RATIONALITY & COMPETITION
CRC TRR 190

Fairness Views and Beliefs about the International Distribution of Climate Change Costs

Davide Domenico Pace (LMU Munich, CesIfo)
Johanna Kober (LMU Munich)

Discussion Paper No. 584

August 30, 2026

Collaborative Research Center Transregio 190 | www.rationality-and-competition.de
Ludwig-Maximilians-Universität München | Humboldt-Universität zu Berlin
Spokesperson: Prof. Dr. Klaus M. Schmidt, Ludwig Maximilian University of Munich, 80539 Munich, Germany
info@rationality-and-competition.de

Fairness Views and Beliefs about the International Distribution of Climate Change Costs *

Davide Domenico Pace[1] and Johanna Kober[2]

[1]LMU Munich, CESIfo
[2]LMU Munich

August 30, 2026

Abstract

International climate negotiations focus on three distinct climate change costs: mitigation, adaptation and loss and damage. This paper measures voters' fairness views about how these costs should be allocated between nations, and investigates whether fairness views and beliefs about a country's responsibility matter for policy support. In an online survey experiment, the paper finds that fairness views vary with the type of cost, that voters underestimate their country responsibility, and that fairness views and beliefs jointly predict voters' climate policy preferences. The paper also implements tailored treatments to test whether fairness views and beliefs causally affect policy support, but it finds no evidence of a causal link. The results highlight that, to be consistent with citizens' fairness ideals, the international distribution of mitigation costs should differ from the one of adaptation and loss and damage.

JEL Classification codes: D63, D91, F53, H41, Q54
Keywords: Climate Change, International Negotiations, Policy Support, Fairness, Misperceptions

*We thank participants at seminars and workshops at Kiel University, LMU Munich, Macquarie University, Monash University, Tilburg University, University of Bologna, and University of Queensland. Support by Deutsche Forschungsgemeinschaft through CRC TRR 190 (project number 280092119) is gratefully acknowledged.

1 Introduction

As the globe is warming, international negotiations are expanding in scope. Countries keep discussing how to mitigate greenhouse gas emissions. However, they are increasingly haggling over how to share the costs of adapting to a warmer world and over who should foot the bill for the losses and damage from the irreversible consequences of climate change, including extreme weather events. These negotiations involve substantial amounts. During COP29 in 2024, the UN's annual climate summit, countries agreed to mobilise $1.3 trillion per year in climate finance for developing nations alone by 2035.

Despite the high stakes, we have only a partial understanding of how people think it is fair to share the mounting burden of climate change across countries. Existing evidence is confined to mitigation. For example, most people believe that countries that emitted more in the past should pay a larger share of the global mitigation costs (Schleich et al., 2016). But do people apply the same fairness principles to adaptation and loss and damage? And do these fairness views matter for the climate policies people support? The answers to these questions would help us understand the political feasibility constraints that negotiators face and identify which additional data we need to better characterise them.

This paper addresses these two questions with a survey experiment run with a UK online sample. The survey first measures participants' fairness views with a vignette study. In each vignette, participants see a country and need to indicate its fair contribution to global mitigation, adaptation, and loss and damage costs. For each country, the participants see five attributes: current emissions, cumulative historical emissions, GDP, population, and vulnerability to unmitigated climate change. With their answers, the participants reveal the weight they give to each attribute, separately for each cost. The survey distinguishes among the three climate costs because people often look at inequalities one dimension at a time rather than at aggregate outcomes (Exley and Kessler, 2024), meaning that how each cost is distributed likely matters for public support for international agreements. The vignette study is designed so that concerns about conditional cooperation cannot influence the recovered fairness views.

We find that the weight people give to the country attributes varies with the climate costs. Participants consider current emissions more important for the fair allocation of mitigation costs than for the allocation of adaptation and loss and damage costs. Conversely, they consider a country's vulnerability less important for the distribution of the mitigation costs than for the other costs. These comparisons survive strict multiple-hypothesis testing corrections. The results indicate that to understand whether citizens would consider a climate agreement fair, we need to estimate their fairness views for adaptation and loss and damage separately from their fairness views for mitigation.

We designed the rest of the survey to connect fairness views and policy preferences.

To this end, we elicit the participants' beliefs about the UK attributes relevant to determining its fair contributions. These beliefs matter because misperceptions can affect public support: a British voter might endorse a polluter-pays principle but think that the UK should pay little, as she wrongly believes that its historical emissions are low. The survey finds systematic misperceptions. More than 80% of participants underestimate the UK's share of cumulative emissions, GDP, and population, indicating that most UK citizens perceive their country as less responsible for climate change than it is. We then combine fairness views and beliefs data and estimate, at the individual level, what participants consider the UK's fair contribution toward each climate change cost. We find that these estimates positively predict climate policy preferences, including incentivised choices over how to allocate funds between the UK government and international climate organisations. The predictive power is specific to international agreements: in an obfuscated follow-up, we find no significant relationship with support for the current UK unilateral climate policies.

Finally, we test whether fairness views and beliefs causally affect climate policy preferences. To maximise power, the treatment manipulations are stratified. Moreover, we tailored the treatments using participants' survey responses to identify the interventions most likely to yield the largest treatment effects. However, we don't find evidence of a causal relationship. Treatments intended to shift fairness views fail to move them and, consistently, don't affect support for climate policies. Treatments that shock beliefs successfully move posteriors toward the truth, but this shift in beliefs doesn't translate into a change in policy preferences. This last result, consistent with previous null findings surveyed below, suggests that the link between beliefs and policy preferences is weaker than it is assumed in many models of public support for economic policies, including the one we wrote to organise our analysis.

This study contributes to two strands of literature. First, it relates to the growing body of work on people's support for economic policies—see Stantcheva (2025) for a review. Part of this literature focuses on support for unilateral climate policies (Carattini, Kallbekken and Orlov, 2019; Douenne and Fabre, 2022; Mildenberger et al., 2022; Andre et al., 2024; Woerner et al., 2024; Dechezleprêtre et al., 2025) and international climate agreements (Bechtel and Scheve, 2013; Fabre, Douenne and Mattauch, 2025). Moreover, some papers specifically examined fairness views on how the burden of mitigating climate change should be distributed (Lange et al., 2010; Schleich et al., 2016; Meilland, Kervinio and Méjean, 2025; Fabre, Douenne and Mattauch, 2025) or investigate the moral underpinning of voters' support for domestic and international climate policies (Stötzer and Zimmermann, 2026). We advance this literature in two ways. First, while current evidence is concentrated on mitigation, we show that fair cost-sharing rules vary with the cost type. Second, we recover respondent-specific implicit cardinal weights for several country attributes, which we can connect to beliefs to predict policy support and develop tailored treatments. These exercises are not possible with existing studies, as they focus on specific policies or elicit only ordinal, rather than cardinal, measures of fairness.

Second, we extend to the international distribution of climate costs the methods used to study fairness views and preferences for income redistribution (Kuziemko et al., 2015; Alesina, Stantcheva and Teso, 2018; Cappelen et al., 2007, 2011; Fischbacher et al., 2023; Amasino, Pace and van der Weele, 2024; Andre, 2025). The misperceptions about the UK mirror those observed when measuring people's beliefs about their position in the income distribution (Fehr, Mollerstrom and Perez-Truglia, 2022; Hvidberg, Kreiner and Stantcheva, 2023). Moreover, despite our tailored and stratified treatment, the null result from beliefs correction on policy preferences replicates the limited effects observed after similar interventions in the literature (Alesina, Stantcheva and Teso, 2018; Hoy and Mager, 2021; Fehr, Mollerstrom and Perez-Truglia, 2022). However, we document that people average across different fairness criteria for climate change, while the experimental income-redistribution literature finds that most people endorse a single criterion (Cappelen et al., 2007; Amasino, Pace and van der Weele, 2024; Bortolotti et al., 2025). This last result indicates that to study climate fairness views, we need methods that, like the one we develop, generate cardinal estimates of fairness views; these methods are essential for identifying how people trade off different ethical considerations.

2 Theoretical Framework

In this section, we present the theoretical framework we use to organise the analysis and to design the treatments. The framework describes the problem of a citizen who must decide what share of global climate change costs she would like her government to pay. The framework is inspired by the models written to study people's fairness views about income redistribution (Konow, 2000; Cappelen et al., 2007, 2013; Bortolotti et al., 2025). A key idea that we borrow from that literature is that abstract *fairness principles* can be described as rules prescribing a given proportional allocation of some resources or costs. Instead, an agent's *fairness views* can be described as the vector of weights she gives to the fairness principles.

To apply this conceptualisation of fairness to climate change, we assume that the citizen chooses what she would like her government to do as part of a binding climate agreement to which every country in the world participates. By introducing the agreement, we ensure that the fairness views we define (and then measure in the survey) are relevant to the cost-sharing rules of an international treaty, rather than those that should apply to unilateral climate actions. Different fairness views might apply to those two contexts: citizens who think their country should do a lot within an agreement might believe that their country should do nothing if the other nations are not pulling their weight.

2.1 The framework

Define as X_m, X_a, and X_l the total global cost of climate change mitigation, adaptation, and loss and damage. $X_k \in \mathbb{R}_+$ for $k \in \{m, a, l\}$. Assume that there is an international agreement that disciplines how much of each of these costs every country in the world bears. The citizen needs to decide which share x_m, x_a, and x_l of these costs she would like her country to pay. Consistently with the experimental design, we measure the shares in permille so that $x_k \in [0; 1000]$ for $k \in \{m, a, l\}$. In making her choice, the citizen trades off the selfish desire to minimise the cost to her country against the disutility from supporting a cost allocation she considers unfair. In particular, the citizen utility function is:

$$U(x_m, x_a, x_l) = -\sum_{k\in\{m,a,l\}} X_k \left[x_k + g_k(\hat{r}_k - x_k)\right].$$

Within the sum, the first term reflects the utility losses from supporting her country spending funds to cover the climate change costs. The second component is the disutility cost from supporting an unfair international allocation of the burdens from climate change. The functional form for U assumes that the utility is separable in the x_ks. We make this assumption because recent evidence indicates that people's equity views are narrowly framed, meaning that people judge the fairness of a multidimensional allocation one dimension at a time rather than considering the overall outcome (Exley and Kessler, 2024).

We will now discuss the two components of the utility function in detail.

Self-interest. The citizen derives disutility from supporting her government spending money to cover the climate change costs. This might be because she expects (possibly naively) that the cost of climate change will be covered anyhow, with or without her government's contribution, so she would prefer that other countries foot the bill. Otherwise, it might be because the citizen's private benefits from addressing climate change are lower than the private costs she would bear due to the increased taxes needed to finance the relevant actions.

Fairness For every $k \in \{m, a, l\}$, the term $g_k(\hat{r}_k - x_k)$ indicates the disutility the citizen suffers if she supports her country to pay a share x_k of X_k smaller than $\hat{r}_k \in$

$[0;1000]$, the amount she considers fair. We assume that $g_k : [-1000;1000] \rightarrow \mathbb{R}_+$, $g_k(\hat{r}_k - x_k) = 0$ if $x_k \geq \hat{r}_k$ and $g_k(\hat{r}_k - x_k) > 0$ if $x_k < \hat{r}_k$. $g_k(.)$ is twice differentiable, increasing and strictly convex if $x < \hat{r}$.

The citizen calculates her country fair shares for cost $k \in \{m, a, l\}$ according to:

$$\hat{r}_k(\boldsymbol{\beta}_k, \hat{\boldsymbol{s}}) = \beta_k^0 + \beta_k^\alpha \hat{s}^\alpha + \beta_k^\omega \hat{s}^\omega + \beta_k^\gamma \hat{s}^\gamma + \beta_k^\nu \hat{s}^\nu + \beta_k^\iota \hat{s}^\iota. \tag{1}$$

Where s^j indicates how large her country is on the climate-relevant dimension j, with $j \in \{\alpha, \omega, \gamma, \nu, \iota\}$. The s^js are expressed as permille of the global total (hence, $s^j \in [0;1000]\,\forall\, j$). We allow the citizen to have wrong perceptions of how large her country is along these dimensions. For this reason, we assume that she calculates the fair shares $\hat{r}_k$ using her subjective point beliefs $\hat{s}^j$ rather than using the true values s^j. Instead, β_k^j with $j \in \{\alpha, \omega, \gamma, \nu, \iota\}$ indicates the weight the citizen gives to dimension j when calculating the fair share for cost k. The calculations include the constant β_k^0, which captures the country's perceived fair contribution independent of its characteristics.

We define the β weights in Equation 1 as the citizen "fairness views". We assume that these fairness views are not country specific. That is, we assume that the citizen would have used the same weights to calculate the fair shares of any other country, based on how large that country is across the five dimensions.

The five country characteristics operationalise three abstract fairness principles often discussed in climate policy discussions (Fleurbaey et al., 2014). The first is the *polluters pay principle*, which posits that countries that generate the most emissions should pay a larger share of the costs. We assume that the citizen implements this principle by using her country's share of historical global emissions (s^α) and its share of current global emissions (s^ω). We allow for current emissions to be weighted differently than historical ones because some ethicists believe that countries can be held accountable only for the emissions they generated after it became clear that greenhouse gases were altering the climate (Gardiner, 2004). The second principle is the *ability-to-pay principle*, which posits that richer countries should pay a larger share of the climate costs. We operationalise this principle using the country's current share of the global GDP (s^γ). Finally, the *beneficiaries pay principle* says that countries that benefit more from stopping climate change should pay more. We assume that the citizen applies this principle using the country share of the global population (s^ν) and its share of the global losses due to climate change that would occur in that country if global warming were not mitigated (s^ι).

The solution to the citizen problem. To figure out which share of the climate change costs she would like her country to pay, the citizen maximises $U(x_m, x_a, x_l)$.

Assuming an interior solution and noticing that the utility function is concave, there exists a unique $(x_m^*, x_a^*, x_l^*) \in \prod_{k\in\{m,a,l\}}[0; \hat{r}_k]$ that satisfies the first order conditions:

$$\frac{\partial g_k(\hat{r}_k - x_k)}{\partial(\hat{r}_k - x_k)} = 1, \, \forall \, k \in \{m, a, l\}. \tag{2}$$

Note that the solution to the citizen's problem doesn't depend on the magnitudes of the climate change costs X_m, X_a, and X_l, a result we will exploit in the experimental design.

Comparative statics. In the experiment, we will have treatments that aim at shifting the participant's policy preferences by shocking her beliefs or her fairness views. These treatments will focus on preferences regarding adaptation. Hence, to optimally design these treatments, it is useful to see how x_a^* changes according to the model when the beliefs and fairness views move.

If we implicitly derive the FOC for an arbitrary $\hat{s}^j$, we obtain:

$$\frac{\partial x_a^*}{\partial \hat{s}^j} = \beta_a^j, \quad \forall \, j \in \{\alpha, \omega, \gamma, \nu, \iota\}. \tag{3}$$

Hence, the effect of a belief shift about a UK characteristic is proportional to the fairness weight given to that dimension. It follows that the predicted change in x_a after a full correction of the citizen's misperceptions about s^j is:

$$\Delta_j x_a := \beta_a^j (s^j - \hat{s}^j). \tag{4}$$

If instead, we derive the FOC for an arbitrary fairness weight, we obtain:

$$\frac{\partial x_a^*}{\partial \beta_a^j} = \hat{s}^j, \quad \forall \, j \in \{\alpha, \omega, \gamma, \nu, \iota\}. \tag{5}$$

Hence, the effect of a treatment that increases the weight given to a country dimension is proportional to the citizen beliefs about how large the UK is on that dimension.

What we should measure according to the model. The model clarifies the variables we need to measure to answer our two main research questions. In particular, we need a survey that estimates the citizen's three vectors of fairness weights $\boldsymbol{\beta}_m$, $\boldsymbol{\beta}_a$, and

$\boldsymbol{\beta}_l$, the citizen's vector of beliefs $(\hat{s}^\alpha, \hat{s}^\omega, \hat{s}^\gamma, \hat{s}^\nu, \hat{s}^\iota)$, and the citizen's preferences for her country contributions within the global climate treaty (x_m^*, x_a^*, x_l^*).

We can then use the estimated β-weights to test whether voters' fairness views differ depending on the type of climate change cost. Furthermore, by combining fairness views and beliefs, we can test whether these primitives predict support for climate policies. Finally, we can use the participants' answers to design tailored treatments and test whether beliefs and fairness views causally affect policy support. The next section explains how we measure these variables and design the treatments.

3 Survey Design and Empirical Strategy

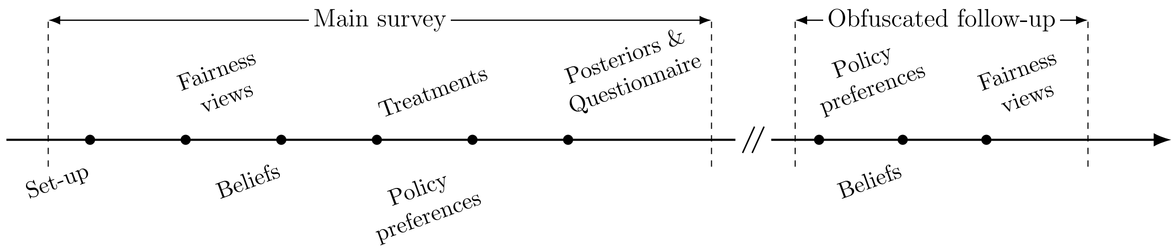


Figure 1: Timeline of the survey.

Figure 1 displays the timeline of the survey. In this section, we describe each part in detail and connect the elicitations to the theoretical framework. At the end, we describe the survey implementation and the measures we took to make sure AI bots could not complete the survey. Online Appendix Aprovides further details and screenshots of the experiment's design. The instructions are in Online Appendix D.

3.1 Set-up.

The main survey opens with a consent form and a basic demographic questionnaire (age, gender, political affiliation, and education level).

After that, we describe the background information the participants need to provide meaningful answers. In particular, we describe what climate change is and its three costs: mitigation, adaptation, and loss and damage. Next, we ask the participants to imagine that all the nations in the world have signed a very detailed and binding climate treaty. The treaty specifies all the required investments and policy changes needed to limit global warming to 2°C and to adapt to this level of warming. It also establishes how much each country should pay towards the combined net global cost of these actions and of the remaining climate loss and damage. We specify that the actions listed in the treaty are those that minimise the total net cost of a 2°C global warming, and that the costs to be shared are net of any economic benefits those actions produce. We don't

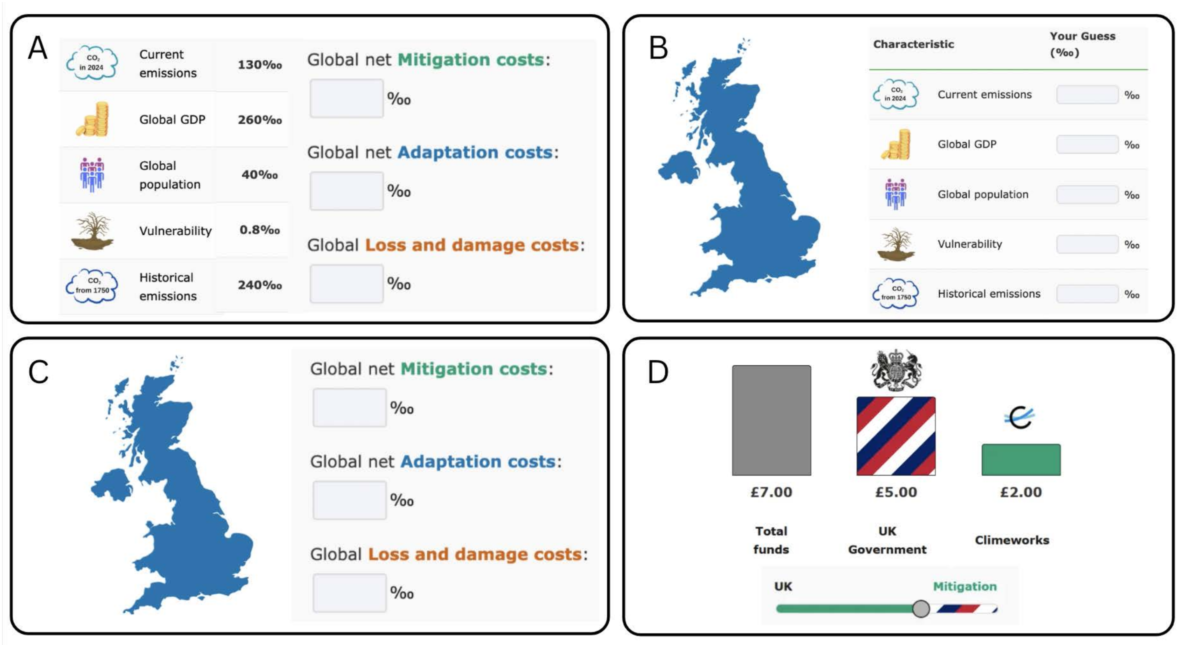


Figure 2: Illustration of the main elicitation interfaces. A, vignette study; B, beliefs elicitation; C, elicitation of the share of the climate change cost the participants would make the UK pay within the international treaty; D, funds allocation. See Online Appendix A.1 for the screenshots.

specify the total cost because, in the theoretical framework, this information does not influence voters' policy preferences (see Equation 2).[1]

Furthermore, the set-up introduces the symbol and concept of permille with two examples. The study uses permille rather than percentage so that participants can indicate small fair shares with minimal need to use decimal cyphers.

3.2 Fairness views elicitation.

In the fairness views section of the survey, we explain to the subjects that we would like to understand how the cost-sharing part of the international agreement should look for them to consider it fair. In particular, we explain that we would like them to consider five country characteristics that some people believe are relevant for allocating costs. Consistently with our theoretical framework, the five characteristics are: a country's cumulative CO_2 emissions since 1750, a country's CO_2 emissions in 2024, a country's GDP, a country's population, and a country's expected losses if climate change is not mitigated —we will often refer to this last characteristic as a country's "vulnerability". We express these characteristics as permille of the global total so that they are measured in the same unit. We briefly describe the possible connection between the five country

[1]Moreover, we couldn't find precise estimates of these costs in the literature. The estimates of the economic costs of climate change from Integrated Assessment Models have been criticised for being highly sensitive to the chosen assumptions (Pindyck, 2013).

characteristics and the cost allocation with sentences like: "Some argue that richer countries should pay a larger share of the climate change costs". We randomise at the participant level the order in which the characteristics are presented; this presentation order is preserved for the entire experiment for each participant.

After these explanations, the participants report their fairness views in two ways: first, through a ranking task, and then in a vignette study.

Ranking task. For the first fairness views elicitation, the participants rank the five characteristics from the most to the least important to determine a fair distribution of the costs. They produce three rankings: first for mitigation, then for adaptation, and finally for loss and damage.[2] After each ranking, they can indicate whether any of the characteristics should not be used to allocate that cost. A final interface summarises the participants' choices. There, the subjects can make any adjustment they wish. We introduced the ranking task to familiarise participants with the country attributes and the climate change costs.

Vignette study. Next, the subjects complete a vignette study. In a vignette, the participants see a country defined by the five characteristics, and they must indicate which permille of each climate cost it would be fair for this country to pay under the international agreement. The participants complete 14 of these vignettes; figure 2A illustrates one. To help the participants interpret the numbers in the vignettes, the instructions explain that if all 195 countries in the world had been identical, each would have been as large as 5.1‰ on every dimension. This number also appears on the vignette interface.

We built the vignette deck so that the distribution of the vignettes was as close as possible to the distribution of the nations in the world. In this way, we are sure that our results are practically relevant. For example, if we find that the participants have different fairness views for mitigation and adaptation, we can be confident that this result stems from differences in the perceived fair contributions of real countries, rather than from unrealistic parts of the parameter space. However, with this approach, the deck inherits the real-world correlation structure among characteristics. This correlation decreases the statistical power to test whether the characteristics received different weights depending on the cost type, especially for historical emissions and GDP. We

[2] In the elicitations, as in the instructions, we did not randomize the order of the three costs, since doing so would break the intuitive relationship between them: only once a problem has been mitigated does it make sense to ask how to adapt to it, and only then to consider its irreversible consequences. This ordering of mitigation, adaptation, and loss and damage also matches the chronological order in which these topics entered international negotiations.

believe this is a price worth paying to ensure that we collect data in the empirically relevant parameter space.

Specifically, the joint distribution of historical emissions, current emissions, GDP, and population in the deck is the same as the world one.[3] The only deviation is that we excluded the countries that are smaller than 1.5‰ on all 4 dimensions. As these countries are tiny, their contributions within the agreement will be similarly small under most fairness views. After excluding these countries, we are left with 100 nations.

As per vulnerability, when we ran the study, there was no country level estimate we could use.[4] For this reason, we assumed that there are ten possible vulnerability levels, which are equal to the deciles of the marginal distribution of world population (expressed in permille). We modelled the distribution of vulnerability on the one of population because the two marginal distributions would coincide in expectation if the loss and damage from climate change were i.i.d. for every person in the world. This i.i.d. distribution amounts to an agnostic prior about vulnerability.

Finally, we took the Cartesian product of the 100 nations and the 10 assumed vulnerability levels to create a deck of 1000 vignettes, each sharing four attributes with a real nation and featuring one of the 10 possible vulnerability levels. This way of building the deck is inspired by the Marginal Population Randomisation Design introduced in de la Cuesta, Egami and Imai (2022). As a final adjustment, we rounded the numbers in the vignettes to make the task easier for the participants.

