

THE ELEMENTS OF POLITICAL PERSUASION: CONTENT, CHARISMA AND CUE*

Torun Dewan, Macartan Humphreys and Daniel Rubenson

Political campaigns employ complex strategies to persuade voters to support them. We analyse the contributions of elements of these strategies using data from a field experiment that randomly assigned canvassers to districts, as well as messaging and endorsement conditions. We find evidence for a strong overall campaign effect and show effects for both message-based and endorsement-based campaigns. However, we find little evidence that canvassers varied according to their persuasive ability or that endorser identity matters. Overall the results suggest a surprisingly muted role for idiosyncratic features of prospective persuaders.

1. The Elements of Political Persuasion

Persuasion has long been studied by economists and political scientists who seek to understand how different strategies in advertising or political campaigning lead agents to update their beliefs and ultimately change their behaviours.¹ Political campaigns aim to persuade voters of the merits of a particular alternative, be it a party, a candidate or a policy position. They do so using multiple channels: specific messages to enhance the profile of a candidate or position; individual canvassers to get the message across in a compelling way; and endorsements by public figures to enhance appeal. Do such campaigns matter and which, if any, of their core elements are effective at persuasion?

Finding answers to these questions, although central to understanding these political phenomena, is difficult. First, it is hard to separate the impact of a campaign from other factors that influence opinion. Second, campaigns typically target specific groups of voters, who may or may not be more likely to support the position even in the absence of a campaign. Third, what works with one group might not work with another, so campaigns may tailor their appeal to suit distinct voters. And there is a final complication: campaigns require media – whether distributed by individual canvassers, or print or other channels – and so it is often hard to know whether it is the message or the medium which is persuasive. In sum, real-world aspects of political campaigns make

* Corresponding author: Torun Dewan, Department of Government, London School of Economics. Email: t.dewan@lse.ac.uk.

We thank members of the BC Citizens' Assembly and Fair Voting BC, especially Arjun Singh, Wendy Bergerud and Maxwell Anderson. We are also very grateful to members of the BC-STV campaign. The campaign manager, Susan Anderson-Behn, as well as Maggie Gilbert and David Gagnon provided invaluable input and cooperation. Thanks to the canvassers and enumerators involved in carrying out the project. We thank our research project managers, Stewart Prest and Pierce O'Reilly, for handling the coordination of enumerators. Don Green and David Epstein gave thoughtful comments on the project design and John Bullock and Jon Eguia provided helpful comments on our analysis. We thank seminar participants at the 2009 APSA meetings and at Oxford and NYU. Thanks to Kjell Rubenson, Antje Ellermann and Alan Jacobs for help during the field work and to Ross Jenkins, Tina Narang, Ali Cirone and Sarah Khan for terrific research assistance. We thank Columbia University, the London School of Economics and Ryerson University for financial support.

¹ See Della Vigna and Gentzkow (2010) for an overview.

it difficult to evaluate the impact of different elements that might contribute to overall persuasiveness.

For these reasons, regression analysis of observational data from a political campaign – one that used different messages, campaigners and endorsements – would be unlikely to provide unbiased estimates that identify the key parameters of interest. Rather, the prospect of obtaining useful estimates of different elements of persuasion may hinge upon being able to directly manipulate them, to analyse the relevant counterfactual conditions.

An empirical design that can, in principle, deal with these issues involves two steps: the first randomises contact with a campaign; and the second randomises which elements of a campaign are employed. The former separates the overall campaign effect from contextual factors. The latter separates the impact of the message or endorsement from audience characteristics. This structure, if implemented correctly, would yield the required data and allow for evaluation of the impact on political outcomes of different elements of a campaign, conditional on locality and the intrinsic attributes of particular canvassers. In this study we describe the implementation of such an experiment in the field and evaluate its outcomes. We believe it to be the first field experiment that looks at political attitudes and that accounts for different persuasive elements of a campaign.

To implement this strategy we worked directly with British Columbians for Single Transferable Vote (BC-STV) who campaigned to change the electoral system in the province of British Columbia (BC), Canada, from first past the post to single transferable vote (STV). A referendum on the reform would take place on 12 May 2009. BC-STV planned a door-to-door campaign using a variety of messages that they hoped would convince contacted households to support the proposed change. They also wished to advertise the fact that the proposal was supported by some high-profile public figures. In our collaboration we randomly assigned canvassers to different localities, as well as messages and information about endorsements within those localities. This allowed us to evaluate which messages and endorsements were effective; and to observe differences in the persuasiveness of individual canvassers, when fixing message and endorser conditions.

The first of our treatments consisted of a set of messages. The campaign devised two political messages that they believed would enhance the prospects of a yes vote. The first argued that a change in the electoral system would increase fairness of outcomes, and the second that a switch to STV would increase voter choice and accountability. A second treatment consisted of a set of endorsements: the names and pictures of four leading public figures advocating support for the proposed reform.

Canvassers were randomly assigned to different local voting areas. Within each area a random subset of households received a ‘placebo’ message in which voters were told that a referendum would take place on the prescribed date. A random subset of households received, in addition, a treatment in the form of an official communication from the campaign. The message treatments involved one of two messages (either ‘fairness’ or ‘accountability’) communicated orally by the canvasser and backed up by a campaign flyer outlining its key points. The endorser treatment consisted of a campaign flyer that contained a picture with the name of the endorser. Some treatments involved both message and endorser: the respondent received a campaign

flyer with a picture of the endorser and the key points of the selected message that was also communicated orally.

The control group consisted of households that received only information about the proposed referendum with no official communication from the campaign (the placebo condition). We explicitly did not introduce a condition in which a campaigner advocates for a yes vote without providing any reasoning in the form of a message or an endorsement. We excluded this 'null campaign' to ensure that our results were not driven by the imposition of an unnatural condition. Such concerns for ecological validity aside, we were also concerned by a possible negative impact: allowing canvassers to lobby for a position, while prohibiting them from making any case for it, might lead voters to infer that the campaign had no arguments to make.

The outcome of interest is the intention of a yes vote elicited from a survey of households visited by the campaign. A comparison with the households under the placebo condition allows for an assessment of different types of campaign: a campaign that uses messages, one that uses individual endorsements and one that uses both. Our core results are that both message-based campaigns and endorser-based ones proved effective in increasing support for the proposed reform. Indeed, both had a similar effect: a six percentage point increase in the intention to vote yes.

We can also assess the marginal effect of message content given endorsement (and *vice versa*), as well as the effects of different messages and different endorsers. To make inferences about these marginal effects, and in the absence of a null campaign condition, we require an assumption that a null campaign has null effects. In making this assumption we find a marginal effect of messages of around 6 percentage points, with that of the accountability message around 9 percentage points. Analysing interaction effects suggests that message-based campaigns and endorser-based ones act more as substitutes than complements, although statistically this result is not strong.

These results are conditional on locality and the canvasser delivering the message. A related question is whether – when fixing the message and/or endorsement – it matters which canvasser made contact with the voter. We test whether there were significant differences observed in outcomes across households visited by different canvassers. Somewhat surprisingly, and even though the set of canvassers deployed to persuade voters was diverse in age and background, it made little difference with whom voters made contact. Specifically, and in contrast to existing studies that we discuss below, when controlling for the content of the message delivered we were unable to reject the hypothesis that individuals deployed to persuade voters did not vary in their ability to do so.

Our study distinguishes between different elements of persuasion in campaigns. One is the content of campaign messages. Another is the contact made by individuals associated with the campaign who vary in latent traits such as their charisma, or their communicative ability as in Dewan and Myatt (2007). A final element is due to prominent people who, although associated with the campaign, make no direct contact with voters. Political scientists refer to the effect of 'cues' whereby voters take a stance on a position, independent of its merits, because of its association with a particular individual or party. Our findings suggest that whereas content and cues are important and, indeed, can be seen as substitutes for each other, latent traits such as

charisma are perhaps less important than previously thought. In the next Section, we relate our advances and findings to previous studies in this area.

2. Content, Charisma and Cue

Our analysis relates to a growing field of experimental literature on voter mobilisation that suggests that differences in the content of messages – or arguments – can have effects on behaviour (Arceneaux and Nickerson, 2005, 2010; McNulty, 2005; Gerber *et al.*, 2008; Panagopoulos, 2009). A seminal contribution is by Gerber and Green (2000) who evaluate the impact of information and message content in the context of a get-out-the-vote campaign in New Haven, Connecticut. A core finding is that, relative to a control group that was not contacted, intention to vote increased by around 9 percentage points in a treatment group that received information about the election. A similar effect was recorded for a group that received information and a message emphasising the importance of voting due to ‘civic duty’ or ‘neighbourhood solidarity’. The focus on turnout makes that study different from ours. The designs are also different in that their control group, unlike ours, receives no information about the election. Moreover, our design builds on that earlier study by also assessing the effects of randomly matching canvassers to households.

Closely related to our study is a recent article by Kendall *et al.* (2013). They look at how voters respond to information, randomly distributed by mail and cold calling by phone, about candidates’ valence and ideological disposition in the context of an Italian mayoral election. Their analysis of precinct-level official vote shares shows that voters reacted differently to different types of message and to the medium: voters responded positively (around 4 percentage points) to campaign messages about valence when contacted by mail and phone; with smaller effects recorded when voters received a message with ideological content, or when contact was by mail only. Our study looks at vote intention in a door-to-door campaign where fixed attributes of the person making contact can be important. In randomly matching canvassers to households, we have the advantage of being able to control for such attributes when showing the impact of message-based campaigns.

Other research also explores whether some messages can be more usefully deployed than others. Wantchekon (2003) shows the differential effects of arguments in a 2001 presidential campaign in Benin where candidates used different types of messages in different communities: in some the message emphasised the local benefits of policies; in others the message emphasised national benefits. In a follow-up study, Wantchekon (2009) explored the effect of the ‘national benefits’ message when devised by a group of policy ‘experts’. This treatment group was compared with a control that received a standard party message. Although this work presents persuasive evidence that message content matters, it is unable to assess the independent effects of candidate attributes.²

² The persuasiveness of arguments could of course also depend on whether there are competing arguments and which of these they come up against during the course of a political campaign (Loewen and Rubenson, 2011; Loewen *et al.*, 2012).

Going beyond the content of campaign messages, recent studies have looked at the impact of individuals associated with a political campaign. In a field experiment, Humphreys *et al.* (2006) randomly assigned discussion leaders to groups during a national-level public deliberation process on the island state of São Tomé and Príncipe. The aim was to estimate the extent to which they influenced discussion outcomes. In their study, as in ours, the individuals were not well known in the community and so any impact was likely due to latent traits such as charisma. The estimated effects were pronounced: leader characteristics often predicted up to 50% of variation in outcomes. As the authors acknowledge, however, these effects could be due to different preferences among the set of leaders and hence to different arguments made during deliberations. An advantage of our study is that we control the message a canvasser delivers. We thereby disassociate the intrinsic appeal of the leader from the effect of the argument she makes. In doing so, and in contrast to those earlier findings, we cannot reject the null that the identity of the person delivering the message does not matter to the outcome.

We also allow for the impact of individuals who did not have any contact with voters. Their effect could not be due to their ability to present a case compellingly, but may instead reflect that voters rely on 'information shortcuts' or 'heuristics' when deciding which position to adopt on an issue (Kahneman *et al.*, 1982; Lupia, 1994). An individual may be persuaded of the desirability of an outcome by virtue of the characteristics of the person endorsing it. A large body of work looks at the relationship between election outcomes and 'cues' taken by voters (Kinder *et al.*, 1980; Conover and Feldman, 1989). Recent contributions in the experimental literature include Druckman *et al.* (2010), who distinguish the effect of cues (from other framing effects) in a controlled environment where subjects were exposed to campaign material and footage from debates in the Republican primary for the open seat in the 5th congressional district in Massachusetts. Malhotra and Kuo (2007) manipulate partisan cues in a survey experiment where respondents were asked to apportion blame to public officials for property damage and loss of life in New Orleans after Hurricane Katrina. Brader and Tucker (2010) develop and implement survey experiments in Britain, Hungary and Poland and show that subjects who identify themselves as partisans express policy preferences in line with those of their party leaders when prompted to do so.

A key advantage of our design is that we can analyse cues alongside other elements of a campaign. Recent studies have also sought to distinguish between elements of political persuasion by manipulating party cues and policy-relevant information in controlled settings: Arceneaux (2008) looks at policy and party positions on abortion; Berinsky (2009) studies attitudes on military intervention by manipulating information on casualty rates and partisan cues; Rahn (1993) studied student evaluations of debates between candidates for election to the Minnesota legislature and manipulated information on the candidates' positions and party affiliation and Riggle *et al.* (1992) exposed subjects to information about a candidate's voting record and manipulated both party affiliation and the policy-relevant information. In an important contribution, Bullock (2011) surveys this and other earlier work and argues that this literature offers no clear conclusions over the relative magnitude of the effects. He claims that the standard view, expressed perhaps most strongly by Cohen (1990), that partisan cues have a dominant effect on attitudes cannot be supported by

the combined evidence. In his work, in which partisans received detailed newspaper articles about health care in Wisconsin, Bullock, finds that ‘party cues are influential, but partisans in these experiments are generally affected at least as much – and sometimes much more – by exposure to substantial amounts of policy information’ (Bullock, 2011, p. 512). Likewise, we find that a cue-based campaign has similar effects to a content-based one.

