias in the aggregate. *American Economic Review*, 113(7):1933–1966.

Esponda, I., Vespa, E., and Yuksel, S. (2020). Mental models and learning: The case of base-rate neglect. Technical report, Tech. rep.

Farhi, E. and Gabaix, X. (2019). Online appendix for "optimal taxation with behavioral agents".

Fremstad, A. and Paul, M. (2019). The impact of a carbon tax on inequality. *Ecological Economics*, 163:88–97.

Goulder, L. H., Hafstead, M. A., Kim, G., and Long, X. (2019). Impacts of a carbon tax across us household income groups: What are the equity-efficiency trade-offs? *Journal of Public Economics*, 175:44–64.

Hastings, J. and Shapiro, J. M. (2012). Mental accounting and consumer choice: Evidence from commodity price shocks. Technical report, National Bureau of Economic Research.

Ho, B. (2021). Time to give up on a carbon tax? https://news.climate.columbia.edu/2021/04/26/retire-carbon-tax/.

Horowitz, J. K., Cronin, J.-A., Hawkins, H., Konda, L., and Yuskavage, A. (2017). *Methodology for analyzing a carbon tax.*

Kahneman, D., Knetsch, J. L., and Thaler, R. H. (1991). Anomalies: The endowment effect, loss aversion, and status quo bias. *Journal of Economic perspectives*, 5(1):193–206.

Kallbekken, S. and Aasen, M. (2010). The demand for earmarking: Results from a focus group study. *Ecological economics*, 69(11):2183–2190.

Kallbekken, S., Kroll, S., and Cherry, T. L. (2011). Do you not like pigou, or do you not understand him? tax aversion and revenue recycling in the lab. *Journal of Environmental Economics and Management*, 62(1):53–64.

Kallbekken, S. and Sælen, H. (2011). Public acceptance for environmental taxes: Self-interest, environmental and distributional concerns. *Energy Policy*, 39(5):2966–2973.

Klenert, D., Mattauch, L., Combet, E., Edenhofer, O., Hepburn, C., Rafaty, R., and Stern, N. (2018). Making carbon pricing work for citizens. *Nature Climate Change*, 8(8):669–677.

Kopp, B. (2022). Tweet from aug. 7, 2022: Environmental economists should probably take a moment of humble reflection... https://www.twitter.com/.

Kotchen, M. (2022). Taxing externalities: Revenue vs. welfare gains with an application to us carbon taxes. Technical report, National Bureau of Economic Research.

Kuziemko, I., Marx, N. L., and Naidu, S. (2023). "compensate the losers?" economic policy and partisan realignment in the us. Technical report, National Bureau of Economic Research.

Lacetera, N., Pope, D. G., and Sydnor, J. R. (2012). Heuristic thinking and limited attention in the car market. *American Economic Review*, 102(5):2206–2236.

Lohmann, P. M., Gsottbauer, E., Doherty, A., and Kontoleon, A. (2022). Do carbon footprint labels promote climatarian diets? evidence from a large-scale field experiment. *Journal of Environmental Economics and Management*, 114:102693.

Maestre-Andrés, S., Drews, S., and van den Bergh, J. (2019). Perceived fairness and public acceptability of carbon pricing: a review of the literature. *Climate policy*, 19(9):1186–1204.

Metcalf, G. E. (2021). Carbon taxes in theory and practice. *Annual Review of Resource Economics*, 13:245–265.

Metcalf, G. E. (2023). Five myths about carbon pricing. Technical report, National Bureau of Economic Research.

Mildenberger, M., Lachapelle, E., Harrison, K., and Stadelmann-Steffen, I. (2022). Limited impacts of carbon tax rebate programmes on public support for carbon pricing. *Nature climate change*, 12(2):141–147.

Montgomery, W. D. (1972). Markets in licenses and efficient pollution control programs. *Journal of economic theory*, 5(3):395–418.

Muller, L., Lacroix, A., and Ruffieux, B. (2019). Environmental labelling and consumption changes: A food choice experiment. *Environmental and resource economics*, 73:871–897.

Murray, B. and Rivers, N. (2015). British columbia's revenue-neutral carbon tax: A review of the latest "grand experiment" in environmental policy. *Energy Policy*, 86:674–683.

Nielsen, K. and Rehbeck, J. (2022). When choices are mistakes. *American Economic Review*, 112(7):2237–2268.

Nordhaus, W. D. (2001). Global warming economics.

Oprea, R. (2022). Simplicity equivalents. *Working Paper*.

Pigou, A. C. (1924). *The Economics of Welfare*. Macmillan.

Puviani, A. (1903). *Teoria Dell'illusione Finanziaria*. Milan: Remo Sandon.

Rees-Jones, A. and Taubinsky, D. (2016). *Heuristic perceptions of the income tax: Evidence and implications for debiasing*. National Bureau of Economic Research.

Rees-Jones, A. and Taubinsky, D. (2020). Measuring "schmeduling". *The Review of Economic Studies*, 87(5):2399–2438.

Rivers, N. and Schaufele, B. (2015). Salience of carbon taxes in the gasoline market. *Journal of Environmental Economics and management*, 74:23–36.

Rivers, N. and Shaffer, B. (2022). How do households spend carbon tax rebates? *Available at SSRN 4054023*.

Sælen, H. and Kallbekken, S. (2011). A choice experiment on fuel taxation and earmarking in norway. *Ecological Economics*, 70(11):2181–2190.

Sausgruber, R. and Tyran, J.-R. (2005). Testing the mill hypothesis of fiscal illusion. *Public choice*, 122(1):39–68.

Sausgruber, R. and Tyran, J.-R. (2011). Are we taxing ourselves?: How deliberation and experience shape voting on taxes. *Journal of Public Economics*, 95(1-2):164–176.

Shultz, G. P. and Halstead, T. (2018). The dividend advantage. *Climate Leadership Council*.

Starr, J., Nicolson, C., Ash, M., Markowitz, E. M., and Moran, D. (2023a). Assessing us consumers' carbon footprints reveals outsized impact of the top 1%. *Ecological Economics*, 205:107698.

Starr, J., Nicolson, C., Ash, M., Markowitz, E. M., and Moran, D. (2023b). Income-based us household carbon footprints (1990–2019) offer new insights on emissions inequality and climate finance. *PLOS Climate*, 2(8):e0000190.

Steg, L., Dreijerink, L., and Abrahamse, W. (2006). Why are energy policies acceptable and effective? *Environment and behavior*, 38(1):92–111.

Thaler, R. H. (1990). Anomalies: Saving, fungibility, and mental accounts. *Journal of economic perspectives*, 4(1):193–205.

Thaler, R. H. (1999). Mental accounting matters. *Journal of Behavioral decision making*, 12(3):183–206.

Tietenberg, T. H. (1973). Controlling pollution by price and standard systems: a general equilibrium analysis. *The Swedish Journal of Economics*, pages 193–203.

Yamazaki, A. (2017). Jobs and climate policy: Evidence from british columbia's revenue-neutral carbon tax. *Journal of Environmental Economics and Management*, 83:197–216.

A Appendix 1: Theory Supplement

A.1 General Framework

Choosing a pricing policy as a voter-consumer requires thinking of two stages recursively:

1. Voter expresses policy preference.

2. Consumers make consumption decisions given a policy.

All policy proposals and status quo policy conditions correspond to a value, τ, which when zero implies no policy, when positive implies a tax, and when negative implies a subsidy. Voters value the policy shift $\Delta\tau$, when forming a view of a proposal and the strength of that view. For a rational voter:

$$V_i(\Delta\tau \mid \tau_0) = V_i(\tau_1) - V_i(\tau_0)$$

When $V_i(\Delta\tau \mid \tau_0) = V_i(\tau_1) - V_i(\tau_0) > 0$, the voter supports the proposal, when $V_i(\Delta\tau \mid \tau_0) < 0$, we support the status quo, and otherwise we are indifferent. In this paper we consider $V_i(\Delta\tau \mid \tau_0)$ our primary object of interest (i.e. how do subjects sign it and perceive its magnitude relative to the rational full attention benchmark?)[45]

[45]You might ask, 'Why have the object of interest be the value of a shift $V_i(\Delta\tau \mid \tau_0)$, and not simply one's comprehensive welfare under the new policy; $U_i(\tau_1)$?' For two important reasons: (1) In a political context, people are asked about preferences about a policy change, but not to explicitly quantify their welfare at present or under a proposal. So, $U_i(\tau_0)$ and $U_i(\tau_1)$ may be rather strange objects. Furthermore, (2) both in the field and in the lab, asking people to estimate their before and after welfare is leading them to a kind of calculation that they may not typically make when considering a proposal. In general, we don't want to assume $V_i(\Delta\tau \mid \tau_0) = V_i(\tau_1) - V_i(\tau_0)$.

$$V_i(\Delta\tau \mid \tau_0) = \pi_i(c_i(\tau_1)) \;-\; \pi_i(c_i(\tau_0)) \;+\; \sum_{j\in J} (\pi_i(c_j(\tau_1)) \;-\; \pi_i(c_j(\tau_0)))$$

Here, $\pi()$ is an argument representing net gain or loss to voter-consumer i from actions from consumer i and any consumer j from the set of J consumers. Because actions are in dominant strategies and a function of revealed policy outcomes, there are no interaction terms between the private net gain of subsequent consumption choice of any two agents. Hence I can represent the gain or loss from one's own consumption choices and the consumption choices of others as additively separable.

I define the portion of your change in welfare that stems from your own consumption choices c_i, in essence $\pi_i(c_i(\tau_1)) \;-\; \pi_i(c_i(\tau_0))$ as a Consumer Effect or *CE*. The portion of your change in welfare that stems from the consumption of others c_j, i.e. $\sum_{j\in J} (\pi_i(c_j(\tau_1)) \;-\; \pi_i(c_j(\tau_0)))$ is Distributed Effect, or *DE* (or additionally an Externality Effect, *EE*, depending on the circumstance. If the consumption of others has no bearing on your welfare, $V_i(\Delta\tau \mid \tau_0) = CE_i((\tau_1, \tau_0))$.

Now generalizing to a behavioral voter:

$$\hat{V}_i(\Delta\tau \mid \tau_0) = \iota \overbrace{(\pi_i(c_i(\tau_1)) \;-\; \pi_i(c_i(\tau_0)))}^{\textit{Consumer Effect } (CE)} + \lambda \overbrace{\sum_{j\in J} (\pi_i(c_j(\tau_1)) \;-\; \pi_i(c_j(\tau_0)))}^{\textit{Distributed Effect } (DE)}$$

Above, ι and λ are weights that might bias a voter on a proposed policy. (Each weight would equal 1 for an unbiased individual.) The CE-DE structure comes from the two

natural components of the problem, but the weights could also be interpreted 'as if.'

To explore the implications of neglect to the Distributed Effects of subsidy and tax policies that we've discussed, I define the policy variable τ as a number that would be added the market price p of a purchasable object. Therefore when $\tau > 0$, it represents a tax, and when $\tau < 0$ it represents a subsidy.

Subsequent to policy outcome, one's own actions are purchase choices. For any given private value function for goods consumed, the net benefit from one's optimal purchases (i.e. personal consumer surplus) can only be the same or lower when policy moves in the direction of taxation or desubsidization. The effective price under the new policy $\tilde{p}_1 = \tau_1 + p$ is greater than than $\tilde{p}_0 = \tau_0 + p$. Therefore, when $\tau_1 - \tau_0 > 0$, $CE \leq 0$. Analagously, under direction of subsidization $\tau_1 - \tau_0 < 0$, $CE \geq 0$.

For the analysis of Distributed Effects, let us define x as the cumulative demand function for all other consumers.[46] $\gamma(\tau_1\ x(\tau_1))$ represents one's share of the cumulative revenue or debt generated by the new policy, as $\gamma(\tau_0\ x(\tau_0))$ represents the same for the status quo. Then $DE = \gamma\left[(\tau_1\ x(\tau_1)) - (\tau_0\ x(\tau_0))\right]$.[47]

Let us consider both the proposed introduction of a new policy and the proposed removal of an existing policy. First, if introducing a tax, $\tau_0 = 0$ and $\tau_1 > 0$. Demand is nonnegative $x() \geq 0$, and therefore $\tau_1 x(\tau_1) \geq 0$. Because $\tau_0 x(\tau_0) = 0$, we know $\tau_1 x(\tau_1) \geq \tau_0 x(\tau_0)$ and that therefore $DE \geq 0$. If introducing a subsidy, $\tau_0 = 0$ and

[46] To formally define the cumulative demand function x() for a set J of rational consumers, given policy τ, price p, and a set of object values L_j for a consumer j
$x(\tau) = \sum_{j \in J} \left(\sum_{l \in L_j} I_{(l > p + \tau\omega)} \right)$ where $\omega = 1 - \frac{1}{N}$. N represents the population size and as it increases, ω approaches 1.

[47] $\gamma > 0$. For example, with equal dividend, $\gamma = \frac{1}{N}$ where $N = \mid J \mid +1$

$\tau_1 < 0$. Therefore, $\tau_1 x(\tau_1) \leq 0$. Because $\tau_0 x(\tau_0) = 0$, $\tau_1 x(\tau_1) \leq \tau_0 x(\tau_0)$ and $DE \leq 0$. If removing an existing tax, $\tau_0 > 0$ and $\tau_1 = 0$. Therefore, $\tau_1 x(\tau_1) \leq \tau_0 x(\tau_0)$ and $DE \leq 0$. If removing an existing subsidy, $\tau_0 < 0$ and $\tau_1 = 0$. Therefore, $\tau_1 x(\tau_1) \geq \tau_0 x(\tau_0)$ and $DE \geq 0$.

Returning to our complete value function, let's consider what happens when introducing a new tax or removing an existing subsidy. Based on the analysis above, in both of these cases $CE \leq 0$ and $DE \geq 0$. If $\iota > \lambda$ for a biased individual,

$$\hat{V}_i(\Delta\tau \mid \tau_0) = \iota CE(\tau_1, \tau_0) + \lambda DE(\tau_1, \tau_0) \leq V_i(\Delta\tau \mid \tau_0) = CE(\tau_1, \tau_0) + DE(\tau_1, \tau_0)$$

Furthermore, as long as either $CE \neq 0$ or $DE \neq 0$, then:

$$\hat{V}_i(\Delta\tau \mid \tau_0) = \iota CE(\tau_1, \tau_0) + \lambda DE(\tau_1, \tau_0) < V_i(\tau_1, \tau_0) = CE(\tau_1, \tau_0) + DE(\tau_1, \tau_0)$$

Analogously, when removing an existing tax, or introducing a new subsidy, $CE \geq 0$ and $DE \geq 0$. So when $\iota > \lambda$:

$$\hat{V}_i(\Delta\tau \mid \tau_0) = \iota CE(\tau_1, \tau_0) + \lambda DE(\tau_1, \tau_0) \geq V_i(\Delta\tau \mid \tau_0) = CE(\tau_1, \tau_0) + DE(\tau_1, \tau_0)$$

When either $CE \neq 0$ or $DE \neq 0$, then:

$$\hat{V}_i(\Delta\tau \mid \tau_0) = \iota CE(\tau_1, \tau_0) + \lambda DE(\tau_1, \tau_0) > V_i(\Delta\tau \mid \tau_0) = CE(\tau_1, \tau_0) + DE(\tau_1, \tau_0)$$

A.2 Equilibrium Effects

In Dal Bó et al. (2018), which explores Equilibrium Effects and shows evidence that individuals do not fully incorporate how a change in policy should affect subsequent behavior, it may be useful to decompose CE and DE even further along a direct v. equilibrium (static v. dynamic) dimension. CE can be decomposed into a Direct Effect and an Equilibrium Effect, , and DE can be decomposed into Distributed Direct Effect and a Distributed Equilibrium Effect.

$$V_i(\Delta\tau \mid \tau_0) = \textit{ Direct Eff } + \textit{ Equilibrium Eff } + \textit{ Distrib Direct Eff } + \textit{ Distrib Equil Eff}$$

$$\textit{Direct Effect} = \pi_i(c_i(\tau_0), \tau_1) - \pi_i(c_i(\tau_0), \tau_0)$$

The Direct Effect, or the static part of the Consumer Effect, represents the net gain or loss from a proposed policy attributable to your own consumption choices, assuming that you would not make any change to your consumption choice in response to the enactment of the proposed policy.

$$\textit{Equilibrium effect} = \pi_i(c_i(\tau_1), \tau_1) - \pi_i(c_i(\tau_0), \tau_1)$$

The Equilibrium Effect, or the dynamic part of the Consumer Effect, represents the net gain or loss from a proposed policy attributable only to the change in your own

consumption choices in response to the enactment of the proposed policy.

$$Distributed\ Direct\ Effect = \sum_{j \in J} (\pi_i(c_j(\tau_0), \tau_1) - \pi_i(c_j(\tau_0), \tau_0))$$

The Distributed Direct Effect, or static part of the Distributed Effect, represents the net gain or loss to you from a proposed policy that is attributable to consumption choices of others, assuming that others would not make any change to consumption choices in response to the enactment of the proposed policy.

$$Distributed\ Equilibrium\ Effect = \sum_{j \in J} (\pi_i(c_j(\tau_1), \tau_1) - \pi_i(c_j(\tau_0), \tau_1))$$

The Distributed Equilibrium Effect represents the net gain or loss from a proposed policy attributable only to the change in others' consumption choices in response to the enactment of the proposed policy.

$$CE = Direct\ Effect + Equil\ Effect = \pi_i(c_i(\tau_1), \tau_1) - \pi_i(c_i(\tau_0), \tau_0)$$

$$DE = Dist\ Direct\ Effect\ +\ Dist\ Equil\ Effect = \sum_{j \in J} (\pi_i(c_j(\tau_1), \tau_1) - \pi_i(c_j(\tau_0), \tau_0))$$

Now let's consider some potential bias terms on computing the shift in welfare resulting from a shift in policy:

$$\hat{V}_i(\Delta\tau \mid \tau_0) = \kappa Direct + \eta Equilibrium + \nu Distrib\ Direct + \theta Distrib\ Equil$$

A.3 Loss Aversion With Mental Accounts

If subjects are irrationally tax averse in real-world and abstract settings, there is an alternative explanation to NoDE. This explanation, Loss Aversion With Mental Accounts, or LAMA, cannot be rationalized by behavioral risk preferences, as there is no uncertainty, only computational constraints.[48] If deciding on whether to support the introduction of a tax, a behavioral voter may value the policy as follows:

$$\hat{V}(\Delta\tau \mid \tau_0) = \begin{cases} CE(\tau_1,\tau_0) + DE(\tau_1,\tau_0), & \text{if } CE(\tau_1,\tau_0) \geq 0 \text{ and } DE(\tau_1,\tau_0) \geq 0 \\ \alpha CE(\tau_1,\tau_0) + DE(\tau_1,\tau_0), & \text{if } CE(\tau_1,\tau_0) < 0 \text{ and } DE(\tau_1,\tau_0) \geq 0 \\ CE(\tau_1,\tau_0) + \alpha DE(\tau_1,\tau_0), & \text{if } CE(\tau_1,\tau_0) \geq 0 \text{ and } DE(\tau_1,\tau_0) < 0 \\ \alpha CE(\tau_1,\tau_0) + \alpha DE(\tau_1,\tau_0), & \text{if } CE(\tau_1,\tau_0) < 0 \text{ and } DE(\tau_1,\tau_0) < 0 \end{cases}$$

Here, a voter is not biased in evaluation of DE, but in either of CE or DE, conditional on taking a negative value. This impedes an accurate calculation of the net effect of the two policies, as evident in past work (Kahneman et al., 1991; Thaler, 1999). Because for tax introductions, $CE \leq 0$ and $DE \geq 0$, the effect of LAMA can result in suboptimal rejection of taxes relative to a no-policy reference point. By forcing subjects to evaluate subsidy proposals ($\tau < 0$) in certain rounds, where $CE > 0$ and $DE < 0$, I am able to evaluate NoDE, while controlling for LAMA.

[48] Oprea (2022) shows subjects make similar aggregate errors in deterministic and lottery settings.

A.4 Penalty & Reward Premium: Other mechanisms

There are two alternative or additional behavioral mechanisms that could result in greater support for Penalty & Reward than for Fee & Dividend that I test in the survey and can ultimately control for in the evaluation of the primary NoDE hypothesis. **False correlation** is a misperception that the dividend under Fee & Dividend policy is strongly correlated with one's own total carbon taxes paid. A heuristic known as **mental budgeting**, when applied to evaluation of Fee & Dividend, may cause individuals to choose policies as if dividends will be allocated by individuals to the taxed consumption category.

Both of these cognitive mechanisms could have effects on policy perception through the "effectiveness in reducing emissions" channel. First, a misperception of correlation between fees and dividends that is rectified by Penalty & Reward. Then, a novel application of a well-established behavioral heuristic, mental budgeting, the scope of which is also reduced in Penalty & Reward.

A.4.1 False correlation

A potential misperception concerning Fee & Dividend is that the dividend is strongly correlated with total carbon taxes that an individual pays. There are at least three ways this mistake could arise:

1. A perception of homogeneity of consumer behavior, i.e. decisions are made as if all consumers act as one or mirror each others' behavior.

2. A misunderstanding of the policy that an individual dividend is a repayment of exactly the individual taxes paid.

3. An exaggerated sense of the influence of one's own tax payments on the size of one's personal dividend.

Consider the following formula for a subjective dividend:

$$\textit{Subjective Dividend} = \alpha\ \tau_{carb}\ q_i^{carb} + (1-\alpha)\ \tau_{carb}\ \bar{q}_{carb}$$

When $\alpha = 0$ an individual is unmistaken and when $\alpha = 1$, they are fully mistaken. The effect of a positive α is to effectively shrink the size of τ. For a fully mistaken individual thinking they will receive the entirety of their paid taxes back in the form of a dividend, it's as if the tax does not exist, and the effective price is the original market price.

In reality, with enough people in the population, there is nearly no correlation ($\frac{1}{N}$) between an individual's consumption choices and their personal dividend. But it may be difficult for people to recognize this. For example, suppose the ex ante perception of a policy proposal corresponds to $\alpha = \frac{1}{2}$. The perceived effectiveness of reducing consumption of the bad would be equivalent to if a tax of $\frac{\tau}{2}$ were implemented, instead of τ. This would make offering a dividend seem counterproductive to the goal, like tying one hand behind one's back.[49]

Penalty & Reward has no dividend, and therefore there is no potential to form a false correlation. By directly taxing or rewarding, the step of computing and comparing the incentive effects of two pieces, fees and dividends, is absent.

[49]This is the kind of mistake that could be corrected through experience. However, an unfortunate feature of policy selection is an intrinsic difficulty to learn through experience. Aversion to a policy due to a misperception is not easily corrected when learning is endogenous.

A.4.2 Mental Budgeting

Descriptions of Fee & Dividend sometimes provoke the following skeptical response: *If you give the carbon tax revenue back to people, they will just spend it on fossil fuels. How will that accomplish anything?* (Rivers and Shaffer, 2022). Some of these people would prefer to see the money spent on green projects or other forms of government spending.

The intuition that dividends will be allocated to the problematic consumption category appears related to concept with a long history in behavioral economics, mental budgeting. Consumers often use fixed mental budgets for categories of consumption.[50] Meant to help manage the complexity of allocating income to many spending categories, this disciplining technique can be counterproductive and in violation of the classic consumer optimization problem.

A version of mental budgeting that I explore here involves voters or policy evaluators imagining that consumers would deposit the dividend in the mental budget allocated to the taxed category, for example, gasoline. Perhaps the dividend, as a compensatory measure, is perceived by some as buffering consumers from making necessary changes to their consumption patterns. They think consumption choices are made strictly in terms of category budgets, without regard to prices. Then, they perceive that consumers treat their regular dividend as "gas money."

This notion is well represented by Terrence Corcoran, a Canadian financial newspaper columnist, who argued that the rebate "negates the incentive" to cut emissions fostered by the tax. Corcoran's argument is cited in Rivers and Shaffer (2022), which investigates

[50] Mental budgeting is part of a broader literature called mental accounting. For a deeper look: Thaler (1990, 1999); Hastings and Shapiro (2012); Farhi and Gabaix (2019).

empirically how carbon dividends in British Columbia have been spent. What they discover is in sharp contrast to Corcoran's view: relatively little of dividends are spent on fossil fuels. Related work details how the Fee & Dividend tax policy in British Columbia was more effective than assumed, given its modest size, and that consumers responded to taxes far more than would be expected from evidence of responses to market price changes (Rivers and Schaufele, 2015; Murray and Rivers, 2015), findings that concur with Andersson (2019) , which evaluates a standard carbon tax in Sweden.[51]

Because Fee & Dividend has had few adopters, there is a lack of empirical research on this matter beyond British Columbia. Arguably the notion that dividends would be spent disproportionately on the taxed category is unlikely, and even less likely to persist over time.

To illustrate how a mental budgets notion would impact perceived effectiveness, I formally model it below.

In the absence of any policy:

$$\dot{q}_i^{carb} = q_i^{carb}(p_{carb},\ p_{clean},\ w_i)$$

$$\dot{q}_i^{clean} = q_i^{clean}(p_{carb},\ p_{clean},\ w_i)$$

In optimal response to a carbon Fee & Dividend policy :

$$\hat{q}_i^{carb} = q_i^{carb}(p_{carb} + \tau_{carb},\ p_{clean},\ w_i + \tau_{carb} * \bar{q}_{carb})$$

[51]Murray and Rivers (2015) and Yamazaki (2017) also provide evidence of a negligible or slightly positive effect on the aggregate economy and employment. While a majority of the public opposed the policy ahead of time, after three years a majority of the public supported the policy.

$$\hat{q}_i^{clean} = q_i^{clean}(p_{carb} + \tau_{carb},\ p_{clean},\ w_i + \tau_{carb} * \bar{q}_{carb})$$

Perceived response to a proposed carbon Fee & Dividend policy:

$$\tilde{q}_i^{carb} = q_i^{carb}(p_{carb} + \tau_{carb},\ p_{clean},\ w_i) + \frac{\tau_{carb} * \bar{q}_{carb}}{p_{carb} + \tau_{carb}}$$

$$\tilde{q}_i^{clean} = q_i^{clean}(p_{carb} + \tau_{carb},\ p_{clean},\ w_i)$$

Making the assumptions that $\frac{dx}{dw} > 0$ and $\frac{dx}{dp} < 0$, and that the optimal bundle under Fee & Dividend is an interior solution, we can conclude the following:

$$\hat{q}_i^{carb} < \tilde{q}_i^{carb}$$

$$\hat{q}_i^{clean} > \tilde{q}_i^{clean} > \dot{q}_i^{clean}$$

Aggregating the optimal responses we would find that:

$$\hat{Q}_i^{carb} < \dot{Q}_i^{carb}$$

$$\hat{Q}_i^{carb} < \tilde{Q}_i^{carb}$$

$$\hat{Q}_i^{clean} > \tilde{Q}_i^{clean} > \dot{Q}_i^{clean}$$

To illustrate further, let's consider a Cobb-Douglas case of imperfect substitutes:

$$U(q_{carb}, q_{clean}) = q_{clean}^{\alpha} q_{carb}^{1-\alpha} \qquad q_{clean}p_{clean} + q_{carb}p_{carb} = w$$

Since $q^*_{clean} = \alpha\frac{w}{p_{clean}}$ and $q^*_{carb} = (1-\alpha)\frac{w}{p_{carb}}$, the Cobb Douglas case is a special one. Here, the use of mental budgets, in this case αw and $(1-\alpha)w$, turn out to cause no deviation from optimality for any uncompensated changes in price, e.g., a tax. Units purchased of the taxed category will drop, but the budget in the numerator remains.

$$q^{*tax}_{clean} = \frac{\alpha w}{p_{clean}}$$

$$q^{*tax}_{carb} = \frac{(1-\alpha)w}{p_{carb} + \tau_{carb}}$$

But the dividend, which I label D, can cause a mistaken view for those thinking with mental budgets. The optimal bundle now becomes $q^{*fd}_{clean} = \frac{\alpha(w+D)}{p_{clean}}$ and $q^{*fd}_{carb} = \frac{(1-\alpha)(w+D)}{p_{carb}+\tau_{carb}}$.

How much carbon emissions would this reduce relative to no policy? (Here I use the superscript "sq" to denote the status quo in which there was no policy.)

$$\sum_i q^{*sq}_{i,carb} = \frac{(1-\alpha)\sum_i w}{p_{carb}}$$

$$\sum_i q^{*fd}_{i,carb} = \frac{(1-\alpha)(\sum_i w_i + \sum_i D)}{p_{carb} + \tau_{carb}} = \frac{(1-\alpha)(\sum_i w_i + \tau_{carb}\sum_i q^{*fd}_{i,carb})}{p_{carb} + \tau_{carb}}$$

$$\sum_i q^{*fd}_{i,carb} = \frac{(1-\alpha)\sum_i w}{p_{carb} + \tau_{carb} + \tau_{carb}(\alpha - 1)} = \frac{(1-\alpha)\sum_i w}{p_{carb} + \alpha\tau_{carb}} < \frac{(1-\alpha)\sum_i w}{p_{carb}} = \sum_i q^{*sq}_{i,carb}$$

But a mental-budgeting-afflicted consumer's bundle is:

$$\tilde{q}_{clean} = \frac{\alpha w}{p_{clean}}$$

$$\tilde{q}_{carb} = \frac{(1-\alpha)w + D}{p_{carb} + \tau_{carb}}$$

$$\tilde{q}_{clean} - q^*_{clean} = -\alpha D$$

$$\tilde{q}_{carb} - q^*_{carb} = \alpha D$$

Aggregated:

$$\sum_i \tilde{q}^{*fd}_{i,carb} = \frac{(1-\alpha)\sum_i w_i + \tau_{carb}\sum_i \tilde{q}^{*fd}_{i,carb}}{p_{carb} + \tau_{carb}}$$

$$\sum_i \tilde{q}^{*fd}_{i,carb} = \frac{(1-\alpha)\sum_i w_i}{p_{carb}} = \sum_i q^{*sq}_{i,carb}$$

The mental-budget affected voter overestimates a consumer's carbon emissions in Fee & Dividend by the portion of the dividend that would optimally be spent on clean consumption. And so we see under the Cobb-Douglas conditions, the mental budget notion turns Fee & Dividend from an emission reducer to a policy with precisely no effect on emissions.

Penalty & Reward addresses this by not having a dividend. Consider a relatively high carbon consumer. They are taxed above the population mean threshold and given no compensatory funds to allocate or misallocate. Consider a relatively low consumer

who receives a payment. This payment is not compensation for a tax, because for the low consumer, no tax exists. Therefore to allocate the reward payment to additional fossil fuel consumption does not seem logical nor an application of a known behavioral phenomenon. In fact, consumers may be averse to spending rewards on fossil fuels at all since such a choice will reduce the size of the next reward.

B Appendix 2: Diagnostic Experiment

B.1 Instruction summary

- Think of each round in isolation: a randomly drawn one determines your pay.
- In any round: you coexist with 1000 robots, and all 1,001 of you are each provided a menu of 4 objects and a 420 franc budget, always sufficient to buy either 1, 2, 3 or all 4 objects on the menu.[52]
- The only value to you of buying an object is its resale value. On each object you buy, your profit (or loss) is the resale value minus the price you paid. These rules also apply to the robots, whose resale values may differ from yours.
- Your choice is whether to accept or reject a proposal to introduce or remove a feature within the round that might help or hurt you. (See Figure 2 for an example.) Your decision about the feature is implemented, and we then automatically select the number of objects to buy that maximize your profit, given the feature decision you

[52]Francs are the experiment's currency units to facilitate arithmetic. 20 francs = $1.

made and do the same for the robots, as well. You are paid your final total of francs, each worth $0.05, for a random round.

- Next, we want you to assess the value of your choice, in francs, relative to the choice you rejected. Within the round, how many francs are you willing to exchange for your choice? Your incentives are to be completely honest and accurate here.

- More detail: 1 in 20 subjects will be randomly selected and receive 200 extra francs. If you are one of these subjects you must decide how many of these 200 francs you are willing to pay to receive your choice. A number will be drawn between 1 and 200. If the number of francs you were willing to offer is above the randomly drawn number, you will get your choice and we will deduct the value of the randomly drawn number, in francs from your final total. If the randomly drawn number between 1 and 200 is above the number of francs you are willing to offer, then you keep the 200 francs but your payout from purchasing objects will be using the Feature Choice that you rejected, not the one that you selected.

B.2 Implementation details

Subjects spent approximately 60 minutes in UC Santa Barbara's experimental economics laboratory on an incentivized experiment delivered via Qualtrics and earned on average $21.86. There was no time limit given, and some subjects spent as much as 90 minutes. For the diagnostic study detailed here, 56 subjects participated.

B.3 Results Supplement

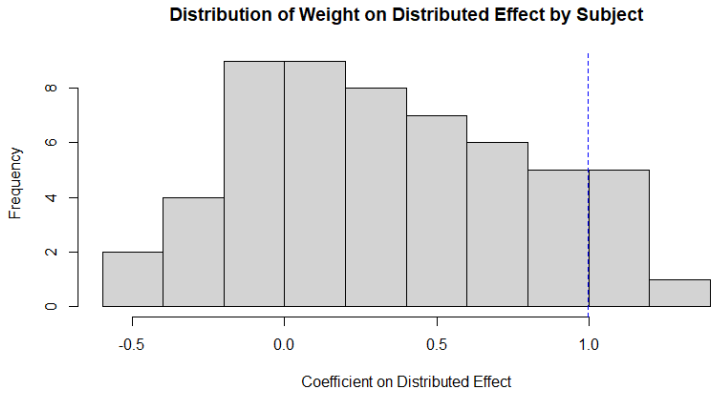


Figure 8: The distribution by individual lab subject of coefficient on Distributed Effect from regression in Equation 3.

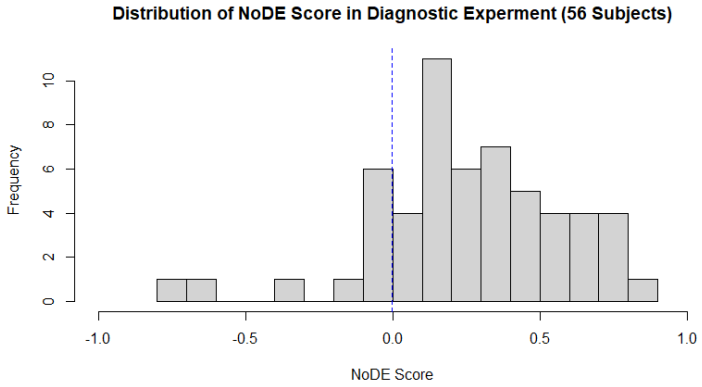


Figure 9: Distribution of subject NoDE Score applying methodology in Section 5.2 to Diagnostic Experiment data. To compare with survey participants' NoDE Score distribution, see Figure 12, lower right.

Table 4: Regressions from Diagnostic Experiment with Consumer and Distributed Effects

	Dependent variable:			
	Valuation of feature proposal			
	(1)	(2)	(3)	(4)
Constant	−3.416	5.062	−2.564	5.501
	(3.033)	(5.949)	(3.465)	(6.844)
Consumer Effect	0.978	0.866	1.155	1.113
	(0.117)	(0.119)	(0.168)	(0.202)
Distributed Effect	0.314***	0.210***	0.533***	0.500***
	(0.062)	(0.073)	(0.113)	(0.158)
Negative term magnitude (LAMA)		−0.258**		−0.043
		(0.103)		(0.173)
Policy magnitude (Policy Bias/Aversion)		0.200*		0.187
		(0.119)		(0.141)
Deviation magnitude (Default Bias/Aversion)		−0.227		−0.427
		(0.262)		(0.402)
Analysis on Subset With No Equilibrium Effect			X	X
Observations	1,680	1,680	632	632
Clusters	56	56	56	56

Notes: Result from estimating Equation 3 across all rounds in columns 1-2 and in the subset of rounds where Equilibrium Effects are absent in columns 3-4. Given the regression specifications, the rational benchmark for Consumer Effect, and Distributed Effect are 1. Rational benchmark for Negative term magnitude, Policy magnitude, and Deviation magnitude are 0. Negative coefficient for Negative term magnitude implies LAMA. *p<0.1; **p<0.05; ***p<0.01.

Table 5: Regressions from Diagnostic Experiment With Subeffects

	Dependent variable:	
	Valuation of feature proposal	
	(1)	(2)
Constant	−4.133	3.943
	(3.299)	(5.720)
Direct Effect	1.142	1.049
	(0.151)	(0.151)
Equilibrium Effect	0.809*	0.779*
	(0.126)	(0.151)
Distributed Direct Effect	0.523***	0.410***
	(0.103)	(0.112)
Distributed Equilibrium Effect	0.093***	−0.011***
	(0.065)	(0.069)
Negative term magnitude (LAMA)		−0.225**
		(0.101)
Policy magnitude (Policy Bias/Aversion)		0.042
		(0.105)
Deviation magnitude (Default Bias/Aversion)		−0.210
		(0.283)
Observations	1,680	1,680
Clusters	56	56

Notes: Result from estimating Equation 4 across all rounds. Given the regression specifications, the rational benchmark for Direct Effect, Equilibrium Effect, Distributed Direct Effect, and Distributed Equilibrium Effect are 1. Rational benchmark for Negative term magnitude, Policy magnitude, and Deviation magnitude are 0. Negative coefficient for Negative term magnitude implies LAMA. *p<0.1; **p<0.05; ***p<0.01.

B.4 Instructions and Comprehension Check

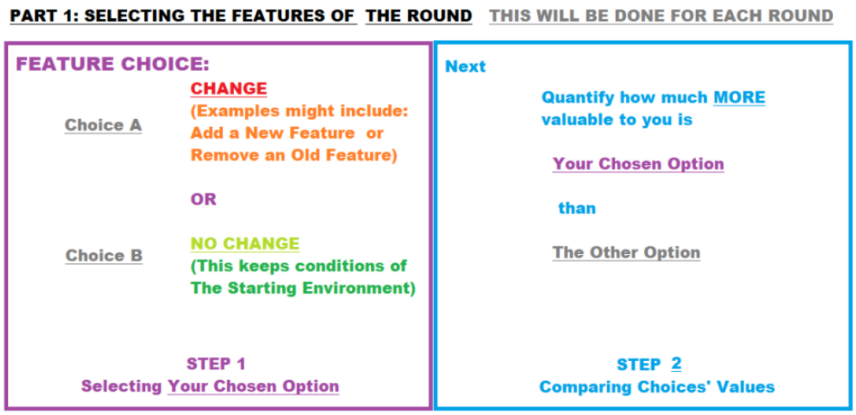
In this experiment, you will experience several **rounds** of choice tasks.

In **each round** you must imagine yourself as a resident of a country, in which the residents consist only of **you** and **1000 robots**. You are not playing with the other humans in the room today. Each of you is playing in your own decision environment that consist of just you and your own set of 1,000 robots.

In **each round**, the residents (all 1,001 of you) will each be provided a 420 Franc budget and menus of **4 objects**. (For reference, $1 = 20 Francs).

After you are done, one round is **randomly drawn**, and the choices you had made there will be used to determine your payment. This means you will **not** be able to save money from your budget from one round to the next.

Here is a first look at the process you will encounter in each round:



If you pick "Choice A" it is Your Chosen Option; "Choice B" is The Other Option.

If you pick "Choice B" it is Your Chosen Option; "Choice A" is The Other Option.

