ECONtribute
MARKETS & PUBLIC POLICY

ECONtribute
Discussion Paper No. 274

Political Activists are Not Driven by Instrumental Motives: Evidence from Two Natural Field Experiments

Anselm Hager
Lukas Hensel
Johannes Hermle
Christopher Roth

January 2024

www.econtribute.de

UNIVERSITÄT BONN
Universität zu Köln

Funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy – EXC 2126/1-390838866 is gratefully acknowledged.

Political Activists are Not Driven by Instrumental Motives: Evidence from Two Natural Field Experiments*

Anselm Hager Lukas Hensel

Johannes Hermle Christopher Roth†

Abstract

Are political activists driven by instrumental motives such as making a career in politics or mobilizing voters? We implement two natural field experiments in which party activists are randomly informed that canvassing is i) effective at mobilizing voters, or ii) effective for enhancing activists' political careers. We find no effect of the treatments on activists' intended and actual canvassing behavior. The null finding holds despite a successful manipulation check and replication study, high statistical power, a natural field setting, and an unobtrusive measurement strategy. Using an expert survey, we show that the null finding shifted Bayesian posterior beliefs about the treatment's effectiveness toward zero. The evidence thus casts doubt on two popular hypothesized instrumental drivers of political activism—voter persuasion and career concerns—and points toward expressive benefits as more plausible motives.

*The study was pre-registered at the AEA RCT registry (number 0002417).

†Anselm Hager, Humboldt-Universität zu Berlin, anselm.hager@gmail.com; Lukas Hensel, University of Oxford, lukas.hensel@economics.ox.ac.uk; Johannes Hermle, University of California, Berkeley, j.hermle@berkeley.edu; Christopher Roth, University of Cologne and ECONtribute. Roth: Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy– EXC 2126/1-390838866. The research described in this article was approved by the University of Oxford Department of Economics Research Ethics Committee, Ref No: R52651/RE004.

What motivates citizens to become active in politics? Scholars have long contended that political engagement is partly a product of instrumental[1] considerations. While voting is hard to rationalize on instrumental grounds (Gelman, Silver and Edlin, 2012), the case is quite different for engagement in a political campaign. Here, party supporters may well be motivated by their own effect on the election, given that local races are often close. Participation in a campaign is also a quintessential means to advance one's political career (Black 1972, 146, Abramowitz, McGlennon and Rapoport 1983, 1008, Fox and Lawless 2005, 653). But, are such instrumental, outcome-oriented considerations actual drivers of political activism?

This paper studies whether party activists can be motivated on instrumental grounds. We present evidence from two field experiments conducted with a major European party. Study 1 took place during a national election and explored the role of individual effectiveness. We randomized whether 1,184 party activists received true information about the effectiveness of canvassing in mobilizing voters. Study 2 took place during a state-level election and assessed the role of career ambitions. Here, we randomized whether 1,885 party activists received information that key party leaders had once been active canvassers, accelerating their political careers. Our outcomes are respondents' intended and actual canvassing behavior, which we unobtrusively measure via the party's canvassing app.

We find that neither the treatment highlighting the effectiveness of canvassing, nor the treatment underlining that canvassing can advance one's career affected party supporters' political engagement. Intended and actual canvassing measured over two weeks (Study 1) and four weeks (Study 2) are highly similar across the treatment and control groups. The null findings are informative for six reasons. First, both studies are well-powered. Second, we find convergent evidence for intentions and behavior. Third, both studies included a successful

[1]We define "instrumental considerations" as a focus on the outcome of an activity (here, persuading voters or advancing one's career). By contrast, "expressive considerations" focus on the process of the activity itself (e.g., meeting friends, expressing one's identity). We note that this dichotomy is well-established in studies of voting behavior (dating back to the foundational accounts of Downs 1957; Butler and Stoke 1974; and Fiorina 1977) as well as the vast literature on party identification (Franklin and Jackson 1983; Green, Palmquist and Schickler 2008; Greene 2002), but has also found its way into studies of activism (Huddy, Mason and Aarøe, 2015).

manipulation check. Fourth, the results are unlikely to be affected by social desirability bias since we use unobtrusive behavioral data coupled with information sent out on behalf of the party. Fifth, we find no meaningful effect heterogeneity. Sixth, we collected expert and lay people's prior beliefs on the likely effects and show that our data moves posterior beliefs toward zero. Taken together, the evidence thus casts doubt that political activists can be motivated on instrumental grounds, pointing to expressive motives as the more likely drivers (Hager et al., 2022).

Motivation

Why do citizens become politically active? Scholars of political behavior argue that political engagement brings about both instrumental as well as expressive benefits (Gordon and Babchuk, 1959). Instrumental benefits accrue from the outcome of political engagement (Hansen, 1985). Expressive benefits accrue from the process of activism itself (Gerber et al., 2016). Given that political activism provides a public good, scholars commonly reject the notion that instrumental motives drive political engagement (Olson, 1965; Gerber, Green and Larimer, 2008). This holds particularly true for voting, given that one vote is unlikely to sway an election (Gelman, Silver and Edlin, 2012). The empirical evidence, however, remains mixed (Enos and Fowler, 2014). Bursztyn et al. (2023), for instance, find that turnout is significantly higher in close races, while Gerber et al. (2020) find no such link.

Whether instrumental motives explain other, more high-stake forms of political engagement remains an open question. The lack of evidence is surprising given that the impact of volunteering in a political campaign is orders of magnitude larger than that of casting a ballot. A good canvasser convinces hundreds of people to vote. In the setting studied in this paper, roughly 40 percent of races were close (within a 5-point margin). Canvassers may thus well have had a noticeable impact. A second instrumental motive specific to participation in campaigns is the advancement of one's political career.[2] Active engagement on the campaign

[2]NB: advancing one's career is *not* a public good, but it still constitutes an instrumental goal.

trail signals commitment to the party, strengthens activists' networks and affords grassroots credibility, which may translate into being nominated and elected to party-internal or public office.

More so than voting, political activism may therefore well be driven by instrumental considerations. Indeed, Figure A.1 provides tentative support that instrumental considerations spark activism: We show a clear correlation between political activists' canvassing effort and their beliefs about i) the effectiveness of canvassing (Figure A.1a) and ii) the importance of canvassing for activists' political careers (Figure A.1b).[3] What is more, in our setting survey-evidence confirms that party activists do believe that canvassing is effective at convincing voters as well as advancing one's political career.[4] And, an analysis of 3.5 million Tweets from the country of study further showcases that canvassing is endorsed by key party leaders, pointing to its effectiveness at persuading voters and advancing one's career (see Section A). This evidence, however, is not causal. To study both potential instrumental drivers—persuasion and career ambitions—we, therefore, implemented two natural field experiments, which we describe next.

Study 1: Are activists driven by a desire to mobilize voters?

Setting and Sample To study whether political activists are driven by a desire to mobilize voters, we cooperated with a large European party and implemented a field experiment during a general election campaign. On behalf of the party, we invited the party's core list of supporters to participate in an online survey via email.[5] The population can best be described

[3]Activists are fundamentally interested in political careers. 79% of activists have "career concerns" in that they either currently hold or are interested in holding a political office.

[4]On average, activists in study 1 think that 28.5% of non-voters reached via canvassing can be convinced to turnout. For study 2, 69% of party activists in the control group think that participating in canvassing is *rather useful*, *useful*, or *very useful* for a political career but only 16% think that it is *very useful*. This suggests that there is room for increasing this perception further through experimental variation.

[5]The same list of supporters was used to recruit participants for the main experiments in Hager et al. (2021) and Hager, Hensel, Hermle and Roth (2023) which were conducted about five weeks prior to this experiment. Hence, 43% of the participants participated in one of the previous experiments. Treatments where fully cross-randomized and controlling for past participation and treatment status does not affect the results.

as "activists," i.e., party supporters that were already active in the campaign or aspired to become active (details on the setting are in Section B). As such, we study both the extensive margin (whether to become active) as well as the intensive margin (whether to increase one's activism). The email and survey used a party template and no reference was made to the research team. 1,184 activists agreed to participate (response rate of 2.4 percent). The survey was distributed two weeks ahead of the election. The descriptive statistics of the sample are in Table A2. The sample broadly maps onto the party's membership statistics, though the sample is younger and includes more men. The sample also, naturally, had high prior levels of political activism with 43 percent having canvassed before and 25 percent having canvassed in the current campaign.

Treatment After administering eight descriptive questions (see Figure A.2), respondents were asked: "*Imagine a canvasser who talks to 100 non-voters. What do you think: How many of these 100 non-voters can the canvasser convince to go vote?*" Thereafter, half the sample was randomly[6] assigned to true information stating that a typical canvasser convinces 16 out of 100 non-voters to turn out.[7] To ease interpretation, the treatment screen also included a figure, which compared respondents' estimates to the true number. The control group was not given this information screen (though the question about their prior beliefs was asked). After the treatment group was given the information, all respondents, including the control group, were asked to imagine a typical canvasser "*in the party's current campaign*" who had spoken to 100 non-voters, and asked again how many non-voters the canvasser could convince to vote. The latter question serves as our measure for respondents' posterior beliefs, i.e., our manipulation check.

[6]Random assignment was implemented during the survey and Table A3 showcases excellent balance. The only significant difference is years of party membership. This variable is included in the pre-registered set of controls and Table A4 shows that our treatment effects do not vary by years of membership.

[7]This is extrapolated from Gerber and Green 2000 who find that canvassing increases in turnout by 8.7 percentage points. This is equivalent to approximately 16% of the 55.5% of potential voters that do not turn out in the control group.

Outcome After eliciting respondents' posterior beliefs, we asked them whether they planned to canvass at all (extensive margin), and, if so, on how many days (intensive margin). After the survey, we also measured respondents' actual canvassing behavior until the election using the party's canvassing smartphone application, which respondents used to register knocked doors (details are in Section B). The app data allows us to unobtrusively measure if respondents went canvassing and, if so, on how many days as well as on how many doors they knocked on. Importantly, the behavioral data was collected before and after the treatment was administered during the entire campaign. We thus include all canvassing activity from the day of the survey until the election and also control for prior canvassing activity. As Table A2 shows, 15 percent of respondents went canvassing, knocking on an average of 32 doors. Finally, we also combine the five outcomes into a standardized canvassing index.[8]

Model To estimate the effect of the treatment on respondents' political engagement, we pre-registered the following model: $y_i = \pi_0 + \pi_1 \mathrm{T}_i + \zeta^T \mathbf{X_i} + \varepsilon_i$.[9] Where y_i is the canvassing outcome of interest. T_i is a dummy taking the value one treated individuals (effectiveness information) and zero otherwise. $\mathbf{X_i}$ is the set of pre-registered controls, which are reported in Table A2. ε_i is the error term. We report robust standard errors.

Manipulation check Column 2 in Table 1 demonstrates that the treatment significantly affected respondents' posterior beliefs. In the aggregate sample, the treatment group reports a posterior belief that is five percentage points higher compared to the control group. More importantly, Table 1 demonstrates that treated respondents who initially underestimated the effectiveness of canvassing, shifted their posterior belief upward by an average of 3 points. By contrast, treated respondents who overestimated the effectiveness of canvassing, shifted their posterior belief downward by an average of 14 points.

[8] Reassuringly, canvassing intentions and behavior are strongly correlated (see Table A5). The distribution of intended and actual days are displayed in Figure A.6.

[9] The pre-analysis plan is ANONYMIZED.

Table 1: Impact of effectiveness treatment on canvassing intentions and behavior

	Manipulation check	**Intentions**		**Behavior**			**Index**
	Persuasion rate	Plans canvassing	Days planned	Any canvassing	Days canvassed	Doors knocked	
Panel A: Pooled sample							
Treatment	-8.041***	-0.008	-0.048	0.011	0.032	4.933	0.017
	(1.043)	(0.025)	(0.165)	(0.016)	(0.079)	(5.943)	(0.046)
Observations	1,138	1,138	1,138	1,157	1,157	1,157	1,138
Panel B: Underestimators							
Treatment	2.540***	-0.021	-0.202	0.024	0.001	1.214	-0.011
	(0.602)	(0.042)	(0.251)	(0.025)	(0.135)	(8.248)	(0.072)
Observations	432	432	432	436	436	436	432
Panel C: Overestimators							
Treatment	-14.266***	0.002	0.053	0.002	0.053	6.404	0.033
	(1.324)	(0.031)	(0.221)	(0.022)	(0.099)	(8.351)	(0.061)
Observations	706	706	706	721	721	721	706

Notes: Table 1 reports results from an OLS regression of the indicated outcome on the effectiveness treatment dummy. Robust standard errors are given in parentheses. All pre-registered control variables are included * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. *Intentions* capture whether a respondent plans to do any canvassing and, if so, on how many days they plan to go canvassing. *Behavior* refers to actual canvassing behavior measured through a smartphone app, capturing whether a respondent engaged in any canvassing and, if so, how many days they went canvassing and on how many doors they knocked. The *Index* combines the intentions and behavior-based data into a standardized index.

