Quantifying the potential persuasive returns to political microtargeting

Ben M. Tappin[a,b,1], Chloe Wittenberg[a] (ID), Luke B. Hewitt[c], Adam J. Berinsky[a] (ID), and David G. Rand[b,d] (ID)

Much concern has been raised about the power of political microtargeting to sway voters' opinions, influence elections, and undermine democracy. Yet little research has directly estimated the persuasive advantage of microtargeting over alternative campaign strategies. Here, we do so using two studies focused on U.S. policy issue advertising. To implement a microtargeting strategy, we combined machine learning with message pretesting to determine which advertisements to show to which individuals to maximize persuasive impact. Using survey experiments, we then compared the performance of this microtargeting strategy against two other messaging strategies. Overall, we estimate that our microtargeting strategy outperformed these strategies by an average of 70% or more in a context where all of the messages aimed to influence the same policy attitude (Study 1). Notably, however, we found no evidence that targeting messages by more than one covariate yielded additional persuasive gains, and the performance advantage of microtargeting was primarily visible for one of the two policy issues under study. Moreover, when microtargeting was used instead to identify which policy attitudes to target with messaging (Study 2), its advantage was more limited. Taken together, these results suggest that the use of microtargeting—combining message pretesting with machine learning—can potentially increase campaigns' persuasive influence and may not require the collection of vast amounts of personal data to uncover complex interactions between audience characteristics and political messaging. However, the extent to which this approach confers a persuasive advantage over alternative strategies likely depends heavily on context.

political persuasion | microtargeting | heterogeneity | randomized experiment | political advertising

Political advocacy groups often spend vast sums of money on advertisements seeking to alter public support for a particular issue or candidate (1, 2), and an increasingly large share of this spending is directed toward online campaigns (3, 4). An oft-cited reason for this trend is that digital advertising enables advocacy groups to more easily and specifically tailor their messaging to different groups (5–7)—a practice frequently described as political microtargeting (8).

Political microtargeting appears to be growing in prevalence (8–10; though see ref. 11) and has caused a great deal of public concern. In an especially salient case, prior to the 2016 U.S. presidential election, the media consulting firm Cambridge Analytica reportedly used unobtrusively collected data from over 50 million Facebook users to target potential voters with political advertisements (12). The subsequent media coverage of this scandal likened Cambridge Analytica's approach to "psychological warfare" (13), and much was written about the seemingly unprecedented capacity of microtargeting to sway voters' opinions, influence elections, and undermine democracy (6, 7, 14–16; though see ref. 17). Moreover, such concerns are not only evident in media narratives but are also reflected in academic discourse (18–20) and public opinion polls (21–23).

Nevertheless, a large body of academic research on political persuasion suggests that political microtargeting's persuasive returns may in fact be limited. Political microtargeting relies on treatment effect heterogeneity—that is, different groups of people responding in different ways to different messages. Yet the results of prominent, large-scale investigations of political persuasion suggest that treatment effect heterogeneity is relatively uncommon and, where it is found, tends to be small; in general, political messages seem to influence the attitudes of different types of people to a broadly similar extent (24–27). On this basis, one might expect—contrary to popular concern—that political microtargeting may not confer a meaningful advantage over nontargeted campaign messaging strategies. Indeed, in line with this view, the few empirical investigations of microtargeting to date have produced inconsistent patterns of results across studies (28–31). Given these

Significance

Despite widespread concerns about political microtargeting, the assumption that this approach confers a persuasive advantage over other messaging strategies has rarely been directly tested. Using two survey experiments focused on issue advocacy in the United States, we found some evidence to support this assumption: in a common advocacy scenario, our microtargeting strategy produced a relatively larger persuasive impact, on average, compared to several alternative messaging strategies. Notably, however, this advantage was observed under highly favorable conditions, and, in a different advocacy scenario, microtargeting's advantage was much more limited. Collectively, our results shed light on the potential impact of political microtargeting on public support for policy issues.

conflicting considerations, the potential returns to political microtargeting remain theoretically ambiguous.

In this paper, we thus directly and systematically evaluate the potential persuasive advantage of political microtargeting across two large studies. In contrast to much of the existing research in this area, in each study, we i) examine a large sample (i.e., dozens) of different persuasive messages, in order to more reliably uncover whether and to what extent different types of people are most effectively persuaded by different types of messages, and ii) use randomized survey experiments in order to rigorously compare the persuasive impact of a microtargeting strategy against two alternative messaging strategies (summarized in Table 1).

The microtargeting strategy that we evaluate here uses a custom machine learning procedure and pretesting data from survey experiments to identify the particular sets of covariates (e.g., age, gender) by which to target messages, as well as determine which messages to show to which individuals. Importantly, though this particular strategy was designed to simulate a plausible approach to message targeting, it may not necessarily reflect the strategies currently in use by campaigns, as campaigns rarely disclose details about their proprietary targeting practices. Instead, the microtargeting strategy that we evaluate here reflects an approach that many campaigns could be implementing (both now and in the future). Indeed, at present, medium-to-large campaigns in the United States likely command the resources needed to successfully execute a similar type of microtargeting strategy (1, 2, 32). In addition, the expertise needed to carry out this strategy already exists at many political consulting firms (33–36). As such, even if our strategy is relatively sophisticated by current standards, it nonetheless sheds important light on microtargeting's potential persuasive returns.

We compare the performance of our microtargeting strategy against two less-sophisticated alternatives. The first, which we call the "naïve" strategy, is the simplest: Individuals are exposed to a message selected at random from a sample of relevant messages, thereby simulating a strategy in which a campaign does not perform any pretesting prior to running its messaging program. The second, which we call the "single-best-message" (SBM) strategy, is one in which a campaign exposes all individuals to the same, top-performing message—specifically, the message that is expected to be the most persuasive on average across the full population, as determined by a pretesting process in which the average effects of various messages are compared in a survey experiment. This latter benchmark is important for several reasons. First, it accounts for the fact that some messages are simply more persuasive than

others on average; that is, comparing microtargeting against an SBM strategy allows us to estimate the benefits of targeting messages, above and beyond showing everyone the one message that is predicted to be most persuasive across the population as a whole. Second, the SBM approach is increasingly used by actual campaigns in the United States, supported by the services of various political media consulting firms (33–36); it is therefore an especially relevant point of comparison when evaluating the possible advantages of a microtargeting strategy.

To assess the persuasiveness of these strategies, we use a two-phase study design, consisting of a calibration phase, followed by an experimental phase (for a visual overview of our study design, see *SI Appendix*, Fig. S1 in *SI Appendix*, section S1). In the calibration phase, we conduct new analyses of existing data (37, 38) from previously published survey experiments that both i) tested a large number of persuasive messages and ii) included individual-level measures of a diverse array of personal and political covariates. We use these datasets to simulate a "pretesting" process wherein a campaign seeks to gauge the efficacy of different appeals before selecting a final set of message(s) to distribute as part of its messaging program. Specifically, we use linear regression to identify which message is likely to be the most persuasive on average across the full population (for the SBM strategy) and machine learning models to predict which messages will be most persuasive to different subpopulations (for our microtargeting strategy). We then use the outputs from this calibration phase as inputs into a new survey experiment conducted specifically for the current project, which provides a critical out-of-sample test of our microtargeting strategy. In this experimental phase, we randomize a new sample of U.S. adults (*n* = 5,284) to one of the three campaign strategies described above or to a control group. This design allows us to rigorously test the extent to which microtargeting offers a persuasive advantage over the alternative messaging strategies considered here.

We investigate the performance of our three messaging strategies across two campaign contexts. In our first study (Study 1), we examine a relatively common context in which a campaign aims to sway public opinion on a single, well-defined policy issue, such as support for a universal basic income (i.e., "policy-centered campaigns"). In our second study (Study 2), by contrast, we explore a less typical setting in which a campaign aims to change people's underlying belief systems, rather than their opinions on a single issue, and thus seeks to identify which of many broad topics (e.g., gun control, immigration, climate change) will maximize the persuasive impact of its messaging program (i.e., "multiissue campaigns"). For instance, an advocacy group attempting to shift public opinion leftward may seek to discover the issues on which people are most persuadable, rather than which messages for a predefined issue elicit the strongest attitude change. To maximize comparability in the experimental phase, we implemented these two studies simultaneously within a single survey (*Materials and Methods*).