Although most of this work has been based on survey experiments, Arceneaux and Kolodny (2009, p. 763) worked directly with a liberal activist group in the 2006 Pennsylvania statehouse elections and attempted to assess the effects of source cues. The group endorsed candidates from the Democratic Party and were involved in canvassing core supporters as well as voters supporting the Republican Party but whom the group deemed to be persuadable. Individuals were randomly assigned to be either contacted by the activist group or not. The authors’ findings suggest that cues matter in that Republicans who were contacted by the group were less likely to support the endorsed candidates. Although the motivation for this study is closely related to ours, ours differs in having a primary focus on distinguishing between persuasive dimensions of campaigns.³

2.1. *The Setting: British Columbians for STV*

To provide answers to our questions we worked directly with the BC-STV in their build up to the referendum on electoral reform of 12 May 2009. We initially made contact with former members of the British Columbia Citizens’ Assembly on Electoral Reform, which was created by the Government of British Columbia in 2003 as an independent, non-partisan assembly of citizens to examine the province’s electoral system and to make recommendations for reform. The 160 members of the assembly conducted an 11-month period of consultations, deliberations and public hearings before advocating the STV in their final report to the people of British Columbia in December 2004 (British Columbia, 2004). Members campaigned across the province for the proposal, with over 800 meetings being led by assembly campaigners. Despite winning a majority in 77 of the 79 districts, the reform narrowly failed to meet the 60% province-wide threshold in the referendum of 17 May 2005. A second referendum on the same proposal was held in May 2009 and our work involved analysing the effects of the campaign in the period leading up to that date.

In correspondence and face-to-face meetings with the campaign, we designed a strategy that would allow us to answer our core questions and provide the campaign with information about the effects of different elements of their strategy. As part of the implementation of this design, they identified a set of canvassers dedicated to persuading voters of the merits of the campaign, a set of messages that they believed would support their position on STV and a set of sites where campaigning would take place. A further critical aspect of the campaign that we were able to exploit was that it

³ In their account: ‘the design does not allow us to tease out the independent effects of the message and the source cue; it only allows us to test whether politically unaware voters use the combination of the source cue and the contextual information contained in the group’s message to behave as if they were politically aware voters’ (Arceneaux and Kolodny, 2009, p. 759).

was endorsed by leading public figures. We asked the campaign to select from among their endorsers a diverse set they believed would have the biggest positive effect on campaign outcomes.

As we make clear below, our design thus allowed us to separate different channels through which leadership could affect the campaign. On the one hand, as we analysed the door-to-door campaign we were able to assess whether the messages the campaign deployed and/or the messengers the campaign deployed were effective. As these individuals were not publicly known figures, their effect is less likely to be due to cue taking and more due to their persuasive ability. By contrast, the campaign endorsers were well known public figures, and, as they had no direct contact with participants, their effect can only have been due to more indirect forms of influence such as cue taking. It was important that these different elements – the messages, messengers and endorsers – were employed by the campaign with our involvement focused primarily on guidance in the randomisation protocols we describe below.

2.2. *The Treatments*

To identify campaign effects with confidence, we used randomised assignment of conditions to ensure that observed differences can be attributed to treatment effects and not due to selection or other confounding effects. In a random subset of households, the campaign delivered a ‘placebo’ message in which voters were simply told that a referendum will take place on 12 May 2009.

The placebo was designed to ensure a level playing field between those who were directly contacted by the campaign and those who were not. Two rationales motivate the use of a placebo. First, in the absence of the placebo, it is possible that one effect of being visited by the campaign is a raised awareness about the referendum in treated houses over and above that which existed in non-treated households. It would be hard for us to identify our quantities of interest, were this the case. Second, in the absence of a placebo we risk biases arising from separate recruitment strategies for subjects in treatment and control. Our study population is defined as the population that activists can make contact with; but if the control population is recruited by enumerators, there is no guarantee that the control subjects belong to the same population. For this reason all households received an informative placebo that contained no information that could be construed as favourable to the yes side. In this way, we can be reasonably certain that any differences in voting intention arose only from differences in the treatments received.

A random subset of households – in addition to receiving the placebo – received one from a set of treatments. These treatments came in the form of a campaign document and a presentation by a campaigner who canvassed these households. The communication combined different elements that the campaign believed might be effective in securing a yes vote from the recipient of the treatment. In particular, canvassers were dispatched to give one of two arguments in favour of STV – an argument that makes the case that STV leads to fairer outcomes and an argument that STV allows for greater choice – and a flyer summarising the main points of the argument. Some of these flyers were also endorsed by leading public STV supporters. Some treatments involved a simple endorsement with no supporting argument.

2.2.1. *The messages*

Two messages were constructed by the campaign to make the case that STV is the right electoral system for British Columbia.⁴ One set of arguments focused on the general claim made in favour of proportional systems that such systems are fairer. This message emphasised that STV translates votes into seat shares more proportionally and reduces the phenomenon of wasted votes. It also explained that plurality rule elections often lead to governments being elected on less than 50% of the vote, and that in the past BC has had governments that have received a lower province-wide vote share than an opposition party.

A second message focused on issues of choice and accountability. Here, the argument was made that parties would be forced to nominate more than one candidate, inducing a choice among different party representatives and a comparison of performance between local districts. The campaign argued that STV would lead to better local representation, for with more local Members of the Legislative Assembly (MLAs), citizens would be able to choose among a larger number of representatives to contact with their concerns (or even choose to visit all local MLAs).

2.2.2. *The canvassers*

A group of activists was trained by the campaign to engage in door-to-door canvassing. This was a central element of the campaign strategy, as it is in many election campaigns. Canvassers shared a commitment to the campaign's goals but varied in several dimensions – including socio-demographic characteristics and political orientation. Just over two thirds of the canvassers who participated in the study were women. Canvassers ranged in age from 19 to 63 years. Roughly 55% had completed a university degree and 40% were full-time students, whereas the rest were employed. The group was ethnically diverse as well as being diverse in terms of political orientation. Half of the canvassers intended to vote for the Green Party in the provincial election, 30% for the left-wing New Democratic Party and the remainder were evenly split between the Liberal Party and the Conservative Party. In addition to this variation in background characteristics and political orientation, canvassers may also have varied on many unobservable dimensions that could affect their persuasiveness. Our design allowed us to identify whether such observed or unobserved differences between canvassers could explain a part of the campaign effect.

2.2.3. *The endorsers*

Some of the messages were also endorsed by leading public figures who supported the campaign.⁵ BC-STV settled on four main endorsers for the purpose of the study: David Suzuki, a well known media figure and environmentalist; Preston Manning, a former politician, founder and leader of the Reform Party; Andrew Coyne, the editor of *Maclean's Magazine* and a well known journalist and Lorne Nystrom, former New Democratic Party Member of Parliament. The list gives some balance on politician and non-politician as well as on ideology, with Manning and Coyne generally considered to

⁴ Images of these messages including the full text can be seen in Figure A1 in Appendix A.

⁵ A screenshot of one of the endorser flyers can be seen in Figure A2 in Appendix A.

be on the right while Suzuki and Nystrom are further on the left of the political spectrum in Canada. Ultimately these were the endorsers the campaign believed would have the biggest impact.

Ex ante, there are a number of reasons to believe that the campaign's endorsers would vary in their effectiveness with regard to persuading individuals to vote for the STV proposal. First, the four endorsers, while all publicly known figures, vary in how well they are known to the general public. It is safe to say that Preston Manning and David Suzuki are known by most Canadians. Manning has been a central figure in Canadian politics for two decades, is arguably responsible for redefining party politics and is the face of western Canadian populism. Suzuki is one of the best known environmentalists and television personalities in the country. In contrast, Nystrom and Coyne are less well known. Second, although all the endorsers supported electoral reform and held the view that STV was preferable to the status quo, they varied in how vocal and public these views were. Third, it is likely that the public's perception of the motivations for the endorsers' support of electoral reform varied. In particular, Manning and Nystrom may have been perceived to have adopted their positions for reasons of political strategy. We might also think that signalling support for electoral reform could be more costly for some endorsers than others. Coyne and Manning, for example, are generally considered to be on the right of the political spectrum. Electoral reform, at least in Canada, is mostly perceived as an issue of the left. Thus, perhaps one would think that Coyne and Manning could pay a price for their public support. It is of course difficult to know to what extent, if at all, this is true, especially given Manning's retirement from party politics and Coyne's general independent minded reputation. Each of these factors suggests that the cue taken by voters could vary according to which (if any) endorsement they received and our design allowed us to isolate this channel of persuasion.

2.3. *Subject Population*

To analyse our quantities of interest we distinguish the general effects of messages and messengers from those specific to particular neighbourhoods. British Columbia has 85 single member electoral districts – 'ridings' – that return representatives to the Provincial Legislature. These districts are broken down further into 'voting areas', analogous to US polling precincts. Voting areas assign voters to a single voting place and ballot box and their population is capped, by legislation, at 400 voters to ensure that a voting officer and voting clerk can administer the votes in one day.

There are 4,799 voting areas in the Greater Vancouver Regional District. In drawing our sample of 216 voting areas for our experiment, we used data from the 2006 Canadian Census to stratify by wealth, education, ethnic demography and geographic size. Wealth and education were divided into two categories, whereas ethnic demography (measured using the ethno-linguistic fractionalisation index) and geographic size were divided into three categories. Sampling was proportionate to voting area population size. The design involved the 216 voting areas in the Greater Vancouver area that are shown shaded in Figure 1.

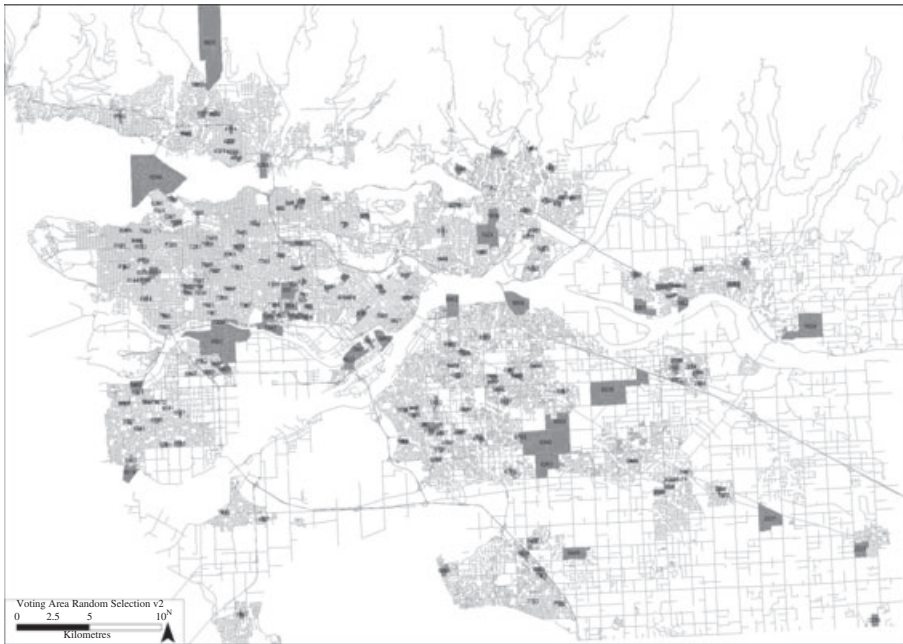


Fig. 1. *Location of Sampled Canvassing Sites*

Note. Areas shaded dark grey are those which were sampled.

Within these areas our aim was not to reach a population representative of all British Columbia voters, nor even those in the Greater Vancouver area, but to construct a sample of households that were likely to be reached by the campaign during door-to-door canvassing. Thus, the aim of the study was to discern campaign effects and to separate different channels of persuasion among a representative sample of those contacted by the campaign.

For this reason the canvassers were encouraged to use their normal canvassing techniques with the exception that there was strict determination of which areas to visit and which messages to deliver to each household. In each voting area, canvassers for the campaign were provided with a map of the area on which a randomly generated starting point and direction were indicated, along with a randomly generated route that they would follow. In a deviation from our original design, canvassers worked in pairs in each area which has the advantage of increasing the number of location–canvasser units. They were told to start on one side of the indicated street and to visit the fifth residential structure to deliver the first message. If they made contact they would then move to the tenth residential structure; if unsuccessful, they would go to the next residential structure and so on. In addition, canvassers were given instructions as to which message/endorser combination to deliver in each household. A typical daily sheet providing the treatment combinations to be delivered is shown in Figure B1, in Appendix B.

