

The roles of social norms and economic reasoning in shaping support for carbon pricing

Pre-analysis Plan

1. Motivation

Climate change is one of the greatest challenges of our times. Economists almost unanimously agree on the fundamental role of carbon pricing in the transition to a decarbonized economy. Yet, in many countries to date, Pigouvian pricing schemes in the environmental context have not been implemented to an extent advocated by scientists, partly due to fear of voters' backlash. Thus, how to overcome resistance and foster support for crucial environmental policies is thus of primary interest for academics and policymakers alike.

In this project, we aim to investigate the joint role of economic reasoning and perceived social norms in shaping attitudes towards carbon pricing. Previous empirical evidence has suggested that many individuals have limited understanding of Pigouvian taxation and that their attitudes towards such policies are imputable to their perceived costs and benefits. Thus, educating people about the merits of carbon pricing can be a vital stepping-stone. However, individuals may not be willing to endorse a cause if they believe that other people do not support it or do not approve of this. Under pluralistic ignorance, this can lead to socially inefficient equilibria. For example, it has been argued for the U.S. that “[c]orrecting misperceived norms of opposition and decoupling policy evaluation from identity concerns would help overcome [...] seemingly insurmountable barriers to bipartisan support for climate policy” (Van Boven et al., 2018).

We are going to run a large-scale online survey experiment in a representative sample of 3,000 U.S. adults. Our main outcome variable will be an incentivised donation decision to elicit support for carbon pricing. To study the causal effects of social norms and economic reasoning, we exogenously provide participants with information through animated video interventions. Participants are randomly allocated to one of four experimental conditions in a 2x2 design. In the first intervention (Norm treatment), we inform participants about descriptive social norms in the U.S. toward carbon neutrality policies. In the second intervention (Policy treatment), we explain in simple terms how carbon pricing (with uniform cash transfers) works according to economic theory, highlighting the Pigouvian mechanisms through which the policy can help to achieve carbon neutrality. Furthermore, the video addresses major concerns that people invoke to justify their lack of support

for carbon pricing, such as distributional concerns (e.g., Carattini et al. 2018). Finally, we test a combined intervention (Norm+Policy treatment) in which participants see both information videos.

These interventions will help us shed light on several questions. Does support for carbon pricing respond more strongly to norm-based appeals or the perceived merit of this policy? Is there heterogeneity in the way people are influenced by these interventions? Importantly, the combined treatment allows us study whether correcting misperceived norms can make individuals more receptive to merit-based information, e.g., due to motivated reasoning or motivated information search. We also plan to collect follow-up data to study persistence of the information interventions and the potential role of motivated memory.

2. Experimental Design

2.1. Setting

We plan to recruit an online sample that is nationally representative of the U.S. adult population across age, gender, and region. The survey is coded in Qualtrics and subjects will be recruited through Prolific, an online survey platform which is widely used to conduct social science experiments.

2.2. Treatments

For the experiment, subjects are randomly assigned to three different informational video interventions and a control group. These video interventions are embedded in the online survey. All subjects (including those in the control group) see a brief animated video that explains how anthropogenic carbon emissions contribute to climate change and that introduces the concept of carbon neutrality.

Subjects in the treatment groups will receive additional information, as described below.

- i) Norm treatment: The first treatment expands the initial animated video by further informing participants of descriptive social norms. In particular, participants will be (truthfully) informed that, according to a representative survey by the Pew Research Centre, 69% of adults in the U.S support the country's carbon neutrality goals.¹
- ii) Policy treatment: The education treatment consists of an additional educational video that explains, in simple terms, the economic mechanisms and consequences of carbon pricing (with lump-sum cash transfers to

¹ <https://www.pewresearch.org/science/2022/03/01/americans-largely-favor-u-s-taking-steps-to-become-carbon-neutral-by-2050/>

recycle revenues). In particular, the video addresses all the five main concerns that drive lack of public support for carbon pricing, as identified in previous literature (see Carattini et al., 2018).