Results Did the effectiveness treatment change respondents' political engagement? Panel A in Table 1 shows that treated individuals did *not* show different canvassing intentions or behavior.[10] Given that we specifically shifted respondents' beliefs, in Panels B and C in Table 1 we analyze the behavior of under- and overestimators separately. The Table confirms our null finding. Underestimators—who learned that canvassing is more effective than previously thought—are not more likely to canvass, and vice versa.[11]

[10]To ascertain robustness, Table A6 shows that the null effect is robust to omitting covariates. Results are also robust to controlling for the number of days canvassed prior to the experiment (Table A7). Table A8 shows that we also do not observe effects on the day or during the week after the treatment, suggesting that decay in treatment effects over time does not explain the results. Figure A.4 confirms this by plotting the cumulative fraction of individuals who canvassed within a ten-day window of the experiment.

[11]The null effect is unlikely to be driven by the fact that the behavioral outcomes are difficult to move. Other studies in the same context with similar survey-based information interventions and very similar outcomes collected through the same application find significant treatment effects on behavioral outcomes (Hager et al., 2021; Hager, Hensel, Roth and Stegmann, 2023; Hager, Hensel, Hermle and Roth, 2023). We should caution, however, that the present studies are not powered to detect small changes in behavior: The minimum detectable effect size at 80 percent power and 5 percent test for the binary engagement indicator, for instance, is 4.5 percentage points.

Heterogeneity To assess heterogeneity, we focus on the combined canvassing index and use all available pre-specified control variables to construct subgroups. Figure A.7a plots the coefficients of the treatment dummy and the interaction of the treatment with the respective covariates. Instrumental motives should be particularly strong for individuals who expect a close election as they are more likely to be pivotal. However, we do not find any heterogeneity by expected election closeness. Another potential source of heterogeneity is whether supporters expect their own party or the main competing party to knock on more doors. Again, however, we do not find any heterogeneity along this dimension. More broadly, Figure A.7a shows that there is no significant effect heterogeneity for any of the subgroups.

Study 2: Are activists driven by a desire to advance their careers?

Setting and Sample A year later, we implemented a second field experiment with the same party to study the question of whether activists are motivated by advancing their political careers. This time, the experiment took place during a state-level electoral campaign for the state parliament (further details are in Figure A.3 and in Section B). As in the effectiveness study, we sent out a survey on behalf of the party to its list of supporters, using the same unobtrusive template, and invited (potential) activists to participate in an online survey. 1,885 activists agreed to participate (response rate of 4.4 percent). The survey was distributed four weeks before the election. The descriptive statistics of the sample are given in Table A2. Compared to the first study, the sample of the career study is older and was, on average, less active during the campaign.

Treatment After administering a similar set of background questions, respondents were randomly[12] assigned to information that key party leaders are proud canvassers, which had accelerated their political careers. The specific script read as follows: "*Many members of the [Party] board are proud canvassers. [Name of party leader 1], for instance, was an active*

[12]Balance is demonstrated in Table A9. The control group has slightly more years of party membership, which we control for in the model and demonstrate no heterogeneous effects in Table A10.

canvasser during his youth, which allowed him to make valuable experiences for his political career. [Name of party leader 2], too, was one of the party's first canvassers." The control group was not given any information. In this study, we did not elicit individuals' prior and posterior belief, which did not prove sensible without a specific estimate of the form elicited in the effectiveness study. Rather, we included a distinct manipulation question at the end of the survey. We asked: "*What do you think: How useful is canvassing to make a career in politics?*" The answer choices—scored on a 6-point scale—ranged from "not useful at all" to "very useful."

Outcome After administering the treatment, we collected the exact same outcomes as in the effectiveness study, i.e., canvassing intentions and, after the survey, the behavioral canvassing data until the end of the campaign.

Manipulation check Using the same model as in Study 1, Table 2 demonstrates that the treatment increased respondents' beliefs that canvassing is an essential tool to advance one's political career: treated respondents score 0.11 points higher on the 6-point scale.[13]

Results Did the career prime increase respondents' political engagement? Table 2 shows that respondents did *not* report different canvassing intentions and also did not change their behavior relative to the control group. All estimated coefficients are close to zero. The null finding holds across both the extensive margin and the intensive margin. The coefficients are also highly similar across the self-reported survey data and the behavioral data. The evidence thus casts doubt on whether career ambitions are a meaningful driver of political engagement—despite activists stating that they *are* motivated by advancing their careers (more below; and in Figure A.8). Importantly, Study 2 has a minimum detectable effect size of 0.10 SD for the combined canvassing index, which allows us to rule out a substantively

[13]We replicated the priming experiment using a sample of 600 lay people recruited using prolific—a sample arguably *less* susceptible to the prime (see Table A18 for a description of the sample). Reassuringly, as Table A21 shows, we find a strong treatment effect of the career prime on respondents' beliefs that canvassing advances activists' careers and, at the same time, no effect on canvassing intentions.

meaningful effect. Again, results are robust to omitting covariates (see Table A11).

Table 2: Impact of career treatment on canvassing

	Manipulation check	Intentions		Behavior			Index
		Any canvassing	Days canvassed	Any canvassing	Days canvassed	Doors knocked	
Treatment (Career prime)	0.119* (0.064)	-0.001 (0.018)	-0.081 (0.149)	-0.006 (0.006)	-0.042 (0.041)	-0.891 (0.617)	-0.041 (0.036)
Observations	1819	1,881	1,881	1,885	1,885	1,885	1,881

Notes: The Table reports results from an OLS regression of the indicated outcome on the effectiveness treatment dummy. Outcomes are defined as in Table 1. Robust standard errors are given in parentheses. Demographic control variables are included. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Heterogeneity Do we observe treatment effects for subgroups where the manipulation check is particularly strong or for other theoretically plausible subgroups? Tables A12 and A13 show that the manipulation check worked very well among men and respondents with initially lower career concerns (though the difference in treatment effects is not significant for the latter). But, the Tables also demonstrate that even for these sub samples we do not observe significant treatment effects. More broadly, Figure A.7b assesses treatment effect heterogeneity across all pre-registered covariates for the standardized canvassing index. The Figure confirms that there is no significant treatment effect heterogeneity for any subgroup.

Informativeness of the Null Findings

Statistical power Are our experiments well-powered to detect reasonable effect sizes? To tackle this question, we compare estimated effects and minimum detectable effect sizes in other studies using survey-administered experiments to study political behavior. Table A17 shows that the statistical power in our study compares well to the overall literature and to studies in the same context. The minimum detectable effect sizes on the canvassing index are 0.129 standard deviations (study 1) and 0.101 standard deviations (study 2). This compares favorably with an average detectable effect size of 0.183 across estimates identified in Table A17. The Minimum detectable effect sizes are also smaller than the average estimated effect

size in the literature (0.189 standard deviations) even though this also includes insignificant effects.[14]

Learning Do our null results provide new information about the motives of political activists? To answer this question, we implemented surveys with three relevant samples: 1,107 supporters of the party, 600 lay people, and 54 experts (political scientists working on activism).[15] We first asked all respondents why they believe party supporters canvass. We asked this question to explore if the relevant samples believe that the primed instrumental motives could, conceivably, be increasing activism. As Figure A.8 shows, 71 percent of activists, 86 percent of lay people and 63 percent of experts believe that canvassers are, indeed, driven by the desire to persuade voters. Similarly, 43 percent of activists, 51 percent of lay people and 54 percent of experts believe that canvassers are motivated by the desire to improve their careers (multiple choices were possible).

Even if activists report that they are motivated by persuasion and career considerations, a critic might object that one would still not expect our specific treatments to affect canvassing.[16] To address this concern, we collected expert and lay people's beliefs about the likely treatment effects in order to assess to what degree our results provide new information. In particular, we explained both experimental designs to both samples, provided them with the respective control group means, and then asked them what treatment effects they expected. We then integrated the resulting distribution of prior beliefs with the experimental estimates using Bayesian statistics (details are in Section D).

The results of the Bayesian analysis for the expert sample are presented in Table A19. Two results stand out. First, our evidence provided new information about the expected

[14]Minimum detectable effect sizes in terms of the control mean are also smaller than the average in the literature (30.8% vs 191.6%). The average estimated relative effect is 199.8% of the control mean, substantially above the estimated minimum detectable effect size (taking out two outliers still yields an average estimate of 56.1%). Comparing absolute effect sizes of binary outcomes we find a range of estimated effects of 0.1 to 23 percentage points (2.6 to 23 among statistically significant results). The minimum detectable effect sizes of 4.5 and 1.7 percentage points in this paper fall comfortably in the range of the observed effects.

[15]Details are in the SI

[16]The successful manipulation checks (one of which we replicated once more; see Footnote 5) arguably assuages this criticism.

effect size for the career experiment. The expected treatment effects based on the posterior belief distribution are much closer to zero than those based on the prior belief distribution.[17] For instance, experts believed that the career treatment would raise intentions by 5.2 percentage points, on average. Updating these prior beliefs with the experimental estimates leads to a posterior belief about the ATE of just 0.9 percentage points. Second, the posterior distributions of beliefs about the treatment effects are much tighter than the prior distributions for both studies. For example, the probability that the treatment effect on actual behavior of the effectiveness experiment falls between -2 and 2 percentage points increased from 46.8% to 75.8% for canvassing behavior in the effectiveness experiment. This underlines that our results not only provided information about the mean, but also increased experts' and lay people's certainty about the effect of the persuasion and career primes on political activism.

Conclusion

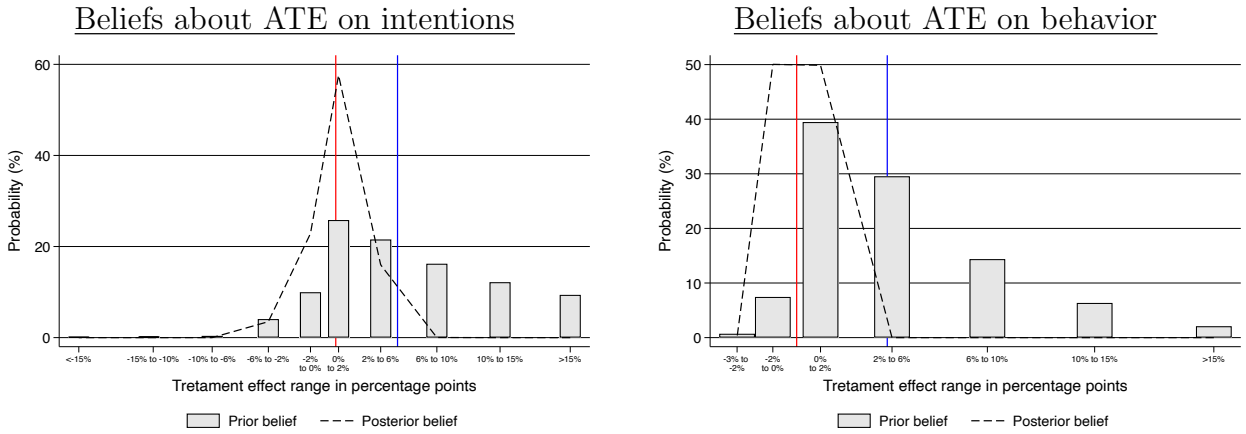
What can the two null findings teach us about political engagement? If taken at face value, scholars are well-advised to continue to direct their focus on "expressive" and social motives. Instrumental motivations that are closely linked to the outcome of political activism are seemingly less relevant. Importantly, this finding held across two very different "instrumental" treatments, which—though both primed outcomes (the election or one's career)—differed in the degree to which the outcome was a public or private good. That is, study 1 primed a public good (affecting an election), while study 2 primed a private good (affecting one's career). The degree to which the good is public (and thus plagued by free-riding) therefore did not seem to play a role. In both studies we found no effect for the instrumental treatments.