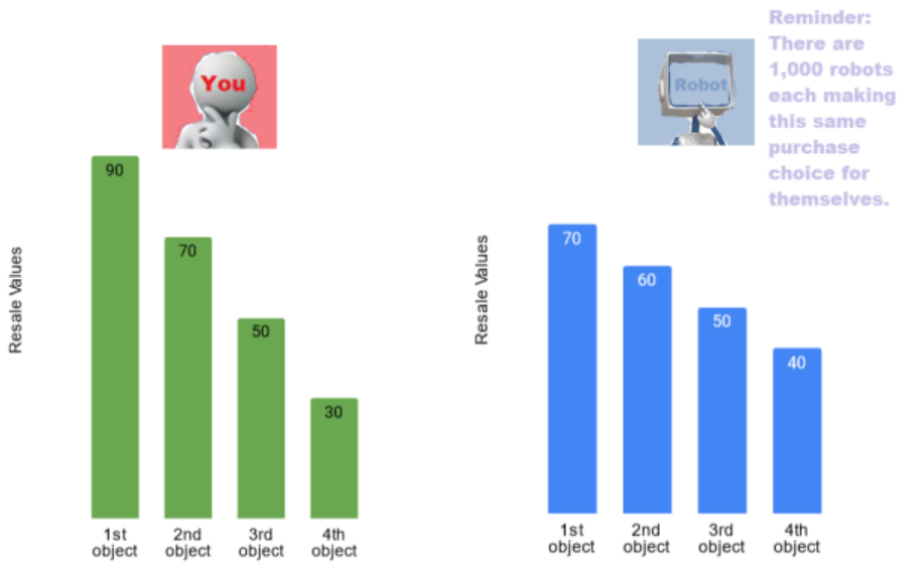
Any participant can buy 0-4 objects in their menu **in every round**, but they can also choose to buy **none**. Your budget of 420 Francs is always enough or more than enough to buy all the objects on your menu. But you may not want to buy them all.

In the round that is selected for payment, whatever you don't spend on objects is money that you keep. But becuse only one round matters, money is not moved between rounds.

There are **unlimited** objects. Even if every robot buys all 4 objects on its menu, it has **no** effect on your ability to buy any or all of the objects on your menu.

Each object carries a **resale value**: this is the only worth of the object to you.
For each object that you buy, you will experience either a **profit or loss** of the **resale value** of that object **minus** the **price** you paid. Reselling is done for you automatically at the end.

EXAMPLE:



Above is only an example of resale values. They will vary in all of the different rounds you will experience.

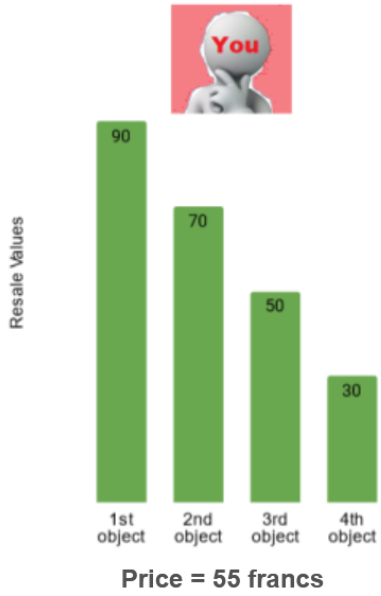
The same rules that apply to you also apply to each of the 1000 robots. But the robots could each have **different resale values** for their 1st, 2nd, 3rd and 4th object **than you**.

The things that are in **common** between you and the robots is the **price and special rules that we call features** (more about features later).

When it is time to make purchase choices, all objects, whether on your menu or on the robots menu, will carry the same price.

The **robots**, like you, would make purchase choices to **maximize their profit**. They are not concerned about helping or hurting you, just as you are not concerned about helping or hurting a robot, which is nothing more than some lines of software programming.

So what will a purchase decision look like? Let's continue with the **example** above:



The price shown is for **each** object that you buy. Let's try it out.

Purchase Choice:

- Buy 1st Object
- Buy 1st and 2nd Object
- Buy 1st, 2nd and 3rd Object
- Buy 1st, 2nd, 3rd and 4th Object
- Buy no object

Comprehension check: How much did you earn from the previous decision?

20 francs

50 francs

160 francs

195 francs

220 francs

240 francs

→

As you can see, once you buy the 1st object, which has the highest resale value, you can't buy it again. The next object to consider is 2nd object.

All robots will have the same resale values as each other, but their resale values will often be different than yours.

Step 2:
In every round, you will be asked about whether you want to **Add or Remove an added feature** to your universe.

Choose what is best for you. In the one round that we randomly select to determine your payment, there is a 95 percent chance we will automatically use **your choice** to determine whether have an added feature or don't have it. The robots don't get a vote. So your choice of feature is highly likely to determine the environment in which the purchase decision (that matters for your pay) will ultimately be made.

Let's revisit our **example**, now with a **choice** to make about a new **feature**. From now on, resale values will be visible by clicking on labeled **buttons**.

It is your choice whether to use these buttons. It is not required.

Feature Choice:

New Feature:
- Whoever purchases an object gets a 20 franc **bonus** for each object that they buy.
- The **cost** of each 20 franc bonus is divided equally and paid by the 1,001 residents of the country.

No Change

Price = 55 francs

Your Resale Values | Robot Resale Values

Assume this round will determine your payment.

Next, suppose we give you an extra 200 francs ($10).
But, instead of giving you the choice you just selected, the choice that you just **rejected** becomes the environment in which you and robots must make purchasing decisions.

Of the 200 francs you were just given, what **portion** are you **willing to trade** to get back to **Your Choice**. You keep the remaining francs.



In other words, what is it worth to you to get your choice vs. the other option.

(Honesty is the best for you here. If you'd like us to prove why, you can click the button below)

More Info

"More info" button opens following text explanation of willingness to pay mechanism:

This method ensures that the optimal answer is your honest opinion of how much more, in francs, you like one choice than the other.

Put a number above your true personal value?
You might overspend for the choice you do like.

Put a number below your true personal value?
You're likelier to get the choice you don't like.

How it works:

When you select between, for example, "Feature X" or "No Change":

There is a 95 out of 100 chance that you'll gete your choice

There is a 5 out of 100 chance that you'll get the other option that you rejected, plus 200 francs.

After you decide your value for your original choice compared to the alternative by using the slider bar, we save the number and call it Y. Next, I randomly draw a number between 1 and 200 and call it X.

If Y > X , then we give you your original choice and deduct only X, which is less than your value Y, so it's a good deal for you still.

If X > Y, then we deduct nothing, and you remain in the environment of the rejected choice, and you keep the 200 francs.

This may seem complicated, but you can ignore the details and trust that it's designed to incentivize you to put your actual true maximum willingness to pay for your choice over the alternative, instead of underreporting or overreporting that value.

Very Important: Please Keep in mind when you make feature choices, that your **final purchase choices** for the round and the robots' final purchase for the round, would be made **after the feature settings are revealed**. In other words, you would know exactly what the feature and prices are.

Finally, to make things easier, after the features are set (based very likely on your selections), we make the optimal purchase decision for you, given all the best and current information (features, prices, etc.). In other words, given the feature that you select, we will make sure your purchase choice is the best one possible.

→

Here are some comprehension questions:

In this experiment, you maximize your pay by:

- Making feature choices to beating the robots.
- Making feature choices that are best for your finances.
- Making feature choices and purchase choices that are good for you and good for robots.

I use the slider bar to ...

- Express my subjective value for my choice in the Purchase Choice segment.
- Express my subjective value for my choice in the Feature Choice segment.
- Express my subjective value for how much **more I value** my choice **than I value** the rejected choice in the Feature Choice segment.

Suppose a specific round is selected for payment. If I didn't spend all of my budget on objects ...

That money I didn't spend disappears and nobody gets it.

That leftover money gets carried over to the next round

I get to keep the money I didn't spend and add it to profit I might have made from any objects that I bought in that round. For the one round randomly selected for my pay, this total amount will determine my $ earnings.

When considering a feature choice, what should I do?

Consider potential benefits only

Consider potential costs only

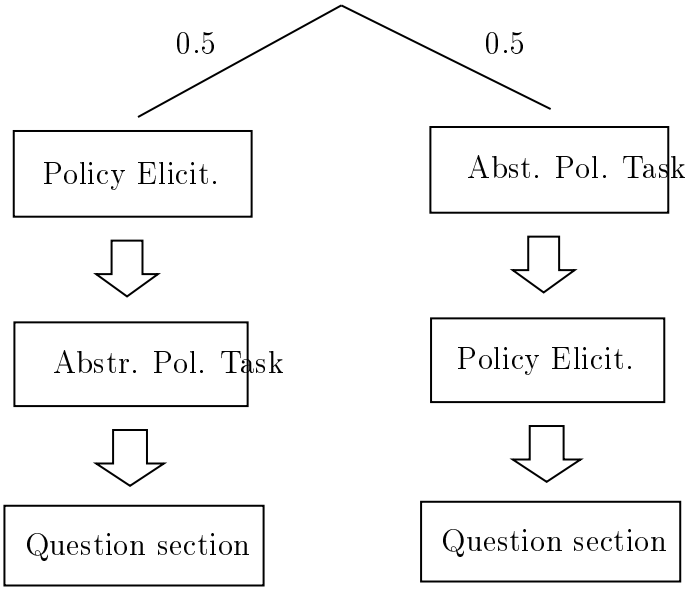
Consider both potential costs and benefits by carefully reading every line in the description.

C Appendix 3: Survey Experiment

C.1 Implementation details

C.1.1 Policy section

The survey consists of three main sections: Policy Elicitation, Policy Lab Task, and Questions section. One of two possible orders is randomly assigned, as shown below.



Every subject encounters all four of the rating questions for each of the two policies. The four prompts, whose sequences are represented in Figure 10, read:

- Overall support: "For each policy, provide your overall degree of support or opposition."
- Impact on emissions rating: "Please rate each policy, solely on its effectiveness for reducing emissions."
- Financial impact on you rating: "Please rate each policy, solely on its financial impact to your household."

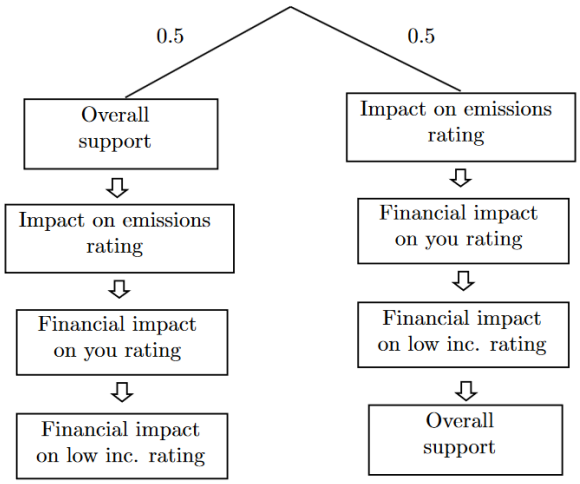


Figure 10: Policy question order: Participants are then randomly assigned to one of two question orders.

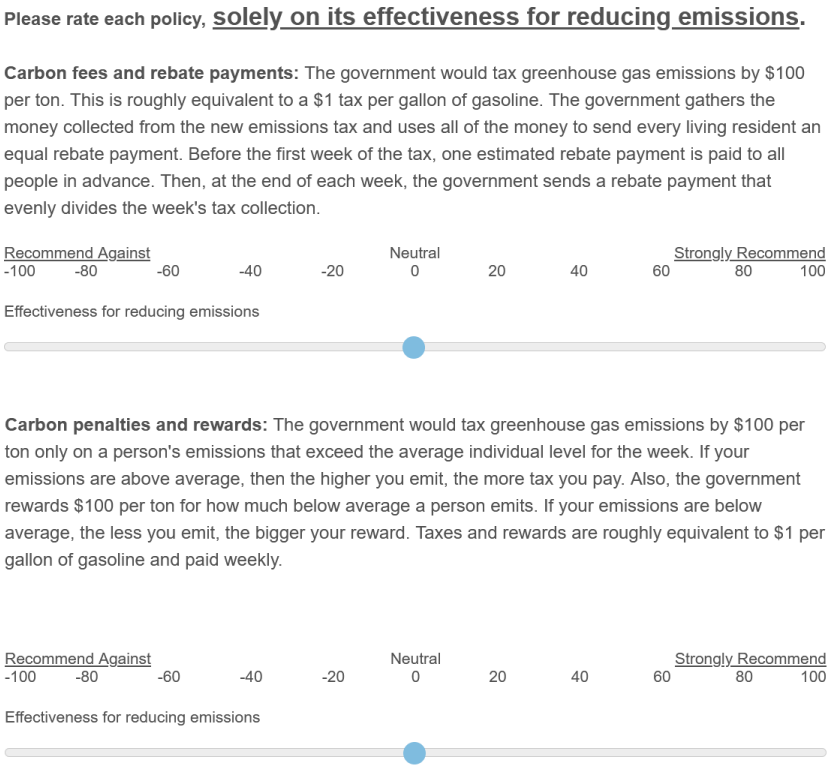


Figure 11: The layout for the rating question pertaining to effectiveness of the policy for reducing emissions. This is only one example of the factor questions that complement the overall support rating. All four questions share this layout.

- Financial impact on low inc. rating: "Please rate each policy solely on its financial impact on low income individuals."

C.1.2 Question Section

What follows is a list of questions that were consistently included in the exit questionnaire sections of each survey distribution. (Age, sex, party affiliation, employment and student status were provided directly through Prolific's screening questionnaire for the participants who agreed to provide and release those details.)

State of residence was asked to monitor the geographic diversity and representativeness of the sample:

In what state do you currently reside? The drop-down menu included 50 states, the District of Columbia and Puerto Rico.

The next panel of six questions measure predicted factors for support for a carbon pricing policy like Fee & Dividend or Penalty & Reward. For the six questions, I had no ex ante reason to think that any of these factors would be predictive of a difference in support between the two policies. In theory, political background, climate beliefs, personal emissions, and the other factors should affect support for the two policies identically.

Here is a 9-point scale on which the political views that people might hold are arranged from very liberal (left) to very conservative (right). Where would you place yourself on this scale? In the numerical score, higher values correspond to further right politically. This response allows a finer and more current measure of partisanship than the party affiliation variable provided through Prolific.

How confident are you in the following statement? Climate change is real and caused by human activity. A slider bar is provided ranging from 0 (Not confident at all) to 100 (Entirely confident).

How much faith and trust do you place in how the government spends money? Responses range on a scale of 1 (Not well at all) to 5 (Extremely well).

What do you guess the effect of a carbon tax on your income (from your job or other sources) would be? Choices are "Positive," "Negative" or "No effect," and the order in which these options are shown is randomized. In the analysis I transform this question into two dummy variables for "Positive" and "Negative."

It is appropriate/inappropriate for the government to sometimes redistribute some resources, including money, from the wealthy, high-earning to others, such as the poor or middle class. A slider bar ranging from -100 (Totally inappropriate) to +100 (Totally appropriate) measures the response.

Consider the average level of carbon emissions in the country. How much higher or lower do you think your emissions total is, relative to the average level? By what percent? I combine the response to the first question of "Higher" or "Lower" with the open-ended percent response to form a numerical composite variable, where responses can range from -1 (implying 100 percent less than the average level, or zero emissions) to +10 (Implying 1,000 percent, or ten times the average level of emissions).

Finally, two questions each measure additional behavioral factors that could theoretically result in difference of support between the two policies.

Do you agree with the following statement? Because carbon taxes in the "Carbon fees and rebate payments" policy make fossil fuels more expensive than before, the payments that people get back from the policy will be spent on their fossil fuel budget. (Participants are shown the text of the "Carbon fees and rebate payments" policy in case the need a reminder.) Choices are "Yes," "No," and "I don't know," which I transform into a dummy variable measuring 1 if the response was either "Yes" or "I don't know." This question is a measure of the application of mental budgeting heuristic detailed in the previous section.

Subjects are shown the text of "Carbon fees and rebate payments," and asked: **True or False? Under the policy, I should not change my behavior because the more fossil fuel tax I pay, the more I'll get back in payments.** If subjects choose "True," a dummy variable takes the value of 1. This question measures whether subjects misunderstand that the dividend in a Fee & Dividend policy is essentially uncorrelated with the taxes paid.

C.2 Results Supplement

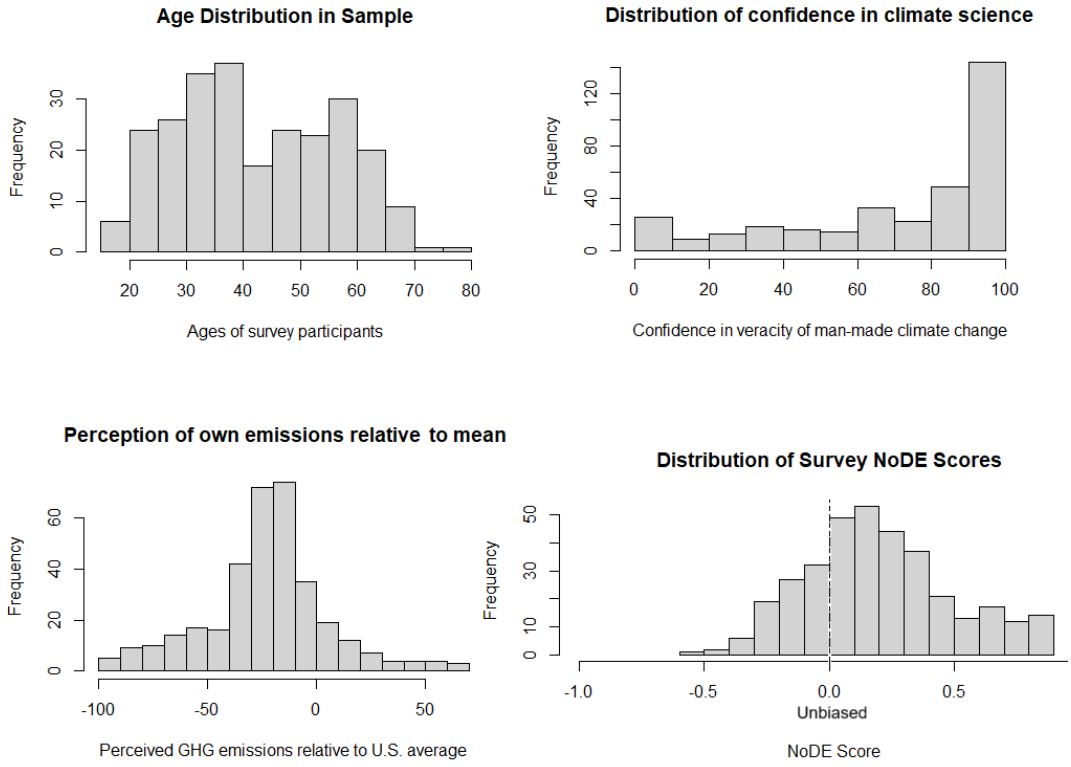


Figure 12: Some descriptive properties that characterize the sample. More participants believe themselves to be below-average emitters than above-average emitters. By far, more subjects demonstrate NoDE, or low weighting to Distributed Effects, in the incentivized, Abstract Policy Task than high weighting to Distributed Effects.

Table 6: Effect of Beh. Factors on Diff. in Support, Extended to Show Controls

	Dep. Var.: Penalty & Reward Premium		
	(1)	(2)	(3)
Constant	35.767***	4.109	−12.877
	(8.245)	(21.035)	(35.353)
NoDE Score	35.310**	29.689**	48.185***
	(14.664)	(15.028)	(17.279)
Correlation error	14.812	16.090	19.612
	(12.267)	(11.896)	(14.010)
Mental budgeting error	16.179**	15.609**	17.880**
	(7.462)	(7.487)	(8.798)
Survey order: experiment first	−13.988*	−14.349*	−18.597*
	(7.966)	(8.092)	(9.740)
"Compute" prompt	−13.092	−16.164*	−2.544
	(9.823)	(9.793)	(11.489)
"Expert" prompt	−4.498	−0.869	1.624
	(10.248)	(11.038)	(13.486)
Relative emissions		0.058	0.094
		(0.134)	(0.148)
Belief in climate change science		0.535***	0.557**
		(0.185)	(0.223)
Left right score		1.241	3.269
		(1.874)	(2.472)
Believe job benefits from carbon tax		−11.910	−6.947
		(10.259)	(11.245)
Believe job hurt from carbon tax		3.153	−3.898
		(9.784)	(11.156)
Trust government use of revenue		−4.206	−2.183
		(4.960)	(5.976)
Support for redistribution		−0.044	−0.027
		(0.096)	(0.117)
Black			−11.620
			(13.795)
Student			16.843
			(11.383)
Unemployed			−13.027
			(12.736)
Male			3.307
			(9.581)
Age			−0.008
			(0.385)
Observations	347	347	253

Notes: The table above results from estimating Equation 5 on the survey participant responses. The dependent variable is the Penalty & Reward premium, i.e. the difference in each subject's support for Penalty & Reward versus Fee & Dividend. *p<0.1; **p<0.05; ***p<0.01.

Table 7: Baseline Study: Policy Support/Opposition

	Dependent variable: Policy Rating (-100 to +100)			
	Fee & Dividend		Penalty & Reward	
	(1)	(2)	(3)	(4)
Constant	−23.447	−31.599	−23.902	−34.249
	(20.523)	(30.245)	(18.101)	(25.061)
NoDE Score	−19.792*	−28.567**	0.790	8.071
	(11.679)	(13.946)	(11.467)	(11.846)
Correlation error	−24.112**	−18.302	1.069	3.891
	(10.294)	(13.307)	(8.689)	(10.250)
Mental budgeting error	−9.151	−6.292	2.057	6.481
	(6.887)	(8.746)	(6.275)	(6.848)
Relative emissions	−0.329***	−0.346**	−0.317***	−0.371***
	(0.121)	(0.139)	(0.113)	(0.120)
Belief in climate change science	0.241	0.260	0.813***	0.817***
	(0.176)	(0.201)	(0.164)	(0.171)
Left right score	−1.608	−1.596	1.198	2.454
	(1.786)	(2.650)	(1.295)	(1.587)
Believe job benefits from carbon tax	10.166	8.970	−2.843	0.342
	(9.971)	(11.709)	(7.399)	(8.359)
Believe job hurt from carbon tax	−19.210**	−15.477	−17.656**	−18.300**
	(8.563)	(9.496)	(7.009)	(7.502)
Trust government use of revenue	7.947	7.930	3.380	8.368*
	(4.911)	(6.170)	(3.822)	(4.375)
Support for redistribution	0.087	0.082	0.064	0.084
	(0.084)	(0.105)	(0.068)	(0.075)
Survey order: experiment first	7.629	16.452*	−9.783	−5.293
	(7.180)	(8.793)	(6.368)	(7.120)
Black		25.485*		10.187
		(13.232)		(7.535)
Student		−6.352		1.209
		(10.805)		(7.606)
Unemployed		22.917*		1.708
		(12.209)		(7.291)
Male		−9.638		−7.363
		(9.766)		(7.493)
Age		0.071		−0.199
		(0.304)		(0.307)
Observations	248	184	248	184
R^2	0.257	0.282	0.341	0.416
Adjusted R^2	0.223	0.213	0.310	0.360

Notes: The table above results from estimating a set of explanatory variables on elicitation of support for each policy, Fee & Dividend in columns 1 and 2, and Penalty & Reward in columns 3 and 4. This estimation is on the baseline participants not receiving a supplemental information treatment. *p<0.1; **p<0.05; ***p<0.01.

Table 8: Predictiveness of Impact factors on overall policy support rating

	Dependent variable: Overall Support (-100 to +100)	
	Fee & Dividend	Penalty & Reward
	(1)	(2)
Constant	-5.274^{**}	3.470
	(2.078)	(3.179)
Emissions impact rating (-100 to +100)	0.398^{***}	0.430^{***}
	(0.058)	(0.075)
Your Finances impact rating (-100 to +100)	0.290^{***}	0.341^{***}
	(0.072)	(0.087)
Low Income impact rating	0.246^{***}	0.202^{***}
	(0.061)	(0.068)
Observations	347	347
R^2	0.608	0.597
Adjusted R^2	0.604	0.594

Notes: The Emissions impact rating was solicited by asking for both policies,"Please rate each policy, solely on its effectiveness for reducing emissions." Financial impact rating asked: "Please rate each policy, solely on its financial impact to your household." And Low Income impact rating: "Please rate each policy solely on its financial impact on low income individuals." *p<0.1; **p<0.05; ***p<0.01.

Table 9: Wilcoxon Rank Sum Test With Continuity Correction (p-value)

Treatment	P&R = F&D	Diff(HNoDE)=Diff(LNoDE)	Diff(Left)=Diff(Right)	P&R(Right)=F&D(Left)
BASELINE	$< 2e-16^{***}$	0.0002^{***}	0.0367^{**}	0.0044^{***}
T:COMPUTE	0.0154^{**}	0.0338^{**}	0.3429	0.151
T:EXPERT	0.0006^{***}	0.1897	0.1992	0.5762

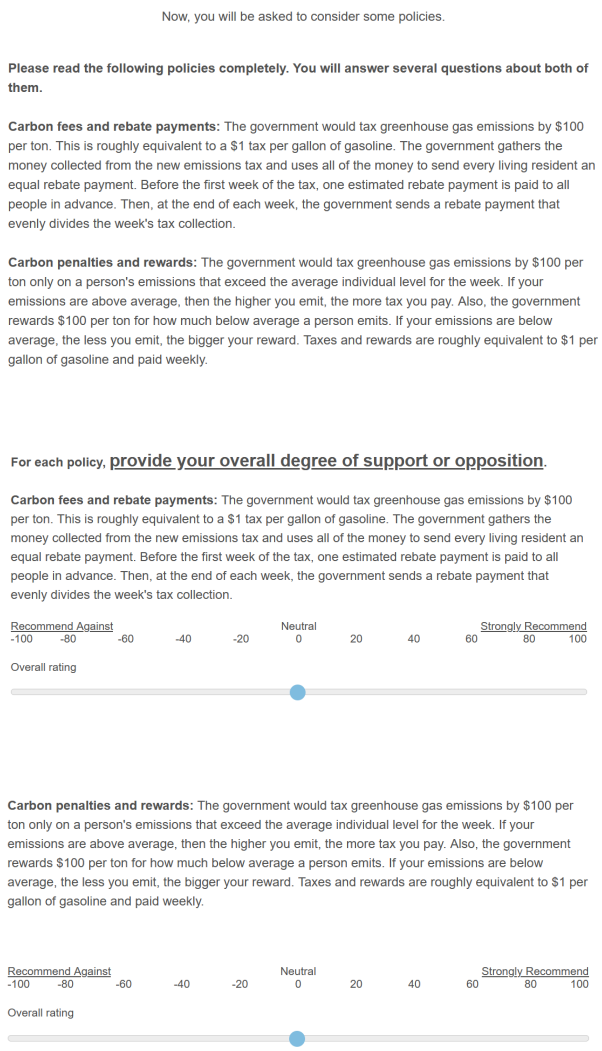
The p-values, $< 2e-16^{***}$ and 0.0002^{***} correspond to Hypotheses 1 (Penalty & Reward is better supported than Fee & Dividend) and Hypothesis 2 (Penalty Reward Premium is higher among high NoDE than among low NoDE subjects), respectively.

Table 10: Robustness to Single Policy Presentation (Between Subjects)

	Dependent variable:
	Support for Policy
Constant (Fee & Dividend)	-18.814^{***}
	(6.580)
Fee & Dividend With "Compute" Treatment	10.814
	(13.719)
Fee & Dividend With "Expert" Treatment	13.510
	(10.962)
Penalty & Reward	37.826^{***}
	(8.159)
Observations	331

Notes: In Table 10, above, subjects either encountered a basic Fee & Dividend policy (represented as the constant term in the regression), the "Compute" assisted Fee & Dividend policy (represented as the first treatment dummy in the regression), the "Expert" assisted Fee & Dividend policy (the second treatment dummy), or Penalty & Reward (the third treatment dummy). Higher support for Penalty & Reward is clearly robust to sequential policy presentation, but the magnitude of the gap in support intensity is a more modest 38 points than the 49 point gap observed in the simultaneous-presentation sample discussed in Section 6. While not statistically significant, the effect magnitude of information treatments is modestly larger in the robustness sample. $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

C.3 Full Survey



Please rate each policy, solely on its effectiveness for reducing emissions.

Carbon fees and rebate payments: The government would tax greenhouse gas emissions by $100 per ton. This is roughly equivalent to a $1 tax per gallon of gasoline. The government gathers the money collected from the new emissions tax and uses all of the money to send every living resident an equal rebate payment. Before the first week of the tax, one estimated rebate payment is paid to all people in advance. Then, at the end of each week, the government sends a rebate payment that evenly divides the week's tax collection.

Recommend Against — Neutral — Strongly Recommend
-100 -80 -60 -40 -20 0 20 40 60 80 100

Effectiveness for reducing emissions

Carbon penalties and rewards: The government would tax greenhouse gas emissions by $100 per ton only on a person's emissions that exceed the average individual level for the week. If your emissions are above average, then the higher you emit, the more tax you pay. Also, the government rewards $100 per ton for how much below average a person emits. If your emissions are below average, the less you emit, the bigger your reward. Taxes and rewards are roughly equivalent to $1 per gallon of gasoline and paid weekly.

Recommend Against — Neutral — Strongly Recommend
-100 -80 -60 -40 -20 0 20 40 60 80 100

Effectiveness for reducing emissions

Please rate each policy, solely on its financial impact to your household.

Carbon fees and rebate payments: The government would tax greenhouse gas emissions by $100 per ton. This is roughly equivalent to a $1 tax per gallon of gasoline. The government gathers the money collected from the new emissions tax and uses all of the money to send every living resident an equal rebate payment. Before the first week of the tax, one estimated rebate payment is paid to all people in advance. Then, at the end of each week, the government sends a rebate payment that evenly divides the week's tax collection.

Recommend Against — Neutral — Strongly Recommend
-100 -80 -60 -40 -20 0 20 40 60 80 100

Financial impact to your household

Carbon penalties and rewards: The government would tax greenhouse gas emissions by $100 per ton only on a person's emissions that exceed the average individual level for the week. If your emissions are above average, then the higher you emit, the more tax you pay. Also, the government rewards $100 per ton for how much below average a person emits. If your emissions are below average, the less you emit, the bigger your reward. Taxes and rewards are roughly equivalent to $1 per gallon of gasoline and paid weekly.

Recommend Against — Neutral — Strongly Recommend
-100 -80 -60 -40 -20 0 20 40 60 80 100

Financial impact to your household

Please rate each policy, solely on its financial impact to low-income individuals.

Carbon fees and rebate payments: The government would tax greenhouse gas emissions by $100 per ton. This is roughly equivalent to a $1 tax per gallon of gasoline. The government gathers the money collected from the new emissions tax and uses all of the money to send every living resident an equal rebate payment. Before the first week of the tax, one estimated rebate payment is paid to all people in advance. Then, at the end of each week, the government sends a rebate payment that evenly divides the week's tax collection.

Recommend Against — Neutral — Strongly Recommend
-100 -80 -60 -40 -20 0 20 40 60 80 100

Financial impact to low-income people



Carbon penalties and rewards: The government would tax greenhouse gas emissions by $100 per ton only on a person's emissions that exceed the average individual level for the week. If your emissions are above average, then the higher you emit, the more tax you pay. Also, the government rewards $100 per ton for how much below average a person emits. If your emissions are below average, the less you emit, the bigger your reward. Taxes and rewards are roughly equivalent to $1 per gallon of gasoline and paid weekly.

Recommend Against — Neutral — Strongly Recommend

-100 -80 -60 -40 -20 0 20 40 60 80 100

Financial impact to low-income people

Now comes a section in which you can earn some extra real money by choosing well.

You will experience 15 independent **rounds.**
(Nothing from one round can affect anything in another round.)

Only one round, which is randomly selected, will be used to determine if you will earn a bonus payment.

In **each round**, you and 1,000 others are called Franklians.
Money is measured in units called francs.

The 1,000 other Franklians are **robots**.

All 1,001 Franklians will each be provided
a 500 franc budget and the option to buy some objects.

Any Franklian (you and the 1,000 robots)
can buy 0, 1, 2, 3, or 4 objects **in any round.**

Everyone's budget is always enough
to buy up to the limit of 4 objects,
and the supply of objects never runs out.

Why might you buy an object?
Each object carries a **resale value**,
which is the only worth of the object to you.

AN EXAMPLE OF YOUR RESALE VALUES:



Price = 55 francs

These resale values can vary between rounds.

Your **profit (or loss)** from **buying** any object
is the **resale value** of that object
minus the **price** you paid.

Reselling is done for you automatically at the end of the round.
The money from reselling any objects that you may have bought
plus unspent money from your budget will go into your **Final Money**.

Each object has the same price. In the example, it's 55 francs per object.

Let's try a purchase decision, using the resale values shown above.

Purchase Choice:

- Buy 1st Object
- Buy 1st and 2nd Object
- Buy 1st, 2nd and 3rd Object
- Buy 1st, 2nd, 3rd and 4th Object
- Buy no object

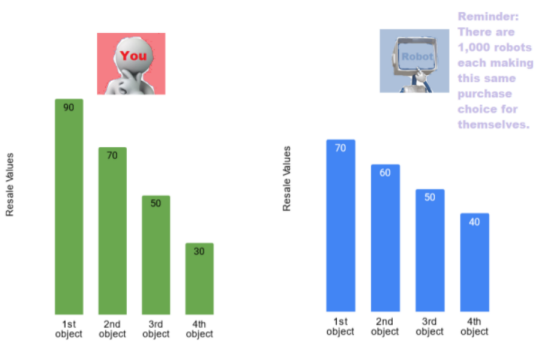
Comprehension check: How much did you earn from the previous decision?

- 20 francs
- 50 francs
- 160 francs
- 195 francs
- 220 francs
- 240 francs

Great! So now you know what buying
(or not buying) objects is like.

The same rules that apply to you
also apply to each of the 1000 robots,
who can also buy objects and affect
their own **Final Money**.

But the robots could (and probably do) have **different resale values**
for their own 1st, 2nd, 3rd and 4th objects than you do.
As a result they may often make different purchase choices than you.



Each robot, like you, would make purchase choices
to maximize their own **Final Money**.
They are not interested in helping or hurting you.

The things that is in common between you
and all of the the robots is the **price per object**
and special rules called **Features.**

A Feature changes the cost of buying an object,
and can affect your **Final Money** in many ways.
You should read all lines of a Feature carefully.

In every round, you will have to make a **Feature Choice**:
Whether you want to **Add** or **Remove** a **Feature**.

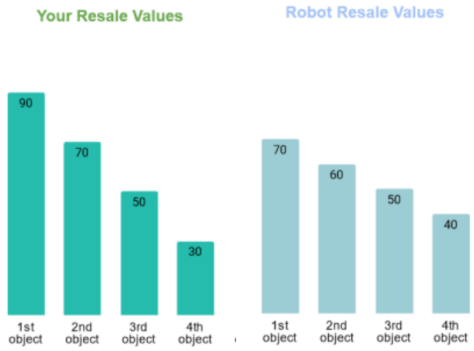
Your choice applies just for that round.

The Feature Choice is the decision on which you will be judged.
Once you decide about the Feature, your choice will be the law for that round.

Then, we will automatically make the best purchase choice
of 0,1,2,3, or 4 objects for you and all of the robots.
(We will only purchase an object, for you or for a robot,
if the resale value of that object is a bigger number
than the total cost of buying an object.)

Let's revisit our example:

Price = 55 francs



Feature Choice:

New Feature:
- Whoever purchases an object gets a 20 franc bonus for each object that they buy.
- The 20 franc cost of each bonus is divided equally and paid for by all 1,001 Franklians, including you.

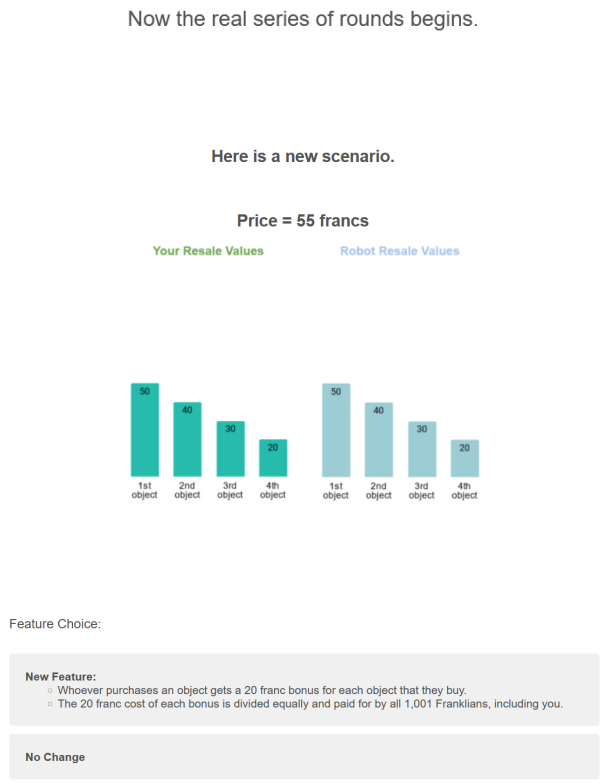
No Change

Note that the Feature had two lines,
both of which would ultimately affect your **Final Money**.

If you are randomly selected to be potentially paid a bonus,
we will select **one round** at random.
Your pay is completely decided by your **Feature Choice** in the selected round.

If you make a Feature Choice that results
in a larger amount of **Final Money** than the other choice,
you will earn a **$1 bonus**.

Note: in any round, the subject first sees the initial scenario alone, then must click "Next" to move to a slide where the initial scenario is accompanied by the Feature Choice being offered in the round. To avoid unnecessary duplication, the images below do not include the initial slide, only the Feature Choice slide, which includes the content of that round's initial slide (resale values, price, and any other initial features.)

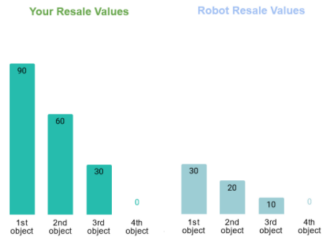


Here is a new scenario.

Price = 55 francs

Existing Feature:

- Whoever purchases an object gets a 20 franc bonus for each object that they buy.
- The 20 franc cost of each bonus is divided equally and paid for by all 1,001 Franklians, including you.



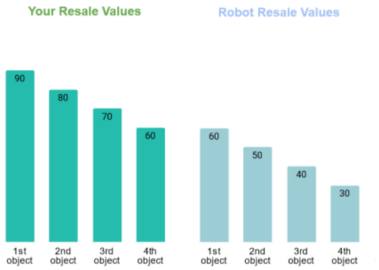
Feature Choice:

Remove 'Existing Feature'

No Change

Here is a new scenario.

Price = 75 francs



Feature Choice:

New Feature:
- Whoever purchases an object pays an additional 20 franc fee for each object that they buy.
- Every 20 franc fee collected is divided and delivered as an equal payment to all 1,001 Franklians, including you.

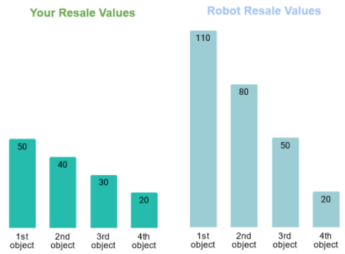
No Change

Here is a new scenario.

Price = 55 francs

Existing Feature:

- Whoever purchases an object pays an additional 20 franc fee for each object that they buy.
- Every 20 franc fee collected is divided and delivered as an equal payment to all 1,001 Franklians, including you.