Study 1: Policy-Centered Campaigns

Calibration Phase. In Study 1, we simulated two issue advocacy campaigns whose goal was to influence public opinion on one of two policy proposals: i) the U.S. Citizenship Act of 2021, a legislative bill proposed by President Joe Biden that includes a host of immigration reforms, and ii) a universal basic income (UBI). For the calibration phase of this study, we analyzed existing data from two large-scale survey experiments (37). In both experiments, respondents were shown a short video about the target policy (i.e., the U.S. Citizenship Act or a UBI) before reporting their level of support for this policy.

Table 1. Summary of the different campaign messaging strategies compared in each of our studies

Campaign messaging strategy	Implementation
(1) Microtargeting	Pretest messages → train machine learning models → show people whichever message is predicted to be best for them personally, based on their demographic and psychological traits
(2) Single best message (SBM)	Pretest messages → show everyone the message that performed the best *on average* in the pretest
(3) Naïve	No pretesting of messages → show people a message selected randomly from the full set of messages

In the U.S. Citizenship Act experiment (n = 17,013), respondents were randomly assigned to one of 26 treatment groups or a control group. Respondents in the control group were shown a brief message conveying basic information about the U.S. Citizenship Act, whereas respondents in the treatment groups were shown a message that included this same information but also contained one of 26 arguments in favor of the Act. The UBI experiment (n = 6,408) followed a similar design. The respondents were randomized to one of 10 treatment groups or a control group; respondents in the control group were shown a short video relaying factual information about a universal basic income, whereas respondents in the treatment groups were shown a longer clip that included this same information, as well as an argument against implementing a UBI. The treatment messages generally spanned a wide range of different rhetorical strategies, including appeals to moral values, such as fairness, loyalty, and sanctity (39, 40); appeals to religious authority; scientific and historical evidence; expert and public opinion; and common sense and ad hominem attacks on people with opposing views (*SI Appendix*, section S2.1.2).

As described above, we treated the data from these previous studies as if they came from a "pretesting" process in which a campaign piloted a set of messages before distributing one or more as part of their messaging program. As we describe below, this practice, increasingly common among political campaigns (33–36), can be used to implement each of the messaging strategies outlined in Table 1.

Microtargeting strategy. For the microtargeting strategy, we used the calibration phase to determine which messages to show different groups in the experimental phase. To do so, we first established the specific set of covariates (e.g., age, gender, and partisanship) by which to partition the population into different audiences. Namely, we determined whether it is most effective to target messages to individuals based solely on their age; their age and gender; their age, gender, and partisanship; or some other combination of covariates. To achieve this first step, we used a two-part crossvalidation procedure consisting of a training and test phase. We started by taking the full set of pretreatment covariates available in each dataset, which included standard demographic variables (e.g., age, gender, party identification) as well as psychological variables (e.g., self-reported moral values; see *Materials and Methods*), and constructed an exhaustive list of covariate "profiles," where each profile corresponds to a unique combination of covariates. For example, age is one covariate profile; gender is another profile; and age and gender is yet another profile, such that the list of covariate profiles covers all possible permutations of the covariate space.

For each covariate profile, we trained a series of generalized random forest models (41) on 75% of the data (the training set) in order to learn the extent of treatment effect heterogeneity across the covariates in that profile. For each trained model, we then identified the message that was predicted to be most persuasive for each subgroup within that covariate profile. For instance, for the covariate profile based on partisanship, we determined which messages were predicted to have the largest persuasive effects for Democrats, Republicans, and Independents. We then used the remaining 25% of the data (the test set) to estimate the persuasive impact of targeting the identified messages to the corresponding subgroups, which allowed us to determine the particular covariate profile that maximized the expected persuasive impact of microtargeting (for full details on this procedure, see *Materials and Methods*).

After running this train–test procedure 250 times, we computed the median of the estimates for each covariate profile across runs

and identified the three top-performing covariate profiles for each dataset, stratified by the number of covariates used for targeting (i.e., 1, 2-3, or 4+ covariates). This stratification allowed us to explore whether more complex targeting procedures performed better in the later experimental phase. We then conducted a final modeling step: For each of the three top-performing covariate profiles, we refit the generalized random forest models to the full dataset, in order to most accurately identify the messages that were expected to be most persuasive for each subgroup. The outcome of this process was a list of predictions of the most persuasive messages for each subgroup within each of our top-performing covariate profiles; we used this list to determine which messages to show to respondents in the experimental phase, described in greater detail below.

Alternative messaging strategies. For our two alternative messaging strategies—the SBM and naïve approaches—the calibration phase was much simpler. For the SBM strategy, we first computed the average treatment effect of each persuasive message (across all respondents), relative to the control group, and then identified the message with the largest estimated effect. The naïve messaging strategy, by contrast, had no pretesting component and thus no calibration procedure; this strategy simply involved showing individuals in the later experimental phase a randomly selected appeal from the full set of potentially persuasive messages. Additional details about these two strategies are available in the *Materials and Methods* section.

Experimental Phase. In the calibration phase of Study 1, we used previously collected data to identify the set of covariate profiles (and associated persuasive messages) for which microtargeting was expected to confer the greatest persuasive advantage over alternative messaging strategies. To assess how well this strategy performed "in the wild," we then conducted a critical out-of-sample test. In this "experimental phase" of our study, we recruited a new sample of U.S. adults (n = 5,284) for an online survey in February 2022. As part of this study, respondents completed three separate experimental modules, presented in random order: i) the U.S. Citizenship Act module, ii) the UBI module, and iii) a multiissue module (summarized later as part of Study 2). Within each module, respondents were randomized to either a control group or one of the three treatment conditions, representing the three messaging strategies outlined in Table 1.

Respondents in the control group of each module were shown the same placebo messages included in the previously analyzed experiments. These placebo messages conveyed basic information about each policy but did not express either support for or opposition to the policy. Respondents in the naïve condition were instead shown a randomly selected message from the set of persuasive appeals examined in the calibration phase (i.e., the 26 treatment messages included in the original U.S. Citizenship Act study or the 10 treatment messages included in the UBI study). Respondents in the single-best-message condition were shown the message that was estimated in the calibration phase to be most effective in swaying policy support, on average, across the full population.

Finally, respondents in the microtargeting condition were shown the message that was expected to be most persuasive to them, given their self-reported demographic and personality traits. In each module, we further randomized respondents in this condition to be targeted at different levels of complexity (with targeting based on one covariate only, two to three covariates, or four covariates). For example, in the U.S. Citizenship Act module, approximately one-third of respondents in the microtargeting condition were targeted based on their party affiliation; one-third
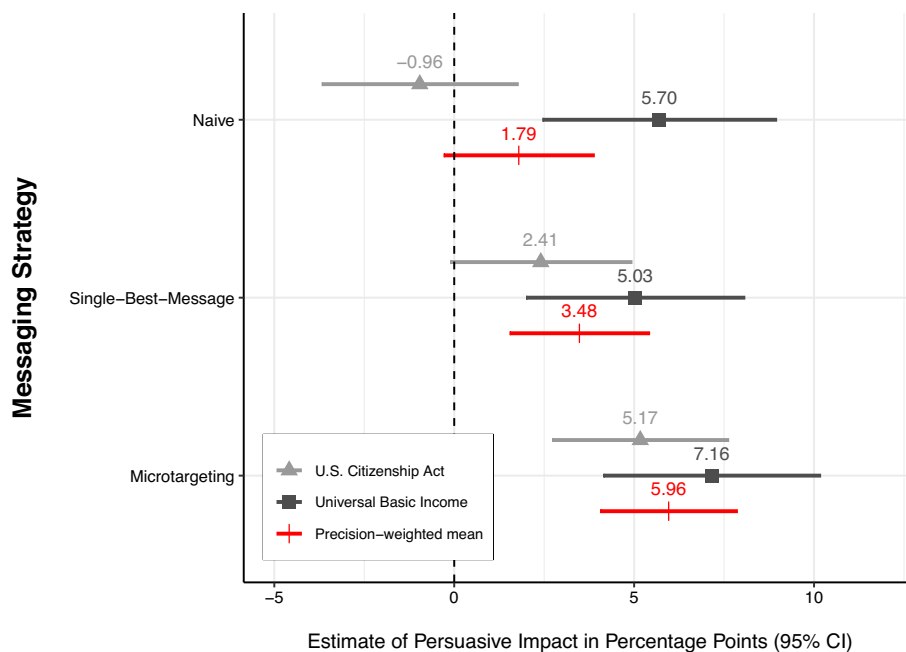
Fig. 1. Political microtargeting increases persuasive impact, relative to the alternative messaging strategies, in Study 1. Results from the experimental phase of Study 1. The first two rows display the estimated persuasive impact of showing people a random message (the naïve strategy) or the best overall message in the calibration phase (the single-best-message strategy), respectively. The third row displays the estimated persuasive impact of microtargeting, collapsed across the three covariate profiles used to target the messages. The three covariate profiles correspond to the "top-performing" profiles for each policy issue (where the expected advantage of microtargeting was largest in the calibration phase). 95% CI are based on robust SEs. *SI Appendix*, section S2.3.2, reports average ratings of policy support across conditions and covariate profiles.

