Dissuasion and polarization vs. persuasion in nano-targeting: experimental evidence from climate policy messaging

Miriam Sorace[1], Thomas Robinson[2], Simon Hix[3], and Joris Frese[3]

[1]University of Reading, UK. [2]London School of Economics and Political Science, UK.
[3]European University Institute, Italy.

Abstract

Micro-targeting is entering a new era: interacting a large set of voter characteristics *with* multiple message sub-components – "nano-targeting" – will become increasingly cheap and accessible. We combine evidence from two survey experiments (N=5,000) and predictive modeling to test the effectiveness of nano-targeted messages on climate policy. We find that, on average, nano-targeting does not perform better than an overall "best message", but that nano-targeted messages were more persuasive than the best message for individuals who were already supportive of climate action. We also discover backlash effects from "mis-targeting", and find an important asymmetry: being randomly assigned messages predicted to be far from one's best policy has stronger dissuasive effects than the persuasive effects of being shown one's best message. Together these results suggest that these technologies may not be practical for most actors, but may be effective for entrepreneurs who aim to dissuade, demobilize or polarize.

Words: 12,148

Introduction

Political micro-targeting – the targeting of political communication on the basis of a wide array of personal characteristics – is allegedly a formidable persuasive weapon. The message-receiver congruence and the cognitive consonance hypotheses from cognitive psychology (Festinger 1957, Petty and Cacioppo 1986) support the assumption that messages should be more persuasive if they closely match one's prior beliefs. Experimental empirical evidence, however, finds that micro-targeting often does not yield significant persuasive effects, and that an 'overall best message' – one that is supported by a majority – often performs as well as, if not better, than micro-targeted messages (Tappin et al. 2023, Hackenburg and Margetts 2024, Argyle et al. 2025).

But what about nano-targeting? We now have the technology to not only assign message units by micro-targeting on the basis of several receiver characteristics: we can also disassemble messages into sub-components and assign unique combinations of those via micro-targeting. This blend of micro-targeting and message sub-units tailoring is what we call 'nano-targeting', and it is now increasingly scalable. This, together with its persuasive potential, underscores the importance of a systematic assessment of nano-targeting as a political communication technology. Uniquely, our study does not simply aim to test the persuasive effects of nano-targeting, but also tests its polarizing and dissuasive effects. The current scholarship on the potential of micro-targeting – and, by extension, nano-targeting – to reinforce pre-existing views and poison democratic discourse is, in fact, chiefly normative or legal in nature (Zuiderveen Borgesius et al. 2018, Lavigne 2021, Kohl and Eisler 2021).

First, we illustrate a novel nano-targeting framework that combines machine learning predictive modeling with message calibration via representative surveys and conjoint analysis. The conjoint design allows to nano-target: i.e. connect message *components* to recipients' characteristics, rather than treating the entire message as a unit. Using training data from conjoint experiments, we leverage eXtreme Gradient Boosting (xgboost) predictive modeling to build nano-targeted messages assigned to the treatment group in a message testing experiment[1]. Secondly, we test: (1) whether respondents seeing cam-

[1] An alternative would be to let an LLM come up with the message dimensions to vary and then rely on its ability to vary these dimensions according to subjects' demographics. We decided against this strategy

paign messages specifically nano-targeted to match their preferences support the message more than respondents seeing the macro-targeted 'overall best message' (the message supported by a majority) or an untargeted message assembled completely at random and separately for each respondent (the control condition); (2) whether seeing a nano-targeted message is more effective for moderates/politically undecideds or for people that already supported the cause; and (3) whether receiving a randomly assigned message that, by chance, was mis-targeted elicits stronger backlash or dissuasion compared to a randomly assigned message that, by chance, aligns well with the individual's predicted preferences. We first fielded a conjoint survey experiment on a nationally representative sample of 2,000 UK adults to train a predictive model, and then fielded a separate three-arm survey experiment on over 3,000 UK adults which tested nano-targeting against (a) a control group of randomly generated messages and (b) an "overall best message" condition.

We exploit the climate protection issue in our nano-targeting application to illustrate and test the technique in the strongest, most realistic, and most ethical way possible. This issue offers the strongest possible test of nano-targeting as it allows to craft highly multi-dimensional messages where there is strong political and demographic variation on what climate message dimensions are supported most. There are indeed multiple possible pathways to reduce emissions and meet the 1.5 degree target (Allen et al. 2018, Bergquist, Mildenberger and Stokes 2020): one could prioritize different sectors among the largest net emitters (energy and industry, construction, farming, transport, trade and finance) and/or pick different policy instruments (ranging from investment type and re-allocation, technological development, phase-outs and bans, regulating standards, and taxation). We know that these different sub-dimensions matter differently for different individuals

for three key reasons. First, it is very unlikely that political campaigns would relinquish control over their messaging nor rely on models whose training process is opaque and not necessarily representative of the time period or electorate in question. Second, as we wanted to *nano*-target, we needed to vary several message sub-dimensions, and have the capacity to determine which specific message feature mattered most for specific individuals. LLM applications tend to vary one message component at a time (typically cognitive or moral foundation frames) (Simchon, Edwards and Lewandowsky 2024, Matz et al. 2024), and there is no research to date on what aspects of the message LLMs assign to different groups. Given the accumulating evidence of LLM biases (Zhou et al. 2024, Yeh et al. 2023), we decided against using LLMs to craft our nano-targeted messages. Third, we wanted to be able to transparently quantify our prediction error. LLMs do not straightforwardly return predicted probabilities and frontier models are typically proprietary and closed-source. We do not think we are missing out on superior message crafting quality either: an important finding from the LLM literature is that LLM-generated messages are not consistently more persuasive than human generated ones (Hackenburg et al. 2025), and that their persuasiveness is further hampered when individuals realize the message is machine-generated (Palmer and Spirling 2023, Schoenegger et al. 2025).

(Bergquist et al. 2022). Secondly, the climate topic offers a realistic set up, since it has a clear uni-directional persuasive end-goal (reaching net zero)[2]. Since political campaigns typically have a fixed, uni-directional end-goal they want to persuade the public to back or entrench in policy (e.g. redistribution in the case of socialist parties, human rights protection in the case of NGOs like Amnesty International, or nativism in the case of radical-right parties) we wanted to exemplify the technique through an issue with an uncontested end-goal: i.e. carbon emissions reduction. Scientists overwhelmingly agree that the best policy to prevent climate change from spiraling into disastrous consequences is a "radical reduction in fossil-fuel- and land-use-related carbon emissions" (Fankhauser et al. 2022 : 16). Finally, the climate topic made it possible to apply nano-targeting ethically: whilst there is high contestation on the best policies to achieve carbon emission reduction, there is typically high public concern for the uni-directional goal of protecting climate by reducing emissions (Fairbrother 2022).[3] We did not want to exert persuasion on study participants where the uni-directional end-goal itself is highly contested and not a clear public good.

Our study makes several contributions to research on political communication and political behavior, as well as to political campaigns practice. First, we introduce and demonstrate the nano-targeting technique. We develop and apply a novel two-step approach that can be applied in future research and by political campaigns: in the *calibration* phase, we train the machine learning model on real, nationally representative, survey data, so it can accurately learn how each of the available individual characteristics (e.g. political and sociodemographic factors) are related to *each specific message content feature*. Then, in the *testing* phase, we administer individually tailored climate policies to a separate sample of respondents. Unlike most existing research in political science on micro-targeting – which either focuses on normative/legal issues or applies standard A/B tests and/or heterogeneous effect tests – we systematically test the persuasive and

[2]Some might argue that our choice to field different policy instruments as levels in the conjoint might make the technique less realistic for, say, party campaigns, if not dangerous for democracy as it would be encouraging parties to promise substantially different policies for different people. Our specific application is aimed to simply exemplify the technique in the strongest way possible, but nothing stops a party campaign from applying it by designing the conjoint with mutually non-exclusive levels and attributes, such as, for example different policy areas as attributes and using the multiplicity of manifesto policy pledges on each policy area as conjoint levels.

[3]Also see: YouGov - Most Europeans Are Worried About Climate Change and YouGov - Eurotrack.

polarizing potential of this technology *against alternative messaging approaches*. Our nano-targeted messages are benchmarked against randomly generated messages and a typical "overall best message" treatment condition. To the best of our knowledge, only Tappin et al. (2023) carry out such benchmarked tests *and* assigned messages based on algorithms trained on a calibration step – but they chiefly test micro-targeting, rather than nano-targeting, and they do not empirically test polarization and dissuasion effects.

Crucially, our study's second contribution addresses this gap: we empirically test key boundary conditions of nano-targeting, such as the role of the persuadable (switchers and/or moderates) vs. the already converted (the 'base'), and the potential for backlash effects from mis-targeting. We find that nano-targeting does not work better on the persuadable (moderates or switchers) but, instead, we discover it works best for individuals who were already pro-climate, thus exercising a polarizing effect. We also find important backlash effects from mis-targeting, and, surprisingly, discover an important asymmetry: being assigned personalized most-disliked messages has stronger dissuasive effects than receiving personalized most-liked messages has persuasive effects. This study is, hence, the first to empirically demonstrate these perverse incentives and inherent democratic dangers of this technology: the limited distinctive persuasive advantages of nano-targeting when compared to an "overall best message strategy" means these technologies are not likely to be useful for mainstream parties or interest groups, but, in contrast, may be effective for political entrepreneurs who aim to dissuade, demobilize and/or polarize.

Political targeting and nano-targeting: benefits vs. risks

Political targeting can have different "depths". At its most shallow, it identifies key segments – such as core supporters vs. undecideds/switchers vs. "never backers", or augmented with some key demographic data, i.e. 'young switcher' vs. 'rural core voter' – which then guide the strategic deployment of campaign resources and messaging. At a deeper, micro-targeted level, a campaign would run predictive models leveraging the interaction of several political and socio-demographic characteristics at once, effectively building 'segment matrices' (Endres 2016). The next step – which is now increasingly scalable, also thanks to advances in AI technologies (Argyle et al. 2025, Hackenburg and Margetts 2024) – is "nano-targeting". Through nano-targeting, it is message *sub-dimensions* – rather than the message as a whole – that get micro-targeted, i.e. calibrated on a model that automatically discovers the relationship between multiple socio-demographic, behavioral and political features of the individual, and multiple features of the message itself.

The premise underpinning any targeting effort is that a personally relevant message is more likely to be appealing to the receiver than a less personally relevant one: the deeper the targeting, the higher the appeal of the message to the individual is likely to be. The research on cognitive dissonance, for example, shows how new items of information are more likely to be processed if they are consistent with prior cognitions (Festinger 1957, Cooper 2007). Messages that are relevant to a receiver's pre-existing cognitive schema have a higher chance of being accepted. The message-receiver congruence hypothesis (Petty and Cacioppo 1986) similarly suggests that a message will be more persuasive if it plays to the cognitive style of the receiver: individuals high in "need for cognition" should prefer factual, logical, well-argued content, while individuals with an emotional cognitive processing style are more persuaded by emotional narratives. Tailoring a message so that it matches the receiver appears to be an important step for effective persuasion, therefore.

Targeting campaigns and messages is thus expected to yield optimization gains. Targeting helps to avoid preaching to the already converted, and to efficiently find individuals who need to be persuaded or who know less and are undecided on an issue. It also helps to package the message appropriately, by discovering which frames and contents are going to play well with the type of voter an actor is trying to persuade. A voter receiving a targeted message is more likely to gain something from interaction with political campaigns, more likely to become engaged and interested in political/policy issues (Matthes et al. 2022), and more likely to turn out to vote (Gerber, Green and Larimer 2008). However, the empirical evidence that micro-targeting yields *persuasive* advantages is thin (Druckman 2024).

Does a deeper version of micro-targeting – i.e. nano-targeting – offer important persuasive advantages over more traditional targeting techniques adopted by political parties, and over crafting messages that persuade "in parallel" (Coppock 2023), for example, across the board? Chiefly, scholars studying the effects of campaign messages have thus far focused on either deep description of parties' targeting practices (Dommett and Kruschinski 2024, Votta et al. 2024, Endres 2016), or applied standard A/B and heterogeneous effects testing rather than micro-targeting, let alone nano-targeting. Hersh and Schaffner (2013) examine campaign messages that explicitly pander to specific sub-groups of the population. Their evidence comes from a series of vignette survey experiments where respondents either see a "pandering" message from a fictional candidate running for Congress or a generic appeal message. They conclude that targeted messages are as persuasive as the generic/ambiguous ones, but that if the respondent is an in-group member that reports high levels of linked fate with the in-group the targeted messages work best. They also show that seeing pandering messages if one is not a member of the in-group can backfire, thus alerting of the potential serious risks of mis-targeting. They deploy a rough measure of targeting – the largely symbolic mention of a sub-group – and do not directly test micro-targeting in the sense of leveraging highly personalized messages based on combinations of multiple demographics. Another important study, by Endres (2020), leverages US Republican National Committee (RNC) 2012 data and shows that ads that were congruent with voters' predicted policy preferences were associated with increased support for Romney and lower support for Obama and/or lower abstention, particularly in the sample of registered Democrats. This evidence hence suggests that micro-targeting works in persuading the other side – which has more room to move – and does not yield extra advantages when applied to voters who already lean Republican. It does not show, however, what happens when individuals are canvassed with generic messages, rather than ads targeted to their issue preferences. Hence, this approach does not test whether micro-targeting yields additional persuasion as opposed to persuasive messages that work overall.

To the best of our knowledge, only Tappin et al. (2023) test micro-targeting appropriately via assignment to messages based on algorithms trained on a calibration step. They also appropriately test the micro-targeted experimental condition against the single best message condition, a mis-targeted condition, and a full control condition. They find that their realistic micro-targeted video ads have only inconsistent persuasive advantages over the other messaging conditions. Hackenburg and Margetts (2024), Argyle et al. (2025), Simchon, Edwards and Lewandowsky (2024) focus on nano-targeted messages generated via AI prompting (rather than via a survey calibration step), and also find that their tailored messaging condition has similar persuasive potential as the generic 'best overall message' condition. They, however, only nano-target on very few message sub-components – chiefly personality-congruent frames.

The available theoretical and empirical discussion of micro-targeting expects this technology to yield persuasive advantages, but current evidence is lacking. On the basis of existing persuasion theory we would expect that nano-targeting will be even more successful in persuading individuals, as several message sub-components will be pinpointed to several individual characteristics at once. From this, we derive the following hypotheses for our climate messaging test case:

H1: (Attitudes) *Nano-targeted pro-net zero messaging will result in a bigger shift towards pro-net zero attitudes and behaviors in the post-test, compared to the best overall message, and compared to the random message.*

H2: (Petition support) *Nano-targeted pro-net zero messaging will result in larger support for the climate protection petition compared to the best overall message, and compared to the random message.*

The risks from the misuse of targeting technologies are, allegedly, manifold. According to the level of individual characteristics used, micro- and nano-targeting can become too invasive and lead to privacy and data protection breaches (Zuiderveen Borgesius et al. 2018, Kohl and Eisler 2021). Psychographic profiling is particularly regarded as ethically questionable and was at the heart of the Cambridge Analytica scandal, where micro-targeting models were calibrated and run on illegally acquired Facebook profile data ahead of the 2016 US Presidential election. In light of such questionable uses of micro-targeting, voters are critical of highly tailored messages, which can lead to backlash effects against perceived manipulation attempts (Gahn 2025). Disinformation and misinformation are other potential risks, particularly if political parties leverage micro- and nano-targeting maliciously and/or pledge different things to different people, selectively hiding their real policy aims from the public (Hillygus and Shields 2008). The final alleged risk is the potential to fragment and polarize the electorate, by increasing issue salience differentially and by reinforcing pre-existing priorities and views (Levy and Razin 2020, Dobber et al. 2023).

The literature on the risks of micro-targeting is largely normative, with limited empirical evidence supporting the claims of harmful effects (Kreiss 2017). Such claims are likely overstated as research from political psychology has shown that campaign efforts rarely lead to radical shifts in public opinion (Kalla and Broockman 2018). Furthermore, political campaigns also have incentives to reach across political divides, and thus to engage with switchers and undecided voters and to attract supporters from electoral competitors, rather than just to reinforce echo chambers. This notwithstanding, we explore the polarizing potential of nano-targeting by checking heterogeneity by strength of political priors. If nano-targeting is mostly persuasive for the politically undecideds (moderates and/or party switchers) then its impact does not flow through further strengthening ideological/partisan priors. Furthermore, we check backlash/dissuasion potential by leveraging mis-targeting arising from the control group: the randomly generated climate messages will in fact result in different degrees of mis-targeting for different respondents.

We hence also test the following hypotheses:

H3: *Nano-targeted pro-net zero messaging will result in a bigger shift towards pro-net zero attitudes and behaviors in the post-test, and larger support for the climate protection petition, compared to the best overall message, and compared to the random message,* ***particularly among switchers, moderates and undecided respondents****.*

H4a: *Mis-targeted pro-net zero messaging will result in a shift against pro-net zero attitudes and behaviors in the post-test, and lower support for the climate protection petition, compared to the best overall message, and compared to the nano-targeted message*

H4b: *The higher the degree of mis-targeting of pro-net zero messaging in the control group, the bigger the shift against pro-net zero attitudes and behaviors in the post-test, and the lower the support for the climate protection petition, compared to the best overall message, and compared to the nano-targeted message.*

Data & Research Design

To test the effectiveness of nano-targeted climate messages, we execute a two-phase experimental protocol with two separate survey companies and two separate respondent pools in the UK. This section provides a broad overview of the research design, and the subsequent sections provide in-depth descriptions of our exact model parameters and estimation strategies in both phases.

The initial step in any nano-targeting application involves mapping the various subdimensions of a message and identifying the various pledges or narrative frames the campaigns wants to use. In our application, we only fixed the overarching 'pro-net zero' persuasion goal of the message, while – to exploit the full political and demographic variation in climate preferences and thus build the strongest possible version of nano-targeting – we purposely tested the full range of potential climate policy instruments,[4] as well as some typical climate narrative frames (Fesenfeld et al. 2020, Huber, Wicki and Bernauer 2020, Bergquist et al. 2022). In our study, we thus primarily vary the type of instrument (market-based vs. command-and-control, push vs. pull) and the targeted sector (energy, transport, building, natural resources and trade). The latter is a feature of climate policy that has not been sufficiently analyzed to date (Fesenfeld et al. 2020), and is another original aspect of our study and a contribution to the climate policy literature. In addition to policy pledges on climate, we test two benefit narratives surrounding climate policy – job creation and energy security – as well as two emotional frames – an optimism and a catastrophism frame.[5] The full list of possible attribute levels is documented in Table 1.

[4] If a political party or candidate wanted to apply this technique to their campaign, they could easily design non-mutually exclusive pledges and narrative frames, to avoid straying from their manifesto and promising fundamentally different things to different electors.

[5] We adopt a deliberately broad conceptualization of persuasion instruments, encompassing not only narrative frames but also formal policy pledges. Political actors often engage in selective emphasis of manifesto pledges to convince electors to vote for them, for example (Budge and Farlie 2025, Huber and Haselmayer 2025). And, as we detail in our findings below, narrative frames are the weakest instruments of persuasion: individuals are more likely to be convinced by policy commitments than simple frames.

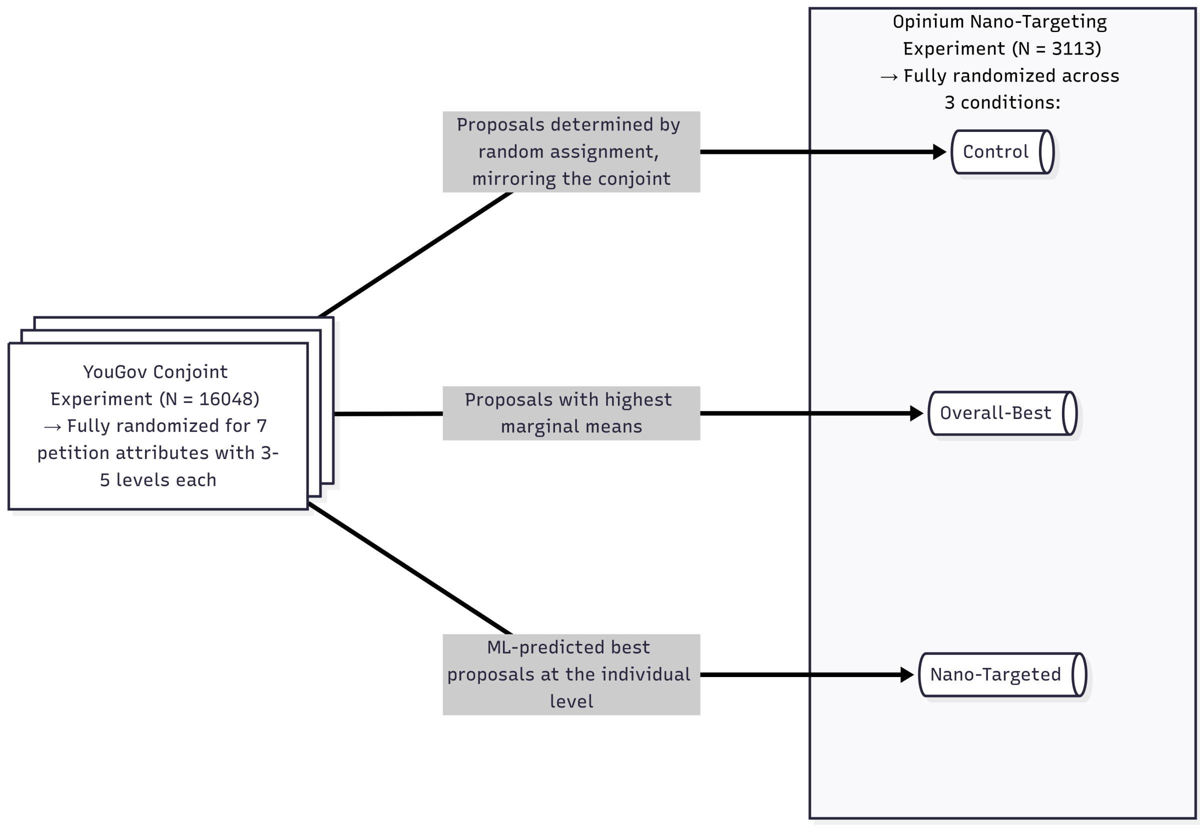
Table 1: Petition attributes and levels

Attribute	**Levels**
Emotional preamble	1. *Negative frame*: "If we do not implement policies to meet the 1.5C global warming target and reach net zero, we will be on the brink of extinction." 2. *Positive frame*: "We are on course to meet the 1.5C global warming target and reach net zero: turning the page on climate change is within our reach." 3. No frame.
Benefit preamble	1. *Security frame*: "Our proposed decarbonization solutions will make us more secure: they will reduce the UK's energy dependence on other countries, which will protect us from global shocks in fuel prices and energy supply. We call on the UK government to meet the 1.5C target by ..." 2. *Economy frame*: "Our proposed decarbonization solutions are good for jobs and the UK's economic growth: they will create tens of thousands of new, sustainable jobs every year. We call on the UK government to meet the 1.5C target by ..." 3. "We call on the UK government to meet the 1.5C target by ..."
Lever 1: Energy & industrial policy	1. Phasing-out the use of all non-renewable energy sources. 2. Increasing taxes across industries that use non-renewable energy sources. 3. Investing £5 billion from the budget (about 0.5 % of government spending) every year to develop green technologies for industry. 4. Reducing taxes across industries that use renewable or safe nuclear energy sources. 5. No statement.
Lever 2: Buildings and Homes	1. Phasing-out all buildings and homes not reaching energy efficiency level A by removing building permissions. 2. Increasing taxes on owners of buildings and homes not reaching energy efficiency level A. 3. Investing £5 billion from the budget (about 0.5 % of government spending) every year to develop energy efficient technologies for homes and buildings. 4. Reducing taxes for owners of buildings and homes that refurbish their properties to meet energy efficiency level A. 5. No statement.
Lever 3: Transport	1. Phasing-out all non-electric vehicles and flights that can be completed by train in an hour. 2. Increasing taxes on all non-electric vehicles and flights that can be completed by train in an hour. 3. Investing £5 billion from the budget (about 0.5 % of government spending) every year to develop eco-friendly fuels and low emission vehicles. 4. Reducing taxes to households and businesses that use electric vehicles. 5. No statement.
Lever 4: Natural Resources	1. Phasing-out soil-degrading materials and practices (e.g. polluting fertilizers, intensive farming, single-crop farming; old/inefficient irrigation systems; deforestation). 2. Increasing taxes and penalties on individuals and businesses that make use of soil-degrading materials and practices. 3. Investing £5 billion from the budget (about 0.5 % government spending) every year to develop soil-preservation materials and practices (e.g. mixed crop-livestock innovations; biotechnologies to increase crop yields; water-saving technologies for reservoirs, dams and irrigation systems; eco-friendly fertilizers, carbon capture and storage processes, re-forestation/re-planting). 4. Reducing taxes on individuals and businesses that adopt soil-preservation materials and practices. 5. No statement.
Lever 5: Trade and Finance	1. Phasing-out trade with, and financial investments in, companies/countries that do not meet strict environmental standards. 2. Increasing taxes/duties on goods and services from companies/countries that do not meet strict environmental standards. 3. Investing £5 billion from the budget (about 0.5 % of government spending) every year to support trading partners in meeting strict environmental standards. 4. Reducing taxes/duties on all businesses and individuals who trade with companies/countries that meet strict environmental standards. 5. No statement.

Note: All attribute levels were fully and independently randomized in the conjoint experiment.