Although designs that work directly with political campaigns come with many advantages, one challenge with designs such as ours is to make sure that canvassers do not deviate from the design when it comes to delivering treatments. If canvassers went

off message, delivered the wrong message or in other ways drifted from the agreed upon protocol, we would have difficulty interpreting our results. We included a number of features to guard against these problems.

First, all canvassers attended a comprehensive training session led by the campaign manager, communications director and the study authors. Canvassers were made aware of how deviating from the design would threaten the experiment – and thus also threaten the campaign's ability to learn about which strategies worked best. During training, canvassers were instructed on the scripts to use for each message, for the endorser condition (and endorser-only condition) as well as the placebo condition. These scripts were developed in collaboration with the campaign and the canvassers, increasing both internal validity and canvasser buy-in. As mentioned above, canvassers were of course encouraged to use their regular campaigning style. Canvassers were also provided with a lengthy manual containing detailed instructions, scripts and protocols, as well as information on what can go wrong, what might lead to false conclusions and what to do in case of mistakes. We provide the full canvasser manual in the supplementary materials.

Second, each day of the study, canvassers were provided with precisely the number of flyers they were to deliver that day (and flyers were sorted according to the randomised order). Therefore, we have good reason to believe that we minimised the risk of canvassers delivering the wrong message and there should be no risk of canvassers simply delivering their preferred message to all households.

Last, canvassers were instructed to record certain data after each house visit (see Appendix B). In particular, they made a note of the start time and length of the visit, the address, the gender and age of the subject, whether they stayed on message (see below) and a general impressionistic rating of the subject's knowledge of STV and the referendum. Each evening these data were entered on a website by canvassers. The data had two purposes. First, they serve as a check on whether treatments were implemented according to our design. Second, the information was necessary in order for our enumerators to carry out survey data collection, as we needed the address of those canvassed as well as a way to identify subjects.⁶ In only less than 5% of cases is there any reporting of misadministration of messages.

Our final sample is composed of 48% women and 52% men. The median age is 48 years and about 48% of the sample had completed a university degree with only 8% of the sample having less than a high school education. The voting profile of the respondent pool largely matches the overall results from the previous referendum on electoral reform held in 2005. Of those reporting, 65% report not having voted in the last referendum and, of those who voted, 58.6% reported having voted yes and 41.4% reported having voted no.

3. Research Strategy

We analyse both the effectiveness of a door-to-door campaign and assess which elements contribute to its success. To analyse the effect of a campaign we need to

⁶ An additional feature of our design was the fact that we randomly assigned enumerators to canvassers to avoid any systematic relationship between canvasser style and enumerator style.

separate its impact from local conditions that may influence the outcome variable of interest. Separating different elements of a door-to-door campaign from each other is difficult. As a message requires a messenger and, *vice versa*, it is difficult to distinguish between their effects. Precisely, the same message may be effective when delivered by one individual but not by another.⁷ Our interest then is in estimating average effects of a messaging strategy, recognising that the medium may alter the interpretation of messages received.

Formally, let the ‘potential outcome’ for individual i in location j be written as:

$$y_i(m, c, e, X_j),$$

where $m \in M = \{\emptyset\} \cup \{1, 2\}$ is the message assigned, $c \in C = \{1, 2, \dots, n_c\}$ is a canvasser dispatched to deliver a message and $e \in E = \{\emptyset\} \cup \{1, 2, \dots, n_e\}$ is an endorser. The set of treatments is given by the product of M , C and E . In addition, y_i can depend on a set of location-specific features, X_j . With some abuse of notation we say $i \in m$ if i receives message condition m (similarly, we write $i \in c$ or $i \in e$) and we let $\pi(m)$, $\pi(e)$ denote the share of subjects in a given condition.

The effect of endorser A for individual i (given message m and canvasser c) is then $y_i(m, c, A, X_j) - y_i(m, c, \emptyset, X_j)$, the effect of endorser A relative to B is $y_i(m, c, A, X_j) - y_i(m, c, B, X_j)$, the effect of message 1 is $y_i(1, c, e, X_j) - y_i(\emptyset, c, e, X_j)$ and so on.

Understood in this way, average message and canvasser effects have a simple interpretation that takes account of the implicit dependency of message effects on characteristics of canvassers and *vice versa*. The average effect of a canvasser or message is the average of these individual effects given some distribution of treatments. Of course, average effects of any treatment may be different for different combinations of treatments and for different localities, j .

The key practical difficulty is that strategic assignment of messaging campaign tactics to prospective voters makes disentangling effects problematic. As we discuss above, there are three related practical problems. First, certain political agents may be prone to making certain types of arguments. This correlation hinders any empirical assessment of the distinct causal effect of canvassers (and their characteristics) and messages on political outcomes of interest. Second, as a canvasser aims to convince voters of a particular viewpoint, and uses her powers of persuasion to that effect, she may also adapt her message to suit her audience. This strategic component of political communication introduces a problem of endogeneity that confounds analysis of canvasser effects. Third, both arguments and canvassers may be selected because they are deemed likely to be effective in a given setting. Thus, a selection bias is introduced because this strategic aspect of campaigns can confound attempts to isolate the causal effect of political agents and the arguments they make. Each of these problems is likely to arise as part of the natural process of political campaigning. Thus, observational data may not allow us to draw an empirical

⁷ Note that in our interpretation of the interaction between a messaging strategy and the medium used, a given message can have heterogeneous effects that depend on the medium; this can be contrasted with an interpretation in which the medium alters the message itself: an overture issued by Nixon to China is a substantially different message than a formally identical one issued by a Democrat.

distinction between the effects of agents and the messages they deploy and across the different channels of persuasion.

Instead, we require independent variation in

- (i) the messages delivered ('content'),
- (ii) the individuals who deliver those messages ('charisma') and
- (iii) the individuals who 'endorse' them ('cue').

The distinction between (ii) and (iii) arises regularly in political campaigns: an individual canvasser may attempt to persuade a voter of the merits of a particular position that is associated with one or more well known public figures, even though the canvasser herself will likely not be known. In principle, this feature of a campaign allows us to distinguish between our dimensions of interest as, in this setting, it

- (i) captures content effects,
- (ii) captures 'charisma' and
- (iii) captures 'cue' effects.

We achieve this variation experimentally by varying M and E at the individual level, within localities and the elements of C at the location level. Thus, we examine individual-level randomisation for M and E and cluster randomisation for C . The exposure is undertaken in 'field conditions'. We turn now to a discussion of the implementation of the research design outlined above, using data from a unique field experiment during this campaign.

3.1. Treatment Assignment

Our design called for 24 canvassers, each visiting a different voting area for each of nine days and visiting 12 households in each. Canvassers were randomly assigned to voting areas and the message and endorser treatments were randomly administered to households within voting areas using a blocked design. This blocking of message and endorsement treatments can be seen in the first subrow in each row of Table 1. Partial blocking at the location level ensured that there were at least five distinct treatment combinations assigned in each location and in most cases there were at least eight. In over 90% of locations, no treatment combination (other than placebo) was assigned more than twice.⁸

In practice, as we show in each of the second and third subrows of Table 1, data available for analysis falls short of this target. The data shortfall arose due to a combination of factors. First, although 2,592 households were targeted for canvassing, engaging sufficient numbers of canvassers for the duration required proved very difficult for the campaign. In part, this was due to the fact that the randomisation strategy required greater time spent travelling than normally needed for canvassing. The effect was that 1,044 rather than 2,592 houses were canvassed according to the campaign, of which 941 locations could be relocated and

⁸ The code for the blocked assignment is available in our replication material.

Table 1
Distribution of Treatments

	Endorser					Total
	None	Coyne	Manning	Nystrom	Suzuki	
<i>Message</i>						
Accountability	216	126	126	126	126	720
	81	47	43	43	49	263
	50	13	14	23	21	121
Fairness	216	126	126	126	126	720
	78	45	46	46	45	260
	42	28	28	22	22	142
None	648	126	126	126	126	1,152
	231	45	42	45	55	418
	122	23	19	23	28	215
Total	1,080	378	378	378	378	2,592
	390	137	131	134	149	941
	214	64	61	68	71	478

Notes. In each subrow the first number gives the target number assigned to treatment under the randomisation scheme. The second number provides the number of reported canvassed areas in identified locations. The third provides the number for which survey data were gathered with responses on the outcome variable of interest.

confirmed by our enumeration teams. The distribution of these 941 households across canvasser and endorser treatment conditions is given in the second subrow of the Table.

There was further data loss at the enumeration stage. A team of enumerators followed up the canvassers with a lag of between two and four days. These enumerators made contact at the door and administered a short survey measuring opinion on various aspects of the referendum issue, vote intention, knowledge about the endorsers, general political knowledge and recall of campaign contact, in addition to basic socio-demographic information.⁹ Critically, enumerators were blind to treatment conditions.

Enumeration was successful – in that the enumerators were able to complete surveys, including responses to the question on voting intention – in 478 cases. Of the remainder, in 157 cases enumerators were unable to find anybody at home after multiple visits (at least 3 per house, often 6); in 66 cases there was no common language; in 150 cases those answering the door refused to respond to the survey in full and in another 92 cases the respondents refused to respond to the key question on voting intentions.

In Appendix F we provide balance tests demonstrating that the randomisation procedure achieved good balance on covariates even in the presence of incomplete message delivery and non-responses. There are, however, risks of biases arising from data loss; we discuss these in Section 7.

⁹ See Appendix C for the complete survey instrument.

3.2. *Estimation*

The experimental procedure allows for relatively simple estimation of treatment effects and allows us in most cases to use methods of design-based inference rather than having to rely on regression specifications.

To examine ‘content’ we compare the mean outcomes among groups assigned specific messages while accounting for location and the assignment of canvassers – elements which are orthogonal to treatment but formed part of our randomisation procedure and are likely correlated with the outcome variables. To do so we generate an independent variable which ‘purges’ the outcome variable of average effects associated with location and the canvassers assigned to those locations. This purged variable is simply the residual from a regression of the outcome on location–canvasser fixed effects. The distribution of both the purged and unpurged variables is given in Table 2.

3.2.1. *Content*

To estimate the effect of message m' compared with m'' we calculated a weighted average of differences in means, using the endorser treatments as blocks and relative block size as weights:

Table 2
Estimated Propensity to Vote Yes Given Message and Endorser Conditions

	Endorser					Total
	None	Coyne	Manning	Nystrom	Suzuki	
<i>Unpurged results</i>						
<i>Message</i>						
Accountability	0.500 (0.071)	0.538 (0.144)	0.357 (0.133)	0.609 (0.104)	0.429 (0.111)	0.496 (0.046)
Fairness	0.405 (0.077)	0.357 (0.092)	0.393 (0.094)	0.500 (0.109)	0.409 (0.107)	0.408 (0.041)
None	0.361 (0.044)	0.348 (0.102)	0.263 (0.104)	0.391 (0.104)	0.571 (0.095)	0.381 (0.033)
Total	0.402 (0.034)	0.391 (0.061)	0.344 (0.061)	0.500 (0.061)	0.479 (0.060)	0.418 (0.023)
<i>Purged results</i>						
<i>Message</i>						
Accountability	0.136 (0.059)	0.193 (0.117)	−0.038 (0.141)	0.265 (0.083)	0.069 (0.082)	0.134 (0.038)
Fairness	0.021 (0.053)	0.036 (0.084)	0.127 (0.073)	0.150 (0.094)	0.116 (0.086)	0.079 (0.033)
No Message	0.000 (0.033)	0.046 (0.099)	−0.015 (0.100)	0.102 (0.098)	0.181 (0.068)	0.038 (0.027)
Total	0.035 (0.026)	0.072 (0.056)	0.045 (0.056)	0.171 (0.053)	0.127 (0.045)	0.074 (0.019)

Notes. Cell entries are the share of subjects who report an intention to vote yes, after accounting for average location and canvasser effects (the average of the quantities y'_i described in subsection 3.2); standard errors in parentheses. Upper panel (‘unpurged’) provides means and standard deviations of the raw data. Lower panel (‘purged’) shows results accounting for canvasser and location fixed effects and setting the placebo condition value to zero. Each reported number in this panel can be interpreted as an estimate of how more or less likely a subject in a given condition is to vote yes compared with a subject in the placebo condition, conditioning on location and canvasser effects.

$$\hat{\tau}_{m',m''} = \sum_{e \in E} \pi(e) \left(\sum_{e \cap m'} \frac{y'_i}{|e \cap m'|} - \sum_{e \cap m''} \frac{y'_i}{|e \cap m''|} \right). \quad (1)$$

The average effect without taking account of blocks can be read directly from the raw data (see Table 2) and, thanks to the randomisation, this also provides an unbiased estimate of treatment effects. Estimated effects reported below differ from these estimated effects only as a function of empirical imbalances in the distribution of messages across endorser conditions. Neyman standard errors are reported, taking account of the blocking by endorser.