We must caution, however, that our study only illuminated the drivers of engagement in one specific political context for one form of activism: canvassing. A skeptic may therefore ask whether instrumental benefits do matter when one studies different parties, different forms of engagement or more local elections. Reassuringly, we found no evidence that the

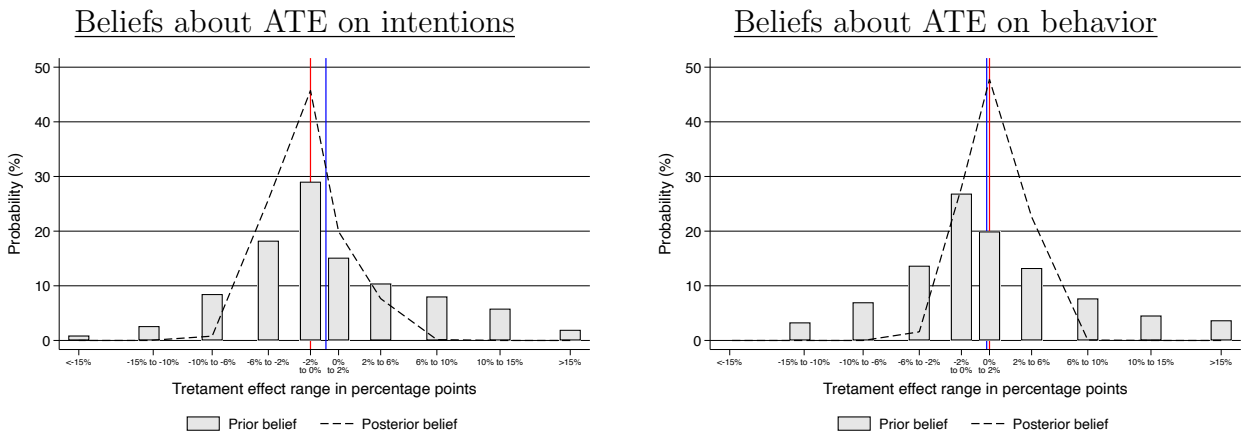
[17]The pattern is even more evident for the laypeople sample where participants expected larger effect sizes for both the career and the effectiveness experiment (see Table A20).

Figure 1: Distribution of expert beliefs about average treatment effects (ATEs)

Panel A: Belief updating in the career experiment



Panel B: Belief updating in the effectiveness experiment



Notes: The Figures display averaged expert beliefs about treatment effects. Grey bars indicate the averaged prior beliefs calculated by averaging the probability mass experts put on each interval. Dashed lines indicate the averaged posterior beliefs obtained by updating averaged prior beliefs using Bayes' rule with a distribution of treatment effects obtained through bootstrapping (10,000 repetitions). Panel A displays beliefs about the effect of the career treatment. Panel B displays beliefs about the effect of the effectiveness treatment. Both panels show effects on canvassing intentions and behavior.

treatments affected other forms of engagement, namely, social media activity (see Table A16). What is more, the fact that we do not find any effect heterogeneity (including for perceived election closeness), makes it difficult to put much trust in instrumental explanations. We should also underline that we found no effects across two different electoral campaigns, in a large Western-European democracy, which uses *both* majoritarian single-member districts as well as proportional party lists. Our evidence thus arguably offers a moderate degree of generalizability.

References

Abramowitz, Alan, John McGlennon and Ronald Rapoport. 1983. "The Party Isn't Over: Incentives for Activism in the 1980 Presidential Nominating Campaign." *The Journal of Politics* 45(4):1006–1015.

Black, Gordon S. 1972. "A Theory of Political Ambition: Career Choices and the Role of Structural Incentives." *American Political Science Review* 66(1):144–159.

Bursztyn, Leonardo, Davide Cantoni, Patricia Funk, Felix Schönenberger and Noam Yuchtman. 2023. "Identifying the Effect of Election Closeness on Voter Turnout: Evidence from Swiss Referenda." *Journal of the European Economic Association* p. jvad038.

Butler, David and Donald Stoke. 1974. *Political Change in Britain: Basis of Electoral Choice*. Springer.

Cantoni, Davide, David Y. Yang, Noam Yuchtman and Y. Jane Zhang. 2019. "Protests as Strategic Games: Experimental Evidence from Hong Kong's Antiauthoritarian Movement." *Quarterly Journal of Economics* 134(2):1021–1077.

Corstange, Daniel. 2016. "Anti-American behavior in the middle east: Evidence from a field experiment in Lebanon." *Journal of Politics* 78(1):311–325.

Downs, Anthony. 1957. "An economic theory of political action in a democracy." *Journal of Political Economy* 65(2):135–150.

Enos, Ryan D and Anthony Fowler. 2014. "Pivotality and turnout: Evidence from a field experiment in the aftermath of a tied election." *Political Science Research and Methods* 2(2):309–319.

Fiorina, Morris P. 1977. "An outline for a model of party choice." *American Journal of Political Science* 21(3):601–625.

Fox, Richard L. and Jennifer L. Lawless. 2005. "To Run or Not To Run for Office: Explaining Nascent Political Ambition." *American Journal of Political Science* 49(3):642–659.

Franklin, Charles H and John E Jackson. 1983. "The dynamics of party identification." *American Political Science Review* 77(4):957–973.

Gelman, Andrew, Nate Silver and Aaron Edlin. 2012. "What Is the Probability Your Vote Will Make a Difference?" *Economic Inquiry* 50(2):321–326.

Gerber, Alan., Gregory Huber and Albert Fang. 2023. "Voting behavior is unaffected by subtle linguistic cues: evidence from a psychologically authentic replication." *Behavioural Public Policy* 7(2):380–394.

Gerber, Alan, Mitchell Hoffman, John Morgan and Collin Raymond. 2020. "One in a million: Field experiments on perceived closeness of the election and voter turnout." *American Economic Journal: Applied Economics* 12(3):287–325.

Gerber, Alan S. and Donald P. Green. 2000. "The Effects of Canvassing, Telephone Calls, and Direct Mail on Voter Turnout: A Field Experiment." *American Political Science Review* 94(3):653–663.

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.

Gerber, Alan S., Gregory A. Huber, David Doherty and Conor M. Dowling. 2016. "Why People Vote: Estimating the Social Returns To Voting." *British Journal of Political Science* 46(2):241–264.

Gordon, C. Wayne and Nicholas Babchuk. 1959. "A Typology of Voluntary Associations." *American Sociological Review* 24(1):22–29.

Green, Donald, Bradley Palmquist and Eric Schickler. 2008. *Partisan hearts and minds*. New Haven, CT: Yale University Press.

Greene, Steven. 2002. "The Social Psychological Measurement of Party ID." *Political Behavior* 24(3).

Hager, Anselm, Lukas Hensel, Christopher Roth and Andreas Stegmann. 2023. "Voice and Political Engagement: Evidence From a Natural Field Experiment." *Review of Economics and Statistics* forthcoming.

Hager, Anselm, Lukas Hensel, Johannes Hermle and Christopher Roth. 2021. "Does Party Competition Affect Political Activism?" *Journal of Politics* 83(4):1681–1694.

Hager, Anselm, Lukas Hensel, Johannes Hermle and Christopher Roth. 2022. "Group Size and Protest Mobilization across Movements and Countermovements." *American Political Science Review* 116(3):1051–1066.

Hager, Anselm, Lukas Hensel, Johannes Hermle and Christopher Roth. 2023. "Political Activists as Free Riders: Evidence From a Natural Field Experiment." *Economic Journal* 133:2068–2084.

Hansen, John M. 1985. "The Political Economy of Group Membership." *American Political Science Review* 79(1):79–96.

Harrison, Glenn W and John A List. 2004. "Field experiments." *Journal of Economic literature* 42(4):1009–1055.

Huddy, Leonie, Lilliana Mason and Lene Aarøe. 2015. "Expressive partisanship: Campaign involvement, political emotion, and partisan identity." *American Political Science Review* 109(1):1–17.

Kreft, Brynde and Kate Orkin. 2020. "Everybody Loves a Winner: A Field Experiment

Providing Information on Polls in South Africa." *Working Paper* .

URL: *https://sites.google.com/site/kateorkin/cv*

Olson, Mancur. 1965. *The Logic of Collective Action: Public Goods and the Theory of Groups.* New York, NY: Schocken Books.

Political Activists are Not Driven by Instrumental Motives: Evidence from Two Natural Field Experiments – Online Appendices Not for Publication

A Twitter analysis

To get a descriptive handle on the relevance of canvassing in the country of study in general as well as a hypothesized relation between canvassing and instrumental drivers of activism (persuading voters and advancing one's career), we analyzed 3.5 million Tweets from members of Parliament.

Table A1: Share of Tweets by own party

Statistic	N	Mean	St. Dev.	Min	Max
percentage	6	16.668	24.147	0.510	55.840

Notes: Table A1 reports the share of Tweets on canvassing sent by our own party relative to competitors.

B Setting and Population

B.1 Setting

To study whether political activists are driven by instrumental motives—in particular, beliefs about their ability to persuade voters and canvassing's ability to improve their careers—we implemented two separate field experiments in a large western European democracy. The country we study implements general, state-level as well as municipal elections. The elections are usually competitive, at least the two elections we studied. Seats to the respective parliaments are allocated using a mix of proportional representation as well as majoritarianism. In particular, in the two elections we studied, citizens had to cast two votes. The first vote was for the local MP, which must receive the plurality of votes (majoritarianism). The second vote is for a party list (proportional representation). Seats in the national or state-level parliaments are then given to all winning MPs with the remaining seats allocated so as to preserve the proportionality of second votes.

We cooperated with one of the two main competing parties during the final months of the electoral campaign. The two studies took place in two different elections. The first study (effectiveness) took place during a general election. The election was widely deemed competitive, though the final result was not razor thin. Importantly, both major parties considered themselves as competitors, but there were a number of additional parties who ultimately made it into parliament. The second study took place during a state-level election. Here, too, the election was widely deemed competitive, though the two major parties ended up rather far apart on election day. As a matter of fact, the race in the state election ended up as a three-way race between three parties, underscoring the competitiveness of the election. Importantly, the state is highly representative of the country of study overall.

We should stress that the context—i.e., a survey with potential activists—within the country of study was not unusual. While the main parties, historically speaking, did not heavily engage in canvassing (unlike, e.g., the U.S. democrats), canvassing was not an entirely new phenomenon. That said, the two campaigns were novel in that they—for the first time—introduced smartphones as a way to record knocked doors. What is more, the party advertised the smartphone app heavily and promoted canvassing as an effective campaign tool through internal communication channels as well as via the media in both elections. The party instructed and encouraged all canvassers to download and use the app. The party also provided the canvassers with training workshops in which party supporters were encouraged to use the app.

To further boost take-up, the smartphone app included a number of gamification elements. Doors were not recorded automatically, but needed to be locked in by canvassers. The unobtrusive, geocoded app data therefore provide a unique lens into the actual canvassing activities of respondents. In the general election (study 1), all of the country's well over 200 constituencies saw canvassing activity, underlining the high level of engagement. In the state-level election (Study 2), all of the state's roughly 70 constituencies saw canvassing activity. During both campaigns, the party headquarters stayed in touch with local canvassers via email, social media, and telephone. One unit of the party was specifically tasked with training, supporting, and motivating local canvassers.

B.2 Population and Sample

How do the two supporter samples compare to the party's full population of members? In order to maintain anonymity, we cannot provide precise figures. Broadly speaking, however, the two samples do match the party's distribution of members regarding gender and geography. However, the samples were both significantly younger than the average party member. The samples were also disproportionately more engaged, which is not surprising and a feature of the study (focusing on activists). In the general election study, the sample includes 12.9 percent of all party supporters who canvassed for the party during the entire campaign. Furthermore, survey respondents were responsible for 21% of all knocked doors during the campaign (as measured with the smartphone app). The sample can thus best be characterized as 'young and highly motivated supporters.' This group is relevant because it includes individuals for whom the party could have hoped to increase engagement. Given the mild nature of the intervention and the relatively high-effort nature of canvassing, the sample characteristic increases our ability to detect treatment effects. Moreover, the young age in our sample also implies that supporters did not face technological barriers to using the smartphone application with which the party organized its canvassing and which we use to obtain unobtrusive behavioral outcomes.