The basic premise of our research design is visualized in Figure 1. In phase one we identify an empirical benchmark for the popularity of various climate policy proposals among various sociodemographic strata by conducting a conjoint experiment with one respondent pool. Then, in phase two, we use a machine-learning algorithm trained on the phase one data to predict onto a different respondent pool which combination of policy proposals is the most likely to appeal to each individual respondent. Finally, with this second respondent pool, we compare – based on random assignment – those nano-targeted petitions to a second treatment arm showcasing the macro-targeted overall-best petition,[6] and to an untargeted control group for which the climate petitions are assembled through the same individual-level randomization protocol as in phase one.[7]

Figure 1: Two-Phase Experimental Nano-targeting Protocol



For **phase one**, the conjoint experiment, we worked with YouGov to survey a nationally representative sample of 2,006 UK respondents.[8] Each respondent evaluated a

[6]In other words, a petition consisting of those proposals with the highest marginal means across the full sample in phase one.

[7]Meaning that different subjects again view different petitions, like in the nano-targeted condition, but without any targeting at play.

[8]We also fielded a pilot study with Prolific in advance to determine the ideal design for the conjoint

total of eight climate petitions (for a total N of 16,048 petition evaluations), each of which contained two brief introductory texts and five climate policy proposals on energy, buildings, transport, resources, and trade. All segments of each petition were fully randomized. An example screenshot documenting how our participants viewed the vignettes is shown in Appendix E. Petition evaluations took two forms in this calibration survey: a binary choice to sign one of two competing petitions; and continuous ratings of individual petitions on a scale from zero to ten.

The conjoint analysis provided us with two sets of relevant outputs: (1) estimates of the most popular petition features in the form of average marginal component effects (AMCEs) and marginal means (MMs), which we used to design the 'macro-targeted' overall-best petition for our alternative treatment arm in phase two; and (2) measures of socio-demographic features of subjects in this sample such as the respondents' age, gender, education, religion, location, income, and vote preferences. We used this information to fit a machine-learning model that predicts petition support based on the combination of varying respondent- and petition-level attributes. In turn, for a given sociodemographic profile, we could predict out ratings over all possible petitions. The highest rated petition became our nano-targeted petitions in the main treatment arm of interest in phase two.

For **phase two**, the nano-targeting experiment, we worked with Opinium to survey a non-representative sample of 3,113 British respondents. Each of these respondents was randomly presented either a control petition (based on the same individual-level randomization protocol as in phase one, so that each respondent would see a bespoke control petition), the macro-targeted petition (based on the highest marginal means in phase one), or a nano-targeted petition (based on the ML-predicted best petition for their sociodemographic profile). Our measures of persuasion consisted of a battery of petition-evaluation questions, a battery of climate-attitudinal questions, and a behavioral measure of political climate advocacy.[9] Linear regression analyses of these outcome measures on our three-pronged treatment variable provided us with the ATEs of macro- and nano-

experiment. The results and conclusions from the pilot study are reported in Appendix D.

[9]Our exact outcome measures and excerpts of what the experiment looked like for respondents are included in Appendix I. Note that for the climate-attitudinal questions, all estimated average treatment effects (ATEs) represent differences in differences: since we surveyed these attitudes before and after the treatment, we subtract the pre-treatment values from the post-treatment values to construct our outcomes.

targeting to test our hypotheses as described above.

Our setup and analyses for both phases were pre-registered ahead of the respective data collection rounds. The anonymized pre-analysis plans are documented in Appendices A and B.

Phase One: Calibrating Nano-targeted Petitions

Conjoint results As a first empirical step, we designed a conjoint experiment where each vignette combined various policy proposals into a climate petition urging the UK government to take action against climate change. Given two brief introductory texts and five climate policy proposals in the areas of energy, buildings, transport, resources, and trade in each petition, we use this data to estimate AMCEs and MMs that serve as a benchmark for the macro-targeted overall-best treatment arm in phase two. We estimate AMCEs with a generalized linear model of the form:

$$Y_{ij} = \alpha \, + \, \sum_{g=1}^{G} \sum_{\ell=2}^{L_g} \beta_{g\ell} \, \mathfrak{D}_{ijg\ell} \, + \, \varepsilon_{ij}.$$

Where

- Y_{ij} is the evaluation of petition j by respondent i,
- $\mathfrak{D}_{ijg\ell} = \ell\}$ indicates that, in petition j, attribute g took level ℓ,
- $\beta_{g\ell}$ is the AMCE of moving attribute g from its baseline level to level ℓ,
- Standard errors are cluster-robust at the respondent-level i.

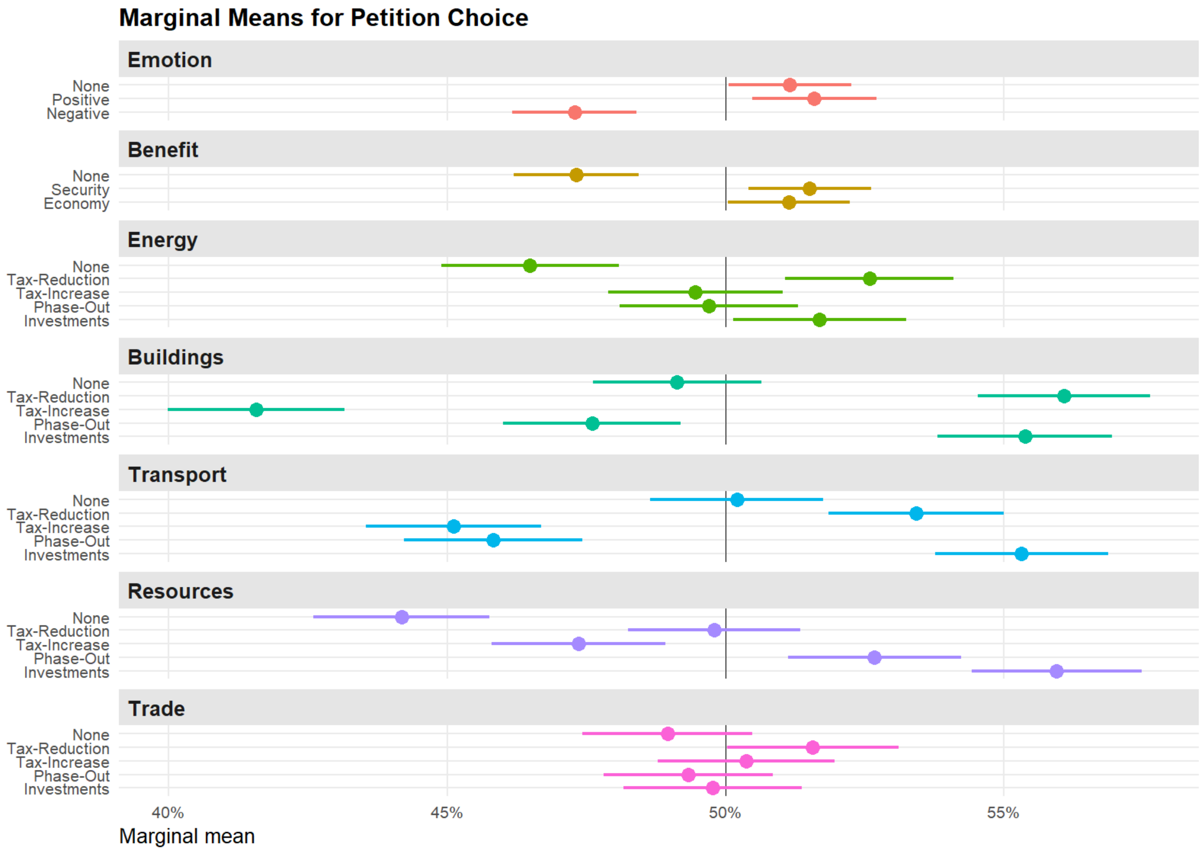
The respective MMs are defined as follows:

$$\mathrm{MM}_{g\ell} \; = \; \mathbb{E}\big[Y_{ij} \mid \mathfrak{D}_{ijg} = \ell\big].$$

Where

- $\mathrm{MM}_{g\ell}$ is the expected petition evaluation Y_{ij} given that for respondent i, the attribute g in petition j took the level ℓ.

Figure 2: Marginal Means for the Conjoint Experiment



Note: Errorbars are 95% confidence intervals based on cluster-robust standard errors (clustered at the respondent-level).

Our sample of 2,006 YouGov respondents evaluated a total of 16,048 climate petitions. The MMs for the full sample are reported in Figure 2 in the form of the percentage of petitions containing a given message that ended up getting signed. Based on the main conjoint analysis results, an attractive climate policy package: (a) offers strong tax incentives on sustainable buildings and for renewable energy; (b) proposes strong investment to make the transport and natural resources sectors more sustainable; (c) begins with a proactive, positive preamble, and a benefits preamble that highlights the energy-security benefits of the policy. We also find a marginal positive effect of tax incentives when it comes to trading with sustainable overseas companies. We derived the 'overall best' climate policy message by simply adding up all components of a climate policy package that have the highest MM value. The respective AMCEs as well as the standard repertoire of conjoint diagnostics are reported in Appendix F.

Prediction model training To nano-target individuals, we used the conjoint data from the first-phase experiment to fit a prediction model that estimates the level of support for petitions with varying policy components *given* a demographic description of each individual. Our approach builds on work modeling heterogeneous treatment effects in experimental studies (Grimmer, Messing and Westwood 2017), and conjoint experiments specifically (Zhirkov 2022, Robinson and Duch 2024), and aims to estimate a prediction function

$$g(A, X) = \mathbb{E}[Y|A, X],$$

which returns the expected rating Y of a petition A given the demographic and pre-treatment features of a subject (X). The advantage of random assignment in the experimental stage is that, by observing choices over many subjects, we do not need to observe every attribute-combination for every subject in order to estimate the marginal effects of these attribute-levels or to make predictions about subjects' behavior (Hainmueller, Hopkins and Yamamoto 2014).

Estimating $\hat{g}$ manually through marginal analyses of each covariate within X would be both inefficient and suboptimal: it would rely on pre-specifying any potential relationships not only between the outcome and covariates but also *between* covariates. This leaves the model prone to large researcher degrees of freedom, and could impinge on the performance of the model as a prediction engine. Instead, we use a non-parametric, machine-learning strategy which allows us to induct associations between subjects' ratings of petitions, the petition attributes, and descriptions of subjects themselves.

Our machine learning approach uses extreme gradient-boosted trees–**xgboost**–a prediction algorithm used for its high prediction accuracy and scalability.[10] In these models, separate weak tree models are fit sequentially to the residuals of the existing model state (Chen and Guestrin 2016).[11] Each "tree" recursively divides the data by splitting it based on an independent variable, to maximize the similarity of outcomes within each result-

[10]Nothing in our design precludes the use of alternative learning strategies. When pre-registering our design, we aimed to balance model complexity against the size of dataset we, and indeed other potential users like advocacy groups or parties, would have access to after the first stage. For example, higher variance strategies like deep learning would unlikely yield greater predictive accuracy on datasets this size and would take considerably longer to tune, potentially leading to delays in fielding the second stage.

[11]Here, "weak" refers to the fact each tree model is constrained to not make too many splits over the data.

ing split of the data. In a sense, therefore, these models perform a more complex form of marginal analysis by both searching for the *best* subsetting variable at each decision point, and performing this search repeatedly over subsets to further refine the predictions. The "boosted" nature of these trees refers to the fact that the data is re-weighted after each tree is fit in order to focus the model's attention on where it is currently making poor predictions.

We trained our **xgboost** model to predict the continuous petition ratings given by subjects, using the attribute-level descriptions of the petitions (as in the conventional explanatory analysis) combined with subject covariate information as the independent variables.[12] In early analyses, we compared the performance of a model using this continuous rating against a model that predicted subjects' binary choices. While difficult to compare on a single scale, we found the performance on the binary ratings was considerably worse (we conjecture because of the dependence between observations in this outcome relative to the forced-choice).[13]

To ensure our **xgboost** model was as predictive as possible, we conducted a 10-fold cross-validation exercise to optimize the models' hyperparameters (aspects of the learning algorithm that are set prior to seeing the data).[14] In total, we considered 10 hyperparameter dimensions.[15] Given the infeasibly large set of possible hyperparameter values these dimensions entail, we used a Tree-structured Parzen Estimator (TPE) sampler to model the contribution of hyperparameter values to the performance of the model, and sample from its posterior distribution to efficiently search the hyperparameter space (Bergstra et al. 2011). We used this sampling technique in conjunction with a Hyperband pruner, which interrupts a trial if the model performance *during* the fitting routine suggests it will not improve on the current best combination (Li et al. 2018). In combination, this routine

[12] The conjunction of splits in the tree models allows the model to identify interactions within and between petition and demographic attributes. For example, a tree might first divide by a subject's age and then, for younger respondents, divide between those that prefer negative or positive emotional framings.

[13] To measure out-of-sample model performance on the binary outcome, we calculated the area under the receiver operator curve (AUC). Our test score on the final, tuned model was 0.6, which is reasonably low–a model assigning random guesses would have an AUC of 0.5.

[14] Prior to cross-validation, we also split out a random 20% of the experimental data for validation and testing purposes. Half this data was used to validate the early stopping of the **xgboost** model, and half was used as a final test set to validate performance after final model training.

[15] These dimensions were: the type of booster model, λ, α, η and γ hyperparameters, and the maximum depth, growth policy, subsampling rate, feature bagging rate, and minimum child node weights for tree models.

allowed us to offset the inherent costs associated with searching over many models, and to run our search routine quickly on a personal computer.[16]

We searched across 1000 model-informed combinations of hyperparameter values to identify the best-performing model. Our final model has a root mean squared error on both cross-validated and test sets of around 1.8, suggesting a good ability to predict the ratings of candidates. This is also considerably better than a "vanilla" linear regression model of subject ratings (using the same predictors), suggesting that the additional complexity of the xgboost algorithm allowed us to learn more about the relationship between subject covariates and their ratings of climate change petitions.

One empirical question we had was how many and what subject-level variables were required to make good predictions. We therefore repeated this cross-validation exercise for three different combinations of subject covariate features: a model containing only two political variables (subjects' vote choice in 2024 and their present vote intention); a model containing only demographic features; and a model containing both sets of descriptors. As shown in Figure G8 in the Appendix, across both cross-validation and test datasets, the model containing all covariates performed best, but only marginally, than the demographic-only model. The political variable model performed considerably worse in cross-validation, although not much worse in terms of test data. This discrepancy suggests the cross-validation routine may be slightly optimistic about the added benefits of having access to more predictors. It is nevertheless the case that all tuned **xgboost** models performed markedly better than a linear regression model on the test data, affirming the decision to use a non-parametric modeling strategy.

Since for nano-targeting we care solely about prediction accuracy, our final prediction model uses all covariate features as predictors (plus the petition attribute-levels). Prior to fielding our second experiment, we received covariate descriptions of the subject pool and, for each subject, predicted their rating of every possible climate change petition ($\mathcal{A}$). We then pre-assign to each subject the attribute-level combination that the models predicts has the highest predicted rating *for them*, i.e.

[16] All models and the hyperparameter searching routine were conducted on an Apple M1 Max MacBook Pro laptop, with the hyperparameter search taking approximately 1 hour to complete for each of the three predictor combinations we tried.

$$d_i^{\text{Nano}} = A_i : \arg\max_{A \in \mathcal{A}} \ \hat{g}(A, X_i).$$

For each potential subject, we also generate a petition treatment containing the unconditional "overall best message":

$$d_i^{\text{Marginal}} = A : \arg\max_{A \in \mathcal{A}} \hat{f}(A),$$

where $\hat{f}$ reflects the fit generalized linear model from the conjoint data, without covariate controls. Hence, this assignment is equivalent to assigning the attribute-levels with the highest marginal means and does not vary across subjects.

Finally, we pre-assign "control" petitions to each potential subject where the attribute-levels are randomly assigned for each individual separately, as was done in the conjoint experiment itself:

$$d_i^{\text{Control}} A : A_i \sim \text{Uniform}(\mathcal{A}).$$

Hence, the control arm is not a singular petition (unlike the overall best message arm). Since we randomly assign each control petition's attribute-levels, we are able to randomise the "distance" between the individuals' ideal petition and the message they actually receive.

Phase Two: Testing Nano-targeted Petitions

Main results In this section, we first document the main results of the nano-targeting experiment with Opinium, including the ATEs of the macro-targeted and the nano-targeted petitions compared to the control petitions on our pre-registered outcome batteries. We then move on to discuss heterogeneity in nano-targeting effectiveness across the pre-registered sociodemographic strata.

Given fully randomized treatment assignment, our estimation of the ATEs on each outcome Y takes the basic form of a linear regression with a constant α, the petition assignment indicator $\mathcal{P}$, and an optional vector of covariates $\mathbb{X}$:

$$Y_i = \alpha + \beta_1 \mathcal{P}_i + \beta_2 \mathbb{X}_i + \varepsilon_i$$

For ease of interpretability, we group our outcomes into binary (Figure 3) and continuous outcomes (Figure 4).[17] All results are documented for the full sample of $\sim$ 3,100 Opinium respondents[18] for models with and without respondent-level control variables, though the inclusion of control variables does not alter the conclusions of any of our models.[19] The left panel of Figure 3 shows that both macro-targeted (overall-best) and nano-targeted messaging are effective at increasing the amount of signatures received, by 5-9 percentage points ($p < 0.05$ across all models with and without control variables, as indicated by the thin 95% confidence intervals). While the overall-best petitions led to an even greater increase in signatures compared to the nano-targeted ones, the difference between the two treatment arms is not statistically significant at the five-percent level, as indicated by the thick 84% confidence intervals.[20] The right panel shows that the two treatments did not cause significant increases in willingness to advocate for climate mitigation policies via e-mails to MPs, neither compared to the control condition nor to each other.

The left panel of Figure 4 documents the effects of targeted petitions on petition evaluations beyond the binary decision to sign or not. Macro-targeted (overall-best) and nano-targeted petitions were evaluated more favorably on average than the control petitions, but this difference is only statistically significant for the overall-best petitions ($p < 0.01$), not the nano-targeted ones. The difference between the two treatment arms themselves is also statistically significant in this case. Neither treatment had a substantive effect on climate-related attitudes more broadly, although the overall-best petition did lead

[17]For the analyses with continuous outcomes, we ran factor analyses on the individual outcomes reported in Appendix K and used the resulting factor scores as outcome variables, capturing the two item batteries for petition evaluations and climate-related attitudes.

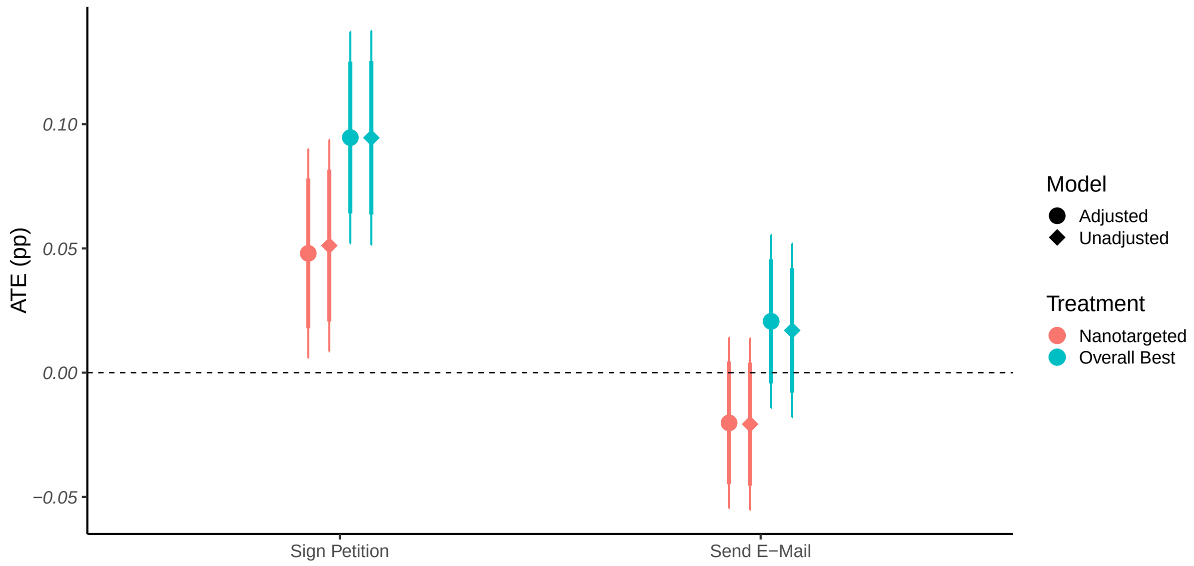
[18]Power analyses indicating that a 2,500 respondent sample would have been sufficient for the detection of small effects of our treatment arms on our outcomes of interest with 90% statistical power are documented in Appendix H. However, due to the uncertainty around unit non-response, Opinium fielded the survey on a larger sample, which led to the slightly increased sample size compared to our requirements and expectations.

[19]We show the full regression tables including coefficients for our control variables (respondent age, gender, locality, social grade, and education) in Appendix J and we document the ATEs for all individual outcomes in Appendix K.

[20]For an explanation of why non-overlapping 84% confidence intervals can be used to visually examine statistical differences between coefficients, see Armstrong II and Poirier (2025).

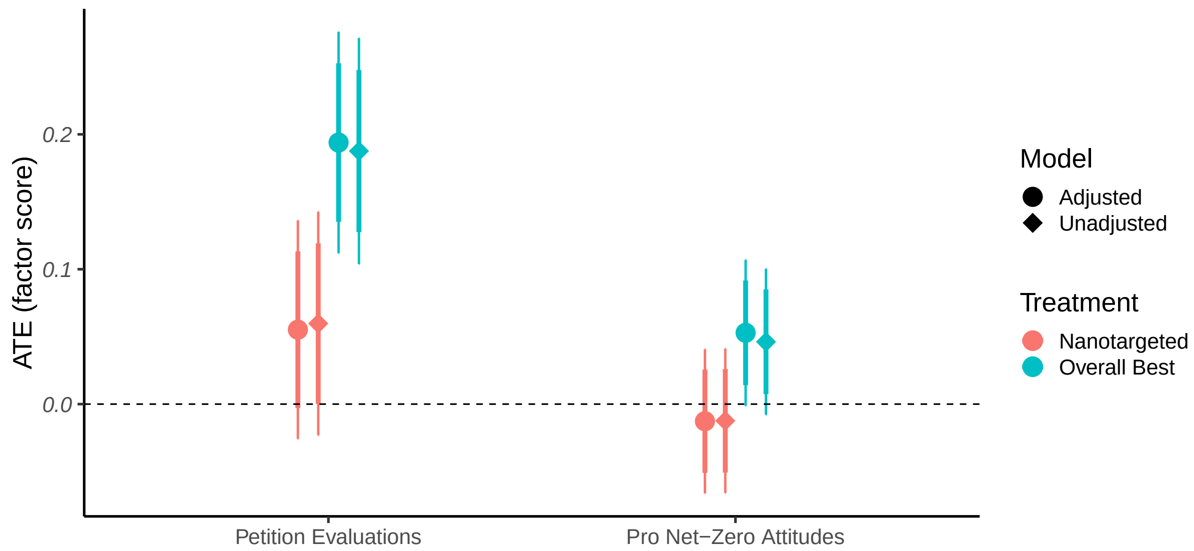
to marginally more climate-friendly responses ($p < 0.1$), as shown in the right panel.[21]

Figure 3: ATEs for the Binary Outcomes



Note: Errorbars are 95% and 84% confidence intervals based on HC1-robust standard errors. The control group serves as the baseline category.

Figure 4: ATEs for the Factor Outcomes



Note: Errorbars are 95% and 84% confidence intervals based on HC1-robust standard errors. The control group serves as the baseline category.

In sum, the ATEs across the full sample of Opinium respondents suggest that nano-targeting is not as effective for persuasion as expected. Across four outcome batteries, nano-targeted petitions outperformed the control petitions only once, for the most basic choice measure of willingness to sign the petition. Furthermore, nano-targeted petitions

[21] For this last outcome battery, we are estimating within-subject changes, because we measured the climate attitudes both before and after treatment exposure.

did not outperform the "overall best" petition for any outcome measure. In fact, when looking at the ATE coefficients, the 'overall best message' treatment arm outperformed nano-targeting across all outcomes – though this unexpected reverse difference is again only statistically significant for one outcome battery. One important finding is that the overall best petition received support from a majority of respondents, with baseline support for the control petition at roughly 43%, and the ATE for the overall best treatment at roughly 9%, moving overall support to 52%. This suggests that there is majority support for net-zero policies, given the right choice of policy instruments.

Why Does Nano-Targeting Under Perform Simpler Strategies?

The results from the first phase of our design were suggestive that nano-targeting should work: on average, the *predicted* ratings for the nano-targeted petitions were 0.82 points higher on a 0-10 scale compared to the predicted ratings for the overall best message (on the second stage sample, $p < 0.001$). Moreover, as shown in Appendix Figure G11, we find that nano-targeted petitions differ substantially (in terms of the chosen attribute levels) from the overall best message: on average, nano-targeted petitions shared about 2.49 of the same attribute levels as shown in the overall best message. The most consistent attribute was the emotional preamble, where approximately 45 percent of nano-targeted petitions matched the overall best message vignette, whereas the proposed trade policy only matched the overall best message level in 19 percent of nano-targeted petitions (see Appendix Figure G12). Within the nano-targeted treatment arm, there is also considerable variance to the distribution of attribute-levels across all facets of the petition. The Hamming distance of nano-targeted petitions was 0.72; in other words, on average, two randomly drawn petitions from this group differed on almost 72% of attributes (Appendix Figures G13 and G14 plot the distribution of levels across attributes for this group).