Feature Choice:

Remove 'Existing Feature'

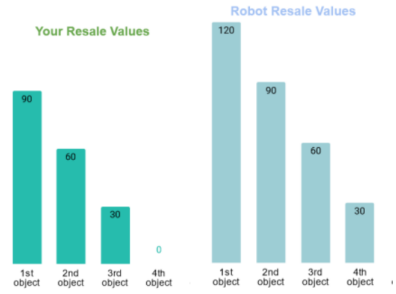
No Change

Here is a new scenario.

Price = 65 francs

Existing Feature:

- Whoever purchases an object pays an additional 20 franc fee for each object that they buy.
- Every 20 franc fee collected is divided and delivered as an equal payment to all 1,001 Franklians, including you.



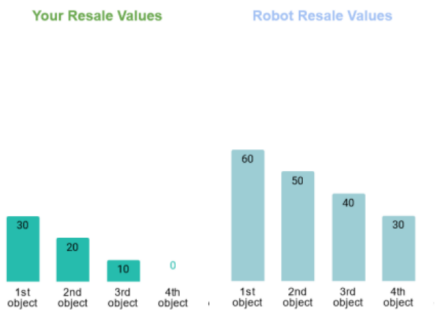
Feature Choice:

Remove 'Existing Feature'

No Change

Here is a new scenario.

Price = 55 francs



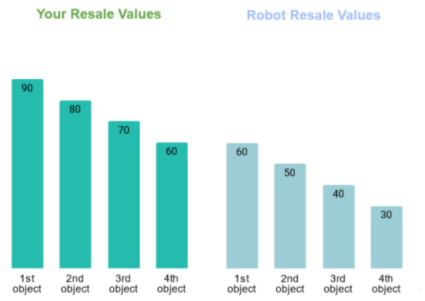
Feature Choice:

New Feature:
- Whoever purchases an object gets a 20 franc bonus for each object that they buy.
- The 20 franc cost of each bonus is divided equally and paid for by all 1,001 Franklians, including you.

No Change

Here is a new scenario.

Price = 85 francs



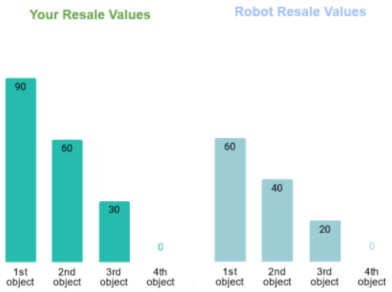
Feature Choice:

New Feature:
- Whoever purchases an object gets a 20 franc bonus for each object that they buy.
- The 20 franc cost of each bonus is divided equally and paid for by all 1,001 Franklians, including you.

No Change

Here is a new scenario.

Price = 45 francs



Feature Choice:

New Feature:
- Whoever purchases an object pays an additional 10 franc fee for each object that they buy.
- Every 10 franc fee collected is divided and delivered as an equal payment to all 1,001 Franklians, including you.

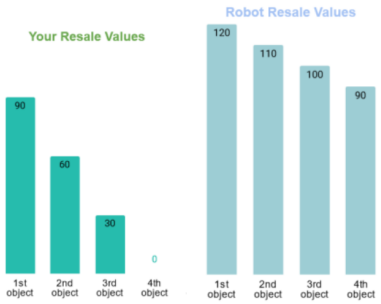
No Change

Here is a new scenario.

Price = 85 francs

Existing Feature:

- Whoever purchases an object gets a 20 franc bonus for each object that they buy.
- The 20 franc cost of each bonus is divided equally and paid for by all 1,001 Franklians, including you.



Feature Choice:

Remove 'Existing Feature'

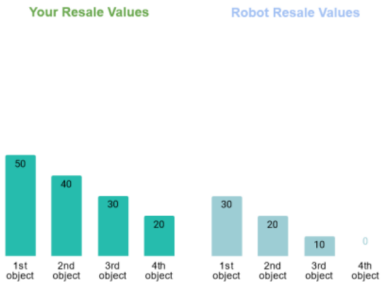
No Change

Here is a new scenario.

Price = 35 francs

Existing Feature:

- Whoever purchases an object pays an additional 20 franc fee for each object that they buy.
- Every 20 franc fee collected is divided and delivered as an equal payment to all 1,001 Franklians, including you.



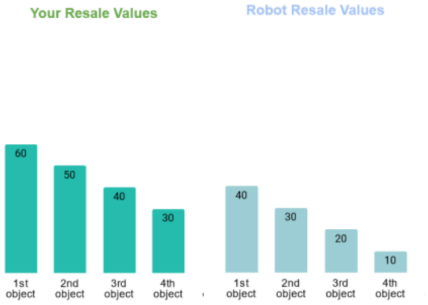
Feature Choice:

Remove 'Existing Feature'

No Change

Here is a new scenario.

Price = 65 francs



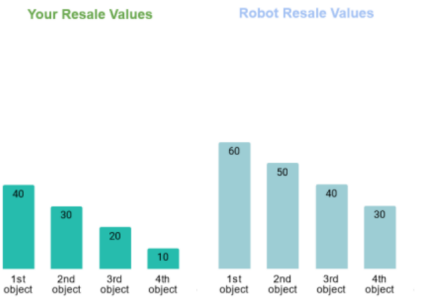
Feature Choice:

New Feature:
- Whoever purchases an object gets a 20 franc bonus for each object that they buy.
- The 20 franc cost of each bonus is divided equally and paid for by all 1,001 Franklians, including you.

No Change

Here is a new scenario.

Price = 65 francs



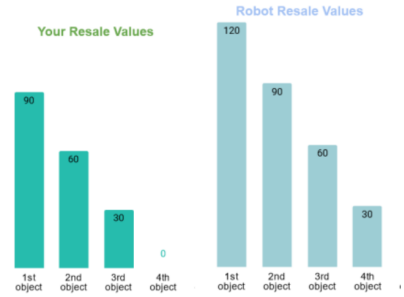
Feature Choice:

New Feature:
- Whoever purchases an object gets a 20 franc bonus for each object that they buy.
- The 20 franc cost of each bonus is divided equally and paid for by all 1,001 Franklians, including you.

No Change

Here is a new scenario.

Price = 65 francs



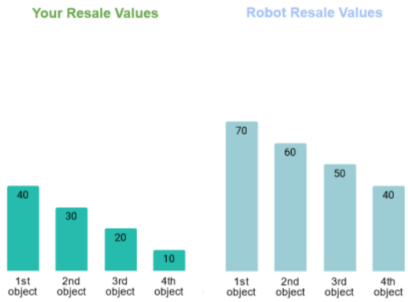
Feature Choice:

New Feature:
- Whoever purchases an object pays an additional 20 franc fee for each object that they buy.
- Every 20 franc fee collected is divided and delivered as an equal payment to all 1,001 Franklians, including you.

No Change

Here is a new scenario.

Price = 45 francs



Feature Choice:

New Feature:
- Whoever purchases an object pays an additional 20 franc fee for each object that they buy.
- Every 20 franc fee collected is divided and delivered as an equal payment to all 1,001 Franklians, including you.

No Change

In which state do you currently reside?

Here is a 9-point scale on which the political views that people might hold are arranged from very liberal (left) to very conservative (right). Where would you place yourself on this scale?

Very liberal 1 2 3 4 5 6 7 8 9 Very conservative

Political Ideology

How confident are you in the following statement? Climate change is real and caused by human activity.

0 10 20 30 40 50 60 70 80 90 100

Your confidence level

How much faith and trust do you place in how the government spends money?

- Not well at all
- Slightly well
- Moderately well
- Very well
- Extremely well

Consider the average level of carbon emissions in the country. How much higher or lower do you think **your** emissions total is, relative to the average level?

- Higher
- Lower

By what percent?

Do you agree with the following statement?
Because carbon taxes in the "Carbon fees and rebate payments" policy make fossil fuels more expensive than before, the payments that people get back from the policy will be spent on their fossil fuel budget.

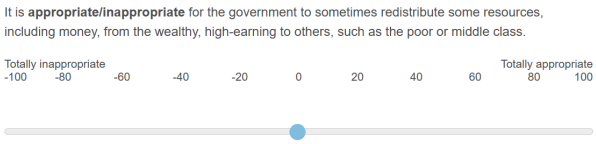
- Agree
- Disagree
- I don't know

True or False?
Under the policy, I should not change my behavior because the more fossil fuel tax I pay, the more I'll get back in payments.

- True
- False

What do you guess the effect of a carbon tax on your income (from your job or other sources) would be:

- Positive
- Negative
- No major effect


It is **appropriate/inappropriate** for the government to sometimes redistribute some resources, including money, from the wealthy, high-earning to others, such as the poor or middle class.
Totally inappropriate
Totally appropriate
-100 -80 -60 -40 -20 0 20 40 60 80 100

D Appendix 4: Implementing Penalty & Reward

When a single corporation or administrator oversees the usage of a the good or service, as is the case for local toll roads, electricity, and water providers, administering congestion and conservation pricing using Penalty & Reward is straightforward. With records of individual usage, no additional data gathering is required. Penalty & Reward can offer a popular, transparent way to accomplish the goal of reducing usage at peak times by helping people see that (1) the goal of the policy is not to "price-gouge" or obtain revenue, and (2) they may stand to gain financially from the policy.

With carbon, the consumer does not usually transact with a sole producer, making observability more difficult. For this reason, Fee & Dividend, with taxes levied at the production of emissions, is considered ideal for implementation.[53] Still, a version of Penalty & Reward can be implemented using carbon credits.[54] Every consumer is issued a quantity of carbon credits determined by the population average consumption level in the nearest prior period. These credits can be thought of as coupons that exempt purchases from the carbon tax at the time of purchase or redeemable at any time for equivalent cash value.[55] This implementation of Penalty & Reward is not game-able. In Penalty & Reward, the credits are based on the population mean consumption in a prior period, a figure no sole consumer can meaningfully influence.

[53]There may be disadvantages to this approach, however, given that tax salience to consumers can deliver outsize reductions in externalities relative to commensurate effects of passed-through prices alone. Evidence can be found in Andersson (2019) and Murray and Rivers (2015).

[54]To distinguish from cap-and-trade, a credit's price is fixed, not quantity.

[55]In the language of Fee & Dividend, these coupons are essentially dividends - they are redeemable for the equivalent amount of money. But perceptually they can also be more readily perceived as buffering consumers from taxes, up to a threshold.

As an example, consider the taxation of transportation emissions through gasoline taxes. Consumers could be issued a card that they use to exempt themselves, up to a point, from taxes on gasoline purchases. Using the card is unnecessary and merely an ex ante acceptability tool, as credits can also be immediately redeemed as cash for their tax savings value. Credits, a form of cash, can be consolidated between consumers. These credits are, in effect, the dividend payment of the prior period.

The approach could be further enhanced by informing consumers of the carbon emissions embedded in broader categories of purchases, as consumers appear to be unaware and respond to this information when provided (Muller et al., 2019; Lohmann et al., 2022). If all purchases accounted for embedded emissions, the tax and environmental costs would be more salient and it would be possible to use carbon credits in a wider range of consumer settings.

Finally, Penalty & Reward can be used as a more effective way of communicating a Fee & Dividend policy. Often policies, including tax policies, are communicated in a way that does not convey literally the somewhat more opaque and complex implementation mechanism (Abeler and Jäger, 2015). Communicating Penalty & Reward, while implementing Fee & Dividend, may not be considered misleading, given the fact that they are isometric and a short experience window could clarify this fact for individuals who doubt it.

Endogenous green preferences

N. R. Vora[†] Guglielmo Zappalà[‡]

Abstract

Low public support often impedes stronger environmental policies, but if they are enacted, do policy preferences change? Using surveys covering 38 countries around the world, we study the dynamics of environmental policies and individual preferences over time. Exploiting within-country, across birth-cohort variation, we document that cohorts exposed to more stringent policies in the past are more supportive of environmental policies at the time of the survey, with the effect largely driven by exposure during early adulthood. This relationship suggests that a society's environmental policy attitudes evolve endogenously, with implications for normative frameworks used in welfare economics.

Keywords: Endogenous preferences; Environmental policy; Environmental preferences; Formative age; Policy support

JEL Classification: D83; H23; H31; Q58

[†]UC Santa Barbara, Department of Economics, CA 93106.

[‡]UC Santa Barbara, emLab and Bren School of Environmental Science & Management, CA 93117.

1 Introduction

Environmental externalities can be effectively addressed through sufficiently stringent policies. However, stringent environmental policies, like carbon taxes, have been obstructed by low public support (Douenne & Fabre, 2022). While public support naturally influences policy outcomes, it is traditionally modelled as fixed, invariant to influence from experience with policies (Stigler & Becker, 1977). But if relationships are bi-directional, i.e., public preferences influence subsequent policy outcomes and policy outcomes influence public preferences, then the resulting endogenous relationship affects the evolution of policy.[1] Understanding how green policy preferences evolve is essential, particularly in light of the increasing urgency to address environmental challenges (Dechezleprêtre et al., 2022). Recent theoretical models allow for individuals' consumption preferences to evolve endogenously, enriching the analysis on the dynamics of various environmental behaviors (Besley & Persson, 2023; Konc et al., 2021; Mattauch et al., 2022). Empirically, however, it remains an open question whether individuals' preferences for green policies evolve differently in response to exposure to stringent policies compared to more lenient ones.

If voter preferences are exogenous and information perfect, personal experience with policies would not shape an individual's views. But, alternatively, voters could update

[1]Previous studies document that public support influences policy outcomes (List & Sturm, 2006). Theoretical and empirical analyses model how bidirectional causation between preferences and policy outcomes complicates the evolution of policies (Alesina & Giuliano, 2015; Alesina & Rodrik, 1994; Besley & Persson, 2019; Gerber & Jackson, 1993). Fuchs-Schündeln and Schündeln (2015) document that support for democracy increases with time under the system, resulting in endogenous political preferences. Acemoglu et al. (2024) find a similar result with the effect stronger in democracies that provide economic growth, peace and political stability, and public goods. Aksoy et al. (2020) find that legal same-sex relationship recognition policies improve attitudes toward sexual minorities in Europe.

their views of the normative appropriateness of policies, once exposed to them. Using global survey data from the Integrated Value Surveys (IVS) and an internationally comparable measure of environmental policy stringency, we test this hypothesis in 38 countries over 20 years. We exploit within-country, cross-cohort variation in individual exposure to environmental policy stringency to determine whether birth cohorts exposed to more stringent environmental policies are more supportive of environmental policies at the time of the survey. We account for potential confounds, such as exposure to environmental and economic conditions, country-specific shocks at the time of the survey, generational variations, and country-specific age-specific factors.

We find that individual experience with more stringent environmental policies strengthens subsequent preference for them, implying endogeneity of preferences. As experience with policy reshapes norms, even societies with tepid initial support for stringent policy find themselves increasingly supportive over time. Increasing exposure to environmental policy stringency by 1 standard deviation increases individual support for government intervention to reduce pollution by 0.18 of a standard deviation. The effect is evident even on specific policy instruments. Tax stringency exposure, and in particular CO_2 and fuel tax, has a strongly positive effect on subsequent preferences for environmental taxes.

All past exposure to policy is not equal in its effect.[2] We find strong evidence that policy exposure during the formative age window of 18-25 has a large impact, in line with a growing literature.[3] Policy stringency in any other age windows does not meaningfully

[2]Studies document the high salience of environmental policies, particularly those regulating air pollution (Eshbaugh-Soha, 2006). Policies are shown to be salient in the business community (Noailly et al., 2021) and among consumers (Rivers & Schaufele, 2015).

[3]The "formative age" or "impressionable years" hypothesis, first studied in Krosnick and Alwin (1989), states that values are formed during early adulthood and do not substantially change afterward. This

alter individual preferences at the time of the interview. When we compare formative age policy exposure to contemporaneous exposure at the time of the interview, we document an asymmetry of reference points (De Neve et al., 2018). Demand for government action is most pronounced among individuals who experienced a reduction in stringency since their formative age window. Finally, we document a larger effect of policy exposure conditional on poor environmental quality (as measured by fine particulate matter and household air pollution).

Armed with the estimated effects of policy stringency exposure on environmental preferences, we produce a simple counterfactual exercise. We select three countries with historically different environmental policy levels (Brazil, US, Sweden) and produce hypothetical global preferences if birth-cohorts had been exposed to each of these policy counterfactuals. We document stark differences in public support under each counterfactual, ranging from 85% of the countries in the sample with lower green preferences than historically observed, had individuals been exposed to Brazil's history of policy stringency, to 77% of countries with higher preferences under Sweden's policy counterfactual.

Our results suggest that strong opposition to certain environmental policies may not reflect lasting preferences. Rather, how preferences evolve will depend in part on which policy outcomes are realized.[4] Though introducing more stringent policies may be seen as unacceptable, a counterfactual exercise in which stringent policy were reversed may

hypothesis has recently been tested on a variety of settings in economics (Carreri & Teso, 2023; Cotofan et al., 2024; Eichengreen et al., 2021; Eichengreen et al., 2024).

[4]Low support has been directly expressed through referenda in Washington State (Anderson et al., 2023), Switzerland (Bornstein & Lanz, 2008; Carattini et al., 2017), and California (Burkhardt & Chan, 2017; Holian & Kahn, 2015; Kahn & Matsusaka, 1997), and social protests in France (Douenne & Fabre, 2020, 2022).

also be unacceptable. If a society's norms can change, we would expect differences in support between an ex-ante proposal and an ex-post review of a policy.[5]

Rising interest in the co-dynamics of climate policy stringency and public support has spurred recent work on modeling endogenous preferences over consumption choices in the economics of climate change (Konc et al., 2021; Mattauch et al., 2022). Country-specific evidence suggests endogenously evolving support for green policies. In British Columbia, initially tepid support for a carbon tax policy that became law grew substantially in the subsequent years (Murray & Rivers, 2015). Vice versa, in France, after a carbon tax failed amid widespread social protest, the popularity of such policies appears to have declined further.[6] But these individual cases make it difficult to draw a comprehensive picture. To the best of our knowledge, our paper is the first cross-national attempt to study endogenous environmental preferences over policy options and estimate the causal effect of experience with policy in a quasi-experimental setting, accounting for a range of potentially important co-factors, including environmental quality levels. Endogenous support alters the predictability of the policy path over time, with implications for notions of paternalism and the moral and political economy of policy selection that we discuss in the conclusion.

[5]Experience could foster comfort with policy settings that are familiar (Furnham & Boo, 2011; Tversky & Kahneman, 1974), and rules and other instruments can influence cultural norms for which individuals assign intrinsic value (Bezin, 2015; Bisin & Verdier, 2001, 2011; Schumacher, 2015).

[6]In Douenne and Fabre (2020), support in France for a carbon tax and rebate was 38% in early 2019. In Dechezleprêtre et al. (2022), 29% of French respondents supported the policy in May 2021.

2 Theoretical Motivation

Why might we expect policy and preferences to evolve endogenously? Policy exposure could change norms and the information set used by individuals. We provide theoretical motivation for our empirical work by exploring each of these mechanisms.

2.1 Norms

Traditionally, individuals' preferences over environmental policy can be represented by a conventional utility function, such that $U(\psi) = \mu(c(\psi), q(\psi))$, where c and q represent consumption and environmental quality, each a function of policy stringency $\psi \in \Psi$ (i.e., the price of polluting via a Pigouvian tax or a shadow price).

In our framework, individuals can form policy norms based on past experience, defined as ψ_R. Deviation of the present policy ψ from the norm ψ_R can affect utility $V[\cdot]$, represented as follows:

$$
\begin{aligned}
V[\psi \mid \psi_R, \alpha, \beta, \gamma] \;=\; & \mu(c(\psi), q(\psi)) && (1)\\
& - \alpha \cdot d_1(c(\psi), c(\psi_R)) && (2)\\
& - \beta \cdot d_2(q(\psi), q(\psi_R)) && (3)\\
& - \gamma \cdot d_3(\psi, \psi_R) && (4)
\end{aligned}
$$

The first term of the utility function conveys a trade-off between utility from traditional consumption and environmental quality; lines (2)-(4) represent the consequences

to utility of deviating from consumption, environmental quality and policy norms (where $\alpha, \beta, \gamma \geq 0$, and d_1, d_2, and d_3 are distance functions). An individual chooses their subjective policy stringency such that

$$\psi^* = \underset{\psi \in \Psi}{\operatorname{argmax}} \, V\left[\psi \mid \psi_R, \alpha, \beta, \gamma\right] \tag{5}$$

Changes in policy norms do not affect the traditional utility element (line 1), but exert an influence on the optimal stringency level through lines (2)-(4). When at least one of $\alpha, \beta, \gamma > 0$, the derivative of ψ^* with respect to the policy norm ψ_R is unambiguously positive under a set of reasonable conditions.[7]:

$$\frac{\partial \psi^*}{\partial \psi_R} > 0 \tag{6}$$

When the current policy level $\psi_t < \psi^*$, an individual would have stronger environmental policy preferences, expressed by more support for government involvement in reducing pollution or increases in environmental taxes.

2.2 Learning

Environmental policy exposure may also affect individual preferences through endogenous learning. Consider a policy level ψ. The value of a policy setting $V(\psi)$ represents the net benefit (or cost). For simplicity, let V(ψ) be a quadratic function, e.g., $-(\psi - b)^2$,

[7]For continuously differentiable distance functions, such as $(\psi - \psi_R)^2$, the implicit function theorem is sufficient to show Result (6) for any interior optimum ψ^*. For a much broader set of distance functions, Result (6) can be shown for any function that satisfies the single crossing condition with respect to ψ_R (Milgrom & Shannon, 1994).

with a unique interior optimum, $\psi^* = b$. A person choosing whether to support more or less government involvement in reducing pollution must consider whether ψ_t is above or below ψ^* without full knowledge of the function $V(\psi)$.[8] Determining this is similar to the econometrician's task of estimating the relationship between ψ and $V(\psi)$ using only sampled observations of each. The sample of ψ is determined by knowledge from policy levels in the past, and $V(\psi)$ is observed from the environmental and economic conditions that ensue. The larger and more varied the sample in ψ, the easier to form an accurate sense of whether it is better to support or oppose a stringency proposal.

We can expect positively endogenous green preferences if the public believes in the absence of past evidence that $\psi_t \geq \psi^*$, when in fact $\psi_t < \psi^*$.

Our main findings could be a result of the norms and the learning channel. We explore the relative importance of these channels empirically in Section 5.3.4.

3 Data

We use individual survey data from the Integrated Value Surveys (IVS), which harmonize the European Values Study (EVS) and the World Value Survey (WVS). These large-scale cross-national surveys gather data on socio-political, environmental attitudes, and other characteristics of individuals (Aghion et al., 2023).

We focus on two measures of environmental preferences: (i) "Government should reduce environmental pollution",[9] (ii) "Increase in taxes if used to prevent environmental

[8] To put this another way, a person must determine $V'(\psi_t)$.

[9] Respondents can interpret the question as asking the appropriate role of government and as support for immediate actions to reduce pollution. Each is relevant to the question of endogeneity of preferences and the implication of such endogeneity on how policy evolves. We show that the survey question is

pollution". For both questions, answers can be 1 ("strongly disagree"), 2 ("disagree"), 3 ("agree") and 4 ("strongly agree"). To exploit all the variation in survey responses, we use the 4-point Likert scale as the main outcome (see Online Appendix A.1 for details).

Our measure of environmental policy exposure comes from the OECD Environmental Policy Stringency Index (EPS), a country-specific and internationally-comparable measure. Stringency is defined as the degree to which environmental policies put an explicit or implicit price on polluting or environmentally harmful behavior (Kruse et al., 2022). The index is constructed by scoring policy stringency on a scale from zero to six and subsequently aggregating the scores into an index, where higher levels are associated with more stringent policies.

Our main variable $policy\ exposure_{bct}$ for an individual born in year b in country c interviewed in year t is defined as:

$$policy\ exposure_{bct} = \frac{1}{t-(b+18)} \sum_{\tau=b+18}^{t} policy\ stringency_{c,\tau} \tag{7}$$

An individual's $policy\ exposure_{bct}$ is the average environmental policy stringency in their country between age 18 and the year of interview. Environmental policy preferences are recorded discontinuously from 1990 to 2010. Respondents interviewed in the same year and country can have different treatment exposure because of variation in their year of birth. To test for the "formative age" hypothesis, we construct the average policy exposure when the respondent was aged 18 to 25 and other eight-year age window exposures starting from the range of the impressionable years. Since certain countries

meaningful and predictive of future policy outcomes in Online Appendix B.

exhibit a positive trend over time, we also construct two alternative measures. First, we de-mean each country's level of environmental policy from the annual cross-sectional average. This procedure accounts for policies implemented at supra-national levels which would contemporaneously make more stringent several countries (e.g., the EU Emissions Trading System). Second, we de-trend the policy stringency from a country-specific linear trend (see Online Appendix A.2 for details).

We also use two indicators from the Environmental Performance Index (Wolf et al., 2022) to measure environmental quality. We consider PM2.5 and Household Air Pollution (HAP) measured as the number of age-standardized disability-adjusted life-years lost per 100,000 people, respectively due to exposure to fine air particulate matter smaller than 2.5 micrometers, and to household air pollution from the use of household solid fuels. We recode the measures such that higher values mean better environmental conditions and construct environmental quality exposure similar to policy exposure (see Online Appendix A.3 for details).

4 Empirical Approach

To test whether experience with policies feeds back into preferences for policies, we adopt an empirical approach that exploits within-country, across-birth-cohort variation in policy stringency, and removes confounds such as local economic, political, and environmental conditions at the time of interview. Our baseline specification is

$$Y_{ibctw} = \beta_1 policy\ exposure_{bct} + X_i'\boldsymbol{\gamma} + Z_{bct}'\boldsymbol{\delta} + \kappa_b + \mu_{ct} + \alpha_w + \theta_c \times age + \varepsilon_{ibctw} \quad (8)$$

where Y_{ibctw} is the answer to one survey question by individual i, born in year b interviewed in year t in country c in survey source w. Our main coefficient of interest is associated with exposure to environmental policy stringency, *policy exposure*, which varies across countries, years of birth, and years of interview. We control for a set of individual covariates X_i', capturing socio-economic characteristics (gender; employment status; education; ten country-specific income decile dummies) in year of interview t. Most importantly, other past experiences could be correlated with environmental policy stringency and influence environmental preferences. To capture these exposures, we control for a vector of country-cohort specific covariates Z_{bct}', including environmental quality (PM2.5 and HAP exposure) and economic recessions (average number of years in which GDP growth contracted by at least 10%).

Our specification accounts for birth-cohort κ_b, country-year of interview μ_{ct}, and survey (EVS or WVS) α_w fixed effects, as well as country-specific age trends $\theta_c \times age$. All regressions are estimated using OLS for ease of interpretation, but similar results are obtained with ordered probit and probit models. We cluster standard errors at the country-year-of-interview level. We use survey sample weights to make the data representative at the country level and limit our sample to individuals born in the same country in which they are interviewed.

Identification Strategy. There are several potential threats to the identification of a causal effect of policy exposure on environmental preferences, which would imply endogeneity of preferences. Spurious correlations may arise due to reverse causality (i.e., countries have more stringent policies because citizens have strong environmental values), or unobserved confounders, such as historic events or economic conditions, that could co-determine individuals' preferences and policy levels in place.

We exploit within-country variation at the birth-cohort level in exposure to environmental policy stringency to establish a plausibly causal impact of experience with environmental policies on green preferences. Below, we detail how our specification addresses a number of potential threats to identify such an effect.

First, age-specific factors could matter if different birth cohorts are exposed to different policies with different probabilities. The global positive trend in environmental policy stringency may suggest that younger generations are more likely to experience more stringent policies. We control for birth-year fixed effects, which account for cohort-specific attitudes so as to compare individuals within the same birth cohort.

Second, contemporaneous levels of environmental policy and quality and any other national and global economic and political conditions may drive differences in preferences. We account for any contemporaneous country-specific characteristics with country-year-of-interview fixed effects. This approach mitigates concerns that the results are driven by other structural time-varying differences between countries and strengthens the assertion that observed differences in attitudes towards environmental policies constitute a change in intrinsic preferences due to differences in the stringency of environmental policy exposure.

Third, there could be heterogeneous generational trends in environmental preferences across countries. Countries could lie on differential trends in the evolution of individual values which can lead to larger differences across generations. To rule out such a possibility, we include country-specific age trends in our specification.

Finally, we also account for the source of the survey fixed effects (World Value Survey or the European Value Study), to account for different sampling methodologies and other differences across the two survey sources.

The set of fixed effects ensures that the identifying variation comes from changes in exposure to environmental policy stringency across birth cohorts within a country interviewed in a given year. Although we saturate our specifications with fixed effects, there could remain confounding past exposures correlated with environmental policies, which we address by controlling for economic conditions and environmental quality.

To emphasize the role of the fixed effects, Online Appendix Figure C1 shows the unconditional correlation between survey answers and environmental policy stringency exposure, negative and statistically significant at the 95% level (Panel a), and the strongly positive and statistically significant relationship between the residual variation in preferences and policy exposure (Panel b). Our set of fixed effects accounts for potential confounders that change the direction of the association between preferences and policies. For instance, consider that a relatively stringent country is likely to have been relatively stringent in the past and have lower pollution levels at time of interview. Low pollution levels should predict weaker support for government action than would be expected in countries with high pollution. Including country-year-of-interview fixed effects allows us to isolate the role of variation in past policy exposure, while removing the confounding

influence of variation in contemporaneous environmental quality, policy stringency, and economic conditions.

5 Results

5.1 Adulthood exposure

Table 1 reports the estimates from our baseline specification. The coefficient on environmental policy stringency exposure is positive and statistically significant, in line with the theoretical prediction. In our preferred and most conservative specification that accounts for economic recessions and environmental quality (column 5), increasing EPS exposure by 1 standard deviation (SD) increases individual support for government to reduce pollution by 0.18 SD. In other words, an increase in policy stringency from the US level (2.91) to Sweden's level (3.61) in 2019 corresponds to a 5.7% increase in the Likert support for government to reduce pollution. Our results provide suggestive evidence of endogenous green policy preferences, highlighting a bidirectional relationship between policy stringency and preferences.[10]

Types of environmental policies. The EPS is a composite index of different environmental policies. Different policy instruments can have different salience. Some policies may attract more political and media attention, others may be more noticeable in perceptions or experiences with cost burden or incidence (Huse & Koptyug, 2022; Rivers & Schaufele, 2015). Using various sub-indices of the EPS index (market-based instruments,

[10]As noted in Section 1, previous work documents the relatively uncontroversial of the two directional relationships: the effect of preferences on policy outcomes. We also show empirical evidence of past preferences predicting future policy outcomes in Online Appendix B.

non-market-based instruments, environmental taxes, and carbon trading schemes), we find no conclusive evidence that the effect of policy stringency on support for government to reduce pollution is solely driven by a specific sub-category of policy instruments (Appendix Figure C2).

Nevertheless, we document heterogeneous effects of environmental policies on preferences for a specific policy instrument, environmental taxes. Figure 1 displays the estimated coefficients on various dimensions of past environmental policy exposure. More stringent exposure to taxes increases support for an increase in tax to prevent pollution. A 1 SD increase in environmental tax stringency exposure is associated with a 0.12 SD increase in the support for a green tax increase. Using the same real-world comparison as before, an increase from the US environmental tax stringency level (0.25) to Sweden's level (3.75) corresponds to a 18% increase in the Likert support for increase in tax to prevent pollution. Importantly, past exposure to non-market green policies has a negative and statistically insignificant relationship with green tax preference.

Results have more pronounced statistical significance and larger magnitude for the stringency of combined carbon and diesel taxation. A 1 SD increase in exposure to these policy instruments increases subsequent tax support by around 0.17 SD. By disentangling CO2 and diesel tax exposure, the effect is mostly driven by carbon tax.

Environmental tax policies are one of the most salient policy instruments (Dechezleprêtre et al., 2022; Douenne & Fabre, 2022). We document that past exposure to them affects preference formation. Thus, specific policy types appear to reflect the dynamics of positive feedback that we document at the composite EPS level. Well-documented initial skepticism about taxes (Anderson et al., 2023; Douenne & Fabre, 2020) could make

tax policies especially well-suited to the positive endogeneity that arises from experience. Exposure to specific policy types (as in, policy design or mechanism) might mediate the effect of experience on preferences.[11]

Robustness. We conduct several tests to probe the robustness of our results, reported in Online Appendix C.2. We find coefficients quantitatively very similar to our baseline with alternative measures of policy stringency de-meaned from the annual cross-sectional average and de-trended from a country-specific linear trend. We examine the influence of starting age, and find that the effect becomes more robust from age 16 across all policy exposure measures. This result supports our hypothesized channel of norm formation, which requires awareness of and influence by policy stringency, and becomes likelier approaching adulthood.[12] Results are also robust to alternative fixed effects and alternative estimation methods, using ordered probit for categorical, and probit and linear probability models for binary outcomes. In the latter exercise, increasing stringency exposure from the US level to Sweden's level in 2019 increases the probability by 5 percentage points (7.5% at the mean) of expressing that government should reduce pollution and of strongly agreeing by 9.4 p.p. (31%).

Falsification test. To mitigate concerns that the effect of environmental policy exposure on environmental preferences conflates the consequences of general social and political changes, we consider attitudinal survey questions unrelated to the environment

[11]Unfortunately, the survey does not ask other questions about support for additional pollution policy types, so we cannot test, e.g., the effect of non-market policy exposure on support for them. Given the limited salience and higher costs of non-market policies, dynamics could be different.

[12]Differently than exposure to democratic institutions which are assumed to be understood by individuals as of age of six (Acemoglu et al., 2024), environmental policies and their effects are posited to be internalized from adult life (Aklin et al., 2013).

on family relationships, societal well-being, and economic values. Using 13 alternative outcomes, we find no effect that is statistically distinguishable from zero, except for one at the 5% level, consistent with sampling variation attributable to multiple tests (see Online Appendix C.3).

5.2 Age windows

Individual preferences are particularly malleable during certain years of life (Severen & Van Benthem, 2022). An expanding literature documents how values, norms, and preferences can be shaped by exposure to conditions and events, most notably in early adulthood. The formative age (or impressionable years) hypothesis states that individual values are formed during a period of great mental plasticity in early adulthood between 18 and 25 years and remain mostly unaltered afterward (Krosnick & Alwin, 1989).[13]

We construct measures of policy stringency exposure in various eight-year age windows, starting from the 18-25 window. Figure 2 reports the coefficients on each eight-year age window, with the formative age window's coefficient in red. This coefficient is the only statistically significant estimate and is comparable in magnitude to the effect of exposure from age 18 to the year of interview, suggesting no evidence of the effect getting smaller as an individual acquires more information throughout their life. A 1 SD increase in formative age exposure to environmental policy stringency is associated with 0.16 SD

[13]Other relevant work in social psychology includes Cutler (1974) and Torney and Hess (1967). Empirical evidence of the impressionable years hypothesis has been documented for attitudes towards migrants (Cotofan et al., 2024), attitudes towards democracy (Magistretti & Tabellini, 2022), confidence in political institutions and leaders (Eichengreen et al., 2024), job preferences (Cotofan et al., 2023), political preferences (Barone et al., 2022), preferences for redistribution (Carreri & Teso, 2023; Roth & Wohlfart, 2018), and trust in science (Eichengreen et al., 2021). The only paper about environmental preferences considers formative age exposure to natural disasters (Falco & Corbi, 2023).

increase in support for government intervention to reduce pollution.

Robustness. In Online Appendix D.1, we probe the robustness of our results to a variety of checks, including binary versions of the outcome, alternative definitions of formative age, de-meaned and detrended exposure measures, and measuring exposure over four-year age windows. A final potential concern is that societal preferences during formative age window confound the role of stringency exposure. To allay this concern, we show that controlling for national average environmental policy preferences during the respondent's formative age does not significantly alter the effect of policy exposure (see Online Appendix D.2). By doing so, we rule out an important alternative explanation: that preferences are shaped not by policy stringency during formative age, but by exposure to societal beliefs that themselves correlated with policy stringency during the window. This bolsters the causal interpretation of stringency exposure.

5.3 Mechanisms

5.3.1 Environmental policy direction

The importance of the impressionable years indicates that distant experiences play a role in shaping preferences today. Nevertheless, policies can vary widely over time, and a natural question arises about whether changes to higher or lower stringency than formative age levels lead to different effects on present-day preferences. Without specifically testing for this heterogeneity, one might incorrectly assume that past exposure shapes preferences symmetrically (De Neve et al., 2018).

To test for asymmetric effects, we compare formative age exposure with the contem-

poraneous level of policy stringency in a binary variable (Δ policy), equal to one if the contemporaneous policy is less stringent than the formative age level (i.e., "Negative"). Figure 3a displays the marginal effects of formative-age policy exposure and shows that this asymmetry exists. The effect of formative-age exposure is more pronounced among individuals interviewed when the policy environment is more lax. For two individuals in the same birth-cohort but interviewed in different years, experiencing a more lenient policy in the year of interview drives a stronger reaction to an increase in stringency exposure in terms of a support for government action against pollution.

We compare exposure during any age window to the current level of policy stringency and fail to recover a similar effect using exposure in any other age window than the formative age (Online Appendix E.1). This result suggests that this specific period has a prominent role as a reference point against which individuals evaluate the current state (Abel, 1990; Coppock & Green, 2016; Roth & Wohlfart, 2018).

5.3.2 Environmental quality role

We also explore whether the effect of more stringent environmental policies on individual support is conditional on environmental quality. We conjecture that more stringent policies where environmental quality is low increase the policies' salience and perceived value to individuals. This, in turn, increases lasting support for government action to reduce pollution.

We construct an indicator if PM2.5 or HAP exposure is below the sample median (i.e., worse environmental quality conditions). Figure 3b shows the heterogeneous marginal effects of EPS exposure by environmental quality. The results confirm that exposure to

stringent policy in predicting support for government pollution reduction is conditional on poor environmental quality. The uninteracted term of exposure to environmental quality below the median is negative and in certain cases not significant. The interacted term identifies the effect of experience of both poor conditions and more stringent policies meant to address them. The significance of this combination shows that stringent policies targeting poor environmental conditions are likelier to build lasting public support for government pollution actions (see Online Appendix E.2 for additional details).

5.3.3 Heterogeneity by socio-demographics

We summarize heterogeneous effects of policy exposure by respondents' characteristics (reported in Online Appendix C.4). There is no substantial heterogeneity across most of the characteristics (gender, employment status, income, political orientation), except for education: lower-educated individuals are more likely than higher-educated ones to have green preferences as a result of more stringent environmental policies. We also document that the effect of policy exposure is driven by individuals less interested in politics and with less confidence in the government. This may be because individuals with limited political knowledge may be more influenced by their formative experiences than those who regularly update their views. This adds evidence that policy outcomes are critical to convey societal norms. Finally, individuals who think the government should take more responsibility are more prone to support government action to reduce pollution and their support for such actions is more responsive to past policy exposure.

5.3.4 Norms and Endogenous Learning

We identify two mechanisms for endogenous green preferences, norms and endogenous learning, and ultimately find stronger evidence for norms. The learning channel would depend on somewhat strong assumptions, as shown in Section 2.2 (additional details in Online Appendix E.3). In essence, a required assumption is that age cohorts behave as if using different information, because directly lived experience of policy impacts is weighted more than indirectly obtained information. To produce our baseline results of positive endogeneity, an additional requirement for the learning channel is that individuals are pessimistic about the net benefits of more stringent policies, and that this systematic pessimism is absent or reduced in cohorts with direct, lived experience of policy stringency in the past, and particularly in early adulthood.