B.3 Ethical considerations

Field experiments are an excellent method for drawing causal inferences. But they also raise tough ethical questions because researchers intervene in (rather than observe) the real world. In our case, ethical considerations were particularly pressing because our study could have had an impact on the election. We therefore carefully considered the ethical dimension of

our study which we want to discuss before concluding. While we obtained ethical approval, we still want to reflect on two particular ethical issues: potential effects on the election and subjects' non-information about participation in an experiment.

First, implementing the survey meant that we intervened in an electoral campaign. Were we justified in doing so? Importantly, the survey among party supporters would have taken place with or without our presence. The party regularly engages its supporters using emails, surveys and phone calls. We simply advised the party on how to best implement the survey. The ultimate decision to launch the survey, however, was made by party officials. There was also no power differential, which could have led the party to feel obligated to implement the survey. At the time, all authors were graduate students and the party is one of Europe's largest with a highly professional team of campaigners.

Second and related, the expected sample size meant that it was exceedingly unlikely for the study to have any effect on the election. Today, we know that this calculation was correct. We do not observe any treatment effects in either experiment. Even taking the point estimates at face value implies that not a single constituency would have elected a different candidate had the study not taken place. All this is not to say that the survey was without *any* effect. We did, after all, intervene in the real world. But it strikes us that the scientific insights—presented above—were sufficiently high to justify our intervention.

Third, the survey did not deceive subjects. Party supporters were provided with truthful information about the effort of the main competitor. If anything, the study thus provided a public good to party supporters. Study participants—who were contacted online—were also entirely free in their decision to participate in the study. The party did not, however, inform subjects that the data would also be used for scientific purposes. This non-information worked in our favor by preserving the natural field setting "where the environment is one where the subjects [...] do not know that they are in an experiment" (Harrison and List, 2004, p. 1014). That said, we hope that i) by avoiding any harm, ii) by allowing subjects to freely choose to participate, and iii) by maintaining the confidentiality of all subjects including the party and country, we were justified to stomach this non-information (decided upon by the party) in order to explore an important question in political science.

C Additional tables and figures

Table A2: Summary statistics

	Study 1 Effectiveness sample			**Study 2** Career sample		
	Mean	*SD*	*Obs.*	*Mean*	*SD*	*Obs.*
Pre-treatment covariates						
Female	0.23	0.42	1,184	0.22	0.42	1,885
Age	45.19	19.43	1,184	57.42	17.65	1,885
Party member	0.87	0.34	1,184	0.97	0.18	1,885
Years of party membership	8.54	13.73	1,184	22.50	17.33	1,885
Canvassed in prior elections	0.43	0.50	1,184	0.55	0.50	1,885
Participated in campaign workshop	0.32	0.47	1,184	0.16	0.37	1,885
Canvassed before survey in current election	0.25	0.43	1,184	0.01	0.10	1,885
Expected vote margin	15.75	7.64	1,184	–	–	–
Expects more knocked doors for own party	0.84	0.37	1,184	–	–	–
Difference in knocked doors (mio)	0.62	0.76	1,184	–	–	–
Perceived visibility of canvassing	–	–	–	3.04	1.51	1,885
Has career concerns	–	–	–	0.78	0.41	1,885
Manipulation check						
Pre-treat belief about persuasion rate	28.56	21.80	1,184	–	–	–
Post-treat belief about persuasion rate	23.87	18.46	1,164	–	–	–
Post-treat belief about career concerns	–	–	–	4.08	1.39	1,819
Outcomes						
Intended canvassing (any)	0.55	0.50	1,164	0.25	0.43	1,881
Intended canvassing (days)	2.35	3.14	1,164	1.29	3.51	1,881
Actual canvassing (any)	0.15	0.36	1,184	0.02	0.15	1,885
Actual canvassing (days)	0.49	1.52	1,184	0.11	0.94	1,885
Actual canvassing (doors)	32.42	113.28	1,184	1.97	14.90	1,885

Notes: The Table presents the summary statistics of the two samples reporting each variable's mean (*mean*), standard deviation (*SD*) and sample size (N). Details are provided in Section E.

Table A3: Balance across treatment and control group (effectiveness study)

	Treatment	Control	Δ	se(Δ)	p(Δ=0)
Male	0.775	0.762	0.014	(0.025)	0.582
Age	44.560	45.816	-1.256	(1.142)	0.272
Is party member	0.849	0.881	-0.033	(0.020)	0.103
Years of party membership	7.547	9.524	-1.976	(0.796)	0.013
Has experience canvassing	0.438	0.424	0.014	(0.029)	0.634
Participated in door-to-door workshop	0.296	0.348	-0.052	(0.027)	0.054
Downloaded app before effectiveness survey	0.382	0.395	-0.014	(0.028)	0.634
Prior Belief: persuasion rate	28.389	28.729	-0.341	(1.268)	0.788
Has canvassed before survey	0.247	0.247	-0.000	(0.025)	1.000
Days canvassed before survey	1.164	1.125	0.039	(0.202)	0.848
Doors visited before survey	59.188	51.807	7.380	(12.208)	0.546
Absolute value of vote margin	15.975	15.526	0.449	(0.444)	0.312
Expects more knocked doors for own party	0.831	0.848	-0.017	(0.021)	0.429
Difference in knocked doors (mio)	0.614	0.619	-0.005	(0.044)	0.904
Number of observations	592	592			

Notes: The Table presents the mean of the indicated variables for the treatment and control group as well as the corresponding p-values of t-tests in order to showcase balance for the effectiveness sample (Study 1).

Table A4: Heterogeneity by years of membership (effectiveness study)

	Manipulation check	Intentions		App Data			Index
	Belief: persuasion rate	Any	Days	Any	Days	Door	Overall
Panel A: Pooled sample							
Treatment	-7.974***	-0.007	-0.037	0.011	0.032	4.867	0.018
	(1.044)	(0.025)	(0.165)	(0.016)	(0.079)	(5.926)	(0.046)
Treatment × years member	-0.005	-0.002	-0.010	0.000	0.003	0.284	-0.000
	(0.069)	(0.002)	(0.012)	(0.001)	(0.004)	(0.290)	(0.003)
Years member	0.005	0.001	0.006	-0.000	-0.003	-0.183	-0.000
	(0.059)	(0.001)	(0.007)	(0.001)	(0.003)	(0.169)	(0.002)
Control mean	27.848	0.561	2.373	0.150	0.483	30.910	-0.000
Observations	1138	1138	1138	1157	1157	1157	1138
Panel B: Underestimators							
Treatment	2.577***	-0.018	-0.199	0.024	-0.008	0.552	-0.013
	(0.611)	(0.042)	(0.253)	(0.026)	(0.137)	(8.388)	(0.072)
Treatment × years member	0.014	-0.002	0.004	0.000	0.004	0.167	0.001
	(0.040)	(0.003)	(0.018)	(0.001)	(0.006)	(0.372)	(0.004)
Years member	0.001	-0.001	-0.010	0.000	-0.004	0.042	-0.002
	(0.030)	(0.002)	(0.011)	(0.001)	(0.007)	(0.242)	(0.003)
Control mean	9.431	0.518	2.261	0.127	0.445	26.850	-0.064
Observations	432	432	432	436	436	436	432
Panel C: Overestimators							
Treatment	-14.139***	0.002	0.060	0.003	0.057	6.856	0.037
	(1.316)	(0.031)	(0.219)	(0.022)	(0.098)	(8.265)	(0.061)
Treatment × years member	0.102	-0.002	-0.018	0.000	0.003	0.379	-0.000
	(0.087)	(0.003)	(0.015)	(0.001)	(0.006)	(0.448)	(0.004)
Years member	-0.082	0.002	0.014	-0.001	-0.002	-0.274	0.001
	(0.068)	(0.002)	(0.010)	(0.001)	(0.003)	(0.229)	(0.002)
Control mean	38.787	0.586	2.439	0.164	0.505	33.312	0.038
Observations	706	706	706	721	721	721	706

Notes: Table A4 presents OLS regressions of the indicated outcomes on the effectiveness treatment dummy and the interaction with years of party membership (centered to have mean zero). Pre-registered control variables are included. Robust standard errors are given in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A5: Correlation between canvassing intentions and behavior

	App Data	
	(1) Effectiveness study	(2) Career study
Panel A: Any canvassing		
Any canvassing intention	0.069***	0.052***
	(0.022)	(0.014)
Control mean	0.150	0.026
Observations	575	963
Panel B: Canvassing days		
Intended days	0.110***	0.084**
	(0.035)	(0.033)
Control mean	0.483	0.120
Observations	575	963

Notes: Table A5 presents the correlations between canvassing intentions and behavior for both studies. Panel A has a dummy for any observed canvassing as outcome. Panel B has the number of observed canvassing days as outcome. Robust standard errors are given in parentheses. The sample is restricted to the control group. All pre-registered control variables are included in the regressions. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A6: Main results, no control variables (effectiveness study)

	Manipulation check	Intentions		App Data			Index
	Belief: persuasion rate	Any	Days	Any	Days	Door	Overall
Panel A: Pooled sample							
Treatment	-7.996***	-0.015	-0.038	0.007	0.019	3.022	0.007
	(1.055)	(0.029)	(0.184)	(0.021)	(0.088)	(6.586)	(0.059)
Control mean	27.848	0.561	2.373	0.150	0.483	30.910	-0.000
Observations	1164	1164	1164	1184	1184	1184	1164
Panel B: Underestimators							
Treatment	2.695***	-0.050	-0.356	0.007	-0.057	-3.671	-0.069
	(0.599)	(0.048)	(0.285)	(0.032)	(0.145)	(9.074)	(0.090)
Control mean	9.431	0.518	2.261	0.127	0.445	26.850	-0.064
Observations	440	440	440	444	444	444	440
Panel C: Overestimators							
Treatment	-14.132***	0.008	0.164	0.007	0.065	7.166	0.056
	(1.297)	(0.037)	(0.239)	(0.027)	(0.112)	(9.026)	(0.078)
Control mean	38.787	0.586	2.439	0.164	0.505	33.312	0.038
Observations	724	724	724	740	740	740	724

Notes: Table A6 presents OLS regressions of the indicated outcomes on the effectiveness treatment dummy. Robust standard errors are given in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A7: Robustness to controlling for days canvassed prior to experiment (Effectiveness study)

	Intentions		App Data			Index
	Any	Days	Any	Days	Door	Overall
Panel A: Effectiveness experiment: pooled						
Treatment	-0.006	-0.065	0.013	0.025	4.094	0.009
	(0.025)	(0.161)	(0.016)	(0.067)	(4.863)	(0.041)
Control mean	0.561	2.373	0.150	0.483	30.910	-0.000
Observations	1138	1138	1157	1157	1157	1138
Panel B: Effectiveness experiment: underestimators						
Treatment	-0.020	-0.185	0.027	0.025	2.946	0.001
	(0.043)	(0.248)	(0.026)	(0.116)	(6.780)	(0.066)
Control mean	0.518	2.261	0.127	0.445	26.850	-0.064
Observations	432	432	436	436	436	432
Panel C: Effectiveness experiment: overestimators						
Treatment	0.004	0.016	0.003	0.028	3.898	0.011
	(0.031)	(0.214)	(0.022)	(0.078)	(6.594)	(0.052)
Control mean	0.586	2.439	0.164	0.505	33.312	0.038
Observations	706	706	721	721	721	706
Panel D: Career concern experiment						
Treatment	0.001	-0.042	-0.004	-0.022	-0.603	-0.024
	(0.018)	(0.146)	(0.006)	(0.038)	(0.600)	(0.035)
Control mean	0.249	1.308	0.026	0.120	2.299	0.000
Observations	1881	1881	1885	1885	1885	1881

Notes: Table A14 presents OLS regressions of the indicated outcomes treatment dummies. Pre-registered control variables with one exception: The dummy variable indicating any canvassing prior to the experiment is replaced by the number of days canvassed prior to the experiment. Robust standard errors are given in parentheses. All pre-registered control variables are included. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A8: Treatment effects shortly after experiment (effectiveness study)