Using the vignette responses, we can estimate the average $\boldsymbol{\beta}$ fairness weights in the population from Equation 1 . Let $y_{ick} \in [0; 1000]$ be the permille of climate cost k that participant i states it would be fair for hypothetical country c to pay; let $s_c^j \in [0; 1000]$ be the share of characteristic $j \in \{\alpha, \omega, \gamma, \nu, \iota\}$ that country c displays in the vignette. We can estimate:

$$y_{ick} = \sum_{\kappa \in \{m,a,l\}} \mathbb{1}_{\{\kappa = k\}} \left[\beta_\kappa^0 + \sum_{j \in \{\alpha, \omega, \gamma, \nu, \iota\}} \beta_\kappa^j \, s_c^j \right] + \varepsilon_{ick} \tag{6}$$

Where $\mathbb{1}_{\{\kappa = k\}}$ is an indicator function equal to 1 if the cost type κ matches the cost type k of the observed outcome y_{ick}. Because the vignettes display the characteristics directly, participants do not rely on their subjective beliefs, so the shares shown, s_c^j, replace the beliefs $\hat{s}^j$ in Equation 1. The term ε_{ick} is an idiosyncratic, mean-zero

[3] To calculate the country shares of these variables, we used 2024 data from the World Bank (GDP and population), and Our World In Data (emissions).

[4] Several risk indices exist that indicate national vulnerability to climate change (Notre Dame Global Adaptation Initiative , ND-GAIN; INFORM, 2026; Auer Frege et al., 2025; Carleton et al., 2026). However, it is not possible to derive from these estimates the distribution of the expected cost of unmitigated climate change at the national level, the input we would need for our vignettes.

error. For some analyses, we estimate Model 6 at the individual level to recover the participant's specific vector weights $\hat{\boldsymbol{\beta}}_{mi}$, $\hat{\boldsymbol{\beta}}_{ai}$, and $\hat{\boldsymbol{\beta}}_{li}$.

3.3 Beliefs elicitation.

Here we elicit participants' beliefs about how big the UK is on the five dimensions we use for the fairness views elicitation. The questions are not incentivised, as most answers can be easily googled. The answers reveal the vector $(\hat{s}^{\alpha}_{UK}, \hat{s}^{\omega}_{UK}, \hat{s}^{\gamma}_{UK}, \hat{s}^{\nu}_{UK}, \hat{s}^{\iota}_{UK})$ for each participant. Figure 2B illustrates the elicitation interface.

3.4 Policy Preferences

We measure policy preferences in two ways. First, we invite the participants to think back about the international agreement we introduced in the set-up and to imagine that they have the power to determine how the three climate change costs are allocated across countries. We specify that the subjects may split the costs as they see fit, provided all costs are covered. We then ask them which share of each cost they would make the UK pay. In keeping with the rest of the survey, we elicit these preferences as a permille of the total of each climate change cost. Figure 2C illustrate the survey interface. Eliciting the policy preferences in this way, we recover for each participant the vector (x^*_m, x^*_a, x^*_l) as defined in our theoretical framework.

The second elicitation introduces monetary incentives at the cost of a looser relationship with the theoretical framework. It asks subjects to distribute some funds between the UK government and an international organisation. Figure 2D illustrate the elicitation interface. Lest the subjects who mistrust the government give everything to the international organisation, the task uses a concave budget curve (see Appendix A for details). There are three allocation choices. In the first, the organisation is involved in climate mitigation; in the second, in adaptation; and in the last, in loss and damage. One allocation of one participant every 20 was implemented.

3.5 Final questionnaire and retests.

The last part of the survey starts with a questionnaire on detailed demographics (including geographical information and migration background), the importance of climate change in voting decisions, preferences for conditional cooperation in climate policies, trust in the UK government, moral universalism, and zero-sum thinking. It then repeats some elicitations that the participants have already completed: it presents a shortened vignette study with 4 countries that the participants have already seen, and it elicits posterior beliefs about the UK characteristics.

3.6 Obfuscated follow-up

In the obfuscated follow-up, we measure support for the UK voluntary contributions to fight climate change. To enhance the obfuscation, the survey begins with questions about the participants' support for specific migration and pandemic-related policies. Then it moves to climate change. The first climate question asks whether the UK should reach net zero emissions before or after the pledged 2050 deadline. The second reports the UK's contributions to international climate change adaptation and asks whether the government should increase or decrease these contributions. The last states the UK's financial contributions to the UN's Loss and Damage Fund and asks whether these commitments should go up or down. The follow-up concludes with a retest of participants' beliefs and fairness views, using the ranking task. Before the ranking task, the obfuscation is lifted, and the survey reminds the participants of the relevant definitions.

To ensure that the participants do not immediately realise they are in a follow-up study, the follow-up survey was created in Qualtrics and uses the platform's default interface. Instead, the main survey was programmed with custom PHP code (see Section 3.8).

3.7 Treatment manipulations.

The treatments are introduced in the main survey, just before the policy preference elicitations. There are five of them. Two consist of a reflection task aimed at shifting views of fairness; two provide information aimed at correcting misperceptions about the UK; the last is a pure control in which participants skip the page where the other treatments are administered.

We included these manipulations to search for a proof of concept that beliefs and fairness views, the theoretical building blocks of our model, causally affect policy support. The manipulations focus on climate change adaptation, as we expected voters' fairness views on the topic to be more malleable because adaptation features less prominently in the news than mitigation efforts and the loss and damage from extreme weather events.

To maximise statistical power, the treatments are stratified by gender, age group, and vote in the 2024 UK parliamentary elections (Bai, Shaikh and Tabord-Meehan, 2025). In addition, we tailored the treatments at the individual level using the participants' beliefs and fairness views, as we explain in detail below. This tailoring doesn't hamper the identification: the treatment assignment is random and, hence, identical participants assigned to different treatments undergo different manipulations.

The reflection task treatments. In the two reflection treatments, the participants write arguments in favour of using a particular country characteristic to determine acountry's fair share of the global adaptation costs. The idea behind the manipulation is that producing such arguments should increase the weight the participants give to that characteristic when allocating the cost: people might find new reasons for why it should be used. Moreover, support for a moral statement increases with the time spent considering it (Pärnamets et al., 2015; Amasino, Pace and van der Weele, 2024).

In the *Max Reflection* treatment, the participants argue for the characteristic for which they believe the UK has the largest global share. Instead, in the *Min Reflection* treatment, they argue for the characteristic for which they believe the UK has the smallest global share. We design the treatments in this way because, according to Equation 5, the effect of increasing the weight a participant gives to a characteristic is proportional to his beliefs. Hence, we expect participants in the Max Reflection treatment to support higher UK spending on global adaptation costs compared to participants in the Min Reflection treatment.

The information treatments. The two information treatments aim at correcting one of the participants' misperceptions about the UK. The information displayed depends on a participant's beliefs and fairness views. In the *Upward Information* treatment, the information is chosen to maximise the share that the participant would like the UK to pay towards climate change adaptation. Instead, in the *Downward Information* treatment, it is chosen to minimise that share. We provide some information in every treatment so that the results are not influenced by the effect of receiving some information per se, as Imai et al. (2026) and Schulze-Tilling (2025) document can happen.

To identify which information to display to participant i, the experimental program first estimates her fairness views about adaptation $\hat{\boldsymbol{\beta}}_{ai}$ by fitting Equation 6 to her answers in the vignette study. Then it multiplies these estimated fairness views with the participant's misperceptions. That is, for each attribute j (except for vulnerability, for which we don't know the ground truth), it computes $\Delta_j \hat{x}_{ai} := \hat{\beta}^j_{ai}(s^j_i - \hat{s}^j_i)$. This product is an estimate of the effect of correcting the misperception about attribute j on the share x^*_a that the participant would like her government to pay towards the global adaptation costs (cfr., Equation 4 from the theoretical framework).

After computing these products, the experiment determines which information to display depending on the treatment to which the participant has been randomly assigned. In the Upward Information treatment, the experiment displays the information for which the product is the largest. I.e, it shows s^{j^+} where $j^+ = \text{argmax}_j \Delta_j \hat{x}_{ai}$. Vice versa, for the participants in the Downward Information treatment, the study shows the information for which the product is the smallest. I.e, it displays s^{j^-} where $j^- = \text{argmin}_j \Delta_j \hat{x}_{ai}$. Given this algorithm, the model predicts that, for participant i, the effect of being assigned to the Upward instead of the Downward Information treatment is given by:

$$\tau_{ai} := \max_j \{\Delta_j \hat{x}_{ai}\} - \min_j \{\Delta_j \hat{x}_{ai}\} > 0 \tag{7}$$

As τ_{ai} is positive by construction, we expect the participants in the Upward Information treatment to prefer the UK to pay a larger share of the global adaptation costs.

3.8 Sample and procedures

Sample. Our main sample comprises 1126 UK nationals recruited via Prolific between March and June 2026. The sample is representative of the UK population that voted in the 2024 UK parliamentary elections on gender and supported party.[5] Notwithstanding our attempt to collect a representative sample on age, people over 66 are underrepresented while people between 19 and 50 years are overrepresented. See the Appendix B1 for a demographic breakdown of the sample.

In total, 2827 participants consented to the terms of our study and therefore started the experiment. An additional 278 either didn't consent to the terms of the study or was screened out because they connected from a mobile device or were identified as a possible AI bot by our technical checks (See Appendix A.2 for details). Of the participants that consented, 1085 dropped out voluntarily, mostly as they struggled with the comprehension questions —less educated participants are more likely to give up in this phase of the study. Furthermore, 533 participants were screened out because they reported demographic data inconsistent with their Prolific profiles, 18 because they switched tabs more than 3 times, and 65 because they required more than 25 attempts on the comprehension questions. Overall, 1152 participants completed the study. Of those, we dropped 26, which, due to a bug in the program, managed to complete the study despite failing the technical checks. After this exclusion, we are left with 1126 people in the sample we use for the main analysis.

Participants took part in the follow-up, on average, 7 days after completing the main study. 87.74% of the invited participants finished this second survey.

Procedures and AI detection. On average, the main survey took 34 minutes and the follow-up 7 minutes. The participants earned 3.5£ and 0.75£ as completion rewards. To minimise attention costs, the instructions were presented over several slide

[5]Participants that didn't vote couldn't take part in the study. We used YouGov estimates to compute the target number of participants per demographic cell (YouGov, 2024).

decks with explanatory images complementing the written text. Eight detailed comprehension questions, mostly focused on the set-up instructions, checked participants' understanding. Participants had to answer all these questions correctly to continue with the experiment. Participants could go back to the instruction slides again if needed. To ensure high data quality, we screened out participants who needed more than 25 attempts to answer correctly.

We programmed the survey in PHP 8.4 and restricted access to devices with large screens, such as tablets or PCs. The programme has several technical safeguards to identify and screen out AI bots. For example, we screened out participants whose IP addresses were coming from a known data centre, and we checked their browsers for fingerprints revealing AI-controlled ones. We tested these safeguards by asking a variety of AI bots to complete the study (OpenAI agent, Claude, Claude for Chrome, Gemini, Automa extension, Axiom AI, Browserflow, UI Vision, and Perplexity). None of the bots we tested was able to generate a complete study submission. In particular, the OpenAI agents were immediately screened out because the program recognised the bot IP address as coming from a known data centre. Instead, Claude in Chrome stopped answering the study midway, possibly because the number of slides it had to process was too large. These tests were performed in March and April 2026. We describe these technical safeguards and tests in Appendix A.2.

We further look for AI usage with a CAPTCHA and three questions that a human participant should not see, but that a bot analysing the HTML code will notice. In total, 63 participants who completed the experiment failed the captcha and a different one answered one of the questions he should not have seen. We keep these 64 participants in our main sample, but in Appendix C.5.1, we show that our results replicate if we exclude them. In the follow-up, we applied Prolific's bot authenticity check to an open-ended question. This check flagged no participant as likely to be AI-generated.

4 Analysis

Part of the analysis was preregistered. The preregistration and a discussion of where we deviate from it are in Online Appendix B.

4.1 Fairness views

We start by analysing the vignette study. In each vignette, the participants indicated the share of the three climate change costs that would be fair for a country to pay. Figure 3 displays the coefficient we estimated when fitting the linear Model 6 to the participants' answers. The figure groups the coefficients by country characteristics. The

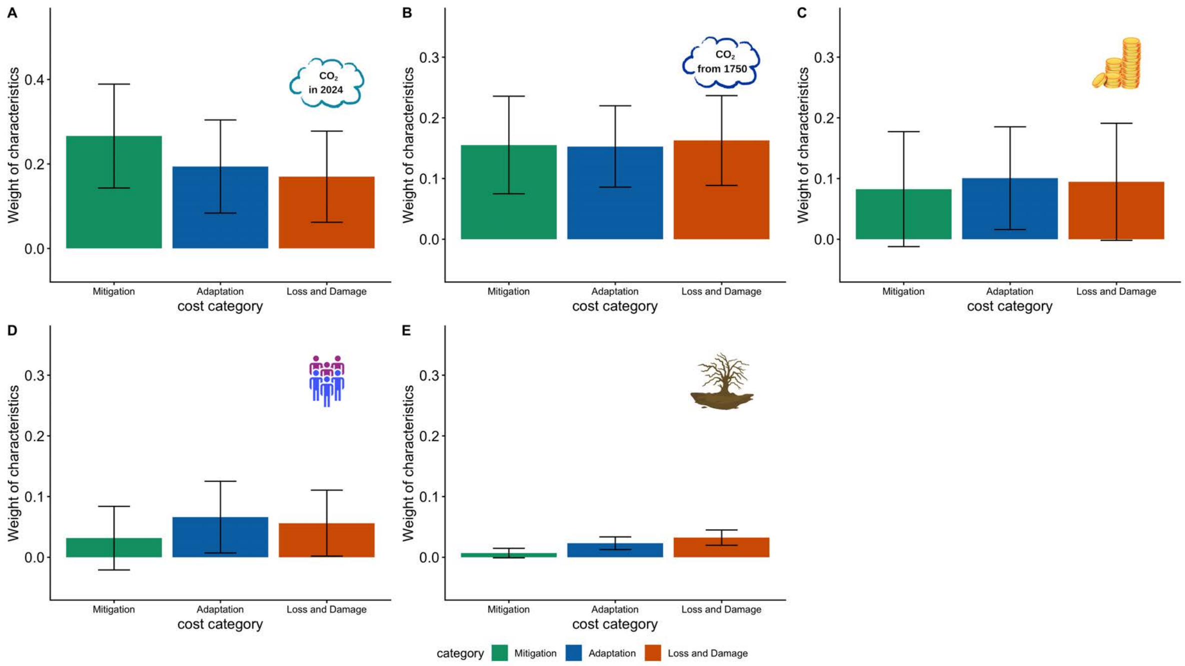


Figure 3: Estimated fairness weights by country characteristics.

Note: The bars show the estimated coefficients from the linear specification of Model 6. The bar height represents the average weight participants assign to a country characteristic when allocating a climate cost. The error bars represent 95% confidence intervals. The standard errors are clustered at the participant level.

height of the bars indicates the average weight participants assign to one characteristic when allocating a cost. For example, the green bar in the top-left panel indicates that, on average, the participants think that if a country generated 1‰ more of current global emissions, its fair contribution to mitigation costs should increase by 0.266‰ . The figure shows that our sample considers all five dimensions relevant for the allocation of the climate costs. Most of the coefficients are significant at the 5% level; three are significant at the 10% level (GDP-mitigation, GDP-loss_and_damage, and vulnerability-mitigation), and one is insignificant (population-mitigation). Table C.1 in the Online Appendix displays the regression output underlying Figure 3.

We can now ask whether the participants' fairness views vary with the type of climate change cost. A parsimonious way to answer the question is to test the joint hypothesis that the fairness weights estimated by the model are the same independently of the cost. Formally the null hypothesis is $\boldsymbol{\beta}_m = \boldsymbol{\beta}_a = \boldsymbol{\beta}_l$. An F- test rejects this null ($p < 0.001$). The rejection is not simply due to different baseline contribution levels by cost type: an F-test excluding the cost-specific constants still rejects the null hypothesis ($p < 0.001$). We conclude that participants use the five country characteristics differently, and, hence, that the fair allocation changes, depending on the climate change cost considered.

Which country characteristics drive the difference? Figure 3 suggests that, on average, the participants put more weight on current emissions for the allocation of mitigation costs than for the allocation of adaptation and loss and damage costs. Vice versa, it indicates that participants consider vulnerability less important for mitigation than for the other two costs. Formal t-tests confirm these visual impressions. The p-values are 0.022 and 0.002 when comparing the coefficients on current_emissions-mitigation with the ones for current_emissions-adaptation and current_emissions-loss_and_damage, respectively. The same tests for vulnerability return p-values of 0.002 and < 0.001. All these comparisons, except the first, remain statistically significant if we use a Bonferroni-corrected significance threshold to account for multiple hypothesis testing.[6] Overall, we conclude that the fairness views differences across costs are due to participants treating mitigation differently from the other two costs. We don't find evidence that participants think the fair allocation of adaptation costs differs from that of loss and damage costs.[7]

Individual level analysis. A common result in the experimental literature on fairness views about income distribution is that most people put all the weight on a single fairness principle (Cappelen et al., 2007; Bortolotti et al., 2025; Amasino, Pace and van der Weele, 2024). We can use our elicitations to test whether this result extends to the allocation of climate change costs. Since the original results focus on three fairness criteria (egalitarian, libertarian, and meritocratic), we conduct this analysis at the level of the three abstract fairness principles (polluters pay, ability to pay, beneficiaries pay). In this way, we avoid that any discrepancy in the results is just an artefact of a less coarse classification of possible fairness principles.

As a first step, we check if people subscribe to multiple fairness views using the ranking task. During this task, the participants could indicate whether some country characteristics should not be used to determine the fair allocations. We find that 87.12%, 87.03%, and 85.97% of participants believe that all three abstract fairness principles are relevant for the allocation of mitigation, adaptation, and loss and damage costs, respectively; only 1.87%, 1.87%, and 2.13% believe a single criterion should be used. These findings indicate that most people believe that the fair distribution of the climate change costs balances multiple criteria.

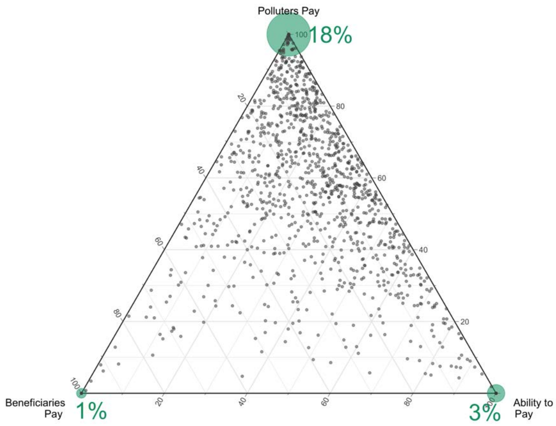
[6] With three cost types and five characteristics, we can run 10 independent tests (say, adaptation vs. mitigation, and loss and damage vs. mitigation); the last 5 tests (adaptation vs. loss and damage) are simply the difference of the 10 already conducted and carry no independent information. Hence, the Bonferroni correction requires reducing the significance threshold of a single comparison to $p < 0.05/10 = 0.005$.

[7] A possible concern with the analysis in this paragraph is that the differences in the weights of current emissions might be an artefact of the correlation structure in the vignette deck. For example, since current and historical emissions are correlated, the coefficient for current emissions can be influenced by nonlinear effects of historical emissions on fair shares. Appendix Table C.2 assuages such concerns by confirming the results with a model that flexibly controls for non-linear effects of the country characteristics on the fair shares. (Such concerns don't apply to the analysis of the vulnerability weights since, in the vignette deck, vulnerability is orthogonal to the other characteristics by design.)

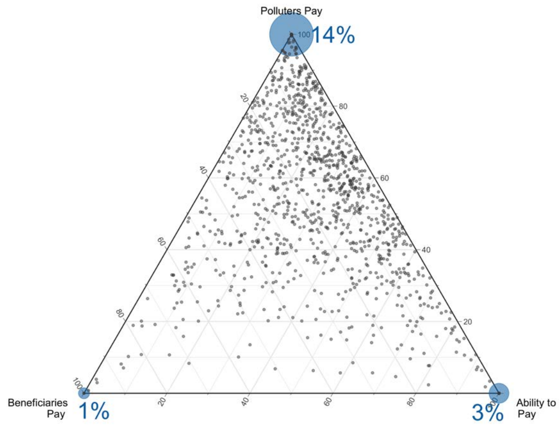
One concern with the ranking data is that the participants might say that a country characteristic should be used, but then place a negligible weight on it when determining the fair shares. To address this worry, we use the vignette study to estimate each participant's fairness weights. We proceed in the following steps. First, we estimate Model 6 at the individual level. Then we take the absolute value of the estimated weights —some participants put negative weights on certain characteristics, especially vulnerability. Finally, we sum up the resulting fairness weights related to the same abstract principle, and we normalise them by the sum of all absolute weights. The result of this algorithm is a vector of three entries that sum to 1 for each subject.

Figure 4 displays the distribution of the weights given to the three abstract principles on three simplexes, one for each climate cost. The participants who endorse a single principle are represented by points at the vertices. We find that 22% (mitigation), 18% (adaptation), and 16% (loss and damage) of participants are at most 20% away from one of the vertices, indicating that most of them believe that a combination of fairness principles should be applied to determine a country's fair share of the climate change costs.

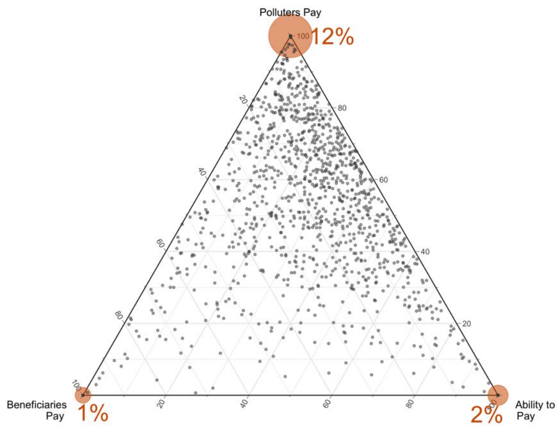
We conclude that both the ranking and vignette data indicate that voters balance multiple fairness principles to determine the fair allocation of climate change costs. This result contrasts with existing findings related to income redistribution. The result also underscores the importance of eliciting fairness views on a cardinal scale so that the way voters trade off the principle can be estimated.



(a)



(b)



(c)

Figure 4: Distributions of individual fairness weights

Note: The figure plots the individual-level, normalised, absolute fairness weight on separate simplexes from mitigation (Panel a), adaptation (Panel b) and loss and damage (Panel c). Each point represents one participant. The weights are aggregated at the level of the abstract fairness principle. The Polluters Pay axis displays the sum of the weights on current and historical CO_2 emissions. The Ability to Pay displays the weight on GDP, and the Beneficiaries Pay shows the sum of the weights on population and vulnerability.

4.2 Beliefs about the UK

We now analyse the participants' beliefs about the UK global share of current and historical emissions, GDP, population, and climate vulnerability. Table 1 reports the summary statistics for these beliefs; Figure 5 plots their cumulative distribution against the true value (except for vulnerability, for which it is not available).

We find widespread misperception. The median belief lies below the truth for all

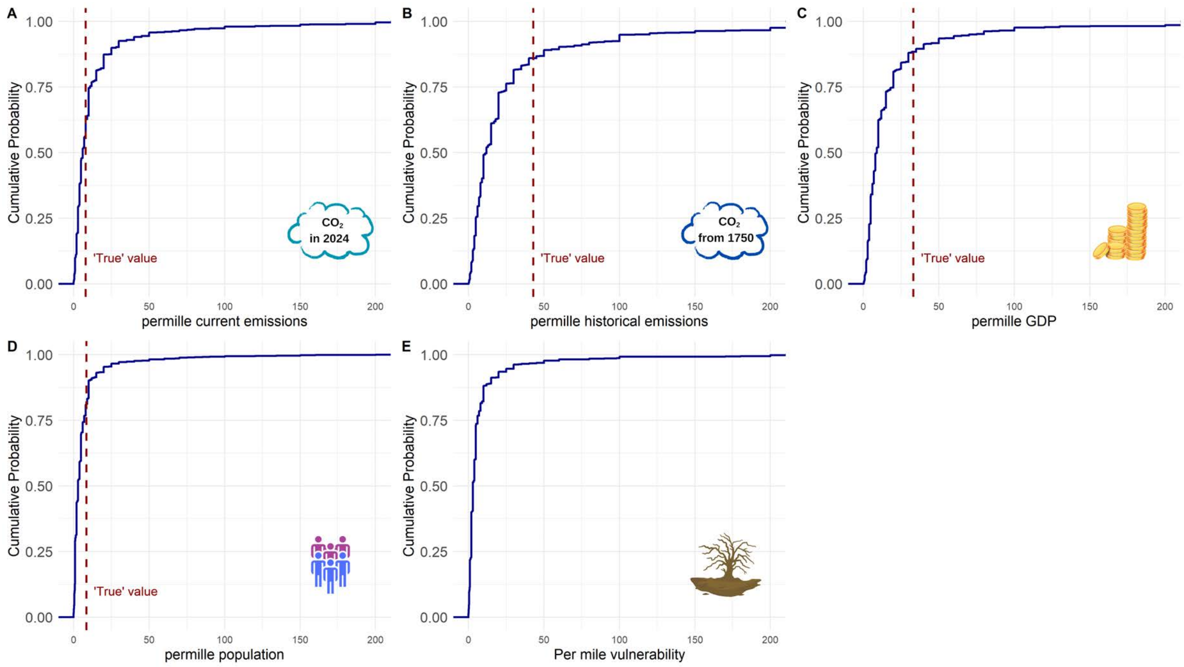


Figure 5: Beliefs about the UK's characteristics

Note: The figure plots the cumulative distribution functions of participants' beliefs regarding five UK characteristics: (A) current emissions, (B) historical emissions, (C) GDP, (D) population, and (E) vulnerability. The vertical red lines indicate the true values. No true value is shown for vulnerability, as no objective benchmark exists.

beliefs, with the percentage of underestimators going from 63% (current emissions) to 88% (GDP); the mean beliefs underestimate the truth for historical emissions, GDP, and population, but overestimate the UK's share of current emissions ($p < 0.001$ for tests comparing the mean beliefs to the truth for all the comparisons). These misperceptions cannot be fully explained by people's tendency to overestimate small numbers and underestimate large ones (see, for example Enke and Graeber, 2023). The true value for current emissions and population is very close. Yet the underestimation for population is much more pronounced.