Were learning an important channel to correct otherwise pessimistic beliefs, a higher degree of policy experimentation would be predictive of support.[14] In a horse race between the first and the second moment of the policy exposure distribution, coefficients on the standard deviation, as well as its interaction with the average, suggest a modest impact of learning that, even when controlled for, leaves the relatively large effect of average exposure unaltered and suggests a robust role for norms.

[14] A literature on learning via policy experimentation discusses the importance of varying policy in order to provide regulators and voters with necessary data to assess causal impacts (Aghion et al., 1991; Warren & Wilkening, 2012; Wieland, 2000; Zhao & Kling, 2003).

6 Counterfactual policy stringency

We use our reduced-form estimates to assess the importance of the uncovered mechanism on the endogeneity of preferences to policy exposure in the environmental political process. We answer the following question: if individuals were exposed to another counterfactual policy level, all else equal, how would the change in policy preferences predicted by our model be? We choose as counterfactuals three countries with notable historically different policy stringency levels: Brazil (cross-cohort average policy stringency exposure is 0.46), United States (average policy exposure 1.75, close to the cross-country average 1.74), and Sweden (average policy exposure 2.57).

We predict the counterfactual change in green preferences using the difference between the observed level of policy stringency and the counterfactual policy level in each of the three countries, and the associated coefficient on EPS (see Online Appendix F for details). Figure 4 shows the model-predicted shares of respondents with *strong* green preferences under each counterfactual. If individuals had past environmental policy exposure as stringent as the US, the average support for green policies would be similar across countries (28.18 observed in the data vis-à-vis 28.11 under the US counterfactual). We observe stark differences in public support if birth cohorts were exposed to a stringency level equal to Brazil's (mean is 14.2) and equal to Sweden's (mean is 38.44). Nine countries under the Sweden counterfactual and four countries under the US counterfactual have a share of public support strictly above 50% (in red). If we simplistically assume a one-to-one mapping between stated green preferences and revealed voting preferences in a majority voting rule system, these cases indicate that countries, upon passing more

stringent environmental policies, might set in motion a dynamic loop where stringent policies feed back into higher future demand for policies.

7 Conclusion

We document that support for environmental policies increases significantly when individuals have been exposed to more stringent environmental policies during their adulthood and especially during their formative age window. Using data across more than thirty-five countries, we compare individuals in the same birth cohort within the same country across different points in time. We also provide evidence that higher exposure to environmental taxes increases subsequent preferences for environmental taxes, and not for other environmental policy instruments. Similarly, stringent environmental policies have no impact on non-environmental preferences. We document that our effect is stronger among cohorts exposed to a more lenient environmental policy mix in the year of the interview relative to their formative age exposure, as well as among cohorts exposed to lower environmental quality.

This paper reveals that the relationship between environmental policy and public support is bidirectional, and thus endogenous. From a public policy perspective, these results indicate that although implementing major environmental policies might be a politically difficult task, individuals will increase their support for such policies over time. The complex co-dynamics of environmental preferences and policies can make equilibrium policy stringency levels more difficult to predict. The path to an equilibrium policy level, if one exists, would likely follow a different and longer path, with important

implications for societal welfare.

Endogenous environmental policy preferences have implications for moral and political economy, as well. Consider a social planner who determines that an unpopular, misunderstood policy will be broadly appreciated in the fullness of time. Imposing policy, even if deemed welfare-improving, in a manner that contradicts the preferences of individuals is considered paternalistic and can draw objections on ethical grounds. Our evidence that appreciation for the policy will change over time has the potential to weaken the premises underlying such objections. Inertia of policy to appropriately respond to new scientific information or material circumstances is not a reflection of deeply rooted, permanent preferences but, at least in part, a result of a dynamic policy familiarity from past experience.

Regulators and politicians who are accountable to the public over medium or long terms may opt to set policy nearer to the current optimum instead of taking a more gradual approach commonly assumed to be better tolerated by opponents to such measures. Moving faster could be more effective in eventually generating the endogenous support that we document.

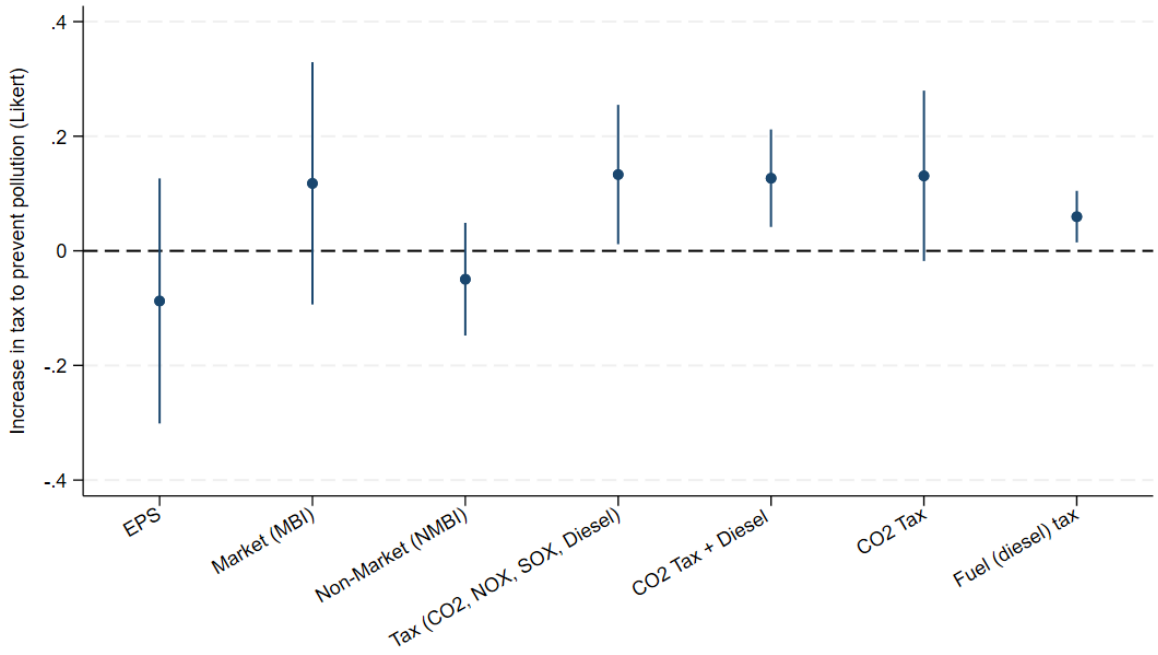
There are important avenues for future research within this area. Given the urgency of climate change, it would be valuable to study climate policy preferences, in isolation from other environmental policy, as data become available. This approach would exploit longitudinal data of public support for carbon taxes and other instruments (e.g., cap-and-trade, subsidies) to evaluate how policy preferences change after an increase or decrease in the stringency of a climate policy.

Table 1: Environmental Policy Stringency (EPS) exposure during adult years

	Government should reduce pollution (Likert)				
	(1)	(2)	(3)	(4)	(5)
Environmental Policy Stringency (EPS) exposure	0.228**	0.225**	0.219**	0.228**	0.236**
	(0.109)	(0.110)	(0.107)	(0.109)	(0.111)
Recession exposure		-0.515	-0.528	-0.534	-0.518
		(0.381)	(0.406)	(0.378)	(0.392)
PM2.5 exposure			0.0144		0.217
			(0.281)		(0.291)
HAP exposure				-0.305	-0.377
				(0.298)	(0.307)
Individual controls	✓	✓	✓	✓	✓
Survey FE	✓	✓	✓	✓	✓
Year of birth FE	✓	✓	✓	✓	✓
Country × Year of interview FE	✓	✓	✓	✓	✓
Country-age linear trends	✓	✓	✓	✓	✓
Mean outcome	2.909	2.909	2.909	2.909	2.909
SD Outcome	0.896	0.896	0.896	0.896	0.896
Mean EPS exposure	0.9	0.977	0.977	0.977	0.977
SD EPS exposure	0.696	0.696	0.696	0.696	0.696
N	16889	16889	16889	16889	16889
adj. R^2	0.1324	0.1324	0.1323	0.1324	0.1323

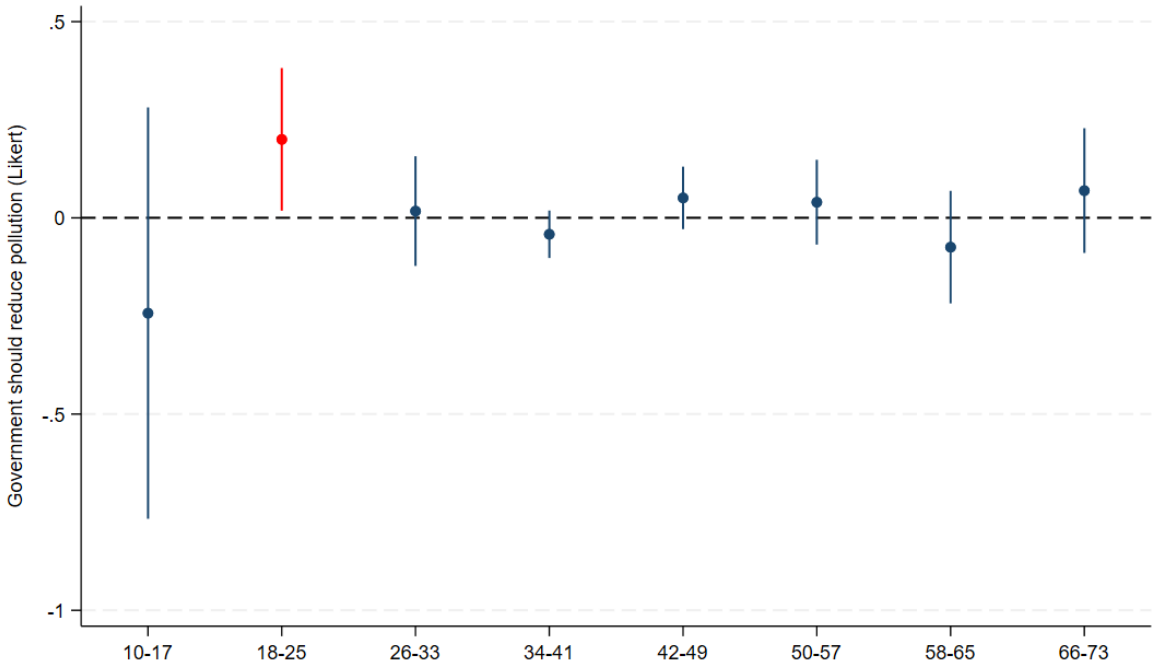
Notes: The outcome variable ranges from 1 to 4 (with higher values reflecting stronger agreement with the statement "Government should reduce environmental pollution"). All regressions control for male dummy, unemployment dummy, 3-category education (Lower, i.e., primary education, Middle, i.e., secondary education, Upper, i.e., tertiary education), 10-class subjective income decile scale. Recession exposure is a dummy variable that takes value one if during her adult years the individual has experienced at least one year in which the real GDP per capita growth rate of its country is at least 10% below the previous year's growth rate. PM2.5 and HAP exposure are respectively the adult years average of age-standardized disability-adjusted life-years lost per 100,000 people due to exposure to fine air particulate matter smaller than 2.5 micrometers (PM2.5) and due to exposure to household air pollution (HAP) from the use of household solid fuels (PM2.5 and HAP exposure variables are inverted such that higher values imply higher quality and divided by 1000 to improve readability of coefficients). All regressions also include year of birth, country-by-year of interview, survey fixed effects, and country-by-age linear trends. Observations are weighted using survey weights. Standard errors are clustered by country-year of interview. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Appendix Table C1 reports the full estimates including controls.

Figure 1: Effect of EPS sub-indices exposure in adulthood on tax increase preferences



Notes: The figure shows the coefficient associated with the average level of sub-index in the x-axis during the adulthood of an individual, using as an outcome the question "Increase in tax to prevent pollution" in the Likert scale. All regressions control for male dummy, unemployment dummy, 3-category education (Lower, i.e., primary education, Middle, i.e., secondary education, Upper, i.e., tertiary education), 10-class subjective income decile scale, recession exposure, PM2.5 and HAP exposure. The regression also includes year of birth, country-by-year of interview, survey fixed effects, and country-by-age linear trends. Observations are weighted using survey weights. Standard errors are clustered at the country-year of interview level. Bins represent the 95% confidence intervals around point estimates. Tabular results are reported in Appendix Table C9.

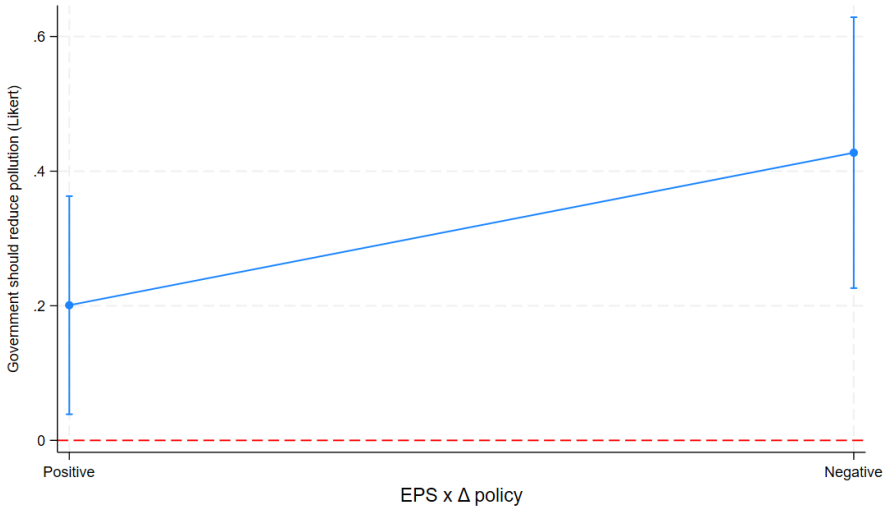
Figure 2: Effect of EPS exposure on environmental preferences by age window



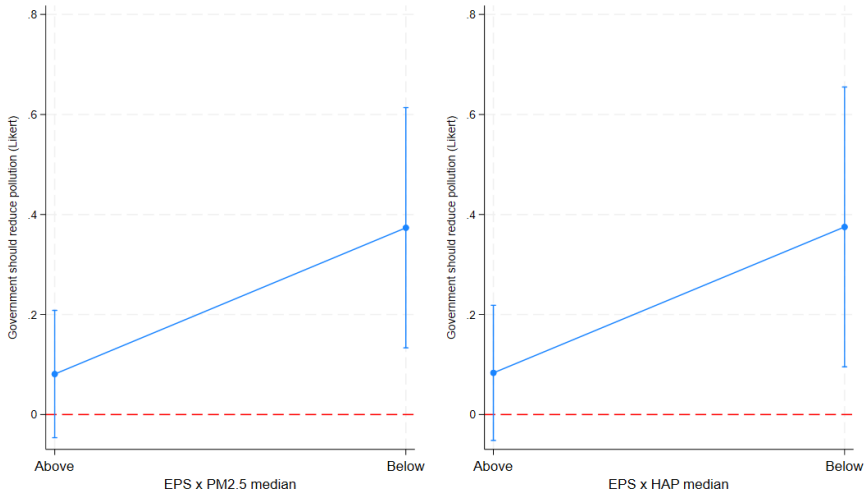
Notes: The figure plots the coefficients associated with the average level of EPS during each eight-year age window reported in the x-axis using the question "Government should reduce pollution" as outcome in Likert scale. The regression controls for male dummy, unemployment dummy, 3-category education (Lower, i.e., primary education, Middle, i.e., secondary education, Upper, i.e., tertiary education), 10-class subjective income decile scale, recession exposure, PM2.5 and HAP exposure. The regression also includes year of birth, country-by-year of interview, survey fixed effects, and country-by-age linear trends. Observations are weighted using survey weights. Standard errors are clustered at the country-year of interview level. Bins represent the 95% confidence intervals around point estimates. Tabular results are reported in Appendix Table D1.

Figure 3: Heterogeneous effects of policy exposure on environmental preferences

(a) Deviations from contemporaneous policy level

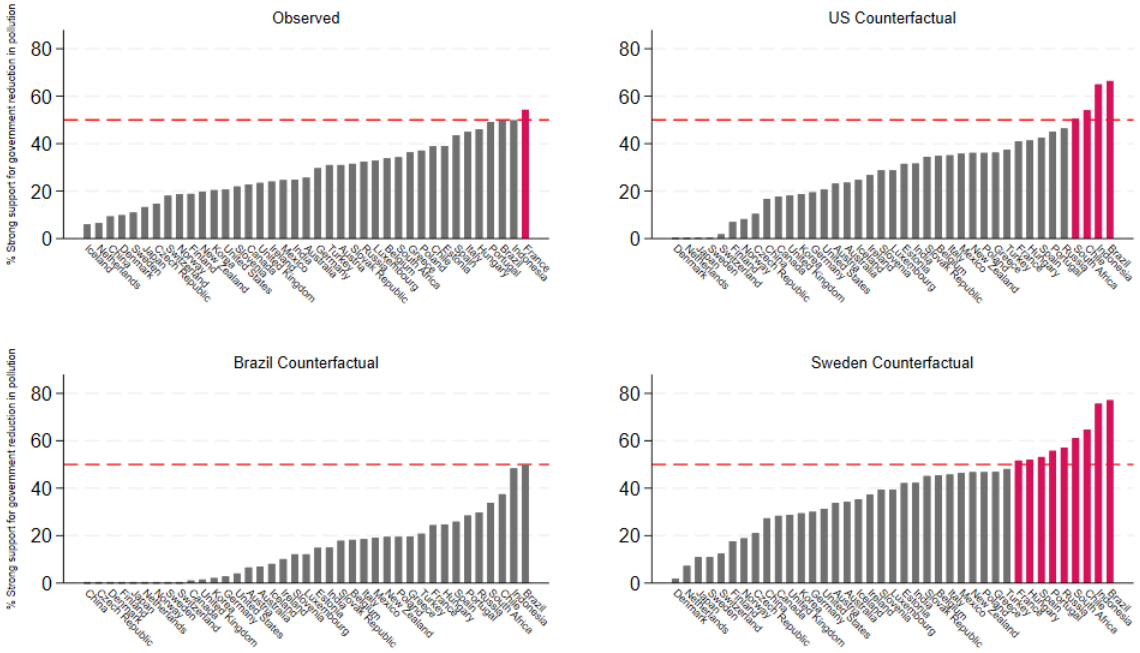


(b) Environmental quality (PM2.5 and HAP) exposure



Notes: The figures plots the marginal effect of EPS exposure in formative age on three different heterogeneity dimensions. Panel (a) interacts EPS exposure with a dummy (Δ policy) equal to one (i.e., negative) when the environmental policy stringency in the year of the interview is lower than the average stringency in the formative age. Panel (b) interacts EPS exposure with a dummy for the exposure to environmental quality measures (PM2.5 and household air pollution - HAP) is above or below the median in the sample. Each regression controls for male dummy, unemployment dummy, 3-category education (Lower, i.e., primary education, Middle, i.e., secondary education, Upper, i.e., tertiary education), 10-class subjective income decile scale, recession exposure, PM2.5, and HAP exposure. The regression also includes the year of birth, country-by-year of interview, survey fixed effects, and country-by-age linear trends. Observations are weighted using survey weights. Standard errors are clustered at the country-year of interview level. Bins represent the 95% confidence intervals around point estimates.

Figure 4: Green preferences for counterfactual policy stringency



Notes: Each histogram reports the share of respondents that strongly agree with the statement "Government should reduce pollution" observed in the country across survey waves (top-left corner histogram) and predicted using three different counterfactual policy experiments. From the top-right corner histogram clockwise, each histogram reports the model-predicted support induced by the difference between the birth-cohort exposure in policy stringency in the country and the exposure in policy stringency experienced by the same birth cohort in the US, in Sweden, and in Brazil. The red dashed line indicates the 50% and countries that are above the line are color-coded in red.

References

Abel, A. B. (1990). *Asset prices under habit formation and catching up with the Joneses* (NBER Working Paper No. 3279). National Bureau of Economic Research.

Acemoglu, D., Ajzenman, N., Aksoy, C. G., Fiszbein, M., & Molina, C. A. (2024). (Successful) Democracies Breed Their Own Support. *The Review of Economic Studies*, *Forthcoming.*

Aghion, P., Bénabou, R., Martin, R., & Roulet, A. (2023). Environmental preferences and technological choices: Is market competition clean or dirty? *American Economic Review: Insights*, *5*(1), 1–19.

Aghion, P., Bolton, P., Harris, C., & Jullien, B. (1991). Optimal learning by experimentation. *The Review of Economic Studies*, *58*(4), 621–654.

Aklin, M., Bayer, P., Harish, S., & Urpelainen, J. (2013). Understanding environmental policy preferences: New evidence from Brazil. *Ecological Economics*, *94*, 28–36.

Aksoy, C. G., Carpenter, C. S., De Haas, R., & Tran, K. D. (2020). Do laws shape attitudes? evidence from same-sex relationship recognition policies in europe. *European Economic Review*, *124*, 103399.

Alesina, A., & Giuliano, P. (2015). Culture and institutions. *Journal of Economic Literature*, *53*(4), 898–944.

Alesina, A., & Rodrik, D. (1994). Distributive politics and economic growth. *The Quarterly Journal of Economics*, *109*(2), 465–490.

Anderson, S., Marinescu, I., & Shor, B. (2023). Can Pigou at the polls stop us melting the poles? *Journal of the Association of Environmental and Resource Economists*, *10*(4), 903–945.

Barone, G., de Blasio, G., & Poy, S. (2022). The legacy of 1968 student protests on political preferences. *Economics Letters*, *210*, 110198.

Barro, R. J., & Ursúa, J. F. (2008). *Macroeconomic crises since 1870* (NBER Working Paper No. 13940). National Bureau of Economic Research.

Besley, T., & Persson, T. (2019). Democratic values and institutions. *American Economic Review: Insights*, *1*(1), 59–76.

Besley, T., & Persson, T. (2023). The political economics of green transitions. *Quarterly Journal of Economics*, *138*(3), 1863–1906.

Bezin, E. (2015). A cultural model of private provision and the environment. *Journal of Environmental Economics and Management*, *71*, 109–124.

Bisin, A., & Verdier, T. (2001). The economics of cultural transmission and the dynamics of preferences. *Journal of Economic Theory*, *97*(2), 298–319.

Bisin, A., & Verdier, T. (2011). The economics of cultural transmission and socialization. *Handbook of Social Economics* (pp. 339–416). Elsevier.

Bornstein, N., & Lanz, B. (2008). Voting on the environment: Price or ideology? Evidence from Swiss referendums. *Ecological Economics*, *67*(3), 430–440.

Burkhardt, J., & Chan, N. W. (2017). The dollars and sense of ballot propositions: Estimating willingness to pay for public goods using aggregate voting data. *Journal of the Association of Environmental and Resource Economists*, *4*(2), 479–503.

Carattini, S., Baranzini, A., Thalmann, P., Varone, F., & Vöhringer, F. (2017). Green taxes in a post-Paris world: are millions of nays inevitable? *Environmental and Resource Economics*, *68*, 97–128.

Carreri, M., & Teso, E. (2023). Economic recessions and congressional preferences for redistribution. *The Review of Economics and Statistics*, *105*(3), 723–732.

Coppock, A., & Green, D. P. (2016). Is voting habit forming? New evidence from experiments and regression discontinuities. *American Journal of Political Science*, *60*(4), 1044–1062.

Cotofan, M., Cassar, L., Dur, R., & Meier, S. (2023). Macroeconomic conditions when young shape job preferences for life. *The Review of Economics and Statistics*, *105*(2), 467–473.

Cotofan, M., Dur, R., & Meier, S. (2024). Does growing up in economic hard times increase compassion? The case of attitudes towards immigration. *Journal of Economic Behavior & Organization*, *218*, 245–262.

Cutler, N. E. (1974). Aging and generations in politics: The conflict of explanations and inference. *Public Opinion and Political Attitudes*, 440–462.

De Neve, J.-E., Ward, G., De Keulenaer, F., Van Landeghem, B., Kavetsos, G., & Norton, M. I. (2018). The asymmetric experience of positive and negative economic growth: Global evidence using subjective well-being data. *Review of Economics and Statistics*, *100*(2), 362–375.

Dechezleprêtre, A., Fabre, A., Kruse, T., Planterose, B., Chico, A. S., & Stantcheva, S. (2022). *Fighting climate change: International attitudes toward climate policies* (NBER Working Paper w30265). National Bureau of Economic Research.

Douenne, T., & Fabre, A. (2020). French attitudes on climate change, carbon taxation and other climate policies. *Ecological Economics*, *169*, 106496.

Douenne, T., & Fabre, A. (2022). Yellow vests, pessimistic beliefs, and carbon tax aversion. *American Economic Journal: Economic Policy*, *14*(1), 81–110.

Eichengreen, B., Aksoy, C. G., & Saka, O. (2021). Revenge of the experts: Will COVID-19 renew or diminish public trust in science? *Journal of Public Economics*, *193*, 104343.

Eichengreen, B., Saka, O., & Aksoy, C. G. (2024). The political scar of epidemics. *The Economic Journal*, *134*(660), 1683–1700.

Eshbaugh-Soha, M. (2006). The conditioning effects of policy salience and complexity on American political institutions. *Policy Studies Journal*, *34*(2), 223–243.

Falco, C., & Corbi, R. (2023). Natural disasters and preferences for the environment: Evidence from the impressionable years. *Economics Letters*, *222*, 110946.

Fuchs-Schündeln, N., & Schündeln, M. (2015). On the endogeneity of political preferences: Evidence from individual experience with democracy. *Science*, *347*(6226), 1145–1148.

Furnham, A., & Boo, H. C. (2011). A literature review of the anchoring effect. *The Journal of Socio-Economics*, *40*(1), 35–42.

Gerber, E. R., & Jackson, J. E. (1993). Endogenous preferences and the study of institutions. *American Political Science Review*, *87*(3), 639–656.

Holian, M. J., & Kahn, M. E. (2015). Household demand for low carbon policies: Evidence from California. *Journal of the Association of Environmental and Resource Economists*, *2*(2), 205–234.

Huse, C., & Koptyug, N. (2022). Salience and policy instruments: Evidence from the auto market. *Journal of the Association of Environmental and Resource Economists*, *9*(2), 345–382.

Kahn, M. E., & Matsusaka, J. G. (1997). Demand for environmental goods: Evidence from voting patterns on California initiatives. *The Journal of Law and Economics*, *40*(1), 137–174.

Konc, T., Savin, I., & van den Bergh, J. C. (2021). The social multiplier of environmental policy: Application to carbon taxation. *Journal of Environmental Economics and Management*, *105*, 102396.

Krosnick, J. A., & Alwin, D. F. (1989). Aging and susceptibility to attitude change. *Journal of Personality and Social Psychology*, *57*(3), 416.

Kruse, T., Dechezleprêtre, A., Saffar, R., & Robert, L. (2022). Measuring environmental policy stringency in OECD countries. (1703). https://doi.org//10.1787/90ab82e8-en

List, J. A., & Sturm, D. M. (2006). How elections matter: Theory and evidence from environmental policy. *The Quarterly Journal of Economics*, *121*(4), 1249–1281.

Magistretti, G., & Tabellini, M. (2022). *Economic integration and the transmission of democracy* (NBER Working Paper No. 30055). National Bureau of Economic Research.

Mattauch, L., Hepburn, C., Spuler, F., & Stern, N. (2022). The economics of climate change with endogenous preferences. *Resource and Energy Economics*, 101312.

Milgrom, P., & Shannon, C. (1994). Monotone comparative statics. *Econometrica: Journal of the Econometric Society*, 157–180.

Murray, B., & Rivers, N. (2015). British Columbia's revenue-neutral carbon tax: A review of the latest "grand experiment" in environmental policy. *Energy Policy*, *86*, 674–683.

Noailly, J., Nowzohour, L. M., & Van Den Heuvel, M. (2021). *Heard the news? Environmental policy and clean investments* (CIES Research Paper No. 70). Graduate Institute of International and Development Studies, Centre for International Environmental Studies.

Rivers, N., & Schaufele, B. (2015). Salience of carbon taxes in the gasoline market. *Journal of Environmental Economics and Management*, *74*, 23–36.

Roth, C., & Wohlfart, J. (2018). Experienced inequality and preferences for redistribution. *Journal of Public Economics*, *167*, 251–262.

Schumacher, I. (2015). The endogenous formation of an environmental culture. *European Economic Review*, *76*, 200–221.

Severen, C., & Van Benthem, A. A. (2022). Formative experiences and the price of gasoline. *American Economic Journal: Applied Economics*, *14*(2), 256–284.

Stigler, G. J., & Becker, G. S. (1977). De gustibus non est disputandum. *American Economic Review*, *67*(2), 76–90.

Torney, J. V., & Hess, R. D. (1967). *The development of political attitudes in children.* Aldine.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases: Biases in judgments reveal some heuristics of thinking under uncertainty. *Science*, *185*(4157), 1124–1131.

Warren, P. L., & Wilkening, T. S. (2012). Regulatory fog: The role of information in regulatory persistence. *Journal of Economic Behavior & Organization*, *84*(3), 840–856.

Wieland, V. (2000). Learning by doing and the value of optimal experimentation. *Journal of Economic Dynamics and Control*, *24*(4), 501–534.

Wolf, M., Emerson, J., Esty, D., & Wending, Z. (2022). Environmental Performance Index. *New Haven, CT: Yale Center for Environmental Law & Policy*.

Zhao, J., & Kling, C. L. (2003). Policy persistence in environmental regulation. *Resource and Energy Economics*, *25*(3), 255–268.

Supplementary Material

"Endogenous Green Preferences"

Table of Contents

A Data Appendix **154**

A.1 Survey Data 154

A.2 Environmental Policy Stringency 155

A.3 Environmental Quality 164

A.4 Descriptive Statistics 165

B Environmental preferences and subsequent policies **167**

C Adulthood Exposure **170**

C.1 Additional results 170

C.2 Robustness 173

C.3 Falsification Test 183

C.4 Heterogeneity 186

D Additional results for formative age **188**

D.1 Robustness 188

D.2 Controlling for preferences during formative window 194

E Mechanisms **198**

E.1 Environmental policy direction 198

E.2 Environmental Quality exposure 201

E.3 Endogenous Learning 206

F Counterfactual Details **209**

A Data Appendix

A.1 Survey Data

The data on environmental attitudes comes from the Integrated Value Surveys (IVS), which harmonize the European Values Study (EVS) and the World Value Survey (WVS). We focus on two measures of environmental preferences that are related to environmental policies and elicit individuals' agreement in a scale from 1 to 4 where responses can be "strongly disagree", "disagree", "agree" or "strongly agree". The first question asks whether "Government should reduce environmental pollution". The second question is "Can you tell me whether you with the following statement: 'I would agree to an increase in taxes if the extra money were used to prevent environmental pollution' ". We also test for robustness by constructing two binary versions equal to one if the respondent chooses either "Agree" or "Strongly agree", and zero otherwise, and a more conservative equal to one only if the respondent chooses "Strongly agree".

The IVS also contains socio-demographic information about the respondents which we use as individual controls in our estimation. In particular, all our regressions control for gender; employment status (defined using a dummy indicating whether the person is unemployed); education (measured using a three-category variable, where lower education means at most primary education, middle education corresponds to at most secondary education, upper education means at least tertiary education). We also control for income in a non-parametric way by including 10 dummies on the country-specific subjective income decile scale.

A.2 Environmental Policy Stringency

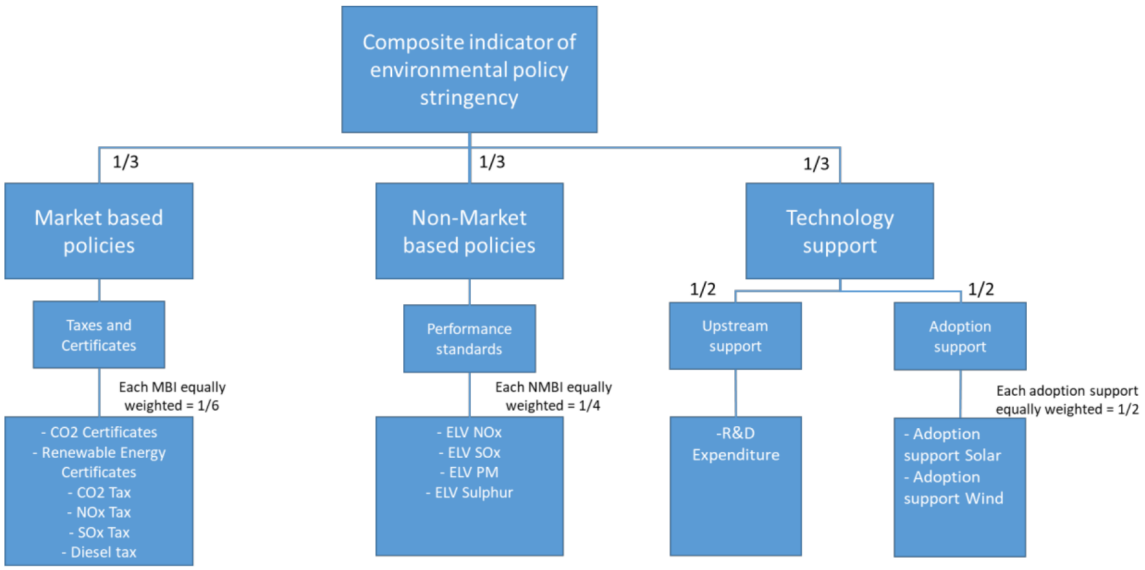
The EPS comprises three sub-indices with equal weight that, in turn, are composed of several policies. The index aggregates the scores given to each policy's stringency on a scale of zero to six. The stringency of environmental policies is measured in different units. As an example, the carbon tax is measured by the tax rate for CO2 emissions, with the raw values in national currency converted to USD/tonne CO2 for international comparison. As another example, the stringency of the fuel (diesel) tax is measured using the tax for a litre of diesel fuel used in transport for industry as a share of the pre-tax diesel price. It is computed by dividing the tax on diesel by the national pre-tax price paid by industry for diesel and the values are converted to USD/Litre.

To aggregate several policy types into a composite index of policy stringency, their stringency is measured on a common scale. For each policy instrument, the raw data is ordered from the least to the most stringent observation across the 1990-2020 period. The lowest score of zero is assigned to observations with no policy in place. The remaining scores are assigned using the distribution of observations that have the policy in place. The highest score of six is assigned to observations with values above the 90th percentile of observations that have the respective policy implemented. To assign the remaining scores, the difference between the 90th and the 10th percentile is divided into five equal bins that define the thresholds (Kruse et al., 2022).

The first sub-index is market-based instruments (MBI) that group policies that put a price on pollution. In particular, it accounts for CO2 Trading Schemes, Renewable Energy Trading Scheme, CO2 Taxes, Nitrogen Oxides (NOx) Tax, Sulphur Oxides (SOx)

Tax, Fuel Tax (Diesel). The second sub-index includes Non-Market Based instruments (NMBI), entailing policies that mandate emission limits and standards: Emission Limit Value (ELV) for nitrogen oxides (NOx); ELV for sulphur oxides (SOx); ELV for Particulate Matter (PM); Sulphur content limit for diesel. The final sub-index, Technology Support (TS), entails that support innovation in clean technologies and their adoption, including Public research and development expenditure (R&D) and Renewable energy support for Solar and Wind. Each component of the sub-index has equal weight within each sub-index (i.e. MBI components have 1/6 weight, NMBI components have 1/4 weight and TS components have 1/2 weight).

Figure A1: Construction of Environmental Policy Stringency Index



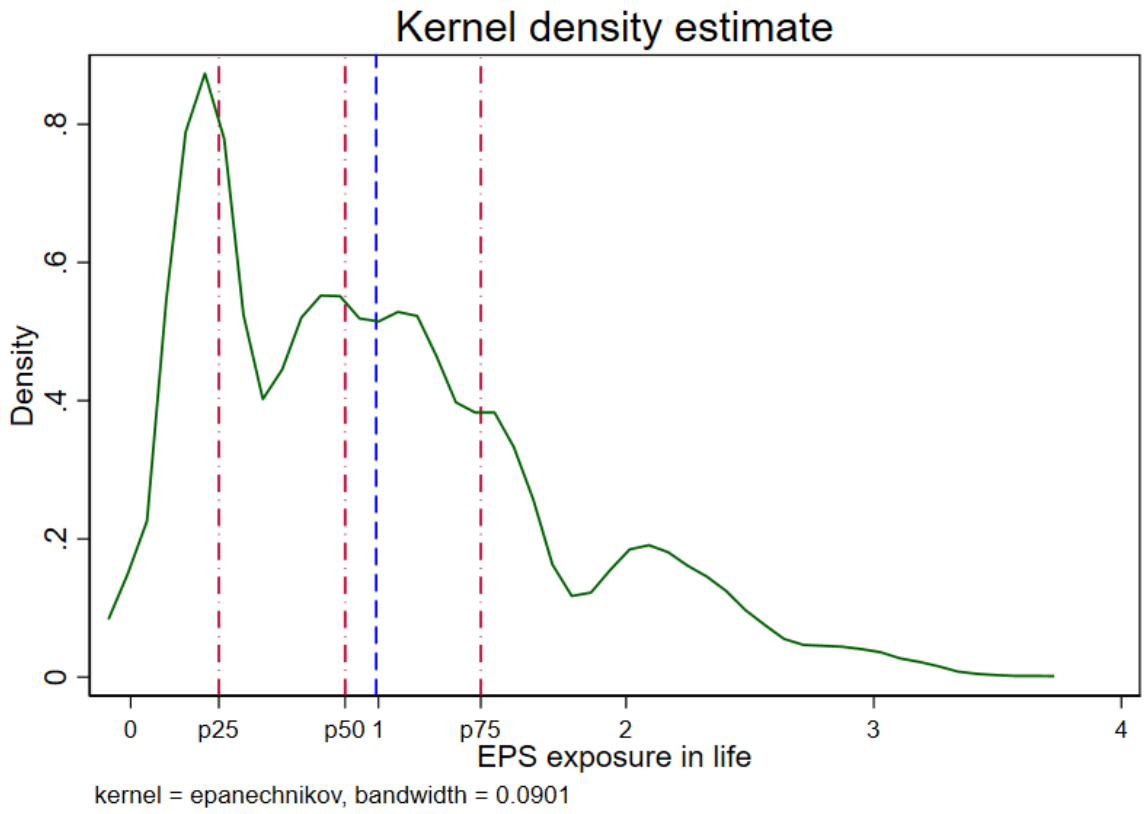
Notes: The figure shows the aggregation structure of the revised EPS index from Kruse et al. (2022).

We construct policy exposure as detailed in the main text in Section 3. Environmental policy preferences are recorded from 1990 to 2010, with certain gaps. Respondents interviewed in the same year and country can have different treatment exposure because

of variation in their year of birth. Figure A2 shows the density distribution of EPS exposure in the final estimation sample. The average stringency is close to one, with the distribution right-skewed. Figure A3 shows the distribution by country. Since most countries exhibit a positive trend over time (Figure A4), we also construct two alternative measures of policy stringency. First, we de-mean each country's level of environmental policy from the annual cross-sectional average. This procedure accounts for policies implemented at supra-national levels which would contemporaneously make more stringent several countries (e.g., the EU Emissions Trading System). Second, we de-trend the policy stringency from a country-specific linear trend. Figure A5 shows the time series of the birth-cohort exposure (averaged across years of interview) for the two alternative measures of EPS exposure for each country.