Note that because we do not have a condition with a campaign that has no message or endorser (i.e. no ‘null campaign’), we are constrained in making inferences about the effect attributable to messages *per se* in message-based campaigns, although we can cleanly assess the differences in effects associated with different messages and the differences between message-based and endorser-based campaigns. Under an assumption which we call ‘null effects of null campaigns’, which means that a campaign with no message and no endorser would have no effect, the effect of messages in a message-based campaign is given simply by the effect of a message-based campaign.¹⁰

3.2.2. Charisma

Evaluating whether voters respond to contact with different canvassers is difficult as there is no natural control group. Nevertheless, even in such cases variation in the effects of different canvassers can be examined. This is the approach used by Humphreys *et al.* (2006). If our canvassers have differential effects, then we should expect that the post-treatment differences in opinion among the population respond in some way to exposure to the different canvassers. In practice and in the application we describe more fully below, canvassers worked in clusters of locations on a given day. To account for this we undertake analysis by defining outcomes at the cluster level. Again as we expect message and endorser treatments to affect outcomes we purge our dependent variables of these effects. In this case, the purged variable, y'_i , is simply the residual from the regression of the outcome on message–endorser fixed effects. Then, for canvasser c in location j we define: $\bar{y}'_{c,j} = (1/|c \cap j|) \sum_{i \in c \cap j} y'_i$ (in calculating this score we exclude the placebo units). To estimate the effects of canvassers across average responses in each location we can then deploy a standard F test to test the hypothesis that average outcomes on $\bar{y}'_{c,j}$ are invariant across canvasser identities.

3.2.3. Cue

For the analysis of cueing effects we can employ both types of analysis. That is, we can examine the effect of the presence of a cue-based campaign (relative to placebo) and we can examine whether there is variation by endorser. The estimate of the effect of the presence of an endorser is done in the same way as the examination of message

¹⁰ We note that this assumption is not innocent as there are ways in which even a null campaign could have an effect, for example, by signalling that there is a constituency that cares about the issue.

content; in this case purging location and canvasser effects and matching on message effects and estimating effects at the respondent level. To test whether it matters which endorser is associated with the campaign, we employ an F test.

The key outcome variable we examine is the subject's expected voting behaviour. In practice, this outcome takes three values for 478 subjects for whom we have data on this question: 'Yes' (42%), 'No' (21%) and 'Don't know' (38%).¹¹ Given the large number reporting that they do not know, it is possible that treatment could lead simultaneously to an increase in the propensity to vote yes and the propensity to vote no. For this reason we examine the effects of treatments on yes votes and no votes separately (and sometimes explicitly on don't know positions).

Finally, we note that for all analyses of endorser and messaging we report the intention to treat effects (ITT). This is a quantity of interest for strategies of persuasion and is well identified; moreover, the actual rates of (reported) non-compliance is very low (<5%) and in most cases not plausibly related to potential outcomes.¹² Although canvassers varied with respect to the days and the number of days they could work, we do not suffer from non-compliance regarding allocation of canvassers to locations as canvassers always visited the regions that they were dispatched to each day. There are no instances we know of in which a given subject was canvassed by one canvasser when they should have been canvassed by another. Nevertheless, as our subject population is defined as the subjects canvassed by canvassers, it is possible that a given subject is canvassed by one canvasser but would not have been canvassed by another had the other been allocated to the same area because of differences in the persistence of canvassers or the times of day they went to sites. Our balance tests do not uncover evidence of such selection effects, however.

4. Results

The basic results for the yes outcome can be read from Table 2. The Table shows the share of subjects who report an intention to vote 'Yes' before (upper section) and after (lower section) accounting for average location and canvasser effects.

The upper panel of the Table shows the raw data and the lower panel shows the 'purged' version of the outcome variable. In the purged version, the outcome in the placebo condition is fixed at zero and so each reported number can be interpreted as an estimate of how more or less likely a subject in a given condition is to vote yes

¹¹ Numbers subject to rounding error.

¹² For each household canvassed we asked canvassers to report whether they went off message or delivered the wrong treatment. In nearly all non-compliance cases described by enumerators the reason was not obviously related to features of the subjects. For example, one canvasser notes: 'I had to change the treatments given to #2 and #3 because I ran out of these handouts;' and another notes: 'I accidentally told the #1 person the fairness message and gave the accountability leaflet (which it should have been)'. In a very small number of cases, non-compliance is plausibly related to potential outcomes (one canvasser noted: 'I went off topic it was a placebo and he said he would definitely vote yes, I accidentally responded "Great!"'; another noted '#3 didn't wait for message' and another noted 'I went off message with #2 because her message was endorser only, and she was commenting on how parties that support childcare issues never get elected [... I mentioned that with STV...] there is a better chance that she would be locally represented by a party that has a childcare platform that she agrees with').

compared with a subject in the placebo condition, conditioning on location and canvasser effects.

Scanning the outcomes in the ‘total’ rows and columns provides initial insights into the effects of messages and endorsers. On average, we see that the propensity for supporting STV is highest among those receiving the accountability message, for example, and for those who received a treatment that contained an endorsement by a prominent individual associated with the campaign.

These aggregate effects are shown in Figure 2, which provides the estimate of message treatments conditioned on endorser treatments and *vice versa*, using the blocked estimator given in (1), as well as the overall campaign effect. The Figure reports the core results for our primary outcome of interest, the decision to vote ‘Yes’ for the reform. In addition, as subjects could report both a ‘No’ position and a ‘Don’t know’ position we report the results for these outcomes as well. Examining these is important as, in principle, any treatment could increase support for a campaign either by shifting uncertain voters into the yes camp, by producing shifts from the no camp to the yes camp or by shifting voters from the no camp to an uncertain position. The campaign, naturally, seeks an increase in the yes vote and a decline in the no vote.

The first result in Figure 2 shows the campaign effect on the likelihood of a yes in the referendum on the introduction of STV in British Columbia by comparison of the likelihood of a yes vote among the treated group (those contacted by the campaign) with the control (those who received the placebo). The estimated magnitude of this effect is in the region of 10 percentage points (CI: 2–18). In a one-tailed test this effect is significant at the 99% level. Looking down the column we see that the effect is associated with a drop in the no support and a slightly larger drop in the numbers

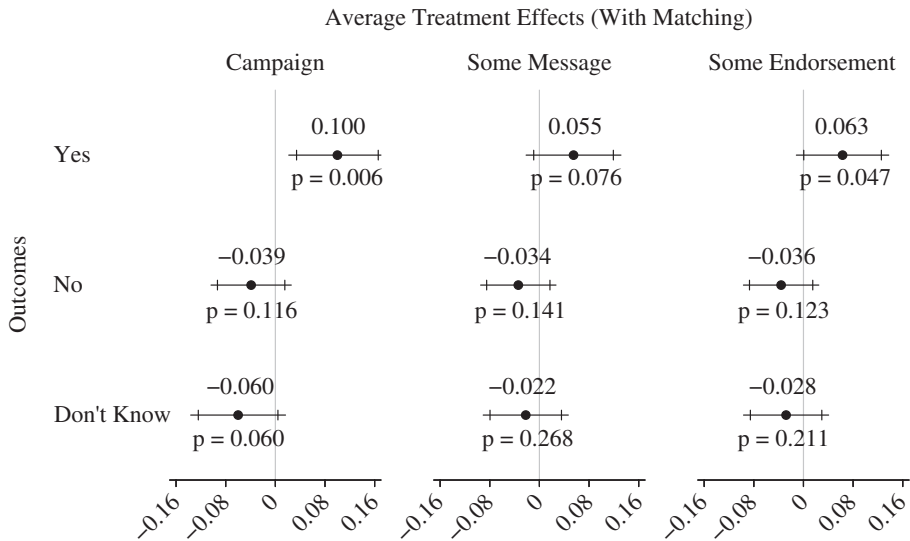


Fig. 2. Average Treatment Effects (With Matching)

Notes. Figure shows estimated average treatment effects for the overall campaign, for message-based campaigns (pooling over types of messages) and for endorser-based campaigns (pooling over endorsers) and three outcome measures (rows). 95% and 90% confidence intervals are indicated as well as p values from one-sided tests.

reporting a ‘Don’t know’ position. We see very similar patterns for both the ‘some endorser’ treatment and the ‘some message’ treatments with each of these posting effects on the order of about 6 percentage points. Finally, we note that the same analysis conducted without taking account of blocks leads to slightly stronger effects (see Figure D1 in Appendix D); however, an analysis that does not take account of canvasser and locational effects produces a considerably noisier estimate. In that case, the estimated effect of the overall campaign and of message-based campaigns drops marginally and attains significance at the 10% level only; and the effect of endorsement falls further and loses significance at conventional levels.

4.1. Content

How do message effects depend on the substantive messages deployed? The upper panel of Figure 3 shows the average effect of content: our estimates of $\hat{\tau}_{m,m'}$ where the

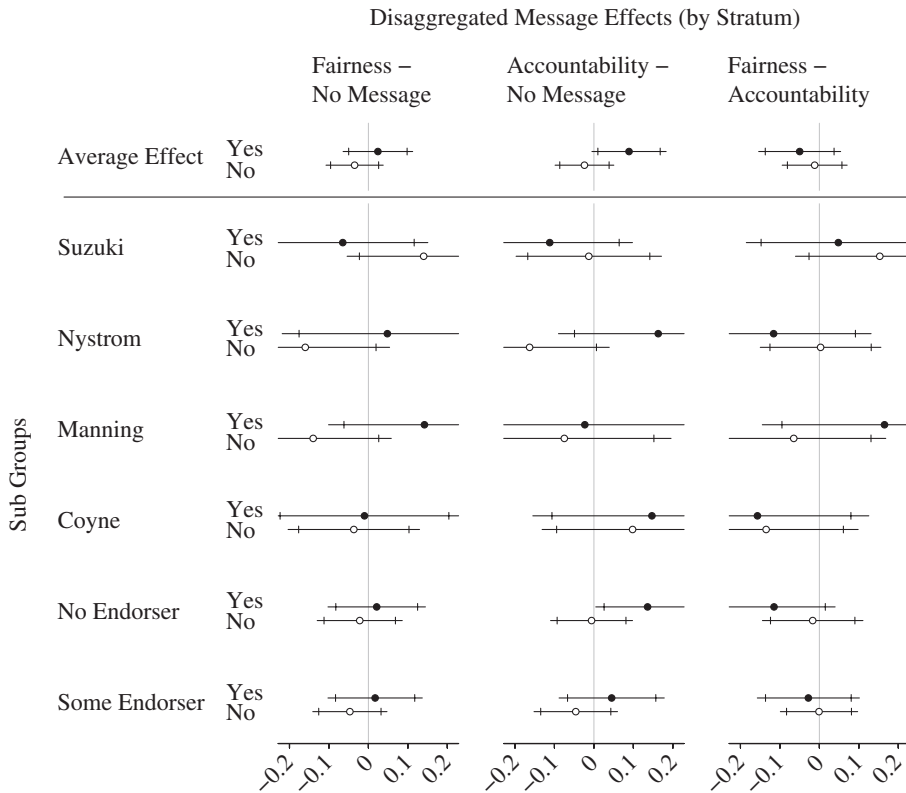


Fig. 3. Message Effects by Stratum

Notes. Figure shows estimated average treatment effects for individual message treatments (columns) and two outcome measures (marked in each row with white and black points). The top of the panel shows the estimated effects and the lower part shows effects by subgroup. 90% and 95% confidence intervals are indicated. The Figure shows strong effects for the accountability message, generated largely from the subgroup that had no endorsements. The fairness effect was weaker but the difference between these two effects is nevertheless small.

vote intention of the treatment group (those who received $m = m'$) is compared with that of the control group (those who received $m = \emptyset$), while conditioning on all geographic data and average canvasser effects. The Figure shows that the message-only effect is driven primarily by reactions to the accountability message. Invoking the assumption of null effects for null campaigns, we estimate the marginal effect of messages across treatments at 6 points (significant at the 90% level in a one-tailed test) which averages an effect from the accountability treatment (9 points, significant at the 95% level in a one-tailed test) and from the fairness treatment (2.4 points, not significant at conventional levels). The difference between these two message effects is itself not statistically significant. Moreover, we note that although the effect on a yes vote appears stronger for the accountability message, there is also evidence of less obvious gains from the fairness message associated with an increase in the don't knows (and a decline in the no vote). Thus, the fairness measure, although it had a weak effect on the yes vote, nevertheless had a moderate effect on the yes vote as a share of yes's and no's.

These marginal effects are averaged over endorser treatments. The lower panel of Figure 3 shows the fundamental effects within endorser strata as well as the average effects across strata for each message and for 'any message'; in addition, it shows the differential effects of the messages across strata. We see some evidence for heterogeneity across strata; however, little of this variation is statistically significant. In general, all messages had a positive effect across endorsers with the exception of Suzuki – for Suzuki we estimate a negative (although not significant) effect of both messages.