- iii) Norm+Policy treatment: This treatment combines the information received in the norms video treatment as well as the policy video treatment.

To hold survey engagement, fatigue, and wage rates roughly constant across groups, subjects in the Control and Norm groups see an animated placebo video whose length is approximately the same as that of the carbon pricing explainer. The topic of the placebo video (loosely) environmentally-related but it does not convey any information on climate change or climate policies. Of course, we cannot rule out that the information from the placebo video may nevertheless have an effect in its own, but this would in fact lead to an underestimation of our treatment effects and thus work against us. However, to mitigate concerns about potential placebo effects, we will show an extremely condensed version of the placebo video to all subjects, including those in the Policy and the Norm+Policy groups.²

As mentioned, the video treatments convey simple and truthful information. Prior to rolling out the interventions climate policy experts and communication experts reviewed the video scripts certifying it was correct and easy to understand.

2.3. Survey structure

The survey will consist of 3 main blocks. In the first block (pre-intervention), we collect information on respondents' personal characteristics and values and elicit baseline beliefs and attitudes. This survey block will contain questions about demographics (e.g., age, education), energy dependence, political affiliations and ideology, and trust in different institutions. We also include survey modules on time, risk, and social preferences, as well as certain values and personality traits (e.g., Schwartz human values scale, social comparison orientation, cognitive reflection). Finally, we elicit participants' baseline beliefs and attitudes about climate change, carbon neutrality, social norms, as well as their views on carbon pricing and other climate policies. Some follow-up questions will also ask participants to state how (un)certain they are about their responses.

² We may run a robustness experiment after the main data collection to cleanly test the effect of placebo information on our outcome variables.

In the second block (intervention), respondents will see one of 4 video intervention variants, depending on which experimental condition they were assigned to (see the previous subsection). Table 1 shows the video treatments for each condition. Following each video, subjects are also asked to provide feedback, i.e., if they perceived the video as interesting, informative, politically biased, etc. To make sure that information in the placebo video has no distorting effect per se, we also show very short versions of the placebo video to all subjects before video 1.

Table 1: Video treatments by experimental condition

Experimental condition	Video 1 (<i>about 1-2 mins</i>)	Video 2 (<i>about 3-4 mins</i>)
Control group	Primer on climate change and carbon neutrality	Placebo video
Norm group	Primer on climate change and carbon neutrality + social norms	Placebo video
Policy group	Primer on climate change and carbon neutrality	Carbon pricing explainer
Norm+Policy group	Primer on climate change and carbon neutrality + social norms	Carbon pricing explainer

In the third block (post-intervention), we will measure how the different intervention videos affected participants' beliefs and attitudes as well as their support for carbon pricing. People will be asked about their views on climate change, their beliefs about carbon footprints, and their attitudes towards environmental policies, in particular carbon pricing. Importantly, we include an incentivized elicitation of environmental attitudes and support for carbon pricing. Participants will be told that by completing the survey, they are automatically enrolled in a lottery to win 100\$. We then give them the choice to donate part of their lottery prize to environmental organizations. One of the organizations will be a bipartisan group that advocates for carbon pricing in the United States, while the other environmental organization will not be related to carbon pricing. We match the donation amount that participants allocate within the survey to reduce "wait-and-see" motives. At the end of the survey, we will include some final demographic questions (e.g., on income and religion) and give the opportunity for participants to express their opinions in an open-ended text question.

2.4. Follow-up

We plan to conduct a follow-up survey that takes place at least 4 weeks after the original experiment to study whether treatment effects persist or fade away. Participants will not receive any additional treatments and we will restrict the survey question mostly to those in the third block of the original survey.

3. Data and variables

3.1 Main outcome variables

As the main outcome variable for our analysis, we use subjects' post-intervention donation choices to environmental organizations from their potential lottery winnings of \$100. These are continuous variables that can take values between \$0 and \$100 with increments of \$0.01. This revealed preferences measure will be used to assess the intervention's impact on pro-environmental behavior and support for carbon pricing. As mentioned, participants will be asked whether they wish to donate part of their lottery winnings to the Climate Leadership Council (CLC), a bipartisan non-profit organization that advocates for carbon pricing policies in the United States. The amount donated to CLC will provide us with an incentivized and fine-grained measure of carbon pricing support.