There are several reasons why the nano-targeted treatment arm might not yield higher support for climate action, on average, compared to the simpler overall best message strategy. There may, for instance, be little benefit to nano-targeting because individuals' preferences over climate change action are homogeneous, such that the overall best message is what many individuals want. In which case, the lower efficacy of nano-targeting is indicative of an "overfit" model that is exaggerating correlations between individuals' characteristics and their preferences over aspects of the climate change petition. Alternatively, it may be that the prediction model is not able to effectively predict individuals' ratings. That is, nano-targeting may work but only when one can accurately model the relationship between features of the message recipient, aspects of the petition, and the outcome.[22]

Although both explanations imply prediction error, we can distinguish them through subset analysis. If the diminished performance stems from hard to predict but heterogeneous preferences then for those we predict well nano-targeting should have an advantage over the overall best message. If, on the other hand, individuals' preferences are homogeneous then even for those we predict well there should be no advantage to nano-targeting the message.

We calculated the prediction error between what our nano-targetting model predicted each subject's petition rating to be and what their observed rating was in fact.[23] This prediction error allows us to isolate whose behaviors in our second sample were more consistent with what our nano-targeting model predicted. We then re-estimated the ATE, sub-setting the data to only include those for whom the absolute prediction error was within a given margin. We estimated these ATEs repeatedly, varying this threshold.

As we show in Appendix Figure M24, for subjects for whom the prediction error is low–where we are able to predict their rating within approximately ± 2 points on a 10-point scale–we find that nano-targeting works significantly *better* than the overall best message. Moreover, for those we can predict within approximately 1 rating point, nano-targeting was about 2.7 times more effective than the overall best message strategy (relative to a random petition). Collectively, this evidence is consistent with an explanation that nano-targeted messages can be more persuasive when the targeting itself is accurate, but

[22]Other potential impacts on the performance of nano-targeting include attitudes changing over time, such that what predicts well in phase one fails to predict well in phase two. We minimized this possibility as much as possible by fielding the phase two experiment within two weeks of our initial conjoint experiment. To the best of our knowledge, there were no major events in the intervening period that would likely cause major changes in people's attitudes or preferences.

[23]The ability to estimate this error is a distinct benefit of this approach relative to LLM-based generation. Not only are predictions exactly replicable, but these predictions are also numerical values rather than a sequence of text tokens (even if those text tokens include digits). With LLM methods, one may be able to proxy this score, but you would be reliant on assuming the resulting token output actually reflects the model's prediction (rather than some semantically and syntactically consistent, but adjacent, construct).

inconsistent with an explanation that preferences are homogeneous. Hence our finding that the overall best message outperforms nano-targeting is likely due to the difficulty in predicting subjects' preferences across the full sample, rather than because nano-targeting as a strategy is itself ineffective.

Given this finding, we also explored for *whom* the nano-targeting was less accurate. Given our underlying model is non-parametric, we took the measure of prediction error and modeled this as a function of the same subject demographics used to nano-target subjects. We then recover Shapley values–a game theoretic measure of a feature's contribution to the outcome (prediction error)–to understand whether, in our sample, the error was greater for certain types of voters. As shown in Appendix Figures M26-M36, we find that those who are older, male, in lower social classes, and supporters of Reform are positively associated with prediction errors. This is suggestive, albeit neither causal nor conclusive, evidence that our model struggles to predict accurately a section of the population that is typically considered harder to reach.[24]

Heterogeneity by Attitude Strength: Is Nano-targeting More Effective on "Persuadables"?

Beyond illustrating and testing a more advanced version of micro-targeting, namely nano-targeting, our study pre-registered the hypothesis that nano-targeting might be particularly effective only among certain types of respondents: the persuadable (switchers and/or moderates). If the evidence supports the hypothesis that nano-targeting is particularly effective among individuals in need of persuasion (in this case on the climate issue) or who are politically undecided, it would mean that the technology has important persuasive applications, and it would be unlikely to further polarize public opinion, as feared by Levy and Razin (2020) and Dobber et al. (2023).

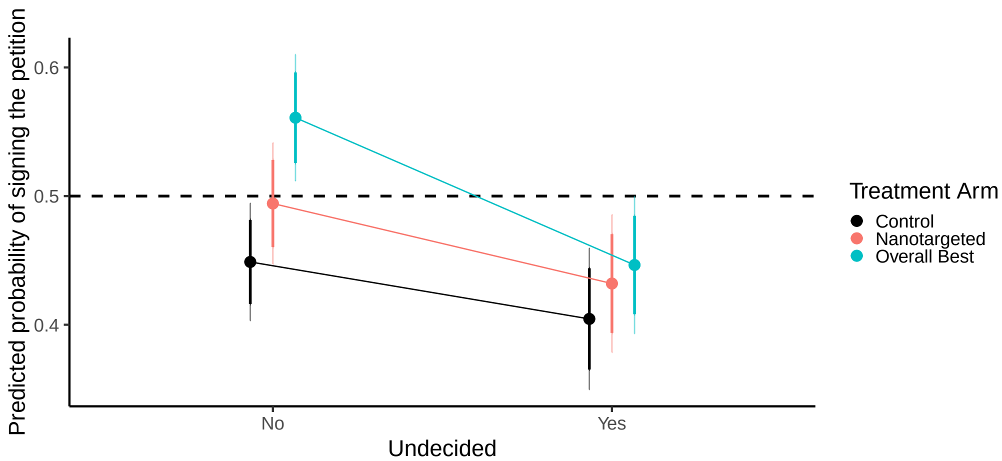
To test H3 we ran additional regression analyses where we interacted our treatment arms with dummy variables indicating (a) party switchers,[25] and (b) respondents who

[24] Importantly, our claim is not that these individuals are less receptive to nano-targeting, but rather that nano-targeting may work less well because these individuals' beliefs and attitudes are harder to predict.

[25] As operationalized by having a different vote intention in 2024 compared to the past vote choice in 2019.

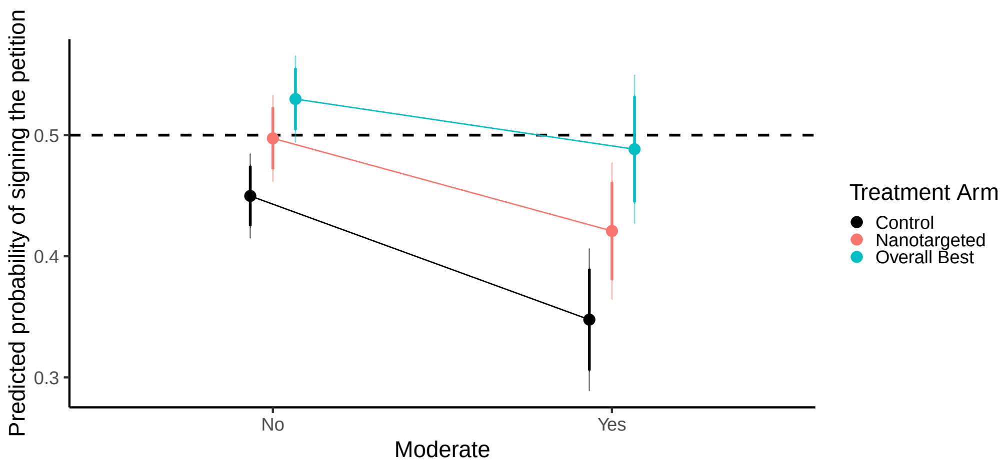
are moderates on pre-test climate attitude items.[26] Figures 5 and 6 depict the predicted probability of supporting the nano-targeted, control, and overall best climate message by party loyalty and climate attitude moderation.

Figure 5: Heterogeneity by Party Loyalty Status



Note: Errorbars are 95% and 84% confidence intervals based on HC1-robust standard errors.

Figure 6: Heterogeneity by Climate Attitude Strength: Climate Moderates



Note: Errorbars are 95% and 84% confidence intervals based on HC1-robust standard errors.

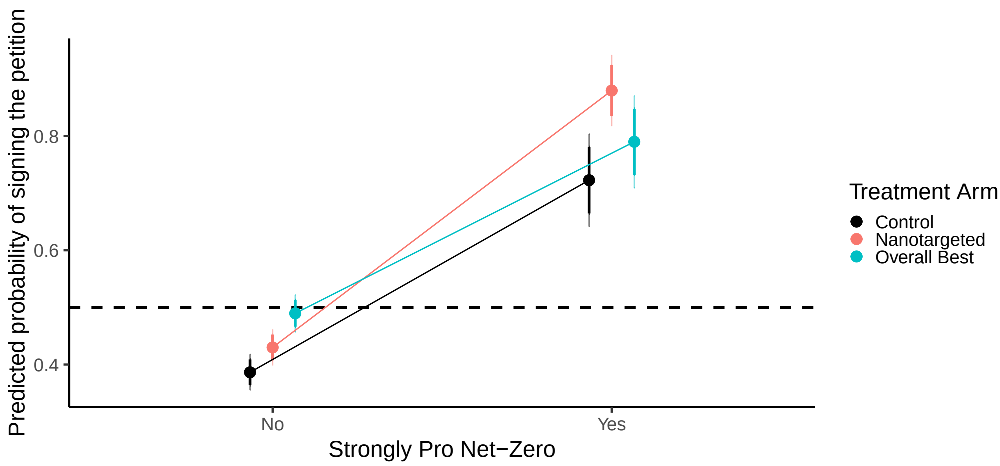
We find that both switchers and moderates are overall less likely to report behavioral support for the climate policy message (for example, via signing a petition) in all conditions. We find the same broad rankings among switchers and non-switchers and moderates and non-moderates: the overall best message performs best, the nano-targeted message is a close second and they both perform better than the randomly assigned (con-

[26] As operationalized by a value between 8 and 12 on a climate index going from 0 to 20, based on the pre-test items asking if it is important to reach net-zero by 2050 and if the government should spend more on climate protection.

trol) messages. We therefore do not find any evidence that nano-targeting is particularly effective on the persuadable (proxied via party switching and attitudinal moderation).

We discover that nano-targeting is, however, more effective among those that already held strongly progressive attitudes on climate. Figure 7 below illustrates the interaction effect between treatment assignment and holding progressive climate stances in the pre-test:[27] the probability of supporting the pro-net zero message goes up significantly more in the nano-targeted group for respondents who were climate progressives already (compared to the overall best message ($p < 0.05$) and compared to the control condition ($p < 0.1$)). Furthermore, the stronger effects of nano-targeting among climate progressives are not merely driven by the higher education of these respondents. In Appendix L, we show that nano-targeted petitions are even more effective among less educated climate progressives than highly educated ones (though the overall-best petition is likewise also more effective for that demographic).[28]

Figure 7: Heterogeneity by Climate Attitude Strength: Climate Progressives



Note: Errorbars are 95% and 84% confidence intervals based on HC1-robust standard errors.

In sum, this is suggestive evidence that nano-targeting may not be useful for persuading voters who are on the fence, but instead could be useful for mobilizing voters who

[27] As operationalised by a value above 18 on a climate index going from 0 to 20, based on the pre-test items asking if it is important to reach net-zero by 2050 and if the government should spend more on climate protection. Thus, anyone operationalised as strongly pro net-zero gave the maximum score for at least one of the two pre-test items.

[28] Separately, we largely do not find substantial differences in model prediction error across education levels (see Appendix Figure M25). On average, the prediction error is positive for the high education group (i.e. those with a higher education level tended to return a more positive rating than predicted) and negative for both mid and low education groups. None of these differences are statistically significantly different from zero. Between groups, only the difference between high and mid education levels is statistically significant ($p = 0.04$).

are already sympathetic towards the policy's broad direction of travel (in here: climate emissions reduction). This result suggests, though, that nano-targeting may exacerbate polarization, as it works best in reinforcing pre-existing attitudes rather than influencing the undecided.

Potential for Backlash: What Happens When You Receive a Mis-Targeted Message?

Another original aspect of our nano-targeting test is the random assignment of climate message attributes in the control group. The randomization allows us to test the consequences of receiving a mis-targeted message of varying degree (H4a and H4b). Recall that we hypothesize that mis-targeted messages should provoke backlash, resulting in decreased support for the specific climate message or a decline in pro-climate positions. In Table 2 we leverage this same control group sample (approximately 1,000 respondents from the Opinium survey) to proxy the degree of mis-targeting via the predicted rating measure. Low predicted ratings mean that the particular message received through the random assignment should have been disliked. For robustness, we consider a continuous, binary, and binned operationalisation of predicted ratings.

Table 2: Association between subjects' observed behavior in Phase II and predicted ratings, for those assigned random petitions

	Rating			Choice		
	(1)	(2)	(3)	(4)	(5)	(6)
Predicted rating	1.10^{***}			0.14^{***}		
	(0.11)			(0.02)		
Predicted rating < 5?		-1.54^{***}			-0.19^{***}	
		(0.18)			(0.03)	
Predicted rating lower tertile			-1.51^{***}			-0.16^{***}
			(0.21)			(0.04)
Predicted rating upper tertile			0.72^{***}			0.13^{***}
			(0.21)			(0.04)
Intercept	-0.28	5.94^{***}	5.48^{***}	-0.29^{**}	0.51^{***}	0.43^{***}
	(0.54)	(0.12)	(0.15)	(0.09)	(0.02)	(0.03)
R^2	0.09	0.07	0.10	0.06	0.04	0.06
Adj. R^2	0.09	0.07	0.10	0.05	0.04	0.05
Num. obs.	1043	1043	1043	1043	1043	1043

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

We find that, indeed, mis-targeting decreases support for the pro-net-zero message received, both in the form of rating and behavioral choice (signing a petition containing those policy measures).[29] Furthermore, what we were not expecting is the asymmetry we find in the control group between individuals who randomly received well-liked policies, and those who received significantly mis-targeted policies. Comparing the coefficients for those in the lower and upper predicted rating tertiles, we find that those who received messages that our model predicts they would dislike reduced their support for the petition, in terms of their rating, over twice as much as those who received messages predicted to be rated highly. In terms of our behavioral measure, although the effects are not as pronounced, those receiving mis-targeted messages were three percentage points less likely to support the petition than those receiving well-targeted messages were to support the petition. In other words, for every 100 voters, mis-targeting dissuades three more individuals than effective targeting persuades them.[30]

Our results provide suggestive evidence that being assigned personalized most-disliked messages has stronger dissuasive effects than receiving personalized most-liked messages is likely to have persuasive effects. This finding is in line with findings from Hersh and Schaffner (2013) on mis-targeting, and from Nicholson (2012) on the stronger effect of out-party cues than in-party cues. The implication of this result is that while nano-targeting may not yield significantly large advantages when applied to persuasion, when applied for dissuasion and de-mobilization from a particular aim, the technique has the potential to be more powerful.

We speculate that this asymmetry in nano-targeting messaging effectiveness may be due to negativity bias arising from loss-aversion or automatic vigilance mechanisms. Losses are more cognitively loaded than wins: negative information acts as an 'attention highlighter' (automatic vigilance theory) (Pratto and John 1991) and triggers stronger emotional reactions than positive information (loss aversion theory) (Kahneman 2011). Given these features of human cognition, it is not surprising that reading disliked climate petitions will move one's climate attitudes more strongly than reading favored climate

[29] Those mis-targeted in the control condition also have a substantially lower average petition rating (3.97) compared to the nano-targeted and overall best message treatment arms (5.44 and 5.87, respectively).

[30] More generally, as shown in Appendix Figure **??**, the median predicted rating for those who declined these petitions was lower than those who accepted them, and below the midpoint of our rating scale.

petitions.

Discussion & Conclusion

This study illustrates and tests the effectiveness of a more advanced version of micro-targeting, which we define as "nano-targeting". These technologies for hyper-personalization of campaign messaging are going to become increasingly scalable and less costly in the age of AI: the systematic empirical testing of their persuasion potential, as well as of their potential for democratic erosion, are pressing concerns. Normative and legal analyses already raised important arguments against micro-targeting on the back of the Cambridge Analytica scandal. However, at the same time, political science research shows only limited persuasive advantages, and limited usage of this technology by mainstream political parties. This study squares this circle by doing several things: (a) testing whether a more precisely calibrated level of micro-targeting would prove more effective relative to more standard campaign tools, such as using an "overall best message"; (b) empirically assessing whether nano-targeting is helpful in reaching the persuadable; (c) testing the extent of backlash that can result from mis-targeting.

We implemented a two phase study. In phase one, we fielded a conjoint survey experiment on a nationally representative sample of 2,000 UK adults to identify which climate change policy positions/messages appeal to particular sets of individuals, as defined by fine-grained combinations of demographic and political characteristics. In phase two, we used a standard "vignette" message design (with over 3,000 UK respondents) to test these more-precisely targeted policy platforms against the overall best message and randomly assigned messages. We also investigated the effectiveness of this nano-targeting for undecideds/climate moderates as well as the magnitude of the backlash effects from mis-targeting.

Our findings can be summarized as follows. First, from a campaigns perspective, we find that nano-targeting does not necessarily yield higher rates of support than the overall best message. And, what is more, while nano-targeting has generally a positive effect relative to showing a random policy package (the control group), this effect is not robust across all outcomes. These results confirm previous research that suggests that fine-grained message targeting is not worth the significant extra costs and risks of this technique.

Second, we find that nano-targeting can work better than the overall best message for individuals whose petition ratings can be modeled with higher accuracy. Indeed, for these individuals, our nano-targeted message increased the probability of supporting climate action by 16 percentage points relative to the overall best message[31]. Nevertheless, there is a set of respondents for whom it is difficult to predict attitudes, and for whom nano-targeting is ineffective, which adds to the costs of nano-targeting and contributes to bringing down the average overall effect of nano-targeting. This suggests that nano-targeting can work if campaigns have high quality data about respondents and the political attitudes of the respondents are stable – in other words the particular individuals to be targeted are easily predictable. Again, though, the contexts in which the preferences of many individuals are so predictable are likely to be rare, and/or will require extremely precise (costly) data.

Third, we do not find support for our third hypothesis on enhanced effects of nano-targeting on undecided and/or moderate individuals. This is further evidence that nano-targeting is not an exceptional persuasion technology, as it fails to more strongly mobilize and persuade the critical constituency that any campaign seeks to influence: the undecided. We furthermore discover that it is individuals who were already progressive on climate protection who exhibited the strongest pro-climate shifts in attitudes and behaviors after being nano-targeted. This discovery, together with the null effect on moderates, contributes compelling evidence that nano-targeting is a polarizing technology.

Last but not least, we find evidence consistent with a strong demobilizing effect when individuals are assigned particularly "bad" messages (that are not aligned with their prior political preferences). This result is consistent across both our binary choice and rating outcomes. We find that most-disliked messages move attitudes (against net zero climate policy actions, in our case) more than most-liked messages. We concede that these messages were not deliberately badly targeted, so we cannot claim that this is strong evidence that negative-targeting is more powerful than positive-targeting. Nevertheless, this result is an important indicator that when applied with the aim of dissuasion and

[31] For those individuals where the prediction error is less than 1.

de-mobilization, nano-targeting has the potential to be quite powerful.

In terms of the efficacy of political campaigns, our results suggest that micro- or nano-targeting can be effective, but only under quite restrictive conditions. Micro- and nano-targeting work best on more predictable groups, for whom predictive modeling is accurate (either due to superior data quality held on them or strong attitude stability), and on core supporters (which is helpful when the campaign aim is mobilization and not persuasion). Moreover, our predictions were made using a relatively "rich" set of covariate descriptors of individuals' demographics and political positions. Political parties may be more constrained in terms of what data they can collect (for example, from electoral registers) which, we suspect, would limit the ability to train effective nano-targeting models. Overall, the benefits of nano-targeting are unlikely to outweigh the costs, even for well funded parties or interest groups, which perhaps explains why there is little evidence so far of the application of micro-targeting techniques by mainstream parties.

More broadly, our findings suggest a positive message for electoral democracy and political campaigns going forward. Parties and interest groups can shift public attitudes with overall best policy platforms, rather than trying to selectively hide or emphasize different pledges with different people. Nano-targeting can work, but only for very specific types of people, only for specific campaign aims, and at high cost. AI offers the prospect of reducing these costs, but the likely benefits of nano-targeting seem low.

That said, we find suggestive evidence of the potential of this technology for democratic erosion. Our study is the first to provide systematic empirical evidence about the heterogeneous effects of micro-targeting. We discover that nano-targeting mostly works on the already converted, which can exacerbate polarization. Last but not least, we find some suggestive evidence that dissuasion-oriented micro-targeting (presenting individuals with contents they most *dislike*) is more effective than persuasion-oriented micro-targeting. Our study suggests that nano-targeting may be more useful to entrepreneurs who aim to dissuade, demobilize and/or polarize, which perhaps explains why extremist and/or foreign interference campaigns have an incentive to use these techniques.

Finally, our results are based on testing the effect of nano-targeting on a particular policy issue: climate policy. Climate policy is a highly salient issue, and the policy choices on this issue combine both distributive conflicts as well as collective action aspects. As such, climate policy is akin to many other contemporary policy issues, which lends support to the generalizability of our findings. Nonetheless, we acknowledge that there are also certain scope conditions in our study – such as the technical nature of the particular policy choices we set out – and these may limit how general our findings are. With this in mind we hope future research will examine the effects of nano-targeting in other salient policy domains, and in other countries and political contexts.

References

Allen, Myles, Opha Pauline Dube, William Solecki, Fernando Aragón-Durand, Wolfgang Cramer, Stephen Humphreys, Mikiko Kainuma et al. 2018. "Special report: Global warming of 1.5 C." *Intergovernmental Panel on Climate Change (IPCC)* 677:393.

Argyle, Lisa P, Ethan C Busby, Joshua R Gubler, Alex Lyman, Justin Olcott, Jackson Pond and David Wingate. 2025. "Testing theories of political persuasion using AI." *Proceedings of the National Academy of Sciences* 122(18):e2412815122.

Armstrong II, David A. and William Poirier. 2025. "Decoupling Visualization and Testing when Presenting Confidence Intervals." *Political Analysis* First View:1–6.

Bergquist, Magnus, Andreas Nilsson, Niklas Harring and Sverker C Jagers. 2022. "Meta-analyses of fifteen determinants of public opinion about climate change taxes and laws." *Nature Climate Change* 12(3):235–240.

Bergquist, Parrish, Matto Mildenberger and Leah C Stokes. 2020. "Combining climate, economic, and social policy builds public support for climate action in the US." *Environmental Research Letters* 15(5):054019.

Bergstra, James, Rémi Bardenet, Yoshua Bengio and Balázs Kégl. 2011. "Algorithms for hyper-parameter optimization." *Advances in neural information processing systems* 24.

Budge, Ian and Dennis J Farlie. 2025. *Explaining and predicting elections: Issue effects and party strategies in twenty-three democracies*. Taylor & Francis.

Chen, Tianqi and Carlos Guestrin. 2016. Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*. pp. 785–794.

Cooper, Joel. 2007. *Cognitive Dissonance: 50 Years of a Classic Theory*. Sage Publications.

Coppock, Alexander. 2023. *Persuasion in parallel: How information changes minds about politics*. University of Chicago Press.

Dobber, Tom, Damian Trilling, Natali Helberger and Claes de Vreese. 2023. "Effects of an issue-based microtargeting campaign: A small-scale field experiment in a multi-party setting." *The Information Society* 39(1):35–44.

Dommett, Kate and Simon Kruschinski. 2024. "Data-driven campaigning: Using data, analytics, technology, and personnel for a new era of campaigning?" *The Routledge Handbook of Political Campaigning* pp. 199–212.

Druckman, James N. 2024. "Persuasive political targeting: A remarkably understudied dynamic." *Unpublished manuscript, Northwestern University* .

Endres, Kyle. 2016. "The accuracy of microtargeted policy positions." *PS: Political Science & Politics* 49(4):771–774.

Endres, Kyle. 2020. "Targeted issue messages and voting behavior." *American Politics Research* 48(2):317–328.

Fairbrother, Malcolm. 2022. "Public opinion about climate policies: A review and call for more studies of what people want." *PLoS Climate* 1(5):e0000030.

Fankhauser, Sam, Stephen M Smith, Myles Allen, Kaya Axelsson, Thomas Hale, Cameron Hepburn, J Michael Kendall, Radhika Khosla, Javier Lezaun, Eli Mitchell-Larson et al. 2022. "The meaning of net zero and how to get it right." *Nature Climate Change* 12(1):15–21.

Fesenfeld, Lukas Paul, Michael Wicki, Yixian Sun and Thomas Bernauer. 2020. "Policy packaging can make food system transformation feasible." *Nature Food* 1(3):173–182.

Festinger, Leon. 1957. *A Theory of Cognitive Dissonance.* Stanford University Press.

Gahn, Christian. 2025. "How Much Tailoring is too Much? Voter Backlash on Highly Tailored Campaign Messages." *The International Journal of Press/Politics* 30(4):1071–1095.

Gerber, Alan S, Donald P Green and Christopher W Larimer. 2008. "Social pressure and voter turnout: Evidence from a large-scale field experiment." *American Political Science Review* 102(1):33–48.

Grimmer, Justin, Solomon Messing and Sean J Westwood. 2017. "Estimating heterogeneous treatment effects and the effects of heterogeneous treatments with ensemble methods." *Political Analysis* 25(4):413–434.

Hackenburg, Kobi, Ben M Tappin, Paul Röttger, Scott A Hale, Jonathan Bright and Helen Margetts. 2025. "Scaling language model size yields diminishing returns for single-message political persuasion." *Proceedings of the National Academy of Sciences* 122(10):e2413443122.

Hackenburg, Kobi and Helen Margetts. 2024. "Evaluating the persuasive influence of political microtargeting with large language models." *Proceedings of the National Academy of Sciences* 121(24):1–8.

Hainmueller, Jens, Daniel J Hopkins and Teppei Yamamoto. 2014. "Causal inference in conjoint analysis: Understanding multidimensional choices via stated preference experiments." *Political Analysis* 22(1):1–30.