This analysis then provides some evidence that content matters for influence, but given our power constraints it does not provide strong evidence that the content of that content matters.

4.2. *Charisma*

To analyse the campaign effect that is due to charisma of canvassers, we examine whether those individuals who had been contacted by one or other of the canvassers were more likely to support reform. Note that in effect what we examine is the differences between the effects of different canvassers deployed by the campaign; we do not ask whether it makes a difference for there to be a canvasser or not but whether it makes a difference which individual (from the relevant population) the voter had contact with. We expect that, if charisma effects exist, variance in the opinions of respondents should reflect differential exposure to canvassers. Put another way: a condition for a charisma effect is that some canvassers are more persuasive than others.

To detect these differential effects, should they exist, we estimate a series of models controlling for canvasser fixed effects, using an F test to test the hypothesis that the joint difference between these fixed effects is zero. We examine the same three outcomes (yes, no and don't know) and condition our analysis first on cases in which the campaign made contact (i.e. we exclude placebo cases) and second on cases in which canvassers had a message that they could use to persuade voters. We expect canvasser effects to be stronger under the second condition as in this case they had material to work with. The results are presented in Table 3.

Table 3
Relative Impacts of Canvassers on Outcomes

Condition	Yes	No	Don't know
<i>All</i>			
Share of variance explained by identity of canvasser	0.03 (0.23)	0.03 (0.22)	0.04 (0.18)
<i>Message-based campaigns only</i>			
Share of variance explained by identity of canvasser	0.05 (0.15)	0.02 (0.28)	-0.03 (0.70)

Notes. ***p < 0.01, **p < 0.05, *p < 0.1. Table reports Adj-R² and p-values from F-test.

We see from Table 3 that in no case are we able to reject the null that outcomes are invariant to the identity of canvassers. As expected, the relationship between the canvasser's identity and a yes outcome is somewhat stronger under the message condition than under the campaign condition, but even in this case the relationship is not significant at conventional levels. Our results cast doubt on the view that persuasion depends strongly on the characteristics of those who are persuading. That 'leaders' do matter in opinion formation was one of the strongest findings of the study by Humphreys *et al.* (2006). There, as here, the 'leaders' (in our case, canvassers) were not well known figures in the communities to which they were assigned, so that their impact is likely due either to innate characteristics or the messages they deployed. As mentioned already, that study could not distinguish which. Here by conditioning on message used, our analysis suggests that conditional on content, the characteristics of canvassers play no role.

4.3. Cues

Our third area of investigation examines the effectiveness of campaign endorsements that might provide cues to respondents. Figure 4 reports the differences in outcomes associated with different endorsers. The first five rows and columns form a matrix with data points indicating the propensity to report 'yes' when exposed to a row type relative to the propensity when exposed to a column type. The final column then gives the average effect of exposure to the row type endorsement when compared with all other types.

We have already seen evidence that endorsement-only campaigns are effective. In Figure 2, we saw evidence of a 9 percentage point effect of an endorsement campaign. However, as seen in Figure 4 there is little evidence that it matters which public figure makes the endorsement. On average, Suzuki and Nystrom are associated with more positive results and Coyne and Manning are associated with more negative results. But these individual-level differences are not strong and the difference between the most successful endorser (Nystrom) and the least successful one (Manning) is not itself significant. Table 4 confirms the weakness of the between-endorser relations. For each outcome variable we test the null hypothesis that the identity of the endorser matters and in each case we are unable to reject the null.

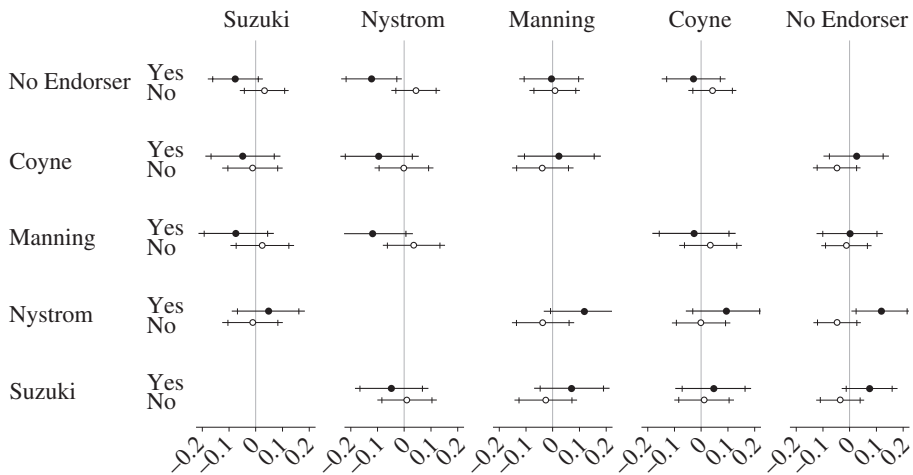


Fig. 4. Endorser Effects (Difference Between Row and Column Conditions)

Notes. Figure shows the difference in average outcomes for different endorsers, indicated in rows, relative to the column condition, including the ‘no endorser’ treatment for outcomes yes and no. Outcomes are purged of canvasser, location and message effects. 95% and 90% confidence intervals are marked.

Table 4

Relative Impacts of Endorsers on Outcomes

	Yes	No	Don't know
Share of variance explained by identity of endorser	0.000 (0.38)	-0.010 (0.92)	-0.001 (0.44)

Notes. ***p < 0.01, **p < 0.05, *p < 0.1. Adj-R² and p-values from F-test.

Grouped together there is some evidence of individual endorser effects. As described above, the four endorsers are balanced in two ways: two of the four are politicians, whereas the other two are media personalities; and two of the four are broadly associated with the left of the political spectrum and two with the right. These two dimensions of balance produce a two-by-two table. Analysing at the level of dimensions rather than individuals, we find that endorsements from politicians and endorsements from the left are more strongly associated with yes votes (see Table 5). The latter of these relations – reflecting the effects of Suzuki and Nystrom – is significant at the 95% level in a simple two-tailed test, without taking account of the multiple comparisons implicit in this Table.

Although suggestive, the evidence for endorser effects remains weak and as, in fact, each dimension only consists of two endorsers, these relations are properly thought of as the effects of two endorsers and not of the dimensions that distinguish them. Although groups of endorsers are associated with more positive effects, it remains the case that we are unable to reject the null that all endorser effects are zero.

Table 5
Propensity to Vote 'Yes' Given Exposure to Endorsers of Different Types

	Media	Political	Total	Difference (political <i>versus</i> media)
Right	0.01	0.01	0.01	-0.00 (0.06)
Left	0.08	0.13	0.10	0.05 (0.07)
Total	0.03	0.07	0.04	0.04 (0.04)
Difference (left <i>versus</i> right)	0.07 (0.05)	0.12 (0.08)	0.09 (0.04)**	

Notes. ***p < 0.01, **p < 0.05, *p < 0.1. Mean of purged variable; standard errors in parentheses.

4.4. Interactions

We have found positive evidence for overall cueing effects as well as for content effects (and for one message in particular). Recent work has found that such effects may complement each other. In a recent study, Bullock (2011) found that for one set of voters the estimated effects of content was identical for one subgroup whether or not they were exposed to cues; for another group, the effect of content was augmented by the cues (this latter effect was weakly significant).

Under the assumption that null campaigns have null effects, our estimates provide more evidence for substitution effects than for complementarities. Figure 3 provides the key evidence. From the top row of the Figure we see that the strongest effects for content are present for the 'accountability message'. The second to last row shows that these effects are especially strong and significant in the 'no endorser' subgroup (8.3 points). They are, however, closer to zero for the 'some endorser' group (2.8 points). This suggests a negative interaction between content and cue; consistent with substitution by voters.¹³ However, although we have evidence that the content effect is present with no endorser and no evidence that it is present with an endorser, we do not have the statistical power to confirm that the interaction is itself significant.

5. Power of Our Test

A concern is whether the weakness of the identified charisma effects reflects the power of our test. We address this question by estimating the probability of a relationship as weak or weaker than what we observe, if in fact the probability of voting 'yes' took the form $q = \underline{q} + \beta_i$, where \underline{q} reflects the average probability of voting yes and $\beta_i \sim N(0, \sigma^2)$ captures the increased or decreased likelihood of voting yes given that one is exposed

¹³ Relaxing the assumption of null effects from a null campaign, a null campaign would need to produce a 5.4 percentage point effect on the propensity to vote yes to produce an estimate consistent with complementarities.

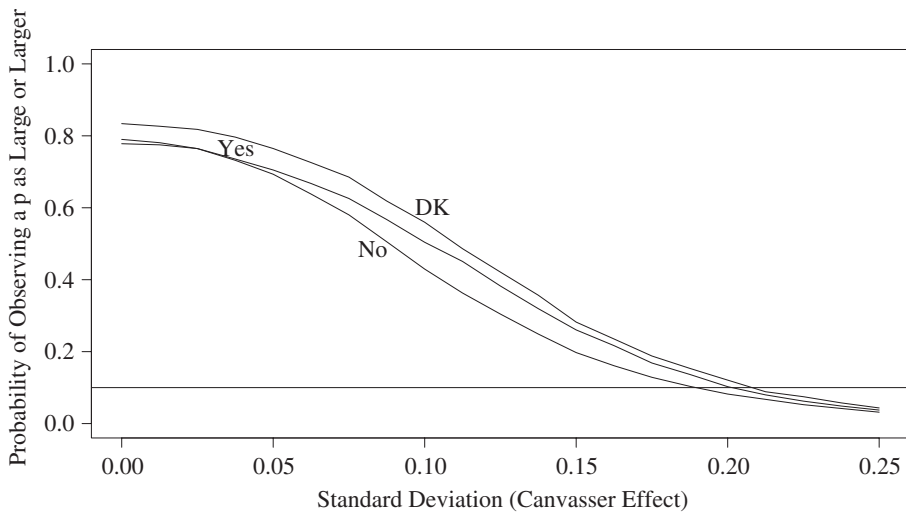


Fig. 5. Power Graph for Variation in Canvasser Effects

Notes. Figure shows the probability of observing evidence of individual charisma effects as weak or weaker than observed in our data for hypothetical values of σ for two outcome variables.

to canvasser i . For a range of values of σ and given the structure of our data, we use simulation to assess the probability of observing a p-value as high or higher than what we in fact observe. The results for values of σ , ranging from 0 to 0.25, are shown in Figure 5.

The Figure suggests that given our data structure, such weak results are reasonably likely for values of σ below about 0.2 and extremely unlikely for values above about 0.25. How big a value is a standard deviation of 0.25 in this context? One way to assess this quantity is to consider a simple structure in which the average outcome was q and that with a 0.5 probability a canvasser is ‘effective’ and produces outcome $q + \alpha$ and with 0.5 probability a canvasser is ‘ineffective’ and produces an outcome of $q - \alpha$. In this case the standard deviation for the distribution on β is simply $(0.5\alpha^2 + 0.5\alpha^2)^{0.5} = \alpha$. Thus, for canvasser effects of this form, a division of canvassers into those in which there is a 0.5 point spread between the effective and ineffective canvassers would be extremely unlikely to produce results so weak. There are approximately even odds, however, of getting results as weak as we find for spreads half as large. A second approach is to compare this quantity to past results. To do this we examined the data used by Humphreys *et al.* (2006) to estimate σ for each of the nine binary items in their study. We found that the average estimated standard deviation of enumerator effects in the nine binary variables examined in their study was 0.34 and that in seven of nine cases the estimated standard deviation was well above this, in fact above 0.5. Thus, although our results are consistent with small differences in effects across canvassers, they are not consistent with effect sizes on the scale of those found in past work.

We can ask a similar question with respect to our finding that, despite an overall effect of endorsement, there is weak evidence suggesting that it matters which individual endorsed the message. Are these non-results simply reflective of the low

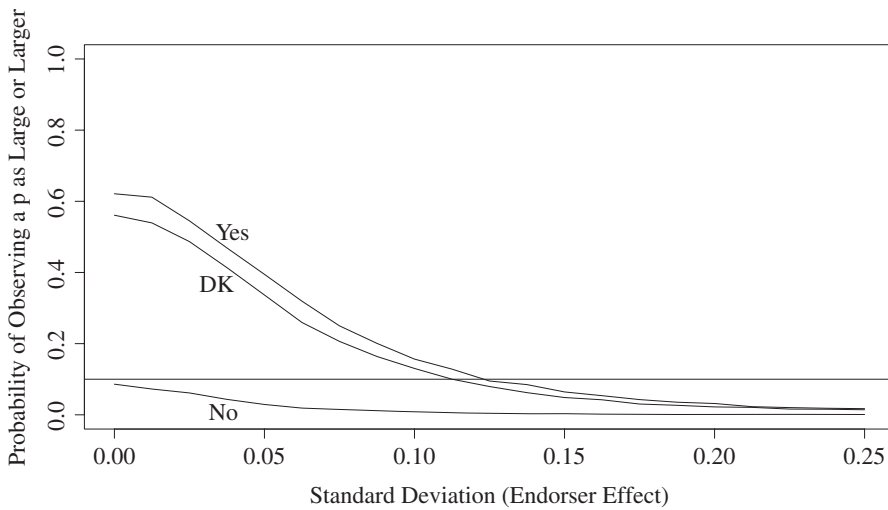


Fig. 6. *Power Graph for Variation in Endorser Effects*

Notes. Figure shows the probability of observing evidence of endorser effects as weak or weaker than observed in our data for hypothetical values of σ .

power of our test? We address the question by estimating the probability that one would find results as weak or weaker for different types of true effects. Employing the same strategy as before we estimate that, with even moderate variation in the effects of endorsers, we would be unlikely to see results as weak as we do. These results are shown in Figure 6.