	One day after treatment			On week after treatment		
	Any	Days	Doors	Any	Days	Doors
Panel A: Effectiveness sample: pooled Treatment						
	0.012	0.014	2.115	0.011	0.011	-1.103
	(0.016)	(0.041)	(3.272)	(0.011)	(0.011)	(2.934)
Control mean						
Observations	0.113	0.235	14.840	0.034	0.034	5.753
Panel B: Effectiveness sample: underestimators	1157	1157	1157	1157	1157	1157
Treatment						
	0.029	0.017	1.256	0.004	0.004	-2.818
	(0.025)	(0.061)	(4.327)	(0.016)	(0.016)	(3.124)
Control mean						
Observations	0.091	0.200	11.768	0.032	0.032	4.591
Panel C: Effectiveness sample: overestimators	436	436	436	436	436	436
Treatment						
	0.003	0.017	2.519	0.013	0.013	-0.440
	(0.021)	(0.056)	(4.569)	(0.014)	(0.014)	(4.493)
Control mean						
Observations	0.126	0.255	16.656	0.035	0.035	6.441
thisstat18	721	721	721	721	721	721

Notes: Table A8 presents OLS regressions of the indicated outcomes on the effectiveness treatment dummy. Columns (1) to (3) display results for canvassing behavior one day after treatment. Columns (4) to (6) display results for canvassing behavior in the first week after treatment. Robust standard errors are given in parentheses. All pre-registered control variables are included. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A9: Balance across treatment and control group (career study)

	Treatment	Control	Δ	se(Δ)	p(Δ=0)
Male	0.767	0.787	-0.019	(0.019)	0.319
Age	56.792	58.022	-1.229	(0.813)	0.131
Is party member	0.970	0.964	0.006	(0.008)	0.480
Years of party membership	21.847	23.124	-1.278	(0.799)	0.110
Has experience canvassing	0.546	0.554	-0.009	(0.023)	0.703
Participated in door-to-door workshop	0.162	0.159	0.003	(0.017)	0.840
Perceived visibility of canvassing	3.054	3.026	0.028	(0.069)	0.682
Has career concerns	0.788	0.772	0.016	(0.019)	0.401
Has canvassed before survey	0.011	0.010	0.001	(0.005)	0.915
Days canvassed before survey	0.012	0.022	-0.010	(0.010)	0.331
Doors visited before survey	0.171	0.492	-0.322	(0.407)	0.430
Number of observations	920	965			

Notes: The Table presents the mean of the indicated variables for the treatment and control group as well as the corresponding p-values of t-tests in order to showcase balance for the career sample (Study 2).

Table A10: Heterogeneity by years of membership (career study)

	Manipulation check	Intentions		App Data			Index
	Belief: role of canvassing	Any	Days	Any	Days	Door	Overall
Treatment	0.119*	-0.000	-0.081	-0.006	-0.042	-0.891	-0.041
	(0.064)	(0.018)	(0.149)	(0.006)	(0.041)	(0.617)	(0.036)
Treatment × years member	0.002	-0.001	-0.009	-0.000	0.000	-0.018	-0.002
	(0.004)	(0.001)	(0.008)	(0.000)	(0.002)	(0.036)	(0.002)
Years member	-0.003	-0.002**	-0.000	0.000	-0.000	0.003	-0.001
	(0.003)	(0.001)	(0.006)	(0.000)	(0.001)	(0.022)	(0.001)
Control mean	4.023	0.249	1.308	0.026	0.120	2.299	0.000
Observations	1819	1881	1881	1885	1885	1885	1881

Notes: Table A10 presents OLS regressions of the indicated outcomes on the career treatment dummy and the interaction with years of party membership (centered to have mean zero). Pre-registered control variables are included. Robust standard errors are given in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A11: Main results, no control variables (career study)

	Manipulation check	Intentions		App Data			Index
	Belief: role of canvassing	Any	Days	Any	Days	Door	Overall
Treatment	0.113*	0.008	-0.036	-0.004	-0.031	-0.670	-0.023
	(0.065)	(0.020)	(0.162)	(0.007)	(0.043)	(0.684)	(0.043)
Control mean	4.023	0.249	1.308	0.026	0.120	2.299	0.000
Observations	1819	1881	1881	1885	1885	1885	1881

Notes: Table A11 presents OLS regressions of the indicated outcomes on the career treatment dummy. Robust standard errors are given in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A12: Treatment effects for males (career study)

	Manipulation check	Intentions		App Data			Index
		Any	Days	Any	Days	Door	Overall
Panel A: males							
Treatment	0.161**	-0.009	-0.102	0.001	-0.050	-0.655	-0.034
	(0.073)	(0.020)	(0.157)	(0.007)	(0.039)	(0.622)	(0.036)
Control mean	4.001	0.239	1.216	0.018	0.100	1.768	-0.039
Observations	1,418	1,463	1,463	1,465	1,465	1,465	1,463
Panel B: females							
Treatment	-0.055	0.033	0.106	-0.034*	0.010	-1.730	-0.054
	(0.135)	(0.041)	(0.350)	(0.018)	(0.108)	(1.816)	(0.103)
Control mean	4.102	0.288	1.649	0.053	0.194	4.257	0.143
Observations	401	418	418	420	420	420	418

Notes: Table A12 presents OLS regressions of the indicated outcomes on the career treatment dummy for the subsample of male respondents. Robust standard errors are given in parentheses. All pre-registered control variables are included. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A13: Treatment effects for respondents with lower career concerns (career study)

	Manipulation check	Intentions		App Data			Index
		Any	Days	Any	Days	Door	Overall
Panel A: no career concerns							
Treatment	0.235*	0.038	0.148	-0.003	-0.005	0.067	0.028
	(0.138)	(0.024)	(0.138)	(0.007)	(0.021)	(0.625)	(0.038)
Treatment × High career concerns	-0.147	-0.051	-0.299	-0.004	-0.048	-1.220	-0.090
	(0.156)	(0.032)	(0.235)	(0.011)	(0.055)	(0.984)	(0.059)
High career concerns	-0.071	0.157***	0.758***	-0.006	-0.010	-0.349	0.131***
	(0.118)	(0.022)	(0.146)	(0.008)	(0.029)	(0.645)	(0.037)
Control mean	4.023	0.249	1.308	0.026	0.120	2.299	0.000
Observations	1819	1881	1881	1885	1885	1885	1881

Notes: Table A13 presents OLS regressions of the indicated outcomes on the career treatment dummy for the subsample of respondents with initially low career concerns and initially high career concerns separately. Initially high career concerns are defined as supporters who either see themselves running for public office (or for a mandate in the party), and those who already have a public office or a mandate within the party. Robust standard errors are given in parentheses. All pre-registered control variables are included. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A14: Robustness to controlling for days canvassed prior to experiment

	Intentions		App Data			Index
	Any	Days	Any	Days	Door	Overall
Panel A: Effectiveness experiment: pooled						
Treatment	-0.006	-0.065	0.013	0.025	4.094	0.009
	(0.025)	(0.161)	(0.016)	(0.067)	(4.863)	(0.041)
Control mean	0.561	2.373	0.150	0.483	30.910	-0.000
Observations	1138	1138	1157	1157	1157	1138
Panel B: Effectiveness experiment: underestimators						
Treatment	-0.020	-0.185	0.027	0.025	2.946	0.001
	(0.043)	(0.248)	(0.026)	(0.116)	(6.780)	(0.066)
Control mean	0.518	2.261	0.127	0.445	26.850	-0.064
Observations	432	432	436	436	436	432
Panel C: Effectiveness experiment: overestimators						
Treatment	0.004	0.016	0.003	0.028	3.898	0.011
	(0.031)	(0.214)	(0.022)	(0.078)	(6.594)	(0.052)
Control mean	0.586	2.439	0.164	0.505	33.312	0.038
Observations	706	706	721	721	721	706
Panel D: Career concern experiment						
Treatment	0.001	-0.042	-0.004	-0.022	-0.603	-0.024
	(0.018)	(0.146)	(0.006)	(0.038)	(0.600)	(0.035)
Control mean	0.249	1.308	0.026	0.120	2.299	0.000
Observations	1881	1881	1885	1885	1885	1881

Notes: Table A14 presents OLS regressions of the indicated outcomes treatment dummies. Pre-registered control variables with one exception: The dummy variable indicating any canvassing prior to the experiment is replaced by the number of days canvassed prior to the experiment. Robust standard errors are given in parentheses. All pre-registered control variables are included. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A15: Treatment effects by election closeness (effectiveness study)

	Manipulation check	Intentions		App Data			Index
	Belief: persuasion rate	Any	Days	Any	Days	Door	Overall
Panel A: Pooled sample							
Treatment	-8.577***	0.034	0.234	0.021	0.007	4.882	0.060
	(1.507)	(0.041)	(0.272)	(0.038)	(0.163)	(11.778)	(0.100)
Treatment × close race	2.464	-0.138*	-0.571	-0.066	-0.013	-6.956	-0.173
	(3.254)	(0.083)	(0.544)	(0.076)	(0.324)	(26.208)	(0.213)
Close race	-2.632	0.009	0.075	0.054	0.106	18.068	0.105
	(2.597)	(0.059)	(0.392)	(0.056)	(0.229)	(18.472)	(0.151)
Control mean	27.848	0.561	2.373	0.150	0.483	30.910	-0.000
Observations	702	702	702	714	714	714	702
Panel B: Underestimators							
Treatment	2.432**	0.003	0.134	0.074	0.012	1.983	0.065
	(0.999)	(0.072)	(0.461)	(0.063)	(0.298)	(18.480)	(0.166)
Treatment × close race	0.630	-0.145	-1.117	-0.196	-0.326	-26.220	-0.410
	(1.779)	(0.143)	(0.857)	(0.126)	(0.560)	(37.469)	(0.341)
Close race	-2.545**	-0.035	0.103	0.129	0.280	22.247	0.176
	(1.116)	(0.102)	(0.662)	(0.094)	(0.436)	(29.597)	(0.259)
Control mean	9.431	0.518	2.261	0.127	0.445	26.850	-0.064
Observations	242	242	242	245	245	245	242
Panel C: Overestimators							
Treatment	-14.405***	0.049	0.285	-0.007	0.002	6.161	0.057
	(1.792)	(0.050)	(0.339)	(0.047)	(0.194)	(15.095)	(0.125)
Treatment × close race	2.596	-0.134	-0.266	0.004	0.164	4.072	-0.041
	(4.021)	(0.102)	(0.698)	(0.095)	(0.397)	(35.018)	(0.271)
Close race	-1.496	0.035	0.058	0.013	0.010	15.901	0.068
	(2.985)	(0.073)	(0.489)	(0.069)	(0.263)	(23.673)	(0.186)
Control mean	38.787	0.586	2.439	0.164	0.505	33.312	0.038
Observations	460	460	460	469	469	469	460

Notes: Table A15 presents OLS regressions of the indicated outcomes on the effectiveness treatment dummy. Close race is defined as having had voteshare margin within 5 percentage points at the last election. Sample restricted to those respondents who i) downloaded the app or ii) who provided their zipcode in the survey. Robust standard errors are given in parentheses. All pre-registered control variables are included. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A16: Treatment effects on social media activity (effectiveness study)

	Social media posting		
	Any	News story	Effort signal
Panel A: Pooled sample			
Treatment	0.009	0.011	-0.006
	(0.012)	(0.012)	(0.011)
Control mean	0.064	0.057	0.046
Observations	1157	1157	1157
Panel B: Underestimators			
Treatment	0.023	0.023	0.005
	(0.018)	(0.018)	(0.016)
Control mean	0.050	0.045	0.032
Observations	436	436	436
Panel C: Overestimators			
Treatment	0.001	0.003	-0.012
	(0.017)	(0.016)	(0.015)
Control mean	0.073	0.065	0.054
Observations	721	721	721

Notes: Table A16 presents OLS regressions of social media campaign activity on the effectiveness treatment dummy. Social media activity is measured as party messages shared through the smartphone application. Robust standard errors are given in parentheses. All pre-registered control variables are included. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A17: Effect sizes and power in survey experiments with political behavior