Hersh, Eitan D. and Brian F. Schaffner. 2013. "Targeted Campaign Appeals and the Value of Ambiguity." *The Journal of Politics* 75(2):520–534.

Hillygus, D Sunshine and Todd G Shields. 2008. *The persuadable voter: Wedge issues in presidential campaigns*. Princeton University Press.

Huber, Lena Maria and Martin Haselmayer. 2025. "Promising links: how parties combine policy issues with group appeals." *West European Politics* 48(7):1732–1759.

Huber, Robert A, Michael L Wicki and Thomas Bernauer. 2020. "Public support for environmental policy depends on beliefs concerning effectiveness, intrusiveness, and fairness." *Environmental Politics* 29(4):649–673.

Kahneman, Daniel. 2011. *Thinking, fast and slow*. macmillan.

Kalla, Joshua L and David E Broockman. 2018. "The minimal persuasive effects of campaign contact in general elections: Evidence from 49 field experiments." *American Political Science Review* 112(1):148–166.

Kohl, Uta and Jacob Eisler. 2021. *Data-driven Personalisation in markets, politics and law.* Cambridge University Press.

Kreiss, Daniel. 2017. "Micro-targeting, the quantified persuasion." *Internet Policy Review* 6(4).

Lavigne, Mathieu. 2021. "Strengthening ties: The influence of microtargeting on partisan attitudes and the vote." *Party Politics* 27(5):965–976.

Levy, Gilat and Ronny Razin. 2020. "Social Media and Political Polarisation." *LSE Public Policy Review* 1(1):1–7.

Li, Lisha, Kevin Jamieson, Giulia DeSalvo, Afshin Rostamizadeh and Ameet Talwalkar. 2018. "Hyperband: A novel bandit-based approach to hyperparameter optimization." *Journal of Machine Learning Research* 18(185):1–52.

Matthes, Jörg, Melanie Hirsch, Marlis Stubenvoll, Alice Binder, Sanne Kruikemeier, Sophie Lecheler and Lukas Otto. 2022. "Understanding the democratic role of perceived online political micro-targeting: longitudinal effects on trust in democracy and political interest." *Journal of Information Technology & Politics* 19(4):435–448.

Matz, SC, JD Teeny, Sumer S Vaid, H Peters, GM Harari and M Cerf. 2024. "The potential of generative AI for personalized persuasion at scale." *Scientific Reports* 14(1):4692.

Nicholson, Stephen P. 2012. "Polarizing cues." *American Journal of Political Science* 56(1):52–66.

Palmer, Alexis and Arthur Spirling. 2023. "Large language models can argue in convincing ways about politics, but humans dislike AI authors: implications for governance." *Political Science* 75(3):281–291.

Petty, Richard E and John T Cacioppo. 1986. *Communication and persuasion: Central and peripheral routes to attitude change.* New York: Springer.

Pratto, Felicia and Oliver P John. 1991. "Automatic vigilance: the attention-grabbing power of negative social information." *Journal of Personality and Social Psychology* 61(3):380.

Robinson, Thomas S and Raymond M Duch. 2024. "How to detect heterogeneity in conjoint experiments." *The Journal of Politics* 86(2):412–427.

Schoenegger, Philipp, Francesco Salvi, Jiacheng Liu, Xiaoli Nan, Ramit Debnath, Barbara Fasolo, Evelina Leivada, Gabriel Recchia, Fritz Günther, Ali Zarifhonarvar et al. 2025. "Large Language Models Are More Persuasive Than Incentivized Human Persuaders." *arXiv preprint arXiv:2505.09662* .

Simchon, Almog, Matthew Edwards and Stephan Lewandowsky. 2024. "The persuasive effects of political microtargeting in the age of generative artificial intelligence." *PNAS Nexus* 3(2):1–5.

Tappin, Ben M, Chloe Wittenberg, Luke B Hewitt, Adam J Berinsky and David G Rand. 2023. "Quantifying the potential persuasive returns to political microtargeting." *Proceedings of the National Academy of Sciences* 120(25):1–10.

Votta, Fabio, Simon Kruschinski, Mads Hove, Natali Helberger, Tom Dobber and Claes de Vreese. 2024. "Who Does (n't) Target You? Mapping the Worldwide Usage of Online Political Microtargeting." *Journal of Quantitative Description: Digital Media* 4.

Yeh, Kai-Ching, Jou-An Chi, Da-Chen Lian and Shu-Kai Hsieh. 2023. Evaluating interfaced llm bias. In *Proceedings of the 35th Conference on Computational Linguistics and Speech Processing (ROCLING 2023).* pp. 292–299.

Zhirkov, Kirill. 2022. "Estimating and using individual marginal component effects from conjoint experiments." *Political Analysis* 30(2):236–249.

Zhou, Jianlong, Heimo Müller, Andreas Holzinger and Fang Chen. 2024. "Ethical ChatGPT: Concerns, challenges, and commandments." *Electronics* 13(17):3417.

Zuiderveen Borgesius, Frederik, Judith Möller, Sanne Kruikemeier, Ronan Ó Fathaigh, Kristina Irion, Tom Dobber, Balazs Bodo and Claes H de Vreese. 2018. "Online political microtargeting: Promises and threats for democracy." *Utrecht Law Review* 14(1):82–96.

Dissuasion and polarization vs. persuasion in nano-targeting: experimental evidence from climate policy messaging

Online Appendix

Miriam Sorace[1], Thomas Robinson[2], Simon Hix[3], and Joris Frese[3]

[1]University of Reading, UK. [2]London School of Economics and Political Science, UK.

[3]European University Institute, Italy.

Contents

A Conjoint Experiment Pre-Analysis Plan

ASPREDICTED

Wharton
CREDIBILITY LAB
UNIVERSITY of PENNSYLVANIA

SRHF Climate Policy Conjoint (#216633)

Author(s)
This pre-registration is currently anonymous to enable blind peer-review.
It has 4 authors.

Pre-registered on: 2025/03/10 - 03:15 AM (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?
What climate policies are preferred by the public, and by which specific sub-group?

3) Describe the key dependent variable(s) specifying how they will be measured.
Forced petition choice and 0-10 rating scale for each petition presented in the conjoint

4) How many and which conditions will participants be assigned to?
The survey experiment is a large fractional factorial experiment (conjoint survey experiment) with 7 attributes having between 3 and 5 levels each. The respondents will be exposed to 4 rounds of a pairwise conjoint, and see 8 total petitions.

5) Specify exactly which analyses you will conduct to examine the main question/hypothesis.
Standard conjoint analyses (AMCEs, MMs, split sample MMs) and predictive modelling (based on the Robinson-Duch heterogeneity estimation strategy). Socio-demographic and political covariates to be collected (and leveraged in the heterogeneous effects analyses): age, gender, education, religious affiliation, rural/urban, income and/or economic insecurity/standards of living, ge24 vote, vote intention.

6) Describe exactly how outliers will be defined and handled, and your precise rule(s) for excluding observations.
We will exclude participants who complete the survey in less than 1 minute. If outlying observations will be detected, robustness analyses including and excluding them will be attempted, and both sets of results will be reported in the main manuscript or its supplementary materials.

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.
2,000 voting age UK respondents (nationally representative YouGov panel)

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.

B Nano-Targeting Experiment Pre-Analysis Plan

SRHF Climate Policy Nanotargeting (#221931)

Author(s)
This pre-registration is currently anonymous to enable blind peer-review.
It has 4 authors.

Pre-registered on: 2025/04/08 - 02:50 AM (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?
H1 (attitudes): Nano-targeted pro-net zero messaging will result in a bigger shift towards pro-net zero attitudes and behaviors in the post-test, compared to the best overall message, and compared to the random (control) message.
H1a(best vs. control): The best overall message will result in a bigger shift towards pro-net zero attitudes and behaviors in the post-test, compared to the random (control) message.
H2 (petition support): Nano-targeted pro-net zero messaging will result in larger support for the climate protection petition compared to the best overall message, and compared to the random (control) message.
H2a: (best vs. control): The best overall message will result in larger support for the climate protection petition in the post-test, compared to the random (control) message.

3) Describe the key dependent variable(s) specifying how they will be measured.
The dependent variables of the study consist of (a) a battery on general attitudes towards net zero and government action on climate (asked pre and post treatment); (b) a battery on climate petition-specific attitudes; (c) a behavioral measure of climate policy interest.

4) How many and which conditions will participants be assigned to?
Participants will be randomly assigned to one of three experimental conditions: (1) the control condition (the respondent will see a randomly generated petition), and (2) the "uniform best message" condition (the petition that we found from a previous conjoint study to be the most supported petition overall in the UK); (3) the nano-targeted petition condition (the respondent will see a petition that, via predictive modeling, we determine to be the most likely to be supported by individuals with their same demographic and political characteristics. The predictive model is calibrated on a previous conjoint experiment which has its own separate pre-analysis plan).

5) Specify exactly which analyses you will conduct to examine the main question/hypothesis.
1. Predicting out on a larger sampling frame potential respondents (around 9k individuals offered by Opinium, to take into account lower than perfect response rates) to be provided by Opinium. On the basis of the predictive model built on the conjoint calibration study (which has its own separate pre-analysis plan), we will pre-assign nano-targeted petitions to all individuals in the sampling frame. If, via randomization, they get assigned to the nano-targeted condition, they will see the nano-targeted message we predicted out for them.
2. Regression models, mean tests and heterogeneous effect analyses (split samples or interaction effects). Should balance checks reveal non-equivalence of the experimental groups the models will be estimated with demographic and political controls.

6) Describe exactly how outliers will be defined and handled, and your precise rule(s) for excluding observations.
We will exclude participants who complete the survey in less than 30 seconds or who otherwise fail Opinium's attention and speeding checks. If outlying observations are detected through standard statistical tests, robustness analyses including and excluding them will be attempted, and both sets of results will be reported in the main manuscript or its supplementary materials.

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.
2500 UK respondents will be recruited, giving us ~833 observations per experimental group. The number has been determined via a mixture of cost and statistical power considerations. The power analysis revealed that we have 90% power to detect small effects of 0.2 standard deviations in our main tests. While 833 participants per experimental group is a very large number, it was important to ensure significant variation in socio-demographic and political characteristics for the nano-targeting to be applied to sufficient combinations of demographic and political characteristics. Furthermore, as we expect the difference between the nano-targeting and the 'best overall message' condition to be small, we require a large sample size. The pre-registered heterogeneous effects analyses (see below) also require additional sample size to be estimated more precisely. Equally, the control group (random petition allocation) will also be used to test the effect of receiving different degrees of dispreferred petitions, which also require a larger sample size.

8) Anything else you would like to pre-register? (e.g., secondary analyses, variables collected for exploratory purposes, unusual analyses planned?)
We'd like to explore whether H1, H1a, H2 and H2a work differently at different levels of political interest, at different levels of past vote and vote intention (and, particularly for switchers vs. non-switchers), for climate moderates or undecideds (from the pre-test battery), and for different educational, age, gender, and economic/social class groups, or for individuals living in different geographies.

Some respondents in the control group, due to the random petition assignment, may see petitions that will be particularly at odds with what the predictive model would assign them as their favored petition. Hence, we also hypothesize that these subjects may backlash away from climate protection positions and will particularly strongly reject the specific petition presented to them.

The core heterogeneous effects and backlash hypotheses we'd like to formally test are formulated as such:

H3: Nano-targeted pro-net zero messaging will result in a bigger shift towards pro-net zero attitudes and behaviors in the post-test, and larger support for the climate protection petition, compared to the best overall message, and compared to the random message, particularly among switchers, moderates and undecided respondents.

H4a: Mistargeted pro-net zero messaging will result in a shift against pro-net zero attitudes and behaviours in the post-test, and lower support for the climate protection petition, compared to the best overall message, and compared to the nano-targeted message.

H4b: The higher the degree of mistargeting of pro-net zero messaging in the control group, the bigger the shift against pro-net zero attitudes and behaviours in the post-test, and the lower the support for the climate protection petition, compared to the best overall message, and compared to the nano-targeted message.

C Ethics

The studies were fully reviewed and approved by the ethics committee of [University name redacted for anonymity purposes] on March 17th 2025, i.e. before our data collection started.

The polling companies involved – YouGov and Opinium – are world-leading companies recognised by the British Polling Council and with extensive experience in anonymisation, consent, privacy and data protection. They dealt with the participant selection and ask for consent at the point of invitation to the panel and to the specific survey. The experimental treatments were designed not to involve deception or psychological harm. All petitions were introduced as hypothetical. No psychometric measure is part of the survey batteries or predictive modelling. The persuasion treatments concern a public good: climate protection. There was no disproportionate or coercive incentive in-built into the panel recruitment process, and participants could withdraw from the surveys at any time. The surveys lasted approximately 8 minutes maximum, thus not requiring unduly punitive time commitments.

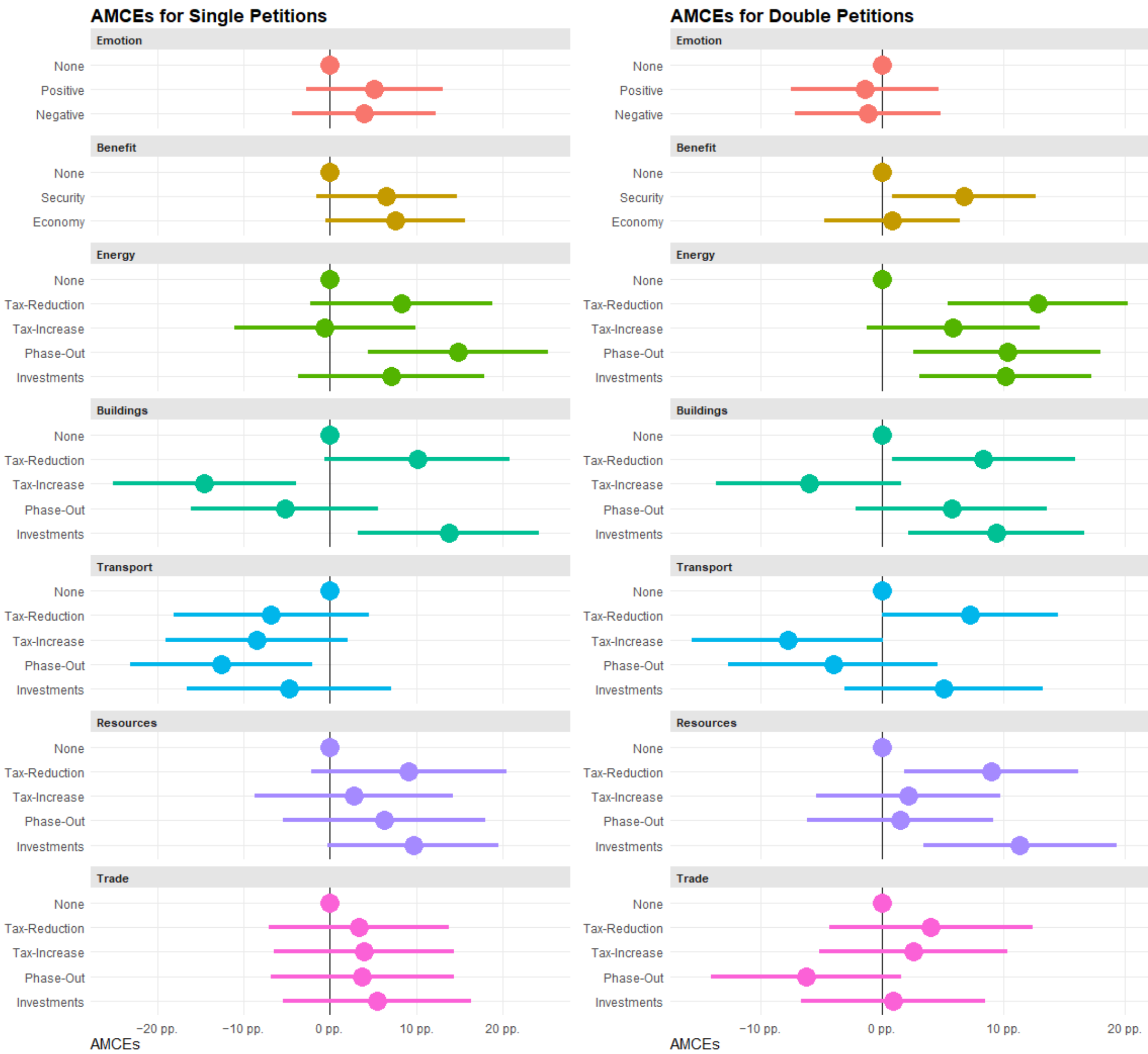
Economic coercion concerns are likewise low to non-existent: survey participants are typically nominally rewarded by both YouGov and Opinium via competitions and point systems. In the case of YouGov: "members are paid for surveys and other data sharing activity in points. The points vary based on the length and complexity of the survey – longer and more complex surveys receive more points. A short, simple survey pays 50-100 points. Points are stored in a member's YouGov Wallet, and once they reach a certain balance – they can convert these points to a reward of their choice" (typically 5000 points mean a £50 reward). See: https://yougov.co.uk/about/panel. In the case of Opinium, the reward is usually 50p per survey, and, as per their FAQ page (see: https://www.opiniumresearch.co.uk/faq.aspxFAQLink8) "for longer surveys we will always aim to increase the incentive for taking part. This money will appear in your Opinium account within 24 hours of survey completion. Upon reaching £25, participants can claim a reward via the panel area. We aim to be as flexible as possible offering payment by bank transfer, Amazon voucher and/or an optional charitable donation".

In terms of data protection, the two companies only provided us with data table of survey answers with a randomly generated numerical ID codes for each respondent. They will send us the completely anonymised data via excel files. Such files will be securely saved on the team's private GitHub repository and then made publicly available for replication purposes after the study is published. Further information on privacy and secure handling of personal data from YouGov's methodology page: "Respondents have the following rights to help protect their privacy and control of their data: They choose the survey invitations they accept and the questions they answer. They can request that we not sell or share their personal data to our clients. (We never sell personal data to data brokers.) They can be notified which categories of personal data we collect and the purposes for which data is being used. They can request a copy of the data that we hold about them. They can ask us to correct any inaccurate data about them. They can request that we delete the personal data we hold about them. They can opt out of cookies, including ones used for targeted advertising and tracking. Occasionally clients might want more details about respondents, such as contact information to ask them follow-up questions. It is up to panelists whether to provide this kind of information; it is never required. When we report findings, they are aggregated to a degree that protects any individual respondent from being identified. For sensitive questions we often provide a response option such as "prefer not to say"; skipping also is often provided as an option. See our research privacy and cookies notice and our consumer health data privacy policy." And on Opinium's survey privacy page: "Opinium Research LLP conducts market research on behalf of its clients. Information supplied by you will be disclosed to RGA in pseudonymised form, and will not be used to identify you, or to contact you without your permission.[...] You have certain rights regarding your personal information, which include: the right to access your personal information; rectify the information we hold about you; erase your personal information; restrict our use of your personal information; and object to our use of your personal information. If the legal basis for our processing your personal data or special category data is consent or explicit consent, as appropriate, you have the right to withdraw your consent at any time. Please note that if you withdraw your consent, the lawfulness of any processing undertaken before such withdrawal will not be affected. To exercise your rights, please contact us at gdpr@opinium.com"

D Prolific Conjoint Pilot

Before we conducted the main conjoint experiment on a nationally representative sample of 2000 British YouGov respondents, we first fielded a small pilot study on 208 British Prolific respondents to 1) test the effectiveness and popularity of our petition attributes and 2) compare the traditional forced choice conjoint format with two vignettes side by side to a more parsimonious format with just one vignette per survey page. The former was merely done to ensure that we could achieve some level of predictive power for the nano-targeting part of our study (which would be difficult if none of our respondents actually had strong preferences between different attribute levels). The latter was done because we wanted to ensure that our main conjoint experiment would be cost effective.

Figure D1: Single and Double AMCEs from the Pilot Conjoint



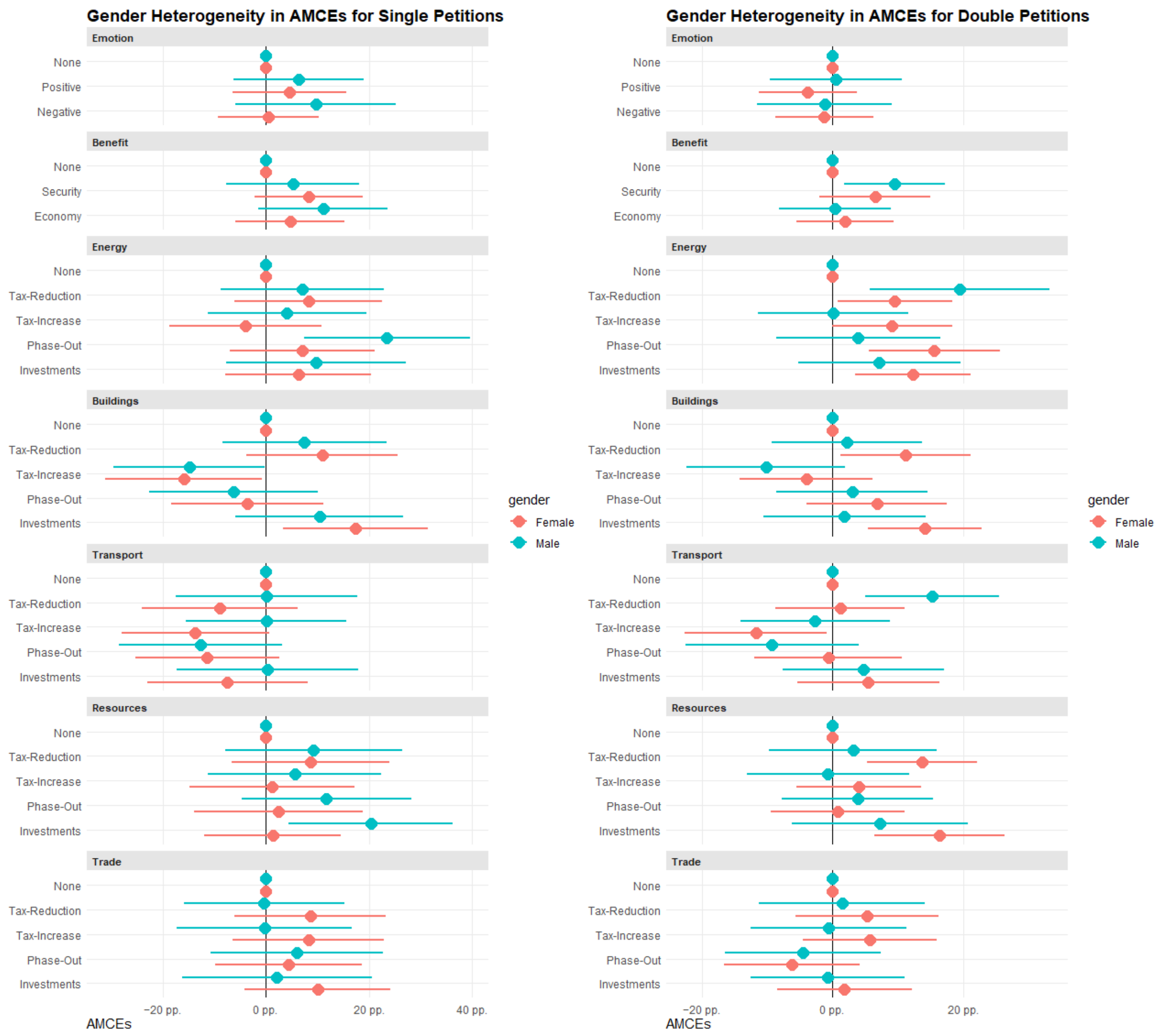
Initially, the double vignette format was our preferred option, because it came with higher statistical power, due to the fact that our costs with YouGov would increase with each additional item and survey page. For this reason, we would have only been able to field four single petitions per respondent with our budget, for an effective sample size of 8000 petitions, while the four double petitions gave us an effective sample size of 16000 petitions at the same prize. However, one might expect that the double petitions could come with some caveats, given that each petition contains lots of text and it might be overwhelming to constantly be exposed to two petitions at once. To ensure that the increase in sample size and statistical power did not come at the cost of decreased attribute effectiveness, decreased heterogeneous effects (which are important for our predictive model), increased petition fatigue, or increased user discomfort, we incorporated four single petitions and four double petitions (for a total of 12 overall petitions) in each pilot survey. The order of formats was randomized to ensure no systematic differences due to overall survey fatigue.

In Figure D1, we first show the AMCEs for each petition attribute level. All levels are always compared to the level with no proposed changes. As can be seen, even on the small sample of 208 respondents, several AMCEs are already statistically significant at the five percent level in both formats, suggesting that the double exposure did not decrease attribute effectiveness (in fact, likely due to the increased statistical power, more AMCEs were statistically significant for the double petitions compared to the single petitions).

Beyond general attribute effectiveness, another important factor to consider was the level of heterogeneity in each format. In Figure D2, we show that there is again no apparent loss in heterogeneity in the double format (we restrict this showcase to the gender dimension simply for illustration purposes, but this finding also holds across other strata).