6. Heterogeneous Effects

Our focus so far has been on features of political actors and not on those of the public they are trying to persuade. However, we can expect that the effectiveness of approaches to persuasion may depend significantly on who the recipients of these messages are. Moreover, if variation in features of the recipients is important, this can have implications for the generality of the results we find here. Recent studies, for example, show heterogeneous effects on persuasion of independent television in Russia according to prior political knowledge (Enikolopov *et al.*, 2011). In examining heterogeneous effects we are motivated by one particular consideration. In previous work, in developing countries there has been some evidence, consistent with what we find here, that message matters (Wantchekon, 2003). Other work, not consistent with our results here, has found very strong effects associated with individual leaders (Humphreys *et al.*, 2006). There are many features that differentiate the setting we examine from that in Humphreys *et al.* (2006); but one obvious one is the education level and, perhaps, political sophistication of respondents. Perhaps, individual ‘leaders’ had such an effect in São Tomé and Príncipe because of the relatively large education gap between leaders and respondents in that setting.

To examine this logic we turned to our data to investigate whether canvasser and endorser effects are stronger or weaker among different types of voters. We divided our voters in two ways. First, we distinguished between a group of ‘low education’ voters (those with secondary education or less; 27%) and ‘high education’ voters (those with tertiary-level education, including both university and vocational; 73%); second, we distinguished between voters based on political sophistication; in particular, we asked respondents to provide a guess of the number of seats in the provincial legislative assembly in Victoria. We classed subjects as ‘sophisticated’ if they estimated between 42 and 128 seats (50%) and ‘less sophisticated’ if they guess outside this range (50%). We have 250 subjects for whom we have data on both of these measures and for these there is a correlation of 0.13, which suggests that these measures, although related (the correlation is significant at the 95% level), pick out substantially different subsets of subjects.

In Figures 7 and 8, we show the estimates of the main message effects and of the ‘some endorser’ treatment for these subgroups. We find that the message effects are generally larger for the sophisticated subjects than for the less sophisticated ones and that the effect of having some endorser is about the same for the two groups. For the education breakdown we find similar effects of messages for the propensity to vote yes; but we also find an increased, although more moderate, effect on the propensity to vote no. The strongest effect of the campaign was to produce some opinion among less educated voters, although that opinion was not necessarily in the desired direction. This effect for the campaign derives especially from the presence of endorsers. In Table 6 we show the effects of canvassers and endorsers for these subgroups. For one in eight analyses there is a moderate suggestion of endorser effects – for low education subjects in particular. Although this result is in our hypothesised direction, more generally the evidence remains weak that it matters which endorser is assigned.

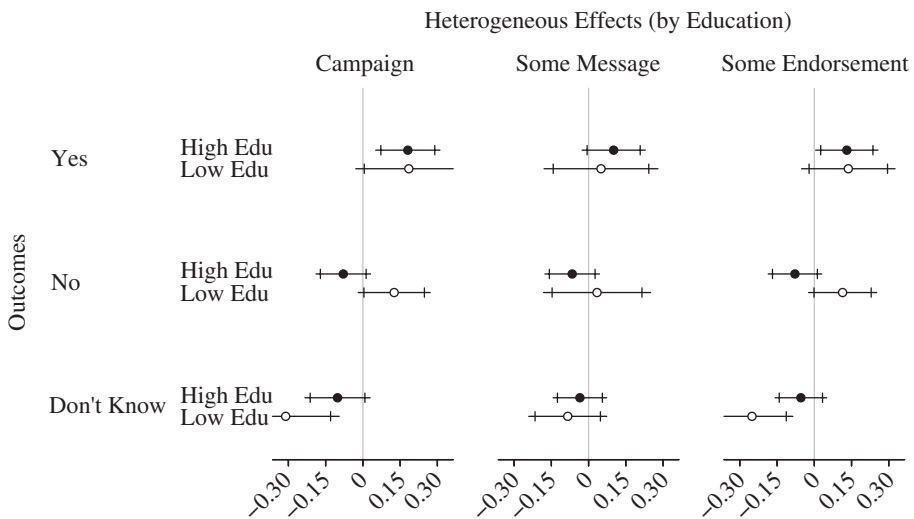


Fig. 7. Main Treatments Broken Down for ‘High Education’ and ‘Low Education’ Respondents
Notes. Columns represent independent variables (treatments) and rows dependent variables (outcomes). Horizontal bars show 90% confidence intervals.

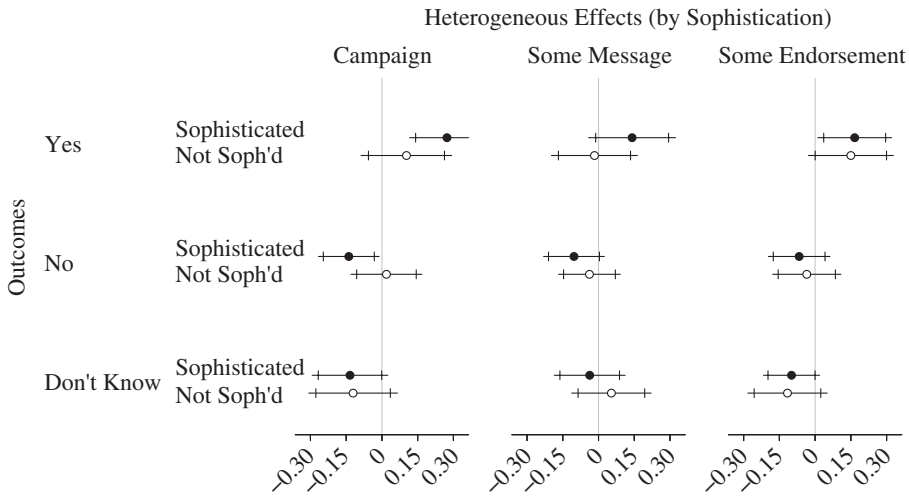


Fig. 8. *Main Treatments Broken Down for ‘Sophisticated’ and ‘Not Sophisticated’ Respondents*
Notes. Columns represent independent variables (treatments) and rows dependent variables (outcomes). Horizontal bars show 90% confidence intervals.

Table 6
Impacts of Canvassers and Endorser on Yes Outcomes (Heterogeneous Effects)

Treatment	Low education	High education	Low sophistication	High sophistication
Canvassers	-0.12 (0.74)	0.02 (0.36)	-0.007 (0.49)	0.04 (0.29)
Endorsers	0.13 (0.08)*	0.02 (0.22)	0.04 (0.19)	0.04 (0.15)

Notes. ***p < 0.01, **p < 0.05, *p < 0.1. Table shows the variation in outcomes that can be attributed to variation in the identity of individual canvassers and individual endorsers. Adj-R² and p-values from F-test.

Overall this analysis does not suggest that the differences in effects we find in British Columbia are due to high levels of education or political sophistication.

7. Missing Data

The low canvasser contact rate (40%) and low enumerator response rate among those contacted by canvassers (51%) limit our ability to perform subgroup analyses of various forms. But does it also introduce bias? Our design is such that we are protected from bias on the message and endorser treatments from the canvasser contact rate – treatment, including placebo treatment, is orthogonal to contact success. However there is a concern that the assignment to treatment could affect response rates for our enumerators; if this occurs it could introduce bias in our analysis.

We examine the data to assess the importance of this concern for each of the dimensions under study. Figure D2, in Appendix D, shows the relation between

treatments and non-responses for all subjects and for the heterogeneous population examined in Section 5.

The evidence for the effects of ‘some endorser’ on the data missing is the most concerning pattern in the data. Here, there is an estimated 7–8 percentage point effect. We cannot determine whether this non-response introduces bias into the results or the direction of bias, if there is any. But we can estimate bounds on effects under extreme assumptions regarding the potential outcomes of those that received an endorsement treatment and subsequently dropped from our sample. These bounds are large, ranging from –17% to 37% and clearly possibly large enough to wipe out the estimated 7% endorsement effect. Probing more deeply we find, however, that only about 2 percentage points of the difference is related to refusal to answer questions, other reasons for missing data seem less plausibly related to the endorsement treatment, including enumerators and respondents sharing no common language (2 points), no one being at home (2.5 points) and enumerators not being able to find the household in the first place (half a point).

For the message treatments (first four columns) we find little cause for concern that treatment may have resulted in greater missing data. Of the 20 estimates only one approaches statistical significance in a two-tailed test – the effect of the fairness message on non-response rates of sophisticated voters.

In addition, contact with the campaign in general is not correlated with non-responsiveness and, moreover, receipt of a particular message does not affect the sample of respondents. Probing further, in Table D1, in Appendix D, we show results of tests to see whether some canvassers, or some endorsers, are more likely than others to lead to high non-response rates. In each case we test the hypothesis that missing data are invariant to the identity of the canvasser or the endorser, conditioning on cases in which subjects were exposed to the campaign (first row) or to substantive messages (second row). In no case is the relation between canvasser or endorser identity and missing data significant at the 90% level, although in one case (for endorsers when we do not condition upon message) the relationship is borderline significant.

8. Conclusion

Previous work in economics and political science has suggested that political campaigns can be effective and that voters’ beliefs are influenced by them. However, campaigns are multifaceted and deploy different elements that each can contribute to its overall persuasiveness: Do such campaigns matter and, if so, which if any of its elements are effective at persuasion? To our knowledge, until now, no study has been able to provide answers to these questions. This is, perhaps, unsurprising as typical campaign strategies will confound attempts to provide answers.

We present findings from a field experiment involving a door-to-door campaign that randomised contact with canvassers, the content of campaign messages and campaign endorsements by leading public figures. Working together with the BC-STV campaign in the May 2009 British Columbia referendum on electoral reform we randomly assigned canvassers to voting areas, who then randomly assigned campaign messages and information about endorsements to households. This allowed us to analyse both message-based and endorsement-based campaigns. More importantly, it allowed us to

ascertain whether message content matters independently of latent traits such as the charisma of the canvasser delivering the message; separate the impact of endorsement-based campaigns that seek to provide voters with cues from that of message-based campaigns that emphasise content and analyse whether it mattered which message or endorsement was used.

Our analysis reveals four findings. First, we provide additional evidence that even simple campaigning of this form persuades voters: where the campaign sought to convince voters, they were more likely to support the proposed reform to the voting rules. Second, we find that both message-based and endorser-based campaigns appear effective to approximately equal extents. Third, we find surprisingly little evidence of differential effects between endorsers, and the evidence for differential effects of messages is statistically weak. Finally, there is little to no evidence suggesting that it mattered which canvasser made contact with voters.

These findings confirm the potential for strong shifts in preferences from relatively simple efforts in political persuasion. Voters responded to arguments and to endorsements but the individual characteristics of persuaders did not play an important role. Recall that the conclusion of an earlier experiment in the field was precisely the opposite: in the study by Humphreys *et al.* (2006, p. 604) the authors argued that '[l]eaders matter profoundly ... Knowing which member of the country's political elite was randomly selected to lead the discussions provides an extraordinarily powerful indicator of what policies the participants in each group ostensibly supported'. There, as here, leaders were not well known figures in the community. That study could not establish whether leaders were effective because of the arguments they selected to employ or because of features that are truly specific to those individuals. Here, we distinguish these components through experimental control and, conditional on messaging strategies, find no evidence that individual characteristics play any role in the success of attempts at persuasion.

Similarly, our examination of cues found evidence for campaign effects, but only weak evidence that the identity of endorsers mattered. These results suggest that in this environment at least, the ability to persuade may not be tied strongly to individual characteristics that operate independent of message content. This conclusion is consistent with two possibilities that we cannot distinguish here precisely because of the controlled nature of our design (Gaines and Kuklinski, 2011). One is that individual effects matter little: through message testing and message selection, persuasive power may be transferred from a centralised campaign to local leaders. The second, more subtle, possibility is that effective leaders may be effective precisely because of their ability to select the right message and so say the right thing at the right time.

Appendix A. Campaign Material

This Appendix provides screenshots of the message flyers and one of the endorser flyers used during the campaigns.