were targeted based on their ideological self-placement and age; and one-third were targeted based on their ideology, age, and endorsement of two sets of moral values. All relevant covariates were measured pretreatment in the survey, and the design and analysis of these experiments was preregistered. Further details regarding the design are reported in the *Materials and Methods* section (*SI Appendix*, section S3).

As shown in Fig. 1, in the aggregate, we observe a persuasive advantage of microtargeting over the alternative strategies. On support for the U.S. Citizenship Act, we estimate that the persuasive impact of microtargeting is over twice as large as that of the single-best-message strategy (5.17 vs. 2.41 percentage points, respectively, $P = 0.016$) and even larger still compared to the naïve messaging strategy (5.17 vs. –0.96, respectively, $P < 0.001$). On support for a UBI, the estimated effects of microtargeting are smaller and not statistically significant, but they point in a similar direction; for this module, the persuasive impact of microtargeting is approximately 40% larger than the single-best-message strategy (7.16 vs. 5.03, respectively, $P = 0.12$) and 25% larger than the naïve strategy (7.16 vs. 5.70, respectively, $P = 0.33$). Given our interest in the general impact of microtargeting (as opposed to microtargeting on a particular policy issue), we also compute precision-weighted means of the differences in conditions across both modules. When doing so, we find that the average persuasive impact of the microtargeting strategy is approximately 70% larger than the single-best-message strategy (5.96 vs. 3.48, respectively, $P = 0.004$) and over 200% larger than the naïve strategy (5.96 vs. 1.79, respectively, $P < 0.001$). In all cases, the reported estimates and *P*-values are based on OLS regression models with robust SEs.

Notably, however, we find no evidence that the persuasive returns to microtargeting are greater for covariate profiles based on a larger number of personal characteristics. In fact, as shown in Fig. 2, the average persuasive advantage of microtargeting over the alternative messaging strategies is qualitatively largest when the messages were targeted by one covariate only—as opposed to

when the messages were targeted by two or four covariates—though the differences are relatively slight. Contrary to conventional wisdom, we therefore find little evidence that targeting is most impactful when using vast amounts of personal data to uncover complex interactions between audience characteristics and different types of messages (7, 8, 12). Instead, this result highlights that relatively basic targeting practices—focused, for instance, solely on partisanship or ideology—can also confer a persuasive advantage over alternative strategies.

In sum, the results of Study 1 suggest that in a typical campaign context, in which an advocacy group seeks to select issue advertisements that will maximize persuasion on a single, well-defined policy topic, showing different messages to different audiences can produce relatively more attitude change than less sophisticated messaging strategies. Importantly, we found that this advantage of microtargeting emerged even when the targeting procedure was coarse—for instance, when showing people different messages on the basis of just their party affiliation. We thus find no evidence here to support the idea that the campaigns who benefit the most from message targeting are those that use large amounts of personal data to unearth complex interactions between audience characteristics and political messaging.

Study 2: Multiissue Campaigns

In Study 2, we explore the generalizability of the microtargeting approach to a different setting, in which campaigns do not focus on a specific policy issue but instead aim to influence people's broader ideology, leveraging whatever issues they can (e.g., by convincing people to generally adopt more liberal or conservative positions). In this context, campaigns may try to identify the types of issues on which people are most persuadable (e.g., gun control, immigration, or climate change), rather than determine which individual messages for a predefined issue elicit the strongest persuasive effects. Here, microtargeting could provide large returns
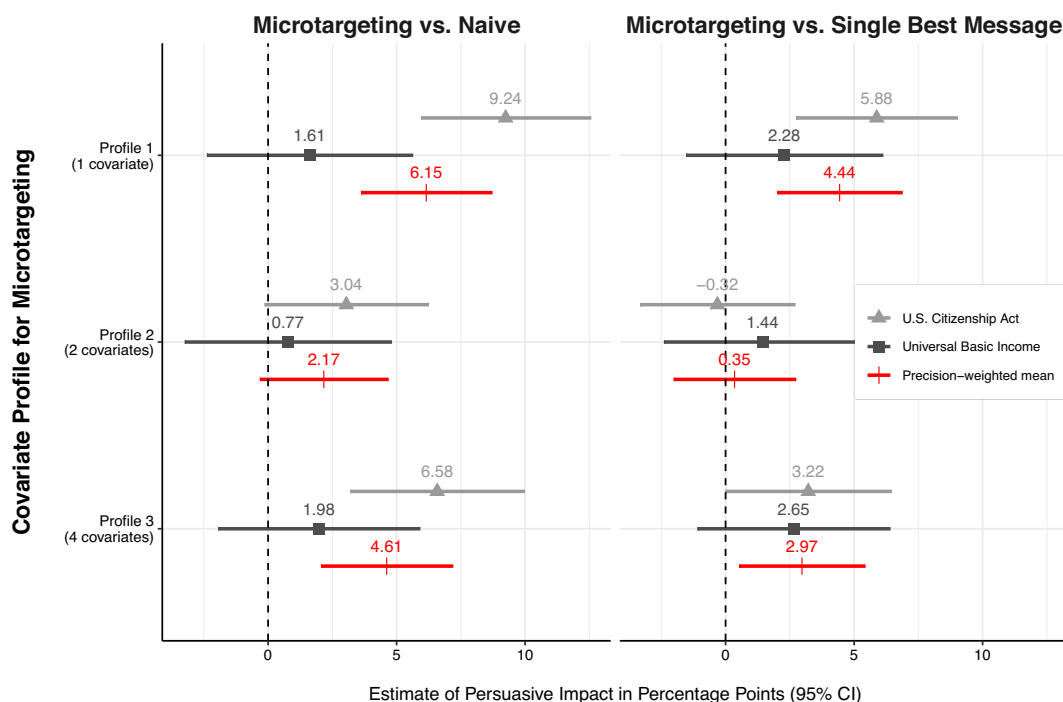
Fig. 2. No evidence that microtargeting by greater (vs. fewer) numbers of covariates increases persuasive impact in Study 1. The left-hand panel shows the average treatment effect of assignment to the microtargeting versus naïve strategy in the experimental phase of Study 1, and the right-hand panel shows the average treatment effect of the microtargeting versus single-best-message condition. In all cases, effect estimates are presented as percentage points and are estimated using OLS models with robust SEs. We show results for each of the two individual modules, as well as a precision-weighted mean calculated across these two modules. For the U.S. Citizenship Act module, respondents in the microtargeting condition were targeted based on either their partisan affiliation (Profile 1); their age and ideological self-placement (Profile 2); or their age, ideological self-placement, and endorsement of two sets of moral values (the "binding" foundations of authority, loyalty, and sanctity and the "individualizing" foundations of care and fairness; Profile 3). For the UBI module, respondents in the microtargeting condition were targeted based on either their endorsement of these "individualizing" moral foundations (Profile 1); their endorsement of both the "individualizing" and "binding" moral foundations (Profile 2); or their age, gender identity, and endorsement of these two sets of moral foundations (Profile 3). For more information about how these covariate profiles were selected, see *Materials and Methods*.

to persuasive impact if different types of people are more or less persuadable on different types of issues. On the contrary, if people tend to be most persuadable on the same types of issues, then it may be better to show everyone messages about the same issue. To assess these competing predictions, we used a similar, two-phase approach as in Study 1, with minor modifications to account for differences in the design of the two studies (e.g., the structure of the control group and the outcome variables, described in greater detail below).