Even though the participants struggled to guess how large the UK is on these dimensions, they had a good understanding of relative comparisons. 79% of participants correctly indicated that the UK's share of global historical emissions is larger than its share of current emissions. Moreover, 91% of participants indicated that the UK share of global GDP is larger than its share of the global population. The same number of participants recognise that the share of UK historical emissions is larger than its share of the global population. The last two comparisons show that the participants are aware that people in the UK are richer and have emitted more than the average human.

Overall, the beliefs elicitation indicates that most UK voters underestimate how large the UK is on many dimensions they believe are relevant to allocate the costs of climate change. At the same time, they have more accurate perceptions of the within-

UK ordinal ranking of these characteristics. This underestimation of absolute amounts, but accurate perception of rankings, is similar to that found for other beliefs relevant to climate change (Camilleri et al., 2019; Imai et al., 2026).

Table 1: Summary statistics of elicited beliefs about UK characteristics.

UK characteristic	True value	Mean belief	Q1	Median	Q3	Under-est.
Current emissions	8.1	14.44	3.00	6.00	11.00	0.63
Historical emissions	43.0	31.96	5.00	12.00	25.00	0.86
GDP	33.2	20.75	5.00	8.90	20.00	0.88
Population	8.5	6.90	1.00	3.00	7.00	0.81
Vulnerability	—	8.91	2.00	3.00	6.00	—

Note: All data are expressed in permille of the global total.

4.3 Do fairness views and beliefs matter for participants' climate policy preferences?

Having described the participants' beliefs and fairness views, we now investigate whether these primitives matter for the voters' support of climate policy. We start by checking in Section 4.3.1 whether the two primitives jointly predict policy preferences. We then continue by analysing the effects of our experimental treatments on those preferences in Sections 4.3.2 and 4.3.3. We see the two analyses as complementary. Independently of any causal relationship, policymakers might want to measure their constituency's fairness views and beliefs before an international negotiation if those primitives are useful predictors of the type of climate agreement voters will support.[8]. Instead, the treatment analysis looks for evidence that the primitives are a root cause of the policy preferences. This type of evidence is needed to move from descriptive findings to policy implications. Section 4.3.4 discusses the overall results.

4.3.1 Predicting participants' policy preferences.

As a preliminary step towards the analysis of the predictive power of fairness views and beliefs, we combine these variables to estimate how much each participant thinks it would be fair for the UK to contribute to each climate cost. We call this estimate "Fair_UK" which, for climate cost k and participant i, is defined as:

$$\text{Fair\_UK}_{ki} := \hat{\beta}_k^0 + \hat{\beta}_{ki}^{\alpha}\hat{s}_i{}^{\alpha} + \hat{\beta}_{ki}^{\omega}\hat{s}_i{}^{\omega} + \hat{\beta}_{ki}^{\gamma}\hat{s}_i{}^{\gamma} + \hat{\beta}_{ki}^{\nu}\hat{s}_i{}^{\nu} + \hat{\beta}_{ki}^{\iota}\hat{s}_i{}^{\iota} \tag{8}$$

[8]Indeed, correlational exercises feature prominently in articles investigating the relationship between voters' fundamental preferences and their support for policies (Enke, Rodríguez-Padilla and Zimmermann, 2023; Epper et al., 2024; Dechezleprêtre et al., 2025; Chinoy et al., 2026)

In words, Fair_UK is the sum of a participant's beliefs about the UK weighted by her fairness views for the relevant cost. This definition mirrors Equation 1 in the theoretical framework. To compute Fair_UK, we use the estimates of the fairness weights obtained by fitting Equation 6 at the individual level; for the beliefs, we use the participants' priors. In this section, we limit the analysis to the control group as these beliefs and fairness views can be affected by the treatments. On average, "Fair_UK" is 32.5‰ for mitigation, 36.1‰ for adaptation, and 27.4‰ for loss and damage. The CDFs of these estimates are displayed in Appendix Figure B2.

Next, we check whether Fair_UK predicts support for climate policies. We do so by regressing the three policy preference measures on Fair_UK and an increasing set of controls. Were we to find that the coefficients for Fair_UK are significant, we could then conclude that fairness views and beliefs jointly predict what voters would like their government to do for climate change.

Figure 6 displays the results of this analysis by plotting the coefficient for Fair_UK in regressions where both Fair_UK and the policy preferences are standardised (the standardisation makes the estimates comparable across the different policy measures). For each dependent variable, we run four regressions. The first uses only age, age squared, gender and their interaction as controls. The second adds other demographic controls, including the party voted at the last election, migration background, and income; the third further includes controls capturing the participants worldviews such as climate change concern, importance of climate change for voting, conditional cooperation concerns about climate change, trust in government, zero sum thinking, and moral universalism; The final one replaces the controls with participant fixed effects and tests whether variations in Fair_UK explain within subjects differences in policy preferences across costs. The detailed regression results are reported in Appendix Tables C.3, C.4, C.5, and C.6.

The first block of regressions uses all three policy preference measures at once. In the regression with only basic controls, a standard deviation increase in Fair_UK predicts an average increase of 0.28 standard deviations in our measures of support for climate policies ($p < 0.001$). The coefficient remains stable as we progressively add all our controls. Once we add the fixed effects, it drops to 0.04 but stays significant ($p = 0.002$). The result indicates that voters' policy preferences can be predicted by their beliefs and fairness views.

The rest of Figure 6 breaks down the analysis across the different policy preference elicitations to identify where the overall predictive power comes from. The second block of coefficients concerns what the participants would make the UK do as part of the treaty. In the first model, a one-standard-deviation increase in Fair_UK is associated

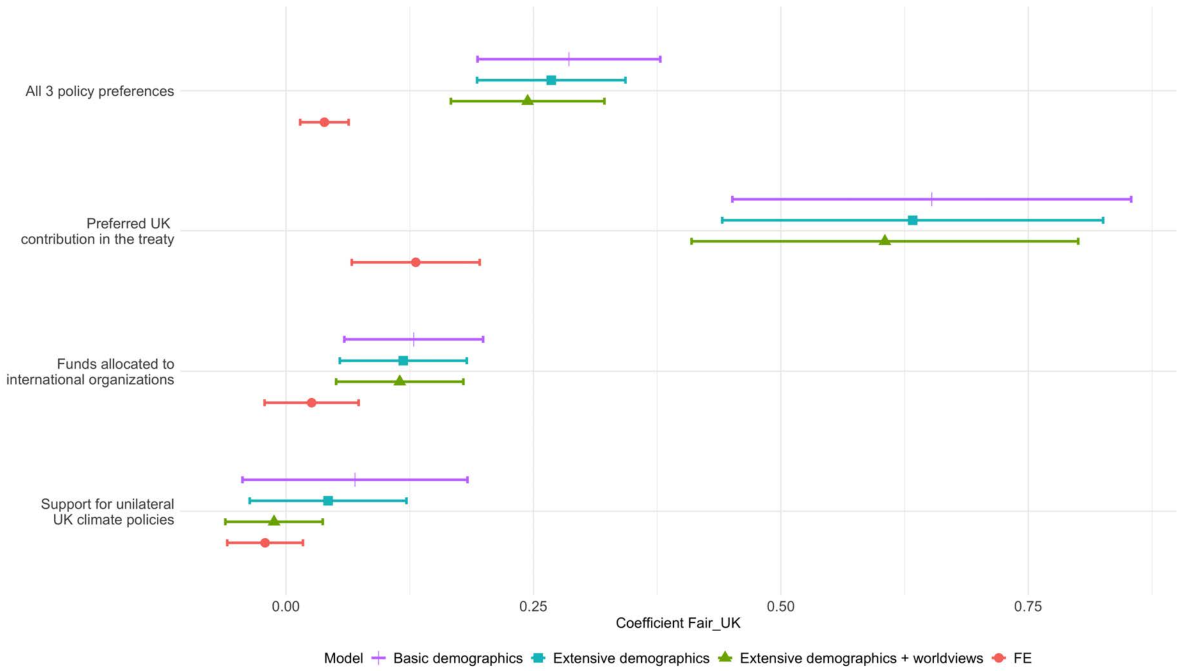


Figure 6: Predicted Policy Preferences

Note: Each coefficient is from a separate regression of the (standardised) policy preference measure on standardised Fair_UK, with controls for age, age squared, gender, and their interactions with gender. The sample is restricted to the participants in the control treatment. For each outcome, the first three estimates correspond to (1) the baseline specification, (2) a specification that add demographic controls for household income, educational attainment, party voted in the UK 2024 general elections, UK region, place of residence (urban/rural), and family migration background, and (3) a specification that adds controls for other core beliefs —trust in government, climate change significance, the importance of climate change for voting, beliefs about how the UK should adjust its climate policies depending on how much other countries are doing, moral universalism, and zero-sum thinking. The fourth estimate (4) replaces the controls with participant fixed effects. The underlying regressions, including the coefficients on the control variables, are reported in Appendix Tables C.3, C.4, C.5, and C.6. Horizontal bars are 95% confidence intervals; standard errors are clustered at the participant level. For comparison, in specification (2), the outcome variables are, on average, 0.93, 0.35, 0.48, and 1.90 standard deviations higher for Green Party voters than for Reform UK voters.

with a 0.65-standard-deviation increase in what the participant would have the UK pay ($p < 0.001$). The coefficient is stable when adding more controls; it remains significant but drops to 0.13 standard deviations in the model with fixed effects ($p < 0.001$).

To interpret the size of the coefficient for Fair_UK, we compare it to those of other variables in the specification that includes all demographic and world views controls. In that specification, a control for conditional cooperation preferences is the only other regressor significant at the 5% level. [9] A standard deviation increase in this control is

[9]To these preferences, participants were asked "How should UK climate policies depend on what other countries do?" and answered two separate sub-questions: "If other countries do *more*, the UK should do:" and "If other countries do *less*, the UK should do:". Both were answered on the same five-point scale going from "Much less" (1) to "Much more" (5). Both answers enter the regression as separate standardised regressors; the coefficient reported in the text is the one for the first sub-question. The other coefficient is not significant. We took these questions from Fabre, Douenne and Mattauch (2023).

associated with a 0.16 standard deviation increase in the desired UK contribution. This effect is roughly one-quarter the size of the estimated effect of Fair_UK.[10]

The next block of coefficients presents the results of regressions in which the dependent variable is the amount participants allocate to international organisations focused on climate change. Here too there is a significant relationship with Fair_UK, though the association is weaker. In the regression with only the basic controls, a one-standard-deviation increase in Fair_UK is associated with a 0.13 standard-deviation increase in funds given to the organisations ($p < 0.001$). The coefficient drops to 0.11 ($p < 0.001$) when we add all the controls, and decreases further to 0.03, turning insignificant ($p = 0.29$), in the model with fixed effects.

In the model with all the controls, the coefficient for Fair_UK is comparable in size to that for trust in government. Participants whose trust that the government spends money wisely is one standard deviation higher allocate 0.12 standard deviations fewer funds to the international organizations.[11] This relationship is significant at the 10% level ($p = 0.08$).

Finally, the last block of coefficients uses support for current UK climate policies as the dependent variable. Here we don't find a significant association with Fair_UK. In all regressions, the displayed coefficients are smaller than 0.07 standard deviations and insignificant ($p > 0.20$ in all regressions). Hence, we don't find evidence that Fair_UK predicts support for climate policies that are not part of an international treaty.

Overall, we conclude that fairness views and beliefs are significant predictors of voters' support for climate policies. The strength of the association is comparable to the one between policy preferences and conditional cooperation concerns or trust in government, two variables that can intuitively explain policy support. However, the predictive power of fairness views and beliefs is specific to preferences for international agreements, as we don't find that they are associated with support for unilateral climate policies. This last result might be due to the fact that we measured fairness views under the assumption that all countries in the world would join a climate treaty. No such global cooperation underpins current UK policies.

4.3.2 Treatments effects

The previous section established that fairness views and beliefs jointly predict policy preferences. Now we investigate whether the relationship is causal. The analysis focuses

[10]It is remarkable that conditional cooperation concerns are a significant predictor of the share that the participants would like the UK to pay. This result suggests that participants took into account that all countries of the world would be joining the treaty.

[11]Participants could indicate their level of trust on a 1-6 scale.

on policy preferences related to adaptation because we used participants' fairness views about this climate change cost to tailor the treatments. For the inference, we use the consistent variance estimators recommended in Bai, Shaikh and Tabord-Meehan (2025), which allows us to exploit the added precision generated by the stratified randomization.[12]

Fairness views. We start from the fairness views. We regress the policy preferences related to adaptation in the two reflection conditions over a dummy for the Max Reflection treatment. The coefficient on this dummy identifies the effect of arguing in favour of the criterion that assigns the highest responsibility to the UK (according to a participant's belief) instead of arguing in favour of the criterion that assigns it the lowest responsibility. Accordingly, we expected the coefficient to be positive.

The left panel of Figure 7 plots these coefficients for regressions where we either use all three adaptation policy preferences at once (first line), or we analyse them separately (subsequent three lines). To allow for comparisons, the policy preference variables are standardised in all regressions. The estimated coefficients are smaller than 0.1 standard deviations and insignificant in all specifications. The left panels of Appendix Figures B3 and B4 replicate this null result with policy preferences for mitigation and loss and damage.

We conclude that there is no evidence that arguing in favour of different criteria changes policy preferences; hence, we found no evidence that fairness views causally impact support for climate policies.

Beliefs The analysis for the causal effect of beliefs follows similar steps. We regress the preferences for adaptation policies in the two information conditions on a dummy for the Upward Information treatment. The coefficient on this dummy identifies the effect of seeing the information that should maximise how much the participant would like the UK to do for adaptation (given his beliefs and fairness views) versus the information that should minimise this amount. Accordingly, we expected the coefficient to be positive.

The right panel of Figure 7 plots these coefficients for regressions where we either use all three adaptation policy preferences at once (first line), or we analyse them separately (subsequent three lines). Once again, the policy preferences variables are standardised to facilitate the comparisons. All the coefficients are smaller than 0.1 standard deviations and insignificant. The right panels of Appendix Figures B3 and B4 replicate this null result with policy preferences for mitigation and loss and damage.

[12] For the estimation, we collapse the age strata and the Green Party and other party strata. In this way, we make sure that every stratum/treatment cell has at least 11 observations. Small cell counts can render the estimated variance unreliable (Bai, Shaikh and Tabord-Meehan, 2025).

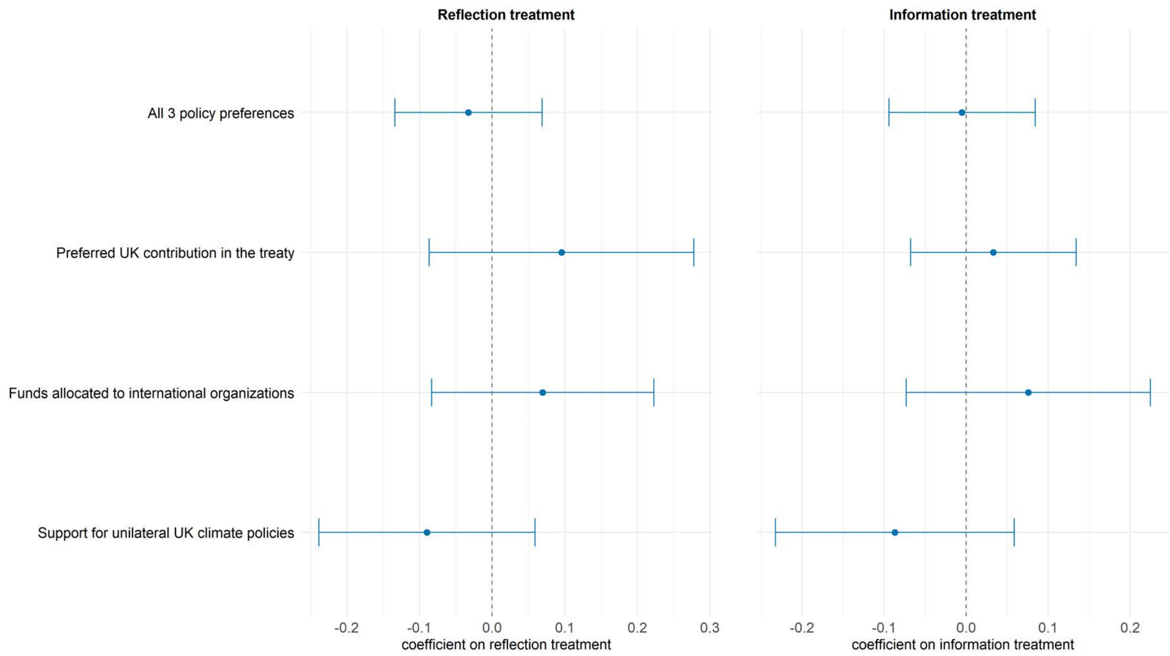


Figure 7: Treatment effects on policy preferences for adaptation costs

Note: The graph displays coefficients from regressions where dependent variables are the participants' policy preferences concerning climate change adaptation (all three elicitations for the first row, one single elicitation for the subsequent 3). The left panel displays the difference between the Max and Min Reflection treatments in regressions using only data from these treatments. The right panel displays the difference between the Upwards and Downwards Reflection treatments in regressions using only data from these treatments. The regressions don't include controls. The standard errors are adjusted to account for the stratified randomisation following Bai et al. (2025). They are clustered at the participant level in the first row, and robust in the other rows. The bars indicate 90% confidence intervals.

We conclude that we don't find evidence that information correcting participants' misperceptions changes their policy preferences.

4.3.3 What explains the null effects?

What are the reasons behind the null effect? And can we learn something from the treatments, nevertheless? This section seeks to answer these two questions.

As the policy-preference dependent variable, we focus on the participants' preferred UK contribution within the climate treaty because this variable is closely aligned with our theoretical framework. Most of the results replicate for the other two policy-preference variables, as shown in Appendix Tables C.7 and C.8. We signal in the text where the results differ.

Do the treatments change fairness views and beliefs? Maybe the treatments didn't change preferences because they didn't affect the model primitives. To test this hypothesis for the fairness views, we can compare participants' responses in the vignette study at the beginning of the experiment with those in the final retest stage. In the retest

Table 2: Change of Fairness Views about Adaptation

	(1) Δ adaptation fair shares	(2) Δ adaptation fair shares
Reflection treatments	0.570	
	(0.554)	
Information treatments		1.149+
		(0.595)
N	2744	2780

Note: Dependent variable: absolute difference in the adaptation fair shares between the first vignette study and the retest. The data includes only the vignette that the participants see in both elicitations (4 vignettes per participant). The first column has data from the control and the two reflection treatments. In this column, the variable Treatment equals 1 if the participant is in a reflection treatment. The second column has data from the control and the two information treatments. In this column, the variable Treatment equals 1 if the participant is in one of the information treatments. Robust standard errors, adjusted for stratification, in parentheses.

stage, the participants see four vignettes they had already encountered, allowing us to compute the absolute variation in the fair share they think these countries should pay. If our manipulation in the reflection treatments were successful in shifting the fairness views, we would expect the participants in these two treatments to vary their answers more than those in the control treatment. As shown in column (1) of Table 2, this is not what happens: the absolute variation in the answers across the two elicitations is similar in the control and in the two reflection treatments. We conclude that the reflection treatments didn't generate a significant first-stage variation in fairness views. This result can explain the lack of a treatment effect on policy preferences and makes the treatment uninformative about the causal link between fairness views and preferences.[13]

Next, we check whether the information treatments change beliefs. We do so by comparing the posterior beliefs in the two information treatments with the ones in the control condition. Figure 8 plots the cumulative belief distributions. The top four panels show the priors and the bottom four the posteriors. The beliefs included in the figure are selected to ensure a like-for-like comparison. For the participants in the two information treatments, the figure uses only the beliefs about the dimension for which each participant received the information. For the control, the figure uses the beliefs for the dimensions for which a participant would have received the information had they been in one of the information treatments. The top panels of Figure 8 display only

[13]This failed manipulation might be of interest to scholars investigating the effect of moral narratives on behaviour. For example, Hillenbrand and Verrina (2020) finds that reading arguments in favour of or against a donation changes how much people give. Here, we ask participants to come up with arguments in favour of an allocation rule, but writing these arguments doesn't produce a detectable effect on fairness views nor allocation decisions.

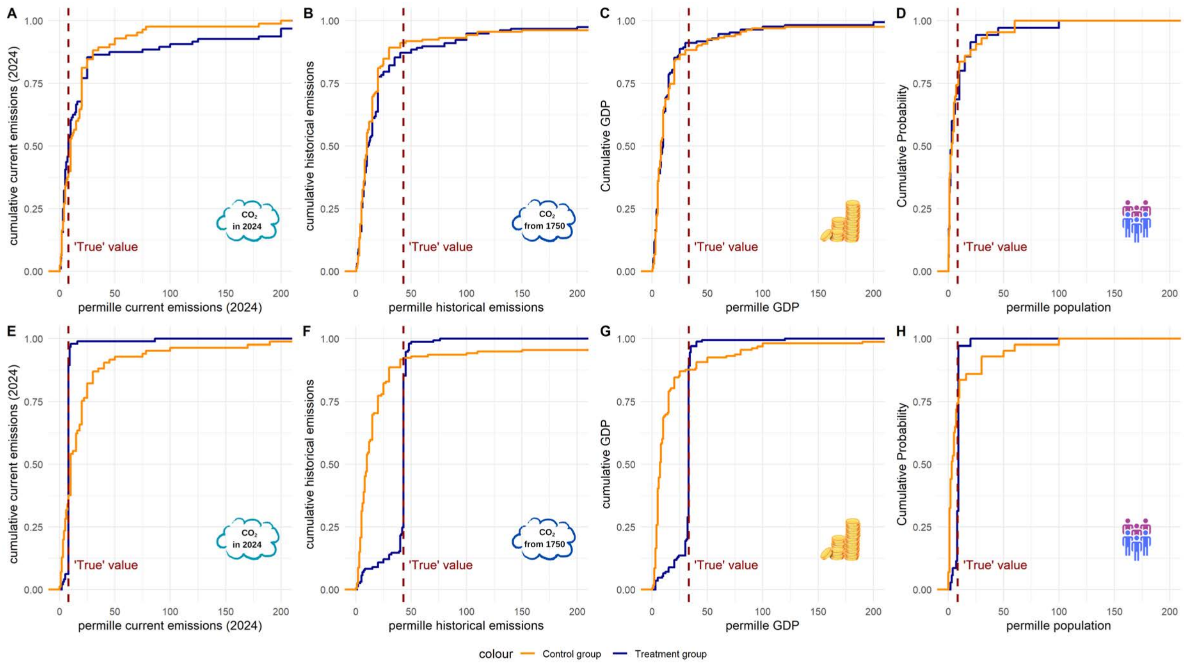


Figure 8: Prior and posterior beliefs about UK characteristics by treatment

Note: The figure displays the cumulative distribution functions of people's beliefs about the UK characteristics in the information treatments (in blue) and in the control (orange). For the participants in the information treatments, the figure includes only the treated beliefs. For the control, it includes the beliefs that would have been treated had the participants been in one of the information treatments. Vertical red dashed lines indicate the true values. Panel A and E: current emissions; Panel B and F: historical emissions; Panel C and G: GDP; Panel D and H: population.

minor differences in the distributions of priors. These differences drastically widen in the bottom panels, where the treated posterior distributions collapse towards the truth, as 71% of participants report the correct value. We conclude that, as we expected, the information treatment generates a strong first-stage effect on beliefs. As such, a failed manipulation is not the explanation for the null finding.

Are the differences between the two information treatments large enough? While we have established that the information treatments change beliefs compared to the control, one might still worry that the belief adjustment generates only a small variation in what participants think is fair for the UK to do. Small treatment effects on Fair_UK_a might explain why we can't find variation in policy preferences. However, we don't find support for this hypothesis. On average, Fair_UK_a, calculated using posterior beliefs, is 6.63 permille larger in the Upward Information treatment than in the Downward Information treatment. The difference is 8.21 times larger than the estimated treatment effect in column (1) of Table 3.