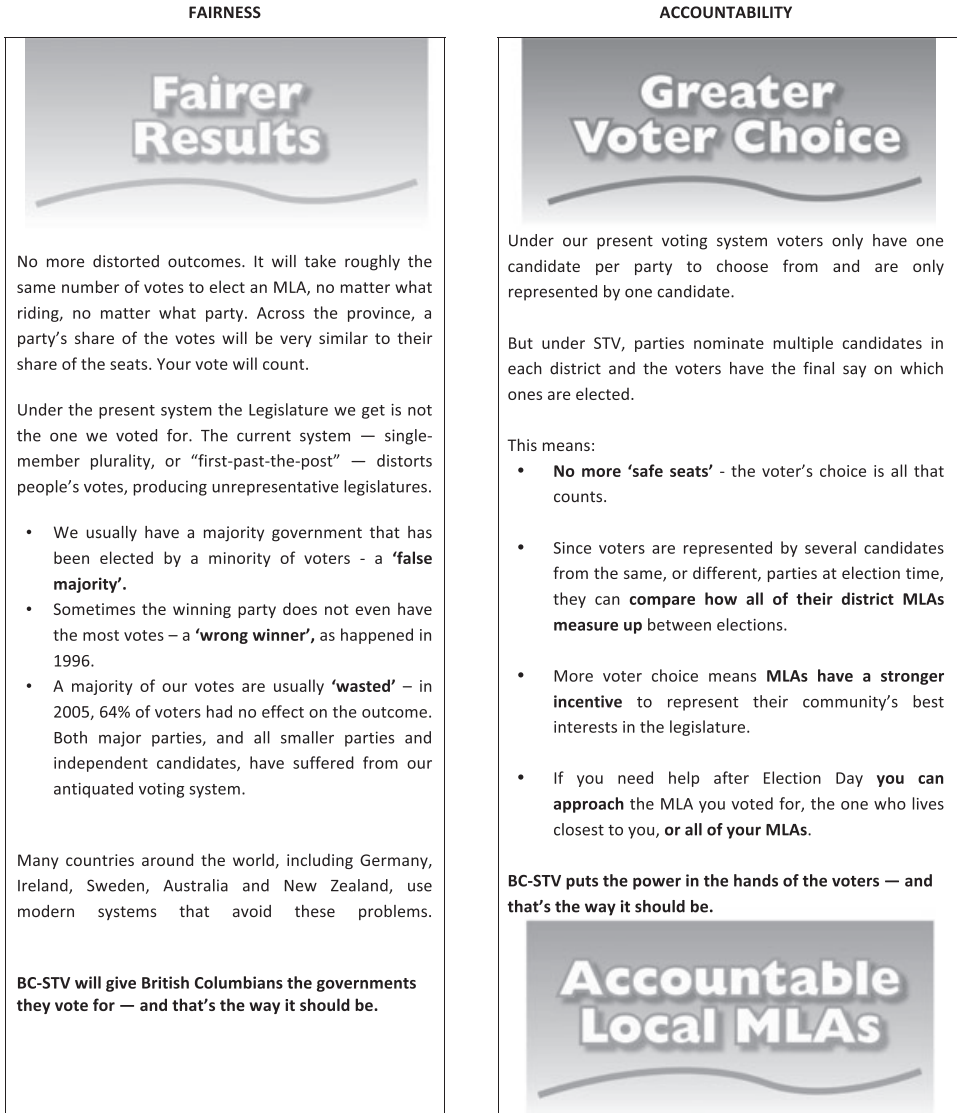


Fig. A1. Messages

Appendix C. Survey Instrument

MARCH 2009 STV VOTER SURVEY		Riding: _____	Survey ID: _____	Non Response: <input type="radio"/> No one at home <input type="radio"/> No common language <input type="radio"/> Refused to respond				
Enumerator ID: _____	Voting Area: _____	Date: _____	Language: _____					
[Can I speak to a [male/female/you] eligible voter in this household?]								
CONSENT SCRIPT: Hello, I am undertaking a survey on behalf a research team from Ryerson, Columbia and LSE universities. We are trying to understand attitudes of people in British Columbia towards the upcoming referendum on the way that British Columbians elect their representatives. There are only 20 questions and we estimate that the survey will take approximately 10 minutes. This is an anonymous survey. I do not need to ask or record any names, and while I have your address, this information will not be associated with responses.								
To respondent: [Repeat non bracketed part above] Please know that you do not have to answer any questions if you don't want to and you can stop this interview at any stage if you wish to. Are you willing to take part in this survey?								
Q 1	Have you heard of the referendum on British Columbia's electoral system?	<input type="radio"/> ¹ Yes	<input type="radio"/> ² No					
*Q 2	Do you expect to vote in the May provincial elections ?	<input type="radio"/> Not likely <input type="radio"/> Likely <input type="radio"/> Very likely	<input type="radio"/> ⁴ Other					
Q 3	Do you expect to vote in the referendum to change British Columbia's voting system to STV ?	<input type="radio"/> Not likely <input type="radio"/> Likely <input type="radio"/> Very likely	<input type="radio"/> ⁴ Other					
Q 4	If not likely: Why not?	<input type="radio"/> Unavailable <input type="radio"/> I don't understand issues enough <input type="radio"/> Vote has no effect <input type="radio"/> I am indifferent						
Q 5	<i>If likely/ very likely:</i> Do you expect to vote for or against the proposed reform? <i>Otherwise:</i> If you were to vote, would you vote for or against the proposed reform?	<input type="radio"/> Likely to vote FOR <input type="radio"/> Likely to vote AGAINST <input type="radio"/> Leaning towards FOR <input type="radio"/> Leaning towards AGAINST						
Q 6	People's opinions often change in the run-up to a referendum. Compared to one week ago would you say that you lean more towards a yes (for reform), more towards a no, or have you not changed your opinion?	<input type="radio"/> No Change <input type="radio"/> More Yes <input type="radio"/> More No <input type="radio"/> ⁴ Other						
Q 7	How strongly do you agree or disagree with the following statements:	Agree strongly	Agree	Disagree	Disagree strongly	DK	Ref'd	
	A Our present voting system works well and does not need to be changed	<input type="radio"/> ^{A1}	<input type="radio"/> ^{A2}	<input type="radio"/> ^{A3}	<input type="radio"/> ^{A4}	<input type="radio"/> ^{A5}	<input type="radio"/> ^{A6}	
	B STV would lead to a more diverse representation in the legislative assembly.	<input type="radio"/> ^{B1}	<input type="radio"/> ^{B2}	<input type="radio"/> ^{B3}	<input type="radio"/> ^{B4}	<input type="radio"/> ^{B5}	<input type="radio"/> ^{B6}	
	C Under STV the results of elections would better reflect the way people vote	<input type="radio"/> ^{C1}	<input type="radio"/> ^{C2}	<input type="radio"/> ^{C3}	<input type="radio"/> ^{C4}	<input type="radio"/> ^{C5}	<input type="radio"/> ^{C6}	
	D STV would lead to weak political parties	<input type="radio"/> ^{D1}	<input type="radio"/> ^{D2}	<input type="radio"/> ^{D3}	<input type="radio"/> ^{D4}	<input type="radio"/> ^{D5}	<input type="radio"/> ^{D6}	
	E STV would make politicians more responsive to the needs of voters	<input type="radio"/> ^{E1}	<input type="radio"/> ^{E2}	<input type="radio"/> ^{E3}	<input type="radio"/> ^{E4}	<input type="radio"/> ^{E5}	<input type="radio"/> ^{E6}	
	F STV is too complicated a system	<input type="radio"/> ^{F1}	<input type="radio"/> ^{F2}	<input type="radio"/> ^{F3}	<input type="radio"/> ^{F4}	<input type="radio"/> ^{F5}	<input type="radio"/> ^{F6}	
Q 8	Which of these is the most important consideration for you when deciding whether to vote in favor or against STV:	___ [item A - F]. Reread list from Q 7 <input type="radio"/> ¹ Other						
Q 9	For each of these parties, do you think that most members of the parties support or oppose the reform?	<input type="radio"/> ¹ Lib Support <input type="radio"/> ² Lib Oppose	<input type="radio"/> ³ NDP Support <input type="radio"/> ⁴ NDP Oppose	<input type="radio"/> ⁵ Gn Support <input type="radio"/> ⁶ Gn Oppose				
*Q 10	i. Do you know who X (Y) is?	**ii. Do you believe that X's (Y's) political positions are generally consistent with yours?	**iii. Do you believe that X (Y) has a good understanding of the issues on STV?	**iv. Do you know whether X (Y) supports or opposes the reform?	** v. If yes: How do you know this?			
	X <input type="radio"/> ¹ Yes <input type="radio"/> ² No	<input type="radio"/> ¹ Yes <input type="radio"/> ² No	<input type="radio"/> ¹ Yes <input type="radio"/> ² No	<input type="radio"/> ¹ Supp <input type="radio"/> ² Opp <input type="radio"/> ³ DK	<input type="radio"/> ¹ campgn <input type="radio"/> ² otr			
	Y <input type="radio"/> ¹ Yes <input type="radio"/> ² No	<input type="radio"/> ¹ Yes <input type="radio"/> ² No	<input type="radio"/> ¹ Yes <input type="radio"/> ² No	<input type="radio"/> ¹ Supp <input type="radio"/> ² Opp <input type="radio"/> ³ DK	<input type="radio"/> ¹ campgn <input type="radio"/> ² otr			
Q 11	Have you or others in your household been contacted by people arguing in favor or against STV?	<input type="radio"/> ¹ No <input type="checkbox"/> ² You, In Favor <input type="checkbox"/> ³ Others, In Favor <input type="checkbox"/> ⁴ You, Against <input type="checkbox"/> ⁵ Others, Against						
Q 12	*How did you vote in the LAST referendum for STV ?	<input type="radio"/> Voted against <input type="radio"/> In favor of STV <input type="radio"/> Did not vote <input type="radio"/> Other						
Q 13	*Mark the respondent's gender :	<input type="radio"/> Male <input type="radio"/> Female						
Q 14	*Year of Birth	19 ___						
Q 15	* Employment	<input type="radio"/> unemployed <input type="radio"/> working for pay, <input type="radio"/> selfemployed <input type="radio"/> student <input type="radio"/> retired <input type="radio"/> working at home						
Q 16	* Education level	<input type="radio"/> < high school <input type="radio"/> completed high school <input type="radio"/> some univ/cllge <input type="radio"/> othr 3rd lvl <input type="radio"/> cmlpltd uni degree						
Q 17	*How many eligible male and female voters live in this household?	___ male ___ female						
Q 18	*What were the ethnic/ regional origins of your ancestors? [Mark as many as apply]	<input type="checkbox"/> Abrg, <input type="checkbox"/> Brtsh <input type="checkbox"/> Fr <input type="checkbox"/> Oth Eur <input type="checkbox"/> Afr <input type="checkbox"/> MENA <input type="checkbox"/> S.As.n <input type="checkbox"/> E.As.n <input type="checkbox"/> W.As.n <input type="checkbox"/> LaAm&Car <input type="checkbox"/> 12.N.Am						
Q 19	*For which party do you expect to vote in the upcoming provincial elections?	<input type="radio"/> BC Liberal Party <input type="radio"/> New Democratic Party <input type="radio"/> Green Party <input type="radio"/> Cons <input type="radio"/> None <input type="radio"/> Other						
Q 20	*Can you tell me approximately how many seats there are in the provincial legislative assembly in Victoria?	_____						
Contact Can we contact you for a very short follow up after the vote? If yes, can you give us your email address or telephone number and a contact name? Your number will not be shared with third parties or linked to any responses you have given here.		<input type="radio"/> Yes <input type="radio"/> No	Name _____				Email and/or No. _____	

* If answer to Q1 is "No" then only ask starred questions ** Ask double starred questions only if answer to Q10 (i) is Yes