In addition to the incentivized donation choices, we will also look at stated level of support for introducing carbon pricing (with uniform cash transfers) as a policy in the United States on a 4-point Likert scale (strongly support, somewhat support, somewhat oppose, strongly oppose). For political economy considerations, the full distribution of views in the population matters, so we will examine effects not only on the share of individuals who support the policy, but also on the share of individuals who strongly oppose the policy.

3.2 Secondary outcome variables

Participants will also have the chance to donate to a second environmental organization whose work is not connected to carbon pricing (e.g., the National Wildlife Federation). We will make use of this second donation option to assess extensive margin (i.e., general pro-environmental attitudes) versus intensive margin (i.e., support specifically for carbon pricing) effects of our interventions, as well as potential spillover effects.

As additional outcome variables, we will include stated support for other climate policies on discrete scales. These measures will be collected after the video interventions. Furthermore, we assess the effects that the videos exert on subjects' general environmental attitudes and their reasoning about carbon pricing (mechanisms, effectiveness, distributional effects, economic impacts, etc.) through Likert scale questions. We

will also assess the effect of our videos on posterior beliefs about social norms about carbon neutrality in the United States on continuous scales.

3.3 Treatment variables

We will create indicator variables representing each of our treatments:

- $Norm_i$: takes the value of 1 if participant i received the social norm video intervention (Norms and Norms+Policy treatment) and 0 otherwise
- $Policy_i$: takes the value of 1 if participant i received the policy video intervention on carbon pricing (Policy and Norm+Policy treatment) and 0 otherwise
- $Norm_i \times Policy_i$: the interaction of $Norm_i$ and $Policy_i$ which takes the value of 1 only when participant i received both interventions (Norms+Policy treatment).

We code our variables in this way to easily test for interaction effects between information on social norms and information on carbon pricing.

3.4 Other variables

Additional explanatory variables are collected through baseline data from the first block of the survey and, in the case of potential mediator variables, also the third block of the survey. These variables include the following.

- **Socio-demographics:** These include age, gender, county of residence, ethnicity, religion, and level of education.
- **Economic situation:** Here we include questions on household income, energy dependency (e.g., car usage), concerns about inflation, employment status, and financial distress.
- **Political orientation:** These include measures of party affiliation, political ideology, affective partisanship on a feeling thermometer, and views on taxation and redistribution.
- **Climate change views:** A set of questions adapted from the Yale Program on Climate Change Communication will be used to elicit baseline environmental attitudes (Howe et al. 2015). We further include a question on whether respondents support the U.S. taking steps to become carbon neutral by 2050, as well as the perceived norm about support for carbon neutrality.

- **Carbon emissions inequality:** We will ask participants (post-intervention) to estimate the average carbon footprints of high- and low-income individuals in the United States, as well as their own carbon footprint.
- **Views on specific climate policies:** A set of Likert-scale questions will be used to measure baseline support for different climate policies, including carbon pricing with uniform cash transfers, as well as some more specific questions on knowledge about carbon pricing, beliefs about its effectiveness in reducing carbon emissions, and reasoning about economic mechanisms and consequences of carbon pricing (e.g., distributional effects, economic impact).
- **Trust in institutions:** This includes a collection of Likert-scale questions on people's trust in certain institutions or groups of people, ranging from government, major companies, and environmental organizations to scientists and journalists.
- **Conspiracy mentality:** We will include a set of items to capture people's general tendency to believe in conspiracy theories.
- **Risk, time, and social preferences:** We will adapt survey items from Falk et al. (2018) to measure participants' time and risk preferences as well as social preferences (altruism, generalized trust, positive and negative reciprocity). Altruism may be an underlying mechanism for donations to climate-oriented organizations, given that people believe it is a good cause. The degree of an individual's reciprocal behavior can give insight into their potential reaction to information on others' behaviors and willingness to cooperate.
- **Ambiguity attitudes:** Ambiguity aversion is elicited through a subset of items from the need for closure scale (Roets & Van Hiel, 2011) which measures the degree to which a person has a desire for certainty.
- **Social comparison orientation:** We will use a condensed version of the social comparison orientation scale.
- **Human values:** We will include survey items on conformity and universalism from the Schwartz basic human values scale.
- **Cognitive reflection:** To measure the tendency to override intuition with analytical thinking, we use an alternative version of the cognitive reflection test (Thomson & Oppenheimer, 2016) that avoids noise through numerical ability and familiarity with the original test by Frederick (2005).
- **Representativeness heuristic:** We include an estimation task for the share of red-haired people in Ireland as proxy for subjects' tendency to overweight probabilities based on the representativeness heuristic (Bordalo et al. 2016).