Outcome	Effect size			Minimum detectable effect size			Source
	absolute	% of control mean	standard deviation	absolute	% of control mean	standard deviation	
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Studies in this paper							
Canvassing index			0.017			0.129	Study one - pooled
Any canvassing	0.011	0.073	0.031	0.045	0.299	0.125	Study one - pooled
Doors canvassed	4.933	0.16	0.046	16.64	0.538	0.154	Study one - pooled
Canvassing index			-0.041			0.101	Study two
Any canvassing	-0.006	-0.231	-0.038	0.017	0.646	0.106	Study two
Doors canvassed	-0.891	-0.029	-0.055	1.728	0.056	0.107	Study two
Panel B: Studies in the same context							
Canvassing index			-0.093			0.132	Hager, Hensel, Hermle and Roth (2023) - underestimators
Any canvassing	-0.013	-0.105	-0.039	0.045	0.364	0.136	Hager, Hensel, Hermle and Roth (2023) - underestimators
Doors canvassed	-14.388	-0.375	-0.088	21.949	0.572	0.135	Hager, Hensel, Hermle and Roth (2023) - underestimators
Any canvassing	0.026	8.667	0.476	0.020	6.533	0.359	Hager, Hensel, Roth and Stegmann (2023)
Doors canvassed	1.207	12.573	0.691	0.986	10.267	0.564	Hager, Hensel, Roth and Stegmann (2023)
Any canvassing	-0.032	0.200	0.087	0.042	0.263	0.115	Hager et al. (2021)
Doors canvassed	0.024	0.001	0.000	19.718	0.577	0.137	Hager et al. (2021)
Panel C: Other survey experiments with behavioral outcomes							
Verified protest attendance	0.026	1.182	0.179	0.0448	2.036	0.309	Hager et al. (2022) - left-wing sample
Verified protest attendance	-0.006	-1.500	-0.081	0.0196	4.900	0.265	Hager et al. (2022) - right-wing sample
Self-reported protest attendance	-0.027	-1.000	-0.167	0.021	0.793	0.132	Cantoni et al. (2019) - Pooled treatment
Refused (political) survey participation	0.230	2.300	0.767	0.191	1.915	0.638	Corstange (2016) - American Embassy treatment
Election turnout	0.014	0.019	0.032	0.081	0.109	0.187	Kreft and Orkin (2020) - information treatment
Election turnout	0.088	0.118	0.202	0.090	0.120	0.206	Kreft and Orkin (2020) - ANC ahead; ANC supporters
Election turnout	-0.001	-0.001	-0.002	0.006	0.007	0.019	Gerber et al. (2020) - 2010 experiment; IV estimation
Election turnout	-0.026	-0.034	-0.061	0.056	0.074	0.131	Gerber, Huber and Fang (2023) - pooled estimate
Panel D: Comparison of average effect sizes and power							
Mean absolute value across panel A		0.099	0.038		0.308	0.12	
Mean absolute value across panels B and C		1.998	0.189		1.916	0.183	

Notes: Table A17 compares estimated effect sizes and minimum detectable effect sizes in this paper to selected studies in the literature. Panel B includes all studies using information delivered through surveys and studying the same type of outcome data as our study (canvassing behavior measured using a smartphone application. Panel C includes experimental studies first made public after 2014 that use interventions administered through surveys to study actual political behavior outside the digital sphere. This implies that we do not include studies using other means of treatment administration (e.g. mail, email, or phone calls). We also do not include studies that study pure online behavior (e.g. signing of petitions or social media posts) as they are arguably easier to shift than 'offline' behavior. We identified studies using a survey of experts in the field of political behavior rather than using literature databases to ensure that we also include recent, and unpublished work. Bolded outcomes in column 1 indicate statistical significance at the 5% level. Columns 2 to 4 indicate different effect sizes. Columns 5 to 7 indicate different minimum detectable effect sizes at 80% power and 5% test size. Panel A displays statistics for studies in this paper. Panel B displays statistics for other studies run in the same context and the same outcome measurement. Panel C displays statistics for studies that study the impact of survey experiments on other studies. Panel D compares average effect sizes and power across studies in this paper and in the literature.

Table A18: Summary statistics for Prolific sample

	Mean	SD	Median	Min.	Max.	Obs.
Predetermined variables						
Female	0.54	0.50	1.00	0.00	1.00	600
Age	28.44	9.01	26.00	18.00	69.00	600
Has no party preference	0.09	0.29	0.00	0.00	1.00	600
Consider career with party	2.61	1.21	2.00	1.00	5.00	600
Consider career as elec. official	2.36	1.16	2.00	1.00	5.00	600
Outcome variables (control)						
Manipulation check: usefullness of canvassing for career	3.64	1.23	4.00	1.00	6.00	301
Intention: any canvassing	0.08	0.28	0.00	0.00	1.00	301
Intention: canvassing days	0.64	2.79	0.00	0.00	28.00	301

Notes: Table A18 presents summary statistics for the Prolific sample used in Table A21. Career questions are measured on a five-point Likert scale with the following options: 1 "No, definitely not" 2 "No, rather not" 3 "I am not sure" 4 "Yes, rather yes" 5 "Yes, definitely". The manipulation check is the answer to the following question: "What do you think? How useful is canvassing for a political career?" Answers are recorded on a six-point Likert scale: 1 "Not at all useful" 2 "Not useful " 3 "Rather not useful" 4 "Rather useful" 5 "Useful" 6 "Very useful".

Table A19: Prior and posterior beliefs

	Experts					
	Prior			*Posterior*		
	ATE	Var	$P(\lvert ATE\rvert \leq 2)$	ATE	Var	$P(\lvert ATE\rvert \leq 2)$
Effectiveness prime on intentions	0.1	37.54	44.20%	-1	5.53	65.67%
Effectiveness prime on behavior	0.8	37.18	46.82%	1	3.58	75.75%
Career prime on intentions	5.2	36.49	35.78%	0.8	3.22	80.48%
Career prime on behavior	3.8	16.51	46.93%	0	1	99.92%

Notes: Table A19 provides summary statistics of experts' prior beliefs on the average treatment effect (ATE) and the implied posterior beliefs. ATE refers to the expected average treatment calculated using the averaged distribution of beliefs. Prior beliefs are updated using a distribution of treatment effects obtained through repeated rerandomization of treatment assignment (10,000 repetitions). Section D describes the Bayesian methodology in detail.

Table A20: Prior beliefs (Prolific Sample)

	Lay people	
	Average belief	Data
	ATE	*ATE*
Career prime on intentions	15.3	-0.1
Career prime on behavior	13.9	-0.6
Effectiveness prime on intentions	-7.1	-0.8
Effectiveness prime on behavior	7.7	1.1

Notes: Table A20 the expected and actual effect sizes for a sample of 600 individuals recruited through Prolific.

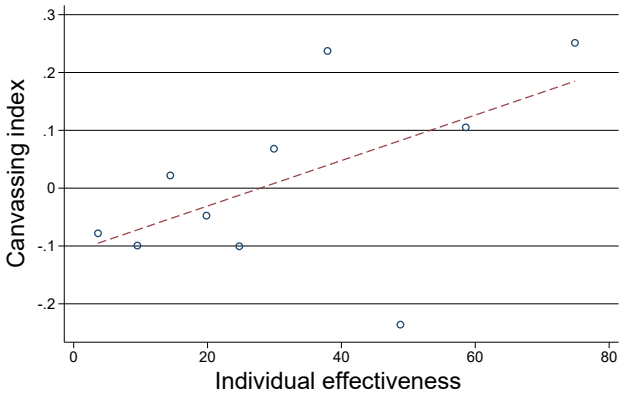
Table A21: Replication of career study

	Manipulation check (z)	Any canvassing	Days canvassing
Treatment effect	0.262***	0.011	0.048
	(0.078)	(0.023)	(0.224)
Control group mean	0.000	0.083	0.641
Observations	600	600	600

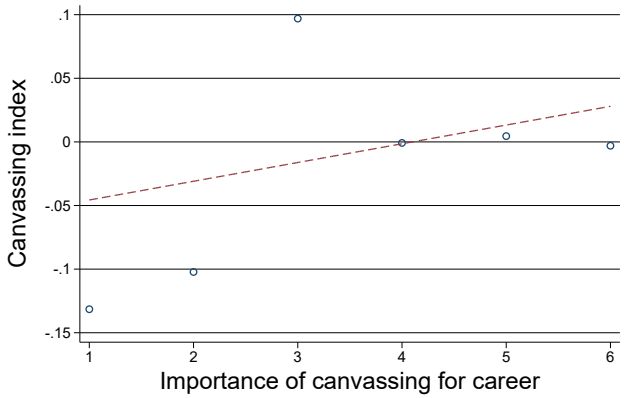
Notes: Table A21 presents the results for replication of the career concern study. The sample was recruited using Prolific and restricted to residents in the same country as the main study. The table shows OLS regressions of the indicated outcomes on the career treatment dummy. The manipulation check are standardizes answers to the following question: "What do you think? How useful is canvassing for a political career?" Answers are recorded on a six point Likert scale: 1 "Not at all useful" 2 "Not useful " 3 "Rather not useful" 4 "Rather useful" 5 "Useful" 6 "Very useful". Robust standard errors are given in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Figure A.1: Instrumental motives and canvassing

(a) Effectiveness (Study 1)

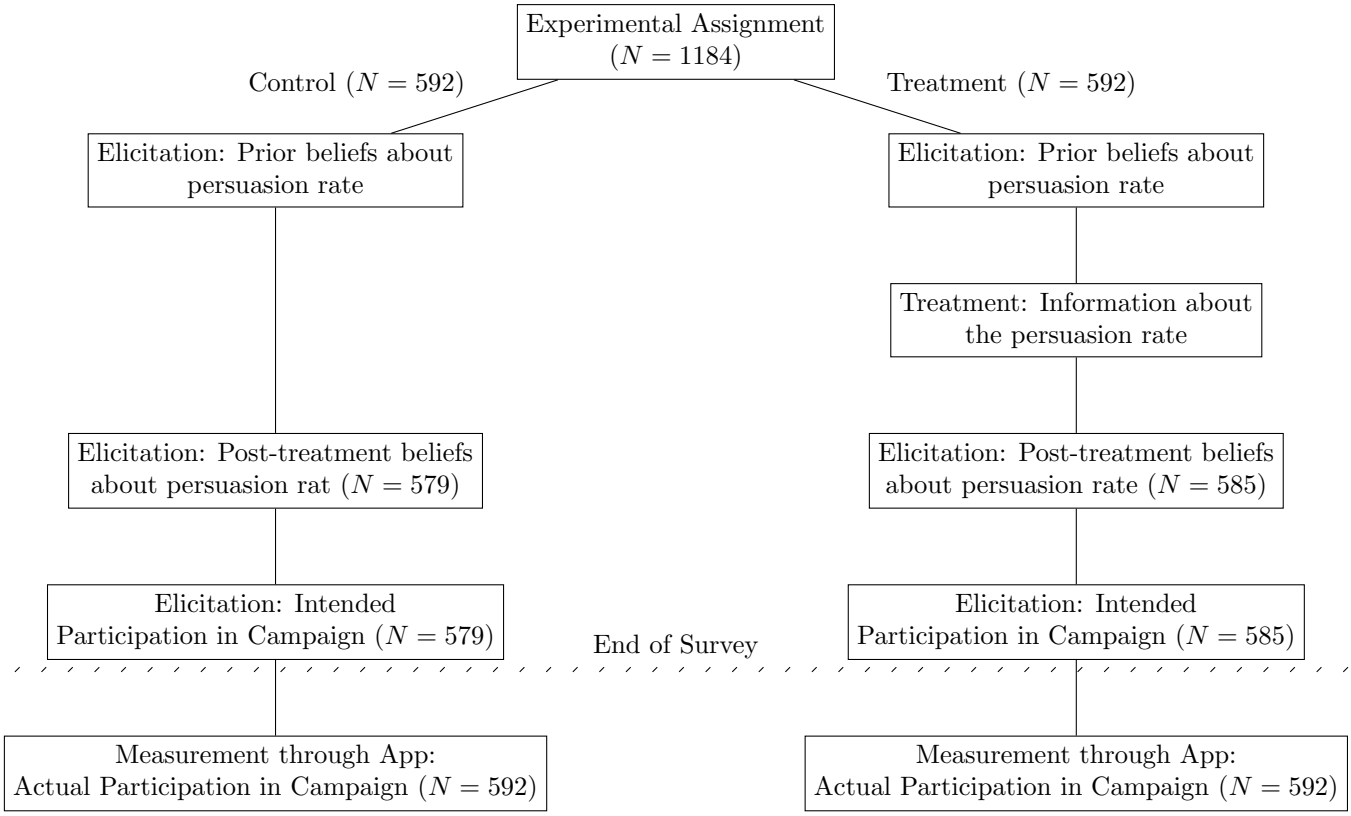


(b) Career ambitions (Study 2)