Fig. C1. Survey Instrument

Appendix D. Robustness

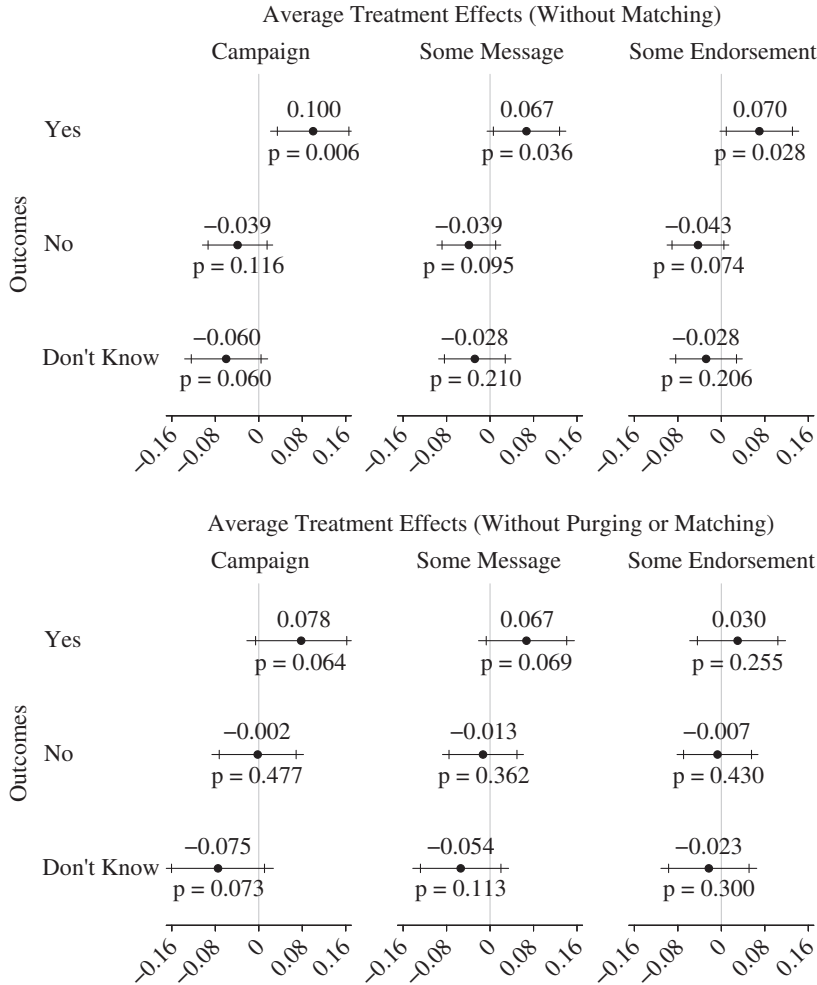


Fig. D1. *Average Treatment Effects (Without Matching (upper panel) and Without Purging or Matching (lower panel))*

Notes. Figure shows estimated average treatment effects for the overall campaign, for message-based campaigns (pooling over types of messages) and for endorser-based campaigns (pooling over endorsers) and three outcome measures (rows). 95% and 90% confidence intervals are indicated as well as p-values from one-sided tests.

Table D1
Relative Impacts of Campaigners and Endorsers on Missing Data

Condition	Canvasser	Endorser
Campaign	0.03 (0.26)	0.004 (0.105)
Message	0.02 (0.33)	-0.003 (0.710)

Notes. ***p < 0.01, **p < 0.05, * p < 0.1. Table reports Adj-R² and p-values from F-test.

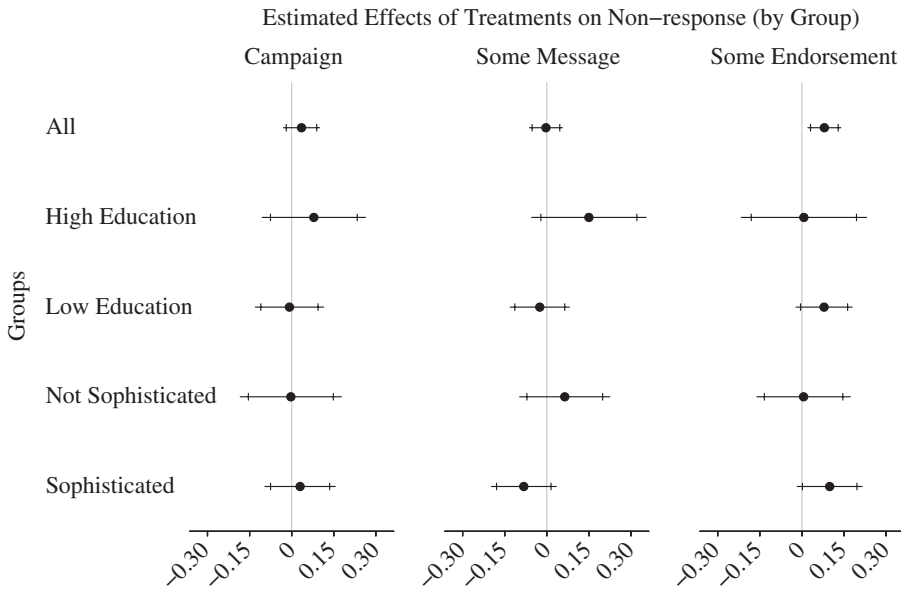


Fig. D2. *Main Treatments Broken Down by Non-response*

Notes. Columns represent independent variables (treatments) and rows dependent variables (outcomes). Horizontal bars show 90% confidence intervals.

Appendix E. Treatment Conditions

Table E1
Distribution of Treatments

Canvasser	Endorser treatment				Message treatment			Total	
	A	B	C	D	None	None	1		2
1	16	16	18	13	45	48	28	32	108
2	16	13	21	13	45	47	32	29	108
3	17	15	14	17	45	48	31	29	108
4	16	13	18	16	45	49	29	30	108
5	15	17	16	15	45	48	30	30	108
6	15	18	16	14	45	48	31	29	108
7	12	18	17	16	45	50	29	29	108
8	14	17	15	17	45	47	28	33	108
9	18	17	12	16	45	47	29	32	108
10	19	16	13	15	45	47	32	29	108
11	17	14	16	16	45	49	31	28	108
12	19	14	13	17	45	48	30	30	108
13	13	19	17	14	45	47	31	30	108
14	16	17	16	14	45	49	30	29	108
15	13	16	18	16	45	48	29	31	108
16	14	14	18	17	45	49	29	30	108
17	17	18	12	16	45	46	31	31	108
18	18	15	15	15	45	50	30	28	108
19	19	14	16	14	45	46	31	31	108

Table E1
(Continued)

Canvasser	Endorser treatment					Message treatment			Total
	A	B	C	D	None	None	1	2	
20	19	13	17	14	45	48	29	31	108
21	14	18	13	18	45	49	30	29	108
22	13	19	14	17	45	50	29	29	108
23	14	13	16	20	45	48	29	31	108
24	14	14	17	18	45	46	32	30	108
Total	378	378	378	378	1,080	1,152	720	720	2,592

Appendix F. Balance

Here, we provide balance tests to test the joint hypothesis that our randomisation was carried out without administrative errors and that non-delivery of messages or non-response to survey enumerators did not introduce imbalance on relevant covariates across treatment conditions. The Table F1 reports results from ANOVA models run on each treatment where we test for balance on the following individual-level covariates: gender, age, employment and education. As can be seen in the Table, these variables do not differ between treatment conditions; with only one in 12 tests approaching significance at the 90% level.

Table F1
Balance Tests

	Canvasser	Treatment Message	Endorser
<i>Covariate</i>			
Gender	-0.002 ($n = 419$) (0.554)	-0.0003 ($n = 421$) (0.354)	-0.002 ($n = 421$) (0.967)
Age	-0.004 ($n = 334$) (0.527)	-0.034 ($n = 336$) (0.770)	0.028 ($n = 336$) (0.257)
Employment	0.0003 ($n = 244$) (0.409)	0.0140 ($n = 246$) (0.136)	0.014 ($n = 246$) (0.133)
Education	-0.001 ($n = 240$) (0.443)	0.014 ($n = 242$) (0.097)	-0.010 ($n = 242$) (0.908)

Notes. The first row in each subrow reports the adjusted R^2 and n from ANOVA tests (one ANOVA for each treatment-covariate combination). p-values are given in parentheses.

London School of Economics

Columbia University

Ryerson University and Research Institute of Industrial Economics (IFN)

Additional Supporting Information may be found in the online version of this article:

Data S1.

References

Arceneaux, K. (2008). 'Can partisan cues diminish democratic accountability?', *Political Behavior*, vol. 30(2), pp. 139–60.

- Arceneaux, K. and Kolodny, R. (2009). 'Educating the least informed: group endorsements in a grassroots campaign', *American Journal of Political Science*, vol. 53(4), pp. 755–70.
- Arceneaux, K. and Nickerson, D. (2005). 'Even if you have nothing nice to say, go ahead and say it: two field experiments testing negative campaign tactics', Paper presented at the Annual Meeting of the American Political Science Association.
- Arceneaux, K. and Nickerson, D. (2010). 'Comparing negative and positive campaign messages evidence from two field experiments', *American Politics Research*, vol. 38(1), pp. 54–83.
- Berinsky, A.J. (2009). *In Time of War: Understanding American Political Opinion from World War II to Iraq*, Chicago, IL: University of Chicago Press.
- Brader, T.A. and Tucker, J.A. (2010). 'Follow the leader: party cues, policy opinion, and the power of partisanship in three multiparty systems', Working Paper, New York University.
- British Columbia (2004). 'Making every vote count: the case for electoral reform in British Columbia', Final report of the BC Citizens' Assembly on Electoral Reform.
- Bullock, J. (2011). 'Elite influence on public opinion in an informed electorate', *American Political Science Review*, vol. 105(3), pp. 496–515.
- Cohen, J. (1990). 'Things I have learned (so far)', *American Psychologist*, vol. 45(12), pp. 1304–12.
- Conover, P. and Feldman, S. (1989). 'Candidate perception in an ambiguous world: campaigns, cues and inference processes', *American Journal of Political Science*, vol. 33(4), pp. 912–40.
- Della Vigna, S. and Gentzkow, M. (2010). 'Persuasion: empirical evidence', *Annual Review of Economics*, vol. 2, pp. 643–69.
- Dewan, T. and Myatt, D.P. (2007). 'Leading the party: coordination, direction and communication', *American Political Science Review*, vol. 101(4), pp. 825–43.
- Druckman, J.N., Hennessy, C.L., Kristi, C.S. and Webber, J. (2010). 'Competing rhetoric over time: frames versus cues', *Journal of Politics*, vol. 72(1), pp. 136–48.
- Enikolopov, R., Petrova, M. and Zhuravskaya, E. (2011). 'Media and political persuasion: evidence from Russia', *American Economic Review*, vol. 101(7), pp. 3253–85.
- Gaines, B.J. and Kuklinski, J.H. (2011). 'Experimental estimation of heterogeneous treatment effects related to self-selection', *American Journal of Political Science*, vol. 55(3), pp. 724–36.
- Gerber, A.S. and Green, D.P. (2000). 'The effects of canvassing, direct mail, and telephone contact on voter turnout: a field experiment', *American Political Science Review*, vol. 94(3), pp. 653–63.
- Gerber, A.S., Green, D.P. and Larimer, C.W. (2008). 'Social pressure and voter turnout: evidence from a large-scale field experiment', *American Political Science Review*, vol. 102(1), pp. 33–48.
- Humphreys, M., Masters, W. and Sandbu, M.E. (2006). 'The role of leadership in democratic deliberations: results from a field experiment in São Tomé and Príncipe', *World Politics*, vol. 58(4), pp. 583–622.
- Kahneman, D., Slovic, P. and Tversky, A. (1982). *Judgement Under Uncertainty: Heuristics and Biases*, Cambridge: Cambridge University Press.
- Kendall, C., Nannicini, T. and Trebbi, F. (2013). 'How do voters respond to information? Evidence from a randomized campaign', Discussion Paper No. 7340, IZA.
- Kinder, D.R., Peters, M.D., Abelson, R. P. and Fiske, S. T. (1980). 'Presidential prototypes', *Political Behavior*, vol. 2(4), pp. 315–37.
- Loewen, P.J. and Rubenson, D. (2011). 'Democratic competition increases political participation', Working Paper, University of Toronto.
- Loewen, P.J., Rubenson, D. and Spirling, A. (2012). 'Testing the power of arguments in referendums: a Bradley–Terry approach', *Electoral Studies*, vol. 31(1), pp. 212–21.
- Lupia, A. (1994). 'Shortcuts versus encyclopedias: information and voting behavior in California insurance reform elections', *American Political Science Review*, vol. 88(1), pp. 63–76.
- Malhotra, N. and Kuo, A.G. (2007). 'Attributing blame: the public's response to hurricane Katrina', *Journal of Politics*, vol. 70(1), pp. 1–16.
- McNulty, J.E. (2005). 'Phone-based GOTV – what's on the line? Field experiments with varies partisan components, 2002–2003', *The Annals of the American Academy of Political and Social Science*, vol. 601(1), pp. 41–65.
- Panagopoulos, C. (2009). 'Partisan and nonpartisan message content and voter mobilization: field experimental evidence', *Political Research Quarterly*, vol. 62(1), pp. 70–6.
- Rahn, W. (1993). 'The role of partisan stereotypes in informational processing about political candidates', *American Journal of Political Science*, vol. 37(2), pp. 472–96.
- Riggle, E.D., Ottati, V.C., Wyer, R.S., Kuklinski, J. and Schwarz, N. (1992). 'Bases of political judgements: the role of stereotypic and nonstereotypic information', *Political Analysis*, vol. 14(1), pp. 67–87.
- Wantchekon, L. (2003). 'Clientelism and voting behavior: evidence from a field experiment in Benin', *World Politics*, vol. 55, pp. 399–422.
- Wantchekon, L. (2009). 'Can informed public deliberation overcome clientelism? Experimental Evidence from Benin', Working Paper, NYU.