Calibration Phase. In the calibration phase, we conducted new analyses of the data from the first survey experiment reported in ref. 38. In that experiment, U.S. respondents (n = 3,990; 9,712 total observations) were asked about three policy issues, drawn from a list of 36 issues, where each issue had its own distinct outcome variable. For example, one of the policy issues advocated for free legal representation for undocumented children, and another argued that the federal minimum wage should be increased to $15/h. For each of their assigned policy issues, respondents were then assigned to either a treatment or control group. Respondents in the treatment group were asked to watch a video ad corresponding to the relevant policy issue and then report their attitudes toward this issue (e.g., whether they thought the minimum wage should be increased). In contrast, respondents in the control group simply reported their attitudes without viewing the ad. For most issues, there was only one corresponding ad. However, for a small number of issues, there were several different ads (e.g., multiple videos about raising the minimum wage). As a result, we examined a total of 48 treatment messages, spanning 36 unique policy issues.

Unlike in Study 1, where the treatment messages were created by members of our team strictly for research purposes, the Study 2 messages were professionally produced video ads drawn from the Peoria Project's database of left-leaning messages (*SI Appendix*, section S2.2.2). As such, the ads always advanced progressive issue positions. The ads in Study 2 also varied along a larger set of dimensions than the messages in Study 1; in addition to spanning a wide range of issues, the messages covered an array of argument types (e.g., narrative vs. informational), messengers (e.g., politicians, experts, or laypeople), and editing styles (e.g., use of music and visuals). As in Study 1, the pretreatment covariates included basic demographic (e.g., age, gender, party identification) and psychological variables (e.g., political knowledge, cognitive reflection).

We used the same crossvalidation procedure described in the calibration phase of Study 1 to identify the set of individual-level covariates (and corresponding messages) that maximized the expected persuasive advantage of microtargeting, relative to the alternative messaging strategies in Study 2, pooling across the multiple experimental trials for each respondent (for more information, see *Materials and Methods*).

Experimental Phase. To assess the relative persuasiveness of microtargeting over the alternative strategies in Study 2, we examine the multiissue module of our out-of-sample validation experiment (n = 5,284), described above. In this module, respondents were first randomized to one of the three messaging strategies—naïve, single-best-message, or microtargeting—before being assigned to either a treatment or control group with equal probability (as described in the *Materials and Methods* section, this treatment
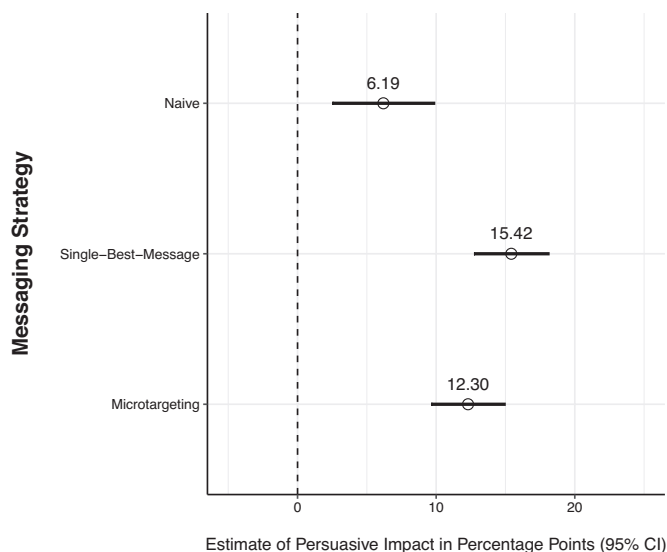
Fig. 3. Political microtargeting increases persuasion over the naïve but not the single-best-message strategy in Study 2. Results from the experimental phase of Study 2. The first two rows display the estimated persuasive impact of showing people a random message (the naïve strategy) or the best overall message in the calibration phase (the single-best-message strategy), respectively. The third row displays the estimated persuasive impact of microtargeting, collapsed across the three covariate profiles used to target the messages. The three covariate profiles correspond to the "top-performing" profiles (where the expected advantage of microtargeting was largest in the calibration phase). 95% CI are based on robust SEs. *SI Appendix*, section S3.4.2, reports average ratings of policy support across conditions and covariate profiles.

assignment procedure differed from the approach used in Study 1 because the outcome variables were specific to each message in Study 2). For each condition, we thus estimated the difference in reported policy support among respondents who were versus were not assigned to view a given message, which corresponded to a randomly selected message in the naïve condition; the overall top-performing message in the single-best-message condition; or a personally targeted message in the microtargeting condition.

As shown in Fig. 3, we estimate that the persuasive impact of microtargeting is directionally (although not significantly) smaller than the SBM strategy (12.30 vs. 15.42 percentage points, respectively, $P = 0.11$) but is approximately twice the size of the naïve messaging strategy (12.30 vs. 6.19, respectively, $P = 0.008$). This pattern holds regardless of the number of covariates that were used to target respondents (*SI Appendix*, section S3.4.2). In summary, we find no evidence that targeting different issue attitudes for different people is more impactful than simply distributing to everyone a single, high-quality (pretested) ad that speaks to the same issue. However, the microtargeting strategy clearly outperforms a naïve strategy in which campaigns broadcast one or more issue ads without engaging in any pretesting or message targeting.

Discussion

Much has been written about the seemingly unprecedented power of political microtargeting to sway voters' opinions, influence elections, and undermine democracy. Yet the assumption that political microtargeting confers a persuasive advantage over other more conventional messaging strategies has seldom been directly tested. In this paper, we conducted such a test and found that our microtargeting approach outperformed alternative messaging strategies—under certain conditions. In a study representing a typical issue advertising context, in which campaigns advocate for or against a specific policy proposal, we estimated that our

microtargeting strategy exceeded the persuasive impact of alternative messaging strategies by an average of 70% or more—driven primarily by messaging on a single policy issue (the U.S. Citizenship Act). Notably, however, we found no evidence that targeting messages by more covariates amplified persuasive impact in this setting; that is, the persuasive advantage of microtargeting did not depend on complex interactions between people's personal characteristics and the different types of messages. Moreover, in a less conventional advertising context, the advantage of microtargeting was more limited. Thus, while our results indicate that message targeting has the potential to improve campaigns' persuasive impact over alternative messaging strategies, this advantage may be highly context dependent.

In addition to informing ongoing debates about microtargeting, our results expand on previous academic research by directly comparing the effectiveness of different messaging strategies using a large number of distinct political messages. For example, while some previous studies investigate treatment effect heterogeneity across multiple political messages, they typically do not distinguish the effectiveness of campaigns that employ microtargeting from those that employ other data-driven messaging strategies, such as the single-best-message approach (24, 42–44). Moreover, even when past studies have investigated the persuasive returns to microtargeting, these studies have often examined very small samples of persuasive messages (typically ≤ 5), which may explain why the patterns of results vary considerably across studies (28–31). By contrast, our design incorporated dozens of different messages, enabling us to more reliably uncover whether, and to what extent, different types of people are more effectively persuaded by different messages.

Furthermore, while past research on consumer behavior has identified large returns to microtargeting outside the political domain (e.g., ref. (45)), this work faces threats to internal validity given the nonrandom assignment procedures used by social media companies to distribute advertising (5, 46). More broadly, the returns to microtargeting in nonpolitical contexts (47) may not neatly generalize to political contexts (or vice versa), given fundamental differences in the underlying psychology and behaviors of interest across these domains. For example, politics often implicates people's deeply held values and group identities, which may cause individuals' attitudes to respond in systematically different ways to persuasive messaging (48). Moreover, the consumer behavior context offers easily observable metrics to use for optimization (e.g., clicks, conversions), whereas political microtargeting often requires measuring attitudes and behaviors that are not so easily inferred from on-platform metrics.

A key question, however, concerns the ecological validity of our design—most notably, how representative our particular microtargeting strategy is of the approaches used by real campaigns. Although campaigns rarely disclose their targeting strategies, our microtargeting approach relies on a message pretesting method that is increasingly available to, and used by, actual campaigns (33–36). Furthermore, while our machine learning procedure may exceed the current capabilities of small and (perhaps) medium-sized campaigns, it is feasible for large, well-funded political campaigns—for example, those that can hire the relevant expertise in the form of data scientists and consultants. More generally, however, regardless of the current sophistication of political campaigns, our focus in this study is the potential persuasive advantage of political microtargeting. That is, we assess the possible gains campaigns could achieve by using a plausible approach to microtargeting.

