Online Appendices Not for Publication

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

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 A	41:	Share	of	Tweets	by	own	party	V
					•/		1 V	

Statistic	Ν	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

	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 (anv)	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

Table A2: Summary statistics

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.

Treatment	Control	Δ	$\operatorname{se}(\Delta)$	$p(\Delta{=}0)$
0.775	0.762	0.014	(0.025)	0.582
44.560	45.816	-1.256	(1.142)	0.272
0.849	0.881	-0.033	(0.020)	0.103
7.547	9.524	-1.976	(0.796)	0.013
0.438	0.424	0.014	(0.029)	0.634
0.296	0.348	-0.052	(0.027)	0.054
0.382	0.395	-0.014	(0.028)	0.634
28.389	28.729	-0.341	(1.268)	0.788
0.247	0.247	0.000	(0.025)	1.000
1.164	1.125	0.039	(0.202)	0.848
59.188	51.807	7.380	(12.208)	0.546
15.975	15.526	0.449	(0.444)	0.312
0.831	0.848	-0.017	(0.021)	0.429
0.614	0.619	-0.005	(0.044)	0.904
592	592			
	Treatment 0.775 44.560 0.849 7.547 0.438 0.296 0.382 28.389 0.247 1.164 59.188 15.975 0.831 0.614 592	Treatment Control 0.775 0.762 44.560 45.816 0.849 0.881 7.547 9.524 0.438 0.424 0.296 0.348 0.382 0.395 28.389 28.729 0.247 0.247 1.164 1.125 59.188 51.807 15.975 15.526 0.831 0.848 0.614 0.619 592 592	Treatment Control Δ 0.775 0.762 0.014 44.560 45.816 -1.256 0.849 0.881 -0.033 7.547 9.524 -1.976 0.438 0.424 0.014 0.296 0.348 -0.052 0.382 0.395 -0.014 28.389 28.729 -0.341 0.247 0.247 0.000 1.164 1.125 0.039 59.188 51.807 7.380 15.975 15.526 0.449 0.831 0.848 -0.017 0.614 0.619 -0.005 592 592 592	TreatmentControl Δ se(Δ)0.7750.7620.014(0.025)44.56045.816-1.256(1.142)0.8490.881-0.033(0.020)7.5479.524-1.976(0.796)0.4380.4240.014(0.029)0.2960.348-0.052(0.027)0.3820.395-0.014(0.028)28.38928.729-0.341(1.268)0.2470.000(0.025)1.1641.1250.039(0.202)59.18851.8077.380(12.208)15.97515.5260.449(0.044)0.8310.848-0.017(0.021)0.6140.619-0.005(0.044)

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

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).

	Manipulation check	Inter	tions		App Data	ı	Index
	Belief: persuasion rate	Any	Days	Any	Days	Doors	Overall
Panel A: Pooled sample							
Treatment	-8.040***	-0.008	-0.047	0.011	0.032	4.923	0.017
	(1.044)	(0.025)	(0.165)	(0.016)	(0.079)	(5.948)	(0.046)
Treatment \times years member	-0.012	-0.002	-0.011	0.000	0.003	0.291	-0.000
	(0.068)	(0.002)	(0.012)	(0.001)	(0.004)	(0.290)	(0.003)
Years member	0.012	0.001	0.007	-0.000	-0.003	-0.191	0.000
	(0.058)	(0.001)	(0.007)	(0.001)	(0.003)	(0.168)	(0.002)
0 + 1	07 0 40	0 501	0.070	0.150	0.400	00.010	0.000
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	2.535^{***}	-0.019	-0.205	0.024	-0.004	0.986	-0.012
	(0.610)	(0.042)	(0.254)	(0.026)	(0.138)	(8.458)	(0.073)
Treatment \times years member	0.006	-0.002	0.003	0.000	0.005	0.256	0.001
	(0.039)	(0.003)	(0.018)	(0.001)	(0.006)	(0.379)	(0.004)
Years member	0.012	-0.001	-0.008	0.000	-0.005	-0.078	-0.002
	(0.029)	(0.002)	(0.011)	(0.001)	(0.006)	(0.259)	(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	-14.207***	0.001	0.042	0.002	0.054	6.636	0.033
	(1.317)	(0.031)	(0.220)	(0.022)	(0.098)	(8.292)	(0.061)
Treatment \times years member	0.101	-0.002	-0.018	0.000	0.003	0.373	-0.001
	(0.087)	(0.003)	(0.015)	(0.001)	(0.006)	(0.450)	(0.004)
Years member	-0.083	0.002	0.014	-0.001	-0.002	-0.275	0.001
	(0.068)	(0.002)	(0.010)	(0.001)	(0.003)	(0.229)	(0.002)
0 + 1	20 707	0 500	0.490	0.104	0 505	99.910	0.020
Control mean	38.787	0.586	2.439	0.164	0.505	33.312	0.038
Observations	706	700	700	(21	(21	(21	706