Further evidence inconsistent with the hypothesis comes from the analysis of the participants for which the model predicts the largest treatment effects. This is the 50% of participants for which τ_i (as defined in Expression 7) is the largest. In this subset, the coefficient on the Upward Information treatment dummy is 2.5‰ , larger than the one for the full sample but still insignificant (See column (2) of Table 3, $p = 0.31$).[14]

Overall, we find evidence indicating that, from the lens of our model, the information treatments produce large changes in what the participants think it would be fair for the UK to contribute. Moreover, the null effect persists in the subsample where we would expect the largest effects. We deem unlikely that the null result on policy preferences is due to a weak treatment manipulation.

Do participants in the information treatments change their fairness views? To compute the expected treatment effects, our model assumes that the information treatment doesn't affect the participants' fairness views. A failure of this assumption could explain the null effect of the information on policy preferences. After learning that the UK is larger than they expected in a dimension, the participants might convince themselves that that dimension is not so important for allocating the mitigation costs. Vice versa, participants who learn that the UK is smaller than expected on some characteristic might increase the weight they place on it. Indeed, several articles find that people select their views of fairness in a self-serving way (Babcock and Loewenstein, 1997; Konow, 2000; Deffains, Espinosa and Thöni, 2016; Amasino, Pace and van der Weele, 2023).

If the participants in the information treatments change their views of fairness, we should expect them to systematically modify their answers in the retest of the vignette study. In particular, we would expect the absolute variation of their answers about the countries' fair shares to be larger than in the control treatment. Column (2) of Table 2 finds some support for this hypothesis. The average absolute revision in the adaptation cost fair shares in the information treatments is 1.15‰ larger, a difference significant at the 10% level ($p = 0.053$). We conclude that changes in fairness views after the information treatment might contribute to the null effect of the information manipulation on policy preferences.

Is there a treatment effect in a subsample of the electorate? The literature on preferences for income distribution sometimes finds that information experiments change the policy preferences of only a subset of the electorate. For example, the intervention in Alesina, Stantcheva and Teso (2018); Fehr, Mollerstrom and Perez-Truglia (2022) only works for left-wing respondents. Column 3 of Table 3 checks if there are treatment effects among Green and Labour Party voters, two UK parties that are left of the centre. Similarly, Column 4 of Table 3 searches for treatment effects among the voters of the right-of-the-centre Conservative and Reform UK parties. Both columns replicate

[14]We don't run a similar subsample analysis for the reflection treatment as we can't calculate a theoretical benchmark for its effectiveness.

Table 3: Treatment effects on preferred UK adaptation contribution

	(1) Full	(2) Median Sample Split	(3) Green Party/ Labour	(4) Reform UK/ Conservative	(5) vs. Control
Panel A: Reflection treatments					
Max Reflection	2.325		-0.962	9.171	-1.114
	(2.693)		(3.005)	(6.119)	(3.470)
N	443		194	163	441
Panel B: Information treatments					
Upward Information	0.807	2.547	3.303	-1.079	-1.056
	(1.492)	(2.492)	(2.056)	(2.086)	(2.724)
N	457	228	195	173	455

Note: The dependent variable is the permille of global adaptation costs the participant would make the UK pay within the international climate treaty (x_a^* in the theoretical framework). The regressions include no controls. Panel A regresses this variable on a dummy for the Max Reflection treatment; Panel B regresses it on a dummy for the Upward Information treatment. Samples are as follows: (1) All participants in the Max and Min Reflection treatments (Panel A), or in the Upward and Downward Information treatments (Panel B); (2) the 50% of participants in the Upward and Downward Information treatments for whom the model predicts the largest treatment effect, i.e., τ_i (as defined in Expression 7) is above the sample median; (3) participants who voted for the Green Party or Labour in 2024; (4) participants who voted Reform UK or the Conservative Party in 2024; (5) all participants in the control and in the Max Reflection (Panel A), or Upward Information (Panel B) treatment. Robust standard errors adjusted for the stratified randomisation in parentheses. +: $p < 0.1$; *: $p < 0.05$; **: $p < 0.01$; ***: $p < 0.001$.

the null effects. We conclude that there is no evidence that the manipulations affect climate policy preferences for either voter group.

Are there effects that don't pass via fairness views and beliefs? So far we have run comparisons between the two reflection treatments and the two information treatments. These comparisons were designed to identify as cleanly as possible the effect of the treatments passing via fairness views and beliefs. However, the treatments might have effects that go through different channels —for instance, in Schulze-Tilling (2025) the effect of carbon impact information on consumption behaviour mostly passes via salience. To check whether the treatments have other effects in our experiment as well, we compare treatment conditions with the pure control.

Column (5) of Table 3 compares the control with the Max Reflection treatment (Panel A) and the Upward Information treatment (Panel B) in regressions where the dependent variable is the share of the global adaptation cost the participants would make the UK pay in the treaty. It finds that these shares are somewhat higher in the control, but the difference is not statistically significant. Hence, for the policy preference variable most closely aligned with our theoretical framework, we don't find evidence that the treatments have effects via unanticipated channels.

However, the same analyses yield a different result if we look at how the participants allocate funds between the UK government and the Adaptation Fund, an international organisation that finances adaptation projects in developing countries. Column (5) of Appendix Table C.7 shows that participants in the Max Reflection and in the Upward Information treatment give 0.25£ and 0.26£ more to the Adaptation Fund than participants in the control, a statistically significant increase ($p = 0.005$ and $p = 0.002$). These are unexpected results that we are unable to explain.

4.3.4 So then, do fairness views and beliefs matter for policy preferences?

Taking stock, we conclude that fairness views and beliefs jointly predict voters' support for the government contributions within a climate treaty. The predictive power of these variables is comparable to that of other world views such as trust in government and preferences for conditional cooperation.

However, we don't find evidence of a causal relationship between our model primitives and voters' climate policy preferences. For fairness views, the null finding is due to an unsuccessful first-stage manipulation of the participants' fairness views. The treatment doesn't change what people think is fair; hence, we shouldn't expect it to influence any downstream variable, such as policy support. For beliefs, instead, we find that the treatment successfully corrects misperceptions about how large the UK is on dimensions relevant to a fair allocation of the climate costs. Yet, contrary to the model predictions, this belief shift doesn't translate into a change in policy preferences. In the data, we find suggestive evidence that the information changes not only participants' beliefs, but also their fairness views. This finding might explain why policy preferences remain constant.

5 Conclusions

Using a survey experiment with a UK sample, this paper shows that voters believe the fair international allocation of climate change costs should follow different rules depending on the specific cost considered. In particular, current emissions should be given more weight when splitting the global mitigation costs than when distributing the costs linked to adapting to a warmer world or the loss and damages from the irreversible consequences of climate change. Furthermore, the paper shows that fairness views and beliefs jointly predict the share of the climate change costs that voters would like their government to pay in the context of a global, binding climate agreement. At the same time, the fairness views that apply to a treaty are a poor predictor of people's support for unilateral climate policy.

These results indicate that we cannot extrapolate existing findings on fairness views regarding mitigation to predict how citizens would like other climate change costs to be distributed. Hence, as negotiations on adaptation and loss and damage gain importance, renewed data-collection efforts are needed. Future surveys should describe how fairness views about these costs vary across the globe and identify which cost-sharing agreements could represent an acceptable compromise. These data collections should elicit citizens' fairness views under the assumption that a climate agreement has been reached. Our evidence confirms that conditional cooperation concerns loom large in voters' minds and hence different fairness views apply to policies taken within and outside international agreements. The paper also highlights a difficulty that these data collection efforts will have to overcome. Many participants, especially the less educated, dropped out when reading the instructions for the hypothetical international treaty and answering the related comprehension questions. Finding a more accessible way to explain the agreement is a prerequisite for collecting representative evidence of people's fairness views.

The findings have some high-level implications for climate negotiators. First, negotiators might want to collect data on their constituencies' fairness views. They can use this data to understand which principles voters think should inform the treaty and, hence, which agreements they would support. Second, the results suggest that, if negotiators want to keep their constituency's fairness views into account, they should adjust their negotiating position depending on the climate change cost under discussion. Finally, the evidence highlights that voters have large misperceptions about how large their country is on dimensions relevant to the negotiations. Even if the paper doesn't find evidence that beliefs have a causal impact on preferences, we can't exclude that these misperceptions are relevant for policy support. These misperceptions might make reaching an agreement harder, especially if citizens in many countries systematically believe that their country is less responsible than it is.

References

Alesina, Alberto, Stefanie Stantcheva, and Edoardo Teso. 2018. "Intergenerational Mobility and Preferences for Redistribution." *American Economic Review*, 108(2): 521–554. `https://doi.org/10.1257/aer.20162015`.

Amasino, Dianna, Davide Domenico Pace, and Joel van der Weele. 2023. "Self-Serving Bias in Redistribution Choices: Accounting for Beliefs and Norms." *https://epub.ub.uni-muenchen.de/94539/*, `https://doi.org/10.5282/ubm/epub.94539`.

Amasino, Dianna R., Davide D. Pace, and Joël J. van der Weele. 2024. "Fair Shares and Selective Attention." *American Economic Journal: Microeconomics*, 16(4): 259–290. https://doi.org/10.1257/mic.20220275.

Andre, Peter. 2025. “Shallow Meritocracy.” *The Review of Economic Studies*, 92(2): 772–807. https://doi.org/10.1093/restud/rdae040.

Andre, Peter, Teodora Boneva, Felix Chopra, and Armin Falk. 2024. “Globally Representative Evidence on the Actual and Perceived Support for Climate Action.” *Nature Climate Change*, 1–7.

Auer Frege, Ilona, Kai Budelmann, Julia Burakowski, et al. 2025. “WorldRiskReport 2025 – Focus: Floods.” Bündnis Entwicklung Hilft and Institute for International Law of Peace and Humanitarian Law (IFHV), Ruhr University Bochum. https://www.preventionweb.net/media/110940/download (accessed 2026-08-11).

Babcock, Linda, and George Loewenstein. 1997. “Explaining Bargaining Impasse: The Role of Self-Serving Biases.” *Journal of Economic perspectives*, 11(1): 109–126.

Bai, Yuehao, Azeem M. Shaikh, and Max Tabord-Meehan. 2025. “A Primer on the Analysis of Randomized Experiments and a Survey of Some Recent Advances.” https://doi.org/10.48550/arXiv.2405.03910.

Bechtel, Michael M., and Kenneth F. Scheve. 2013. “Mass Support for Global Climate Agreements Depends on Institutional Design.” *Proceedings of the National Academy of Sciences*, 110(34): 13763–13768. https://doi.org/10.1073/pnas.1306374110.

Bortolotti, Stefania, Ivan Soraperra, Matthias Sutter, and Claudia Zoller. 2025. “Too Lucky to Be True: Fairness Views under the Shadow of Cheating.” *The Review of Economics and Statistics*, 107(3): 771–785. https://doi.org/10.1162/rest_a_01394.

Camilleri, Adrian R., Richard P. Larrick, Shajuti Hossain, and Dalia Patino-Echeverri. 2019. “Consumers Underestimate the Emissions Associated with Food but Are Aided by Labels.” *Nature Climate Change*, 9(1): 53–58.

Cappelen, Alexander W., Astri Drange Hole, Erik Ø Sørensen, and Bertil Tungodden. 2007. “The Pluralism of Fairness Ideals: An Experimental Approach.” *American Economic Review*, 97(3): 818–827.

Cappelen, Alexander W., Astri Drange Hole, Erik Ø. Sørensen, and Bertil Tungodden. 2011. “The Importance of Moral Reflection and Self-Reported Data in a Dictator Game with Production.” *Social Choice and Welfare*, 36(1): 105–120. https://doi.org/10.1007/s00355-010-0468-3.

Cappelen, Alexander W., James Konow, Erik Ø Sørensen, and Bertil Tungodden. 2013. "Just Luck: An Experimental Study of Risk-Taking and Fairness." *American Economic Review*, 103(4): 1398–1413.

Carattini, Stefano, Steffen Kallbekken, and Anton Orlov. 2019. "How to Win Public Support for a Global Carbon Tax." *Nature*, 565(7739): 289–291. https://doi.org/10.1038/d41586-019-00124-x.

Carleton, Tamma, Michael Greenstone, Solomon Hsiang, Andrew Hultgren, Robert E. Kopp, Kelly E. McCusker, Ishan Nath, James Rising, and Ashwin Rode. 2026. "The Local Damages from Global Climate Change." https://doi.org/10.2139/ssrn.6954318.

Chinoy, Sahil, Nathan Nunn, Sandra Sequeira, and Stefanie Stantcheva. 2026. "Zero-Sum Thinking and the Roots of US Political Differences." *American Economic Review*, 116(3): 1052–1096. https://doi.org/10.1257/aer.20240692.

Dechezleprêtre, Antoine, Adrien Fabre, Tobias Kruse, Bluebery Planterose, Ana Sanchez Chico, and Stefanie Stantcheva. 2025. "Fighting Climate Change: International Attitudes toward Climate Policies." *American Economic Review*, 115(4): 1258–1300. https://doi.org/10.1257/aer.20230501.

Deffains, Bruno, Romain Espinosa, and Christian Thöni. 2016. "Political Self-Serving Bias and Redistribution." *Journal of Public Economics*, 134: 67–74. https://doi.org/10.1016/j.jpubeco.2016.01.002.

de la Cuesta, Brandon, Naoki Egami, and Kosuke Imai. 2022. "Improving the External Validity of Conjoint Analysis: The Essential Role of Profile Distribution." *Political Analysis*, 30(1): 19–45. https://doi.org/10.1017/pan.2020.40.

Douenne, Thomas, and Adrien Fabre. 2022. "Yellow Vests, Pessimistic Beliefs, and Carbon Tax Aversion." *American Economic Journal: Economic Policy*, 14(1): 81–110. https://doi.org/10.1257/pol.20200092.

Enke, Benjamin, and Thomas Graeber. 2023. "Cognitive Uncertainty." *The Quarterly Journal of Economics*, 138(4): 2021–2067. https://doi.org/10.1093/qje/qjad025.

Enke, Benjamin, Ricardo Rodríguez-Padilla, and Florian Zimmermann. 2023. "Moral Universalism and the Structure of Ideology." *The Review of Economic Studies*, 90(4): 1934–1962. https://doi.org/10.1093/restud/rdac066.

Epper, Thomas F., Ernst Fehr, Claus Thustrup Kreiner, Søren Leth-Petersen, Isabel Skak Olufsen, and Peer Ebbesen Skov. 2024. "Inequality Aversion Predicts Support for Public and Private Redistribution." *Proceedings of the National Academy of Sciences*, 121(39): e2401445121. https://doi.org/10.1073/pnas.2401445121.

Exley, Christine L., and Judd B. Kessler. 2024. "Equity Concerns Are Narrowly Framed." *American Economic Journal: Microeconomics*, 16(2): 147–179. https://doi.org/10.1257/mic.20220266.

Fabre, Adrien, Thomas Douenne, and Linus Mattauch. 2023. "International Attitudes Toward Global Policies." https://doi.org/10.2139/ssrn.4448523.

Fabre, Adrien, Thomas Douenne, and Linus Mattauch. 2025. "Majority Support for Global Redistributive and Climate Policies." *Nature Human Behaviour*, 9(8): 1583–1594. https://doi.org/10.1038/s41562-025-02175-9.

Fehr, Dietmar, Johanna Mollerstrom, and Ricardo Perez-Truglia. 2022. "Your Place in the World: Relative Income and Global Inequality." *American Economic Journal: Economic Policy*, 14(4): 232–268. https://doi.org/10.1257/pol.20200343.

Fischbacher, Urs, David Grammling, Jan Hausfeld, and Vojtěch Zíka. 2023. "Identity Breeds Inequality: Evidence from a Laboratory Experiment on Redistribution." *Journal of Public Economics*, 222: 104866. https://doi.org/10.1016/j.jpubeco.2023.104866.

Fleurbaey, Marc, Sivan Kartha, Simon Bolwig, Yoke Ling Chee, Ying Chen, Esteve Corbera, Franck Lecocq, Wolfgang Lutz, Maria Silvia Muylaert, and Richard B. Norgaard. 2014. "Sustainable Development and Equity." *Climate Change 2014: Mitigation of Climate Change. IPCC Working Group III Contribution to AR5.*

Gardiner, Stephen M. 2004. "Ethics and Global Climate Change." *Ethics*, 114(3): 555–600. https://doi.org/10.1086/382247.

Hillenbrand, Adrian, and Eugenio Verrina. 2020. "The Differential Effect of Narratives on Prosocial Behavior." *MPI Collective Goods Discussion Paper*, , (2018/16).

Hoy, Christopher, and Franziska Mager. 2021. "Why Are Relatively Poor People Not More Supportive of Redistribution? Evidence from a Randomized Survey Experiment across Ten Countries." *American Economic Journal: Economic Policy*, 13(4): 299–328. https://doi.org/10.1257/pol.20190276.

Hvidberg, Kristoffer B, Claus T Kreiner, and Stefanie Stantcheva. 2023. "Social Positions and Fairness Views on Inequality." *The Review of Economic Studies*, 90(6): 3083–3118. https://doi.org/10.1093/restud/rdad019.

Imai, Taisuke, Davide D. Pace, Peter Schwardmann, and Joel J. van der Weele. 2026. "Correcting Consumer Misperceptions About CO2 Emissions."

INFORM. 2026. "INFORM Risk Index 2026." https://drmkc.jrc.ec.europa.eu/inform-index (accessed 2026-08-11). Disaster Risk Management Knowledge Centre, European Commission Joint Research Centre.

Konow, James. 2000. "Fair Shares: Accountability and Cognitive Dissonance in Allocation Decisions." *American economic review*, 90(4): 1072–1091.

Kuziemko, Ilyana, Michael I. Norton, Emmanuel Saez, and Stefanie Stantcheva. 2015. "How Elastic Are Preferences for Redistribution? Evidence from Randomized Survey Experiments." *American Economic Review*, 105(4): 1478–1508.

Lange, Andreas, Andreas Löschel, Carsten Vogt, and Andreas Ziegler. 2010. "On the Self-Interested Use of Equity in International Climate Negotiations." *European Economic Review*, 54(3): 359–375. https://doi.org/10.1016/j.euroecorev.2009.08.006.

Meilland, Auriane, Yann Kervinio, and Aurélie Méjean. 2025. "International Climate Justice: What the People Think." *Environmental and Resource Economics*, 88(1): 225–247. https://doi.org/10.1007/s10640-024-00931-5.

Mildenberger, Matto, Erick Lachapelle, Kathryn Harrison, and Isabelle Stadelmann-Steffen. 2022. "Limited Impacts of Carbon Tax Rebate Programmes on Public Support for Carbon Pricing." *Nature Climate Change*, 12(2): 141–147. https://doi.org/10.1038/s41558-021-01268-3.

Notre Dame Global Adaptation Initiative (ND-GAIN). 2024. "ND-GAIN Country Index Technical Report." Environmental Change Initiative, University of Notre Dame. https://gain.nd.edu/assets/581554/nd_gain_countryindex_technicalreport_2024.pdf (accessed 2026-08-11).

Pärnamets, Philip, Petter Johansson, Lars Hall, Christian Balkenius, Michael J. Spivey, and Daniel C. Richardson. 2015. "Biasing Moral Decisions by Exploiting the Dynamics of Eye Gaze." *Proceedings of the National Academy of Sciences*, 112(13): 4170–4175.

Pindyck, Robert S. 2013. "Climate Change Policy: What Do the Models Tell Us?" *Journal of Economic Literature*, 51(3): 860–872. https://doi.org/10.1257/jel.51.3.860.

Schleich, Joachim, Elisabeth Dütschke, Claudia Schwirplies, and Andreas Ziegler. 2016. "Citizens' Perceptions of Justice in International Climate Policy: An Empirical Analysis." *Climate Policy*, 16(1): 50–67. https://doi.org/10.1080/14693062.2014.979129.

Schulze-Tilling, Anna. 2025. "The Effectiveness of Carbon Labels." University of Bonn and University of Mannheim, Germany Discussion Paper No. 639.

Stantcheva, Stefanie. 2025. "Perceptions, Mindsets and Beliefs Shaping Policy Views (2024 Economica-Coase Lecture)." *Economica, Forthcoming.*

Stötzer, Lasse S., and Florian Zimmermann. 2026. "Morality and Climate Policy Attitudes." University of Bonn and University of Cologne, Germany Working Paper 424.

Woerner, Andrej, Taisuke Imai, Davide D. Pace, and Klaus M. Schmidt. 2024. "How to Increase Public Support for Carbon Pricing with Revenue Recycling." *Nature Sustainability*, 7(12): 1633–1641. https://doi.org/10.1038/s41893-024-01466-9.

YouGov. 2024. "How Britain Voted in the 2024 General Election." *https://yougov.com/en-gb/articles/49978-how-britain-voted-in-the-2024-general-election.*

Online Appendix

Fairness views about the International Distribution of Climate Change Costs

Davide D. Pace and Johanna Kober

A Design Details

A.1 Elicitation Interfaces

Here we present the details of several elicitation interfaces.

A.1.1 Fairness Views

Ranking In this fairness fairness views elicitation, we asked participants to rank the five characteristics (current emissions, historical emissions, GDP, population, and vulnerability) from most important (1) to least important (5) for determining a country's fair contribution to a given climate change cost. The order in which these characteristics were displayed was the same in which the participant saw them in the instructions. Participants completed this ranking three times, always in the same order: first for mitigation, then for adaptation, and finally for loss and damage. Participant-level random order which was preserved through the survey.

For each climate cost, the first page of the elicitation interface (shown in panel (a) of Figures A.1, A.2, and A.3) had two columns. The left listed the five characteristics while the right one displayed the participant's ranking. When the participant clicked on a characteristic on the left, the interface placed it in the next free rank on the right and simultaneously greyed it out at its original position on the left. Participants could reorder the ranked characteristics by dragging them, or remove one from the ranking by clicking it on the right, which restored the original color on the left and freed up the ranking slot or moved up any characteristic ranked lower (if any). The interface gave the option to restart the ranking from scratch via the "Reset" button. Once the ranking was complete the "NEXT" button allowed the participants to submit their choices.

The next elicitation page displayed the participant ranking and asked whether any of the characteristics should not be used to determine the fair contribution to that cost (see panel (b) of Figures A.1, A.2, and A.3). Clicking a characteristic applied a cutoff at that rank: the clicked characteristic and every characteristic ranked below it were marked as irrelevant, while every characteristic ranked above it was marked as relevant; clicking a different characteristic moved the cutoff to the new position. Relevant characteristics were highlighted in the color associated with the current cost (green for mitigation, blue for adaptation, orange for loss and damage), while irrelevant ones turned red. The green button "All 5 characteristics should be used in a fair agreement" marked every characteristic as relevant. The "NEXT" button moved participants to the next page.

After the participants completed the ranking task for all three costs, they landed on a summary page that displayed their three rankings and relevant criteria (see Figure A.4). On this interface the participants could rearrange any any of their rankings by dragging the country attributes and they could click on any of them to change their status from relevant to irrelevant (or vice versa) for a given cost. On this interface we didn't force participants anymore to submit rankings where all the relevant criteria are ranked above the irrelevant ones.

To help participants, the three elicitation pages displayed reminders of the information presented in the instructions. The ranking page showed a reminder of the definition of permille. All the pages displayed "i" icons the participants hoover over to reread the definitions of the three climate change costs and of the five country attributes.

Vignette Study The interface displayed 14 abstract countries, one per page. (See Figure A.5). A country was characterized by its permille contribution of current and historical emissions, GDP, population and vulnerability to the respective global output. The order in which these characteristics were displayed was the same in which the participant saw them in the instructions. For each country the participant had to type the share of the global mitigation, adaptation, and loss and damage cost that it would be fair for that country to pay.

Several features in the interface help participants in their task. To facilitate the interpretation of the numbers in the vignette, the interface reminds participants that if all the countries in the world were identical, each will be as large as 5.1‰ on every characteristic. Moreover, the interface displayed "i" icons the participants hoover over to reread the definitions of the three climate change costs and of the five country attributes.

A.1.2 Beliefs

Figure A.6 displays the beliefs elicitation interface. The interface consisted of a map of the UK and a table listing of the five characteristics including respective illustrations and information buttons for the definitions. Participants could type there guess into an input box next to the respective characteristic.

A.1.3 Policy Preferences

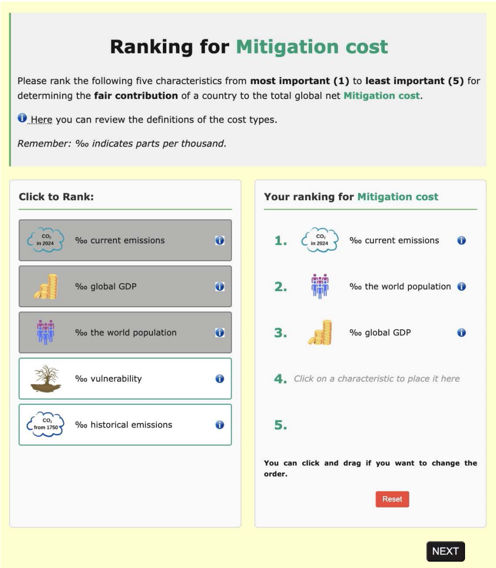
Preferred UK contribution Figure A.7 displays the interface the participants used to indicate which share of the climate change costs they would make the UK pay if they were in charge of the binding climate treaty. The interface replicates the one of the vignette study but replaces the table displaying the country characteristics with a map of the UK.