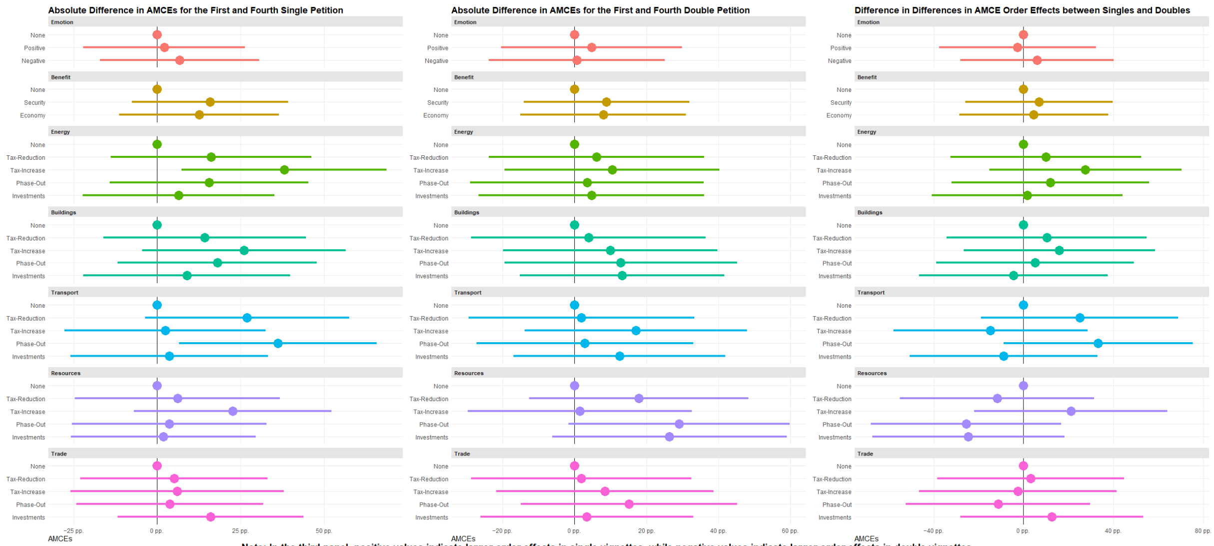
To test for differential petition fatigue, we estimated the difference in AMCEs between the first and the fourth petition separately for both formats, and then the difference in differences in AMCEs between the two formats. This allowed us to assess whether any order effects were at play, and whether those differed between the two formats. In Figure D3, we show that only two of the attribute levels for the single petitions were subject to

Figure D2: Gender Heterogeneity in the Pilot Conjoint



significant petition order effects, and none of the attribute levels for the double petitions. In the rightmost column, we can see that the difference (between the formats) in differences (between the first and last petition) in AMCEs is not statistically significant for a single level. Descriptively, order effects are larger for single petitions across 15 levels and larger for double petitions across 9 levels.

Figure D3: Order Effects in the Pilot Conjoint



Note: In the third panel, positive values indicate larger order effects in single vignettes, while negative values indicate larger order effects in double vignettes.

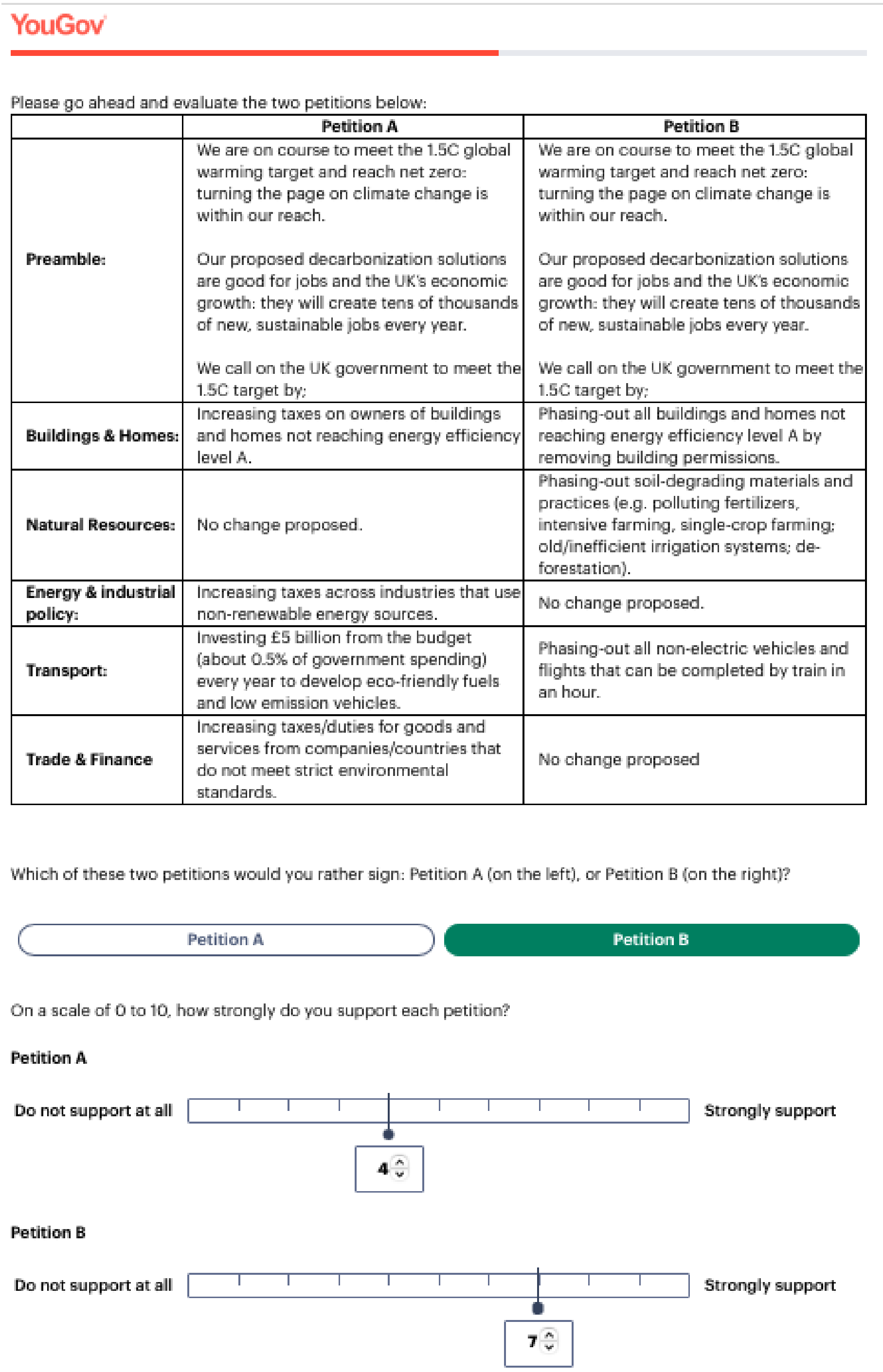
Lastly, we also shared the pilot survey with several survey experts to collect outside opinions on user comfort. While split in their judgments, the majority of those experts also recommended for us to stick with the double petitions.

Based on the increased cost effectiveness and the resulting increase in statistical power, the pilot finding that there were no substantial downsides to the double petitions, and the outside opinions from several experts, we ultimately decided to stick with our original plan of fielding the main YouGov survey using the double petition format.

E YouGov Conjoint Excerpt

Below, we show an example conjoint table as it was shown to our YouGov respondents. Below the table, we also show the questions used to measure our main outcomes, as they were asked of our YouGov respondents: the binary petition choice, and the continuous petition ratings.

Figure E4: Conjoint Experiment Example


YouGov

Please go ahead and evaluate the two petitions below:

	Petition A	Petition B
Preamble:	We are on course to meet the 1.5C global warming target and reach net zero: turning the page on climate change is within our reach. Our proposed decarbonization solutions are good for jobs and the UK's economic growth: they will create tens of thousands of new, sustainable jobs every year. We call on the UK government to meet the 1.5C target by:	We are on course to meet the 1.5C global warming target and reach net zero: turning the page on climate change is within our reach. Our proposed decarbonization solutions are good for jobs and the UK's economic growth: they will create tens of thousands of new, sustainable jobs every year. We call on the UK government to meet the 1.5C target by:
Buildings & Homes:	Increasing taxes on owners of buildings and homes not reaching energy efficiency level A.	Phasing-out all buildings and homes not reaching energy efficiency level A by removing building permissions.
Natural Resources:	No change proposed.	Phasing-out soil-degrading materials and practices (e.g. polluting fertilizers, intensive farming, single-crop farming; old/inefficient irrigation systems; de-forestation).
Energy & industrial policy:	Increasing taxes across industries that use non-renewable energy sources.	No change proposed.
Transport:	Investing £5 billion from the budget (about 0.5% of government spending) every year to develop eco-friendly fuels and low emission vehicles.	Phasing-out all non-electric vehicles and flights that can be completed by train in an hour.
Trade & Finance	Increasing taxes/duties for goods and services from companies/countries that do not meet strict environmental standards.	No change proposed

Which of these two petitions would you rather sign: Petition A (on the left), or Petition B (on the right)?

Petition A | Petition B

On a scale of 0 to 10, how strongly do you support each petition?

Petition A

Do not support at all — Strongly support

4

Petition B

Do not support at all — Strongly support

7

F YouGov Conjoint Diagnostics

For the main conjoint experiment (N = 16048 petitions), we consulted several diagnostics to determine whether a) the various attribute levels were indeed properly randomized as discussed with YouGov, whether b) our results may be affected by question fatigue or order effects due to the long petition texts (of which each respondent read eight), and whether c) our results are robust to the inclusion of control variables and survey-weights.

In Table F1 below, we show that the conjoint randomization worked flawlessly. All attribute levels were shown to respondents at almost identical rates.

Table F1: Frequency of Attribute Levels

Lever	Level 1	Level 2	Level 3	Level 4	Level 5
emotion	5315	5399	5334		
benefit	5321	5346	5381		
energy	3220	3219	3161	3160	3288
buildings	3283	3234	3191	3132	3208
transport	3195	3261	3199	3196	3197
resources	3203	3235	3191	3220	3199
trade	3217	3098	3294	3224	3215

In Figure F5 below, we show that there were no significant order effects. To do this, we ran two tests: in the lefthand plot, we show that - across 24 coefficients - there were no significant differences in AMCEs between the first and the last petition (across all 2006 respondents). In the righthand plot, we further show that there were no significant differences in AMCEs between the left-sided petitions and the right-sided petitions (as all respondents were always shown two petitions at the same time). There, only one of 24 comparisons is marginally statistically significant (which is within the expected range for false positives).

In Figure F6, we show that our conjoint results are robust to the inclusion of a battery of control variables including the respondents gender, education, region, social class, income, religion, and vote choice (lefthand plot), and to the inclusion of a set of survey weights provided by YouGov to ensure representativeness across the UK electorate (righthand plot). Across both alternative specifications, the results remain comparable to the raw AMCEs reported in Figure F7 below, and to each other.

Figure F5: Conjoint Order Effects Across (Left) and Within (Right) Survey Pages

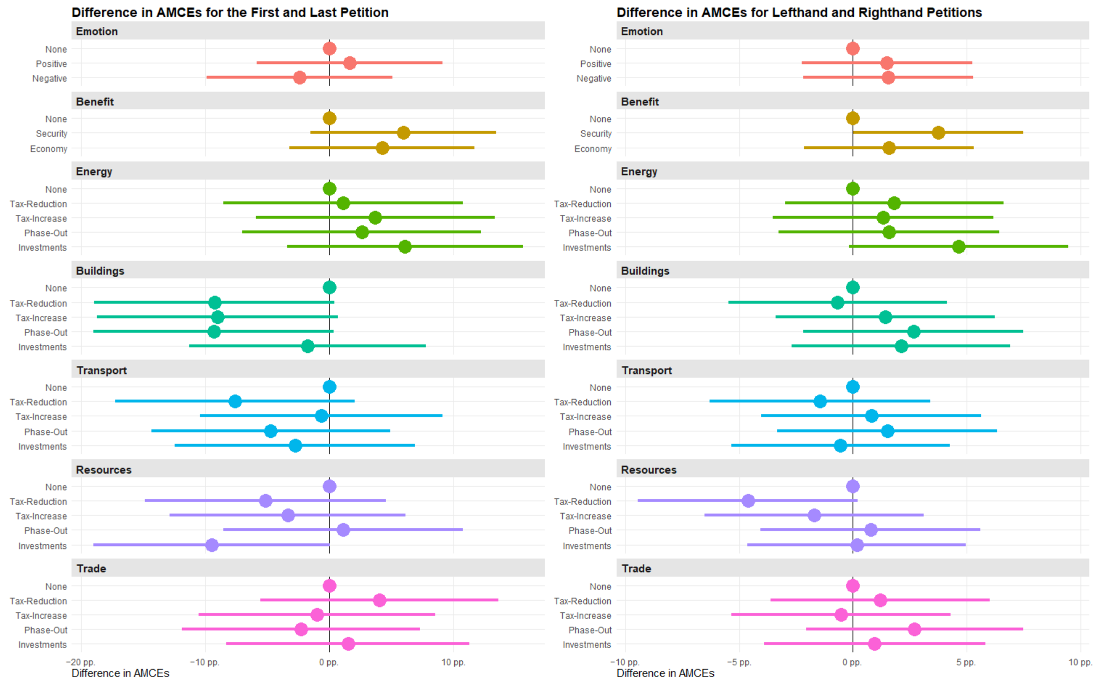


Figure F6: Conjoint Results with Controls (Left) and Survey-Weights (Right)

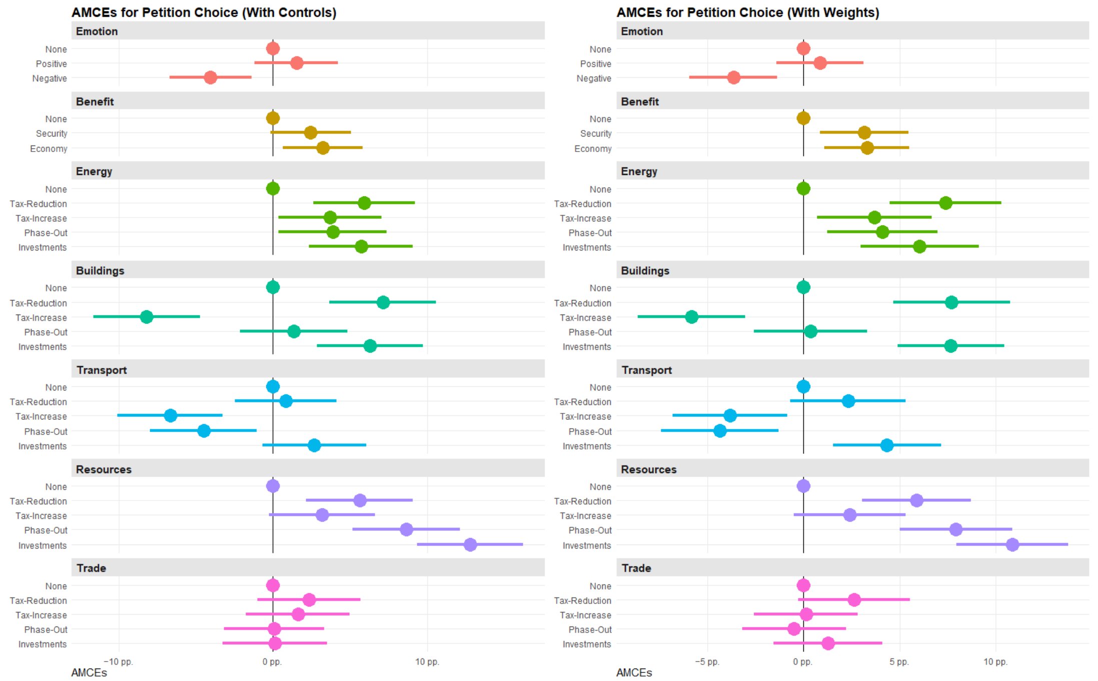
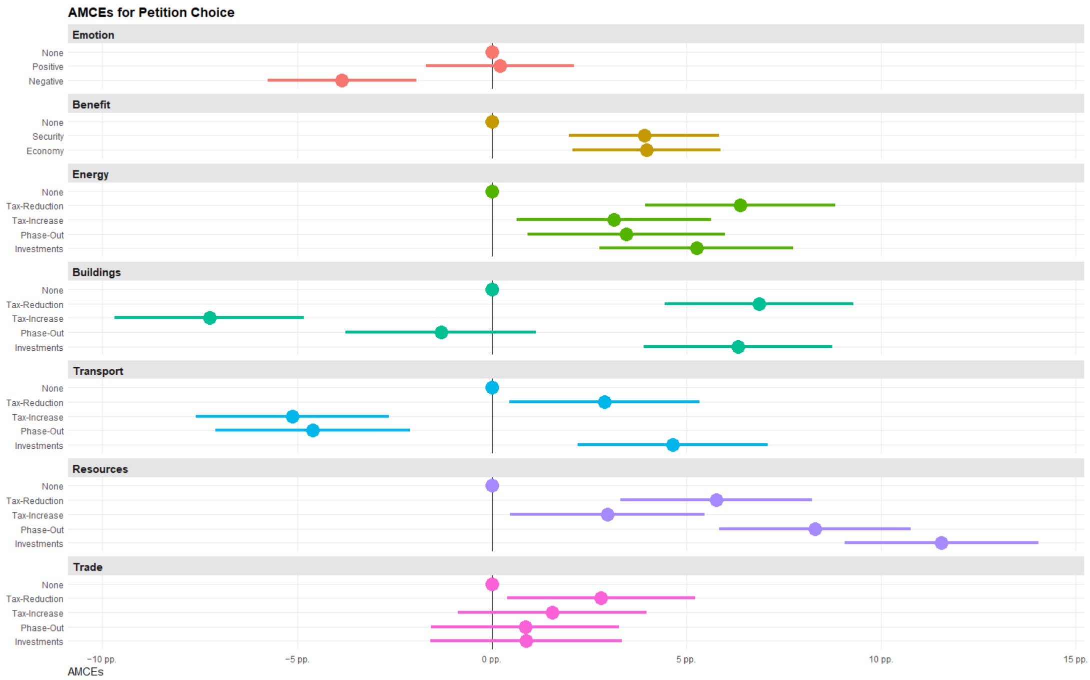
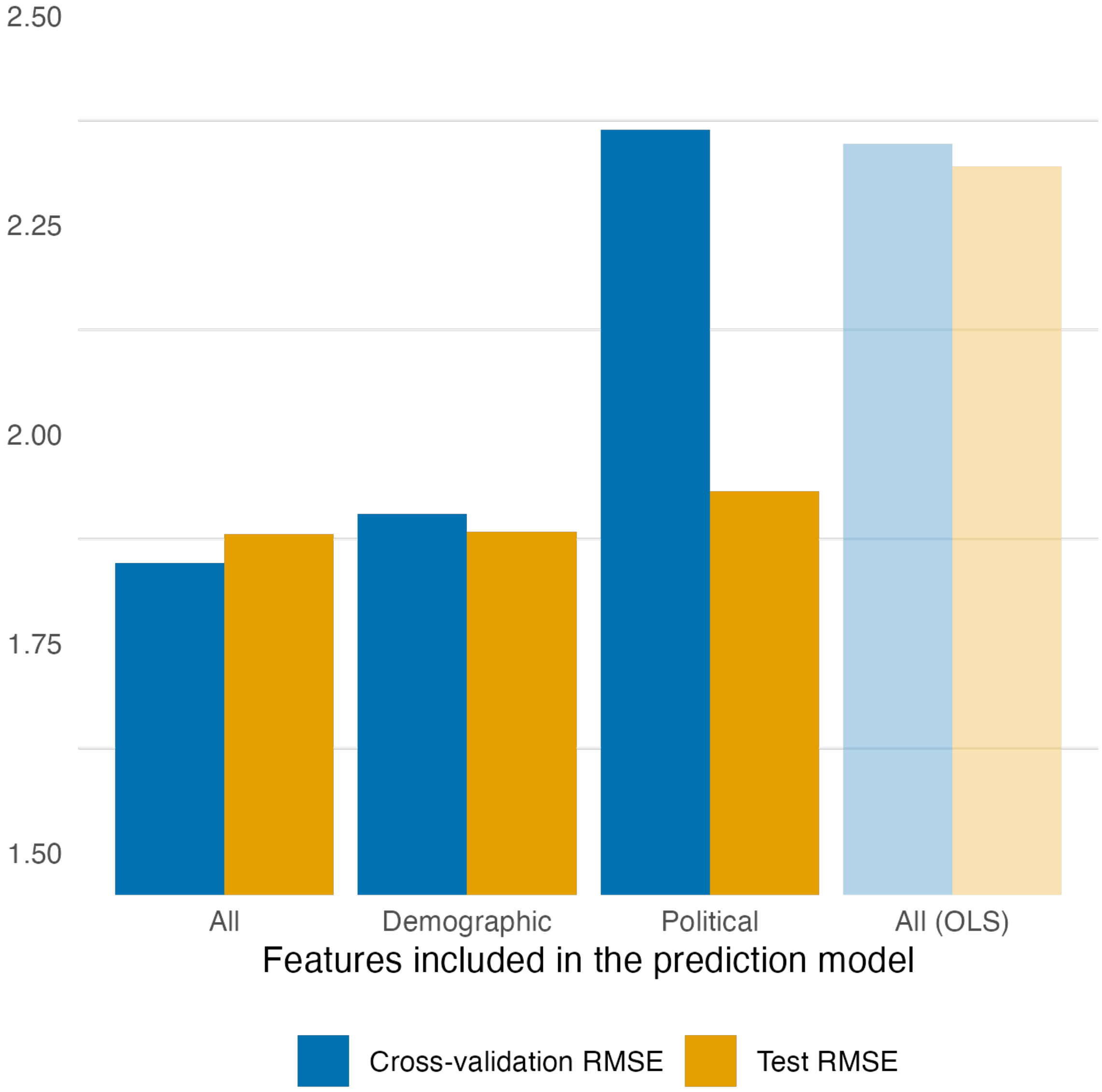


Figure F7: Raw AMCEs



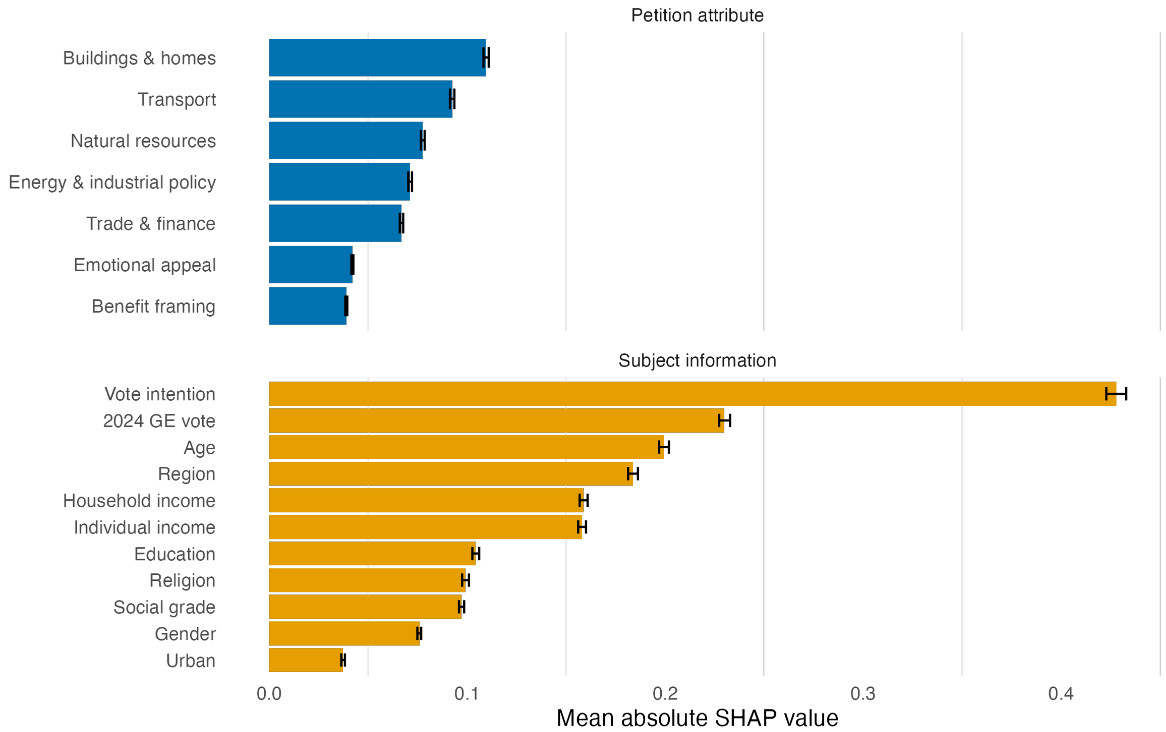
G Nano-Targeting Model Diagnostics

Figure G8: Prediction Error for **XGBoost** Models Trained on Different Subsets of Features



Note: All models include all petition attribute variables. An OLS model fit on all variables is included for comparison. The y-axis has been truncated for clarity.

Figure G9: Marginal Contribution of Each Petition Attribute and Subject Covariate to Final Model Predictions



Notes: Estimates are shown with 95% confidence intervals.