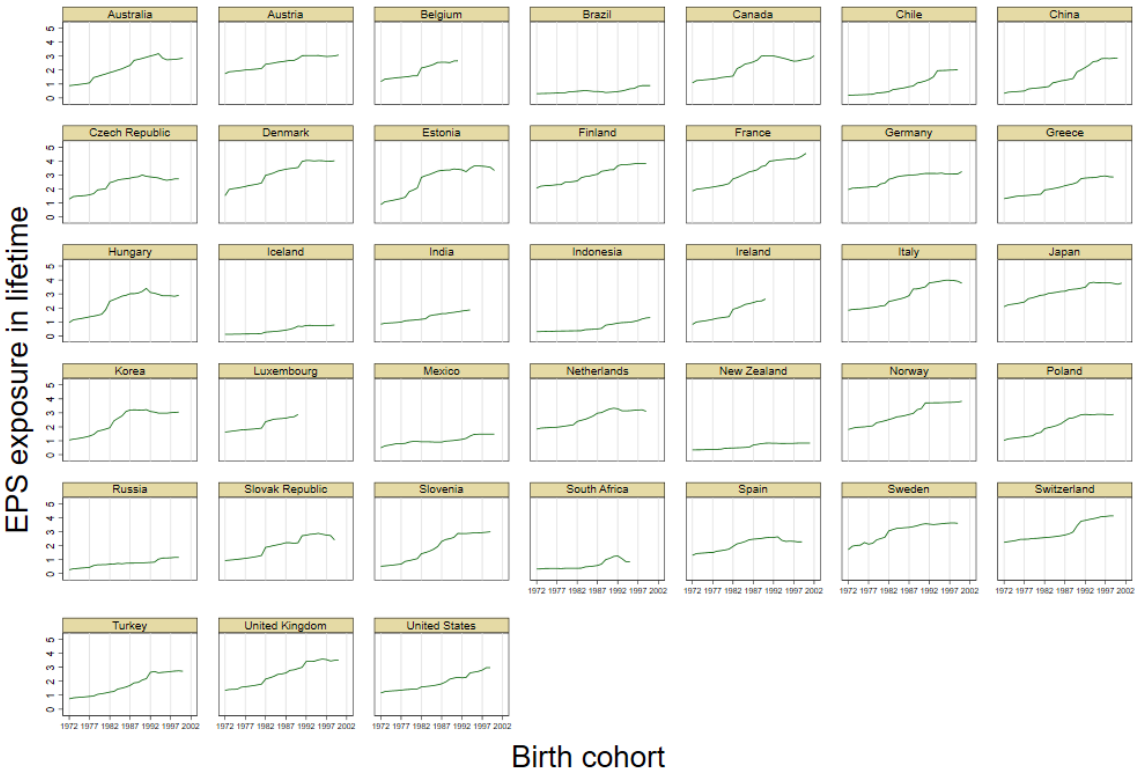
To test for the "formative age" hypothesis, we construct the average policy exposure when the respondent was aged 18 to 25. Appendix Figure A6) displays the formative age exposure for each birth cohort (averaged across years of interview). We construct the other eight-year age window exposures starting from the range of the impressionable years. For the subset of individuals who are either too young or too old, we use all available years over the 8-year formative age window. Results are robust to dropping these individuals. We also test for the robustness of the main results using alternative definitions of formative age (Table D2), alternative exposure measures that are de-meaned and de-trended (Table D3), and using four-year age windows to construct exposures in Figure D2.

Figure A2: EPS average exposure density distribution



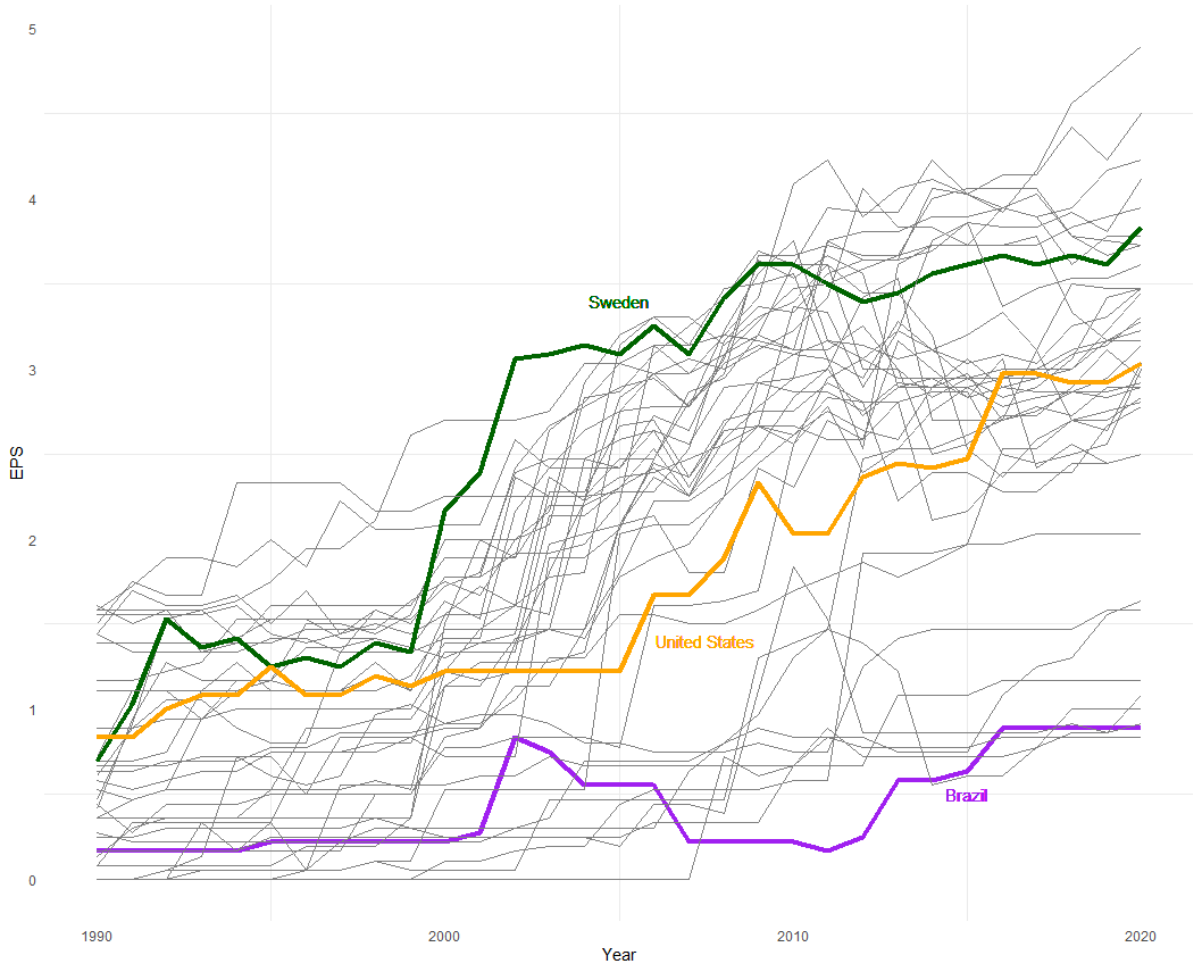
Notes: The figure shows the density distribution of EPS exposure during life in the final estimation sample, with red dot-dashed lines indicating the 25th, 50th, 75th percentile and the blue dashed line indicating the mean.

Figure A3: Country-specific EPS exposure by birth cohort



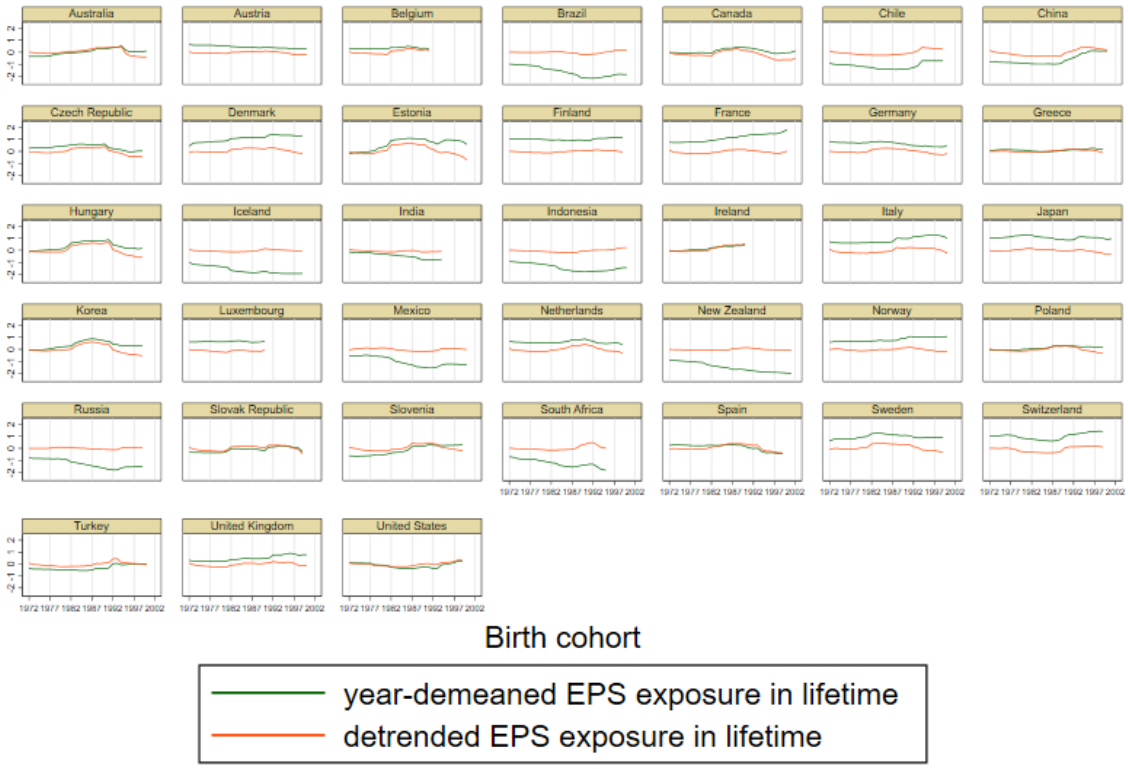
Notes: The figure shows the time series of adulthood mean EPS exposure by cohort for each country over the 30 years available.

Figure A4: Country-specific EPS time series



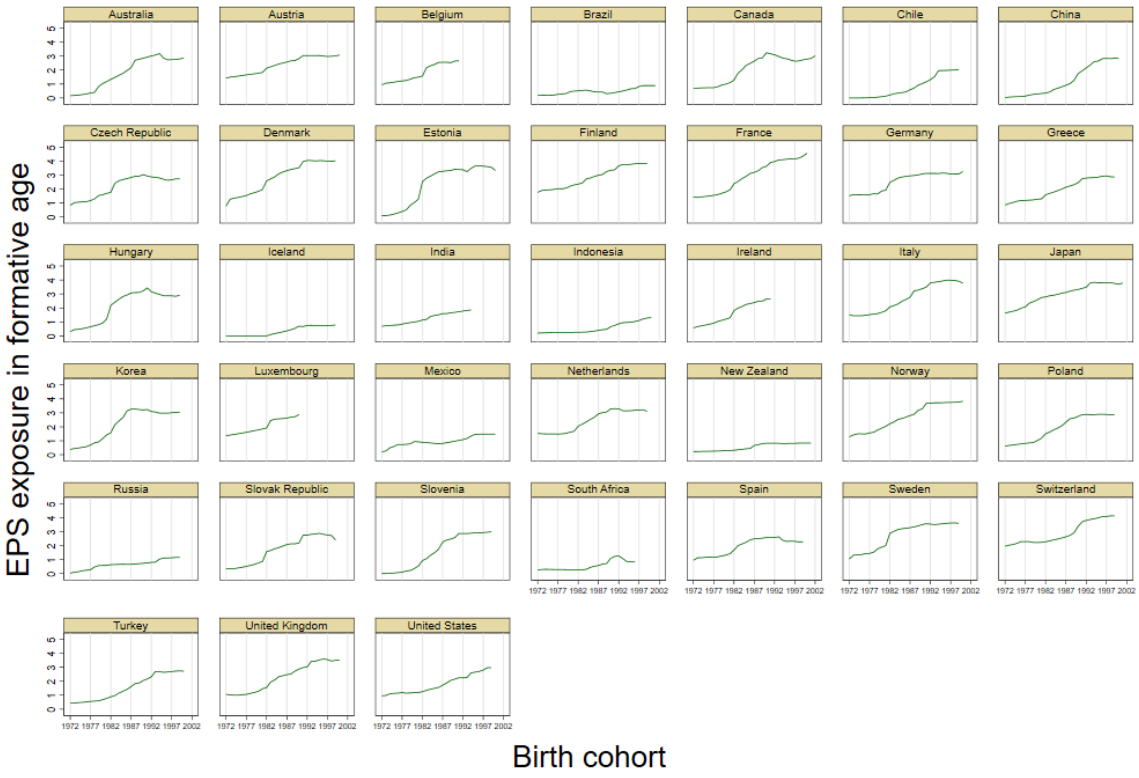
Notes: The figure shows the time series of EPS in each country over time.

Figure A5: Country-specific relative EPS exposure in adult life by birth cohort



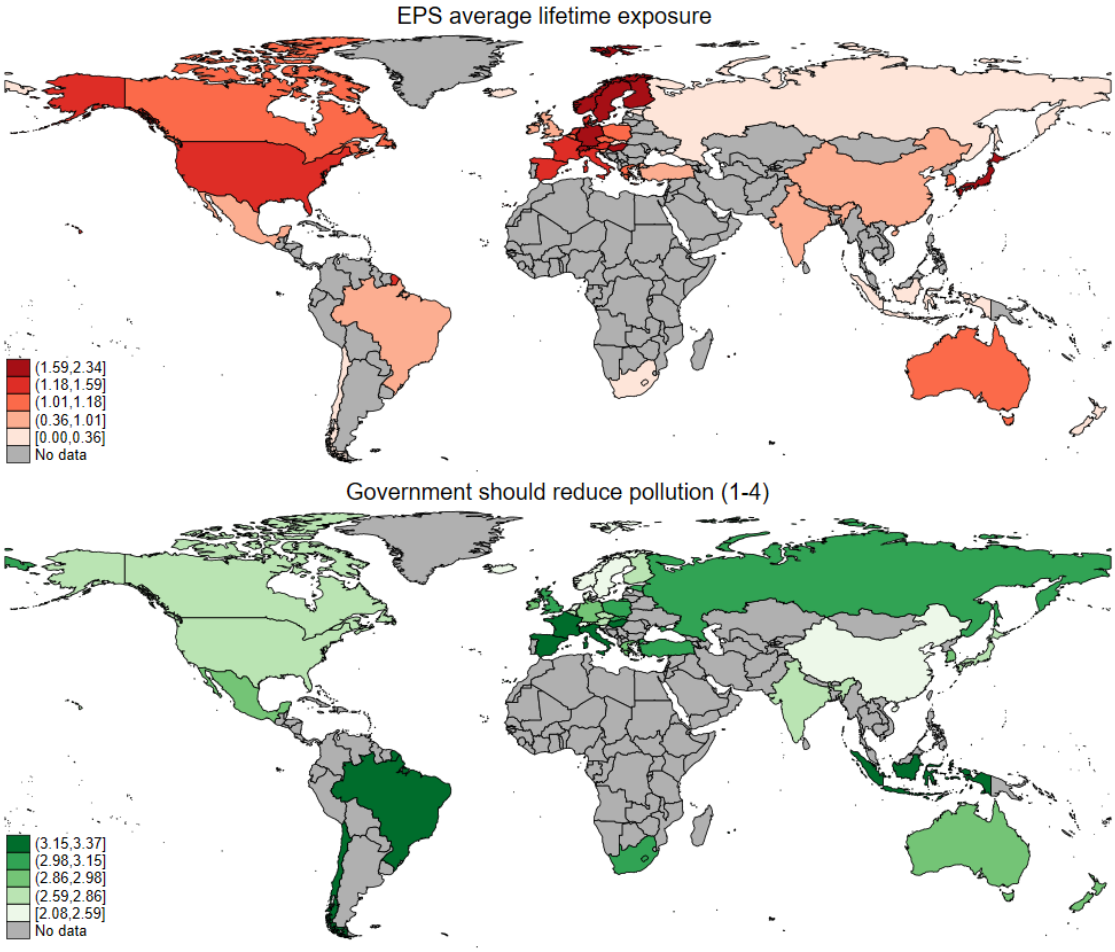
Notes: The figure shows the time series of adulthood mean year-demeaned (in green) and detrended (in red) EPS exposure by cohort for each country over the 30 years available.

Figure A6: EPS exposure in formative age (18-25) by country across birth cohorts



Notes: The figure shows the average EPS exposure across birth cohorts in each country in the IVS sample. Since EPS data are available from 1990 onwards, the oldest birth cohort is those who were born in 1972.

Figure A7: Country pooled average of EPS average exposure across cohorts in baseline sample



Notes: The world map shows the pooled country mean in the final estimation sample of EPS average exposure.

A.3 Environmental Quality

We measure environmental quality using the Environmental Performance Index (EPI) (Wolf et al., 2022), a comprehensive index of global sustainability which combines 40 performance indicators categorized into 11 issues to assess the performance in environmental quality issues. These issues encompass climate change performance, environmental health, and ecosystem vitality, offering a comprehensive overview of a nation's environmental policies and practices. The indicators serve as benchmarks, measuring how closely countries align with established environmental targets and providing insight into areas for improvement to advance sustainability efforts worldwide.

The EPI employs a diverse set of indicators across various issue categories to assess environmental performance comprehensively. The issue categories include climate change mitigation, air quality, sanitation and drinking water, heavy metals, waste management, biodiversity and habitat, ecosystem service, For instance, in the climate change category, indicators such as projected greenhouse gas emissions in 2050, GHG emissions per capita, and CO2 growth rate are evaluated, with each indicator weighted according to its significance. In the environmental health category (HLT), air quality indicators including PM2.5 exposure and household solid fuels are given substantial weight, reflecting the critical importance of addressing air pollution for public health.

The EPI utilizes a weighted scoring system to reflect the relative importance of each issue category and indicator. For example, in the climate change category, mitigation efforts (CCH) hold the highest weight (36.3%), with indicators such as projected GHG emissions and CO2 growth rate receiving significant emphasis. In contrast, household air

pollution (HAD) and PM2.5 exposure (PMD) are weighted at 38% and 47%, respectively, within the air quality subcategory of environmental health (HLT), underscoring their significance in assessing environmental health risks.

In our empirical analysis, we focus on two of the most salient indicators of environmental quality and that can be directly related to air pollution issues as in the World Value Survey questions. We focus on household air pollution from solid fuels (HAP) and ambient particulate matter pollution (PMD). HAD is measured by the number of age-standardized disability-adjusted life-years lost per 100,000 persons due to exposure to household air pollution from solid fuels. Similarly, PMD exposure is quantified by the number of age-standardized disability-adjusted life-years lost per 100,000 persons due to exposure to fine particulate matter smaller than 2.5 micrometers (PM2.5). These indicators provide crucial insights into the health impacts of air pollution, informing environmental health policies and interventions.

A.4 Descriptive Statistics

Table A1 reports the descriptive statistics for the main variables in the final baseline estimation sample. Environmental Quality exposure measures are expressed in terms of the average number of age-standardized disability-adjusted life-years lost per 100 million people, and in the estimation sample the measures of environmental quality are transformed for ease of interpretation such that higher values are associated with better environmental conditions. Recession exposure takes value one if the national GDP growth rate dropped by at least 10% during the adult years of the individual (Barro &

Ursúa, 2008). We use GDP data from the World Bank - World Development Indicators to compute GDP growth.

Table A1: Summary statistics

	N	mean	SD	min	max
Government should reduce environmental pollution	16889	2.909	0.896	1	4
Environmental Policy Stringency Exposure					
EPS	16889	0.977	0.696	0	3.639
Market (MBI)	16889	0.696	0.547	0	4
Non-Market (NMBI)	16889	1.568	1.219	0	5.250
Tax (CO2, NOX, SOX, Diesel)	16889	1.001	0.771	0	4
CO2 Tax + Diesel Tax	16889	1.718	1.168	0	5.038
CO2 (Tax/Cap-and-Trade)	16889	0.238	0.655	0	5
CO2 (Tax/Cap-and-Trade) + Diesel Tax	16889	1.179	0.810	0	4.667
Fuel (Diesel) Tax	16889	3.060	1.862	0	6
Environmental Quality Exposure					
PM2.5	16889	1184.094	679.706	148.837	2302.446
HAP	16889	847.243	1419.208	0.429	7043.224
Recession exposure	16889	0.010	0.037	0.000	0.286
Number of countries	38				

Notes: Summary statistics are computed using the final estimation sample.

B Environmental preferences and subsequent policies

One of our primary results is the positive effect of past policy exposure on subsequent environmental preferences. This result, when combined with evidence of the effect of preferences on subsequent policy, can be interpreted as a case of positive endogeneity of environmental preferences, which has been previously documented in the case of democratic values and institutions (Acemoglu et al., 2024; Besley & Persson, 2019; Fuchs-Schündeln & Schündeln, 2015).

We confirm the effect of preferences on policy outcomes in our setting, the relatively uncontroversial of the bidirectional relationship that we study in our paper. This exercises also has the added benefit of confirming that our survey measures of environmental preferences are meaningfully related to economically relevant policy outcomes. We test for the association between them at the country level. We regress the change in EPS in country c in year t on the weighted-average environmental preferences at the country level recorded at time t, estimating the following equation

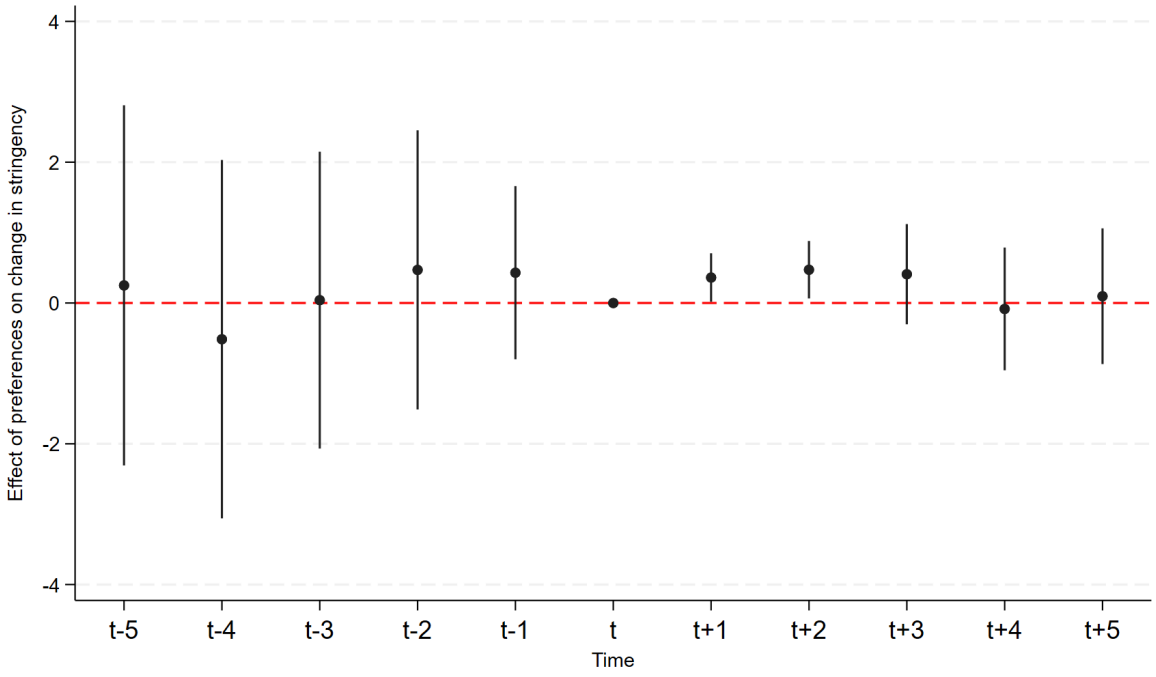
$$\begin{aligned}(EPS_{c\tau} - EPS_{ct}) =& \beta_\tau preferences_{ct} + \alpha_c + \mu_t + \varepsilon_{ct} \\ & \forall \tau \in \{-5, -4, .., -1, 1, ..5\}\end{aligned} \tag{B.9}$$

In doing so, for each regression we estimate using Equation (B.9), we obtain a coefficient β_τ to see if our main survey question predicts policy levels in the near future. Figure B1 plots the β_τ's coefficients obtained. We find that our primary survey measure of environmental policy preferences significantly predicts subsequent changes to environmental stringency for up to two years following the year of survey t. As a placebo test, we

also estimate the association between preferences and past policy deviations from contemporaneous policy levels (for $\tau < 0$). Reassuringly, there is no statistically significant relationship between the survey measure and changes to policy stringency levels in any of the five prior years to the survey.

This exercise allays concerns on the relevance of our environmental preferences and policy measures. Moreover, these results do not affect the internal validity of our approach for two main reasons. First, we consider simultaneous changes between preferences on policy deviations. Our baseline empirical approach focuses on changes in past policy exposure and the effect on subsequent long-lasting policy preferences measured at the time of the survey. Second, to the extent that past policy preferences may still be a confound to the effect of past policy stringency on subsequent preferences, our results are robust to controlling for aggregate country-level preferences during each individual's formative age, as reported in Table D4.

Figure B1: Event-study of environmental survey preferences on changes in policy stringency levels



Notes: The figure reports the coefficients associated with environmental preferences on changes in environmental policy stringency. Each coefficient is obtained from a regression of environmental preferences on a lag/lead of changes in policy stringency with respect to the level of policy stringency contemporaneous to the year of the survey. Environmental preferences are measured as the country-average level of agreement with the question "Government should reduce pollution" on a scale from 1 to 4. The regression also controls for country and year fixed effects. Bins represent the 95% confidence intervals.

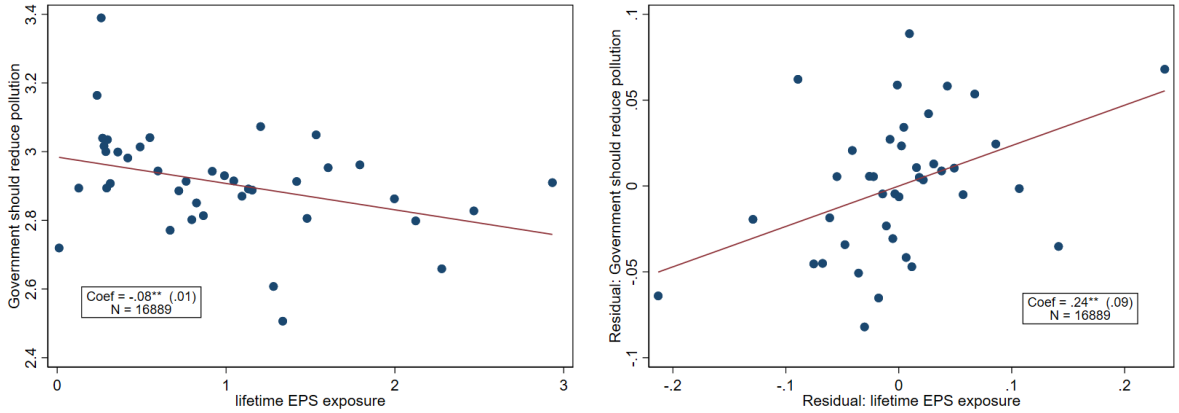
C Adulthood Exposure

C.1 Additional results

Figure C1: Binned scatter plot of environmental preferences and EPS exposure

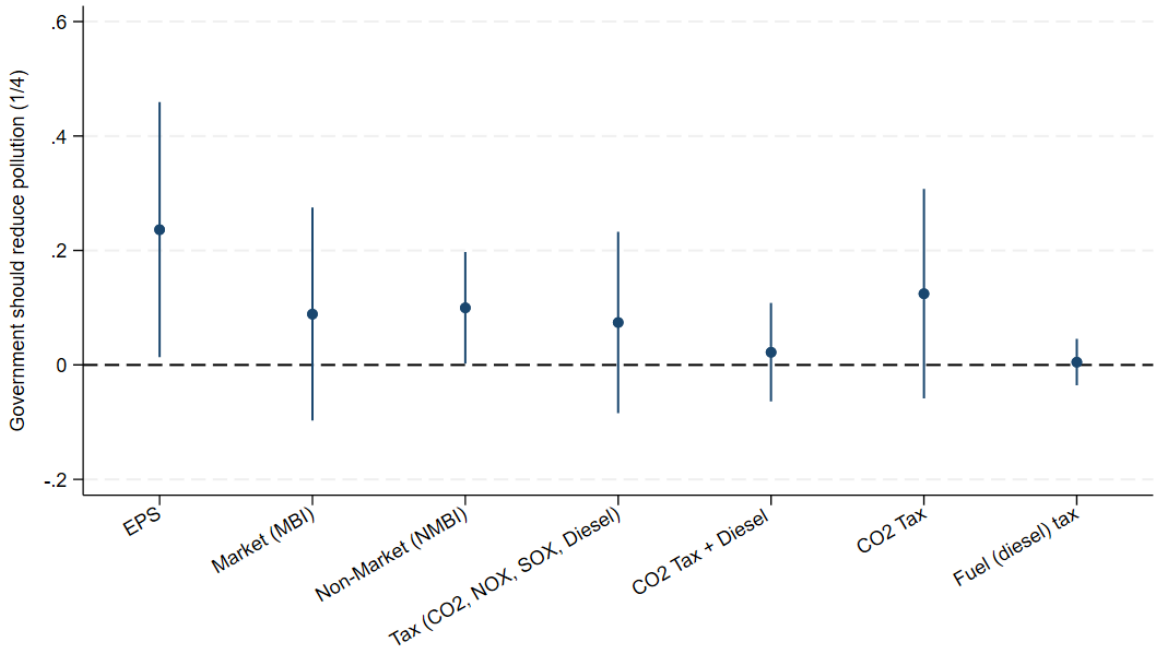
(a) Partial correlation

(b) Residualized partial correlation



Notes: The baseline estimation sample is split into 40 equal-sized bins. Panel (a) shows the raw partial correlation of each mean of that bin between EPS exposure and the government's role in reducing pollution. In Panel (b) each data point shows the mean residual of that bin after controlling for year of birth, country-by-year of interview, survey fixed effects, and country-by-age linear trends. Observations are weighted by the survey weights.

Figure C2: Effect of EPS sub-indices exposure in adulthood on government's role to reduce pollution



Notes: The figure shows the coefficient associated with the average level of sub-index in the x-axis during the adulthood of an individual, using as outcome the question "Government should reduce pollution" in the Likert scale. All regressions control for male dummy, unemployment dummy, 3-category education (Lower, i.e., primary education, Middle, i.e., secondary education, Upper, i.e., tertiary education), 10-class subjective income decile scale, recession exposure, PM2.5 and HAP exposure. The regression also includes year of birth, country-by-year of interview, survey fixed effects, and country-by-age linear trends. Observations are weighted using survey weights. Standard errors are clustered at the country-year of interview level. Bins represent the 95% confidence intervals around point estimates. Tabular results are reported in Appendix Table C8.

Table C1: Environmental Policy Stringency (EPS) exposure during adult life (with controls' estimates)

	Government should reduce pollution (Likert)				
	(1)	(2)	(3)	(4)	(5)
Environmental Policy Stringency (EPS) exposure	0.228**	0.225**	0.219**	0.228**	0.236**
	(0.109)	(0.110)	(0.107)	(0.109)	(0.111)
Recession exposure		-0.515	-0.528	-0.534	-0.518
		(0.381)	(0.406)	(0.378)	(0.392)
PM2.5 exposure			0.314*		0.354**
			(0.186)		(0.173)
HAP exposure				-0.198***	-0.221**
				(0.0713)	(0.0905)
Male	0.0268	0.0267	0.0267	0.0269	0.0269
	(0.0183)	(0.0183)	(0.0183)	(0.0183)	(0.0183)
Education: Reference category: Lower education					
Middle	-0.0716**	-0.0718**	-0.0718**	-0.0718**	-0.0718**
	(0.0307)	(0.0307)	(0.0307)	(0.0307)	(0.0307)
Upper	-0.243***	-0.243***	-0.243***	-0.243***	-0.243***
	(0.0433)	(0.0433)	(0.0433)	(0.0433)	(0.0433)
Unemployed	0.0660***	0.0661***	0.0660***	0.0656***	0.0656***
	(0.0221)	(0.0221)	(0.0221)	(0.0221)	(0.0221)
Income deciles: Reference category: Bottom decile					
2nd	-0.00528	-0.00519	-0.00521	-0.00526	-0.00524
	(0.0361)	(0.0362)	(0.0362)	(0.0363)	(0.0363)
3rd	-0.0444	-0.0443	-0.0443	-0.0441	-0.0441
	(0.0391)	(0.0391)	(0.0391)	(0.0392)	(0.0392)
4th	-0.0779**	-0.0779**	-0.0779**	-0.0776**	-0.0775**
	(0.0339)	(0.0338)	(0.0339)	(0.0339)	(0.0339)
5th	-0.0277	-0.0281	-0.0281	-0.0279	-0.0278
	(0.0321)	(0.0321)	(0.0321)	(0.0322)	(0.0322)
6th	-0.0415	-0.0418	-0.0418	-0.0414	-0.0414
	(0.0392)	(0.0392)	(0.0392)	(0.0393)	(0.0393)
7th	-0.0853**	-0.0857**	-0.0857**	-0.0854**	-0.0853**
	(0.0401)	(0.0400)	(0.0400)	(0.0401)	(0.0401)
8th	-0.0671*	-0.0671*	-0.0671*	-0.0668*	-0.0667*
	(0.0385)	(0.0385)	(0.0385)	(0.0386)	(0.0386)
9th	-0.151***	-0.152***	-0.152***	-0.152***	-0.152***
	(0.0553)	(0.0553)	(0.0553)	(0.0554)	(0.0554)
10th	-0.0983*	-0.0990*	-0.0991*	-0.0987*	-0.0985*
	(0.0555)	(0.0555)	(0.0554)	(0.0555)	(0.0555)
Survey FE	✓	✓	✓	✓	✓
Year of birth FE	✓	✓	✓	✓	✓
Country × Year of interview FE	✓	✓	✓	✓	✓
Country-age linear trends	✓	✓	✓	✓	✓
Mean outcome	2.909	2.909	2.909	2.909	2.909
SD Outcome	0.896	0.896	0.896	0.896	0.896
N	16889	16889	16889	16889	16889
adj. R^2	0.132	0.132	0.132	0.132	0.132

Notes: The outcome variable ranges from 1 to 4 (with higher values reflecting stronger agreement with the statement "Government should reduce environmental pollution"). All regressions control for male dummy, unemployment dummy, 3-category education (Lower, i.e., primary education, Middle, i.e., secondary education, Upper, i.e., tertiary education), 10-class subjective income decile scale. Recession exposure is a dummy variable that takes value one if during the formative age the individual has experienced at least one year in which the real GDP per capita growth rate of its country is at least 10% below the previous year's growth rate. PM2.5 and HAP exposure are respectively the average of age-standardized disability-adjusted life-years lost per 100,000 people due to exposure to fine air particulate matter smaller than 2.5 micrometers (PM2.5) and due to exposure to household air pollution (HAP) from the use of household solid fuels (PM2.5 and HAP exposure variables are inverted such that higher values imply higher quality and divided by 1000 to improve readability of coefficients). All regressions also include year of birth, country-by-year of interview, survey fixed effects, and country-by-age linear trends. Observations are weighted using survey weights. Standard errors are clustered by country-year. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

C.2 Robustness

We find coefficients quantitatively very similar to our baseline with alternative measures of policy stringency de-meaned from the annual cross-sectional average and de-trended from a country-specific linear trend (Appendix Table C2). We examine the influence of starting age, and find that the effect becomes more robust from age 16 across all policy exposure measures, including demeaned and detrended (Appendix Table C3). This result supports our hypothesized channel of norm formation, which requires awareness of and influence by policy stringency, and becomes likelier approaching adulthood.[15]

We also adopt alternative estimation methods, using ordered probit (Appendix Table C4), probit and linear probability models for binary outcomes (Appendix Tables C5 and C6). Results remain quantitatively comparable. In the latter exercise, increasing stringency exposure from the US level to Sweden's level in 2019 increases the probability by 5 percentage points (7.5% at the mean) of expressing that government should reduce pollution and of strongly agreeing by 9.4 p.p. (31%). Results are also robust to alternative fixed effects (Appendix Table C7) and to "leave-one-country-out" (Appendix Figure C3).

[15]Differently than exposure to democratic institutions which are assumed to be understood by individuals as of age of six (Acemoglu et al., 2024), environmental policies and their effects are posited to be internalized from adult life (Aklin et al., 2013).