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

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

	App Da	ta
	(1)	(2)
	Effectiveness study	Career study
Panel A: Any canvassing		
Any canvassing intention	0.069^{***} (0.022)	0.052^{***} (0.014)
Control mean Observations	$0.150 \\ 575$	$0.026 \\ 963$
Panel B: Canvassing days		
Intended days	$\begin{array}{c} 0.110^{***} \\ (0.035) \end{array}$	0.084^{**} (0.033)
Control mean Observations	$0.483 \\ 575$	$\begin{array}{c} 0.120\\ 963 \end{array}$

Table A5: Correlation between canvassing intentions and behavior

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

	Manipulation check	Inter	ntions		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

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

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

	Inter	tions		App Data	ı	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

Table A7: Robustness to controlling for days canvassed prior to experiment

Notes: Table A7 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. Panel A to C show effects in the effectivenes experiment. Panel D shows effects in the career experiment. Robust standard errors are given in parentheses. All pre-registered control variables are included. * p < 0.10, ** p < 0.05, *** p < 0.01

	One day	y after tre	eatment	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

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

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

	Treatment	Control	Δ	$\operatorname{se}(\Delta)$	$p(\Delta=0)$
male_m	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			

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

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).

	Manipulation check	Inter	tions	App Data			Index
	Belief: role of canvassing	Any	Days	Any	Days	Door	Overall
Treatment	$\begin{array}{c} 0.117^{*} \\ (0.065) \end{array}$	$0.003 \\ (0.020)$	-0.058 (0.161)	-0.005 (0.007)	-0.035 (0.043)	-0.767 (0.681)	-0.032 (0.042)
Treatment \times years member	$0.002 \\ (0.004)$	-0.000 (0.001)	-0.001 (0.008)	-0.000 (0.000)	$\begin{array}{c} 0.001 \\ (0.002) \end{array}$	$0.005 \\ (0.037)$	$\begin{array}{c} 0.000\\ (0.002) \end{array}$
Years member	0.003 (0.003)	-0.004*** (0.001)	-0.016^{***} (0.005)	-0.001*** (0.000)	-0.004^{**} (0.002)	-0.078^{***} (0.029)	-0.007^{***} (0.002)
Control mean Observations	4.023 1819	0.249 1881	$1.308 \\ 1881$	$0.026 \\ 1885$	$0.120 \\ 1885$	$2.299 \\ 1885$	$0.000 \\ 1881$

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

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

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

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

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

	Manipulation check	Inten	tions		App Data		Index
		Any	Days	Any	Days	Door	Overall
Panel A: Men							
Treatment	0.163^{**}	-0.008	-0.095	0.002	-0.047	-0.562	-0.029
	(0.073)	(0.020)	(0.157)	(0.007)	(0.038)	(0.604)	(0.036)
Control mean	4.001	0.239	1.216	0.018	0.100	1.768	-0.039
Observations	1418	1463	1463	1465	1465	1465	1463
Panel B: Women							
Treatment	-0.051	0.032	0.177	-0.033*	0.055	-1.432	-0.030
	(0.135)	(0.041)	(0.338)	(0.017)	(0.100)	(1.750)	(0.098)
Control mean	4.102	0.288	1.649	0.053	0.194	4.257	0.143
Observations	401	418	418	420	420	420	418

Table A12: Treatm	ent effects for	males (career	study)
-------------------	-----------------	---------------	--------

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 stu	dy	r)
---	----	----

	Manipulation check	Inter	Intentions		App Data		
		Any	Days	Any	Days	Door	Overall
Panel A: no career concerns							
Treatment	0.234^{*}	0.036	0.138	-0.005	-0.010	-0.098	0.020
	(0.138)	(0.024)	(0.136)	(0.008)	(0.023)	(0.702)	(0.040)
Treatment \times High career concerns	-0.141	-0.047	-0.236	0.001	-0.016	-0.638	-0.057
	(0.156)	(0.033)	(0.228)	(0.011)	(0.050)	(1.016)	(0.059)
High career concerns	-0.074	0.154^{***}	0.723***	-0.010	-0.029	-0.727	0.110***
-	(0.118)	(0.022)	(0.142)	(0.009)	(0.029)	(0.776)	(0.040)
	1.000	0.040	1.000	0.000	0.100	0.000	0.000
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