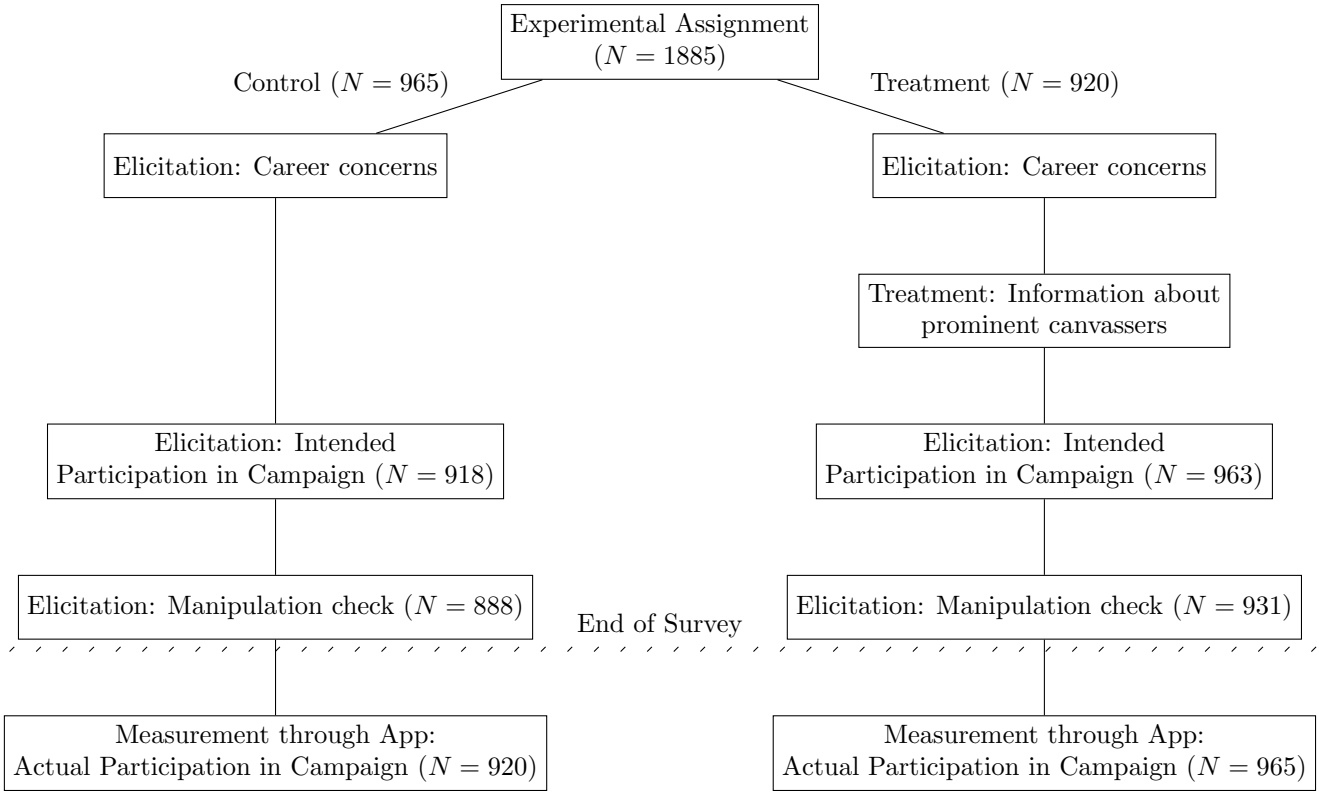
Notes: The Figures display the relationship between respondents' canvassing effort and intentions (a standardized index combining both) and their beliefs about i) individual effectiveness (percentage of persuaded voters; Figure A.1a) and ii) the importance of engagement for one's political career (1-6 scale from "not useful at all" to "very useful"; Figure A.1b). Bins contain deciles for Study 1, and all possible values for Study 2.

Figure A.2: Experimental design (effectiveness study)



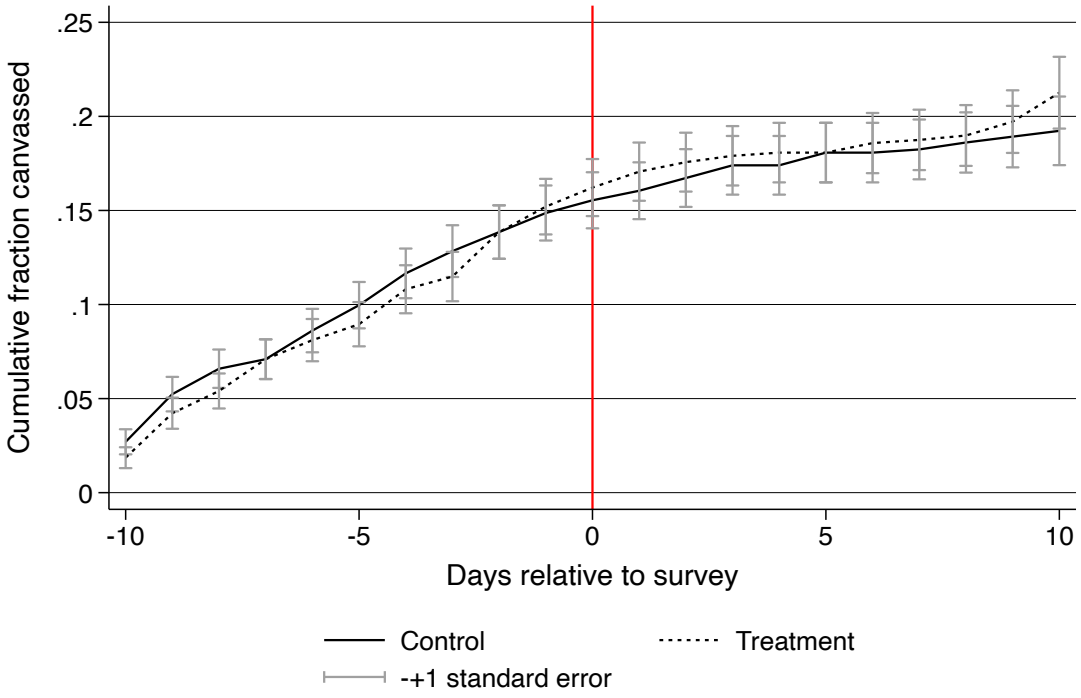
Notes: Figure A.2 illustrates the experimental design for the effectiveness study. The experiment took place during a national election campaign in a large western European country.

Figure A.3: Experimental design (career study)



Notes: Figure A.3 illustrates the experimental design for the effectiveness study. The experiment took place during a regional election campaign in a large western European country.

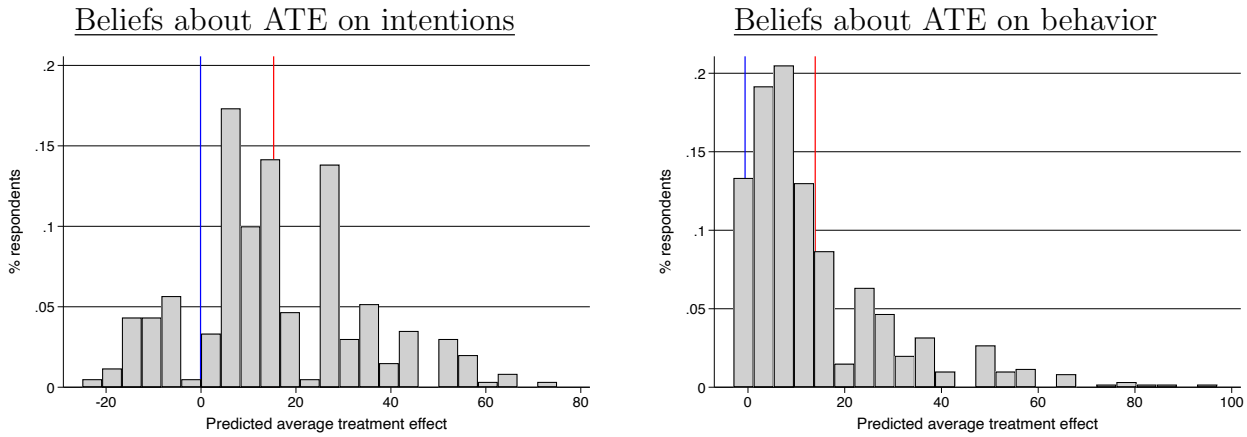
Figure A.4: Fraction canvassed over time (effectiveness study)



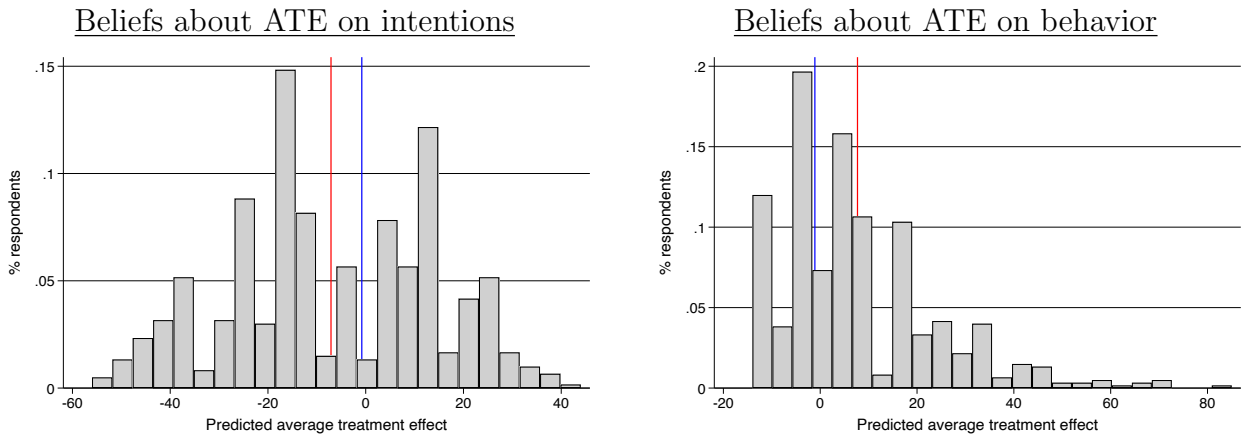
Notes: Figure A.4 displays the cumulative fraction of individuals who canvassed in a 10 day window around the treatment administration.

Figure A.5: Distribution of laypeople beliefs about average treatment effects (ATEs)

Panel A: Belief updating in the career experiment



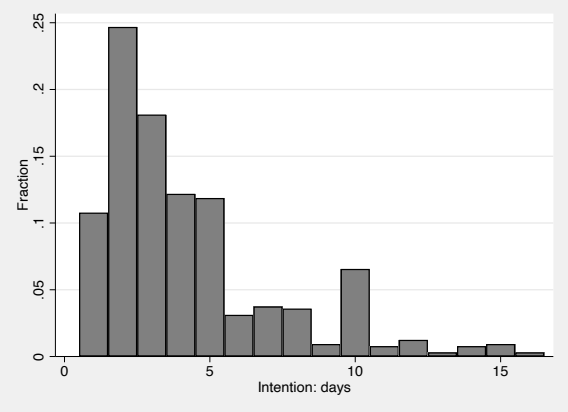
Panel B: Belief updating in the effectiveness experiment



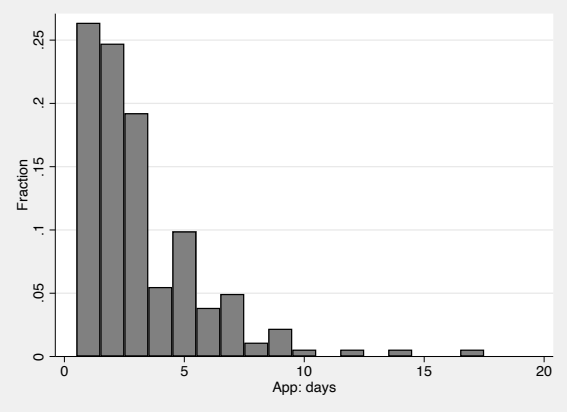
Notes: Figures A.5 display averaged laypeople beliefs about treatment effects. Grey bars indicate the averaged prior beliefs calculated by averaging the probability mass experts put on each interval. Dashed lines indicate the averaged posterior beliefs obtained by updating averaged prior beliefs using Bayes' rule with a distribution of treatment effects obtained through bootstrapping (10,000 repetitions). Panel A displays beliefs about the effect of the career treatment. Panel B displays beliefs about the effect of the effectiveness treatment. Both panels show effects on canvassing intentions and behavior. The sample consists of 600 individuals recruited using Prolific.