Table C2: EPS exposure with relative measure (starting from 18 years)

	Government should reduce pollution (Likert)			
	(1)	(2)	(3)	(4)
Panel A: Year-demeaned relative exposure				
EPS average exposure	0.190*	0.182*	0.190*	0.194*
	(0.109)	(0.106)	(0.107)	(0.109)
Panel B: Country-detrended exposure				
EPS average exposure	0.220*	0.214**	0.224**	0.231**
	(0.110)	(0.107)	(0.109)	(0.111)
Recession exposure	✓	✓	✓	✓
PM2.5 exposure		✓		✓
HAP exposure			✓	✓
Individual controls	✓	✓	✓	✓
Survey FE	✓	✓	✓	✓
Year of birth FE	✓	✓	✓	✓
Country × Year of interview FE	✓	✓	✓	✓
Country-age linear trends	✓	✓	✓	✓

Notes: The outcome variable ranges from 1 to 4 (with higher values reflecting stronger agreement with the statement "Government should reduce environmental pollution"). All regressions control for male dummy, unemployment dummy, 3-category education (Lower, i.e., primary education, Middle, i.e., secondary education, Upper, i.e., tertiary education), 10-class subjective income decile scale. Recession exposure is a dummy variable that takes value one if during adulthood the individual has experienced at least one year in which the real GDP per capita growth rate of its country is at least 10% below the previous year's growth rate. PM2.5 and HAP exposure are respectively the adult years average of age-standardized disability-adjusted life-years lost per 100,000 people due to exposure to fine air particulate matter smaller than 2.5 micrometers (PM2.5) and due to exposure to household air pollution (HAP) from the use of household solid fuels (PM2.5 and HAP exposure variables are divided by 1000 to improve readability of coefficients). All regressions also control for survey, year of birth, country-by-year of interview fixed effects and country-age linear trends. Observations are weighted using survey weights. Standard errors are clustered by country-year. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table C3: EPS exposure in adult life with different starting ages

	Government should reduce pollution			
	(1)	(2)	(3)	(4)
Panel A: Baseline exposure measure				
EPS average exposure	0.226	0.367	0.345*	0.398***
	(0.349)	(0.240)	(0.185)	(0.140)
Panel B: Year-demeaned relative exposure				
EPS average exposure	0.263	0.467**	0.331*	0.400**
	(0.324)	(0.224)	(0.187)	(0.152)
Panel C: Country-detrended exposure				
EPS average exposure	0.201	0.353	0.335*	0.390***
	(0.351)	(0.241)	(0.184)	(0.141)
Starting age	14 years	15 years	16 years	17 years
Recession exposure	✓	✓	✓	✓
EQ exposure	✓	✓	✓	✓
Individual controls	✓	✓	✓	✓
Survey FE	✓	✓	✓	✓
Year of birth FE	✓	✓	✓	✓
Country × Year of interview FE	✓	✓	✓	✓
Country-age linear trends	✓	✓	✓	✓
N	11730	13129	14449	15605

Notes: The outcome variable ranges from 1 to 4 (with higher values reflecting stronger agreement with the statement "Government should reduce environmental pollution"). All regressions control for male dummy, unemployment dummy, 3-category education (Lower, i.e., primary education, Middle, i.e., secondary education, Upper, i.e., tertiary education), 10-class subjective income decile scale. Recession exposure is a dummy variable that takes value one if during adulthood the individual has experienced at least one year in which the real GDP per capita growth rate of its country is at least 10% below the previous year's growth rate. PM2.5 and HAP exposure are respectively the adult years average of age-standardized disability-adjusted life-years lost per 100,000 people due to exposure to fine air particulate matter smaller than 2.5 micrometers (PM2.5) and due to exposure to household air pollution (HAP) from the use of household solid fuels (PM2.5 and HAP exposure variables are divided by 1000 to improve readability of coefficients). All regressions also control for survey, year of birth, country-by-year of interview fixed effects and country-age linear trends. Observations are weighted using survey weights. Standard errors are clustered by country-year. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table C4: Ordered Probit: EPS average exposure during adulthood

	Government should reduce pollution (Likert)			
	(1)	(2)	(3)	(4)
Environmental Policy Stringency (EPS) exposure	0.129*	0.130*	0.130*	0.355**
	(0.0703)	(0.0704)	(0.0676)	(0.157)
Recession exposure	0.00229	-0.00384	-0.0186	-0.00549
	(0.0565)	(0.0574)	(0.0557)	(0.0514)
PM2.5 exposure		-0.0838		0.564**
		(0.158)		(0.275)
HAP exposure			0.201***	-0.203**
			(0.0494)	(0.0942)
Individual controls	✓	✓	✓	✓
Survey FE	✓	✓	✓	✓
Year of birth FE	✓	✓	✓	✓
Country × Year of interview FE	✓	✓	✓	✓
Country-age linear trends	✓	✓	✓	✓
Mean Outcome	2.923	2.923	2.923	2.923
SD Outcome	0.90	0.90	0.90	0.90
N	16889	16889	16889	16889

Notes: The outcome variable ranges from 1 to 4 (with higher values reflecting stronger agreement with the statement "Government should reduce environmental pollution"). All regressions control for male dummy, unemployment dummy, 3-category education (Lower, i.e., primary education, Middle, i.e., secondary education, Upper, i.e., tertiary education), 10-class subjective income decile scale. Recession exposure is a dummy variable that takes value one if during adulthood the individual has experienced at least one year in which the real GDP per capita growth rate of its country is at least 10% below the previous year's growth rate. PM2.5 and HAP exposure are respectively the adult years average of age-standardized disability-adjusted life-years lost per 100,000 people due to exposure to fine air particulate matter smaller than 2.5 micrometers (PM2.5) and due to exposure to household air pollution (HAP) from the use of household solid fuels (PM2.5 and HAP exposure variables are divided by 1000 to improve readability of coefficients). All regressions also control for survey, year of birth, country-by-year of interview fixed effects and country-age linear trends. Observations are weighted using survey weights. Standard errors are clustered by country-year. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table C5: Environmental Policy Stringency (EPS) exposure with binary outcomes. Probit estimates.

	Government should reduce pollution			
	(1)	(2)	(3)	(4)
Panel A: Agree & Strongly Agree				
EPS exposure	0.265*	0.191	0.275*	0.221
	(0.145)	(0.141)	(0.141)	(0.143)
Recession exposure	0.000541	-0.0105	-0.00919	-0.0142
	(0.0666)	(0.0728)	(0.0684)	(0.0720)
PM2.5 exposure		-1.186**		-0.824
		(0.579)		(0.536)
HAP exposure			-0.983*	-0.714
			(0.524)	(0.530)
EPS Marginal effect	0.095*	0.068	0.098**	0.079
	(0.052)	(0.050)	(0.050)	(0.050)
Mean outcome	0.67	0.67	0.67	0.67
SD Outcome	0.47	0.47	0.47	0.47
Panel B: Strongly Agree				
EPS exposure	0.368**	0.434**	0.366**	0.447**
	(0.181)	(0.182)	(0.184)	(0.190)
Recession exposure	-0.0316	-0.0227	-0.0303	-0.0227
	(0.0635)	(0.0602)	(0.0621)	(0.0609)
PM2.5 exposure		0.990**		1.143*
		(0.504)		(0.647)
HAP exposure			0.213	-0.238
			(0.452)	(0.470)
EPS Marginal effect	0.125**	0.147**	0.124**	0.152**
	(0.061)	(0.062)	(0.062)	(0.065)
Mean outcome	0.30	0.30	0.30	0.30
SD Outcome	0.46	0.46	0.46	0.46
N	16889	16889	16889	16889

Notes: All regressions are estimated using a Probit model and control for male dummy, unemployment dummy, 3-category education (Lower, i.e., primary education, Middle, i.e., secondary education, Upper, i.e., tertiary education), 10-class subjective income decile scale. Recession exposure is a dummy variable that takes value one if during adulthood the individual has experienced at least one year in which the real GDP per capita growth rate of its country is at least 10% below the previous year's growth rate. PM2.5 and HAP exposure are respectively the adult years average of age-standardized disability-adjusted life-years lost per 100,000 people due to exposure to fine air particulate matter smaller than 2.5 micrometers (PM2.5) and due to exposure to household air pollution (HAP) from the use of household solid fuels (PM2.5 and HAP exposure variables are divided by 1000 to improve readability of coefficients). All regressions also control for survey, year of birth, country-by-year of interview fixed effects and country-age linear trends. Observations are weighted using survey weights. Standard errors are clustered by country-year. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table C6: Environmental Policy Stringency (EPS) exposure with binary outcomes. OLS estimates.

	Government should reduce pollution			
	(1)	(2)	(3)	(4)
Panel A: Agree & Strongly Agree				
Environmental Policy Stringency (EPS) exposure	0.0862*	0.0634	0.0895**	0.0718
	(0.0441)	(0.0431)	(0.0429)	(0.0438)
Recession exposure	0.00137	-0.00180	-0.000247	-0.00203
	(0.0202)	(0.0219)	(0.0206)	(0.0217)
PM2.5 exposure		-0.355*		-0.259
		(0.178)		(0.173)
HAP exposure			-0.265*	-0.179
			(0.154)	(0.158)
Mean outcome	0.67	0.67	0.67	0.67
SD Outcome	0.47	0.47	0.47	0.47
Panel B: Strongly Agree				
Environmental Policy Stringency (EPS) exposure	0.113*	0.132**	0.112*	0.134**
	(0.0573)	(0.0577)	(0.0584)	(0.0596)
Recession exposure	-0.00756	-0.00490	-0.00708	-0.00493
	(0.0217)	(0.0208)	(0.0213)	(0.0209)
PM2.5 exposure		0.298*		0.310
		(0.174)		(0.193)
HAP exposure			0.0797	-0.0232
			(0.151)	(0.142)
Mean outcome	0.30	0.30	0.30	0.30
SD Outcome	0.46	0.46	0.46	0.46
N	16889	16889	16889	16889

Notes: All regressions control for male dummy, unemployment dummy, 3-category education (Lower, i.e., primary education, Middle, i.e., secondary education, Upper, i.e., tertiary education), 10-class subjective income decile scale. Recession exposure is a dummy variable that takes value one if during adulthood the individual has experienced at least one year in which the real GDP per capita growth rate of its country is at least 10% below the previous year's growth rate. PM2.5 and HAP exposure are respectively the adult years average of age-standardized disability-adjusted life-years lost per 100,000 people due to exposure to fine air particulate matter smaller than 2.5 micrometers (PM2.5) and due to exposure to household air pollution (HAP) from the use of household solid fuels (PM2.5 and HAP exposure variables are divided by 1000 to improve readability of coefficients). All regressions also control for survey, year of birth, country-by-year of interview fixed effects and country-age linear trends. Observations are weighted using survey weights. Standard errors are clustered by country-year. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table C7: Environmental Policy Stringency (EPS) exposure and environmental preferences. Alternative specifications.

	Government should reduce pollution (Likert)				
	(1)	(2)	(3)	(4)	(5)
Environmental Policy Stringency (EPS) exposure	0.227**	0.181**	0.175**	0.127**	0.124**
	(0.108)	(0.0898)	(0.0860)	(0.0527)	(0.0521)
Recession	✓	✓	✓	✓	✓
EQ controls	✓	✓	✓	✓	✓
Individual controls	✓	✓	✓	✓	✓
Survey FE	✓	✓	✓	✓	✓
Country × Year of interview FE	✓	✓	✓	✓	✓
Year of birth FE	✓	✓	✓	✓	✓
Country-age linear trends	✓	✓	✓	✓	✓
Age FE		✓		✓	
Continent × Age FE			✓		✓
Continent-age linear trends				✓	✓
Mean Outcome	2.909	2.909	2.909	2.909	2.909
SD Outcome	0.896	0.896	0.896	0.896	0.896
N	16889	16889	16889	16889	16889
adj. R^2	0.132	0.133	0.134	0.132	0.133

Notes: The outcome variable ranges from 1 to 4 (with higher values reflecting stronger agreement with the statement "Government should reduce environmental pollution"). All regressions control for male dummy, unemployment dummy, 3-category education (Lower, i.e., primary education, Middle, i.e., secondary education, Upper, i.e., tertiary education), 10-class subjective income decile scale, recession exposure and PM2.5 and HAP exposure. Observations are weighted using survey weights. Standard errors are clustered by country-year. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table C8: Exposure to EPS sub-indices during adulthood on government intervention preferences

	Government should reduce pollution (Likert)						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Policy exposure	0.236**	0.0889	0.0998**	0.0742	0.125	0.00492	0.00889
	(0.111)	(0.0930)	(0.0487)	(0.0791)	(0.0914)	(0.0202)	(0.0650)
Policy	EPS	Market	Non-Market	Tax	CO2+Diesel	CO2	Diesel
Recession exposure	✓	✓	✓	✓	✓	✓	✓
Environmental Quality exposure	✓	✓	✓	✓	✓	✓	✓
Individual controls	✓	✓	✓	✓	✓	✓	✓
Survey FE	✓	✓	✓	✓	✓	✓	✓
Year of birth FE	✓	✓	✓	✓	✓	✓	✓
Country × Year of interview FE	✓	✓	✓	✓	✓	✓	✓
Country-age linear trends	✓	✓	✓	✓	✓	✓	✓
Mean Outcome	2.909	2.909	2.909	2.909	2.909	2.909	2.909
SD Outcome	0.896	0.896	0.896	0.896	0.896	0.896	0.896
Mean Exposure	0.977	0.696	1.568	1.001	1.718	0.377	3.060
SD Exposure	0.696	0.547	1.219	0.771	1.168	1.202	1.862
N	16889	16889	16889	16889	16889	16889	16889
adj. R^2	0.132	0.132	0.132	0.132	0.132	0.132	0.132

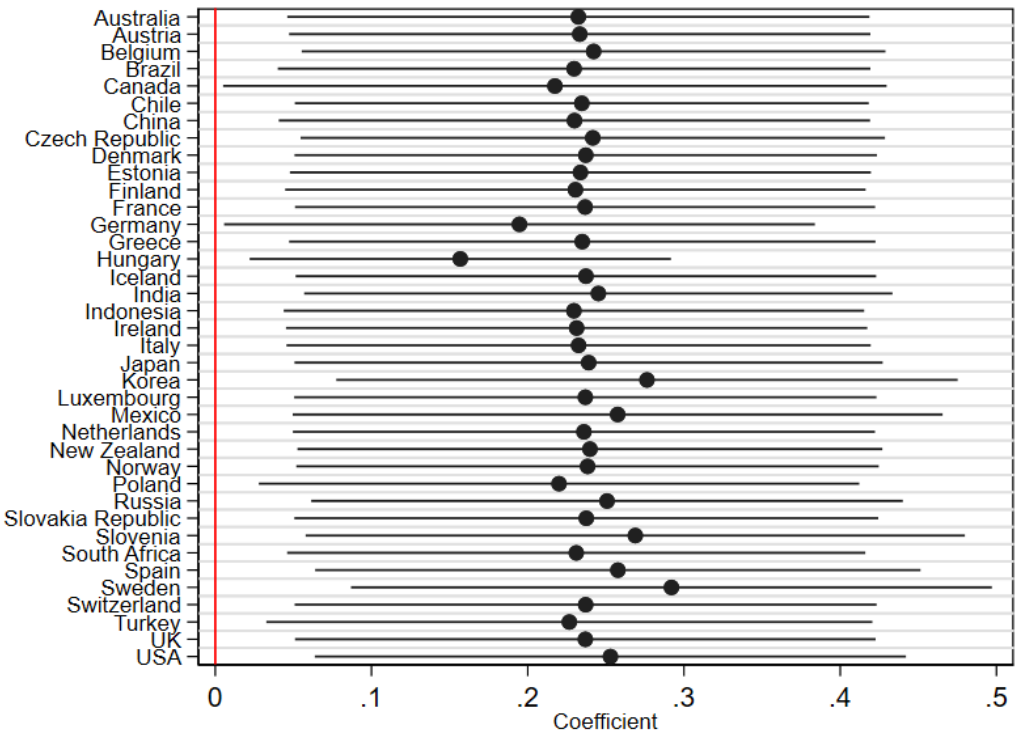
Notes: All regressions control for male dummy, unemployment dummy, 3-category education (Lower, i.e., primary education, Middle, i.e., secondary education, Upper, i.e., tertiary education), 10-class subjective income decile scale. Recession exposure is a dummy variable that takes value one if during adulthood the individual has experienced at least one year in which the real GDP per capita growth rate of its country is at least 10% below the previous year's growth rate. Environmental Quality exposure are PM2.5 and HAP exposure, respectively the adult years average of age-standardized disability-adjusted life-years lost per 100,000 people due to exposure to fine air particulate matter smaller than 2.5 micrometers (PM2.5) and due to exposure to household air pollution (HAP) from the use of household solid fuels. All regressions also control for survey, year of birth, country-by-year of interview fixed effects and country-age linear trends. Observations are weighted using survey weights. Standard errors are clustered by country-year. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table C9: Exposure to EPS sub-indices during adulthood on tax increase preferences

	Increase in tax if used to prevent pollution						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: Likert							
Policy exposure	-0.0875	0.118	-0.0495	0.133**	0.127***	0.131*	0.0596**
	(0.107)	(0.106)	(0.0494)	(0.0611)	(0.0428)	(0.0747)	(0.0226)
Mean Outcome	2.587	2.587	2.587	2.587	2.587	2.587	2.587
SD Outcome	0.859	0.859	0.859	0.859	0.859	0.859	0.859
Panel B: Binary (Agree & Strongly Agree)							
Policy exposure	-0.0474	0.0968**	-0.0274	0.0828**	0.0794***	0.0904**	0.0368***
	(0.0520)	(0.0482)	(0.0247)	(0.0325)	(0.0237)	(0.0388)	(0.0129)
Mean Outcome	0.581	0.581	0.581	0.581	0.581	0.581	0.581
SD Outcome	0.493	0.493	0.493	0.493	0.493	0.493	0.493
Panel C: Binary (Strongly Agree)							
Policy exposure	-0.00174	-0.0139	0.00483	0.0136	0.0282	0.0111	0.0144
	(0.0342)	(0.0453)	(0.0154)	(0.0278)	(0.0187)	(0.0387)	(0.00995)
Mean Outcome	0.133	0.133	0.133	0.133	0.133	0.133	0.133
SD Outcome	0.339	0.339	0.339	0.339	0.339	0.339	0.339
Policy	EPS	Market	Non-Market	Tax	CO2+Diesel	CO2	Diesel
Recession exposure	✓	✓	✓	✓	✓	✓	✓
Environmental Quality exposure	✓	✓	✓	✓	✓	✓	✓
Individual controls	✓	✓	✓	✓	✓	✓	✓
Survey FE	✓	✓	✓	✓	✓	✓	✓
Year of birth FE	✓	✓	✓	✓	✓	✓	✓
Country × Year of interview FE	✓	✓	✓	✓	✓	✓	✓
Country-age linear trends	✓	✓	✓	✓	✓	✓	✓
Mean Exposure	0.914	0.671	1.443	0.971	1.683	0.374	2.991
SD Exposure	0.689	0.545	1.206	0.776	1.202	1.217	1.918
N	20480	20480	20480	20480	20480	20480	20480
adj. R^2	0.090	0.090	0.090	0.090	0.090	0.090	0.090

Notes: All regressions control for male dummy, unemployment dummy, 3-category education (Lower, i.e., primary education, Middle, i.e., secondary education, Upper, i.e., tertiary education), 10-class subjective income decile scale. Recession exposure is a dummy variable that takes value one if during adulthood the individual has experienced at least one year in which the real GDP per capita growth rate of its country is at least 10% below the previous year's growth rate. Environmental Quality exposure are PM2.5 and HAP exposure, respectively the adult years average of age-standardized disability-adjusted life-years lost per 100,000 people due to exposure to fine air particulate matter smaller than 2.5 micrometers (PM2.5) and due to exposure to household air pollution (HAP) from the use of household solid fuels. All regressions also control for survey, year of birth, country-by-year of interview fixed effects and country-age linear trends. Observations are weighted using survey weights. Standard errors are clustered by country-year. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Figure C3: EPS exposure - leave-one-country-out



Notes: We plot the marginal effect of EPS exposure in adulthood in 38 different regressions where we exclude each time a different country as reported in the y-axis. The regression controls for male dummy, unemployment dummy, 3-category education (Lower, i.e., primary education, Middle, i.e., secondary education, Upper, i.e., tertiary education), 10-class subjective income decile scale, recession exposure, PM2.5 and HAP exposure. The regression also includes year of birth, country-by-year of interview, survey fixed effects, and country-by-age linear trends. Observations are weighted using survey weights. Standard errors are clustered at the country-year of interview level. Bins represent the 90% confidence intervals around point estimates.

C.3 Falsification Test

We conduct a falsification test to ensure that the effect of environmental policy exposure does not simply conflate changes in broader social and political attitudes and only affects environmental preferences. We use 13 alternative outcomes (detailed in Table C10) and regress each of these on our main measure of environmental policy stringency. In Table C11, we find no effect that is statistically distinguishable from zero, except for one at the 5% level, consistent with sampling variation attributable to multiple tests. This result allays concerns on general social and political changes conflating changes in environmental preferences.[16]

[16]We also note that another ideal placebo exercise would test the cross-cohort parallel trends underlying assumption (Acemoglu et al., 2024) by regressing pre-birth "exposure" to environmental policy stringency. Unfortunately, the short time series of EPS data starting in 1990 does not allow us to characterize pre-birth exposure in our estimation sample

Table C10: Description of other values used as outcomes in the falsification exercise

Variable	Description	Categories
A025	Respect and love for parents	1 = Respect if earned 2 = Neither 3 = Always respect
A026	Parents responsibilities to their children	1 = Parents have a life of their own and should not be asked to sacrifice their own well-being for the sake of their children 2 = Neither 3 = Parents' duty is to do their best for their children even at the expense of their own well-being
A048	Abortion when woman not married	0 = Disapprove 1 = Approve
A049	Abortion if not wanting more children	0 = Disapprove 1 = Approve
A165	Most people can be trusted	0 = Can´t be too careful 1 = Most people can be trusted
C001	Jobs scarce: Men should have more right to a job than women	1 = Disagree 2 = Neither 3 = Agree
C002	Jobs scarce: Employers should give priority to (nation) people than immigrants	1 = Disagree 2 = Neither 3 = Agree
C036	To develop talents you need to have a job	1 = Strongly agree 2 = Agree 3 = Neither agree nor disagree 4 = Disagree 5 = Strongly disagree
C037	Humiliating to receive money without having to work for it	1 = Strongly agree 2 = Agree 3 = Neither agree nor disagree 4 = Disagree 5 = Strongly disagree
C038	People who don´t work turn lazy	1 = Strongly agree 2 = Agree 3 = Neither agree nor disagree 4 = Disagree 5 = Strongly disagree
C039	Work is a duty towards society	1 = Strongly agree 2 = Agree 3 = Neither agree nor disagree 4 = Disagree 5 = Strongly disagree
D018	Child needs a home with father and mother	0 = Tend to disagree 1 = Tend to agree
E124	Respect for individual human rights nowadays	1 = There is a lot of respect for individual human rights 2 = There is some respect 3 = There is not much respect 4 = There is no respect at all

Table C11: Placebo test using other outcomes

	(1) A025	(2) A026	(3) A048	(4) A049	(5) A165	(6) C001	(7) C002	(8) C036	(9) C037	(10) C038	(11) C039	(12) D018	(13) E124
EPS exposure	0.0853	0.0505	0.0839	0.0173	-0.000162	0.0495**	0.0356	-0.0455	0.0602	-0.0216	0.000166	-0.0112	-0.0283
	(0.107)	(0.0977)	(0.0805)	(0.0712)	(0.0169)	(0.0249)	(0.0342)	(0.0545)	(0.0621)	(0.0588)	(0.0407)	(0.0273)	(0.0273)
Recession exposure	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
EQ exposure	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Individual controls	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Survey FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Year of birth FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Country × Year of interview FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Country-age linear trends	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Mean Outcome	2.404	2.451	0.569	0.550	0.325	1.616	2.332	2.204	2.714	2.384	2.368	0.797	2.294
SD Outcome	0.915	0.851	0.495	0.497	0.468	0.841	0.864	1.093	1.194	1.143	1.048	0.402	0.821
N	21602	21803	12514	12744	69889	71638	71332	38440	38276	50954	50862	31194	44904
adj. R^2	0.158	0.083	0.129	0.137	0.187	0.276	0.169	0.127	0.088	0.108	0.070	0.157	0.160

Notes: All regressions control for male dummy, unemployment dummy, 3-category education (Lower, i.e., primary education, Middle, i.e., secondary education, Upper, i.e., tertiary education), 10-class subjective income decile scale. Recession exposure is a dummy variable that takes value one if during adulthood the individual has experienced at least one year in which the real GDP per capita growth rate of its country is at least 10% below the previous year's growth rate. PM2.5 and HAP exposure are respectively the adult years average of age-standardized disability-adjusted life-years lost per 100,000 people due to exposure to fine air particulate matter smaller than 2.5 micrometers (PM2.5) and due to exposure to household air pollution (HAP) from the use of household solid fuels. All regressions also include year of birth, country-by-year of interview, survey fixed effects, and country-by-age linear trends. Observations are weighted using survey weights. Standard errors are clustered by country-year. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

C.4 Heterogeneity

Table C12: Environmental Policy Stringency and Individual Preferences for Government Intervention: Heterogeneous Effects

	Government should reduce pollution (Likert)						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
EPS exposure	0.263**	0.276**	0.326***	0.274**	0.277**	0.225**	0.210**
	(0.114)	(0.116)	(0.118)	(0.115)	(0.116)	(0.104)	(0.103)
Uninteracted Term 1	0.00127	0.0625**	-0.0387	0.0859	-0.0775	0.0567	-0.0883
	(0.0310)	(0.0303)	(0.0479)	(0.0647)	(0.0761)	(0.0443)	(0.0765)
Uninteracted Term 2			-0.110*				
			(0.0634)				
Interaction Term 1	0.0260	0.00406	-0.0380	0.0111	-0.0201	0.00451	0.0968
	(0.0249)	(0.0325)	(0.0362)	(0.0384)	(0.0611)	(0.0456)	(0.0597)
Interaction Term 2			-0.138***				
			(0.0467)				
Interaction Term	Male (=1)	Unemployed (=1)	Education (Middle=1; Upper=2)	Poor (=1)	Rich (=1)	Left-wing (=1)	Right-wing (=1)
Recession exposure controls	✓	✓	✓	✓	✓	✓	✓
EQ exposure controls	✓	✓	✓	✓	✓	✓	✓
Individual controls	✓	✓	✓	✓	✓	✓	✓
Survey FE	✓	✓	✓	✓	✓	✓	✓
Year of birth FE	✓	✓	✓	✓	✓	✓	✓
Country × Year of interview FE	✓	✓	✓	✓	✓	✓	✓
Country-age linear trends	✓	✓	✓	✓	✓	✓	✓
Mean Outcome	2.922	2.922	2.922	2.922	2.922	2.915	2.915
SD Outcome	0.900	0.900	0.900	0.900	0.900	0.900	0.900
N	16889	16889	16889	16889	16889	13719	13719
adj. R^2	0.132	0.132	0.133	0.132	0.132	0.121	0.121

Notes: The outcome variable ranges from 1 to 4 (with higher values reflecting stronger agreement with the statement "Government should reduce environmental pollution"). All regressions control for male dummy, unemployment dummy, 3-category education (Lower, i.e., primary education, Middle, i.e., secondary education, Upper, i.e., tertiary education), 10-class subjective income decile scale. Recession exposure is a dummy variable that takes value one if during adulthood the individual has experienced at least one year in which the real GDP per capita growth rate of its country is at least 10% below the previous year's growth rate. PM2.5 and HAP exposure are respectively the adult years average of age-standardized disability-adjusted life-years lost per 100,000 people due to exposure to fine air particulate matter smaller than 2.5 micrometers (PM2.5) and due to exposure to household air pollution (HAP) from the use of household solid fuels (PM2.5 and HAP exposure variables are inverted such that higher values imply higher quality). Poor is a dummy variable equal to one if individuals report belonging to the first two lowest deciles of the income scale, and zero otherwise. Rich is a dummy variable equal to one if individuals report belonging to the first two highest deciles of the income scale, and zero otherwise. Left-wing is a dummy variable equal to one if individuals report belonging to the first two steps in a ten-point political scale that goes from one (Left) to ten (Right). Right-wing is a dummy variable equal to one if individuals report belonging to the last two steps in a ten-point political scale that goes from one (Left) to ten (Right). All regressions also include year of birth, country-by-year of interview, survey fixed effects, and country-by-age linear trends. Observations are weighted using survey weights. Standard errors are clustered by country-year. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table C13: Environmental Policy Stringency and Individual Preferences for Government Action: Heterogeneous Effects

	Government should reduce pollution (Likert)			
	(1)	(2)	(3)	(4)
EPS exposure	-0.0126	-0.00539	0.197*	0.220*
	(0.0852)	(0.113)	(0.113)	(0.112)
Interest in politics				
Somewhat interested	0.0204			
	(0.0519)			
Not very interested	0.0486			
	(0.0522)			
Not at all interested	0.0312			
	(0.0494)			
Somewhat interested× EPS exposure	0.0788*			
	(0.0398)			
Not very interested × EPS exposure	0.121**			
	(0.0465)			
Not at all interested× EPS exposure	0.220***			
	(0.0533)			
Confidence: The Government (Baseline: A great deal)				
Quite a lot		0.0440		
		(0.0653)		
Not very much		0.0201		
		(0.0724)		
None at all		0.0270		
		(0.0815)		
Quite a lot × EPS exposure		0.0208		
		(0.0605)		
Not very much × EPS exposure		0.0610		
		(0.0780)		
None at all × EPS exposure		0.187**		
		(0.0761)		
Government vs People Responsibility				
Government responsibility			0.0722**	
			(0.0347)	
Government responsibility × EPS exposure			0.0674**	
			(0.0323)	
People responsibility				-0.0529
				(0.0620)
People responsibility × EPS exposure				0.0586
				(0.0501)
Recession exposure	✓	✓	✓	✓
EQ exposure	✓	✓	✓	✓
Individual controls	✓	✓	✓	✓
Survey FE	✓	✓	✓	✓
Year of birth FE	✓	✓	✓	✓
Country × Year of interview FE	✓	✓	✓	✓
Country-age linear trends	✓	✓	✓	✓
Mean Outcome	2.926	2.950	2.922	2.922
SD Outcome	0.896	0.892	0.900	0.900
N	15414	12102	16646	16646
adj. R^2	0.132	0.120	0.136	0.132

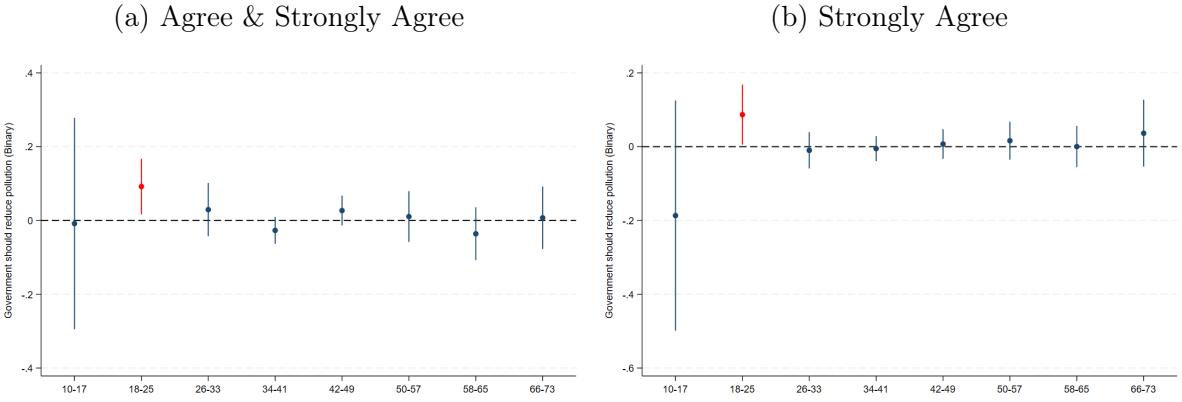
Notes: All regressions control for male dummy, unemployment dummy, 3-category education (Lower, i.e., primary education, Middle, i.e., secondary education, Upper, i.e., tertiary education), 10-class subjective income decile scale. All regressions control for survey, country-year of interview, year of birth FE. Recession exposure is a dummy variable that takes value one if during adulthood the individual has experienced at least one year in which the real GDP per capita growth rate of its country is at least 10% below the previous year's growth rate. PM2.5 and HAP exposure are respectively the adult years average of age-standardized disability-adjusted life-years lost per 100,000 people due to exposure to fine air particulate matter smaller than 2.5 micrometers (PM2.5) and due to exposure to household air pollution (HAP) from the use of household solid fuels. Observations are weighted using survey weights. Standard errors are clustered by country-year. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

D Additional results for formative age

In this Section, we test for the robustness of the baseline results in Section 5.2. Results are robust to the use of binary versions of the outcome (Appendix Figure D1), alternative definitions of formative age (Appendix Table D2), de-meaned and detrended exposure measures (Appendix Table D3), and measuring exposure over four-year age windows (Appendix Figure D2). In Section D.2, we describe an additional robustness test in which we account for country-level formative-age preferences as a measure of a second-order beliefs channel.

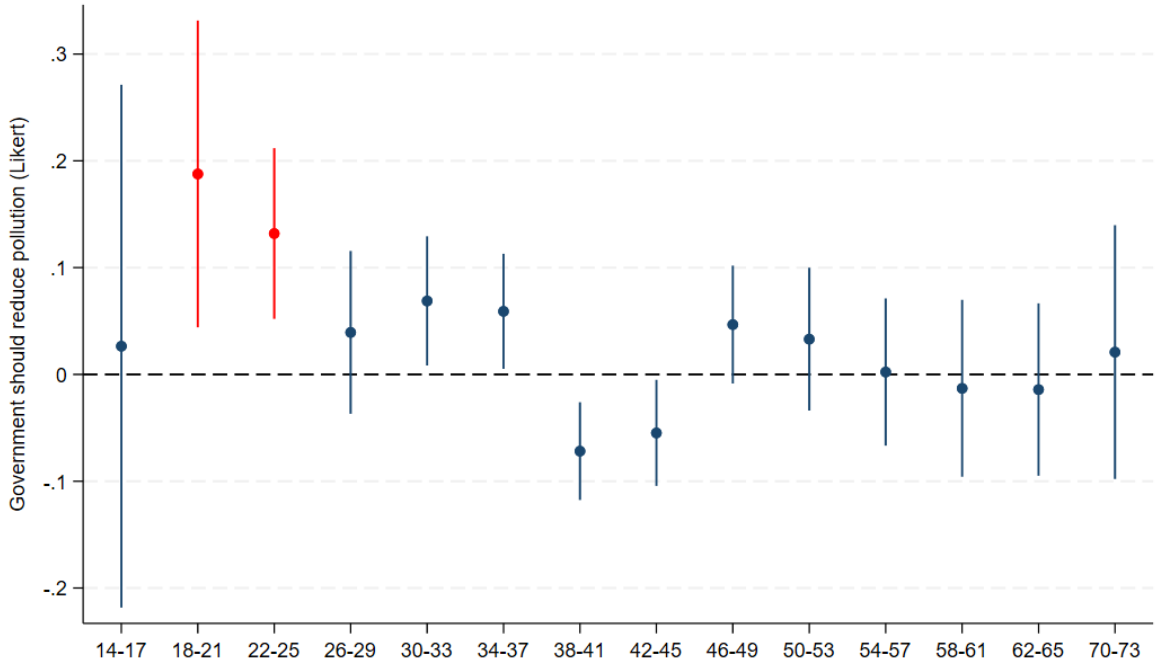
D.1 Robustness

Figure D1: Effect of EPS exposure in other age windows using binary outcomes



Notes: The figure plots the coefficients associated with the average level of EPS during each eight-year age window reported in the x-axis using the question "Government should reduce pollution" as outcome recoded in two different binary versions. In panel (a), the binary variable takes value one if respondents state either agree or strongly agree. In panel (b), the binary takes value one if respondents state strongly agree. The regression controls for male dummy, unemployment dummy, 3-category education (Lower, i.e., primary education, Middle, i.e., secondary education, Upper, i.e., tertiary education), 10-class subjective income decile scale, recession exposure, PM2.5 and HAP exposure. The regression also includes year of birth, country-by-year of interview, survey fixed effects, and country-by-age linear trends. Observations are weighted using survey weights. Standard errors are clustered at the country-year of interview level. Bins represent the 95% confidence intervals around point estimates. Tabular results are reported in Table D1.

Figure D2: Policy stringency effect on environmental preferences by four-year age window exposure



Notes: The figure plots the coefficients associated with the average level of EPS during each four-year age window reported in the x-axis using the question "Government should reduce pollution" as outcome in Likert scale. The regression controls for male dummy, unemployment dummy, 3-category education (Lower, i.e., primary education, Middle, i.e., secondary education, Upper, i.e., tertiary education), 10-class subjective income decile scale, recession exposure, PM2.5 and HAP exposure. The regression also includes year of birth, country-by-year of interview, survey fixed effects, and country-by-age linear trends. Observations are weighted using survey weights. Standard errors are clustered at the country-year of interview level. Bins represent the 95% confidence intervals around point estimates.

Table D1: Effect of Environmental Policy Stringency (EPS) exposure by age windows on environmental preferences

	Government should reduce pollution							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Likert								
EPS exposure	-0.243	0.207**	0.0173	-0.0417	0.0509	0.0399	-0.0747	0.0695
	(0.262)	(0.0914)	(0.0698)	(0.0302)	(0.0397)	(0.0539)	(0.0716)	(0.0793)
Age window	10-17	18-25	26-33	34-41	42-49	50-57	58-65	66-73
Mean outcome	2.962	2.912	2.899	2.901	2.879	2.884	2.944	3.000
SD Outcome	0.875	0.896	0.912	0.909	0.906	0.917	0.905	0.888
Panel B: Binary (Agree & Strongly Agree)								
EPS exposure	-0.00840	0.0920**	0.0293	-0.0270	0.0269	0.0105	-0.0361	0.00717
	(0.143)	(0.0377)	(0.0360)	(0.0182)	(0.0201)	(0.0343)	(0.0358)	(0.0423)
Age window	10-17	18-25	26-33	34-41	42-49	50-57	58-65	66-73
Mean outcome	0.698	0.667	0.654	0.654	0.647	0.649	0.679	0.708
SD Outcome	0.459	0.471	0.476	0.476	0.478	0.477	0.467	0.455
Panel C: Binary (Strongly Agree)								
EPS exposure	-0.187	0.0868**	-0.00985	-0.00535	0.00722	0.0161	0.000245	0.0365
	(0.156)	(0.0406)	(0.0246)	(0.0170)	(0.0201)	(0.0256)	(0.0279)	(0.0453)
Age window	10-17	18-25	26-33	34-41	42-49	50-57	58-65	66-73
Mean outcome	0.314	0.303	0.307	0.307	0.295	0.302	0.324	0.342
SD Outcome	0.464	0.460	0.461	0.461	0.456	0.459	0.468	0.475
Recession exposure	✓	✓	✓	✓	✓	✓	✓	✓
EQ exposure	✓	✓	✓	✓	✓	✓	✓	✓
Individual controls	✓	✓	✓	✓	✓	✓	✓	✓
Survey FE	✓	✓	✓	✓	✓	✓	✓	✓
Year of birth FE	✓	✓	✓	✓	✓	✓	✓	✓
Country × Year of interview FE	✓	✓	✓	✓	✓	✓	✓	✓
Country-age linear trends	✓	✓	✓	✓	✓	✓	✓	✓
N	6623	16889	20300	20211	17904	14522	11179	7788

Notes: All regressions control for male dummy, unemployment dummy, 3-category education (Lower, i.e., primary education, Middle, i.e., secondary education, Upper, i.e., tertiary education), 10-class subjective income decile scale. Recession exposure is a dummy variable that takes value one if during adulthood the individual has experienced at least one year in which the real GDP per capita growth rate of its country is at least 10% below the previous year's growth rate. PM2.5 and HAP exposure are respectively the adult years average of age-standardized disability-adjusted life-years lost per 100,000 people due to exposure to fine air particulate matter smaller than 2.5 micrometers (PM2.5) and due to exposure to household air pollution (HAP) from the use of household solid fuels. All regressions also include year of birth, country-by-year of interview, survey fixed effects, and country-by-age linear trends. Observations are weighted using survey weights. Standard errors are clustered by country-year. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table D2: Environmental Policy Stringency (EPS) exposure and environmental preferences. Alternative definitions of formative age.