	Manipulation check	Inten	tions		Index		
	Belief: persuasion rate	Any	Days	Any	Days	Door	Overall
Panel A: Pooled sample							
Treatment	-8.663***	0.039	0.241	0.030	0.021	7.375	0.081
	(1.476)	(0.036)	(0.245)	(0.032)	(0.156)	(10.948)	(0.083)
Treatment \times close race	2.252	-0.096	-0.437	-0.045	0.123	1.164	-0.087
	(3.273)	(0.067)	(0.484)	(0.060)	(0.290)	(23.793)	(0.165)
Close race	-1.670	-0.017	0.023	0.052	0.082	16.040	0.083
	(2.575)	(0.046)	(0.332)	(0.044)	(0.205)	(16.581)	(0.114)
Control mean	27 848	0.561	2.373	0.150	0.483	30.910	-0.000
Observations	689	689	689	701	701	701	689
Panel B: Underestimators							
Treatment	2.124^{**}	0.041	0.372	0.068	-0.071	-0.663	0.078
	(1.053)	(0.068)	(0.452)	(0.056)	(0.323)	(18.400)	(0.152)
Treatment \times close race	0.763	-0.093	-1.117	-0.107	0.205	7.611	-0.155
	(1.869)	(0.128)	(0.817)	(0.106)	(0.557)	(35.688)	(0.287)
Close race	-2.496**	-0.047	0.178	0.122^{*}	0.213	17.740	0.148
	(1.123)	(0.080)	(0.521)	(0.073)	(0.408)	(27.073)	(0.195)
Control mean	9.431	0.518	2.261	0.127	0.445	26.850	-0.064
Observations	239	239	239	242	242	242	239
Panel C: Overestimators							
Treatment	-14.400***	0.036	0.209	0.011	0.063	9.039	0.079
	(1.801)	(0.043)	(0.310)	(0.040)	(0.178)	(14.181)	(0.103)
Treatment \times close race	2.138	-0.081	-0.157	-0.021	0.110	3.201	-0.041
	(4.009)	(0.082)	(0.633)	(0.076)	(0.353)	(32.078)	(0.212)
Close race	-0.750	0.004	0.014	0.020	0.015	13.588	0.060
	(2.970)	(0.058)	(0.430)	(0.056)	(0.236)	(21.175)	(0.144)
Control moon	29 797	0.586	9.420	0.164	0.505	22 210	0.028
Observations	450	450	⊿.459 450	459	459	459	450
0.0001 (0010110	-100	100	100	100	100	-100	-100

Table A14: Treatment effects by election closeness (effectiveness study)

Notes: Table A14 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

	Social media posting						
	Any	News story	Effort signal				
Panel A: Pooled sample							
Treatment	0.017	0.018	-0.001				
	(0.014)	(0.014)	(0.012)				
Control mean	0.064	0.057	0.046				
Observations	1157	1157	1157				
Panel B: Underestimators							
Treatment	0.035^{*}	0.034^{*}	0.011				
	(0.021)	(0.021)	(0.017)				
Control mean	0.050	0.045	0.032				
Observations	436	436	436				
Panel C: Overestimators							
Treatment	0.006	0.008	-0.008				
	(0.019)	(0.018)	(0.016)				
Control mean	0.073	0.065	0.054				
Observations	721	721	721				

Table A15: Treatment effects on social media activity (effectiveness study)

Notes: Table A15 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

Outcome		Effect size			Minimum detectable	e 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 beha	avioral outc	omes					
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 a	nd 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	

Table A16: Effect sizes and power in survey experiments with political behavior

Notes: Table A16 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 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.

	Mean	SD	Median	Min.	Max.	Obs.
Predetermined variables						
Prove la	0 54	0 50	1.00	0.00	1.00	CO1
Female	0.54	0.50	1.00	0.00	1.00	001
Age	28.45	9.00	26.00	18.00	69.00	601
Has no party preference	0.09	0.29	0.00	0.00	1.00	601
Consider career with party	2.62	1.21	2.00	1.00	5.00	601
Consider career as elec. official	2.36	1.16	2.00	1.00	5.00	601
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

Table A17: Summary statistics for Prolific sample

Notes: Table A17 presents summary statistics for the Prolific sample used in Table A20. 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".

	$\mathbf{Experts}$						
	Prior			Posterior			
	ATE	Var	$P(ATE \le 2) \mid$	ATE	Var	$P(ATE \le 2)$	
Effectiveness prime on intentions	0.1	37.54	44.20%	-1	5.44	66.54%	
Effectiveness prime on behavior	0.8	37.18	46.82%	1	3.68	75.10%	
Career prime on intentions	5.2	36.49	35.78%	0.9	3.25	80.05%	
Career prime on behavior	3.8	16.51	46.93%	0	1	99.88%	

Table A18: Prior and posterior beliefs

Notes: Table A18 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.

	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		

Table A19: Prior beliefs (Prolific Sample)

Notes: Table A19 the expected and actual effect sizes for a sample of 600 individuals recruited through Prolific.

	Manipulation check (z)	Any canvassing	Days canvassing
Treatment effect	$\begin{array}{c} 0.268^{***} \\ (0.078) \end{array}$	0.010 (0.023)	0.045 (0.224)
Control group mean Observations	$\begin{array}{c} 0.000\\ 601 \end{array}$	0.083 601	0.641 601

Table A20: Replication of career study

Notes: Table A20 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

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.



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

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.





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 A19 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 (-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 experts 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.

¹⁸Results 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 can vassing. Individuals who do not plan to canvass are coded as zero.

¹⁹Per 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, 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