Nevertheless, it is still important to carefully consider the generalizability of our results to political campaigning in the real world. On the one hand, our estimates may understate the relative

persuasive advantage of political microtargeting. For example, social media algorithms typically optimize for engagement (e.g., outbound clicks, view time, or conversion rates), which could boost persuasive impact overall, since engaging with a message is necessary (if not sufficient) for persuasion. Additionally, from an academic standpoint, we had an unusually large amount of data with which to train our machine learning models (tens of thousands of observations and dozens of different messages from randomized survey experiments), but well-funded political campaigns often command far more resources and thus can potentially leverage much larger datasets for pretesting and model calibration (2). As a result, they may be able to train models at a finer covariate resolution—potentially further enhancing the persuasive impact of microtargeting beyond what we can observe here.

On the other hand, there are several reasons to think our estimates may overstate the relative persuasive advantage of political microtargeting. First, we tested these messaging strategies using survey, rather than field, experiments. This approach allowed us to directly and precisely measure the covariates used for targeting our messages. In practice, however, most campaigns lack access to such high-quality information about people's demographic and psychological profiles. This limited capacity to directly measure individuals' personal characteristics may lead to mis-targeting and therefore diminish microtargeting's efficacy (29). Second, to the extent that campaigns rely on social media companies to execute their digital advertising strategies, the persuasive advantage of microtargeting may be attenuated by these companies' idiosyncratic procedures for distributing ads (5, 46). Third, although a survey experimental approach enabled us to rigorously control the context in which our messages were delivered, it also required respondents to engage with content they might have otherwise ignored in real life. For example, a recent digital field experiment reported that approximately half the time disseminated political ads were not even seen by their intended audience (49)—let alone viewed in full. Though outside the bounds of the present research, future work should thus explore whether individuals are more likely to engage with targeted versus nontargeted political messages.

Our conclusions are also limited by our focus on a small set of issues (and their corresponding treatment messages). Individuals' susceptibility to persuasive appeals may vary widely across different issues and advertisements (50, 51). The generalizability of our findings to other political issues and sets of messages therefore remains an open question. For example, the issues we focused on in Study 1 were both relatively low salience. In contexts where individuals have stronger prior beliefs (52) or are exposed to countermessaging (53, 54)—namely, the contexts in which this type of resource-intensive, data-driven campaigning is most likely to occur—persuasive effects are likely to be much smaller across all messaging strategies. It is possible that these smaller absolute effect sizes could also result in a smaller relative advantage for microtargeting—for instance, by making it more statistically challenging to reliably detect and exploit treatment effect heterogeneity. Future work that applies our methodology to a wider variety of contexts, including candidate- versus issue-centered campaigns and campaigns outside the United States, can speak to these questions.

Finally, even if microtargeting confers a sizable persuasive advantage over alternative messaging approaches, it still may not be the optimal strategy for all campaigns. Specifically, our microtargeting strategy likely requires more time, money, and data science expertise to deploy than those of the alternative messaging strategies considered here. For many campaigns, it could be the case that deploying one of these other messaging strategies—such as the single-best-message approach—would generate greater persuasive impact per dollar spent. Future research that

analyzes the cost-effectiveness of political microtargeting could shed light on this question.

Aside from its potential persuasive advantage, the widespread use of microtargeting could also undermine trust in democratic institutions and decrease political participation (9), given that many members of the public voice strong opposition to this practice (21–23, 55). Some of these fears rest on the concern that microtargeting could allow politicians to covertly promise different and/or mutually exclusive policies to different segments of the population (56). Furthermore, data privacy remains a central priority to activists and voters alike. Microtargeting often involves gathering and leveraging substantial amounts of personal data (8, 57), which may lead to individuals' data being improperly obtained (as was the case in the Cambridge Analytica scandal). Notably, however, we reiterate that we found no evidence that the persuasive advantage conferred by microtargeting requires the accumulation of vast amounts of personal data; rather, targeting messages by one covariate alone was enough to generate the persuasive benefit we observed—a tactic that considerably predates digital campaigning. Thus, although campaigns may aim to collect a large volume of personal data for the purposes of accurately identifying and reaching out to the "right" people (e.g., swing voters), our results suggest that these data may be less fruitful when it comes to identifying which messages to show these audiences to maximize persuasive impact.

Furthermore, the discrepant results for Studies 1 and 2 highlight other potential limits to microtargeting's influence. In particular, while the results of Study 1 (particularly for the U.S. Citizenship Act) suggest that microtargeting can confer a sizable relative advantage when campaigns aim to influence attitudes toward a specific policy issue, the results of Study 2 suggest that microtargeting offers fewer benefits for campaigns seeking to shift people's broader ideological commitments by finding the issues on which they are most persuadable. These discrepant results could emerge because people are least knowledgeable about the same types of issues—for example, those that are less frequently covered by the media—and tend to be most strongly persuadable for these low-information issues (52, 58). Indeed, in Study 2, the messages predicted to have the largest effects on political attitudes were often ones that dealt with relatively obscure policies such as net neutrality.

Nevertheless, there are other potential explanations for Study 2's divergent findings. For example, the pretest data in Study 2 came from a study that used a between- and within-subjects design, whereas our original follow-up study used a fully between-subjects design. In addition, the data for Study 2's calibration phase were collected almost a year before the follow-up experiment was fielded; to the extent that the persuasiveness of different messages (both overall and across groups) varies over time, campaigns may need to perform multiple rounds of testing to ensure their models remain up-to-date. Future work should further investigate whether our Study 2 result is an anomaly or whether issue advocacy campaigns indeed stand to gain less by microtargeting messages that speak to different policy issues.

In sum, our results indicate that, in some settings, political microtargeting can generate sizable persuasive returns over alternative messaging strategies. Importantly, however, microtargeting's persuasive advantage may be highly context dependent, and we found no conclusive evidence that this advantage requires campaigns to amass vast amounts of personal data to uncover complex interactions between political messages and personal characteristics. Collectively, these results offer information about the possible persuasive impact of political microtargeting at a time when regulation of this practice is an active area of policy debate in many countries.

Materials and Methods

The design and analysis of our out-of-sample validation experiment (summarized in the experimental phases of Studies 1-2) was preregistered prior to data collection (https://osf.io/kvd9t).

Data Sources.

Study 1 data. In the calibration phase of Study 1, we drew on existing data from two survey experiments previously conducted by members of our research team (37). In the first experiment, we compared 26 treatment messages providing arguments in favor of the U.S. Citizenship Act to a control message that provided contextual information about this proposed bill. In the second experiment, we instead compared 10 treatment messages that opposed a universal basic income (UBI) to a control message that simply explained the basic concept of a UBI. Of note, the original experiments also included messages that were opposed to the U.S. Citizenship Act and supportive of a UBI. However, we excluded these messages here, due to resource constraints associated with the experimental phase of the study.

As part of the calibration phase, we defined a list of covariate profiles, corresponding to all unique permutations of the pretreatment covariates measured in the original studies. In Study 1, we used eight covariates to construct a total of 255 profiles: age (in years), gender (female/male), party identification (Democrat/Independent/Republican), ideology (measured on a 1 to 7 scale, binned into: liberal/moderate/conservative, where "leaners" are coded as ideologues), religiosity (1 to 7 scale), and respondents' endorsement of three different moral values (MF1, MF2, MF3). The exact wording of these items is available in *SI Appendix*, section S3.2. Respondents' moral values were inferred from their answers to 12 questions from the Moral Foundations Questionnaire, which was developed to measure endorsement of several core moral foundations, including the "binding" foundations of loyalty, authority, and sanctity; the "individualizing" foundations of care and fairness; and the foundation of "liberty" (39, 59). We used principal component analysis to identify which of the three dimensions each of the 12 questions loaded onto most strongly and then averaged the responses to the questions in each dimension. The resulting loadings conformed closely to theoretical expectations (MF1 = endorsement of binding foundations; MF2 = individualizing; MF3 = liberty, see *SI Appendix*, section S2.1.1). Sample demographics for Study 1 are reported in *SI Appendix*, section S2.1.1.