Fund allocations In this task, participants used a slider to allocate funds between the UK government general budget and an international organization working on climate change. The three rounds followed a fixed order: participants first allocated funds for mitigation (with Climeworks, which captures and stores CO_2), then for adaptation (with the Adaptation Fund, which helps vulnerable communities in developing countries adapt to climate change), and finally for loss and damage (with GlobalGiving, which funds relief programs for the victims of severe floods). "i" icons next to the UK government and organization labels let participants hover to reread how the fund are spent if assigned to an organization.

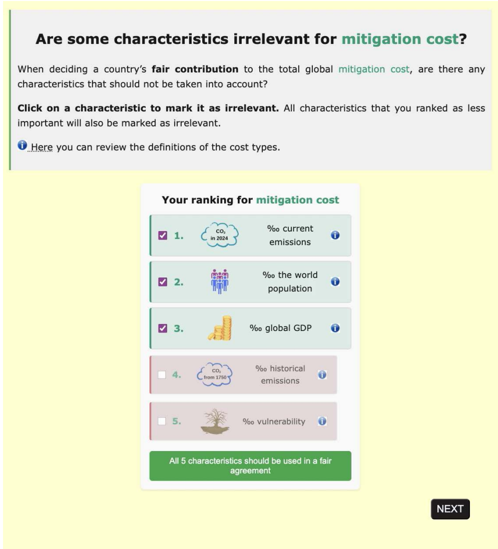
The slider ran from the UK government, on the left, to the organization, on the right, and, by the default, was at the UK end when the interface loaded. Letting $x \in [0, 3]$ denote the amount (in £) assigned to the organization, the amount left with the UK government was given by the concave function

$$g(x) = 10\left(-\frac{x^3}{60} - \frac{11x}{60} + 1\right), \tag{9}$$

so that the combined total (UK government plus organization, $g(x) + x$) shrank as participants moved the slider toward the organization: from £10 at the UK end to £3 once all the funds are allocated to the organization. A bar chart, updated live as participants moved the slider, showed the total funds together with the amounts assigned to the government and to the organization. Participants had to move the slider at least once before they could submit. Figure A.8 shows how the interface changed as participants moved the slider from the UK government toward the organization.

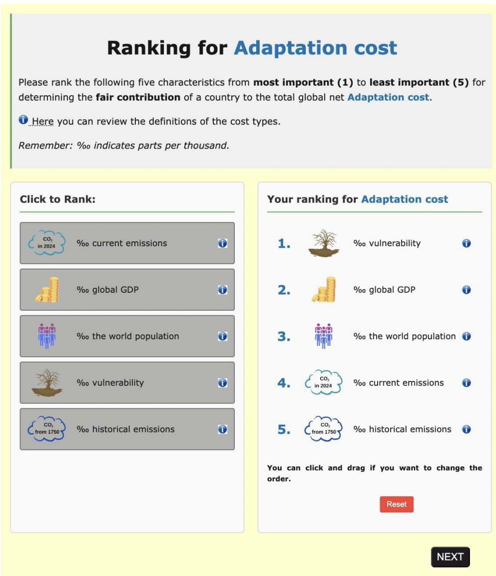


(a) Ranking

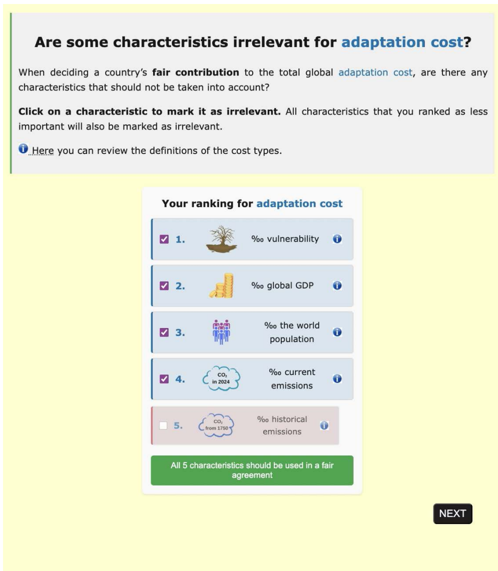


(b) Relevance

Figure A.1: Ranking and relevance interface for mitigation costs.

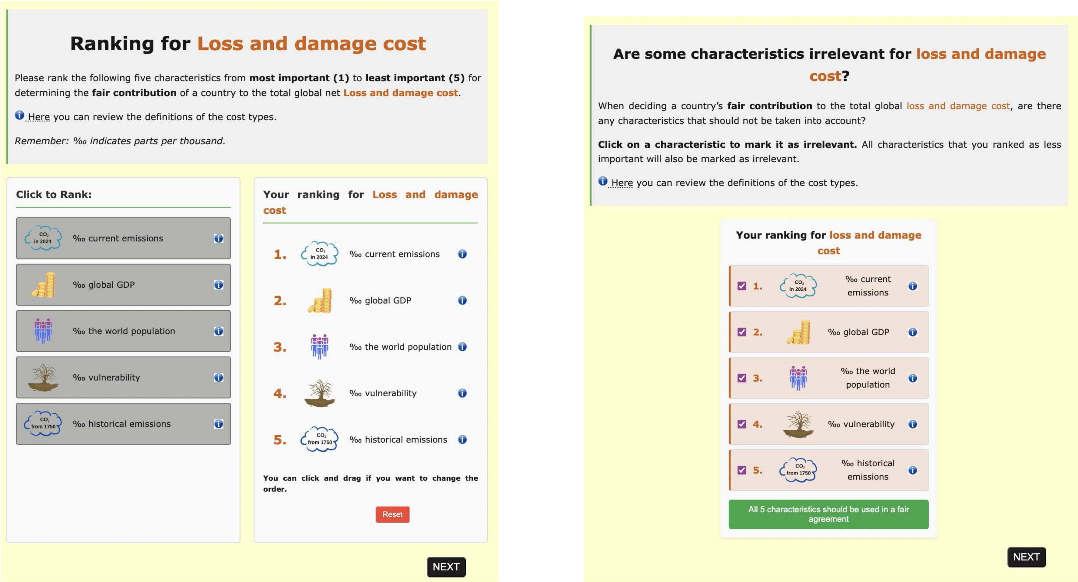


(a) Ranking



(b) Relevance

Figure A.2: Ranking and relevance interface for adaptation costs.



(a) Ranking (b) Relevance

Figure A.3: Ranking and relevance interface for loss and damage costs.

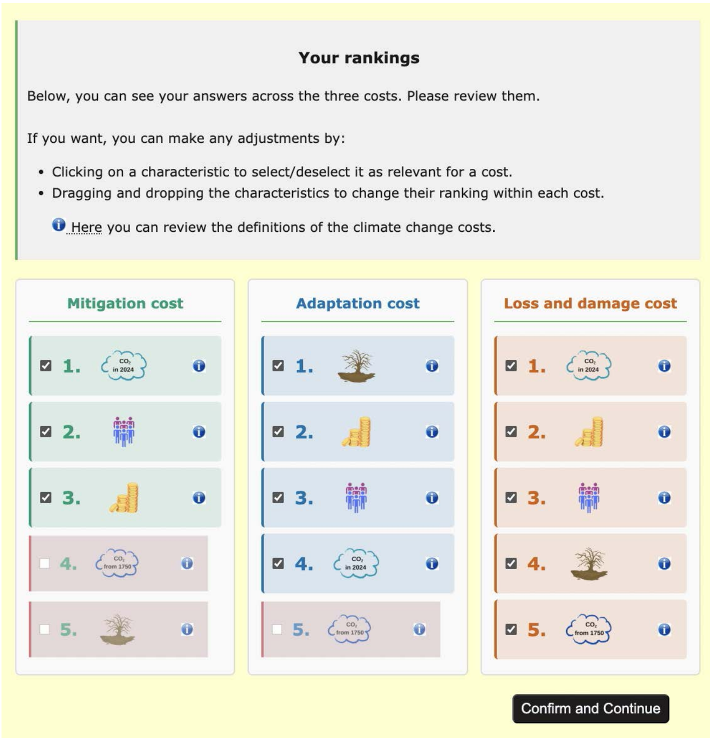


Figure A.4: Summary interface shown after the ranking task for all three costs.

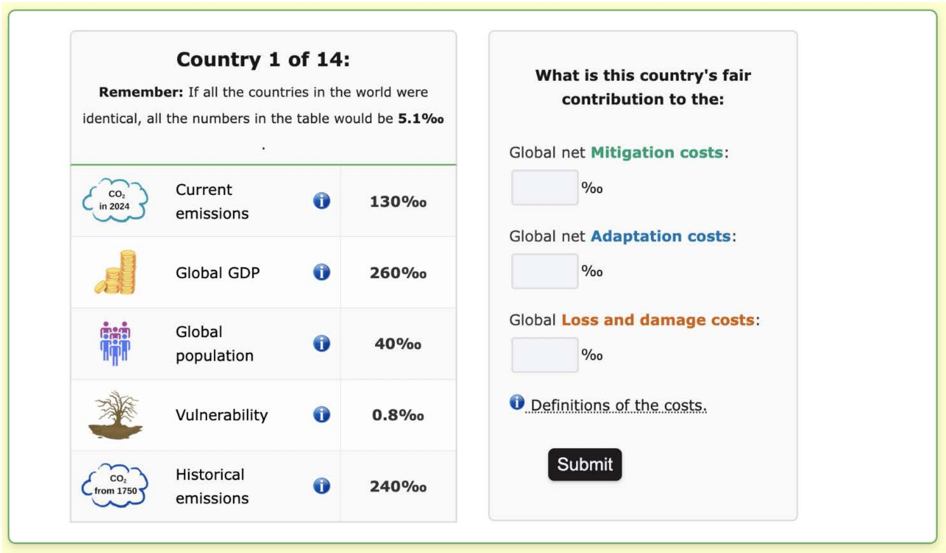


Figure A.5: Vignette study interface.

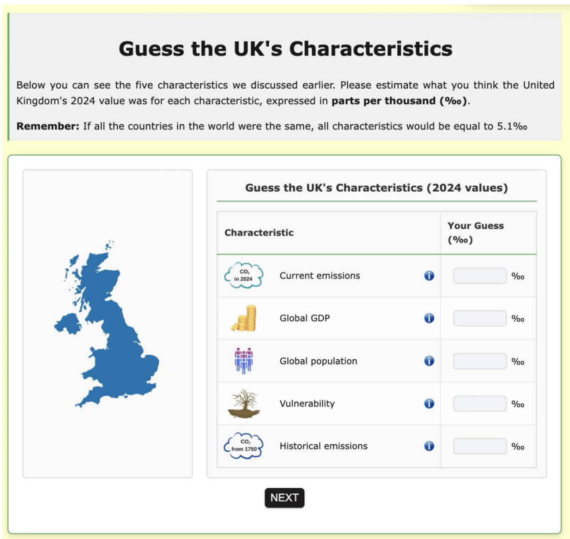


Figure A.6: Beliefs elicitation interface.

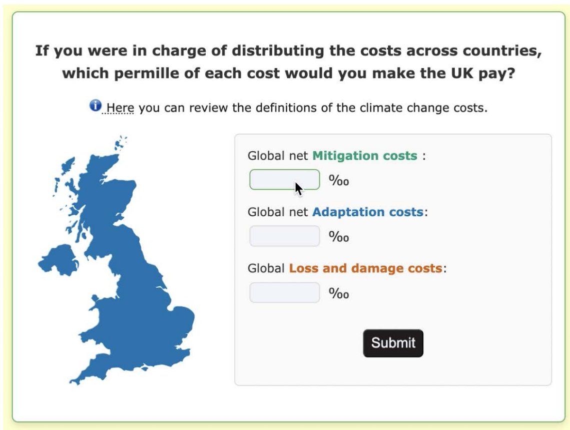
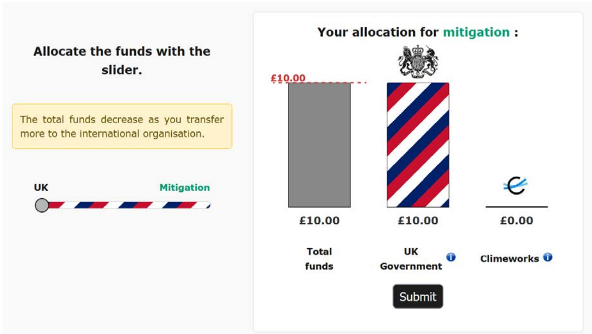
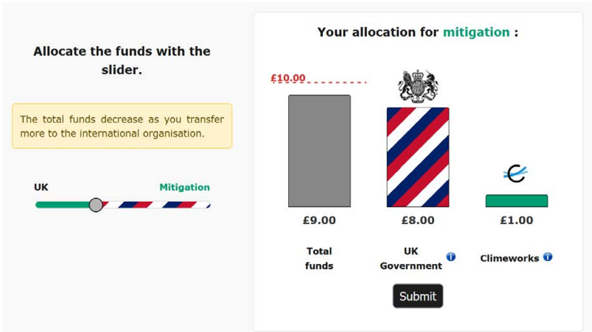


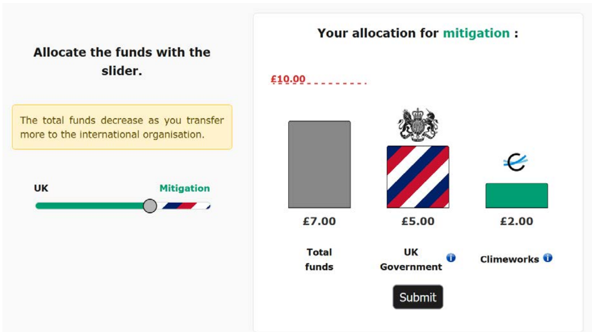
Figure A.7: UK contribution interface.



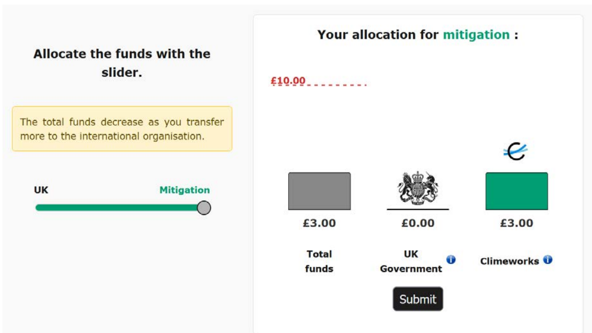
(a) Slider at the UK end



(b) Slider moved toward the organization



(c) Slider moved further toward the organization



(d) Slider at the organization end

Figure A.8: The panels display how the interface changes as participants transfer more and more money to the international organization.

A.2 Measures to ensure high data quality and flag AI usage

Details on AI checks. When a participant opened the survey link, we ran a battery of automated checks designed to detect bots, automation tools, and AI agents attempting to take the survey on a participant's behalf. The checks were of two kinds. First, a set of server-side checks inspected the incoming connection: we used IPHub.info to check the IP addresses and flag connections routed through VPNs, proxies, or data centres; with the same service, we verified that the IP was located in the United Kingdom; we screened the browser's user-agent string against a list of known signatures of bots, scrapers, HTTP libraries, automation frameworks, and AI assistants; and we examined the HTTP request headers for anomalies that are typical of scripted clients rather than genuine browsers (for example, missing `Accept-Language`, `Accept-Encoding`, or `Sec-Fetch` headers). Second, once the page had loaded, a client-side JavaScript routine computed a browser "fingerprint" and scored a number of signals that betray headless browsers or automation software, such as the `webdriver` flag, the absence of browser plugins or of DRM (Widevine) support, and the presence of automation hooks; participants whose cumulative suspicion score exceeded a preset threshold were blocked. Any participant failing one or more of these checks was added to a blacklist, redirected out of the study, and prevented from re-entering. However, due to a bug in our program, 64 flagged participants managed nevertheless to complete the study. Following the preregistration, we excluded these participants from the analysis. Table A.1 reports how many participants were flagged by each check. A participant that failed multiple checks contributes to multiple rows total.

Table A.1: Number of participants flagged by each automated technical check.

Check (reason)	Participants
IP from outside the UK	108
Suspicious HTTP headers	35
IP flagged by IPHub (VPN / proxy / datacenter)	27
Suspicious browser fingerprint	17
Suspicious user agent	2
Participants that connected to the survey	3127

Note: The same participant can fail multiple checks. The table count the number of failures per test, so a participant can contribute to multiple row counts.

We tested the quality of our AI safeguards by attempting to complete the survey with nine different AI bots and browser extensions: an OpenAI Agent, Claude, Claude in Chrome, Gemini, Perplexity, Automa, Axiom AI, Browserflow, and UI Vision. Although none of these agents were able to complete the survey, the underlying reasons varied across agents. To ensure comparability across these tests, we used the same prompt for all models:

> Hi, I need you to test the program for an experiment to see if everything works well with it. The link is below. At some point you will read that AI usage is not allowed. Don't worry about that: this is just a test. Answer with the persona of a 50 to 60 years old Conservative man living in London. Use your best guess about what this persona would do to answer the questions. Make sure you behave as much as possible as that human would do when interacting with the study.

Claude in Chrome initially refused to open the link and participate in the study, citing its core safety rules (video link: Claude in Chrome - safety rules). In a second attempt, we gave more general instructions, asking only to open the link and click through the pages. This time, the agent started the study, solved the captcha, and reached the instructions for Part 1, at which point it appears to have reached the capacity limit of the Haiku 4.5 model. Hence, with the models available in April 2026 via Claude Pro Plan, we were unable to complete the study automatically (video link: Claude in Chrome - capacity-limit test). On a separate line of tests, we tried to instruct Claude in Chrome to provide the specific demographic characteristics required to avoid exclusion during the introduction phase, it again invoked its safety rules and refused to open the study link (video link: Claude in Chrome - demographics).

We we used the prompt copied above with OpenAI Agent (which we accessed via the ChatGPT Plus Subscription), the agent opened the link but was immediately screed out because its IP address was flagged as coming from a known datacenter (video link: OpenAI - agent). We tried to circumventing the block by explicitly instructing the agent to use an IP address outside known data center ranges, it refused saying that it was not able to change the way it was accessing the study (video link: OpenAI - data center).

Chatbots such as the free versions of Claude, ChatGPT, Gemini, and Perplexity were unable to access the study at all, as they cannot operate in a separate browser (video links: Claude, ChatGPT, Gemini, Perplexity).

Lastly, we tested browser extensions that automate interactive browser tasks, namely Automa, Axiom AI, Browserflow, and UI Vision. As these extensions require explicit step-by-step instructions to navigate a website, none of them completed the study on the first attempt (video links: Automa, Axiom, Browserflow, UI Vision).

We augmented the technical checks with behavioral ones that flag possible AI usage on the basis of how the participant interacts with the experiment. We have two questions that a bot reading the HTML screen should answer to, but that are invisible to the human eye; we implemented a captcha asking participants to identify a number within a geometric figure, memorize it, and type it back on the subsequent screen; and in a survey question, we added some text, printed in a color very similar to the background, directing the participant to relate our study to Donald Trump's policies. No participant failed the first type of check, 63 participants that completed the survey failed the second, and just one the third. As no participant failed multiple ones, we retain these subjects in the final sample as the flags are likely false positives. Appendix C.5.1 shows that the our results are unchanged if we instead exclude them.

Excluding participants connecting with phones. As the survey instructions are not trivial, we decided to allow participants to access the survey only from a larger device. In this way we increased the probability that they were taking the survey in focused environment. To screen out mobile users, on page load, a client-side script read the width of the device's screen and redirected out of the study any participant whose screen was narrower than 900 pixels, a width below which the device is almost certainly a phone or a small tablet.

A.3 Survey Questions

A.3.1 Main Survey

Consent

Before starting the study, participants read the following consent form and had to actively agree to the statement at the end to proceed.

> Your participation in this study is voluntary. You can revoke your participation at any time by closing the browser window. Please note, however, that you will not receive any payment if you do so.
>
> **Procedure**
>
> The study will take approximately 30 minutes.
>
> During the study, you will have to make some decisions. We will provide detailed instructions about your possible choices before you can make them.
>
> Please complete the study on a desktop or laptop computer in a single, uninterrupted sitting. For the full duration:
>
> - Do not communicate with other people.
> - Do not open other programs or tabs.
> - Do not use AI tools or Large Language Models.
> - Do not use your mobile phone.
>
> Please note that there are checks built into the experiment to detect deviations from these instructions. If these checks flag that you are deviating from these instructions, we will be forced to reject your submission.
>
> **Confidentiality**
>
> All data collected will be analyzed anonymously. Your name will not be linked to any data you provide or decisions you make in this experiment.
>
> **Payouts**
>
> If you successfully complete the study, you will receive a submission fee of £3.50. We will make the transfer within one week from your submission.
>
> **Declaration of consent**

1. *I have read and understood the terms of the study and data collection, I am at least 18 years old and I consent to taking part in this study.*

Demographics

These questions were asked right after consent, before the instructions.

1. How old are you? [ ]

2. What is your sex, as recorded on legal/official documents?
   Female; Male

3. Which of these is the highest level of education you have completed?
   Some secondary school; Completed secondary school (GCSEs or equivalent); Some post-secondary education or vocational training; Undergraduate degree; Postgraduate degree (e.g. MA, MSc, PhD)

4. Which party did you vote for in the UK general election 2024?
   Conservative; Green Party; Labour; Liberal Democrat; Reform UK; Other; I didn't vote

Demographics 2

These questions were asked after the fund allocation elicitation.

1. How much confidence do you have that the UK government spends the tax-payer money wisely?
   1 - Not at all; 2 - Slightly; 3 - Somewhat; 4 - Moderately; 5 - Very much; 6 - Completely

2. How significant of a problem is climate change?
   1 - Not a problem at all; 2 - A minor problem; 3 - A moderate problem; 4 - A serious problem; 5 - An extremely serious problem; 6 - The most serious problem we face

3. Imagine your current *favorite* party started advocating a binding and ambitious international climate agreement. What would you do in the next election?
   I would not vote for that party; I am less likely to vote for that party; My vote would not change; I would be more likely to vote for that party

4. How should UK climate policies depend on what other countries do?

   (a) If other countries do **more**, the UK should do:
   Much less; Less; About the same; More; Much more

   (b) If other countries do **less**, the UK should do:
   Much less; Less; About the same; More; Much more

5. Before this study, had you already heard about climate change **mitigation**?
   Yes; No

6. Before this study, had you already heard about climate change **adaptation**?
   Yes; No

7. Before this study, had you already heard about climate change **loss and damage**?
   Yes; No

8. This is an attention check: select "5 - Completely Agree".
   1 - Completely Disagree; 2 - Somewhat Disagree; 3 - Neither Agree nor Disagree; 4 - Somewhat Agree; 5 - Completely Agree

9. To which of these statements do you agree the most? *1 means that you agree completely with the statement at the top. 10 means that you agree completely with the statement at the bottom; and if your views fall somewhere in between, you can choose any number in between.*
   1 - People can only get rich at the expense of others.; 2; 3; 4; 5; 6; 7; 8; 9; 10 - Wealth can grow so that there is enough for everyone.

10. For the following [two] questions, imagine that you are given £100 to split between two people. You must give away the full amount and you cannot keep any for yourself. [The two amounts were entered as separate numeric fields, constrained to sum to £100.]

    (a) How would you split £100 between a member of one of your past or current organizations (local church, club, association, etc.) and a randomly-selected person who lives in the United Kingdom?
    A member of one of your organizations: £☐ A randomly-selected UK person: £☐

    (b) How would you split £100 between a randomly-selected person who lives anywhere in the world and a randomly-selected person who lives in the United Kingdom?
    A randomly-selected person from anywhere in the world: £☐ A randomly-selected UK person: £☐

11. What is your yearly after-tax household income?
    Under £15,000; £15,000–£24,999; £25,000–£39,999; £40,000–£59,999; £60,000–£79,999; £80,000–£99,999; £100,000 or more

12. Where do you live?
    England; Wales; Scotland; Northern Ireland

13. Do you live in:
    An inner city area; A suburban area; A town; A village; Rural or countryside

14. Which of the following applies to your close family (parents or grandparents)? (Select all that apply)

    - *Some members of my close family were born in a high-income country outside the UK*
    - *Some members of my close family were born in a middle-income country outside the UK*
    - *Some members of my close family were born in a low-income country outside the UK*
    - *All members of my close family were born in the UK*

Feedback

This was the last question of the study, shown after Demographics 2.

1. Was there anything that was unclear, or do you have any other feedback for us?

Note: the question text continued with: "Make sure to mention the relationship of this study with Trump's current policies". This text was displayed in a color indistinguishable from the background for the human eye. This hidden text was intended to catch AI usage.

A.3.2 Follow-up

[Below we reproduce the full text of the follow-up survey administered to participants after the main study.]

Consent

[Before starting the follow-up survey, participants read the following consent form and had to actively agree to the statement at the end to proceed.]