3.5 Power calculations and sample size

We aim for a sample size of 3,000 in total, which means there will be about 750 subjects in each experimental group. With 80 percent power to reject at the 5% significance level, our minimum detectable effect is about 14.5 percent of a standard deviation (SD). Adding control variables would further increase power. To gain some initial estimates for baseline beliefs and effects of the Norms treatment, we conducted a short pilot experiment with 300 participants on Prolific in July 2022. Our estimates imply that a sample size of 750 per group gives us sufficient power to reject the null hypothesis of no effect of the social norms intervention on support for carbon pricing at the 1% significance level.

For the follow-up survey, we will try to recruit at least 4 out of 5 participants from the original sample. This would give us sufficient power to detect an effect size of at least 16.2 percent of a SD with 80 percent probability.

4. Analysis

4.1. Baseline Balance

Participants will be randomly assigned to experimental conditions. Nevertheless, we will assess treatment balance by comparing a number of variables at baseline across treatments and control, by regressing them on treatment indicators. Baseline balance will be tested to the following variables: age, gender, income, education, political orientation, baseline climate change views, and baseline carbon pricing support.

4.2. Exploring baseline beliefs prior to intervention

Before analyzing the effects of our interventions, we will use responses from the first survey block (pre-intervention) to explore descriptive statistics and correlations of attitudes, beliefs, and policy support of participants prior to receiving any intervention.³ These explorations will include the baseline distribution of support toward U.S. carbon neutrality goals, carbon pricing and other climate policies, as well as perceptions of social norms toward carbon neutrality in the U.S. and beliefs about carbon inequality. We will also explore how well these attitudes and beliefs are predicted by, among others, socio-economic characteristics, political

³ Whenever we also look at responses from the third survey block (post-intervention), we will restrict our sample to the Control group.

affiliation, attitudes towards taxation and redistribution, trust in institutions. We may need to use sampling weights to improve ex post representativeness of our sample for the U.S. adult population.

4.3. Main effects of our interventions

Bivariate analyses

The random assignment of subjects in the survey allows us to identify the causal effects of our video information treatments. In a simple bivariate analysis, we will first compare average donations to the carbon pricing charities, total average donations to both environmental charities, and stated support for carbon pricing across experimental conditions.

Main specifications

For our formal empirical analyses of the impact of different types of information on the main outcome, we will estimate the following statistical model using ordinary least squares (OLS):

$$y_i = \alpha + \beta_N \cdot Norm_i + \beta_P \cdot Policy_i + \beta_{NP} \cdot Norm_i \times Policy_i + \mathbf{x}_i' \boldsymbol{\gamma} + \varepsilon_i$$

Where y_i is the outcome variable of interest for individual i , $Norm_i$ and $Policy_i$ are the indicator variables equal to one for individuals assigned to receive either the norms video intervention or the policy video intervention and 0 otherwise, \mathbf{x}_i is a vector of control variables, and ε_i is the error term. The outcomes are measured at the individual level. The coefficients of interests are the estimated treatment effect of the social norm video intervention in isolation (β_N), the effect of the educational video on carbon pricing (β_P), and the interaction effect of combining information on social norms and carbon pricing (β_{NP}). **Covariates** that will be progressively added to the regression equation include:

- Sociodemographic characteristics: e.g., age, gender, ethnicity, region, education, religion
- Economic situation: e.g., employment status, household income, perceived financial distress, dependency on fuel and energy
- Political attitudes: e.g., party affiliation, ideology, affective partisanship
- Baseline climate change views: e.g., beliefs and attitudes toward climate change, baseline support of carbon neutrality goals, perceived social norms

Note that by virtue of randomization, these covariates are statistically independent of treatment assignment and therefore not strictly required for our analyses. However, adding them as control variables can increase statistical power and correct for slight imbalances across groups due to finite sample size. Relatedly, we may

additionally present results that use sampling weights for better ex post representativeness of the U.S. adult population.

4.4 Hypotheses: Inference

The following statistical hypotheses will be tested for the main outcome variable using standard t-tests:

1. **Effect of social norms information:** Assignment to the norms video treatment has no effect on the outcome relative to assignment to the control group, $\beta_N = 0$.
2. **Effect of explaining carbon pricing:** Assignment to the educational video intervention group has no effect on the outcome relative to assignment to the video control intervention group, $\beta_P = 0$.
3. **Interaction effect:** Combining information on social norms and carbon pricing has the same effect on the outcome as the sum of the norm and educational effects, $\beta_{NP} = 0$.

4.5 Heterogeneity of Treatment Effects

We will also test for heterogeneity of treatment effects based on individual characteristics and pre-intervention attitudes and beliefs. The outcome variables will remain the same as above and we will test for heterogeneity by interacting the treatment indicators with interacting variables of interest.

The following factors will be explored in relation to their impact on the heterogeneity of treatment effects:

- **Socio-economic background.** Specifically, we will test for heterogeneity by age, gender, education, and household income.
- **Political affiliation.** We will examine differential effects of the treatment across political affiliations. An indicator variable equal to one for those that are Republican or leaning Republican will be used to partition the sample.
- **Baseline climate views.** In particular, we will check whether the treatment effects are more (or less) pronounced among people who are more alarmed about anthropogenic climate change as compared to people who are skeptical about climate change.
- **Misperceptions and uncertainty of beliefs.** If the effects of our interventions are partly driven by (rational) belief updating, we could expect to find stronger effects for individuals with larger misperceptions and/or uncertainty in pre-intervention beliefs. We will therefore test for heterogeneous

effects with regard to misperceptions of (and uncertainty about) carbon neutrality norms, as well as uncertainty and lack of previous knowledge about carbon pricing.

- **Trust in institutions.** Varying levels of trust can lead to discounting of new information. Specifically, in relation to carbon pricing and donations to organizations supporting carbon pricing, distrust of both scientists and politicians and associated institutions can lead to a suspicion that experts do not state the costs of a carbon tax or that governments would break their promise to pay the dividend (Douenne & Fabre, 2022; Beiser-McGrath et al., 2019).
- **Personality characteristics.** In particular, we test if information on social norms have larger effects on people with stronger compliance values and social comparison orientation. Furthermore, we will check if ambiguity attitudes and universalism values moderate the effect of information on belief updating.

5. Potential Mechanisms:

To the extent that we will observe effects of our information treatments, it will be discussed and explored which behavioral mechanisms may be driving our empirical results, i.e., through which channels the video interventions affect our outcome variables of interest. Note that different behavioral channels may be at play for different treatment condition. Some potential mechanisms are listed below:

- **General climate-related beliefs and attitudes.** One potential channel through which our information videos might affect self-reported and incentivized policy support is by changing individual's general beliefs and attitudes toward climate change. For example, if the treatments lead to an increase in understanding or belief in anthropogenic climate change and its consequences, individuals would also become more willing to support pro-environmental causes and policies. Although we believe that this is unlikely to be the main channel, we will test for this by comparing effects on post-intervention beliefs about anthropogenic climate change (compared to baseline climate beliefs) across treatment groups.
- **Perceived social acceptability/desirability.** A slightly more subtle effect of our Norm intervention may be that subjects change their views on the degree to which coming out in favor of environmental protection and climate policies is perceived as socially accepted or desired. This also links to the concepts of pluralistic ignorance and conditional cooperation, and it could change individual's behavior even holding constant their private beliefs.
- **Economic reasoning about carbon pricing:** The educational video on carbon pricing is aimed at improving people's understanding of how carbon pricing (with cash transfers) can reduce carbon

emission and which other economic and distributional effects may be at play. Thus, we will examine if the treatments have an effect on post-intervention beliefs and reasoning about different aspects of carbon pricing.

- **Perceived political feasibility.** The effect of the video interventions on policy support could be driven by individual beliefs regarding the feasibility of the U.S. achieving carbon neutrality. In particular, people may increase their optimism about the possibility of (bipartisan) political efforts against climate change and thus become more willing to donate to the carbon pricing organization. We will therefore test if our interventions induce subjects to be more convinced about the feasibility of the carbon neutrality goal. Similarly, receiving information about the workings of carbon pricing could increase people's trust in government's ability to address climate change.
- **Motivated beliefs and/or information search.** Respondents may seek information or update their beliefs in a way that is consistent with their initial views or values instead of integrating new information for a more accurate conclusion. The latter can also work through motivated memory and selective recall (Zimmermann, 2020). Motivated cognition may be one potential channel through which interaction effects between the Norm and the Policy treatment can arise. To test this, we will compare subjects' engagement with (and perceptions of) the carbon pricing explainer in Norm+Policy compared to the Policy group, as well as their post-intervention reasoning about carbon pricing. We will further explore these channels through views on climate change, measures of climate denial, and conspiracy mentality. Note that it may be difficult to cleanly separate motivated reasoning from rational (Bayesian) behavior.

References

- Beiser-McGrath, L. F., & Bernauer, T. (2019). Could revenue recycling make effective carbon taxation politically feasible?. *Science advances*, 5(9), eaax3323.
- Bordalo, P., Coffman, K., Gennaioli, N., & Shleifer, A. (2016). Stereotypes. *The Quarterly Journal of Economics*, 131(4), 1753-1794.
- Carattini, S., Carvalho, M., & Fankhauser, S. (2018). Overcoming public resistance to carbon taxes. *Wiley Interdisciplinary Reviews: Climate Change*, 9(5), e531.
- Douenne, T., & Fabre, A. (2022). Yellow vests, pessimistic beliefs, and carbon tax aversion. *American Economic Journal: Economic Policy*, 14(1), 81-110.

Falk, A., Becker, A., Dohmen, T., Enke, B., Huffman, D., & Sunde, U. (2018). Global evidence on economic preferences. *The Quarterly Journal of Economics*, 133(4), 1645-1692.

Frederick, S. (2005). Cognitive reflection and decision making. *Journal of Economic perspectives*, 19(4), 25-42.

Howe, P. D., Mildenerger, M., Marlon, J. R., & Leiserowitz, A. (2015). Geographic variation in opinions on climate change at state and local scales in the USA. *Nature climate change*, 5(6), 596-603.

Roets, A., & Van Hiel, A. (2011). Item selection and validation of a brief, 15-item version of the Need for Closure Scale. *Personality and individual differences*, 50(1), 90-94.

Thomson, K. S., & Oppenheimer, D. M. (2016). Investigating an alternate form of the cognitive reflection test. *Judgment & Decision Making*, 11(1).

Van Boven, L., Ehret, P. J., & Sherman, D. K. (2018). Psychological barriers to bipartisan public support for climate policy. *Perspectives on Psychological Science*, 13(4), 492-507.

Zimmermann, F. (2020). The dynamics of motivated beliefs. *American Economic Review*, 110(2), 337-61.