Figure G10: Distribution of Difference in Predicted Ratings Under Treatment Assignments

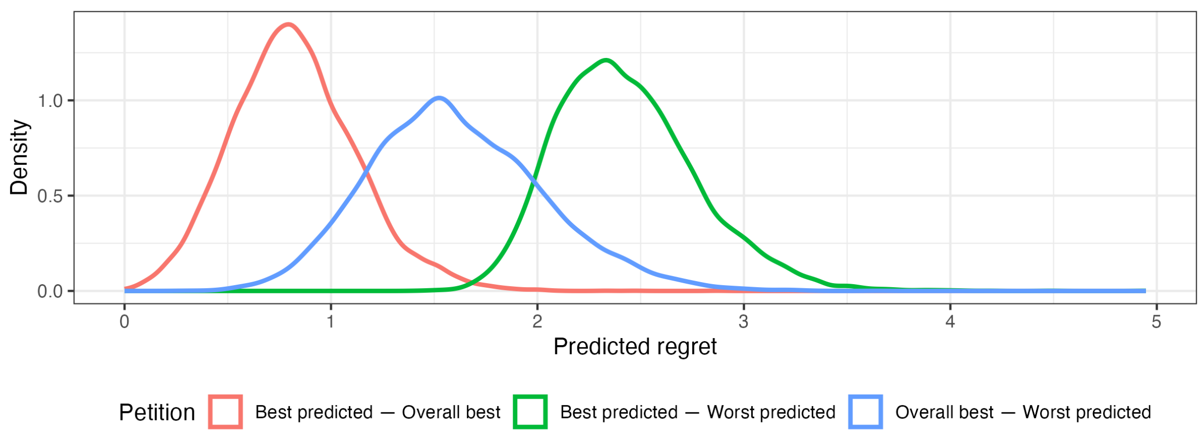


Figure G11: Number of times nano-targeted petitions contained attribute levels present in the overall best message petition

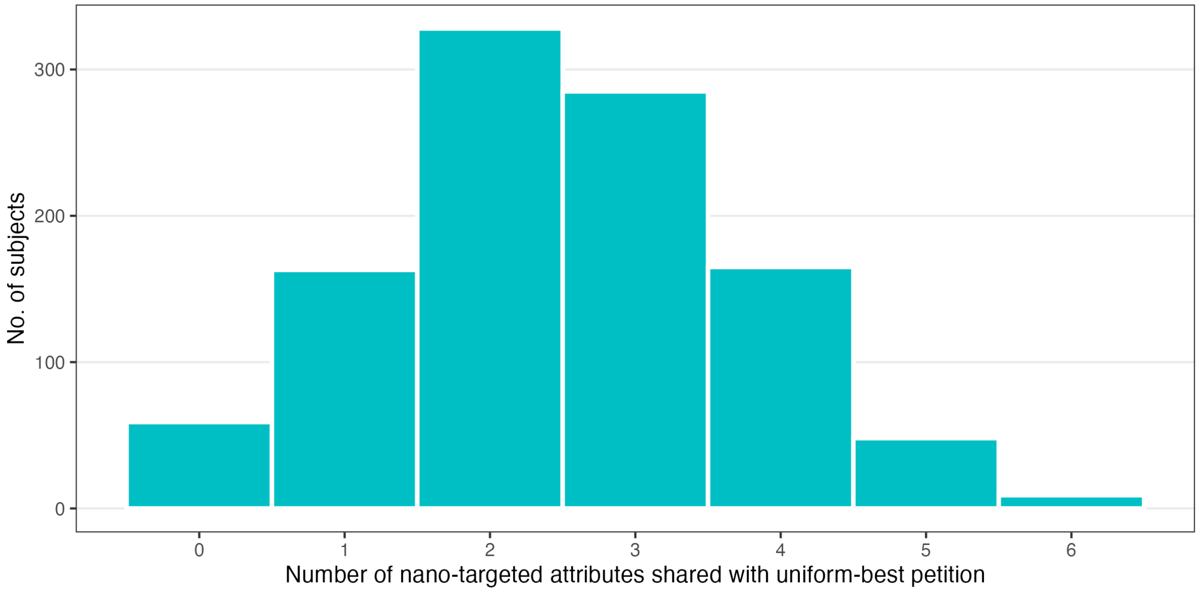
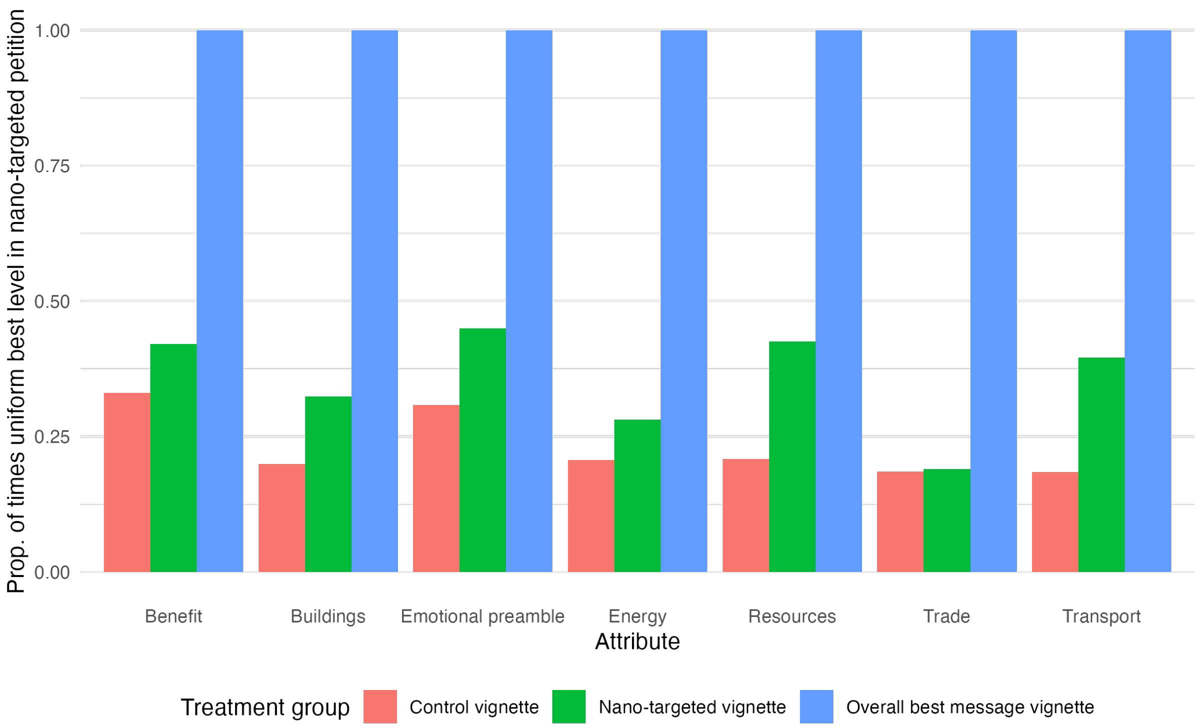


Figure G12: Proportion of times individually-assigned petitions contained attribute levels also present in the overall best message petition



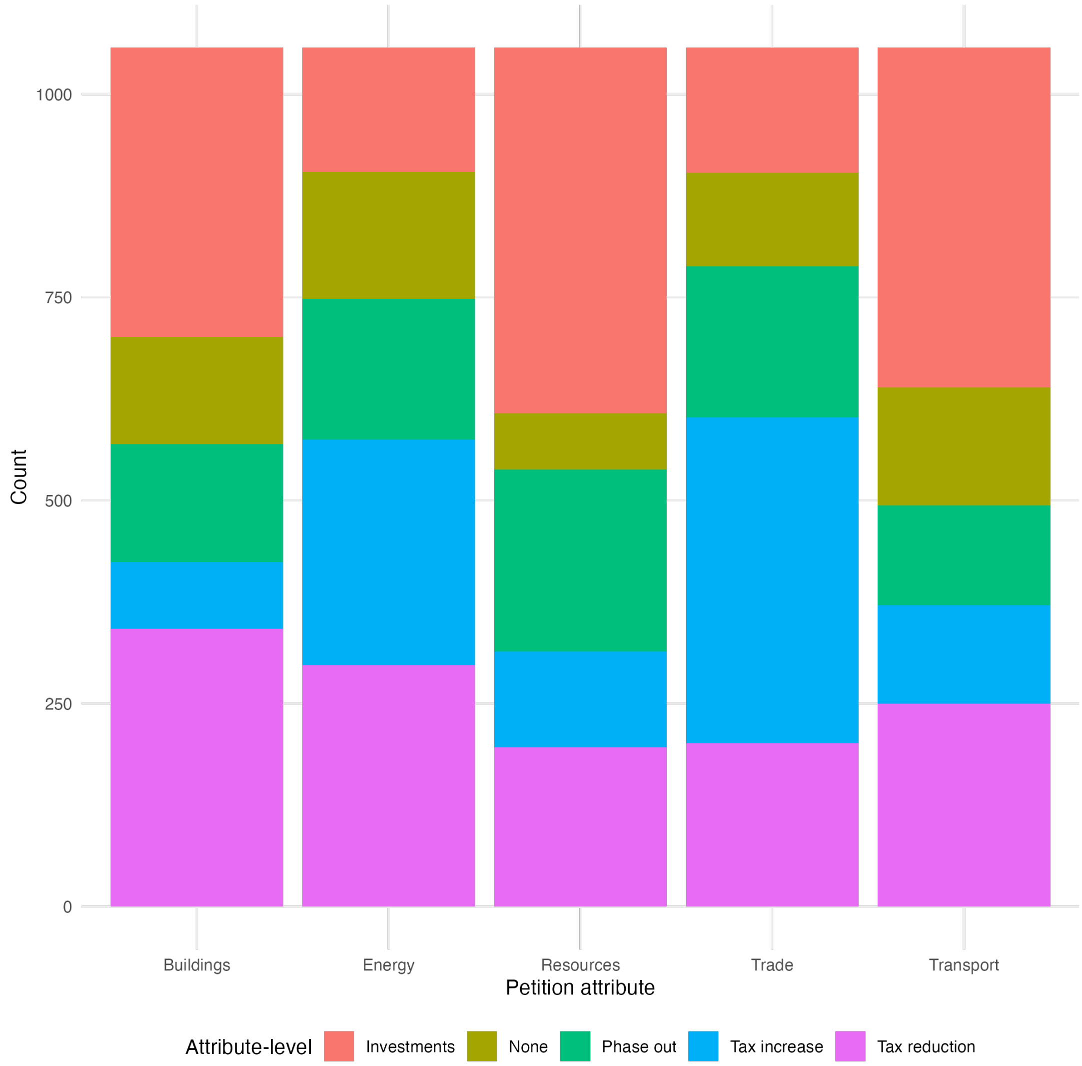


Figure G13: Distribution of assigned main attribute-levels to *nano-targeted* subjects in the Opinium sample

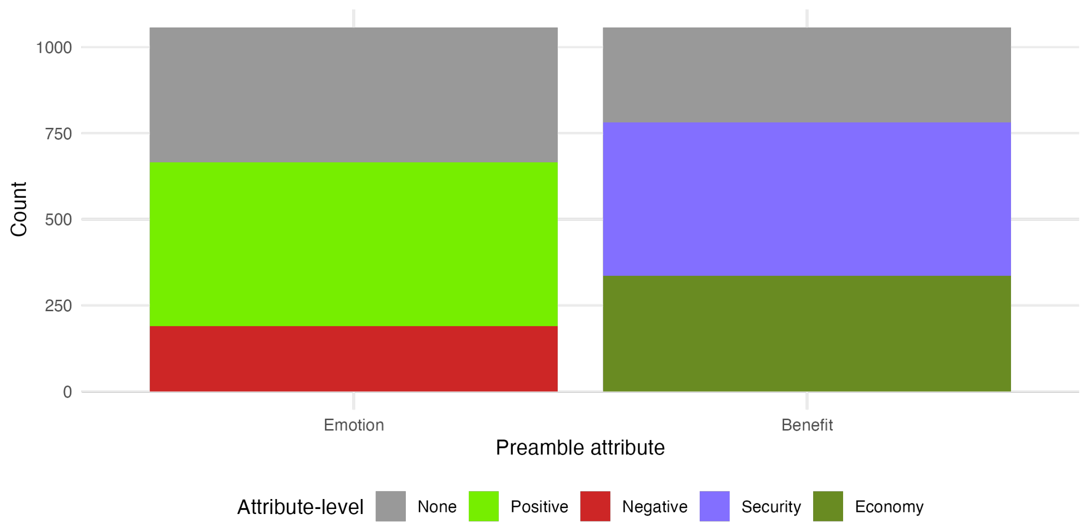


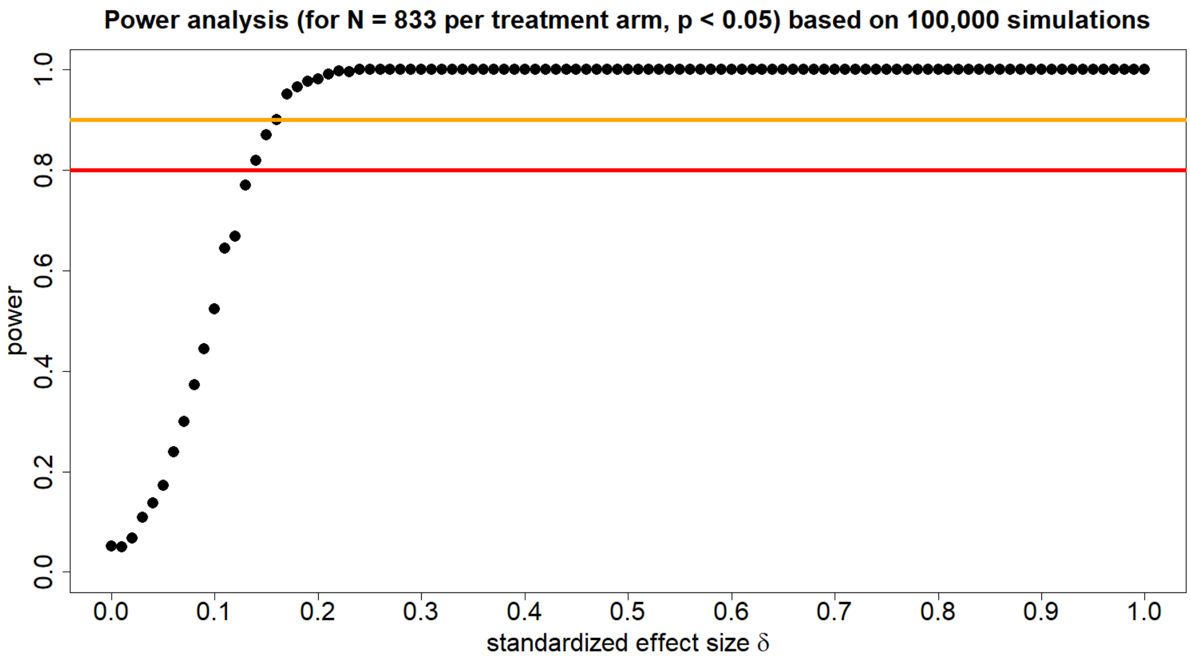
Figure G14: Distribution of assigned preamble attribute-levels to *nano-targeted* subjects in the Opinium sample

H Opinium Power Analysis

Before fielding our main experiment with Opinium, we conducted a series of simulation-based power analyses for our intended sample size of 2500 respondents (i.e., 833 respondents for each of the three treatment arms). Specifically, we tested across 100,000 simulations (with 166,600,000 simulated data points), how much statistical power our setup would grant us to detect one hundred different effect sizes between 0.01 standard deviations and 1 standard deviation. In the resulting power plot below, the horizontal red line marks the conventional threshold for 80% statistical power, while the horizontal yellow line marks the threshold for 90% statistical power. Each black dot represents the mean power of 1000 simulations for a given effect size.

At 833 respondents per treatment arm, with a cutoff for statistical significance at $p < 0.05$, and with heteroskedasticity-robust standard errors in our regressions, we can expect >90% statistical power to detect very small effect sizes above 0.15 sd, and >99% statistical power to detect small effect sizes above 0.25 sd.

Figure H15: Power Simulations



I Opinium Experiment Excerpts

To illustrate how the final experiment appeared to participants, we include several screenshots from the Opinium survey in this Appendix. Figure I16 shows an example petition and the first outcome question (the binary choice). Figure I17 shows an example petition and the rating outcome. Figure I18 shows an example petition and the follow-up petition evaluation questions. Figure I19 shows the three climate-policy questions shown before and after the petition. Figure I20 shows the final question of the survey, which we used for our behavioral outcome. If respondents selected "Yes", they were given instructions and referred to an official website of the UK parliament, as shown in Figure I21.

Figure I16: Petition and Binary Choice

Please read the below petition carefully and answer the questions below:

If we do not implement policies to meet the 1.5C global warming target and reach net zero; we will be on the brink of extinction. Our proposed decarbonization solutions will make us more secure: they will reduce the UK's energy dependence on other countries; which will protect us from global shocks in fuel prices and energy supply. We call on the UK government to meet the 1.5C target by ...

Energy & Industrial Policy

- Increasing taxes across industries that use non-renewable energy sources

Buildings & Homes

- Reducing taxes for owners of buildings and homes that refurbish their properties to meet energy efficiency level A

Transport

- Reducing taxes to households and businesses that use electric vehicles

Natural resources

- No change proposed

Trade & Finance

- Reducing taxes/duties on all businesses and individuals who trade with companies/countries that meet strict environmental standards

Would you sign this petition?

Please select one answer

Yes

No

← Next →

Figure I17: Petition and Rating Outcome

Here is the same petition again. Please read the below petition carefully and answer the questions below:

If we do not implement policies to meet the 1.5C global warming target and reach net zero; we will be on the brink of extinction. Our proposed decarbonization solutions will make us more secure: they will reduce the UK's energy dependence on other countries; which will protect us from global shocks in fuel prices and energy supply. We call on the UK government to meet the 1.5C target by ...

Energy & Industrial Policy

- Investing £5 billion from the budget (about 0.5% of government spending) every year to develop green technologies for industry

Buildings & Homes

- Phasing-out all buildings and homes not reaching energy efficiency level A by removing building permissions

Transport

- Reducing taxes to households and businesses that use electric vehicles

Natural resources

- No change proposed

Trade & Finance

- Investing £5 billion from the budget (about 0.5% of government spending) every year to support trading partners in meeting strict environmental standards

On a scale of 0 to 10, where 10 is strongly support and 0 is strongly oppose, how strongly do you support this petition?

0 1 2 3 4 5 6 7 8 9 10

← Next →

Figure I18: Petition and Evaluation Outcomes

Op
Opinium

Here is the same petition again. Please read the below petition carefully and answer the questions below:

If we do not implement policies to meet the 1.5C global warming target and reach net zero; we will be on the brink of extinction. Our proposed decarbonization solutions will make us more secure: they will reduce the UK's energy dependence on other countries; which will protect us from global shocks in fuel prices and energy supply. We call on the UK government to meet the 1.5C target by ...

Energy & Industrial Policy

- Investing £5 billion from the budget (about 0.5% of government spending) every year to develop green technologies for industry

Buildings & Homes

- Phasing-out all buildings and homes not reaching energy efficiency level A by removing building permissions

Transport

- Reducing taxes to households and businesses that use electric vehicles

Natural resources

- No change proposed

Trade & Finance

- Investing £5 billion from the budget (about 0.5% of government spending) every year to support trading partners in meeting strict environmental standards

On a scale from 0 to 10, where 10 means strongly agree and 0 means strongly disagree, please indicate how much you agree or disagree with each of the following statements?

I would be willing to pay an additional 1% in taxes to help finance the policies outlined in this petition

0 1 2 3 4 5 6 7 8 9 10

The policies outlined in the petition address climate change in the best possible way

0 1 2 3 4 5 6 7 8 9 10

I would vote for a party that implemented the policies outlined in the petition

0 1 2 3 4 5 6 7 8 9 10

Next

Figure I19: Climate-Policy Questions

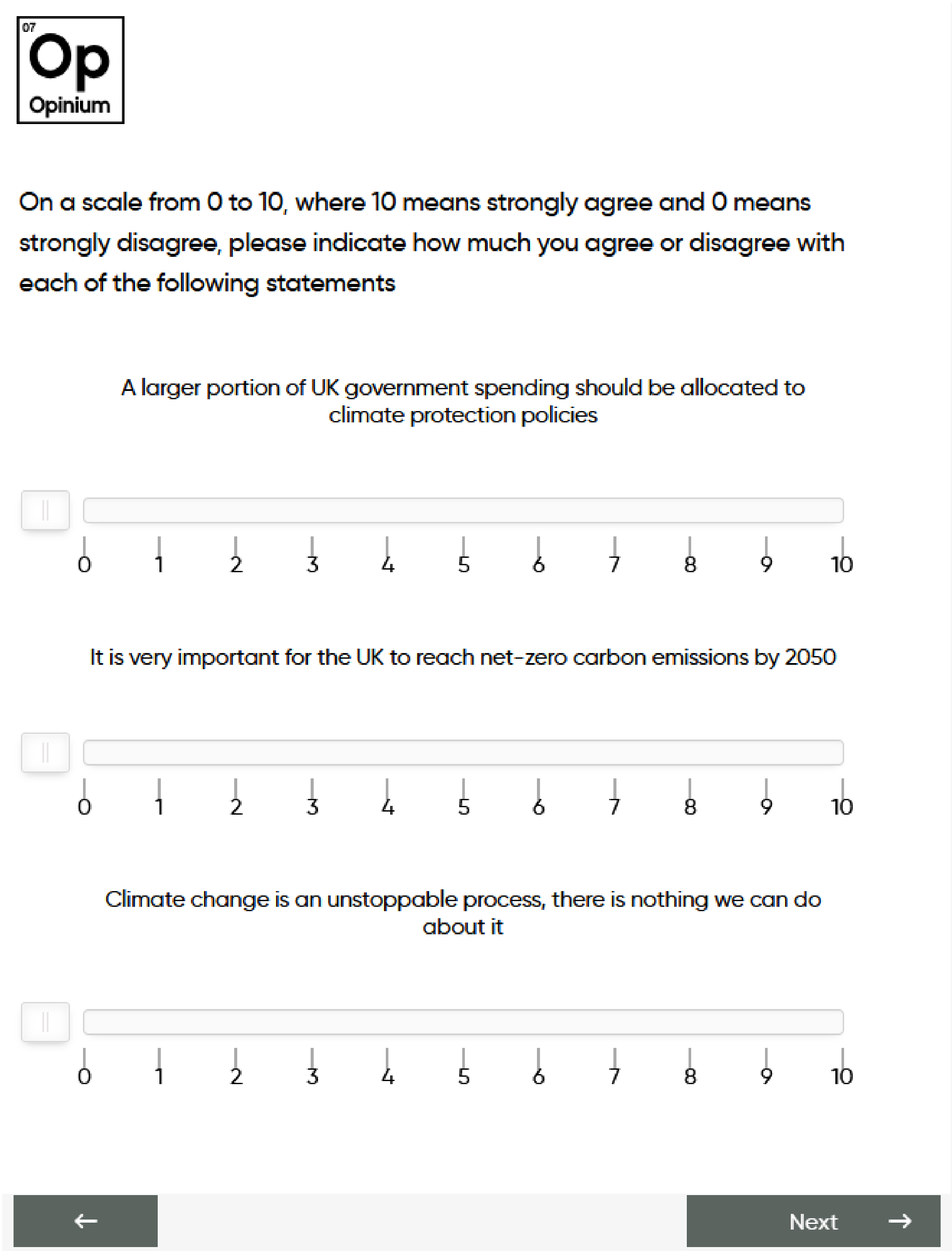


Figure I20: Behavioral Outcome Question

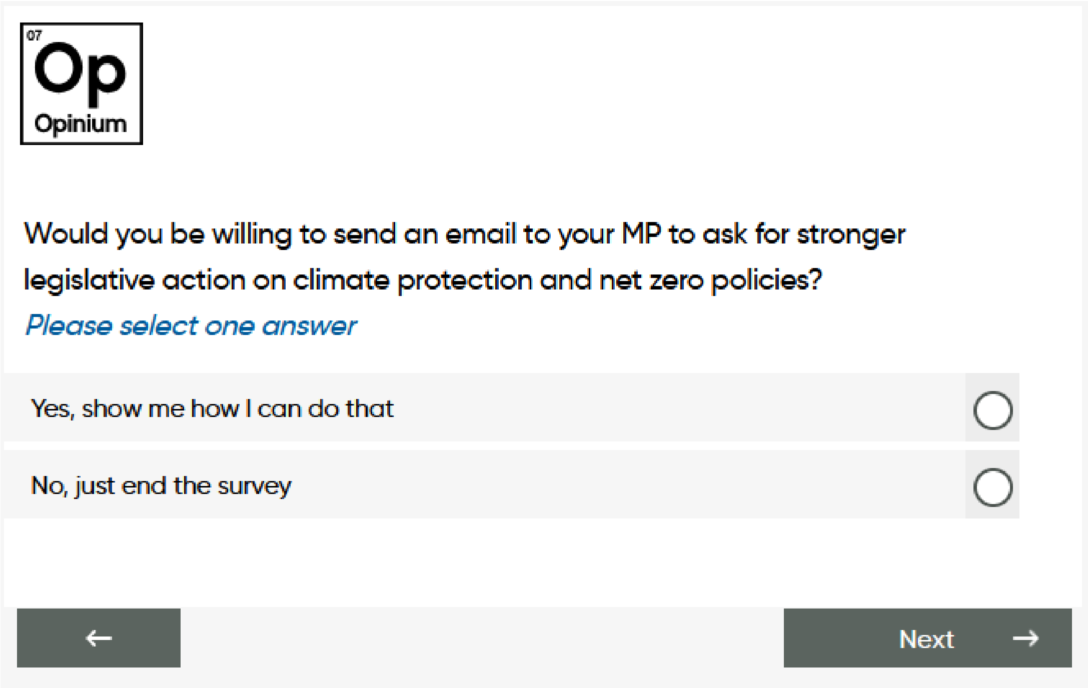
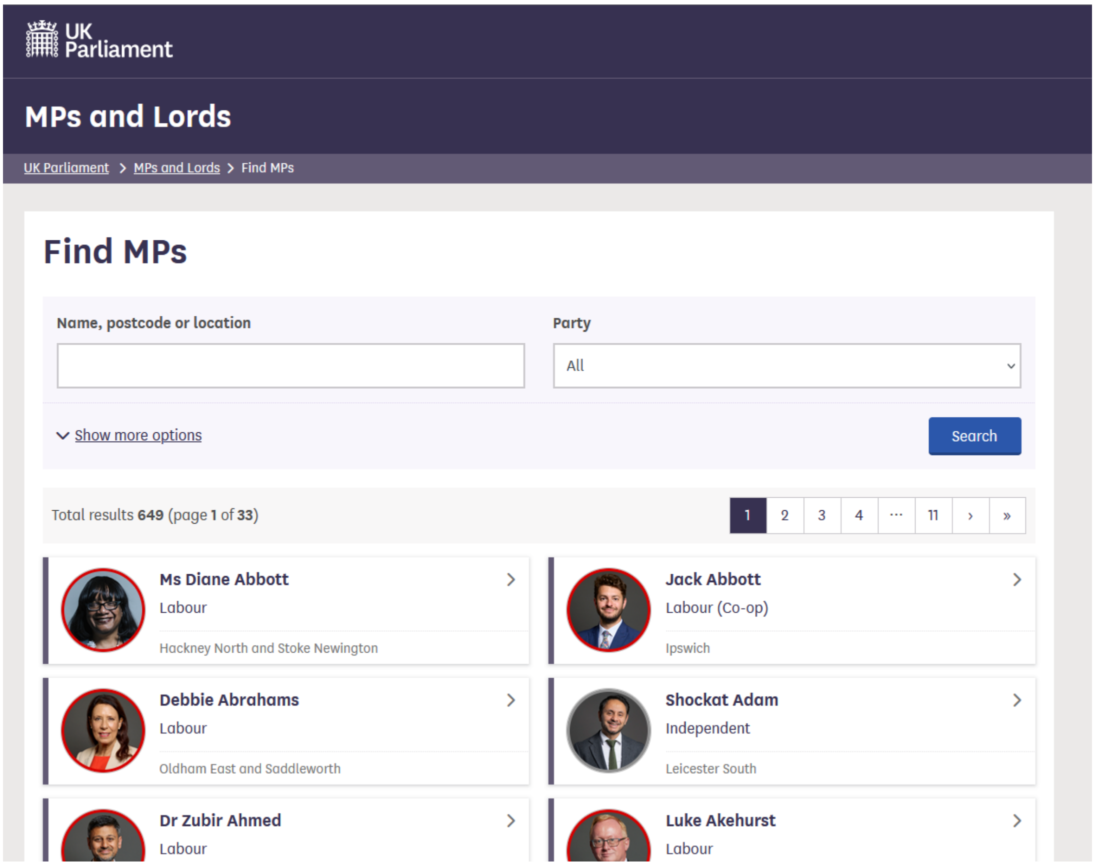