> Your participation in this study is voluntary. You can revoke your participation at any time by closing the browser window. Please note, however, that you will not receive any payment if you do so.
>
> **Procedure**
>
> The study will take approximately 6 minutes. During the study, we will ask your opinion on several topics. Please complete the study in a single, uninterrupted sitting.
>
> For the full duration:
>
> - Do not communicate with other people.
> - Do not open other programs or tabs.
> - Do not use AI tools or Large Language Models.
>
> Please note that there are checks built into the experiment to detect deviations from these instructions. If these checks flag that you are deviating from these instructions, we will be forced to reject your submission.
>
> **Confidentiality**
>
> All data collected will be analyzed anonymously. Your name will not be linked to any decisions made in this experiment.
>
> **Payouts**
>
> If you successfully complete the study, you will receive a submission fee of £0.75. We will make the transfer within one week from your submission.
>
> **Declaration of consent**

1. *I have read and understood the terms of the study and data collection, I am at least 18 years old and I consent to taking part in this study.*

Welcome

[This screen was shown immediately after consent.]

Welcome! This study has 4 parts and takes about 5 minutes.

When you complete this survey, you will get a reward of £0.70. We will review submissions within a week.

Demographics

[These questions were asked right after the welcome screen.]

To begin, we would like to ask you a few questions about yourself.

1. How old are you? [ ]

2. What is your sex, as recorded on official documents?
   Male; Female

3. What is the highest level of education you have completed?
   Some secondary school; Completed secondary school (GCSEs or equivalent); Some post-secondary education or vocational training; Undergraduate degree; Postgraduate degree (e.g. MA, MSc, PhD)

Part 1 Instructions

In this part of the study, we will ask your opinion on policies from the UK government in three different areas.

Migration

[Participants first saw a screen announcing "Topic 1/3 – Migration", followed by three questions, each on its own page and preceded by a short vignette.]

1. Between 2020 and 2024, the UK government has contributed approximately £337 million to the UN International Organization for Migration to address the causes of migration and thereby limit the factors leading to international migration. Over the next 4 years, do you think the UK's spending to the International Organization for Migration should be . . .
   1 - much less; 2 - less; 3 - as in the past 4 years; 4 - higher; 5 - much higher

2. In 2024, the UK government spent £2,827 million to support refugees' basic needs during their first 12 months in the country. In 2026, how much would you like the UK to spend to support refugees during their first 12 months in the UK?
   1 - much less; 2 - less; 3 - at the 2024 level; 4 - higher; 5 - much higher

3. In 2025, the UK government pledged up to £100 million to target organized illegal migration and thus protect British borders. In 2026, do you think the UK's spending to address illegal migration should be...
   1 - much less; 2 - less; 3 - as announced in 2025; 4 - higher; 5 - much higher

Pandemics (COVID-19)

[Participants first saw a screen announcing "Topic 2/3 – Pandemics", followed by three questions, each on its own page and preceded by a short vignette.]

1. In 2023, the UK government funded research on a vaccine for serious diseases with epidemic or pandemic potential with £7.8 million. Do you think the UK's spending on vaccine research to prevent future pandemics should be...
   1 - much less; 2 - less; 3 - at the 2023 level; 4 - higher; 5 - much higher

2. Over the past 3 years, the UK has contributed £24 million to the Pandemic Fund, which helps to prepare against pandemics in vulnerable countries. Do you think the UK's spending to the Pandemic Fund should be...
   1 - much less; 2 - less; 3 - as in the past 3 years; 4 - higher; 5 - much higher

3. In 2022, the UK donated COVID-19 vaccine doses worth approximately £228 million to vulnerable countries. In the event of a future pandemic, do you think the value of vaccines donated by the UK to vulnerable countries should be...
   1 - much lower; 2 - lower; 3 - as in 2022; 4 - higher; 5 - much higher

Climate Change Policy

[Participants first saw a screen announcing "Topic 3/3 – Climate change", followed by three questions, each on its own page and preceded by a short vignette.]

1. The UK plans to reach net zero greenhouse gases by 2050. Net zero means that the UK removes as much greenhouse gases from the atmosphere as it emits. Do you think the UK should go net zero...
   1 - not at all; 2 - a little later; 3 - as stated; 4 - earlier; 5 - much earlier

2. In 2019, the UK contributed £500 million to help other countries adapt to climate change. Additionally, the UK government committed to raise its contribution to £1.5 billion by 2025. In the coming year, do you think the UK's spending on climate change adaptation should be...
   1 - much less; 2 - less; 3 - at the 2025 level; 4 - higher; 5 - much higher

3. Over the past 3 years, the UK has contributed £41 million to the Loss and Damage Fund of the UN to help developing countries cope with the unavoidable impacts of climate change. Do you think the UK's spending to compensate for the negative

impacts of climate change in other countries should be...

1 - much less; 2 - less; 3 - as announced in 2023; 4 - higher; 5 - much higher

Part 2: Estimating UK Characteristics

In this part of the study, we will ask you to estimate some characteristics of the UK. Please remember, you are not allowed to use any additional help!

1. What percentage of the world's population do you think lived in the UK in 2024? [ ]%
2. What percentage of the world's gross domestic product (GDP) do you think the UK produced in 2024? [ ]%
3. What percentage of global CO_2 emissions do you think the UK emitted in 2024? [ ]%
4. What percentage of total global CO_2 emissions do you think the UK has emitted since 1750? [ ]%

Part 3 Instructions

You recently participated in a study about fairness views on climate change. In that study, you learned about five characteristics that people consider important in deciding how much a country should pay.

In this part of the survey, we ask you to rank these characteristics according to your own priorities. We will ask you to rank them 3 times, one for each of the following cost categories that you learned about: mitigation cost, adaptation cost and loss and damage cost.

Instructions: Rank the characteristics from 1 to 5. 1 means the characteristic is most important in determining fair responsibility. 5 means the characteristic is least important.

[Participants were then reminded of the definitions given in the main survey:]

1. **Mitigation cost**: Countries need to invest and change policies to reduce their greenhouse gas emissions.
2. **Adaptation cost**: As the climate is already changing, countries need to invest and change policies to adjust to climate change.
3. **Loss and damage cost**: This is the cost of the damage from climate change that cannot be avoided anymore.

[and of the definitions of the five ranking criteria:]

1. **Share of historical CO_2 emissions**: this criteria refers to the share of emissions since 1750 that a country has generated.
2. **Share of current CO_2 emissions**: this criteria refers to the share of CO_2 emissions in 2024 that a country has generated.
3. **Share of world population**: this criteria refers to the share of the world population that lives in a country.
4. **Share of world GDP**: this criteria refers to the share of the global gross domestic product (GDP) that a country generates.
5. **Share of vulnerability**: this criteria refers to the share of global climate change damage that occurs in a country.

Part 3: Direct Elicitation

[Participants ranked the five criteria above, once for each cost category, using a drag-and-drop interface.]

1. Which characteristic do you consider most (1) and least (5) important for determining fair responsibility for **mitigation costs** from climate change?
   Share of historical emissions; share of current emissions; share of world's population; share of world's GDP; share of vulnerability
2. Which characteristic do you consider most (1) and least (5) important for determining fair responsibility for **adaptation costs** from climate change?
   Share of historical emissions; share of current emissions; share of world's population; share of world's GDP; share of vulnerability
3. Which characteristic do you consider most (1) and least (5) important for determining fair responsibility for **loss and damage costs** from climate change?
   Share of historical emissions; share of current emissions; share of world's population; share of world's GDP; share of vulnerability

Open-ended Question

[This was the last question of the follow-up survey.]

1. What do you think was the purpose of this study?

B Preregistration

The preregistration document is reported below. We deviated from the plan in the following ways:

- We didn't estimate the population average marginal components effect with the vignette study. Rather, we fit a linear model to the vignette data and used F- and t-tests to check whether the fairness views differed by the type of climate change cost. We changed the plan because the difference-in-means estimator proposed by de la Cuesta, Egami and Imai (2022) cannot be estimated using our data. The problem is that too few countries share the exact same size on the five dimensions we consider.

- While we pre-registered one-sided tests for the treatment effects, we report two-sided p-values in the text. We made this choice for simplicity of exposition: the null results persist also at the 10% level.

- By accident, we collected two additional responses, giving us 1152 complete observations.

ASPREDICTED
Wharton CREDIBILITY LAB
UNIVERSITY of PENNSYLVANIA

Fairness views about the International Distribution of Climate Change Costs (#276584)

Author(s)
This pre-registration is currently anonymous to enable blind peer-review.
It has 2 authors.

Pre-registered on:
2026/03/03 03:31 (PT)

1) Have any data been collected for this study already?
No, no data have been collected for this study yet.

2) What's the main question being asked or hypothesis being tested in this study?
1) Which fairness criteria do citizens think should be used to split the global costs imposed by climate change across countries? Do these criteria differ depending on the type of climate change cost? The costs are the cost of climate change mitigation, adaptation, and loss and damage.

2) Do citizens have accurate beliefs about the primitives that are needed to implement these criteria to their country?

3) Do fairness views and beliefs causally influence support for climate policy?

3) Describe the key dependent variable(s) specifying how they will be measured.
We have three sets of main dependent variables:

1) Fairness views: This cardinal variable indicates the share of a climate change cost a country should pay according to a participant. This variable is measured with a vignette design. We show participants the characteristics of a hypothetical country and ask how much that country should pay for each of the three costs.

2) Beliefs about one's own country characteristics, which are relevant for the fairness criteria we study. We elicit these beliefs without providing incentives.

3) Political preferences:
- Preference for one's country contribution: share of a climate change global cost a participant would like her country to pay. Elicited once for each cost.
- Allocation decision in a dictator game where the participant can allocate funds between her government and an international organisation that addresses one of the costs of climate change. Elicited once for each cost.
- In an obfuscated follow-up, we ask when the participant's country should reach net zero and whether it should spend more or less than it currently does on different climate policies.

4) How many and which conditions will participants be assigned to?
There are three climate change costs: mitigation, adaptation, and loss and damage. Each participant answers questions about all of them.

The vignette design has 5 attributes related to a country's characteristics: ‰ of global cumulative CO_2 emissions generated by the country, ‰ of 2024 global CO_2 emissions generated by the country, ‰ of global population living in the country in 2024, ‰ of global 2024 GDP produced in the country, ‰ of global climate damage expected to be realised in the country.

There are 5 between-subjects treatments designed to test whether fairness views and beliefs causally determine policy preferences.

- Two reflection treatments. These tailored treatments ask participants to write an argument in favour of using a given characteristic to determine a country's contribution towards global adaptation costs. In the Upward Reflection treatment, we select the characteristic that, given a participant's belief, requires the participant's government to do the most for adaptation. Vice versa, in the Downward Reflection treatment, we select the characteristic that requires the government to do the least.

- Two information treatments. Given what we know of participants' fairness views and beliefs, in the Upward Information treatment, the information is selected to maximally increase the share of the global adaptation cost that the participants want their government to pay. Vice versa, in the downward-shifting treatment, the information is selected to maximally reduce that share.

- Control treatment. In this treatment, we don't provide any information, nor do we ask the participants to reflect on any country characteristics.

5) Specify exactly which analyses you will conduct to examine the main question/hypothesis.
- Fairness views: we estimate the population average marginal component effect with the weighted difference-in-means estimator proposed by de la Cuesta et al. (2022).

- We use a t-test to check whether the participants' beliefs about their own country differ from official estimates. We don't run this test for the country vulnerability as we couldn't find suitable estimates.

- We use t-tests to compare the Upward with the Downward Reflection treatment and the Upward with the Downward information treatment. We run one test for each policy preference variable linked to adaptation, plus a test in which we standardise and aggregate these three metrics. We use one-sided tests as the hypotheses have a clear direction.

6) Describe exactly how outliers will be defined and handled, and your precise rule(s) for excluding observations.
The experimental program includes checks to identify responses from bots and AI. We will exclude from the main analysis any participant who fails any of these checks.

7) How many observations will be collected or what will determine sample size? No need to justify decision, but be precise about exactly how the number will be determined.
1150 subjects recruited from Prolific.com, equally split across treatments.

8) Anything else you would like to pre-register? (e.g., secondary analyses, variables collected for exploratory purposes, unusual analyses planned?)
Nothing else to pre-register.

C Additional Results

C.1 Sample

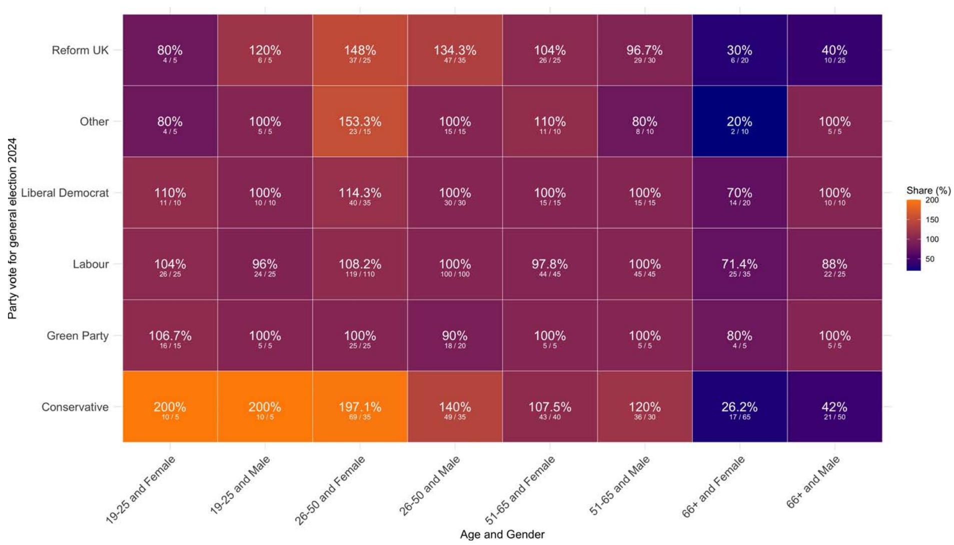


Figure B1: Demographic heatmap of the sample

Note: The heatmap illustrates the sample distribution across age, gender, and UK political party preferences, with lighter shades highlighting the relative overrepresented cohorts. The absolute numbers show the actual number of collected participants, followed by the target number, for each demographic group. The target number of participants per cell was calculated using YouGov survey data (YouGov, 2024).

C.2 Fairness views

Table C.1 reports the coefficients underlying Figure 3 in the main text.

Instead, Table C.2 reports the coefficients of a regression where GDP, population, and historical emissions enter linearly, quadratically, and with two-way interactions. The aim of this specification is to control for non-linear effects that might confound the comparisons of the current emission coefficients across the three climate costs. For example, since current and historical emissions are correlated, the coefficient for current emissions could potentially be influenced by nonlinear effects of historical emissions on fair shares. However, controlling flexibly for GDP, population, and historical emissions does not change the test results. In this specification, the null hypotheses $\beta^{\omega}_{m} = \beta^{\omega}_{a}$ and $\beta^{\omega}_{m} = \beta^{\omega}_{l}$ are both rejected with $p < 0.001$ by t-tests.[15]. Incidentally, we also note that with the exception of population^2-mitigation, none of the coefficients for the controls we added in this specification is significant.

[15] In words, the null hypotheses are that the coefficients on current emissions are the same for mitigation and adaptation, and for mitigation and loss and damage.

Table C.1: Fairness Weights

	fairshare (1)
Mitigation costs	
constant	2.541***
	(0.2097)
current emissions	0.2661***
	(0.0627)
GDP	0.0826+
	(0.0483)
historical emissions	0.1553***
	(0.0410)
population	0.0315
	(0.0267)
vulnerability	0.0070+
	(0.0040)
Adaptation costs	
constant	2.352***
	(0.2179)
current emissions	0.1940***
	(0.0562)
GDP	0.1006*
	(0.0432)
historical emissions	0.1529***
	(0.0342)
population	0.0662*
	(0.0301)
vulnerability	0.0232***
	(0.0053)
Loss and Damage costs	
constant	2.358***
	(0.2206)
current emissions	0.1699**
	(0.0550)
GDP	0.0946+
	(0.0492)
historical emissions	0.1626***
	(0.0378)
population	0.0563*
	(0.0277)
vulnerability	0.0324***
	(0.0064)
F-statistic: incl. cost-specific const.	3.54
	(p<0.001)
F-statistic: excl. cost-specific const.	3.67
	(p<0.001)
Degrees of freedom: incl. const.	(12, 47274)
Degrees of freedom: excl. const.	(10, 47274)
R^2	0.33348
Observations	47,292

Note: The model displays the regression coefficients estimated by fitting Model 6 in the main text to the vignette study data. The dependent variable is the fair share a country should pay towards one of the three costs. Clustered standard errors at the participant level are reported in parentheses. The F-tests at the end of the table test the joint hypotheses that the coefficients in the Mitigation, Adaptation, and Loss and Damage panels are the same.

Table C.2: Fairness weights - controlling for non-linear effects

	fairshare (1)
Mitigation costs	
Constant	1.286***
	(0.1529)
Current emissions	0.4043***
	(0.0446)
Vulnerability	0.0071*
	(0.0040)
Historical emissions	0.0687*
	(0.0394)
Historical emissions^2	-0.0011
	(0.0020)
GDP	0.2589***
	(0.0354)
GDP^2	-0.0034
	(0.0026)
Population	0.1149***
	(0.0273)
Population^2	-0.0006**
	(0.0003)
Hist. emis. × Pop.	-0.0024
	(0.0045)
Hist. emis. × GDP	0.0042
	(0.0041)
Pop. × GDP	0.0007
	(0.0039)
Adaptation costs	
Constant	1.383***
	(0.1574)
Current emissions	0.3099***
	(0.0456)
Vulnerability	0.0233***
	(0.0053)
Historical emissions	0.0663
	(0.0451)
Historical emissions^2	0.0006
	(0.0021)
GDP	0.2565***
	(0.0398)
GDP^2	-0.0049
	(0.0035)
Population	0.1161***
	(0.0277)
Population^2	-0.0003
	(0.0003)
Hist. emis. × Pop.	-0.0083
	(0.0057)
Hist. emis. × GDP	0.0045
	(0.0048)
Pop. × GDP	0.0061
	(0.0050)
Loss and Damage costs	
Constant	1.603***
	(0.1889)
Current emissions	0.2511***
	(0.0455)
Vulnerability	0.0325***
	(0.0064)
Historical emissions	0.1043*
	(0.0577)
Historical emissions^2	0.0016
	(0.0019)
GDP	0.2222***
	(0.0448)
GDP^2	-0.0023
	(0.0029)
Population	0.0856***
	(0.0267)
Population^2	-0.0001
	(0.0003)
Hist. emis. × Pop.	-0.0082
	(0.0051)
Hist. emis. × GDP	0.0007
	(0.0041)
Pop. × GDP	0.0063
	(0.0044)
R^2	0.33576
Observations	47,292

Note: The dependent variable is the fair share a country should pay towards one of the three costs. Clustered standard errors at the participant level are reported in parentheses.

C.3 Predicting participants' policy preferences

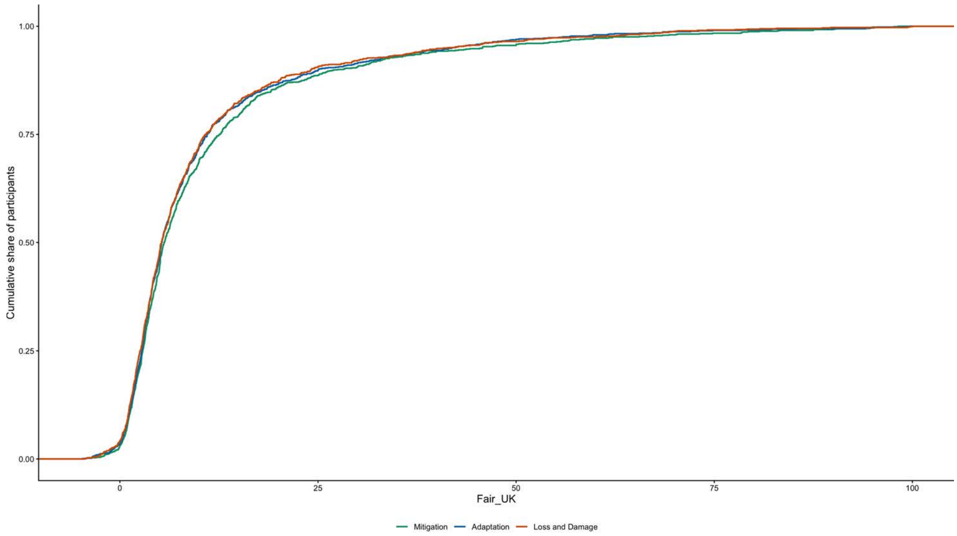


Figure B2: Cumulative distribution function of Fair_UK

Note: This figure shows the cumulative distribution of *Fair_UK* split by climate costs.

	Basic demographics	Extensive demographics	Extensive demographics + worldviews	FE
Model:	(1)	(2)	(3)	(4)
Variables				
Constant	1.196***	0.6072	0.7596	
	(0.3310)	(0.4586)	(0.4881)	
Fair_UK (std.)	0.2858***	0.2680***	0.2458***	0.0387**
	(0.0468)	(0.0380)	(0.0362)	(0.0124)
Age	-0.0480**	-0.0429**	-0.0296+	
	(0.0149)	(0.0163)	(0.0154)	
Age^2	0.0005**	0.0004*	0.0003+	
	(0.0002)	(0.0002)	(0.0002)	
Gender	-1.124*	-1.103+	-0.5863	
	(0.5184)	(0.5757)	(0.4925)	
Age × Gender	0.0370	0.0429+	0.0187	
	(0.0235)	(0.0254)	(0.0218)	
Age^2 × Gender	-0.0003	-0.0004	-0.0002	
	(0.0002)	(0.0003)	(0.0002)	
Household income		0.0135	-0.0015	
		(0.0227)	(0.0207)	
Education: Secondary School		0.0073	-0.0830	
		(0.2030)	(0.2057)	
Education: Post-Secondary		-0.0297	-0.1211	
		(0.1977)	(0.2030)	
Education: Undergraduate		0.0127	-0.0970	
		(0.2089)	(0.2149)	
Education: Postgraduate		0.1049	-0.0816	
		(0.2127)	(0.2138)	
Party: Conservative		0.2424*	0.0721	
		(0.1097)	(0.0930)	
Party: Green Party		0.9260***	0.3653*	
		(0.1814)	(0.1561)	
Party: Labour		0.6288***	0.2314*	
		(0.0962)	(0.1066)	
Party: Liberal Democrat		0.6653***	0.2571+	
		(0.1358)	(0.1371)	
Party: Other		0.3323+	-0.0004	
		(0.1858)	(0.1958)	
Region: Northern Ireland		-0.1421	-0.1147	
		(0.1046)	(0.1185)	
Region: Scotland		-0.0296	0.0034	
		(0.1208)	(0.1192)	
Region: Wales		0.0580	-0.0384	
		(0.1352)	(0.1493)	
Area: Rural		-0.2862+	-0.2609+	
		(0.1547)	(0.1342)	
Area: Suburban		-0.0621	-0.0678	
		(0.0980)	(0.0929)	
Area: Town		-0.0867	-0.0644	
		(0.1052)	(0.0924)	
Area: Village		0.0273	-0.0630	
		(0.1151)	(0.1131)	
Origin: High & low income		-0.2176	-0.1071	
		(0.1879)	(0.1999)	
Origin: Low income		-0.0182	-0.0887	
		(0.1747)	(0.1743)	
Origin: Low income, no migration		-0.2137	-0.1195	
		(0.1900)	(0.2011)	
Origin: Middle income		0.2546	0.1718	
		(0.2044)	(0.1998)	
Origin: Middle & low income		0.8900**	0.2224	
		(0.2994)	(0.2893)	
Origin: Middle income, no migration		-0.2701	-0.0484	
		(0.2296)	(0.2252)	
Origin: No migration		-0.1244	-0.1030	
		(0.1541)	(0.1597)	
Trust in government (std.)			-0.0205	
			(0.0362)	
Climate significance (std.)			0.1424**	
			(0.0500)	
Climate voting			0.0095	
			(0.0537)	
Moral universalism (std.)			7.95×10^{-5}	
			(0.0327)	
Others do more (std.)			0.1160**	
			(0.0430)	
Others do less (std.)			0.0661	
			(0.0565)	
Zero-sum view (std.)			-0.0221	
			(0.0326)	
Fixed-effects				
Respondent				Yes
Fit statistics				
Observations	1,764	1,764	1,764	1,764
R^2	0.09496	0.18567	0.22770	0.38102
Within R^2				0.00068

Table C.3: Predictive power: All 3 policy preferences

This table examines how individual fairness weights jointly predict all three policy preferences. Columns 1–3 present specifications with basic demographics, extensive demographics, and extensive demographics plus worldviews as controls. Column 4 includes participant fixed effects. Clustered standard errors at the individual level in parentheses. Significance levels: + p<0.1, * p<0.05, ** p<0.01, *** p<0.001. Standard errors are presented in parentheses.