Figure A.6: Instrumental motives and canvassing

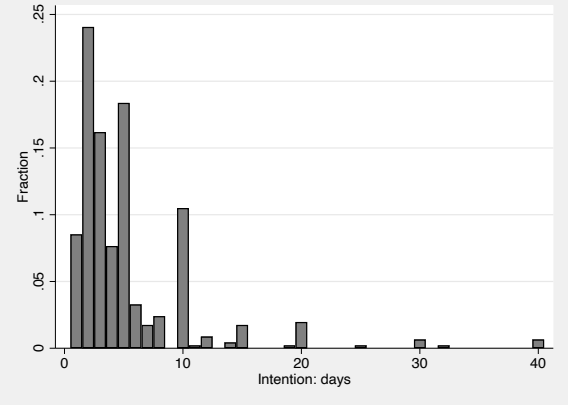
(a) Intended days (effectiveness study)



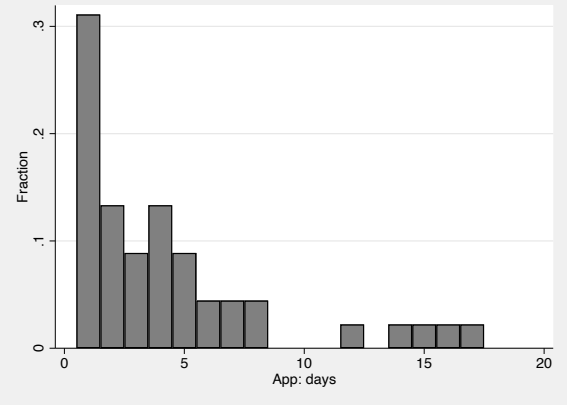
(b) Recorded days (effectiveness study)



(c) Intended days (career study)



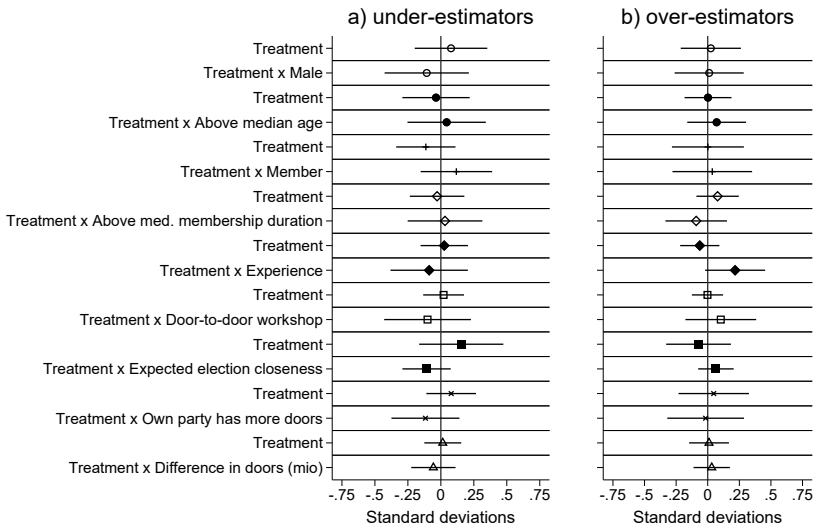
(d) Recorded days (career study)



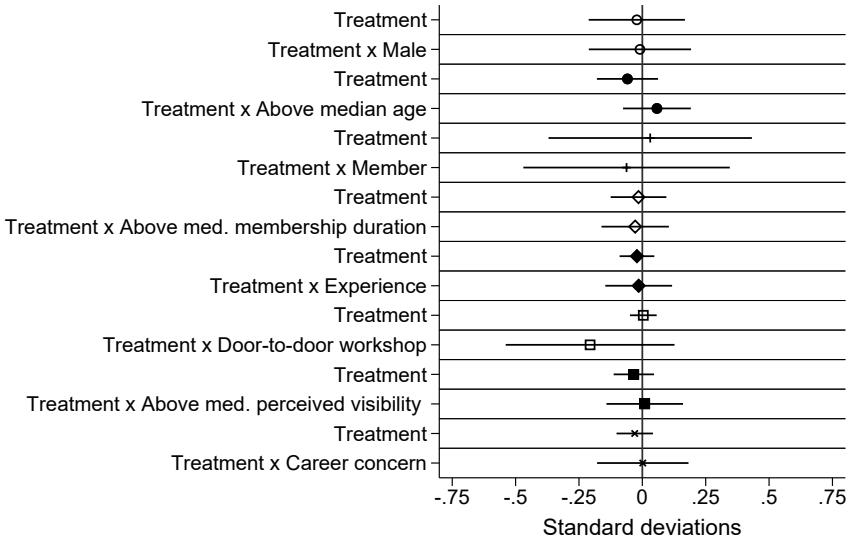
Notes: Figure A.6 displays the distribution of intended and actual canvassing days across both studies. Figures A.6a and A.6b show distributions for the effectiveness study. Figures A.6c and A.6d show distributions for the career study. All distributions are restricted to non-zero values.

Figure A.7: Treatment effect heterogeneity

(a) Effectiveness (Study 1)

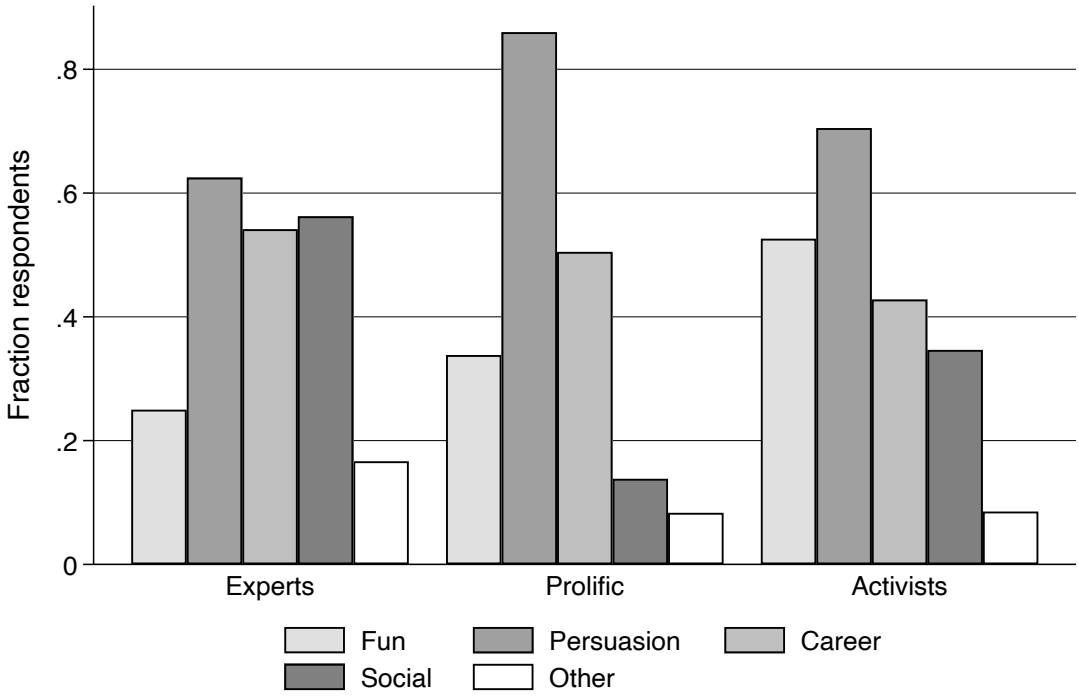


(b) Career ambitions (Study 2)



Notes: The Figures display the estimation results of heterogeneous treatment effects on the pre-specified index of canvassing activity for the two experiments. Specifically, we run the main equation including the heterogeneity variable and interact it with the treatment indicator. The index is defined as the standardized sum of the five standardized measures of canvassing intentions and actual canvassing behavior. All estimates are obtained conditional on the pre-specified control variables. All pre-specified heterogeneity dimensions are shown. "Expected election closeness" is measured as the absolute value of the expected difference in vote share between the supporters' party and the main competing party (in 10% units). "Own party has more doors" is a dummy for whether a respondent expects supporters of her own party to knock on more doors. "Difference in doors" is the difference in the expected number of knocked doors by the respondents' own party and the main competing party (in million doors). The vertical lines indicate 95% confidence intervals.

Figure A.8: Perceived motives of canvassers



Notes: Figure A.8 displays the perceived motives of canvassers elicited from three different samples. In particular, we recruited 600 lay people using Prolific, 1,007 political activists from the same party as well as 54 political scientists working in the activism space.

D Bayesian analysis

To assess to what degree the presented experimental estimates changed our beliefs, we elicited prior beliefs about the treatment's likely effect from lay people as well as experts.

The average distribution of prior beliefs about the two treatments' effects (i.e., the effectiveness prime as well as career prime) on the two outcomes (i.e., canvassing intentions and actual behavior) are presented in Figure 1 for experts. In Figure A.5 we show the distribution of expected treatment effects for lay people.

The Figures yield three findings. First, experts expected average treatment effects in the career experiment of 5.2 percentage points on intentions, and 3.8 percentage points on behavior, on average. Second, experts expected only small average treatment effects in the effectiveness experiment of 0.8 percentage points on intentions, and 0 percentage points on behavior, on average (this may not be surprising, given that we informed the experts about the manipulation check, which showed that for some respondents beliefs were corrected upwards and for some downwards). Third, Table A20 shows that lay people, broadly speaking, expected rather large average treatment effects in both experiments. In the career experiment, they expect 15.3 percentage points on intentions and 13.9 percentage points on behavior. They expect a negative treatment effect on intentions (-7.1 percentage points) and a positive effect on behavior (7.7 percentage points). Taken together, the evidence thus showcases i) that lay people and experts expect at least some treatment effects. Both points underscore the usefulness of the experiment in tightening and changing expert and lay people's beliefs.

To estimate the impact of the information provided by our experiments on experts' uncertainty, we estimate a Bayesian posterior for the experimental sample for both experiments as well as both outcomes. For this purpose, we assume a uniform distribution of beliefs within each elicited treatment effect bin we used in the expert survey. We also generate a distribution of treatment effects using the bootstrap with 10,000 repetitions and collapse the distribution in the same bins. We then use this data to calculate posterior belief distributions using Bayes theorem.[18] Figure 1 display the results of this exercise. We clearly see a compression of belief distributions which implies a reduction in uncertainty about the actual treatment effect. This shows that even for the effectiveness experiment—where the experiments did not move priors by much (at least among experts, the case is different among lay people)—there was a large decrease in uncertainty about the range of possible treatment effects.

[18] Results are highly similar when using STAN.

E Variable description

Outcome variables

In line with our pre-analysis plan, we created an unweighted index based on the five main outcome variables, which we z-score using the mean and the standard deviation of the control group. The five variables are the following:

- Our main outcome of interest is canvassing effort exerted between the completion of the survey and the election. We make use of three variables based on the number of doors people knock on:
  - 1) Whether people knock on any door as registered through the app.
  - 2) The number of doors people knock on as registered through the app.[19]
  - 3) The number of days on which people knocked doors as registered through the app.
- In addition, we use two self-reported canvassing measures, which we collected as part of the survey:
  - 4) A binary variable capturing whether a respondent plans to engage in canvassing during the election campaign.
  - 5) The number of days on which respondents plan to go canvassing. Individuals who do not plan to canvass are coded as zero.

[19]Per our pre-analysis plan, this variable is winsorized at the 99 percentile to deal with outliers.

Control Variables

We estimate all regression models with a list of controls, $\mathbf{X_i}$, which might predict canvassing. Variables indicated with $*$ are only measured in the effectiveness experiment (Study 1). The control variables are the following:

- Party membership (taking the value 1 for members)
- Number of years of party membership (taking the value 0 for non-members)
- Age
- Sex (taking the value 1 for men)
- Whether a participant has participated in a canvassing training workshop
- Whether a participant had already downloaded the online application before the survey
- Whether a participant had participated in canvassing before the current election
- Whether a participant had already canvassed during the current election

* The difference in respondents' beliefs regarding the election result of their own party and the main competing party
* Respondents' beliefs about whether members of their own party or members of the main competing party will knock more doors
* The difference in respondents' beliefs regarding the number of doors members of their own party and members of the main competing party will knock during the election campaign