Figure I21: Find Your MP Website (Behavioral Outcome)



J Opinium Regression Tables

Table J2: ATEs for Binary Outcomes

	Dependent variable:			
	Sign Petition		Send E-Mail	
	(1)	(2)	(3)	(4)
Nanotargeted	0.051^{**}	0.048^{**}	-0.021	-0.020
	(0.022)	(0.021)	(0.017)	(0.017)
Overall Best	0.095^{***}	0.095^{***}	0.017	0.021
	(0.022)	(0.022)	(0.018)	(0.018)
Age		-0.004^{***}		-0.001^{*}
		(0.001)		(0.0005)
Male		-0.024		0.006
		(0.018)		(0.015)
Suburban Area		0.044^{*}		-0.023
		(0.023)		(0.019)
Urban Area		0.114^{***}		0.030
		(0.026)		(0.022)
Social Grade B		-0.058^{*}		0.019
		(0.034)		(0.030)
Social Grade C1		-0.022		-0.027
		(0.035)		(0.030)
Social Grade C2		-0.066		0.009
		(0.040)		(0.034)
Social Grade D		-0.041		-0.051
		(0.044)		(0.035)
Social Grade E		-0.036		0.026
		(0.042)		(0.036)
Low Education		-0.063^{**}		-0.074^{***}
		(0.031)		(0.024)
Mid Education		-0.053^{**}		-0.055^{***}
		(0.021)		(0.017)
Constant	0.425^{***}	0.650^{***}	0.207^{***}	0.285^{***}
	(0.015)	(0.051)	(0.013)	(0.042)
Observations	3,113	3,108	3,113	3,108
R^2	0.006	0.038	0.001	0.018
Adjusted R^2	0.005	0.034	0.001	0.014
Note:		*p<0.1; **p<0.05; ***p<0.01		

Table J3: ATEs for Factor Outcomes

	Dependent variable:			
	Climate Attitude Factor		Petition Evaluation Factor	
	(1)	(2)	(3)	(4)
Nanotargeted	−0.012	−0.013	0.060	0.055
	(0.027)	(0.027)	(0.043)	(0.042)
Overall Best	0.046*	0.053*	0.188***	0.194***
	(0.028)	(0.028)	(0.042)	(0.042)
Age		−0.002**		−0.010***
		(0.001)		(0.001)
Male		0.033		−0.085**
		(0.023)		(0.035)
Suburban Area		0.017		0.079*
		(0.029)		(0.046)
Urban Area		0.017		0.178***
		(0.033)		(0.051)
Social Grade B		0.001		−0.059
		(0.040)		(0.066)
Social Grade C1		0.021		−0.129*
		(0.041)		(0.068)
Social Grade C2		−0.067		−0.186**
		(0.053)		(0.079)
Social Grade D		−0.003		−0.146*
		(0.058)		(0.084)
Social Grade E		0.008		−0.067
		(0.049)		(0.081)
Low Education		−0.043		−0.174***
		(0.045)		(0.058)
Mid Education		−0.030		−0.145***
		(0.026)		(0.040)
Constant	−0.011	0.062	−0.081***	0.596***
	(0.019)	(0.060)	(0.030)	(0.095)
Observations	3,113	3,108	3,113	3,108
R^2	0.002	0.007	0.006	0.057
Adjusted R^2	0.001	0.003	0.006	0.053
Note:				*p<0.1; **p<0.05; ***p<0.01

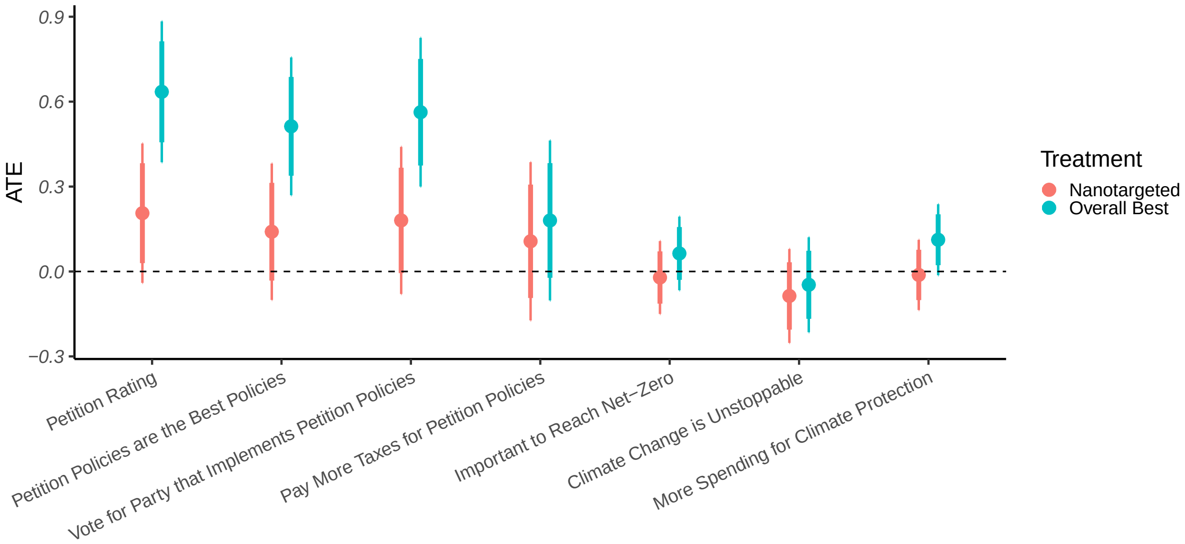
Table J4: Interaction Models

	Sign Petition		
	(1)	(2)	(3)
Nanotargeted	0.045	0.048*	0.044*
	(0.033)	(0.025)	(0.022)
Overall Best	0.112***	0.080***	0.103***
	(0.034)	(0.025)	(0.023)
Switcher	−0.044		
	(0.037)		
Nano*Switcher	−0.018		
	(0.052)		
Best*Switcher	−0.070		
	(0.052)		
Moderate		−0.102***	
		(0.036)	
Nano*Moderate		0.026	
		(0.049)	
Best*Moderate		0.061	
		(0.051)	
Pro-Net-Zero			0.336***
			(0.047)
Nano*Pro			0.113*
			(0.068)
Best*Pro			−0.036
			(0.070)
Constant	0.449***	0.450***	0.386***
	(0.023)	(0.018)	(0.016)
Observations	2,262	3,113	3,113
R^2	0.010	0.011	0.057
Adjusted R^2	0.008	0.009	0.056
Note:	*p<0.1; **p<0.05; ***p<0.01		

K Opinium Individual Outcomes

In this Appendix we present the coefficients for all continuous outcomes in the main experiment. The four petition-related outcomes have been summarized as "Petition Evaluations" through a factor analysis in the main manuscript. Likewise, the three climate-attitudinal outcomes have also been summarized as "Climate Attitudes" through a factor analysis in the main manuscript.

Figure K22: Coefficients for Individual Outcomes

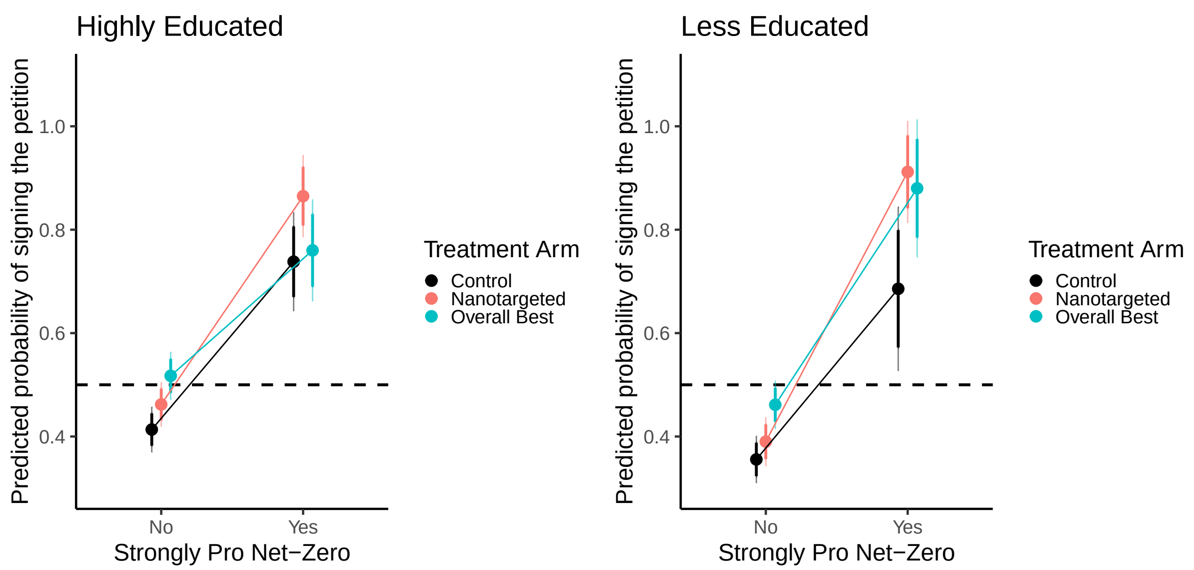


L Opinium Educational Heterogeneity

In this Appendix we present the predicted probabilities of signing a petition for subgroups of climate progressives and non-progressives, further split into sub-samples of highly educated and less educated respondents. As can be seen, the nano-targeted petitions are the most popular among both highly and less educated climate progressives. For highly educated climate progressives, the probability of signing a nano-targeted petition is roughly 12pp higher than the probability of signing a control petition (while the probability of signing the best-overall petition is roughly 2pp higher than signing a control petition). For less educated climate progressives, the probability of signing a nano-targeted petition is roughly 22pp higher than the probability of signing a control petition (while the probability of signing the best-overall petition is roughly 19pp higher than signing a control petition).

The strong effects of nano-targeting among climate progressives are not merely driven by their higher average education or superior cognitive abilities. In fact, the effects are even stronger among less educated climate progressives (though the effects of the overall-best petition are likewise also higher among less educated climate progressives). This suggests that other factors are behind the increased effectiveness of nano-targeting among climate progressives, such as the higher salience of the climate topic for them, or a stronger conviction to bring about change.

Figure L23: Heterogeneity by Climate Attitude Strength and Education



M Opinium Regret Analysis

In this Appendix we present the results of comparing the realized behavior of Opinium subjects against our predicted results, given the machine learning model trained in YouGov data.

Figure M24: Estimated Effects of Treatment, Limiting Data to Subjects Whose Predicted Behavior is Within a Given Bound

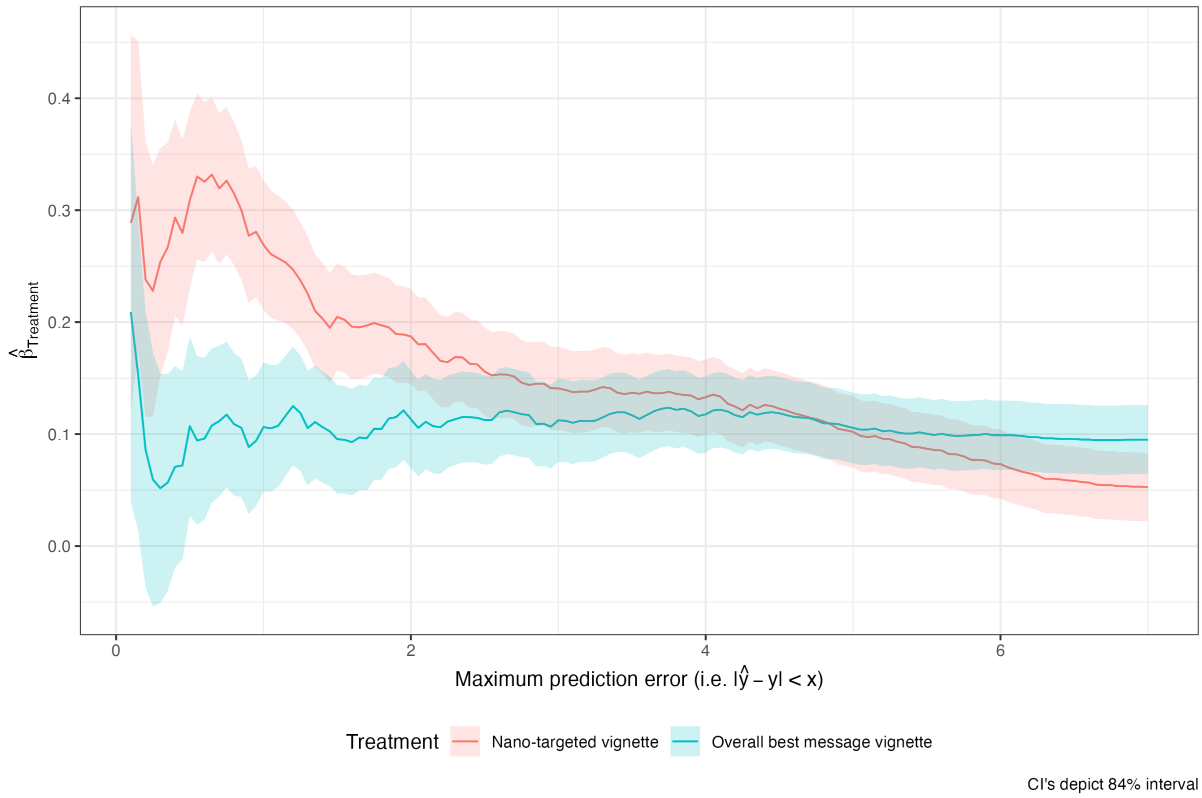


Table M5: Summary of Prediction Errors Comparing ML Model Predictions to Observed Measures in the Opinium Sample

Treatment group	Average prediction error	t	RMSE
Control vignette	0.239**	2.74	2.82
Nano-targeted vignette	-0.764***	-9.02	2.86
Uniform best message vignette	0.495***	6.08	2.64

Figure M25: Distribution of prediction errors by treatment group and *education* level

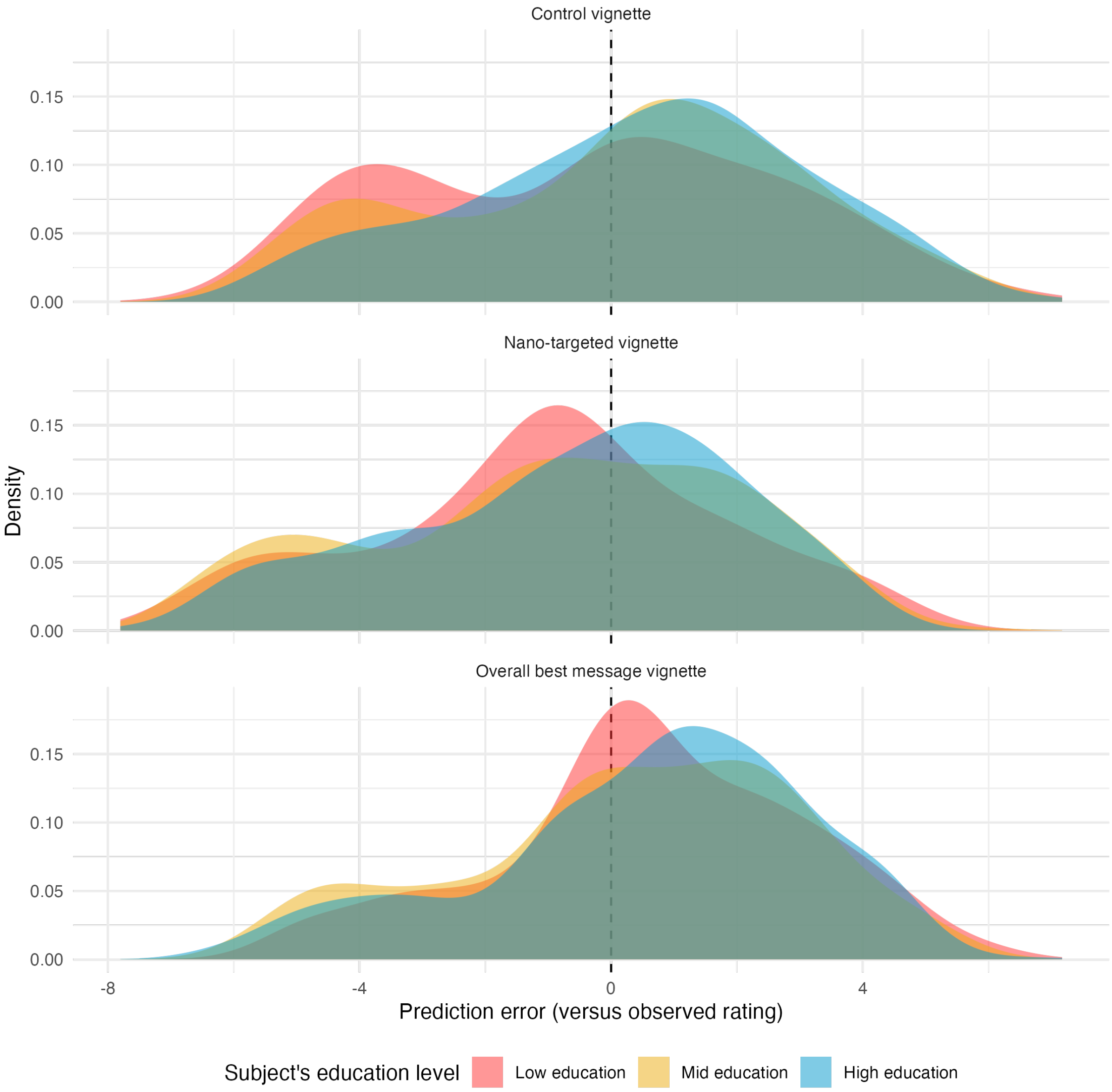


Figure M26: Age SHAP Values Predicting *Opinium Prediction Errors* Using Covariates

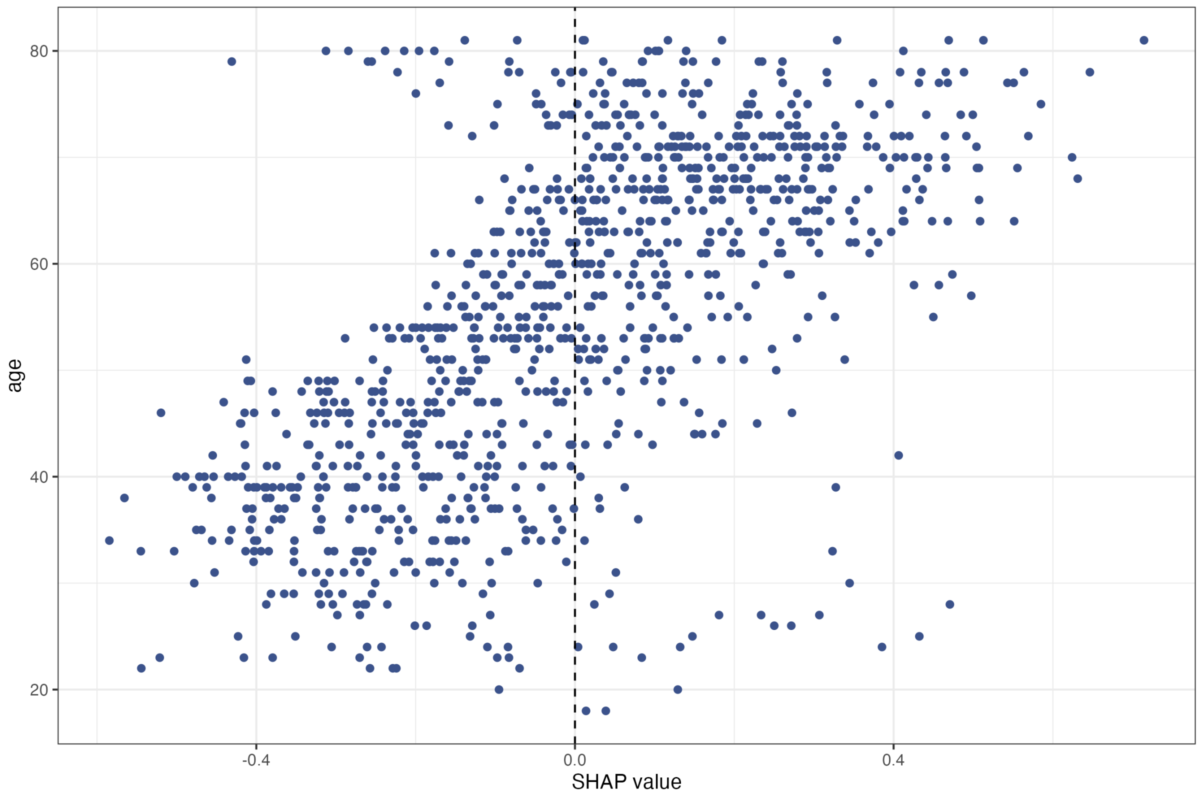


Figure M27: Education SHAP Values Predicting *Opinium Prediction Errors* Using Covariates

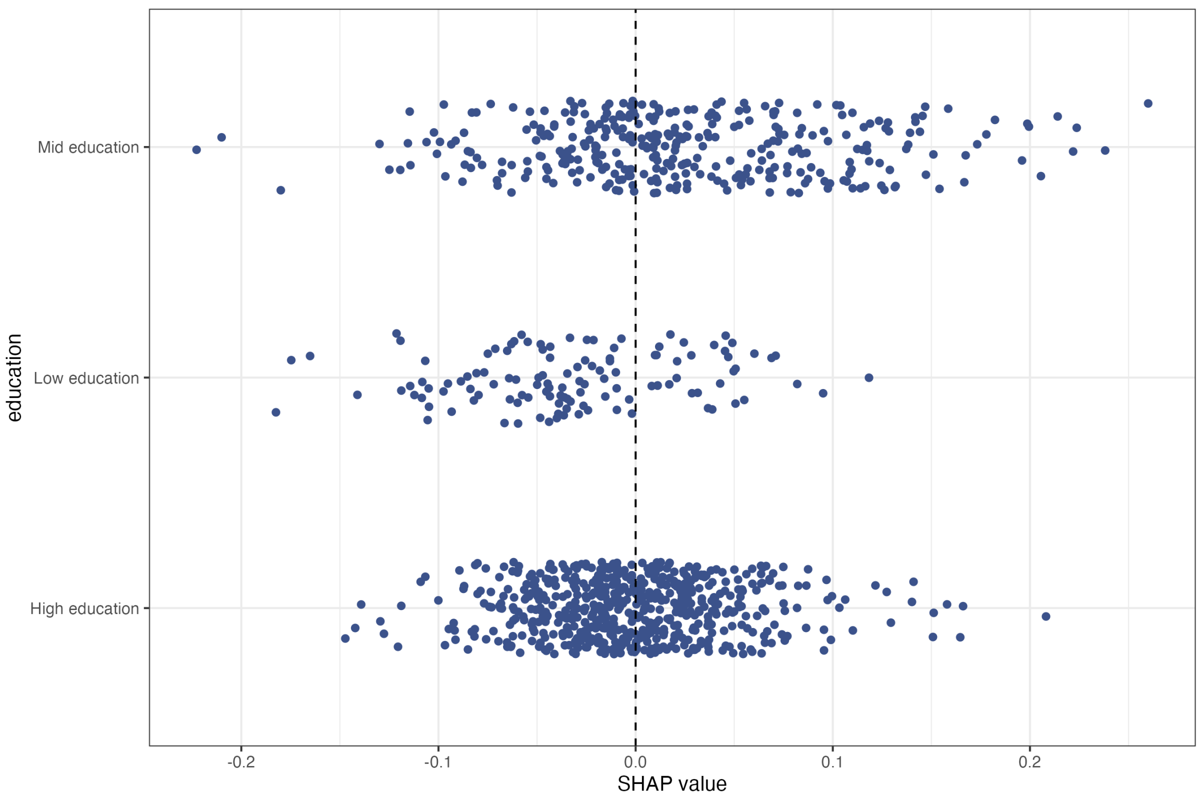


Figure M28: Gender SHAP Values Predicting *Opinium Prediction Errors* Using Covariates