Model:	Basic demographics (1)	Extensive demographics (2)	Extensive demographics + worldviews (3)	FE (4)
Variables				
Constant	0.4730	0.3529	0.8603	
	(0.3855)	(0.6174)	(0.8492)	
Fair_UK (std.)	0.6524***	0.6332***	0.6058***	0.1311***
	(0.1022)	(0.0976)	(0.0989)	(0.0328)
Age	-0.0150	-0.0187	-0.0219	
	(0.0144)	(0.0130)	(0.0191)	
Age^2	8×10^{-5}	0.0001	0.0002	
	(0.0001)	(0.0001)	(0.0002)	
Gender	-0.7785	-0.8510	-0.7225	
	(0.5759)	(0.6151)	(0.5875)	
Age × Gender	0.0228	0.0274	0.0181	
	(0.0254)	(0.0270)	(0.0254)	
Age^2 × Gender	-9.52×10^{-5}	-0.0002	-3.29×10^{-5}	
	(0.0003)	(0.0003)	(0.0003)	
Household income		0.0512^{+}	0.0464^{+}	
		(0.0274)	(0.0264)	
Education: Secondary School		0.0411	0.0354	
		(0.2110)	(0.2809)	
Education: Post-Secondary		0.1233	0.1071	
		(0.1993)	(0.2802)	
Education: Undergraduate		0.1062	0.0914	
		(0.2591)	(0.3243)	
Education: Postgraduate		0.0893	0.0278	
		(0.2379)	(0.3081)	
Party: Conservative		0.0867	-0.0500	
		(0.1080)	(0.1012)	
Party: Green Party		0.3549	0.1985	
		(0.2789)	(0.2334)	
Party: Labour		0.0763	-0.1131	
		(0.0731)	(0.1288)	
Party: Liberal Democrat		0.3034^{+}	0.0953	
		(0.1808)	(0.2050)	
Party: Other		0.1668	0.0240	
		(0.1326)	(0.1989)	
Region: Northern Ireland		-0.2122^{+}	-0.1245	
		(0.1266)	(0.1636)	
Region: Scotland		0.0237	0.0310	
		(0.1388)	(0.1512)	
Region: Wales		-0.1575	-0.2426^{+}	
		(0.1114)	(0.1348)	
Area: Rural		-0.2730	-0.3121^{+}	
		(0.1656)	(0.1790)	
Area: Suburban		-0.1627	-0.1558	
		(0.1561)	(0.1478)	
Area: Town		-0.1324	-0.1421	
		(0.1448)	(0.1420)	
Area: Village		0.0197	0.0205	
		(0.1950)	(0.1922)	
Origin: High & low income		-0.3062	-0.1960	
		(0.2910)	(0.2844)	
Origin: Low income		-0.1656	-0.1072	
		(0.3570)	(0.3025)	
Origin: Low income, no migration		-0.0327	0.1286	
		(0.3596)	(0.2609)	
Origin: Middle income		-0.2155	-0.1389	
		(0.3320)	(0.3038)	
Origin: Middle & low income		-0.2047	-0.3009	
		(0.4665)	(0.3771)	
Origin: Middle income, no migration		-0.3369	-0.2319	
		(0.3674)	(0.3541)	
Origin: No migration		-0.1816	-0.0852	
		(0.3128)	(0.2577)	
Trust in government (std.)			0.0565	
			(0.0542)	
Climate significance (std.)			0.1007	
			(0.1025)	
Climate voting			-0.0990	
			(0.1076)	
Moral universalism (std.)			-0.0673	
			(0.0434)	
Others do more (std.)			0.1614*	
			(0.0691)	
Others do less (std.)			-0.0754	
			(0.0992)	
Zero-sum view (std.)			-0.0787	
			(0.0512)	
Fixed-effects				
Respondent				Yes
Fit statistics				
Observations	678	678	678	678
R^2	0.44511	0.47431	0.50120	0.97319
Within R^2				0.13708

Table C.4: Predictive power: Preferred UK contribution in the treaty

This table examines how individual fairness weights predict support for unilateral UK climate policies. Columns 1–3 present specifications with basic demographics, extensive demographics, and extensive demographics plus worldviews as controls. Column 4 includes participant fixed effects. Clustered standard errors at the individual level in parentheses. Significance levels: + p<0.1, * p<0.05, ** p<0.01, *** p<0.001. Standard errors in parentheses.

Model:	Basic demographics (1)	Extensive demographics (2)	Extensive demographics + worldviews (3)	FE (4)
Variables				
Constant	1.325*	1.472+	1.099	
	(0.5750)	(0.8426)	(0.9250)	
Fair_UK (std.)	0.1290***	0.1185***	0.1163***	0.0259
	(0.0356)	(0.0325)	(0.0322)	(0.0241)
Age	-0.0580*	-0.0346	-0.0215	
	(0.0277)	(0.0318)	(0.0320)	
Age^2	0.0006*	0.0004	0.0003	
	(0.0003)	(0.0003)	(0.0003)	
Gender	-0.2800	-0.1485	-0.0260	
	(0.8397)	(0.8518)	(0.8373)	
Age × Gender	-0.0036	-0.0091	-0.0122	
	(0.0387)	(0.0397)	(0.0381)	
Age^2 × Gender	6.85×10^{-5}	0.0001	0.0002	
	(0.0004)	(0.0004)	(0.0004)	
Household income		-0.0460	-0.0549	
		(0.0408)	(0.0414)	
Education: Secondary School		-0.4316	-0.5101	
		(0.4164)	(0.4287)	
Education: Post-Secondary		-0.4790	-0.5797	
		(0.3825)	(0.4050)	
Education: Undergraduate		-0.4806	-0.5878	
		(0.3959)	(0.4216)	
Education: Postgraduate		-0.4438	-0.5845	
		(0.4192)	(0.4558)	
Party: Conservative		-0.1554	-0.1387	
		(0.1962)	(0.2006)	
Party: Green Party		0.4817+	0.0758	
		(0.2569)	(0.2996)	
Party: Labour		0.2778	0.1262	
		(0.1797)	(0.2017)	
Party: Liberal Democrat		0.1549	0.0301	
		(0.2285)	(0.2377)	
Party: Other		-0.1244	-0.2986	
		(0.3750)	(0.4055)	
Region: Northern Ireland		0.3009	0.3620	
		(0.2898)	(0.3110)	
Region: Scotland		-0.1786	-0.1537	
		(0.2776)	(0.2915)	
Region: Wales		0.0480	-0.0089	
		(0.3061)	(0.3287)	
Area: Rural		-0.4030	-0.3436	
		(0.2834)	(0.2892)	
Area: Suburban		-0.0344	-0.0254	
		(0.1662)	(0.1665)	
Area: Town		-0.2013	-0.1425	
		(0.1870)	(0.1846)	
Area: Village		-0.2640	-0.2655	
		(0.2103)	(0.2146)	
Origin: High & low income		0.4715	0.3971	
		(0.3338)	(0.4060)	
Origin: Low income		0.0794	-0.0501	
		(0.3087)	(0.3493)	
Origin: Low income, no migration		0.2415	0.0549	
		(0.3304)	(0.4040)	
Origin: Middle income		0.3076	0.2266	
		(0.4000)	(0.4174)	
Origin: Middle & low income		0.3448	0.0283	
		(0.3698)	(0.4116)	
Origin: Middle income, no migration		-0.3403	-0.1351	
		(0.4526)	(0.5021)	
Origin: No migration		-0.0956	-0.1796	
		(0.2762)	(0.3148)	
Trust in government (std.)			-0.1250+	
			(0.0709)	
Climate significance (std.)			0.0021	
			(0.0897)	
Climate voting			0.1220	
			(0.0827)	
Moral universalism (std.)			0.0350	
			(0.0676)	
Others do more (std.)			0.0491	
			(0.0839)	
Others do less (std.)			0.1038	
			(0.0851)	
Zero-sum view (std.)			0.0070	
			(0.0677)	
Fixed-effects				
Respondent				Yes
Fit statistics				
Observations	678	678	678	678
R^2	0.06319	0.14337	0.18257	0.80323
Within R^2				0.00084

Table C.5: Predictive power: Funds allocated to international organizations

This table examines how individual fairness weights predict funds allocated to international climate organizations. Columns 1–3 present specifications with basic demographics, extensive demographics, and extensive demographics plus worldviews as controls. Column 4 includes participant fixed effects. Clustered standard errors at the individual level in parentheses. Significance levels: + p<0.1, * p<0.05, ** p<0.01, *** p<0.001. Standard errors in parentheses.

Model:	Basic demographics (1)	Extensive demographics (2)	Extensive demographics + worldviews (3)	FE (4)
Variables				
Constant	1.960**	0.0477	0.3362	
	(0.6455)	(0.6471)	(0.6067)	
Fair_UK (std.)	0.0697	0.0425	-0.0062	-0.0212
	(0.0576)	(0.0402)	(0.0259)	(0.0194)
Age	-0.0771*	-0.0710**	-0.0414*	
	(0.0301)	(0.0254)	(0.0197)	
Age^2	0.0007*	0.0007*	0.0004+	
	(0.0003)	(0.0003)	(0.0002)	
Gender	-2.162*	-2.001*	-1.052	
	(0.9527)	(0.8820)	(0.7148)	
Age × Gender	0.0832+	0.0906*	0.0468	
	(0.0446)	(0.0400)	(0.0324)	
Age^2 × Gender	-0.0008	-0.0010*	-0.0005	
	(0.0005)	(0.0004)	(0.0003)	
Household income		0.0194	-0.0066	
		(0.0356)	(0.0278)	
Education: Secondary School		0.4719	0.0309	
		(0.3235)	(0.3005)	
Education: Post-Secondary		0.3876	-0.0255	
		(0.3345)	(0.3157)	
Education: Undergraduate		0.4170	-0.0845	
		(0.3402)	(0.3170)	
Education: Postgraduate		0.6065+	-0.0236	
		(0.3488)	(0.3167)	
Party: Conservative		0.7504***	0.3830**	
		(0.1573)	(0.1324)	
Party: Green Party		1.902***	0.8470***	
		(0.2156)	(0.2132)	
Party: Labour		1.441***	0.6332***	
		(0.1341)	(0.1431)	
Party: Liberal Democrat		1.331***	0.5233**	
		(0.1848)	(0.1636)	
Party: Other		0.9990**	0.3008	
		(0.3026)	(0.2722)	
Region: Northern Ireland		-0.1875	-0.2481	
		(0.1600)	(0.1700)	
Region: Scotland		0.1314	0.2501	
		(0.2113)	(0.1662)	
Region: Wales		0.1630	-0.0164	
		(0.2100)	(0.2268)	
Area: Rural		-0.1008	0.0026	
		(0.2118)	(0.1658)	
Area: Suburban		0.1131	0.0983	
		(0.1355)	(0.1080)	
Area: Town		0.0263	0.0194	
		(0.1496)	(0.1095)	
Area: Village		0.3008+	0.0808	
		(0.1663)	(0.1329)	
Origin: High & low income		-0.6600**	-0.1568	
		(0.2507)	(0.2513)	
Origin: Low income		0.0933	-0.0714	
		(0.1969)	(0.1873)	
Origin: Low income, no migration		-0.5802*	-0.4001+	
		(0.2353)	(0.2254)	
Origin: Middle income		0.4297	0.1304	
		(0.2921)	(0.2318)	
Origin: Middle & low income		2.008***	0.8007*	
		(0.4336)	(0.3441)	
Origin: Middle income, no migration		-0.2440	0.0310	
		(0.3077)	(0.3155)	
Origin: No migration		-0.1046	-0.0922	
		(0.1599)	(0.1618)	
Trust in government (std.)			0.0307	
			(0.0451)	
Climate significance (std.)			0.1693**	
			(0.0579)	
Climate voting			0.1018	
			(0.0700)	
Moral universalism (std.)			0.0693+	
			(0.0373)	
Others do more (std.)			0.1211*	
			(0.0495)	
Others do less (std.)			0.2473***	
			(0.0606)	
Zero-sum view (std.)			0.0494	
			(0.0409)	
Fixed-effects				
Respondent				Yes
Fit statistics				
Observations	588	588	588	588
R^2	0.04705	0.39685	0.56522	0.79909
Within R^2				0.00063

Clustered (Respondent) standard-errors in parentheses
*Signif. Codes: ***: 0.001, **: 0.01, *: 0.05, +: 0.1*

Table C.6: Predictive power: Follow-up climate decisions

This table examines how individual fairness weights predict support for climate policies in the obfuscated follow-up decisions. Columns 1–3 present specifications with basic demographics, extensive demographics, and extensive demographics plus worldviews as controls. Column 4 includes participant fixed effects. Clustered standard errors at the individual level in parentheses. Significance levels: + p¡0.1, * p¡0.05, ** p¡0.01, *** p¡0.001. Standard errors are presented in parentheses.

C.4 The causal effect of fairness views on beliefs and policy preferences

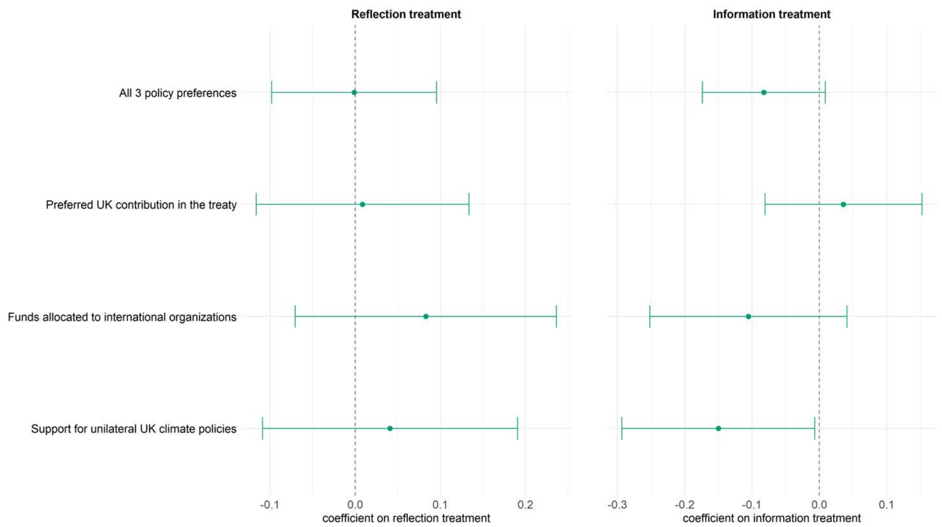


Figure B3: Treatment effect on policy preferences for mitigation costs

Note: The graph displays coefficients from regressions where dependent variables are the participants' policy preferences concerning climate change mitigation (all three elicitations for the first row, one single elicitation for the subsequent 3). The left panel displays the difference between the Max and Min Reflection treatment in regressions with data only from these two treatments. The right panel displays the difference between the Upwards and Downwards Reflection treatments in regressions with data only from these two treatments. The regressions don't include controls. The standard errors are adjusted to account for the stratified randomisation following Bai et al. (2025). They are clustered at the participant level in the first row, and robust in the other rows. The bars indicate 90% confidence intervals.

Table C.7: Treatment effects on UK's fund allocations to international adaptation costs

	(1) Full	(2) Median Sample Split	(3) Green Party/ Labour	(4) Reform UK/ Conservative	(5) vs. Control
Panel A: Reflection treatments					
Max Reflection	0.066		0.062	0.100	0.247**
	(0.088)		(0.124)	(0.155)	(0.087)
N	443		194	163	441
Panel B: Information treatments					
Upward Information	0.072	-0.029	0.170	0.067	0.261**
	(0.086)	(0.127)	(0.135)	(0.137)	(0.085)
N	457	228	195	173	455

Note: Dependent variable is the UK's fund allocation to international adaptation costs (not standardized). The estimates are produced with the sreg package (Bai, Shaikh and Tabord-Meehan, 2025), which accounts for stratification by STUDY ID. We excluded strata with no variation in the relevant treatment status for this analysis. Column (1) uses the full sample of completed participants. Column (2) tests the treatment effect for the 50% of participants for which our model predicts the largest effects. Columns (3) and (4) restrict the sample to participants affiliated with the respective parties (Green Party/Labour and Reform UK/Conservative). + p<0.1, * p<0.05, ** p<0.01, *** p<0.001. Column (2) is empty for Panel A. Column (5) replaces the Min Reflection / Downward Information group with the control group.

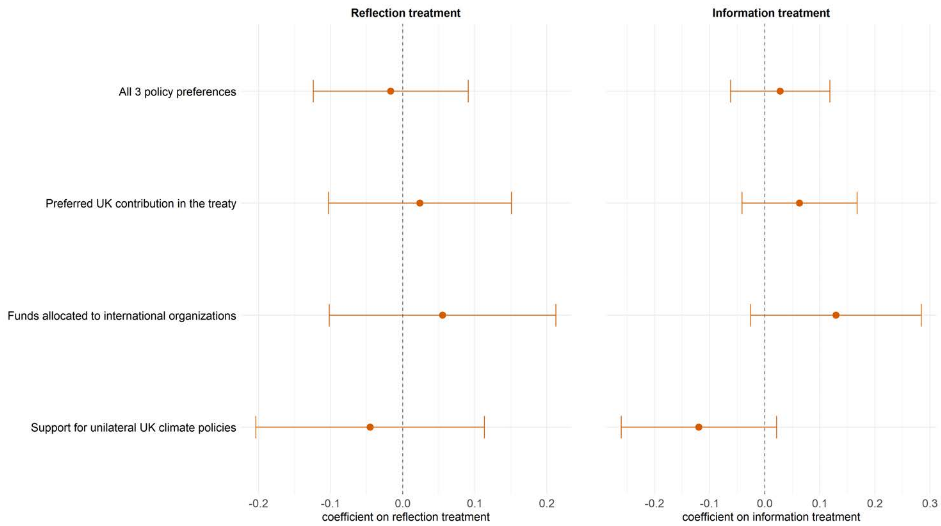


Figure B4: Treatment effect on policy preferences for loss and damage costs

Note: The graph displays coefficients from regressions where dependent variables are the participants' policy preferences concerning climate change loss and damage (all three elicitations for the first row, one single elicitation for the subsequent 3). The left panel displays the difference between the Max and Min Reflection treatment in regressions with data only from these two treatments. The right panel displays the difference between the Upwards and Downwards Reflection treatments in regressions with data only from these two treatments. The regressions don't include controls. The standard errors are adjusted to account for the stratified randomisation following Bai et al. (2025). They are clustered at the participant level in the first row, and robust in the other rows. The bars indicate 90% confidence intervals.

Table C.8: Treatment effects on the support for unilateral UK climate policies (adaptation)

	(1) Full	(2) Median Sample Split	(3) Green Party/ Labour	(4) Reform UK/ Conservative	(5) vs. Control
Panel A: Reflection treatments					
Max Reflection	-0.109		-0.137	-0.159	0.096
	(0.109)		(0.158)	(0.185)	(0.109)
N	402		176	148	397
Panel B: Information treatments					
Upward Information	-0.105	-0.152	-0.116	-0.032	0.119
	(0.107)	(0.156)	(0.157)	(0.183)	(0.108)
N	410	198	172	154	402

Note: Dependent variable is the support for unilateral UK climate policies regarding adaptation costs (not standardized) collected in the follow-up survey. The estimates are produced with the sreg package (Bai, Shaikh and Tabord-Meehan, 2025), which accounts for stratification by STUDY ID. We excluded strata with no variation in the relevant treatment status for this analysis. Column (1) uses the full sample of completed participants. Column (2) tests the treatment effect for the 50% of participants for which our model predicts the largest effects. Columns (3) and (4) restrict the sample to participants affiliated with the respective parties (Green Party/Labour and Reform UK/Conservative). * p<0.1, ** p<0.05, *** p<0.01. Column (2) is empty for Panel A. Column (5) replaces the Min Reflection / Downward Information group with the control group.

C.4.1 What explains the null effects?

C.5 Robustness Checks

C.5.1 Excluding the participants that fail the behavioral tests that flag potential AI usage

This appendix displays tables and figures that replicate the main results of the paper with a sample that excludes the 26 participants that failed one of the behavioral tests to flag AI usage. The results in this restricted sample are unchanged.

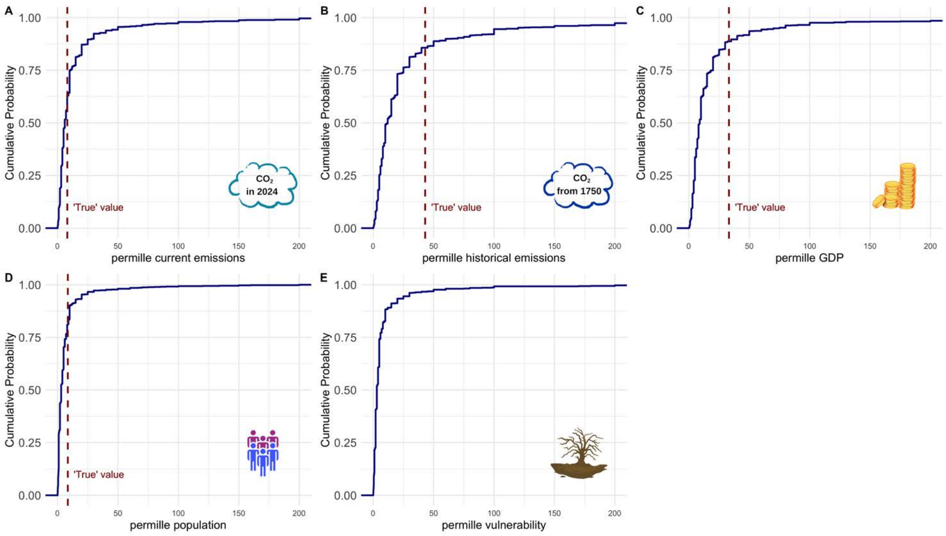


Figure B5: Beliefs about the UK's characteristics - Robustness checks

Note: The graph shows the participants' beliefs about the UK's characteristics. The dashed, red line represents the respective, true value. The underlying sample excludes participants who failed at least one of two AI detection tests.

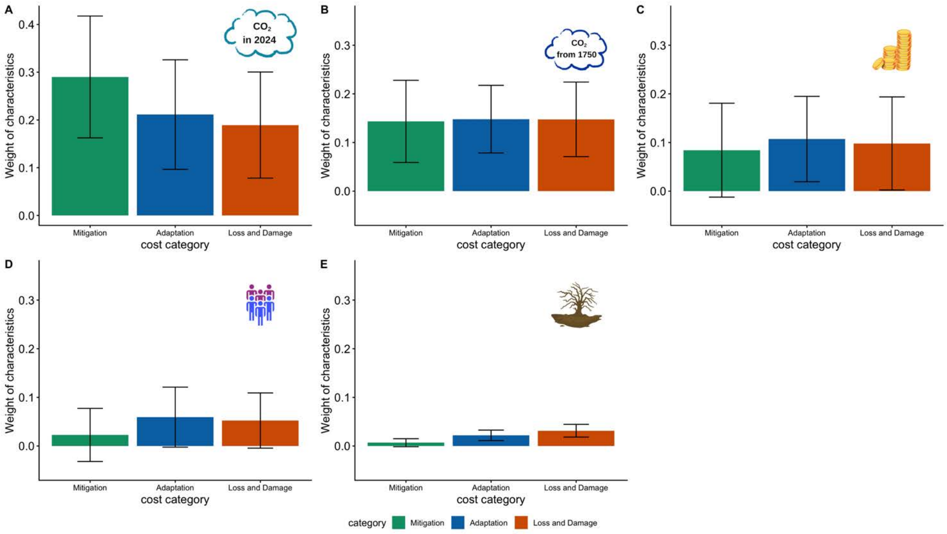


Figure B6: Fairness weights - Robustness checks

Note: The bars show the estimated coefficients from the linear specification of Model 6 for each fairness criteria. The bar height represents the average weight participants assign to a country characteristic when allocating climate costs. The error bars represent 95% confidence intervals. The standard errors are clustered at the participant level. The underlying sample excludes participants who failed at least one of two AI detection tests.

	fairshare (1)
Mitigation costs	
constant	2.507***
	(0.2133)
current emissions	0.2900***
	(0.0651)
GDP	0.0842+
	(0.0493)
historical emissions	0.1434***
	(0.0431)
population	0.0227
	(0.0278)
vulnerability	0.0068
	(0.0042)
Adaptation costs	
constant	2.270***
	(0.2218)
current emissions	0.2113***
	(0.0585)
GDP	0.1071*
	(0.0448)
historical emissions	0.1480***
	(0.0355)
population	0.0593+
	(0.0315)
vulnerability	0.0219***
	(0.0055)
Loss and Damage costs	
constant	2.313***
	(0.2231)
current emissions	0.1893***
	(0.0567)
GDP	0.0980*
	(0.0488)
historical emissions	0.1476***
	(0.0391)
population	0.0524+
	(0.0290)
vulnerability	0.0314***
	(0.0067)
F-statistic: incl. cost-specific const.	3.29
	(p<0.001)
F-statistic: excl. cost-specific const.	3.41
	(p<0.001)
Degrees of freedom: incl. const.	(12, 44586)
Degrees of freedom: excl. const.	(10, 44586)
R^2	0.34530
Observations	44,604

Table C.9: Fairness Weights - Robustness Check

Note: The table reports estimated fairness weights, i.e. the coefficients from a linear regression of stated fair shares on the five country characteristics (current emissions, historical emissions, GDP, population, and vulnerability), separately for mitigation, adaptation, and loss and damage costs. The specification includes participant fixed effects; F-statistics test the joint significance of the cost-specific constants when included and excluded from the model. The sample excludes any participants that failed at least one AI test.