Study 2 data. For the calibration phase of Study 2, we instead used existing data from the first study reported in ref. 38, which included 36 distinct issue areas and 48 treatment messages. To construct our list of 511 covariate profiles, we used a total of nine pretreatment covariates: age (in years), gender (female/male/other gender identity), party identification (Democrat/Independent/Republican, where "leaners" are coded as partisans), ideology (liberal/moderate/conservative, where "leaners" are coded as ideologues), education (college degree/no college degree), race (White/non-White), income (<$50k/$50k-100k/>$100k), performance on a four-item Cognitive Reflection Test (integer from 0 to 4 indicating the number of correct responses), and performance on a four-item political knowledge test (integer from 0 to 4 indicating the number of correct responses). Sample demographics for Study 2 are reported in *SI Appendix*, section S2.2.1.

Calibration Phase.

Training data. The calibration phase of each study had two key goals: i) establish which covariate profile(s) maximize the expected persuasive impact of targeting, and ii) determine which message to show to which individuals in the experimental phase to produce the greatest attitude change. To achieve the first goal, we used a crossvalidation procedure consisting of a training and test phase. First, for the microtargeting strategy, we trained (on 75% of the data) a series of generalized random forest models–one model per treatment–control comparison–to estimate treatment effect heterogeneity across the covariates in that profile, using the grf package in R (41). We then used these models to predict the persuasive effect of each message for each value within the covariate space. For example, if the covariate profile was gender and party and the dataset was the UBI module, we generated 10 predicted persuasion estimates (one per treatment message) for each of six values in the covariate space: male-Democrat, male-Independent, male-Republican, female-Democrat, female-Independent, and female-Republican. We subsequently identified the message that was predicted to be most persuasive for each unique combination of gender and partisanship.

We then benchmarked these estimates against the single-best-message strategy. To do so, we used the training data to identify the treatment message that was

expected to be most persuasive across the population as a whole. Specifically, we used OLS models to estimate the average treatment effect (ATE) of each message in the training data, relative to the control group, and selected the message with the largest overall ATE. Within the training set, we therefore identified the top-performing message for each subgroup within each covariate profile, as well as the top-performing message overall (i.e., averaged across the full set of respondents in the training set).

Test data. We then used the remaining 25% of the data (the test set) to estimate the persuasive impact of targeting messages by each covariate profile, relative both to the control group and to a nontargeted messaging strategy (the single-best-message strategy). Through this process, we obtained a rank ordering of the covariate profiles, in terms of their relative persuasive advantage over alternative messaging strategies. We then used these rankings to select the top-performing covariate profiles to include in the experimental phase of the study.

To implement this procedure, for each covariate profile, we redefined the treatment group in the test set to include only those respondents who were shown the treatment message that was predicted to be most persuasive to them (based on their covariate values). For instance, if the covariate profile was partisanship, we filtered our analysis to Democrats in the test set who were shown the treatment message expected to be most persuasive to Democrats, Republicans in the test set who were shown the message expected to be most persuasive to Republicans, and so on. This approach reduces the size of the treatment group but, crucially, preserves covariate balance between the treatment and control groups (in expectation), because treatment assignment was, by design, independent of the pretreatment covariates. After subsetting the treatment group in this way, we computed the difference in mean policy support between the (filtered) treatment group and the control group, thereby yielding an estimate of the persuasive impact of targeting messages based on that covariate profile.

We then used a similar approach to estimate the performance of the single-best-message strategy. For this process, we employed a similar filtering procedure as above. We first redefined the treatment group to include only those respondents in the test set who were shown the treatment message that was previously identified (in the training set) as the most persuasive on average across all respondents. Then, we calculated the difference in mean policy support between this (filtered) treatment group and the control group. This estimate represents the persuasive effect of showing all respondents the same message that performed the best, on average, across the full sample–absent any targeting.

Finally, for each covariate profile, we computed the difference in the persuasive impact of the microtargeting and single-best-message strategies. To obtain stable estimates of this quantity, we used a Monte Carlo crossvalidation procedure, wherein we split the data into training and test sets 250 times and completed the process described above. For each profile, we computed the median effect estimate across these 250 runs and used these estimates to rank-order the covariate profiles by their estimated persuasive advantage over the single-best-message strategy. As described in the main text, we stratified the covariate profiles based on whether they included 1, 2-3, or 4+ covariates and identified the top-performing covariate profile within each stratum. This stratification process allowed us to explore whether targeting by more (versus fewer) covariates in the experimental phase produced larger persuasive effects.

Message selection. Having identified the top-performing covariate profiles for each dataset, we next estimated a final series of generalized random forest models–just for the top-performing covariate profiles–using the entirety of each dataset. We used the output of these models to generate predictions about which messages would be most persuasive for each unique group in the covariate space (e.g., Democrats, Independents, and Republicans for the partisanship profile), which determined which messages were shown to respondents in the microtargeting condition of the experimental phase. Following a similar logic, for each dataset, we ran OLS models on the full sample of respondents in order to identify the message with the largest overall ATE, which we used to determine which message would be shown to respondents in the single-best-message condition. The reason that we refit these models to the full dataset is because the train–test split procedure was only used to identify the top-performing covariate profiles; after identifying these profiles, we used all the data at our disposal to train the selected models, in order to most precisely uncover the relevant messages to use in the experimental phase of the study.

Experimental Phase. As described in the main text, the experimental phases of Studies 1 and 2 were conducted as part of the same survey experiment and shared a similar design and analysis. Thus, we describe them together below.

Sample. We contracted with Lucid Theorem in February 2022 to recruit a sample of U.S. adults, quota-matched to the national distribution on age, gender, ethnicity, and geographic region (60), with a preregistered sample size of 5,000 respondents. A total of 7,531 respondents entered the survey, 7,238 consented to participate, and 5,632 passed a series of preregistered technical checks confirming their ability to participate. Of these respondents, we removed three cases in which a respondent completed the survey multiple times, retaining only the response with the earliest timestamp (this was not included in our preregistration). In total, N = 5,284 respondents completed at least one experiment module and are therefore included in our analysis. *SI Appendix*, section S3.1, contains additional information about sample demographics. The survey was deemed exempt from requiring ethics approval by the Massachusetts Institute of Technology Committee on the Use of Humans as Experimental Subjects (ID: E-2285).

Research design. The survey consisted of three experiment modules, corresponding to each of the three datasets explored in the calibration phases of Studies 1-2: i) U.S. Citizenship Act (Study 1), ii) UBI (Study 1), and iii) multiissue (Study 2). The modules were completed by respondents in a random order. Prior to completing these modules, the respondents provided informed consent and were required to pass an audiovisual check to ensure they were willing and able to watch video content (*SI Appendix*, section S3.2.1). They then answered a battery of demographic and psychological measures (*SI Appendix*, section S3.2.2), which we used to determine the message they would be shown if they were assigned to the microtargeting condition for any of the three modules. Each module involved a self-contained experiment in which respondents were randomized to one of the several different conditions, where treatment randomization was independent across modules.

For the U.S. Citizenship Act and UBI modules (Study 1), respondents were first randomized to one of four conditions (with assignment probabilities in parentheses): control ($P = 0.2$), naïve ($P = 0.2$), single-best-message ($P = 0.3$), or microtargeting ($P = 0.3$). In the control condition, respondents were shown the same informational video that was used as a placebo message in the pretesting process. In the naïve condition, respondents were shown an issue message that was randomly selected from the full set of treatment messages explored in the calibration phase. In the single-best-message condition, all respondents viewed the exact same message: the video that had the largest estimated ATE in the calibration phase. Finally, in the microtargeting condition, respondents received whichever message was predicted to be most persuasive to them personally, based on their pretreatment covariates and the output of the generalized random forest models trained in the calibration phase (see also *SI Appendix*, section S3.3.1). To determine which set of covariates to use, respondents in this condition were additionally randomized to one of the three top-performing covariate profiles (corresponding to either 1, 2-3, or 4+ covariates), as identified in the corresponding calibration phase, with equal probability. After watching their assigned video, all respondents then rated their support for the target policy (i.e., the U.S. Citizenship Act or a UBI) using the exact same items as in the calibration phase.