	Government should reduce pollution (Likert)							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
EPS exposure	0.398	0.398	0.516**	0.287***	0.392*	0.318	0.198**	0.185*
	(0.246)	(0.246)	(0.242)	(0.0901)	(0.198)	(0.200)	(0.0819)	(0.0944)
Formative age	16-23	16-24	16-25	17-23	17-24	17-25	18-23	18-24
Recession exposure	✓	✓	✓	✓	✓	✓	✓	✓
EQ exposure	✓	✓	✓	✓	✓	✓	✓	✓
Individual controls	✓	✓	✓	✓	✓	✓	✓	✓
Survey FE	✓	✓	✓	✓	✓	✓	✓	✓
Year of birth FE	✓	✓	✓	✓	✓	✓	✓	✓
Country × Year of interview FE	✓	✓	✓	✓	✓	✓	✓	✓
Country-age linear trends	✓	✓	✓	✓	✓	✓	✓	✓
Mean Outcome	2.923	2.923	2.921	2.919	2.911	2.916	2.913	2.915
SD Outcome	0.880	0.880	0.881	0.890	0.886	0.884	0.896	0.893
Mean Exposure	0.964	0.964	0.979	0.859	0.982	0.991	0.879	0.885
SD Exposure	0.719	0.719	0.722	0.664	0.717	0.722	0.673	0.681
N	6064	6064	5919	14805	7418	7185	17215	16282
adj. R^2	0.126	0.126	0.127	0.135	0.136	0.132	0.132	0.135

Notes: All regressions control for male dummy, unemployment dummy, 3-category education (Lower, i.e., primary education, Middle, i.e., secondary education, Upper, i.e., tertiary education), 10-class subjective income decile scale. Recession exposure is a dummy variable that takes value one if during adulthood the individual has experienced at least one year in which the real GDP per capita growth rate of its country is at least 10% below the previous year's growth rate. PM2.5 and HAP exposure are respectively the adult years average of age-standardized disability-adjusted life-years lost per 100,000 people due to exposure to fine air particulate matter smaller than 2.5 micrometers (PM2.5) and due to exposure to household air pollution (HAP) from the use of household solid fuels. All regressions also include year of birth, country-by-year of interview, survey fixed effects, and country-by-age linear trends. Observations are weighted using survey weights. Standard errors are clustered by country-year. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table D3: EPS exposure during formative age on environmental preferences

	Government should reduce pollution (Likert)			
	(1)	(2)	(3)	(4)
Panel A: Baseline exposure measure				
EPS average exposure	0.190**	0.190**	0.206**	0.207**
	(0.0887)	(0.0880)	(0.0917)	(0.0914)
Mean Exposure	0.914	0.914	0.914	0.914
SD Exposure	0.681	0.681	0.681	0.681
Panel B: Year-demeaned relative exposure				
EPS average exposure	0.156	0.183*	0.156	0.183*
	(0.0982)	(0.102)	(0.0980)	(0.102)
Mean exposure	-0.230	-0.230	-0.230	-0.230
SD exposure	0.684	0.684	0.684	0.684
Panel C: Country-detrended exposure				
EPS average exposure	0.160*	0.181*	0.166*	0.190*
	(0.0948)	(0.0982)	(0.0952)	(0.0995)
Mean exposure	-0.112	-0.112	-0.112	-0.112
SD exposure	0.230	0.230	0.230	0.230
Recession exposure	✓	✓	✓	✓
PM2.5 exposure		✓		✓
HAP exposure			✓	✓
Individual controls	✓	✓	✓	✓
Survey FE	✓	✓	✓	✓
Year of birth FE	✓	✓	✓	✓
Country × Year of interview FE	✓	✓	✓	✓
Country-age linear trends	✓	✓	✓	✓
Mean Outcome	2.909	2.909	2.909	2.909
SD Outcome	0.896	0.896	0.896	0.896
N	16889	16889	16889	16889

Notes: The outcome variable ranges from 1 to 4 (with higher values reflecting stronger agreement with the statement "Government should reduce environmental pollution"). All regressions control for male dummy, unemployment dummy, 3-category education (Lower, i.e., primary education, Middle, i.e., secondary education, Upper, i.e., tertiary education), 10-class subjective income decile scale. Recession exposure is a dummy variable that takes value one if during adulthood the individual has experienced at least one year in which the real GDP per capita growth rate of its country is at least 10% below the previous year's growth rate. PM2.5 and HAP exposure are respectively the adult years average of age-standardized disability-adjusted life-years lost per 100,000 people due to exposure to fine air particulate matter smaller than 2.5 micrometers (PM2.5) and due to exposure to household air pollution (HAP) from the use of household solid fuels. All regressions also control for survey, year of birth, country-by-year of interview fixed effects and country-age linear trends. Observations are weighted using survey weights. Standard errors are clustered by country-year. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

D.2 Controlling for preferences during formative window

An alternative explanation to the role of policy stringency in shaping norms is that societal preferences could directly shape both policy stringency and individual norms, resulting in spurious correlation between past stringency and time-of-interview preferences influenced by historical norms. To allay this concern, we show that controlling for national average environmental policy preferences during the respondent's formative age does not meaningfully alter the effect of policy exposure (Table D4). If time of interview preferences were a consequence of conditions prior to the formative window or to social norms independent of stringency levels, then controlling for elicited societal preferences in the formative window should obviate the role of the EPS variable. The fact that controlling for societal preferences in the early part of one's formative years does not meaningfully change the direction or magnitude of our main findings is evidence for a causal interpretation for the role of stringency exposure.

The political success of environmental policy proposals may be informative to individuals of contemporaneous societal attitudes, reducing what is sometimes referred to as pluralistic ignorance. In this sense, second-order beliefs could indeed play an important role, serving as an intermediate variable that helps to explain how policy outcomes shape norms and influence an individual's policy support years later.

In Table D5, we estimate our main results controlling for policy preference only of one's peers at the earliest possible year of the formative age window. This allows us to rule out that cohort preferences are exogenous and formed prior to the formative age window, influencing policy outcomes shaping cohort norms independently.

There is some risk that controlling for peer support of environmental policy in the early part of the formative window, as well as societal support across all ages, could be a bad controls, absorbing some of the hypothesized effect we seek to measure. The fact that our results are robust to these controls contributes to our confidence in the formative age hypothesis and the role of policy stringency. The magnitude of the effect of policy stringency exposure, incorporating societal and peer policy attitudes as controls, is likely to be conservative relative to the true effect.

Table D4: EPS exposure during formative age on environmental preferences controlling for aggregate country-level preferences during formative age

	Government should reduce pollution		
	(1)	(2)	(3)
EPS average exposure	0.172*	0.0766*	0.0837*
	(0.101)	(0.0402)	(0.0423)
Country-level formative-age preferences	-0.0115	-0.0981*	0.0675
	(0.141)	(0.0586)	(0.0856)
Outcome variable	Likert	Strongly agree	Strongly agree & Agree
Recession exposure	✓	✓	✓
PM2.5 exposure	✓	✓	✓
HAP exposure	✓	✓	✓
Individual controls	✓	✓	✓
Survey FE	✓	✓	✓
Year of birth FE	✓	✓	✓
Country × Year of interview FE	✓	✓	✓
Country-age linear trends	✓	✓	✓
Mean Outcome	2.902	0.298	0.662
SD Outcome	0.896	0.457	0.473
N	13569	13569	13569
adj. R^2	0.133	0.082	0.128

Notes: The regression controls for the country-level average support for government action to reduce pollution measured during the earliest available year of the formative age for each individual. The outcome variable ranges from 1 to 4 (with higher values reflecting stronger agreement with the statement "Government should reduce environmental pollution") in column 1. The outcome variable is a dummy equal to 1 if the individual strongly agrees with the statement, and 0 otherwise, in column 2. The outcome variable is a dummy equal to 1 if the individual either strongly agrees or agrees with the statement, and 0 otherwise, in column 3. All regressions control for male dummy, unemployment dummy, 3-category education (Lower, i.e., primary education, Middle, i.e., secondary education, Upper, i.e., tertiary education), 10-class subjective income decile scale. Recession exposure is a dummy variable that takes value one if during adulthood the individual has experienced at least one year in which the real GDP per capita growth rate of its country is at least 10% below the previous year's growth rate. PM2.5 and HAP exposure are respectively the adult years average of age-standardized disability-adjusted life-years lost per 100,000 people due to exposure to fine air particulate matter smaller than 2.5 micrometers (PM2.5) and due to exposure to household air pollution (HAP) from the use of household solid fuels. All regressions also control for survey, year of birth, country-by-year of interview fixed effects and country-age linear trends. Observations are weighted using survey weights. Standard errors are clustered by country-year. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table D5: EPS exposure during formative age on environmental preferences controlling for aggregate country-level preferences of individuals of age 18 during formative age

	Government should reduce pollution		
	(1)	(2)	(3)
EPS average exposure	0.168*	0.0700*	0.0828**
	(0.0946)	(0.0387)	(0.0412)
Country-level formative-age preferences	0.124*	0.0564	0.0442
	(0.0693)	(0.0376)	(0.0342)
Outcome variable	Likert	Strongly agree	Strongly agree & Agree
Recession exposure	✓	✓	✓
PM2.5 exposure	✓	✓	✓
HAP exposure	✓	✓	✓
Individual controls	✓	✓	✓
Survey FE	✓	✓	✓
Year of birth FE	✓	✓	✓
Country × Year of interview FE	✓	✓	✓
Country-age linear trends	✓	✓	✓
Mean Outcome	2.902	0.299	0.661
SD Outcome	0.896	0.458	0.473
N	13377	13377	13377
adj. R^2	0.133	0.083	0.129

Notes: The regression controls for the country-level average support for government action to reduce pollution measured during the earliest available year of the formative age for each individual. The outcome variable ranges from 1 to 4 (with higher values reflecting stronger agreement with the statement "Government should reduce environmental pollution") in column 1. The outcome variable is a dummy equal to 1 if the individual strongly agrees with the statement, and 0 otherwise, in column 2. The outcome variable is a dummy equal to 1 if the individual either strongly agrees or agrees with the statement, and 0 otherwise, in column 3. All regressions control for male dummy, unemployment dummy, 3-category education (Lower, i.e., primary education, Middle, i.e., secondary education, Upper, i.e., tertiary education), 10-class subjective income decile scale. Recession exposure is a dummy variable that takes value one if during adulthood the individual has experienced at least one year in which the real GDP per capita growth rate of its country is at least 10% below the previous year's growth rate. PM2.5 and HAP exposure are respectively the adult years average of age-standardized disability-adjusted life-years lost per 100,000 people due to exposure to fine air particulate matter smaller than 2.5 micrometers (PM2.5) and due to exposure to household air pollution (HAP) from the use of household solid fuels. All regressions also control for survey, year of birth, country-by-year of interview fixed effects and country-age linear trends. Observations are weighted using survey weights. Standard errors are clustered by country-year. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

E Mechanisms

In this section, we report additional results related to Section 5.3 in the main text.

E.1 Environmental policy direction

We report the tabular results of the heterogeneous effect of EPS exposure conditional on the direction of the change in policy with respect to the contemporaneous level (Table E1). We also document that this heterogeneity is specific to the exposure during the formative age, and the effect is not statistically significant in any other age window to which the contemporaneous policy level is compared (Table E2).

Table E1: EPS exposure during formative age and contemporaneous policy level

	Government should reduce pollution (Likert)	
	(1)	(2)
EPS exposure	0.199**	0.274**
	(0.0824)	(0.108)
$\Delta policy < 0$	-0.258***	-0.234***
	(0.0416)	(0.0423)
EPS exposure $\times \Delta policy < 0$	0.229***	0.194***
	(0.0549)	(0.0520)
Exposure	Formative age	Adulthood
Individual controls	✓	✓
Recession exposure	✓	✓
EQ exposure	✓	✓
Survey FE	✓	✓
Year of birth FE	✓	✓
Country × Year of interview FE	✓	✓
Country-age linear trends	✓	✓
Mean Outcome	2.912	2.909
SD Outcome	0.896	0.896
Mean $\Delta policy$	0.126	0.135
SD $\Delta policy$	0.332	0.342
N	17070	16889
adj. R^2	0.133	0.133

Notes: All regressions control for male dummy, unemployment dummy, 3-category education (Lower, i.e., primary education, Middle, i.e., secondary education, Upper, i.e., tertiary education), 10-class subjective income decile scale. Recession exposure is a dummy variable that takes value one if, during her formative age, the individual has experienced at least one year in which the real GDP per capita growth rate of its country is at least 10% below the previous year's growth rate. EQ exposure includes PM2.5 and HAP exposure, which are respectively the formative age's average of age-standardized disability-adjusted life-years lost per 100,000 people due to exposure to fine air particulate matter smaller than 2.5 micrometers (PM2.5) and due to exposure to household air pollution (HAP) from the use of household solid fuels. Observations are weighted using survey weights. Standard errors are clustered by country-year. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table E2: EPS exposure across age windows and contemporaneous policy level

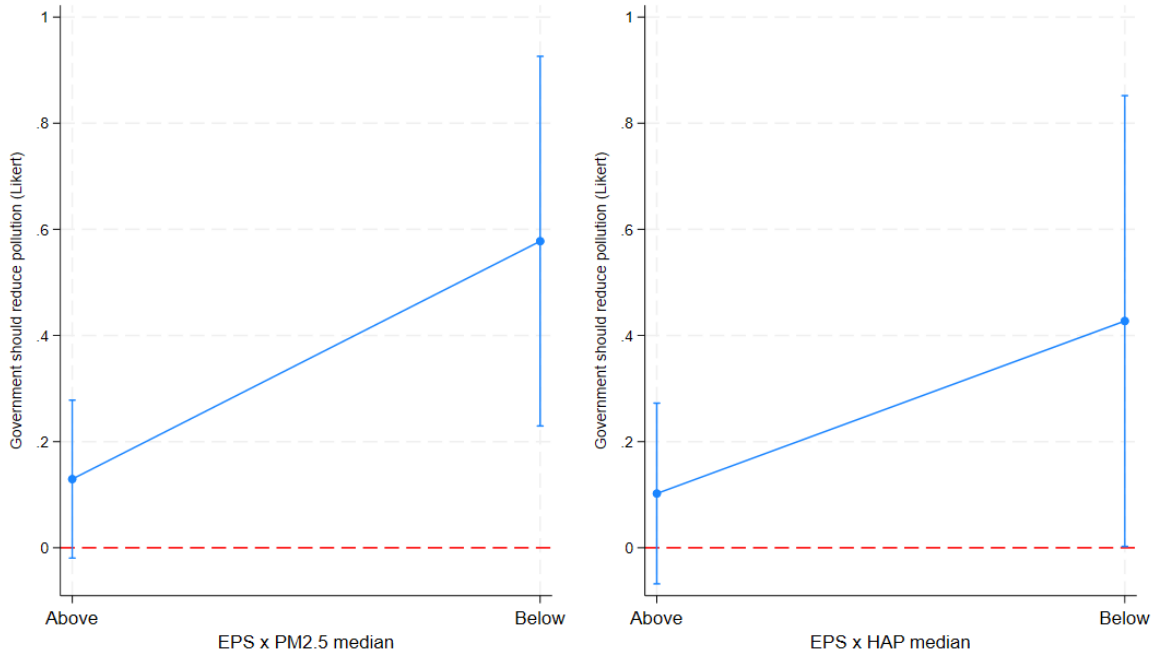
	Government should reduce pollution (Likert)						
	(1)	(2)	(3)	(4)	(5)	(6)	(7+)
EPS exposure	0.199**	0.0153	-0.0539	0.0576	0.0360	-0.0746	0.0582
	(0.0824)	(0.0721)	(0.0326)	(0.0403)	(0.0491)	(0.0732)	(0.0752)
$\Delta policy < 0$	-0.258***	-0.130	-0.0834*	0.0816	0.115	0.103	-0.110
	(0.0416)	(0.143)	(0.0468)	(0.0613)	(0.0959)	(0.0926)	(0.110)
EPS exposure $\times \Delta policy < 0$	0.229***	0.124	0.117*	-0.0948	-0.0819	-0.0814	0.122
	(0.0549)	(0.140)	(0.0654)	(0.0575)	(0.123)	(0.0821)	(0.119)
Age window	18-25	26-33	34-41	42-49	50-57	58-65	66-73
Individual controls	✓	✓	✓	✓	✓	✓	✓
Recession exposure	✓	✓	✓	✓	✓	✓	✓
EQ exposure	✓	✓	✓	✓	✓	✓	✓
Survey FE	✓	✓	✓	✓	✓	✓	✓
Year of birth FE	✓	✓	✓	✓	✓	✓	✓
Country × Year of interview FE	✓	✓	✓	✓	✓	✓	✓
Country-age linear trends	✓	✓	✓	✓	✓	✓	✓
Mean Outcome	2.912	2.899	2.901	2.879	2.884	2.944	3.000
SD Outcome	0.896	0.912	0.909	0.906	0.917	0.905	0.888
Mean $\Delta policy$	0.132	0.123	0.111	0.106	0.090	0.094	0.096
SD $\Delta policy$	0.338	0.328	0.314	0.308	0.286	0.292	0.295
N	17070	20300	20211	17904	14522	11179	7788
adj. R^2	0.133	0.135	0.142	0.162	0.163	0.171	0.168

Notes: All regressions control for male dummy, unemployment dummy, 3-category education (Lower, i.e., primary education, Middle, i.e., secondary education, Upper, i.e., tertiary education), 10-class subjective income decile scale. Recession exposure is a dummy variable that takes value one if during adulthood the individual has experienced at least one year in which the real GDP per capita growth rate of its country is at least 10% below the previous year's growth rate. EQ exposure includes PM2.5 and HAP exposure, which are respectively the adult years average of age-standardized disability-adjusted life-years lost per 100,000 people due to exposure to fine air particulate matter smaller than 2.5 micrometers (PM2.5) and due to exposure to household air pollution (HAP) from the use of household solid fuels. Observations are weighted using survey weights. Standard errors are clustered by country-year. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

E.2 Environmental Quality exposure

We also explore whether the effect of more stringent environmental policies on individual support is conditional on environmental quality. We show that the effect of EPS exposure is stronger when environmental quality exposure is below the median. We probe the robustness of this result when using exposure over adult life (Figure E1 and tabular results in Table E3). The uninteracted term of exposure to environmental quality below the median is negative and in certain cases not significant, suggesting that environmental conditions downgrade individual support of the government's role in reducing pollution and we are not identifying a "poor environmental quality effect". Rather, we are identifying the effect of experience of both poor conditions and more stringent policies meant to address them. This combination induces support for government pollution actions. Similar results are also obtained when estimating the effect of EPS exposure splitting the sample between above and below median of environmental quality (Table E4). The result underlines the importance of stringent policies targeting poor environmental conditions to sustain support for government to reduce pollution.

Figure E1: Policy stringency effect on environmental preferences by environmental quality exposure during adulthood



Notes: The figure shows the marginal effect of EPS exposure in adulthood when exposure to environmental quality measures in the same period is above or below the median in the sample on the question "Government should reduce pollution" as outcome in Likert scale. The regression controls for male dummy, unemployment dummy, 3-category education (Lower, i.e., primary education, Middle, i.e., secondary education, Upper, i.e., tertiary education), 10-class subjective income decile scale, recession exposure. The regression also includes year of birth, country-by-year of interview, survey fixed effects, and country-by-age linear trends. Observations are weighted using survey weights. Standard errors are clustered at the country-year of interview level. Bins represent the 95% confidence intervals around point estimates.

Table E3: EPS and environmental quality exposure during adulthood

	Government should reduce pollution (Likert)	
	(1)	(2)
EPS exposure	0.129*	0.102
	(0.0759)	(0.0869)
Below median EQ	-0.468***	-0.374
	(0.156)	(0.245)
EPS exposure × Below median EQ	0.448***	0.325
	(0.132)	(0.220)
EQ variable	PM2.5	HAP
Recession exposure	✓	✓
Individual controls	✓	✓
Survey FE	✓	✓
Year of birth FE	✓	✓
Country × Year of interview FE	✓	✓
Country-age linear trends	✓	✓
N	16889	16889
adj. R^2	0.133	0.132

Notes: All regressions control for male dummy, unemployment dummy, 3-category education (Lower, i.e., primary education, Middle, i.e., secondary education, Upper, i.e., tertiary education), 10-class subjective income decile scale. Recession exposure is a dummy variable that takes value one if during adulthood the individual has experienced at least one year in which the real GDP per capita growth rate of its country is at least 10% below the previous year's growth rate. Below median EQ is a binary variable equal to one if individuals have been exposed to an environmental quality measure (PM2.5 or HAP) that is strictly below the median exposure value. All regressions also control for survey, year of birth, country-by-year of interview fixed effects and country-age linear trends. Observations are weighted using survey weights. Standard errors are clustered by country-year. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table E4: EPS and EQ exposure during adulthood

	Government should reduce pollution (Likert)			
	(1)	(2)	(3)	(4)
EPS exposure	0.516**	0.170	0.446**	0.0858
	(0.212)	(0.108)	(0.210)	(0.125)
Environmental Quality exposure	-0.0658	-0.127*	-0.0790	0.0152
	(0.0525)	(0.0673)	(0.0586)	(0.0745)
Environmental Quality variable	PM2.5		HAP	
Sample	Below	Above	Below	Above
Recession exposure	✓	✓	✓	✓
Individual controls	✓	✓	✓	✓
Survey FE	✓	✓	✓	✓
Year of birth FE	✓	✓	✓	✓
Country × Year of interview FE	✓	✓	✓	✓
Country-age linear trends	✓	✓	✓	✓
Mean outcome	2.962	2.878	3.005	2.826
SD outcome	0.877	0.923	0.881	0.912
N	8885	8004	9013	7876
adj. R^2	0.095	0.172	0.100	0.159

Notes: All regressions control for male dummy, unemployment dummy, 3-category education (Lower, i.e., primary education, Middle, i.e., secondary education, Upper, i.e., tertiary education), 10-class subjective income decile scale. Recession exposure is a dummy variable that takes value one if during adulthood the individual has experienced at least one year in which the real GDP per capita growth rate of its country is at least 10% below the previous year's growth rate. PM2.5 and HAP exposure are respectively adult years average of age-standardized disability-adjusted life-years lost per 100,000 people due to exposure to fine air particulate matter smaller than 2.5 micrometers (PM2.5) and due to exposure to household air pollution (HAP) from the use of household solid fuels (PM2.5 and HAP exposure variables are inverted such that higher values imply higher quality and divided by 1000 to improve the readability of coefficients). All regressions also include year of birth, country-by-year of interview, survey fixed effects, and country-by-age linear trends. Observations are weighted using survey weights. Standard errors are clustered by country-year. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table E5: EPS and environmental quality exposure during formative age

	Government should reduce pollution (Likert)	
	(1)	(2)
EPS exposure	0.0810	0.0833
	(0.0650)	(0.0690)
Below median EQ	-0.246*	-0.267*
	(0.156)	(0.245)
EPS exposure × Below median EQ	0.293***	0.292***
	(0.109)	(0.144)
EQ variable	PM2.5	HAP
Recession exposure	✓	✓
Individual controls	✓	✓
Survey FE	✓	✓
Year of birth FE	✓	✓
Country × Year of interview FE	✓	✓
Country-age linear trends	✓	✓
N	17070	17070
adj. R^2	0.133	0.132

Notes: All regressions control for male dummy, unemployment dummy, 3-category education (Lower, i.e., primary education, Middle, i.e., secondary education, Upper, i.e., tertiary education), 10-class subjective income decile scale. Recession exposure is a dummy variable that takes value one if during adulthood the individual has experienced at least one year in which the real GDP per capita growth rate of its country is at least 10% below the previous year's growth rate. Below median EQ is a binary variable equal to one if individuals have been exposed to an environmental quality measure (PM2.5 or HAP) that is strictly below the median exposure value. All regressions also control for survey, year of birth, country-by-year of interview fixed effects and country-age linear trends. Observations are weighted using survey weights. Standard errors are clustered by country-year. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

E.3 Endogenous Learning

Exposure to policies can influence norms, as well as information about welfare-relevant conditions under the policy. For example, if my country has spent generations under a constant policy regime with respect to environmental matters, then it can be difficult for me to directly know the costs and benefits of living under a different policy regime. I could listen to the projections of policy theorists or study conditions in other countries, but only living through and contrasting two or more policy regimes will provide me with the most direct form of evidence to inform my subsequent policy preferences.

This form of direct learning may contribute to an endogenous path of support of green policies. But in observational data, the role of norms and learning is difficult to disentangle. We take two approaches to the measurement and control of endogenous learning in this paper: (1) to understand the assumptions necessary for endogenous learning to explain the data that we observe, and (2) to include as a control variable the volatility of policy stringency, a proxy for the direct learning opportunities that individuals experience.

An endogenous learning channel requires some strong assumptions to explain our empirical findings. Our comparison is of two individuals from different birth cohorts interviewed at the same place and time. For learning to play a role, we must first assume that birth cohort variation produces different information sets, in contrast to the standard assumption that any two adults have access to the same information about policy efficacy, regardless of age.

For learning to play an important role in the *positive* endogeneity we empirically ob-

serve, a second assumption is needed: individuals are systematically pessimistic about policy stringency, and that this pessimism is corrected with experience. Imagine that some individuals have experience with stringency while others do not. If everyone has perfect foresight of a policy proposal's impacts, then, absent a role for heterogeneous norms, the observed coefficient on policy exposure would be zero. We do not find this. Now instead imagine that inexperienced individuals tend to be too optimistic about environmental policies ex ante. In that case, experience with policies would leave an individual with systematically weaker support, and we would document negative endogeneity from exposure, i.e., a coefficient less than one. We do not find this. We find a positive coefficient. For the positive endogeneity we observe, for learning to play a role, past experience must have shaped beliefs in a manner that makes government action on pollution more desirable, implying that priors about policy desirability are pessimistic.

Including an individual's experienced policy volatility in our analysis allows us to introduce a proxy for the direct learning, as discussed in Section 2.2. Respondents who have experienced more variability of policy levels would be better equipped to contrast levels and therefore better informed than individuals who experienced relatively stable policy levels. In a horse race between the first and the second moment of the policy exposure distribution, coefficients on the standard deviation, as well as its interaction with the average, suggest a modest impact of learning that, when controlled for, leaves the large effect of average exposure unaltered and suggests a robust role for norms (Table E6).

Table E6: Environmental Policy Stringency mean and standard deviation effects on environmental preferences.

	Government should reduce pollution (Likert)					
	(1)	(2)	(3)	(4)	(5)	(6)
EPS mean exposure	0.236**	0.258*	0.298**	0.207**	0.209**	0.261***
	(0.111)	(0.136)	(0.135)	(0.0914)	(0.0865)	(0.0873)
EPS SD exposure		0.0493	0.204		-0.00666	0.238
		(0.105)	(0.249)		(0.0693)	(0.162)
EPS mean × SD exposure			-0.0854			-0.166
			(0.128)			(0.0999)
Exposure	Adulthood (18-interview year)			Formative age (18-25)		
Recession exposure	✓	✓	✓	✓	✓	✓
EQ exposure	✓	✓	✓	✓	✓	✓
Individual controls	✓	✓	✓	✓	✓	✓
Survey FE	✓	✓	✓	✓	✓	✓
Year of birth FE	✓	✓	✓	✓	✓	✓
Country × Year of interview FE	✓	✓	✓	✓	✓	✓
Country-age linear trends	✓	✓	✓	✓	✓	✓
N	16889	16889	16889	16889	16889	16889
adj. R^2	0.132	0.132	0.132	0.133	0.133	0.133

Notes: Columns (1)-(3) regress the 4-point Likert scale question about "Government should reduce environmental pollution" on the mean and standard deviation of EPS during adulthood (from the year of age 18 to the year of interview) and their interaction term. Columns (4)-(6) replicate the same specification using mean and standard deviation of EPS during the formative age (from the year of age 18 to the year of age 25). All regressions control for male dummy, unemployment dummy, 3-category education (Lower, i.e., primary education, Middle, i.e., secondary education, Upper, i.e., tertiary education), 10-class subjective income decile scale. All regressions control for survey, country-year of interview, year of birth FE. Observations are weighted using survey weights. Standard errors are clustered by country-year. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

F Counterfactual Details

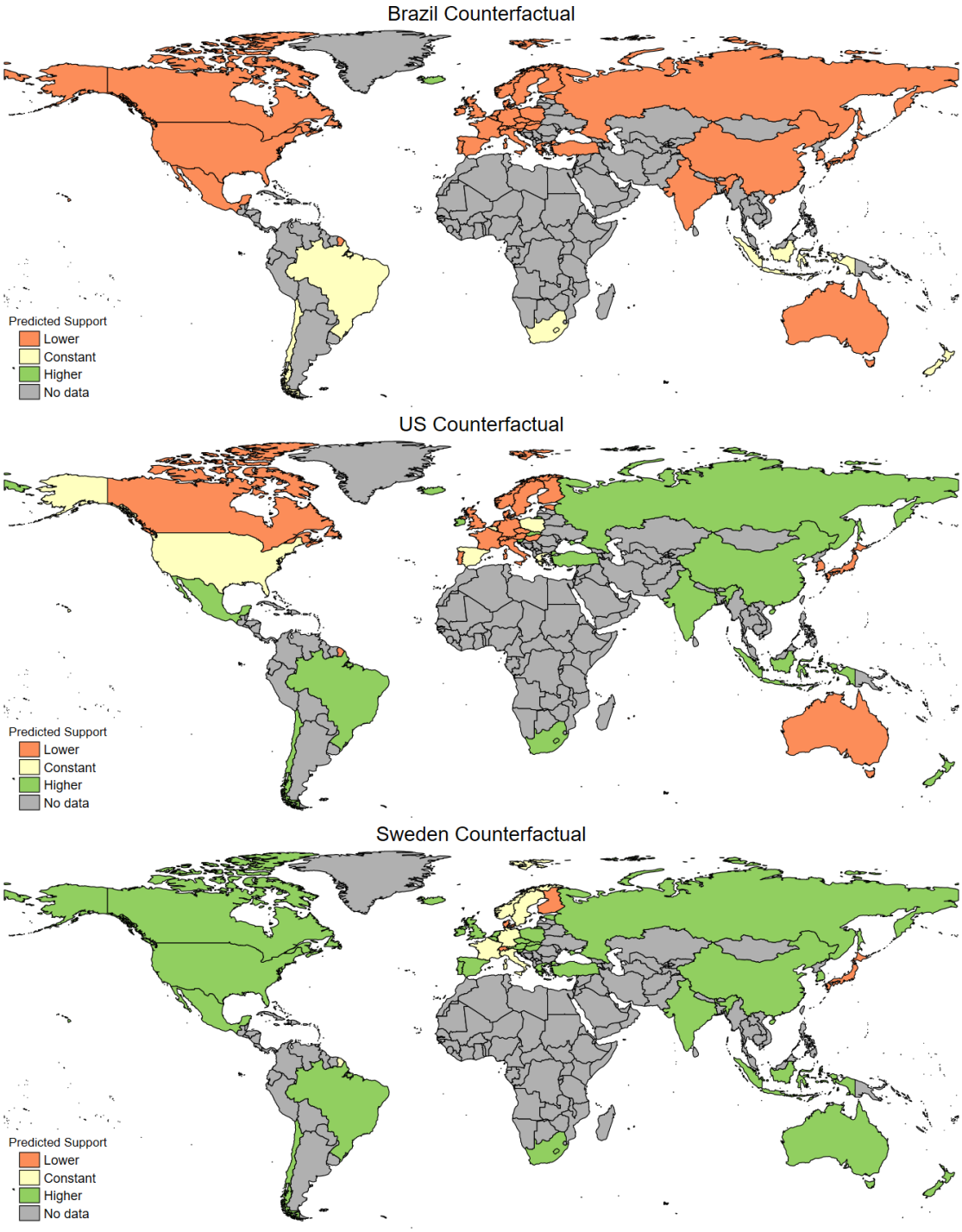
We use our reduced-form estimates to estimate predicted policy preferences under different policy counterfactual scenarios. We answer the following question: if individuals were exposed to another counterfactual policy level, all else equal, how would the change in policy preferences predicted by our model be? We choose as counterfactuals three countries with notable historically different policy stringency levels: Brazil (cross-cohort average policy stringency exposure is 0.46), United States (average policy exposure 1.75, close to the cross-country average 1.74), and Sweden (average policy exposure 2.57).

We take the estimated marginal effect of EPS exposure on the probability to *strongly agree* with the statement "Government should reduce pollution" in column 4, Panel B, Table C6. We choose this approach to define environmental preferences in the most possible conservative manner and to facilitate the interpretation of our results in terms of shares of respondents that strongly support environmental policies.

Using the estimated coefficient on EPS exposure, we predict the counterfactual change in green preferences using the difference between the observed level of policy stringency and the counterfactual policy level in each of the three countries. We take the average across years of interview and report in Figure 4 the model-predicted shares of respondents with *strong* green preferences under each counterfactual.

Figure F1 displays world maps under each counterfactual policy experiment showing which countries have a predicted support for green policies that is lower, constant, or higher than observed in the data.

Figure F1: Predicted support for government reduction in pollution using Brazil, US, and Sweden counterfactual policy stringency exposures



Notes: Each map shows the predicted change in the share of respondents that strongly agree with the statement "Government should reduce pollution", using, respectively, Brazil, US and Sweden's levels of birth-cohorts exposure to past environmental policy stringency, and using the coefficient on EPS in column 4, Panel B, Table C6.

Price, demand and hypothetical thinking

N. R. Vora*

Abstract

When people are asked to make policy choices that affect a good's price – for example, a tax to reduce a negative externality – we traditionally assume that they accurately perceive quantities demanded conditional on price. But failures to consider substitutes and other adaptations to a permanent price shock can result, ex ante, in a perception of inelasticity and policy ineffectiveness. In past surveys, perceptions of ineffectiveness are among the main reasons voters do not support a carbon tax, believing it would fail to cut emissions unless the revenue were spent on offsetting projects. I implement an incentivized survey experiment to test whether increasing time and lowering costs for hypothetical thinking increase subjective perception of responsiveness of demand to price. Results from this design provide suggestive evidence that time and effort costs of hypothetical thinking play a role in shaping perceptions of demand responsiveness.

Keywords: Externalities, carbon tax, political economy, cognitive frictions.

JEL Classification: H23, D91, Q52, Q58

*UC Santa Barbara, Department of Economics, 2127 North Hall University of California, Santa Barbara, CA 93106.

1 Introduction

Forecasting the impact of a policy proposal is a difficult task, even for experts. Yet we implicitly expect people - as voters, politicians, or influencers - to arrive at a fair estimation of these impacts. The voter, in particular, is a concerning case. Votes directly or indirectly form policy but have little material incentive to think carefully, given the extremely small likelihood of affecting the outcome of a ballot measure or election (Downs, 1957). So, we might expect many voters to assess a policy proposal hastily, given the near absence of economic incentives of getting an assessment right, and the real costs in time and effort of attempting to.

One might think, 'So what?' Haste may introduce noise, but barring any systematic mistake, i.e., a directional bias, then on average, the "wisdom of the crowd" effect should imply that a lot of hastily made independent choices will aggregate into a roughly appropriate collective outcome. In this paper, I show that hasty decisions appear to be produce systematically biased assessments that underlie the perception of an important policy proposal's effectiveness.

Using an incentivized survey experiment, I first document that people largely believe aggregate demand for gasoline is nearly inelastic, in contrast to the conclusions of empirical studies. This perception – in essence that demand will not fall meaningfully in response to an increase in price – is relevant to the assessment of proposals that would tax fossil fuels to reduce their use. These corrective taxes have long been supported by economists as an ideal policy tool to tackle issues like climate change, but the usage and magnitude of these taxes have long been below levels deemed optimal by economists

and other experts. The reason is public opposition, in many cases driven by beliefs that corrective taxes on fossil fuels would fail to directly reduce emissions.

Importantly, if beliefs about elasticity are not firmly held but instead conditional on circumstances, such as the the time and effort spent thinking, as I show, then an attitude toward a policy resulting from this "belief" may not represent preferences in a classical sense. I show that longer consideration time and inducement to hypothetical thinking tend to increase perceived elasticity of demand response. I offer a model of how a realistically slow and imperfect recall process would result in biased conditional demand assessment, and not merely added noise.

Past research finds evidence of individual difficulty in assessing conditional demand, in policy and consumer settings. Ineffectiveness is among the most common reasons people give for their opposition or skepticism to carbon taxes, congestion charges, and price-based instruments generally (Baranzini and Carattini, 2017; Baranzini et al., 2017; Carattini et al., 2018; Metcalf, 2023; Douenne and Fabre, 2022). This is despite strong evidence to the contrary that these tools are effective (Andersson, 2019; Metcalf, 2021; Colmer et al., 2022).[1] Despite the evidence, even among those most concerned about climate change, taxes are often perceived effective only as a means to obtaining revenue to be spent on green investments and support is contingent on this use (Baranzini and

[1]While economists' estimates of price elasticity vary, short-run elasticity estimates appear to have converged to around -0.4 (Coglianese et al., 2017; Li et al., 2014; Tiezzi and Verde, 2014). More relevant to this paper, longer-run price elasticity and tax elasticity estimates are closer to between -0.8 and -1 (Hausman and Newey, 1995; Dieler et al., 2015; Li et al., 2014). Studies estimating tax elasticity from policy shocks in British Columbia and Sweden of higher magnitude than -1 (Anderson et al., 2023; Rivers and Schaufele, 2015). Estimates from older studies may not reflect consumer choices available today. The much higher responsiveness to taxes has been attributed to a greater behavioral response to a persistent shock, higher salience, and tax aversion.

Carattini, 2017; Kallbekken and Aasen, 2010).[2] Experience with pricing policies, particularly for congestion charges, appears to improve support (Eliasson and Jonsson, 2011; Janusch et al., 2021; Kallbekken and Sælen, 2011). In Sweden support more than doubled from before to after the implementation of a road tax policy.

Ex ante reasoning often contrasts with understanding gained from experience (Aragones et al., 2005; Esponda et al., 2024). Consider the case of estimating a demand response to a price change in gasoline. People uninformed about elasticity statistics would need to consider number of adaptive responses to the price change. The set of considered responses may not be comprehensive, i.e. a consideration set (Caplin et al., 2011; Hoyer, 1984; Roberts and Lattin, 1997; Eliaz and Spiegler, 2011). Its size would be a function of time and effort used to think hypothetically. Mental availability of gasoline-reduction strategies is limited by imperfect and slow recall, but providing time and aid can improve this recollection process. One such aid to memory is a phenomenon called rehearsal, where information recalled in the past is more likely to be recalled again and used to inform decision making (Mullainathan, 2002).

I incentivize voters to guess the demand response to a real gasoline policy change that occurred in the past and experimentally manipulate decision time and hypothetical thinking using time limits and introspection writing prompts, respectively. I hypothesize that subjects who use less time to make their guesses are likelier to perceive demand as relatively inelastic. Introspection writing prompts, in which subjects are asked to imagine themselves in a world with higher gasoline prices, how behavior might different,

[2]Opposition among climate concerned voters to taxes without earmarking for green projects can result in a political impasse, as other voters may be amenable to taxing externalities if the revenue were recycled into lump sum dividends or used to offset cuts to other taxes.

and what kind of advice they would provide others, could induce a rehearsal of adaptive options, resulting in higher perceived elasticity. In the model I propose, such a rehearsal process makes it easier for subjects to populate their consideration set when then asked to provide their incentivized demand response guess to the real policy-induced price change.

The strongest and most convincing results exclude a small minority of subjects who violate the law of demand by providing a positive demand response guess to a price increase. As I discuss later, it is highly likely that these subjects failed to understand the task, and I am able to show that upon analyzing the responses of the vast majority who obey the law of demand, additional time and introspection prompts have clear, strong effects.

2 Conceptual Framework

For a more formal treatment of this framework, see Appendix A.

Voters are implicitly asked to estimate the policy's effectiveness. To support a policy, they must weigh costs and benefits. For externality pricing, the benefit is the reduction in a negative byproduct of consumption. A voter supports the policy when the benefit is sufficiently high, but determining the size of the benefit requires the voter to become an assessor. An assessor, broadly, is being asked to accurately estimate the aggregate demand response to a higher price.[3]

[3]Consumers adapt to a new price regime by remembering consumption options that are different than habits in the previous price regime, exploring and innovating new options, and learning from experience and peers about best practices. This ex-post adaptation process can happen slowly and is strongly incentivized. Voters, in contrast, make quick, low-stakes decisions, as they are unlikely to be pivotal and cannot justify a large investment of time and effort on an extremely small chance of influencing the price regime.

But this is a difficult quantity for a voter to estimate. In theory, a price change via a policy τ will result in income and substitution effects for each consumer. The choices of an optimizing consumer i at the original price p_0 may no longer be optimal at the new price $p_1 = p_0 + \tau$. Consumer i can choose among a set of behavioral responses, some combination of which could improve the consumer's utility. The optimizing consumer has perfect information about these behavioral adaptations and selects the best bundle of new behaviors from the power set of these adaptations.