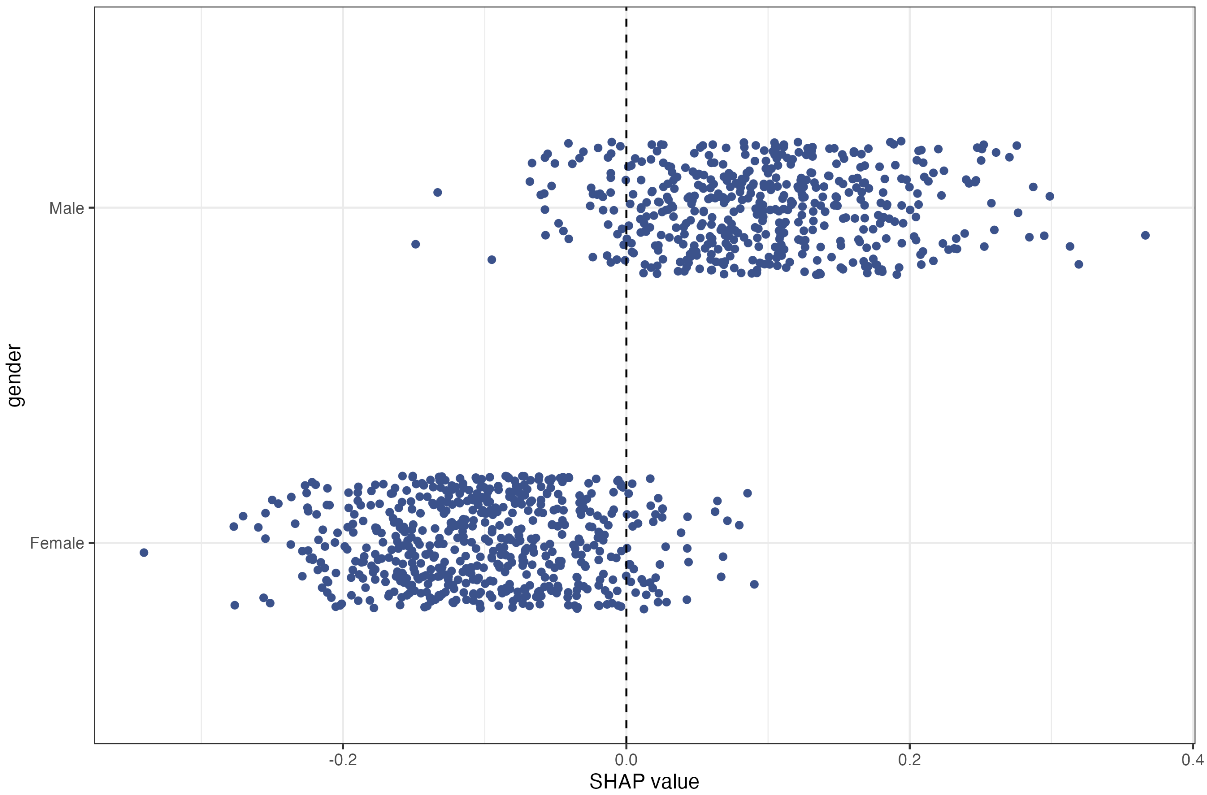


Figure M29: Household Income SHAP Values Predicting *Opinium Prediction Errors* Using Covariates

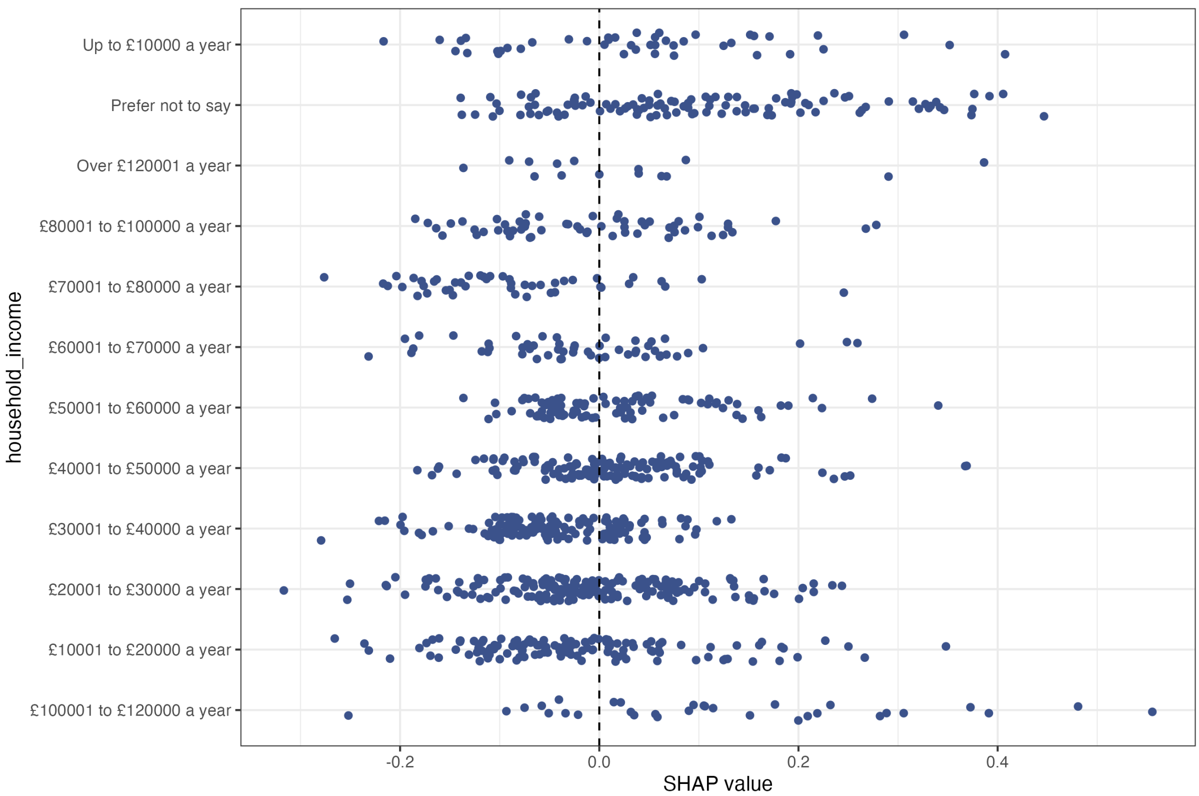


Figure M30: Personal Income SHAP Values Predicting *Opinium Prediction Errors* Using Covariates

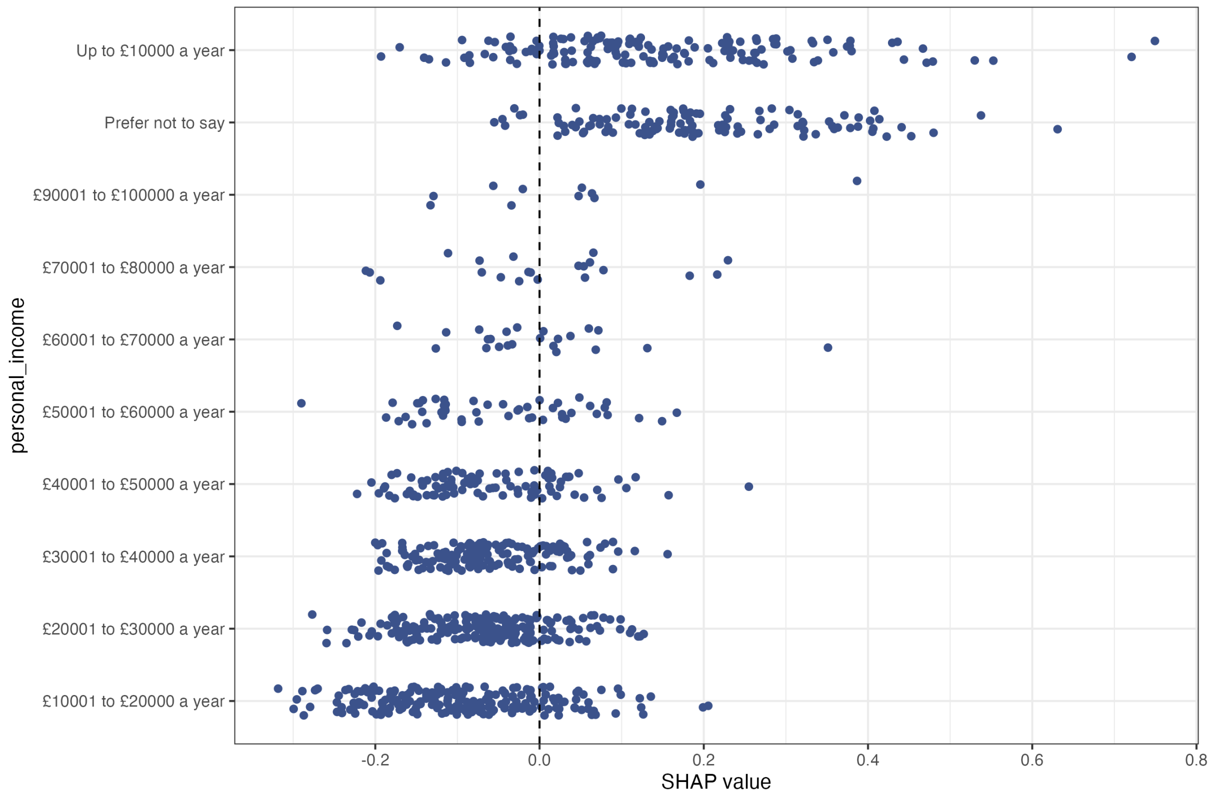


Figure M31: Region SHAP Values Predicting *Opinium Prediction Errors* Using Covariates

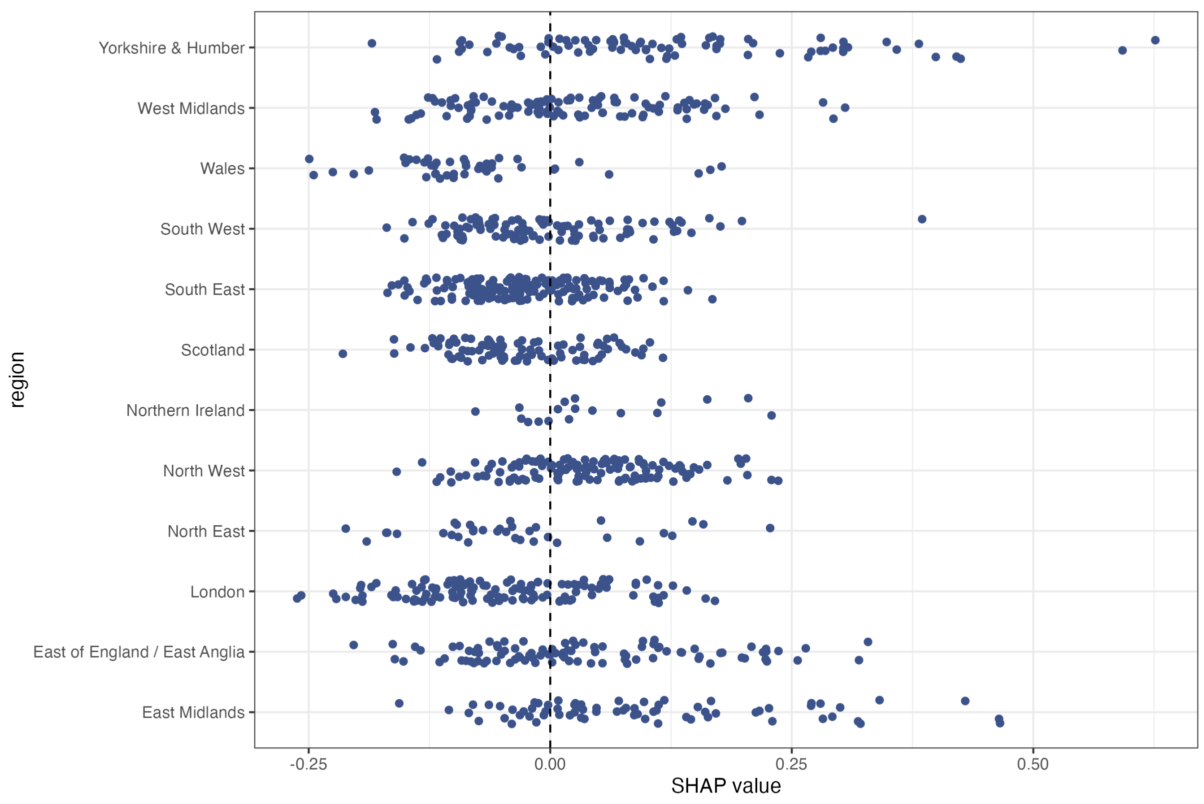


Figure M32: Religion SHAP Values Predicting *Opinium Prediction Errors* Using Covariates

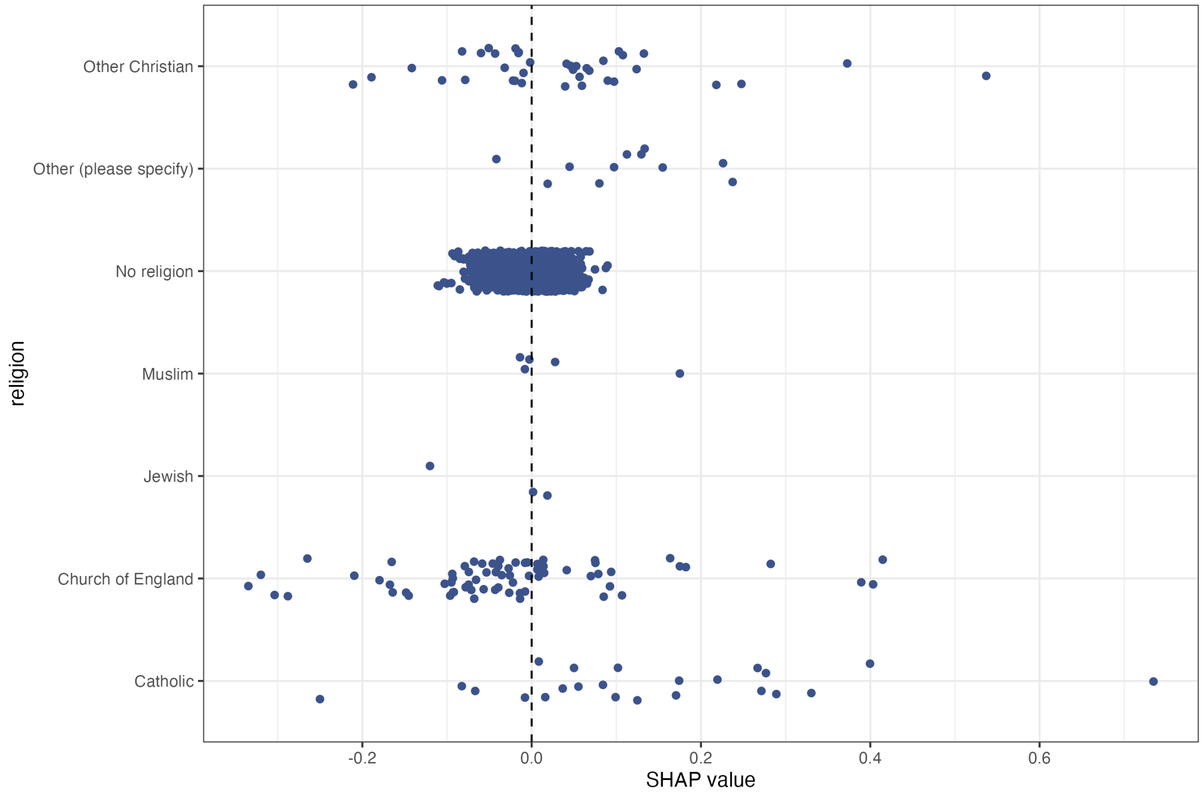


Figure M33: Values Predicting *Opinium Prediction Errors* Using Covariates

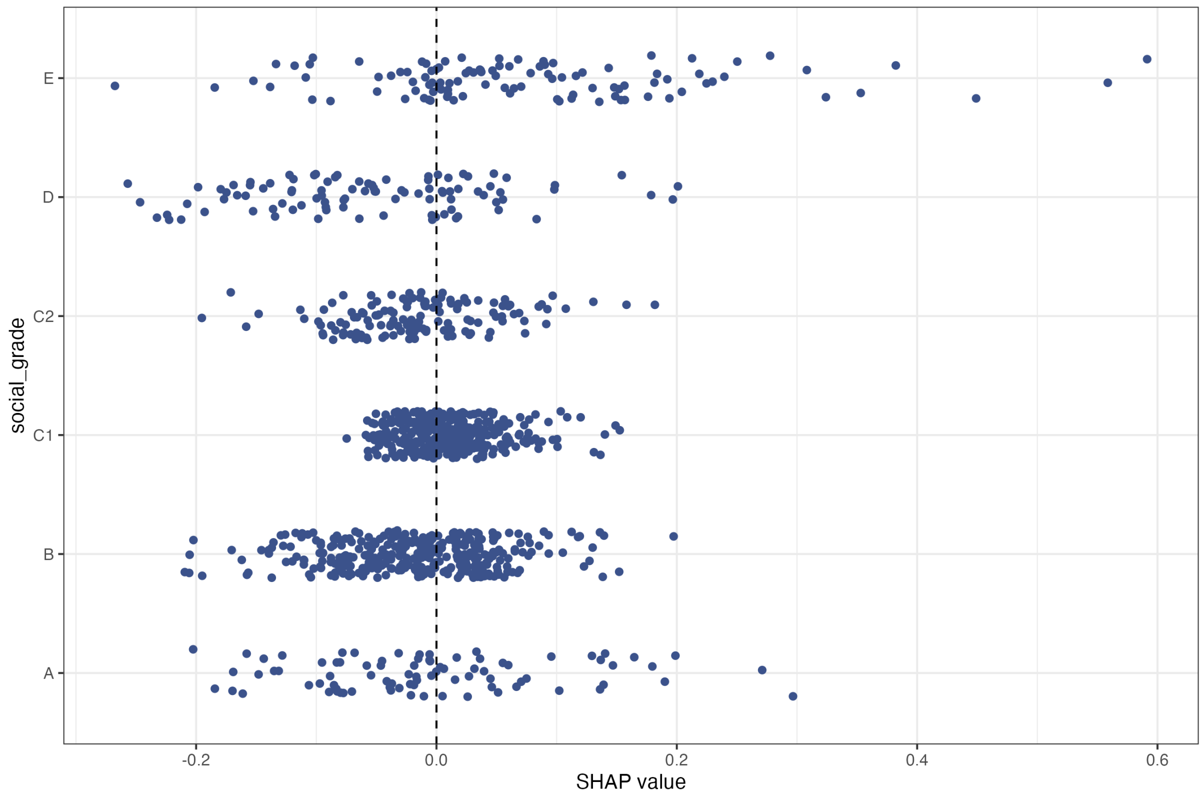


Figure M34: Urban/Rural SHAP Values Predicting *Opinium Prediction Errors* Using Covariates

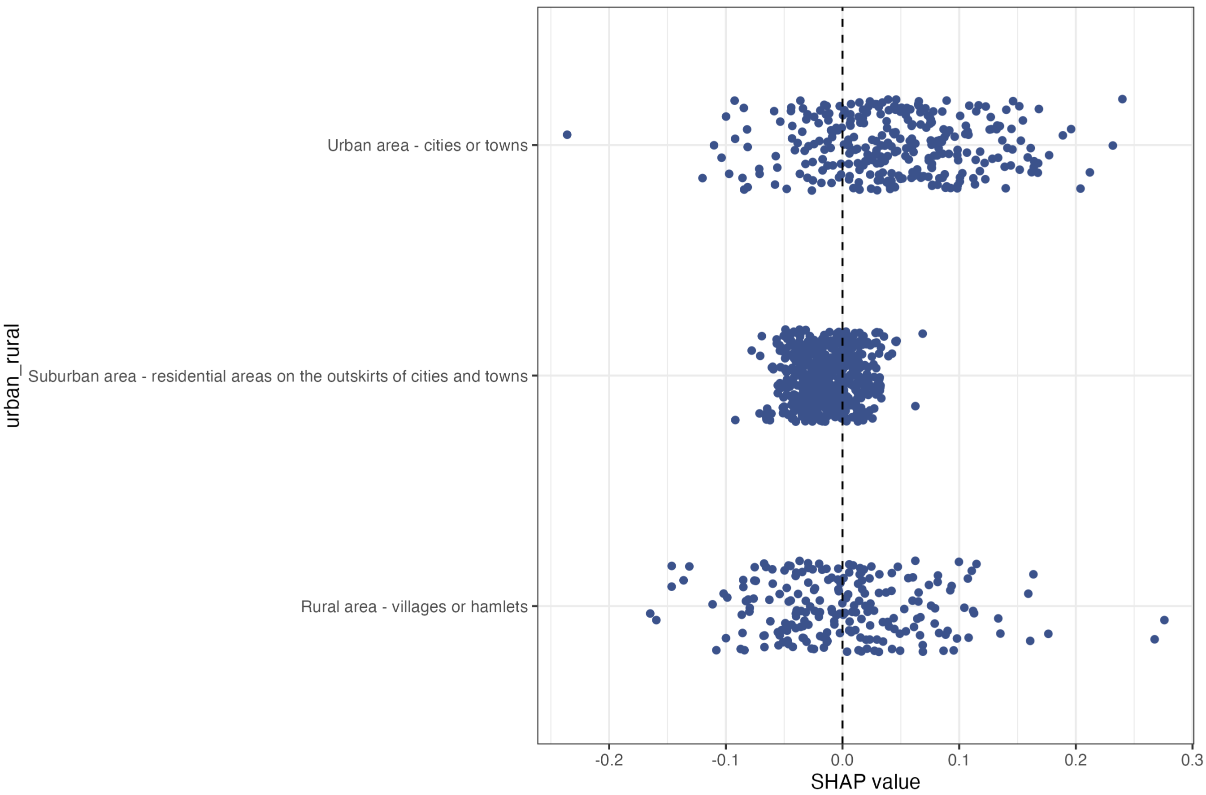


Figure M35: 2024 GE vote SHAP Values Predicting *Opinium Prediction Errors* Using Covariates

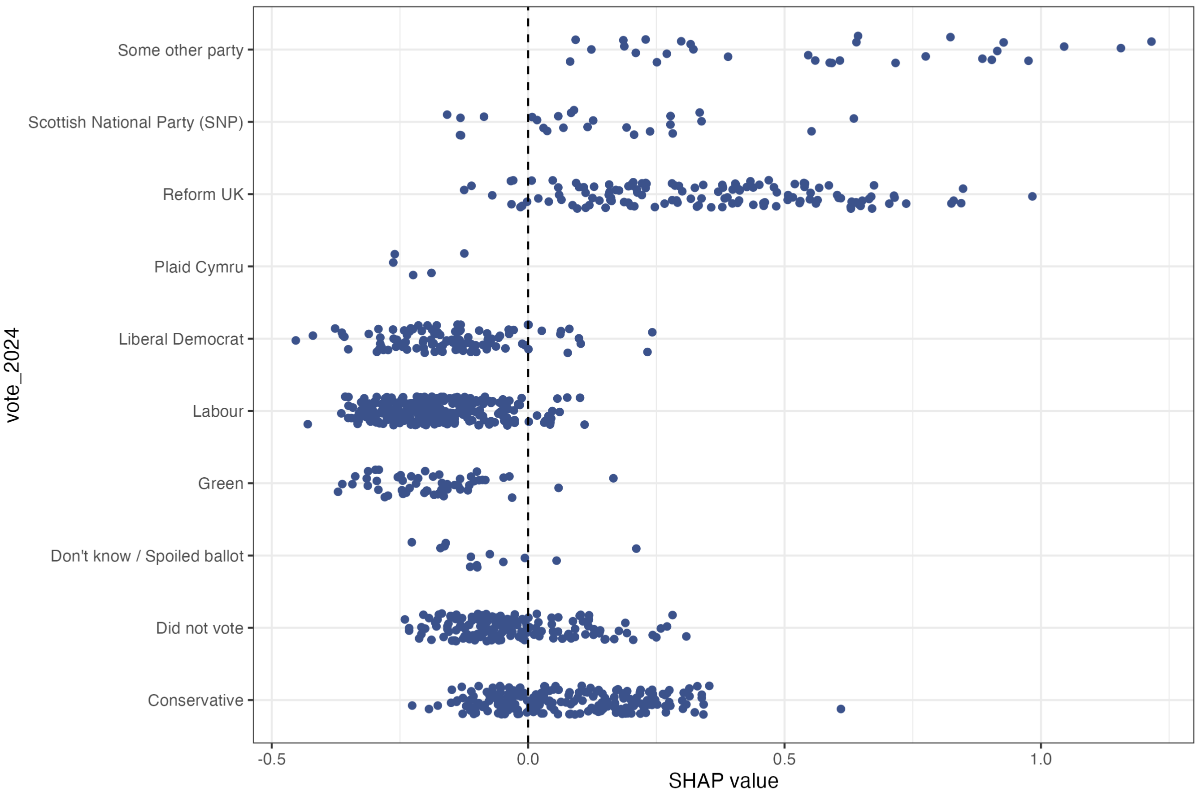


Figure M36: Vote intention SHAP Values Predicting *Opinium Prediction Errors* Using Covariates