For the multiissue module (Study 2), the design was slightly different, given that each message could have a different outcome variable. Respondents were first randomized to one of three conditions: naïve ($P = 0.2$), single-best-message ($P = 0.4$), or microtargeting ($P = 0.4$). In the naïve and single-best-message conditions, respondents were then randomized to the treatment or control group with equal probability, whereas respondents in the microtargeting condition were first randomized to one of the three top-performing covariate profiles–identified in the Study 2 calibration phase–with equal probability before being further randomized into either the treatment or control group (again with equal probability). In the treatment groups, respondents received a message that was determined by their condition: Respondents in the naïve condition received a randomly selected message from the full set of 48 messages examined in the calibration phase, respondents in the single-best-message condition received the message that had the largest estimated ATE in the calibration phase, and respondents in the microtargeting condition received a personalized message identified using the same procedure as in Study 1. After viewing their assigned message, respondents in the treatment group then rated their support for the target policy, measured using a 5-point Likert scale (where higher ratings indicated stronger agreement with the message's position). In contrast, respondents in the control group did not receive any message but were asked to complete the outcome variable corresponding to the message they would have received had they been in the treatment group. All messages and outcome variables are described in detail in *SI Appendix*, section S3.4.1.

For exploratory purposes, in all the three experiment modules, after completing the main outcome variable, respondents were also asked how important the issue in question was to them, and, when they saw a video, whether they would share it with a friend or colleague (*SI Appendix*, sections S3.3.1 and S3.4.1). Balance checks and attrition analysis suggested that randomization was successful and attrition was unproblematic for all the three modules (*SI Appendix*, sections S3.3.3 and S3.4.3).

Analytic strategy. For the U.S. Citizenship Act and UBI modules that comprise Study 1, we estimated the persuasive impact of each messaging strategy using OLS models with dummy variables for each of the three treatment conditions–naïve, single-best-message, or microtargeting–with the reference category set as the control group. To compute the difference between microtargeting and the other two messaging strategies, we also fit two additional OLS models where the reference category was the naïve and single-best-message condition, respectively, where our primary quantity of interest was the coefficient for the microtargeting variable. Following the preregistration, all inferences were based on robust SEs ("HC2" variant).

For the multiissue module in Study 2, the analytic strategy was slightly different, owing to the modified experiment design. We first estimated the persuasive impact of each messaging strategy using separate OLS models–subsetting the data by messaging strategy–that each included a single dummy variable for respondents' message exposure (i.e., assignment to the treatment versus control group). To compare microtargeting to the other two messaging strategies, we used a difference-in-differences approach. Specifically, we fit two additional OLS models (using the full sample of respondents) in which we interacted a dummy variable indicating message exposure with dummy variables indicating assignment to the microtargeting condition (versus the naïve or single-best-message conditions, respectively). A significant positive coefficient on the interaction term indicates that the average effect of assignment to the treatment versus the control group was larger in the microtargeting condition than that in the reference group (i.e., the naïve or single-best-message condition). This analytic approach was necessary because comparing means directly between the various treatment groups would be incoherent; there could be differences in ratings due to different outcome variables being asked across the conditions–rather than the messaging strategies themselves being differentially persuasive. Formal model specifications and results tables are available in *SI Appendix*, sections S3.3.4 and S3.4.4.

Author affiliations: [a]Department of Political Science, Massachusetts Institute of Technology, Cambridge, MA 02139; [b]Sloan School of Management, Massachusetts Institute of Technology, Cambridge, MA 02139; [c]Polarization and Social Change Lab, Stanford University, Stanford, CA 94305; and [d]Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, MA 02139

1. M. E. Reynolds, R. L. Hall, Issue advertising and legislative voting on the affordable care act. *Polit. Res. Q.* **71**, 102–114 (2018).
2. T. N. Ridout, E. F. Fowler, M. M. Franz, Spending fast and furious: Political advertising in 2020. *The Forum* **18**, 465–492 (2021).
3. L. Baum, "Presidential General Election Ad Spending Tops $1.5 Billion" (Wesleyan Media Project, 2020). https://mediaproject.wesleyan.edu/releases-102920/. Accessed 10 September 2022.

4. E. F. Fowler, M. M. Franz, T. N. Ridout, "Online political advertising in the United States" in *Social Media and Democracy: The State of the Field, Prospects for Reform* (Cambridge University Press, 2020), pp. 111–138.
5. M. Ali, P. Sapiezynski, A. Korolova, A. Mislove, A. Rieke, Ad Delivery Algorithms: The Hidden Arbiters of Political Messaging. arXiv [Preprint] (2019). http://arxiv.org/abs/1912.04255 (Accessed 10 September 2022).

6. S. Morrison, Why are you seeing this digital political ad? No one knows! *Vox* (2020). https://www.vox.com/recode/2020/9/29/21439824/online-digital-political-ads-facebook-google. Accessed 10 September 2022.

7. J. C. Wong, "It might work too well": The dark art of political advertising online. *The Guardian* (2018). https://www.theguardian.com/technology/2018/mar/19/facebook-political-ads-social-media-history-online-democracy. Accessed 10 September 2022.

8. Privacy International, Why we're concerned about profiling and micro-targeting in elections. *Privacy International* (2020). http://privacyinternational.org/news-analysis/3735/why-were-concerned-about-profiling-and-micro-targeting-elections. Accessed 10 September 2022.

9. J. Matthes *et al.*, Understanding the democratic role of perceived online political micro-targeting: Longitudinal effects on trust in democracy and political interest. *J. Information Technol. Polit.* **19**, 1–14 (2022).

10. O. Papakyriakopoulos, S. Hegelich, M. Shahrezaye, J. C. M. Serrano, Social media and microtargeting: Political data processing and the consequences for Germany. *Big Data Soc.* **5**, 2053951718811844 (2018).

11. G. Kefford *et al.*, Data-driven campaigning and democratic disruption: Evidence from six advanced democracies. *Party Politics* **29**, 13540688221084040 (2022).

12. M. Hu, Cambridge Analytica's black box. *Big Data Soc.* **7**, 2053951720938091 (2020).

13. C. Cadwalladr, The great British Brexit robbery: How our democracy was hijacked. *The Guardian* (2017). https://www.theguardian.com/technology/2017/may/07/the-great-british-brexit-robbery-hijacked-democracy. Accessed 10 September 2022.

14. BBC, Vote Leave's targeted Brexit ads released by Facebook. *BBC News* (2018). https://www.bbc.com/news/uk-politics-44966969. Accessed 10 September 2022.

15. M. Scott, Cambridge Analytica helped 'cheat' Brexit vote and US election, claims whistleblower. *POLITICO* (2018). https://www.politico.eu/article/cambridge-analytica-chris-wylie-brexit-trump-britain-data-protection-privacy-facebook/. Accessed 10 September 2022.

16. N. Singer, 'Weaponized Ad Technology': Facebook's Moneymaker Gets a Critical Eye. *The New York Times* (2018). https://www.nytimes.com/2018/08/16/technology/facebook-microtargeting-advertising.html. Accessed 10 September 2022.

17. B. Resnick, Cambridge Analytica's "psychographic microtargeting": What's bullshit and what's legit. *Vox* (2018). https://www.vox.com/science-and-health/2018/3/23/17152564/cambridge-analytica-psychographic-microtargeting-what. Accessed 10 September 2022.

18. J. Bayer, Double harm to voters: Data-driven micro-targeting and democratic public discourse. *Internet Policy Rev.* **9**, 1-17 (2020). https://policyreview.info/articles/analysis/double-harm-voters-data-driven-micro-targeting-and-democratic-public-discourse.

19. J. Burkell, P. M. Regan, Voter preferences, voter manipulation, voter analytics: Policy options for less surveillance and more autonomy. *Internet Policy Rev.* **8**, 1–24 (2019).

20. P. Lorenz-Spreen *et al.*, Boosting people's ability to detect microtargeted advertising. *Sci. Rep.* **11**, 15541 (2021).

21. J. McCarthy, U. S. In, Most oppose micro-targeting in online political ads. *Gallup.com* (2020). https://news.gallup.com/opinion/gallup/286490/oppose-micro-targeting-online-political-ads.aspx. Accessed 10 September 2022.

22. Open Rights Group, Public are kept in the dark over data driven political campaigning, poll finds. *Open Rights Group* (2020). https://www.openrightsgroup.org/press-releases/public-are-kept-in-the-dark-over-data-driven-political-campaigning-poll-finds/. Accessed 10 September 2022.