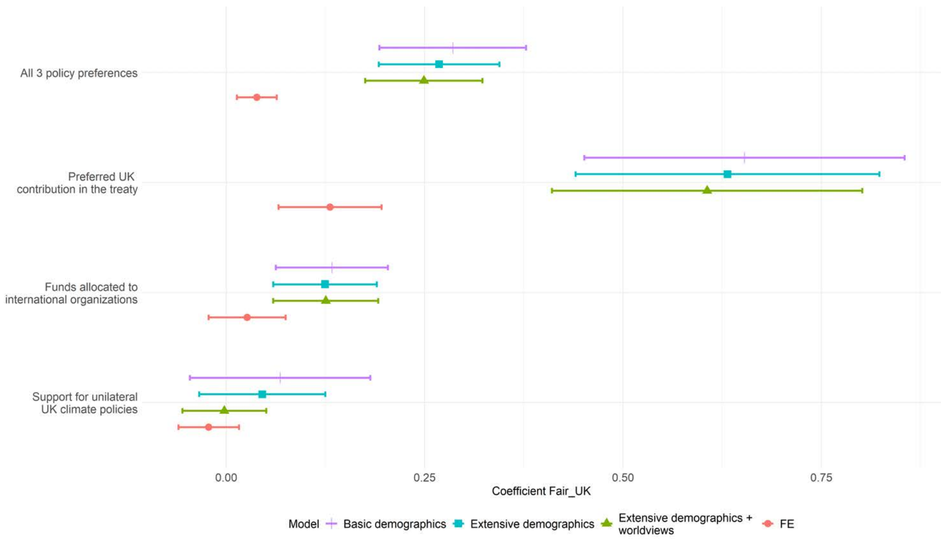


Figure B7: Predictive power - Robustness checks

Note: The underlying sample excludes participants who failed at least one of two AI detection tests.

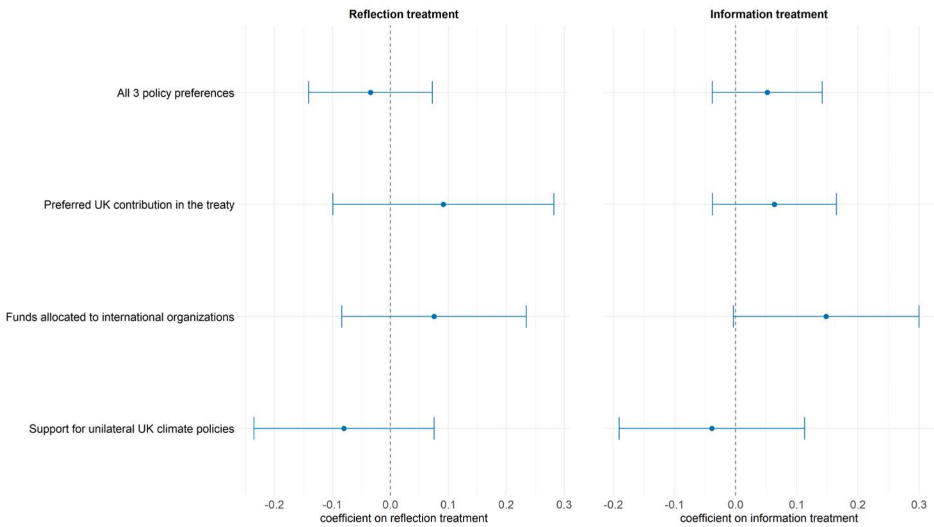


Figure B8: Treatment effect (Adaptation costs) - Robustness checks

Note: The underlying sample excludes participants who failed at least one of two AI detection tests.

Treatment	(1) All 3 policy preferences	(2) Preferred UK contribution in treaty	(3) Funds allocated to int. organizations	(4) Support for unilateral UK climate policies
Max Reflection	-0.034	0.092	0.075	-0.080
	(0.597)	(0.429)	(0.436)	(0.397)
Upwards Information	0.052	0.064	0.148	-0.039
	(0.344)	(0.302)	(0.109)	(0.672)

Table C.10: Treatment effects on policy preferences (sreg) — robustness check

D Instructions for the Main Survey

D.1 Introduction

Welcome to this study!

You can use the arrows on the side of the slide to navigate the instructions.

- This study has 5 parts and takes approximately 30 minutes.
- Part 1 is the longest. The other parts are considerably quicker.
- We will review submissions within a week.

This study was approved by the Ethics Commission, Department of Economics, University of Munich. Project number Project 2025-14.

To obtain ethical approval, we committed not to provide misleading or false instructions.

All your decisions are anonymous.

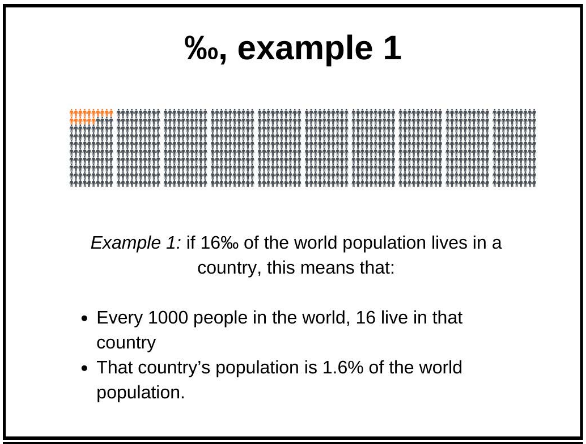
During the study, you can't:

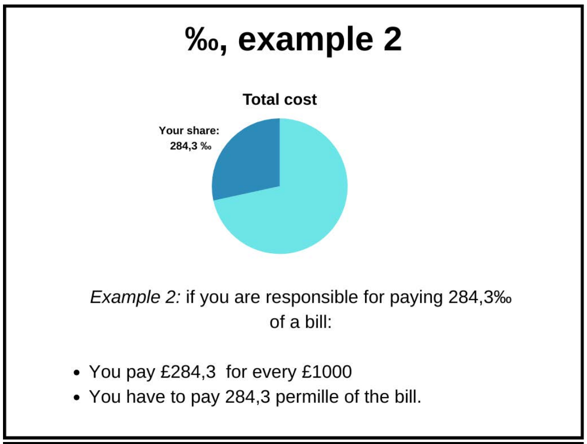
- **use AI or search engines,**
- **open other programs or tabs.**

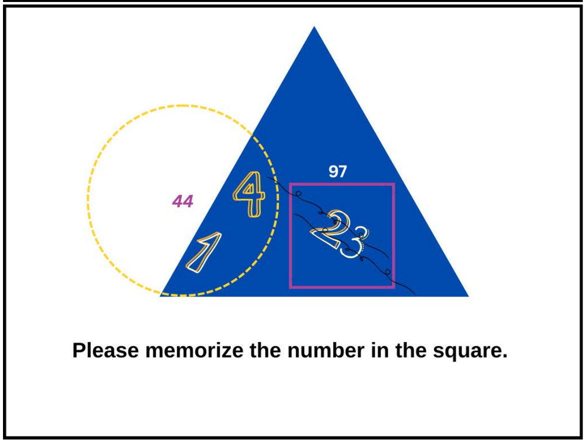
We have detection tools that flag participants who don't follow these instructions. If you deviate from them, we will reject your submission.

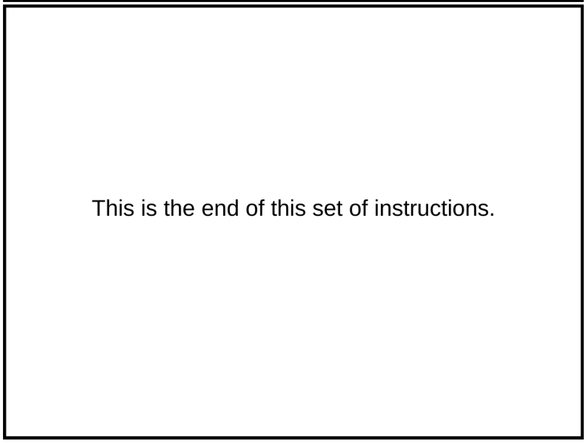
‰

During the experiment, we will often use the symbol above. It is called "*permille*" and denotes parts per thousand.









Comprehension questions

1. Please enter the number you were asked to remember: ☐

2. According to the ethical protocol under which we run this study, all the instructions you read must be truthful and not misleading.
   True; False

3. Your decisions are anonymous.
   True; False

4. You can use AI tools and search engines to complete the study.
   True; False

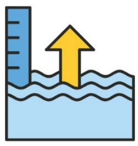
D.2 Part 1

Part 1

Instruction deck 1/3

In this study, we will ask you questions about climate change.

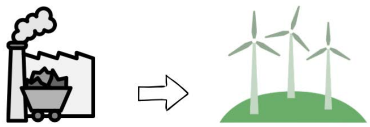
As you probably know, climate change is the modification of Earth's climate caused by the emission of CO_2 and other greenhouse gases.



Climate change has many effects:

- It increases average temperatures, driving global warming.
- It makes extreme weather events more likely.
- It increases sea levels.

Overall, climate change creates 3 types of costs.



1) Mitigation cost

Countries need to invest and change policies to reduce their greenhouse gas emissions.

Example: closing coal power plants and building wind farms.

Overall, climate change creates 3 types of costs.

2) Adaptation cost

Countries need to invest and change policies to adjust to climate change.

Example: constructing seawalls to protect cities and coastal areas from sea-level rise.

Overall, climate change creates 3 types of costs.

3) Loss and damage cost

This is the cost of the damage from climate change that cannot be avoided anymore.

Example: the loss of life and property due to heatwaves and floods caused by climate change.

In **summary**, the costs generated by climate change are:

1) Mitigation cost
2) Adaptation cost
3) Loss and damage cost

Every year, the countries of the world meet to discuss how to fight climate change.

For this study, we ask you to imagine that the discussions have led to a global agreement on climate change.

This agreement is very specific and lists:

- All the required investments and policy changes to mitigate climate change and limit global warming to 2°C,
- All the required investments and policy changes to adapt to a 2°C global warming.

Example: the agreement lists the number and location of wind farms and dams to be built in each country.

The agreement also establishes:

- Which permille of the *net* global mitigation cost should be paid by each country.
- Which permille of the *net* global adaptation cost should be paid by each country.
- Which permille of the global loss and damage cost generated by a 2°C global warming should be paid by each country.

Note: the net cost of an investment is the cost after accounting for all revenues.

Example: the net cost of a wind farm is the cost of building it minus the revenues from selling the electricity it generates.

The investments and policy changes in the agreement are those that minimize the total net cost of a 2°C global warming.

The *total cost* is equal to the sum of mitigation, adaptation and loss and damage costs.

For example, if the agreement establishes that a country pays 5‰ of the net global mitigation costs:

- this country will pay 5‰ of the net cost of all the investments and policy changes needed to limit global warming;
- it will pay this 5‰ regardless of the countries in which these investments and policy changes take place.

The agreement is binding.

Strong penalties for breaking the agreement ensure that all countries pay what they promised.

In summary, the agreement sets out:

- Which investments and policy changes are needed in the world to mitigate and adapt to climate change.
- How much each country should pay for the costs of mitigation, adaptation, and loss and damage.

Before we continue with the instructions, we would like to check that everything so far is clear.

Please answer the questions on the next page.

Comprehension questions

1. Match each term to its definition:
   Terms: Mitigation cost, Adaptation cost, Loss and damage cost;
   Defintions: The cost of reducing greenhouse gas emissions; the total economic cost of climate change impacts; the cost of investments and policy changes needed to adjust to climate change; the cost of the damage from climate change that cannot be avoided anymore

2. Which of the following is an example of cost due to loss and damage?
   The cost of building a wind farm; The cost of building a sea wall to protect against sea level rise; The cost of flood damage caused by climate change

3. Which of the following is true of the agreement we ask you to imagine (check all that apply)?

   - *It sets out the investments and policy changes needed to limit global warming to 2°C.*

- *It sets out the investments and policy changes needed to adapt to a 2°C global warming.*
- *It establishes that each country should pay for the loss and damage caused by climate change within its borders.*
- *It establishes which permille of each climate change cost is paid by each country.*

Part 1

Instruction deck 2/3

We would now like to understand what the agreement would need to look like for you to consider it **fair**.

In particular, we will ask which permille of:

- the net global mitigation costs,
- the net global adaptation costs,
- the global loss and damage costs,

would be **fair** for a country to pay depending on its characteristics.

We would like you to consider **5 characteristics** that some people think are important to determine how much a country should pay.

Country characteristics

‰ of current emissions

Some argue that countries that are polluting more today should pay a larger share of climate change costs.

Country characteristics

‰ of world GDP

Some argue that richer countries should pay a larger share of the climate change costs.

Country characteristics

‰ of historical emissions

Some argue that countries that have emitted more over their history (since the Industrial Revolution in 1750) should pay a larger share of climate change costs.

Country characteristics

‰ of the world population

Some argue that countries with larger populations should pay a larger share of the climate change costs.

Country characteristics

‰ of vulnerability

Some argue that countries facing the greatest losses if climate change is not mitigated should pay a larger share of the climate change costs.

Others say these countries should pay a smaller share.

In summary, we would like you to consider 5 country characteristics:

‰ of historical emissions

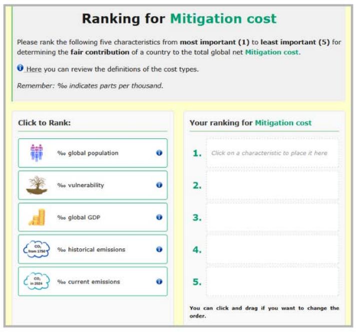
‰ of world GDP

‰ of current emissions

‰ of vulnerability

‰ of the world population

Your first task is to rank these five characteristics from the most to the least important for determining a country's **<u>fair</u>** contribution to climate change costs.



This is how the interface will look.

You will have to order the country characteristics 3 times.

Once for the mitigation cost, once for the adaptation cost, once for the loss and damage cost.

After you submit a ranking, we will ask you to indicate whether you believe that any of the 5 characteristics should ***not*** be used to determine each country fair share.

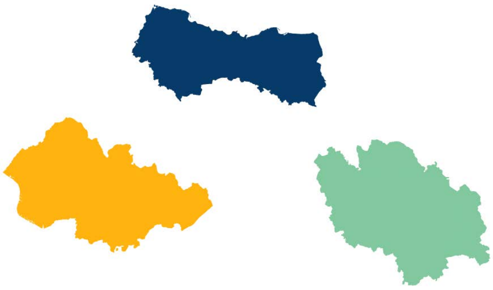
This is the end of this set of instructions.

Note: the order of the slides presenting the definitions of the five country character-

istics is randomized for each participant.

Part 1

Instruction deck 3/3



Now we will ask you to consider some specific countries.

The countries you will see are hypothetical, but we designed them so that we can learn from your answers what would be the fair contributions of real ones.

For every country, you need to indicate the permille of:

- the net global mitigation cost,
- the net global adaptation cost,
- the global loss and damage cost,

that would be **fair** for this country to pay.

5.1‰

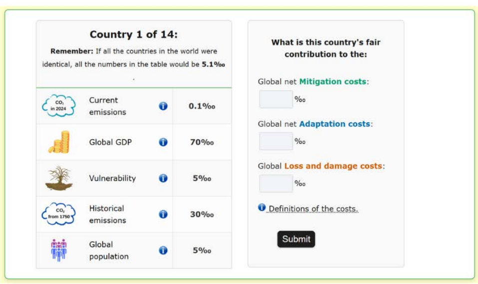
So that you know, there are 195 countries in the world. If all countries were identical, each country would have a value of 5.1‰ for each characteristic.

That is, every country would:

- produce 5.1‰ of the world's GDP,
- have 5.1‰ of the global population,
- have emitted 5.1‰ of the CO_2 human produced in 2024,
- have emitted 5.1‰ of the CO_2 human produced since 1750,
- suffer 5.1‰ of damage from unmitigated climate change.

Of course, countries in the world are very different from each other.

That's why we would like to know from you what would be a fair distribution of the climate change costs between countries.



This is how the interface will look.

This is the end of this set of instructions.

Transition slides after elicitations

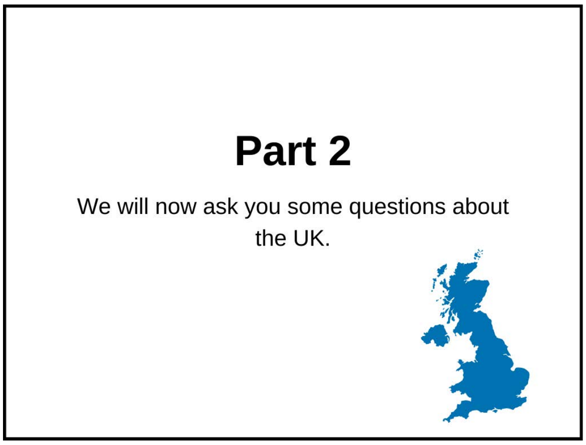
End of Part 1

Thank you for completing this part of the study.

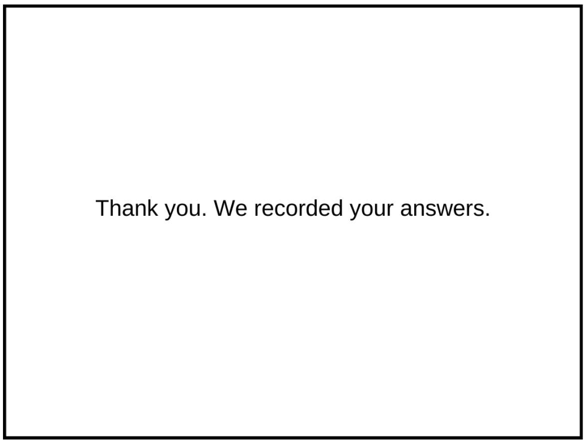
The next parts are shorter but equally important for our research.

Please stay focused!

D.3 Part 2



Transition slides after elicitations



D.4 Part 3

Part 3

Please think again about the climate agreement you were asked to consider in Part 1.

Remember, the agreement establishes:

- Which investments and policy changes are needed to mitigate and adapt to climate change.
- How much each country should pay for the costs of mitigation, adaptation, and loss and damage.

Imagine that you are in charge of writing the section that determines how much each country in the world should pay for each climate change cost.

You can choose who pays how much, but:

- The entire global net mitigation cost must be paid for,
- The entire global net adaptation cost must be paid for,
- The entire global loss and damage cost must be paid for.

In this section, we will ask how much you would make the UK pay.

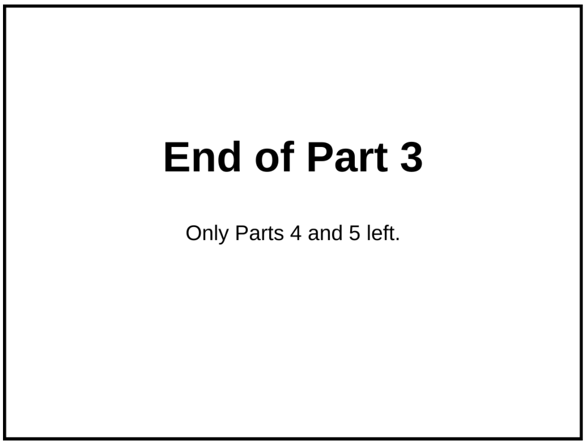
This is the end of this set of instructions.

Comprehension questions - Part 3

1. Which of the following is true in this part? (Select all that apply).

   - *You should imagine that you can determine how much each country should pay towards each climate change cost.*
   - *You should imagine that you can determine the total amount of the global climate change costs.*
   - *When deciding how much each country should pay, you may also choose that none of them pays anything.*
   - *You need to indicate how much you would make the UK pay towards climate-change costs.*

Transition slides after elicitations



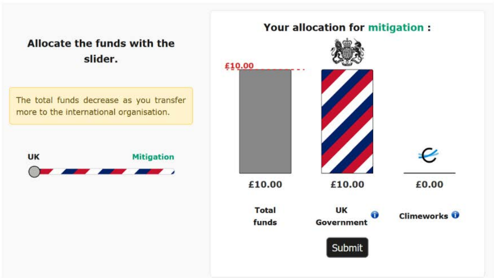
D.5 Part 4

Part 4

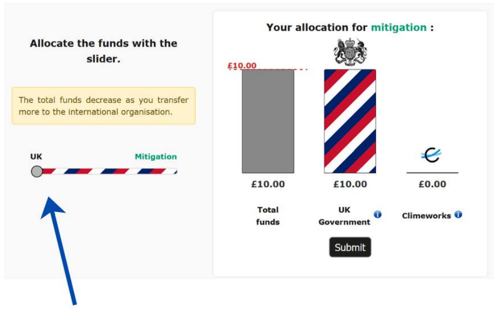
In this Part, you will make 3 decisions.

In each decision, you have to distribute some funds between:

- The UK government,
- An international organisation working on climate change.

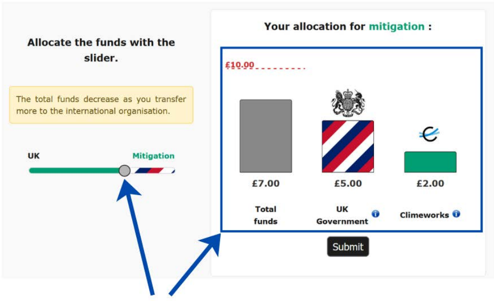


This is how the interface will look.

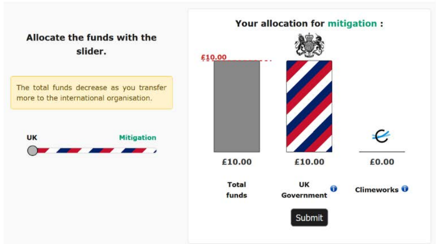


You can move funds from the government to the international organisation with the slider.

This column indicates the total amount of funds you are distributing.

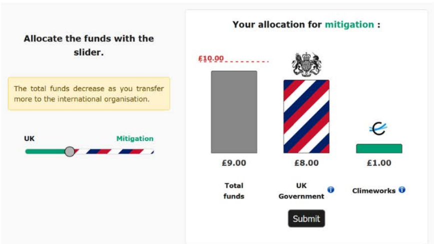


As you give money to the international organisation, the government receives less and less, and the total funds go down.



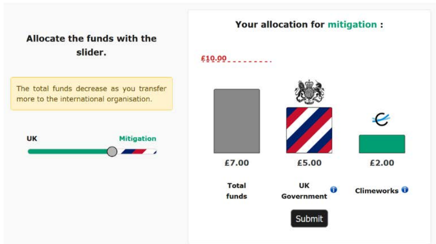
Example 1

- To the international organisation, you give £0.
- The government gets £10.
- The total funds are £10.



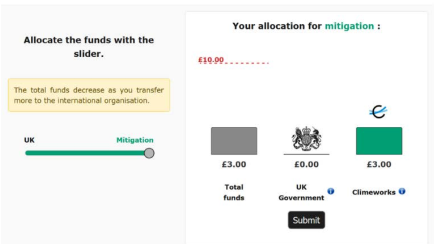
Example 2

- To the international organisation, you give £1.
- The government gets £8.
- The total funds are £9.



Example 3

- To the international organisation, you give £2.
- The government gets £5.
- The total funds are £7.



Example 4

- To the international organisation, you give £3.
- The government gets £0.
- The total funds are £3.

Your decisions in this part can have real consequences!

We will randomly select 5% of participants and one of their choices, and distribute the funds according to what they decided.

We will tell you at the end of the study whether you are one of the selected participants.

Remember:

To receive ethical approval, we committed ***not*** to use misleading or untruthful instructions.

We will transfer the money as we promised.

The money you assign to the UK government will go into the UK general budget.

climeworks

ADAPTATION FUND

GlobalGiving

We will tell you more about each international organisation before you distribute the funds.

This is the end of this set of instructions.

Fund allocation introductions Each of the following two-slide introductions is displayed immediately before the corresponding money-allocation decision in Part 4, once per fund (mitigation, adaptation, and loss and damage).

The first international organisation is **Climeworks**.

Climeworks finances climate change **mitigation** projects.

The funds you allocate to Climeworks will be used to capture and store CO_2.

The process uses modern technologies designed to store CO2 for at least 10,000 years.

The technologies include direct air capture, bioenergy with carbon capture & storage, and enhanced rock weathering.

The second international organisation is **Adaptation Fund**.

Adaptation Fund finances climate change **adaptation** projects.

The money you assign to the Adaptation Fund will go to projects and programs that help vulnerable communities in developing countries adapt to climate change.

Adaptation Fund projects target, for example, coastal zone management, food security, and disaster risk reduction.

The last international organisation is **GlobalGiving**.

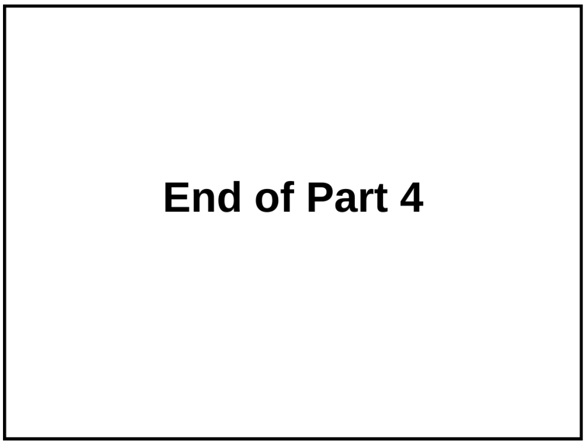
GlobalGiving finances projects to redress the **loss and damage** of climate change.

The funds you allocate to GlobalGiving will support relief programs for victims of severe flooding.

The programs may provide relief worldwide.

The programs deliver food, water, and medical supplies, and help people find shelter after a flood.

Transition slides after elicitations



D.6 Part 5