Adaptations to a price increase can include reductions to certain usages or substitutions. For example, a gasoline tax results in a comprehensive list of potential adaptations: {Switch to an electric car, switch to a hybrid car, work from home, relocate, carpool, use public transportation, use a bicycle, drop low-value trips, combine trips, time driving to avoid congestion, reduce vehicle speed, ...} can be represented as a list S of adaptations $\{a_1, a_2, a_3, ...\}$. Ultimately we envision the fully rational consumer as choosing from the power set of the comprehensive list of feasible adaptations.

The utility maximizing choice α_i for consumer i, from the power set of adaptations, corresponds to a change in gasoline quantity demanded $r_i(\alpha_i)$. A fully rational voter estimating the effectiveness of the policy must find the aggregate demand response, a sum over all consumers: $\sum_{i=1}^{N} r_i(\alpha_i)$.

But a voter with limited hypothetical thinking, which could take the form of imperfect recall or failures of imagination, would use a consideration set of adaptations and not the comprehensive list. When this consideration set $\tilde{S}$ is a proper subset of the comprehensive list S, its use would result in an estimated demand response to a price increase, $\sum_{i=1}^{N} r_i(\alpha_i(\tilde{S}))$ of an equal or lower magnitude than the estimated demand response,

$\sum_{i=1}^{N} r_i(\alpha_i(S))$, from the comprehensive list.

The argument of this paper is that consumers are much better at using the comprehensive list S to make consumer choices than are voters, who are prone to use a consideration set $\tilde{S}$ to gauge the effectiveness of the policy proposal. While voters are often consumers, their ex-ante judgments as voters about their own and others' consumption behavior aren't accurate.

For intuition, if a voter in haste considers only the adaptation of switching to hybrid vehicles, but fails to consider the possibility of consumers working from home or relocating closer to frequent destinations, then reductions will come only from consumers for whom switching to a hybrid is feasible and preferable. Without considering the existence of the other adaptations, the voter's subjective impression of the aggregate demand response, and as a result the effectiveness of the policy, will be smaller. Suppose we are comparing demand response guesses under two consideration sets, $\tilde{S}_j$ and a larger set $\tilde{S}_{j+m}$, where $\tilde{S}_{j+m} \subset \tilde{S}_{j+m}$.

Under Assumption 1 in Appendix A: $r_i(\tilde{S}_{j+m}) \geq r_i(\tilde{S}_j)$ and, when $\alpha_i(\tilde{S}_j) \subset \alpha_i(\tilde{S}_{j+m})$, the inequality is strict: $r_i(\tilde{S}_{j+m}) > r_i(\tilde{S}_j)$. It follows that the aggregate assessment $\sum_{i=1}^{N} r_i(\tilde{S}_{j+m}) > \sum_{i=1}^{N} r_i(\tilde{S}_j)$.

I hypothesize that the process of populating the consideration set with adaptation options from which to choose is a costly one in time and effort. Therefore a voter who assesses a policy proposal quickly or superficially may misjudge its effectiveness. The number of options in the consideration set is likely increasing in consideration time, but this process is unlikely to be linear. Consideration time is likely to have diminishing returns, sometimes referred to as a "fishing out effect" wherein it grows more difficult to

contribute additional elements to the consideration set. In addition, as the consideration set grows larger with consideration time, it may begin to strain the limits of working memory. Using exogenous variation in consideration time, as well as natural endogenous variation, will be key elements of the research design.

Consideration effort is affected by the mental availability of adaptive responses. Rehearsal can reduce these effort costs (Mullainathan, 2002). Because information recalled in the past is easier to retrieve, the design will attempt to induce rehearsal of adaptive options to improve the ease and quality of hypothetical thinking.

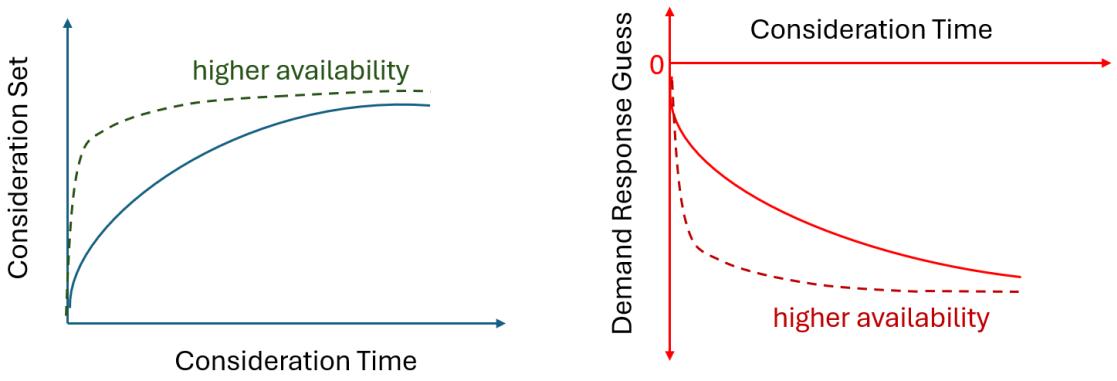


Figure 1: As described above and detailed in Appendix A, a hypothesized process of considering adaptation options to inform a demand response guess would imply a positive relationship between consideration time and the size of the consideration set of adaptation options, as well as a positive relationship between a shock to the mental availability of options and the size of the set. Under a set of formal assumptions detailed in Appendix A, consideration time and higher mental availability of adaptation options would correspond to a demand response guess with higher magnitude, as well, and thus a negative relationship.

3 Design

The goal of the design is measure an individual's perception of demand response to higher prices while experimentally manipulating the time cost and effort cost of populating the consideration set through hypothetical thinking. This is accomplished in two designs. The first design, called "Standard Setting" focuses solely on the the manipulation of consideration time. The second design, "Introspection Setting," introduces introspection prompts prior to eliciting a demand response guess to reduce effort cost of hypothetical thinking via the rehearsal mechanism.

3.1 Standard Setting

To incentivize subjects to provide their true belief of demand response, they are asked:

Between 1990 and 2005, a real country, which we will refer to as Country X, introduced new taxes on purchases of gasoline that gradually made a gallon of gasoline 25% more expensive over those years.

What was the direct effect of the policy on the amount of gasoline consumed per person in Country X?

Please answer below in **percentage terms**.

If you believe the policy had **no effect**, enter **0**.

If you believe the policy **increased** gas consumed, enter the **number** of percent you think it went

up.

If you believe the policy **decreased** gas consumed, enter a **negative number** using the minus symbol.

Numerical text guesses of demand responsiveness, which we refer to as DRG_k for subject k, could not be submitted unless between -100 and +100. The subject, also known as an assessor, is informed that a precisely correct response (nearest percentage point) will be rewarded with a $10 bonus, and answering within 3 percentage points of the correct answer earns them a $1 bonus.

Subjects are never told the true country, Sweden, to prevent temptation to research the answer online. The answer considered precisely correct by the experimenter, -39, is the nearest whole percent to the -39.2% estimate derived from a peer-reviewed and recent empirical study on the impact of Sweden's policies, Andersson (2019). [4] To earn the $1 bonus, subject's must answer -37, -38, -40, -41, or -42.

In the standard setting subjects were randomly assigned to one of three possible time limits within which to provide their demand response guess, 120 seconds, 60 seconds, and 30 seconds. Subjects would see at all times how much time remained. Not answering within the time limit would disqualify subjects for a bonus payment and simply transition to the next block in the survey experiment. For submitted responses, subjects were asked to confirm the direction(up, down, no change) of their response and provide their

[4]It's worth noting that the demand response to taxes found by Andersson (2019) were much larger than would be expected from estimates of price elasticity and larger than tax elasticity estimates in earlier research. It's also worth noting a critique about any specific point estimate is arguably irrelevant to the incentive compatibility of the design. Sweden is an ideal choice for this setting because of the high rate of pass-through and the tight correlation of final market prices with taxes. Between 1990 and 2005, 25% taxes were gradually introduced and tax-inclusive gasoline prices rose approximately 25% in the same span of time. This allows for simpler phrasing when eliciting the demand response guess.

confidence about it. In essence, they are asked the likelihood on a 0-to-100 scale that their response is approximately correct (i.e., within 3 percentage points of the answer).

3.2 Introspection Setting

The "Introspection Setting" introduces to the "Standard Setting" design an open response writing prompt prior to eliciting a demand response guess using the instrument detailed in the previous subsection. In the "Introspection Setting," the first thing that subjects see upon consenting to participate is an open response writing prompt. They are required to write of a length of 70 characters before being allowed to move forward. Each subject is randomly assigned to one of four possible writing prompts, the text of which is provided below:[5]

1. Please take a few moments to consider whether automobile safety has changed (better, worse, or same) over the last 20 years, and provide your honest opinions. (**Baseline**)

2. Imagine yourself in a world in which gasoline prices are much higher than they are today. Please take a few moments to think and then write whatever comes to your mind about this scenario. (**Introspect Broad**)

3. Imagine yourself in a world in which gasoline prices are much higher than they are today. Please take a few moments to think about that. Would people's behaviors change? If so, how would they change and why? If not, why would they not change?

[5]The text was presented to subjects in a more visually appealing way than shown here, and the prompt names, e.g. Baseline, were not included.

(**Introspect Behavior**)

4. Imagine yourself in a world in which gasoline prices are much higher than they are today. Please take a few moments to think about that. Please list some suggestions that people could do to reduce their expenses on gasoline. (**Introspect Advice**)

Another difference from the "Standard Setting," is that now there are just two possible time limits, 120 seconds and 30 seconds, to which subjects are randomly assigned. Apart from this, the "Introspection Setting" maintains the conditions in "Standard Setting," including the wording of the demand response elicitation.

The purpose of these writing prompts is to experimentally vary the extent of hypothetical thinking relevant to demand response that a subject rehearses prior to being asked to guess the "Standard Setting" question. Subjects assigned to Prompt 1 form is a control group to establish a baseline, assuming that Prompt 1 does not induce relevant hypothetical thinking for the subsequent questions. Prompts 2-4 use different approaches to stimulate hypothetical thinking that will be relevant to the demand response guess to come.

3.3 Exit questions

Subjects are asked to provide their current local gas price; an estimate of their personal consumption of gasoline using several ranges; whether they have a car; whether they have an electric vehicle; and whether they believe they are an above, below or average consumer of gasoline. Gender, race, left-right political ideology on a 1-to-7 scale are also obtained. Four economic reasoning questions were also collected, but have not been

incorporated into this draft, as they are intended for a forthcoming empirical model of demand response guesses to be presented in future versions of this paper. I also elicit support for carbon taxes, gasoline taxes, congestion charges, and cigarette taxes.

3.4 Sample

Subjects were recruited on Prolific, with a filtering criterion that 35% of respondents identify as Republican, 35% identify as Democrats, 24% identify as Independent, 3% identify as "Other," and 3% identify as "None." They were not told anything before starting about the subject matter, only that the process would take 5 minutes and that they would be paid a base rate of $1, with the potential to earn up to $10 in bonus.

There were initially 332 Standard Setting subjects and 223 Introspection Setting subjects. Anyone who failed a simple attention check were discarded from the sample, producing samples of 290 and 184 for the Standard and Introspection subject pools, respectively. These subjects demonstrated a reasonably representative of the population in terms of race, gender, and state of residence, although with slight lower levels of participation from underrepresented racial minorities (i.e. respondents not identifying as White or Asian).

4 Results

In the entire sample across 2 settings, the mean and median demand response guess was -5.479% and -4.120%, respectively, in line with the hypothesis that people view gasoline demand as relatively inelastic. Even if excluding subjects who provide nonsensical

positive demand guesses, the mean guess was -10.473%, still relatively inelastic for a 25% price increase. Just 8 of 474 participants had answers near enough to the correct response, -39.2%, to make them eligible for a bonus payment.

Broadly, the results are suggestive of the role of hypothetical thinking. In Table 1, we see that in the "Standard Setting" consideration time is weakly predictive of DRG_k in the hypothesized negative direction. When evaluating the full sample, statistical significance is only observed in specifications including the squared response time. Squared response time's positive coefficient is in line with the hypothesis of diminishing returns to hypothetical thinking. Stringent time limits, in columns 3 and 4, too, have very weak effects on the full sample. Here, and throughout this section, the inclusion of demographic, political, and consumption control variables changes coefficient estimates.

While the signs of the effects are largely as predicted in the full sample, the magnitudes and statistical significance are lacking. This appears due, to some extent, to the 17.2% of the full sample that provide a positive demand response to a price increase. These subjects may not have fully read or been distracted while attending to the question; they may have misinterpreted the question to be asking not about quantity but about spending amounts; or a number of other speculative reasons. My approach for modeling hypothetical thinking does not account for these kinds of errors, nor is it interested in them.

In columns 5-8, the analysis uses the 82.8% subsample with nonpositive demand response guesses. Generally, interpretation can be challenging when subsampling on a function of the outcome variable. As shown in Appendix Table A1, response time variables are not predictive of the likelihood of supplying a positive demand response

guess. This offers some scope to analyze the effects of response time variables on the subset who observe the Law of Demand.

In the Law of Demand, or LoD, sample, we observe more convincing effects of response time. In the linear, column 5, specification, an additional minute of consideration corresponds to an additional 6 percentage points of negative demand response. We continue to see of curvature in column 6. In columns 7-8, we see that an exogenous time limit of 30 seconds shifts demand response higher by more than 5 percentage points, relative to the 120 second baseline time limit. Thus implies that a 30 second time limit shifts response perception from about -10% at the mean to roughly half of that, about -5%.

Table 1: Effect of Time Variables on Demand Response Guess

	Dependent variable:							
	Demand response guess (in percentage points)							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Constant	−2.541	5.318	−4.882	−13.003	−5.236	−0.301	−10.688	−12.980
	(2.385)	(5.477)	(1.457)	(4.628)	(1.808)	(4.407)	(1.096)	(3.420)
Resp. Time	−0.044	−0.380*			−0.100***	−0.309**		
	(0.049)	(0.216)			(0.037)	(0.174)		
Resp. Time^2		0.003				0.002*		
		(0.002)				(0.001)		
30 Sec. Limit			1.798	1.116			5.324***	5.553***
			(2.649)	(2.730)			(1.893)	(1.964)
60 Sec. Limit			−0.012	0.339			−0.641	−0.339
			(2.272)	(2.349)			(1.710)	(1.772)
Group:	All	All	All	All	LoD	LoD	LoD	LoD
Controls:				X				X
Mean Outcome	-4.521	-4.521	-4.521	-4.521	-9.671	-9.671	-9.671	-9.671
Observations	290	290	290	290	240	240	240	240

Notes: The outcome variable, demand response guess, is the subject's response to the question of how much gasoline demand changed per capita as a direct response to 25% increases in gas prices via taxes. Baseline subjects were provided 120 seconds to provide this response, while 30 Second Limit subjects and 60 Second Limit subjects had less time. Within those boundaries, subjects could still submit early. Response time and squared response time, selected by subject within the time limit are included. Controls included a subject's local gas price, their personal consumption of gas, gender, race, ideology, whether they have a car, whether they have an electric vehicle, and 2 dummies for if they believe they are above average consumer of gas or below average consumer. Analysis is provided for all subjects, as well as the subset of subjects whose demand response to a price increase was nonpositive, referred to here as group LoD.
Extended table: Table A4. *p<0.1; **p<0.05; ***p<0.01.

The second study of this paper, "Introspection Setting," provides some compelling evidence of the role of introspection, and its potential to invoke a rehearsal of adaptation options. In Table 2, over the full sample, effects of Introspect Broad are not meaningful. In the Law of Demand subsample that comprises about 87% subjects in this study, the coefficient is negative as predicted.[6] The Introspect Behavior prompt, a rather light intervention, had profoundly significant effects, whether analyzing the full sample or the Law of Demand subsample. In the full sample, holding all else content, encountering an Introspect Behavior prompt reduced demand response guesses by approximately 20 percentage points. Finally, Introspect Advice, the prompt with the highest ex ante likelihood of guiding a subject's thinking, had mildly negative effects (about 6 percentage points) in full sample, lacking statistical significance, and strongly significant negative effects (about 14 percentage points) in the Law of Demand subsample.

Across all specifications, the effect of response time was consistently about 0.3 percentage points lower per second spent with the Demand Response Guess question, or about 18 percentage points lower per minute. These are very meaningful levels and statistical significant across specifications and samples. Interestingly, the interaction of response time and the introspection prompts were all estimated to have a positive sign, suggesting the relationships hypothesized in Section 2. In their instrumental value to producing an accurate demand response guess, availability, via introspection, and consideration time appear to be substitutes.

[6]An outsize number of subjects assigned to Introspect Broad providing positive responses violating the Law of Demand. In Table A2, we can that Introspect Broad somehow the likelihood of positive response to increase to 26% from a base rate of less than 1% for subjects using the neutral prompt about auto safety. It's possible that the phrasing of Introspect Broad produced an unintended effect among subjects.

Table 2: Effect of Introspection Interventions and Response Time on Demand Response Guess

	Dependent variable:			
	Demand response guess (in percentage points)			
	(1)	(2)	(3)	(4)
Constant	2.985	13.656	1.817	6.410
	(5.911)	(11.477)	(4.021)	(8.170)
Introspect Broad	5.658	5.162	−2.787	−3.357
	(8.374)	(8.593)	(5.534)	(5.731)
Introspect Behavior	−21.750***	−21.640**	−10.977*	−12.666**
	(8.213)	(8.547)	(5.814)	(6.069)
Introspect Advice	−6.725	−5.741	−14.685**	−14.100**
	(8.374)	(8.571)	(5.883)	(6.083)
Response time	−0.293*	−0.302*	−0.281***	−0.288***
	(0.153)	(0.156)	(0.103)	(0.106)
Intr. Broad x Time	0.071	0.099	0.075	0.069
	(0.213)	(0.218)	(0.133)	(0.140)
Intr. Behavior x Time	0.464**	0.464**	0.137	0.163
	(0.195)	(0.199)	(0.140)	(0.143)
Intr. Advice x Time	0.149	0.144	0.246*	0.230
	(0.200)	(0.205)	(0.137)	(0.142)
Controls:		X		X
Group:	All	All	Observe Law of Demand	Observe Law of Demand
Observations	184	184	161	161

Notes: The outcome variable, demand response guess, is the subject's response to the question of how much gasoline demand changed per capita as a direct response to 25% increases in gas prices via taxes. Baseline subjects were preempted with an open response prompt on automobile safety before eliciting a demand response guess; Introspect Broad subjects were asked to write openly after imagining living in a world with much higher gasoline prices; Introspect Behavior subjects were asked to write about any behavioral change under higher gas prices; Introspect Advice were asked to provide suggestions under higher gasoline prices. Controls included a subject's local gas price, their personal consumption of gas, gender, race, ideology, whether they have a car, whether they have an electric vehicle, and dummies if they believe they are above average consumer of gas or below average consumer. The group "Observe Law of Demand" is the subset with nonpositive demand responses. *p<0.1; **p<0.05; ***p<0.01.

In other results, shown in Appendix Table A3, we see that there is a weakly significant relationship between the guess value and confidence about it. More confident guesses tend to be lower values of DRG_k, corresponding to a larger negative demand response at the mean.

5 Conclusion

When voters are asked in surveys to assess a policy on a specific criteria, such as its effectiveness in reducing an externality, one should consider confounds to belief elicitation. A public support literature provides a lengthy history of pessimistically low (i.e., inelastic) impressions of the effectiveness of fossil-fuel taxes, but this impression could be misreported or motivated by a distaste for other aspects of the policy, such as its net welfare effect on the voter's household, or its concern with challenges like climate change, which the the voter may reject or not prioritize. These other policy aspects, while irrelevant to the question of a factual demand response, should not be assumed to kept apart from the respondent's answer. By introducing monetary incentives to motivate truthful impressions on a past policy, this study introduces a tool to more accurately reveal a respondent's beliefs. I find that the notion of relatively inelastic demand is unlikely to be attributable solely to misreporting or motivated reporting, because beliefs of relatively inelastic demand could be found even in the presence of meaningful incentives. In the future, incentivized techniques could be used to calibrate unincentivized responses from voters and survey respondents on a broad range of policy matters, in which a concern exists that motivations may confound belief elicitation.

Broadly, the study produces compelling evidence that time constraints and introspection prompts can meaningfully shape beliefs about demand responsiveness. Some of the most promising findings require ex-post conditioning on subjects who align with the Law of Demand. Gasoline is not a Giffen good, and so it seems reasonable to consider whether the subjects defying the Law of Demand are of a type outside the hypothetical thinking framework proposed by this paper and introduce unnecessary noise. It seems plausible that some of these voters misunderstood what they were being asked, perhaps mistaking per capita responses for aggregate responses (population effects), or mistakenly providing a view on per capita gasoline spending vs. demand in gallons of gas (expenditure vs. demand). for In future designs, it may be valuable to elicit per capita demand, aggregate demand, and per capita expenditure. Also, subjects may have provided a guess about a change in per capita demand, and not a change in per capita demand directly attributable to the policy. Similarly, it may be worth eliciting both of these guesses, to reduce the role of misunderstanding and shrink the number of subjects violating the Law of Demand.

Analyses on the effects of response time and introspection prompts do seem strongly suggestive of the role of hypothetical thinking, for the broad sample and especially for subjects obeying the Law of Demand. Another promising direction for future designs would allow a more direct attribution to specific reasoning processes of demand response guesses and policy proposal assessments. The current design tests for evidence of the implications of a hypothesized model, but does not directly test the model. A closer test of the model may be possible by analyzing the open responses and is a promising path for future versions of this study. Additionally, results from other, more heavy-handed designs

may complement the findings here to provide more robust evidence of the consideration model.

As far as implications, using information, trial periods, or introspection might shift support by increasing the mental availability of hypothetical adaptive options. Studies that use information treatments can find it difficult to separate belief updating from demand effects. This project points to the possibility of using introspection prompts, with no information, to estimate a lower bound for how effective information might be, and without provoking demand effects. The idea of easing the costs of hypothetical thinking through introspective rehearsal of memory also implies that trial periods of policies, if feasible, could improve an otherwise biased understanding of policy effects. Finally, introspection prompts themselves may offer a way for voters who mistrust external information to form policy opinions that are more accurate than a typical, hasty assessment.

The results show that demand response guesses are not strongly correlated with policy preferences across the full sample, but are strongly correlated in Democrat or liberal subsamples. Larger (negative) demand response guesses are strongly predictive of support among Democrats and liberals for carbon and gasoline taxes. This has important policy implications as the failure to adopt carbon pricing, congestion pricing, and other similar Pigouvian policies is not simply a matter of winning over enough conservative voters. As discussed earlier, while Republicans and conservatives tend to oppose carbon pricing, Democrats and liberals are also often opposed, despite their economic advantages and flexibility.

Broad disagreement over topics like pollution and climate change is an obstacle to achieving consensus for environmental policies. But even among those highly motivated

to address climate change, the failure to accurately appraise the effectiveness of efficient solutions, like Pigouvian taxes, due to biases from quick, inattentive assessments, presents an altogether different obstacle, and one that may easier to address.

The framework proposed in this paper, where imperfect, slow recall limits consideration sets and results in systematic mistakes, could apply to settings outside of policy selection. Policy selection is a particularly well-suited application because hasty judgments are likelier given the extremely low stakes for any individual voter, despite the large aggregate stakes (Downs, 1957). But we can imagine that an employee who currently saves little and must choose between a defined benefit pension or a retirement plan with generous company matching policy may have difficulty imagining the ways she might curtail spending to save money. An ex ante judgment is a forecast, and it's possible that, in being pessimistic about her ability to save, she may choose a defined benefit option with an inferior ex post outcome than if she had chosen matching plan. Similar errors might be made when choosing among health care plans. A hasty assessment of a low premium, high deductible plan may fail to sufficiently incorporate consideration of list of the possible events and expenses. Sufficient time and willingness to engage in hypothetical thinking, in these instance, could result in better choices.

References

Anderson, S., Marinescu, I., and Shor, B. (2023). Can pigou at the polls stop us melting the poles? *Journal of the Association of Environmental and Resource Economists*, 10(4):903–945.

Andersson, J. J. (2019). Carbon taxes and co2 emissions: Sweden as a case study. *American Economic Journal: Economic Policy*, 11(4):1–30.

Aragones, E., Gilboa, I., Postlewaite, A., and Schmeidler, D. (2005). Fact-free learning. *American Economic Review*, 95(5):1355–1368.

Baranzini, A. and Carattini, S. (2017). Effectiveness, earmarking and labeling: testing the acceptability of carbon taxes with survey data. *Environmental Economics and Policy Studies*, 19:197–227.

Baranzini, A., Van den Bergh, J. C., Carattini, S., Howarth, R. B., Padilla, E., and Roca, J. (2017). Carbon pricing in climate policy: seven reasons, complementary instruments, and political economy considerations. *Wiley Interdisciplinary Reviews: Climate Change*, 8(4):e462.

Caplin, A., Dean, M., and Martin, D. (2011). Search and satisficing. *American Economic Review*, 101(7):2899–2922.

Carattini, S., Carvalho, M., and Fankhauser, S. (2018). Overcoming public resistance to carbon taxes. *Wiley Interdisciplinary Reviews: Climate Change*, 9(5):e531.

Coglianese, J., Davis, L. W., Kilian, L., and Stock, J. H. (2017). Anticipation, tax avoidance, and the price elasticity of gasoline demand. *Journal of Applied Econometrics*, 32(1):1–15.

Colmer, J., Martin, R., Muûls, M., and Wagner, U. J. (2022). Does pricing carbon mitigate climate change? firm-level evidence from the european union emissions trading scheme.

Dieler, J., Jus, D., and Zimmer, M. (2015). Fill'er up! anticipation and inventory effects on fuel demand. *Manuscript, Ifo Institute for Economic Research*.

Douenne, T. and Fabre, A. (2022). Yellow vests, pessimistic beliefs, and carbon tax aversion. *American Economic Journal: Economic Policy*, 14(1):81–110.

Downs, A. (1957). An economic theory of political action in a democracy. *Journal of political economy*, 65(2):135–150.

Eliasson, J. and Jonsson, L. (2011). The unexpected "yes": Explanatory factors behind the positive attitudes to congestion charges in stockholm. *Transport Policy*, 18(4):636–647.

Eliaz, K. and Spiegler, R. (2011). Consideration sets and competitive marketing. *The Review of Economic Studies*, 78(1):235–262.

Esponda, I., Vespa, E., and Yuksel, S. (2024). Mental models and learning: The case of base-rate neglect. *American Economic Review*, 114(3):752–782.

Hausman, J. A. and Newey, W. K. (1995). Nonparametric estimation of exact consumers

surplus and deadweight loss. *Econometrica: Journal of the Econometric Society*, pages 1445–1476.

Hoyer, W. D. (1984). An examination of consumer decision making for a common repeat purchase product. *Journal of consumer research*, 11(3):822–829.

Janusch, N., Kroll, S., Goemans, C., Cherry, T. L., and Kallbekken, S. (2021). Learning to accept welfare-enhancing policies: an experimental investigation of congestion pricing. *Experimental Economics*, 24:59–86.

Kallbekken, S. and Aasen, M. (2010). The demand for earmarking: Results from a focus group study. *Ecological economics*, 69(11):2183–2190.

Kallbekken, S. and Sælen, H. (2011). Public acceptance for environmental taxes: Self-interest, environmental and distributional concerns. *Energy Policy*, 39(5):2966–2973.

Li, S., Linn, J., and Muehlegger, E. (2014). Gasoline taxes and consumer behavior. *American Economic Journal: Economic Policy*, 6(4):302–342.

Metcalf, G. E. (2021). Carbon taxes in theory and practice. *Annual Review of Resource Economics*, 13:245–265.

Metcalf, G. E. (2023). Five myths about carbon pricing. Technical report, National Bureau of Economic Research.

Mullainathan, S. (2002). A memory-based model of bounded rationality. *The Quarterly Journal of Economics*, 117(3):735–774.

Rivers, N. and Schaufele, B. (2015). Salience of carbon taxes in the gasoline market. *Journal of Environmental Economics and management*, 74:23–36.

Roberts, J. H. and Lattin, J. M. (1997). Consideration: Review of research and prospects for future insights. *Journal of Marketing Research*, 34(3):406–410.

Tiezzi, S. and Verde, S. F. (2014). Overreaction to excise taxes: the case of gasoline. *Robert Schuman Centre for Advanced Studies Research Paper No. RSCAS*, 54.

A Appendix Section: Theoretical Framework

Suppose that a consumer will pick and choose from a series of adaptive responses to a price change, and that the comprehensive list of these adaptive responses form the set S. Adopted adaptations depend on consumer-specific considerations, and there is heterogeneity of budget constraints and preferences.

For example, not everyone can or wants to switch to an electric vehicle. Gauging the success of a gasoline tax solely on its ability to convert traditional ICE vehicles into EV drivers will underestimate the effectiveness of the proposal. Consumers who do not switch to EVs may opt to carpool or reduce trips that had higher marginal utility than the market price but lower marginal utility than the tax-inclusive price.

For convenience, we make the assumption that introducing "carpool" in the consideration set will not cause the policy analyst to reassign would-be EV adopters under the smaller consideration set to now stay with ICE vehicles. This is an example of a greedy approximation algorithm, wherein choices once assigned are not reevaluated. A perfect optimization among adaptation options would be computationally intractable as the number of adaptation options grows.

An assessor considers adaptations $a_1, a_2, ..., a_j$, each elements of a consideration set, itself a subset of a comprehensive set S. Individual i's consideration set $\tilde{S}$ can evolve over time with additional consideration time or effort. When $\tilde{S}$ contains j elements, we can index the set as $\tilde{S}_j$. Arrival of element a_{j+1} into the assessor's consideration set $\tilde{S}_j$ results in a new consideration set $\tilde{S}_{j+1}$.

We can assume that the comprehensive set S is a function of the size of τ and

other economic circumstances at the time of introduction. For instance, a small τ may correspond to a small S. For any evaluation problem we take τ is fixed and exogenously provided. The arrival of set elements into the consideration set, i.e. the expansion of $\tilde{S}$ toward the comprehensive set S, requires consideration effort, consideration time, and mental availability of adaptation options. Then we can say that for any assessor for whom the arrival of set elements follows an exogenous ordering (i.e. a particular assessor will always think of EV's first, always followed by hybrids, always followed by carpooling, and so on.), the index j captures how "far" the consideration set has expanded and is increasing in time, effort, and mental availability. Or, to put it differently, given an ordered sequence of elements, $\mid \tilde{S} \mid = f(T, E, A)$.

Expansion of a consideration set impacts the perceived aggregate demand reduction. For example, the arrival of adaptation option a_{j+1} into the consideration set consisting previously only of $a_1, ..., a_j$, or $\tilde{S}_j$ can impact the assessor's perceived aggregate demand reduction as follows.

An assessor considers the aggregate demand response by applying the consideration set of adaptation options to economy's consumers. We denote a consumer using subscript i, where $i \in N$ consumers (or alternatively we can denote a consumer *type* using subscript i, for a total of N representative types).

A consumer i's optimal response α_i, given a set of adaptation options, comprises a subset of these adaptation options.

An assessor uses a set of N functions, denoted α_i, that maps the elements of $\tilde{S}$ to the chosen behaviors of hypothetical consumers (or types) in the economy. We can think of α_i as i's response to adaptation options, given policy τ. So, α_i is hypothetical idiosyncratic

behavior of a consumer, perhaps the result of preferences and circumstances, including the options perceived as available. Of course, it's unlikely that an assessor would consider behavior for a very large number of individual consumers, so here we see why an assessor would likely consider N types that cluster consumers by predictive characteristics.

Consideration-to-assessment process

A demand response assessment occurs dynamically with the mental arrival of adaptation options. The arrival of a new adaptation option a_{j+1} into an assessor's consideration set can cause the assessor to assign the new adaptation to a consumer i's previous set of adaptation responses $\alpha_i(\tilde{S}_j)$, which may be empty. If introducing the new adaptation option to the extant adaptation response set is utility improving, the option is added to the adaptation response set.

Assumption 1 *If* $v_i(\alpha_i(\tilde{S}_j) \cup a_{j+1}) > v_i(\alpha_i(\tilde{S}_j))$, *then* $\alpha_i(\tilde{S}_{j+1}) = \alpha_i(\tilde{S}_j) \cup a_{j+1}$.

Otherwise $\alpha_i(\tilde{S}_{j+1}) = \alpha_i(\tilde{S}_j)$.

Notice that once an adaptation is assigned to a consumer, it is not removed from their response set, i.e. a greedy algorithm.[7] A reduction function r_i counts the amount of gasoline consumption hypothetically voided by a consumer's adaption response set

[7]The use of the tractable greedy algorithm in mapping the consideration set to aggregate reductions introduces error that could over- or underestimate demand responses relative to the truth, as well as order-dependence. As an example, an assessor who considers an electric vehicle first, assigning it to a consumer's response set, would not then assign a subsequently arriving option like carpooling. However given a choice set, {EV, carpool} , suppose the consumer would choose only carpool (and not to switch to an EV). Then greedy algorithm would overstate the consumer's demand reduction. Conversely, the opposite can happen, understating the demand response. An assessor might consider carpooling first, and assign it to the consumer's demand response set. Subsequently, an EV is considered, but not assigned, because the consumer would not switch to an EV conditional on already being a carpooler. Given a choice set, {EV, carpool}, this consumer might choose EV, in which case the heuristic leads to an underestimate of his demand response.

$\alpha_i(\tilde{S}_j)$, thus we write $r_i(\alpha_i(\tilde{S}_j)) = r_i(\tilde{S}_j)$. By design, $\tilde{S}_{j+m} \subset \tilde{S}_{j+m}$, $\forall m > 0$. And because adaptation options result in reductions, $r_i(\tilde{S}_{j+m}) \geq r_i(\tilde{S}_j)$ and, when $\alpha_i(\tilde{S}_j) \subset \alpha_i(\tilde{S}_{j+m})$, the inequality is strict: $r_i(\tilde{S}_{j+m}) > r_i(\tilde{S}_j)$. It follows that the aggregate assessment $\sum_{i=1}^{N} r_i(\tilde{S}_{j+m}) > \sum_{i=1}^{N} r_i(\tilde{S}_j)$.

Remember that $\tilde{S}$, $\alpha_i()$, and $r_i()$ are all perceptual objects of the assessor. If we wish to focus solely on the consideration set channel and assume accuracy in an assessor's α_i and r_i functions, we can compare an assessor's subjective perception of aggregate demand reductions to a true aggregate demand reduction. If $\tilde{S} \subset S$, then $\sum_{i=1}^{N} r_i(S) > \sum_{i=1}^{N} r_i(\tilde{S})$. In other words, a non-comprehensive consideration set results in an assessment falling shy of the true aggregate reduction attributable to the policy.

B Appendix Section: Tables

Table A1: Response Time and Violating Law of Demand

	Dependent variable:		
	Dummy Variable: Violates Law of Demand		
	(1)	(2)	(3)
Constant	0.159***	0.281**	0.049
	(0.053)	(0.123)	(0.161)
Response Time	0.0003	−0.005	−0.004
	(0.001)	(0.005)	(0.005)
Response Time^2		0.00005	0.00004
		(0.00004)	(0.00004)
Local Gas Price			−0.001
			(0.001)
Gas Consumption			−0.007
			(0.009)
Male			0.024
			(0.046)
No Car			0.167**
			(0.082)
Electric vehicle			0.118
			(0.183)
Political ideology			0.013
			(0.009)
Underrep. Race			0.049
			(0.063)
High gas consumer			0.048
			(0.092)
Low gas consumer			0.038
			(0.055)
Observations	290	290	290

Notes: The outcome is a dummy variable indicating whether the subject's demand response guess for a tax increase was positive. Controls include subject's local gas price, their personal consumption of gas, gender, race, ideology, car ownership, electric vehicle ownership, and dummies for being a relatively high or low gas consumer. *p<0.1; **p<0.05; ***p<0.01.

Table A2: Introspection Interventions, Response Time and Violating Law of Demand

	Dependent variable:
	Dummy Variable: Violates Law of Demand
Constant	0.008
	(0.068)
Introspect Broad	0.253***
	(0.069)
Introspect Behavior	0.059
	(0.069)
Introspect Advice	0.105
	(0.070)
Response Time	0.0004
	(0.001)
Observations	184

Notes: The outcome variable is a dummy variable indicating whether the subject's demand response guess was positive for a 25% increase in prices in "Country X." Baseline subjects were preempted with an open response prompt on automobile safety before eliciting the demand response guess; Introspect Broad subjects were asked to write openly after imagining living in a world with much higher gasoline prices; Introspect Behavior subjects were asked to write about any behavioral change under higher gas prices; Introspect Advice were asked to provide suggestions under higher gasoline prices. *p<0.1; **p<0.05; ***p<0.01.

Table A3: Relationship between subjective confidence and demand response guess

	Dependent variable:
	Demand response guess (in percentage points)
Constant	−1.300
	(1.987)
Response confidence (0 to 100)	−0.067*
	(0.036)
Mean Outcome	-4.521
Observations	290

Notes: The outcome variable, demand response guess, is the subject's response to the question of how much gasoline demand changed per capita as a direct response to 25% increases in gas prices via taxes. The response confidence variable is the subject's answer subsequent to the demand response elicitation on how likely they think their demand response guess is approximately correct. *p<0.1; **p<0.05; ***p<0.01.

Table A4: Effect of Time Variables on Demand Response Guess

	Dependent variable:							
	Demand response guess (in percentage points)							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Constant	−2.541	5.318	−4.882	−13.003	−5.236	−0.301	−10.688	−12.980
	(2.385)	(5.477)	(1.457)	(4.628)	(1.808)	(4.407)	(1.096)	(3.420)
Resp. Time	−0.044	−0.380*			−0.100***	−0.309**		
	(0.049)	(0.216)			(0.037)	(0.174)		
Resp. Time^2		0.003				0.002*		
		(0.002)				(0.001)		
30 Second Limit			1.798	1.116			5.324***	5.553***
			(2.649)	(2.730)			(1.893)	(1.964)
60 Second Limit			−0.012	0.339			−0.641	−0.339
			(2.272)	(2.349)			(1.710)	(1.772)
Local gas price				−0.031				−0.015
				(0.041)				(0.028)
Gas consumption				−0.126				−0.019
				(0.424)				(0.318)
Male				−0.212				0.173
				(2.087)				(1.546)
No Car				5.858				1.337
				(3.697)				(2.655)
Electric vehicle				6.052				5.190
				(8.243)				(6.153)
Political ideology				0.607				0.161
				(0.426)				(0.320)
Underrep. Race				1.793				−2.070
				(2.831)				(2.126)
High gas consumer				−0.679				−3.847
				(4.170)				(3.143)
Low gas consumer				1.581				1.783
				(2.472)				(1.876)
Group:	All	All	All	All	LoD	LoD	LoD	LoD
Mean Outcome	-4.521	-4.521	-4.521	-4.521	-9.671	-9.671	-9.671	-9.671
Observations	290	290	290	290	240	240	240	240

Notes: The outcome variable, demand response guess, is the subject's response to the question of how much gasoline demand changed per capita as a direct response to 25% increases in gas prices via taxes. Baseline subjects were provided 120 seconds to provide this response, while 30 Second Limit subjects and 60 Second Limit subjects had less time. Within those boundaries, subjects could still submit early. Response time and squared response time, selected by subject within the time limit are included. Controls included a subject's local gas price, their personal consumption of gas, gender, race, ideology, whether they have a car, whether they have an electric vehicle, and 2 dummies for if they believe they are above average consumer of gas or below average consumer. Analysis is provided for all subjects, as well as the subset of subjects whose demand response to a price increase was nonpositive, referred to here as group LoD. *p<0.1; **p<0.05; ***p<0.01.