23. J. Turow, M. X. D. Carpini, N. A. Draper, R. Howard-Williams, *Americans Roundly Reject Tailored Political Advertising* (Annenberg School for Communication, University of Pennsylvania, 2012), pp. 1–30.

24. A. Coppock, *Persuasion in Parallel* (University of Chicago Press, Chicago, 2022).

25. A. Coppock, E. Ekins, D. Kirby, The long-lasting effects of newspaper Op-Eds on public opinion. *Q. J. Political Sci.* **13**, 59–87 (2018).

26. A. Coppock, S. J. Hill, L. Vavreck, The small effects of political advertising are small regardless of context, message, sender, or receiver: Evidence from 59 real-time randomized experiments. *Sci. Adv.* **6**, eabc4046 (2020).

27. A. Guess, A. Coppock, Does counter-attitudinal information cause backlash? Results from three large survey experiments. *Br. J. Polit. Sci.* **50**, 1–19 (2020).

28. K. Endres, Targeted issue messages and voting behavior. *Am. Polit. Res.* **48**, 317–328 (2020).

29. E. D. Hersh, B. F. Schaffner, Targeted campaign appeals and the value of ambiguity. *J. Polit.* **75**, 520–534 (2013).

30. M. Jacobs-Harukawa, Does Microtargeting Work? Evidence from an Experiment during the 2020 United States Presidential Election. *Github.* https://muhark.github.io/static/docs/harukawa-2021-microtargeting.pdf. Accessed 10 September 2022.

31. B. Zarouali, T. Dobber, G. De Pauw, C. de Vreese, Using a personality-profiling algorithm to investigate political microtargeting: Assessing the persuasion effects of personality-tailored ads on social media. *Commun. Res.* **49**, 1066–1091 (2022).

32. Open Secrets, Fundraising totals: Who raised the most? *OpenSecrets* (2022). https://www.opensecrets.org/elections-overview/fundraising-totals. Accessed 10 September 2022.

33. J. Silberman, Persuasion is the most important part of messaging. Here's the right way to test for it. *Civis Analytics* (2022). https://www.civisanalytics.com/blog/persuasion-is-the-most-important-part-of-messaging-heres-the-right-way-to-test-for-it/. Accessed 10 September 2022.

34. Blue Rose Research, Blue Rose Research testing services (2022). https://wiki.staclabs.io/steven_mccarty/blue_rose_testing_overview_(stac_labs).pdf. Accessed 10 September 2022.

35. M. Wiles, How to know your tv ad will work... before you launch it (2022). https://www.swayable.com/insights/how-to-know-your-tv-ad-will-work-before-you-launch-it/. Accessed 10 September 2022.

36. Data for Progress, Democrats on offense: Messages that Win (2022). https://www.filesforprogress.org/memos/dfp_democrats_on_offense.pdf. Accessed 10 September 2022.

37. L. Hewitt, B. M. Tappin, Rank-heterogeneous effects of political messages: Evidence from randomized survey experiments testing 59 video treatments (2022), 10.31234/osf.io/xk6t3.

38. C. Wittenberg, B. M. Tappin, A. J. Berinsky, D. G. Rand, The (minimal) persuasive advantage of political video over text. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2114388118 (2021), 10.1073/pnas.2114388118.

39. J. Graham *et al.*, Mapping the moral domain. *J. Pers. Soc. Psychol.* **101**, 366–385 (2011).

40. M. Feinberg, R. Willer, Moral reframing: A technique for effective and persuasive communication across political divides. *Soc. Personal. Psychol. Compass* **13**, e12501 (2019).

41. S. Athey, J. Tibshirani, S. Wager, Generalized random forests. *Ann. Statist.* **47**, 1148–1178 (2019).

42. T. Dobber, N. Metoui, D. Trilling, N. Helberger, C. de Vreese, Do (microtargeted) deepfakes have real effects on political attitudes? *Int. J. Press/Politics.* **26**, 69–91 (2021).

43. J. Green *et al.*, Using general messages to persuade on a politicized scientific issue. *Br. J. Polit. Sci.*, 1–9 (2022), https://www.ipr.northwestern.edu/our-work/working-papers/2021/wp-21-45.html.

44. K. Lunz Trujillo, M. Motta, T. Callaghan, S. Sylvester, Correcting misperceptions about the MMR vaccine: Using psychological risk factors to inform targeted communication strategies. *Polit. Res. Q.* **74**, 464–478 (2021).

45. S. C. Matz, M. Kosinski, G. Nave, D. J. Stillwell, Psychological targeting as an effective approach to digital mass persuasion. *Proc. Natl. Acad. Sci. U.S.A.* **114**, 12714–12719 (2017).

46. D. Eckles, B. R. Gordon, G. A. Johnson, Field studies of psychologically targeted ads face threats to internal validity. *Proc. Natl. Acad. Sci. U.S.A.* **115**, E5254–E5255 (2018).

47. J. D. Teeny, J. J. Siev, P. Briñol, R. E. Petty, A review and conceptual framework for understanding personalized matching effects in persuasion. *J. Consum. Psychol.* **31**, 382–414 (2021).

48. D. M. Kahan, "The politically motivated reasoning paradigm, part 1: What politically motivated reasoning is and how to measure it" in *Emerging Trends in the Social and Behavioral Sciences* (Wiley, 2016), pp. 1–16. https://onlinelibrary.wiley.com/doi/abs/10.1002/9781118900772.etrds0417.

49. M. Aggarwal *et al.*, A 2 million-person, campaign-wide field experiment shows how digital advertising affects voter turnout. *Nat. Hum. Behav.* **7**, 332–341 (2023).

50. J. Blumenau, B. E. Lauderdale, The variable persuasiveness of political rhetoric. *Am. J. Polit. Sci.*, **0**, 1–16 (2021), http://benjaminlauderdale.net/files/papers/BlumenauLauderdalePersuasion.pdf.

51. J. N. Druckman, A framework for the study of persuasion. *Annu. Rev. Polit. Sci.* **25**, 65–88 (2021). 10.2139/ssrn.3849077.

52. D. Broockman, J. Kalla, When and why are campaigns' persuasive effects small? Evidence from the 2020 US presidential election *Am. J. Polit. Sci.*, **0**, 1–17 (2022). 10.31219/osf.io/m7326.

53. D. Chong, J. N. Druckman, Framing public opinion in competitive democracies. *Am. Polit. Sci. Rev.* **101**, 637–655 (2007).

54. B. Nyhan, E. Porter, T. J. Wood, Time and skeptical opinion content erode the effects of science coverage on climate beliefs and attitudes. *Proc. Natl. Acad. Sci. U.S.A.* **119**, e2122069119 (2022).

55. E. Herder, S. Dirks "User Attitudes Towards Commercial Versus Political Microtargeting" in *Adjunct Proceedings of the 30th ACM Conference on User Modeling, Adaptation and Personalization*, (Association for Computing Machinery, New York, NY, 2022), UMAP '22 Adjunct, pp. 266–273. https://doi.org/10.1145/3511047.3538027.

56. T. Dobber, C. de Vreese, Beyond manifestos: Exploring how political campaigns use online advertisements to communicate policy information and pledges. *Big Data Soc.* **9**, 20539517221095430 (2022).

57. F. J. Borgesius *et al.*, Online political microtargeting: Promises and threats for democracy. *Utrecht Law Rev.* **14**, 82–96 (2018).

58. M. M. Franz, T. N. Ridout, Does political advertising persuade? *Polit. Behav.* **29**, 465–491 (2007).

59. R. Iyer, S. Koleva, J. Graham, P. Ditto, J. Haidt, Understanding libertarian morality: The psychological dispositions of self-identified libertarians. *PLoS One.* **7**, e42366 (2012).

60. A. Coppock, O. McClellan, Validating the demographic, political, psychological, and experimental results obtained from a new source of online survey respondents. *Res. Politics* **6**, 2053168018822174 (2019).

61. B. M. Tappin, C. Wittenberg, L. B. Hewitt, A. J. Berinsky, D.G. Rand, Data from new randomized survey experiment. *Open Science Framework*. https://osf.io/t3dhe/. Deposited 30 May 2023.