The Impact of Leaders and the Messages they Convey: Evidence From a Field Experiment in British Columbia

Torun Dewan*, Macartan Humphreys† and Daniel Rubenson‡

Paper for presentation at the conference on 'Inequality and Institutions', Nuffield College, Oxford University , May 2009

Abstract. Are leaders effective because of some innate qualities—for example, clarity, trustworthiness or focality—or because of the particular arguments they employ? To analyze these effects systematically we need variation in both messages and leaders. Whilst these conditions are satisfied in many political settings, and in particular during political campaigns, observational studies are likely to produce biased estimates of these effects due to correlation, selection-bias and endogeneity. We describe a hierarchical model which allows us to untangle the effects of leader qualities from those of the arguments leaders use. We implement this model using data from a unique field experiment. Working together with the BC-STV campaign in the May 2009 British Columbia referendum on electoral reform, we randomly assign canvassers to voting areas, who randomly messages to households. We analyze leadership effects in two ways: (i) by estimating the between voting area differences in opinion that are due to the assignment of different canvassers; and (ii) by allowing some of the messages to be endorsed by leading public figures. Neither of these strategies yields any evidence of leadership effects: a fairly substantial overall campaign effect is entirely due to the arguments used by the campaign, not to individuals making the case for reform.[§]

*London School of Economics. Email: t.dewan@lse.ac.uk
†Columbia University. Email: mh2245@columbia.edu
‡Ryerson University. Email: rubenson@ryerson.ca

§We thank members of the Citizens Assembly and Fair Voting BC, especially Arjun Singh, Wendy Bergerud, Maxwell Anderson and members of the BC-STV campaign Susan Anderson-Behn, Maggie Gilbert and David Gagnon as well as the canvassers and enumerators involved in carrying out the project. We thank our research project managers, Stewart Prest and Pierce O Reilly, for excellent research support and for handling the coordination of the enumeration side of implementing this study. Don Green and David Epstein gave thoughtful comments on the project design. Antje Ellermann and Alan Jacobs provided support during the field work in British Columbia and also made useful suggestions on the design. Kjell Rubenson provided us with office space at UBC during the field work. We thank Columbia University, the London School of Economics and Ryerson University for financial support of this research.

1. INTRODUCTION

Do leaders affect public opinion? The answer to this question appears both straightforward and positive, but we should pause before reaching such a conclusion in haste. In fact there is as yet no clear scientific evidence discerning the impact of leaders on perceptions of public issues; therefore, and, somewhat surprisingly perhaps, the role that leaders play in influencing public opinion remains one of the most poorly understood features of political life. The problem is a familiar one: whilst contact with leaders may be correlated with attitudinal change, it is not clear that contact with a leader causes such change.

Indeed, it is hard to separate the impact of a leader from other confounding features that might influence opinion formation. In particular, it is hard to disentangle the impact of a leader from that of the message that she conveys and that of the locality in which speaks. In the event we can discern the impact of a leader on opinion, we should ask whether the same effect would exist if we were to switch the leader's message and/or the location of her address. Moreover in any environment where we can systematically measure the impact of political leaders the questions we should then ask are: How would the opinions of the population have changed if different leaders had been dispatched? How would the opinions of listeners have changed if different arguments had been employed? And would the leader (or message) have had a similar effect in a different location?

Answers to these questions depend on being able to fix the leader (message), whilst allowing the message (leader) to vary. A political campaign presents a natural environment in which to explore this issue. A campaign involves the attempt by some individuals, who might think of as "opinion leaders", to persuade others of the merits of their policy position: as part of the campaign process these opinion leaders adopt messages deemed effective in persuading the public; and a campaign will typically cover a wide geographic area comprised of distinct local entities. Moreover, since most campaigns use a variety of different messages, conveyed to the public and endorsed by different individuals one can, in principle, measure the impact of a campaign on public opinion and break down this overall effect into its constituent parts: the extent of the effect that is due to the arguments the campaign deploys; the talents of those getting the message across; and/or the popularity of leading public figures associated with the campaign.

However, even in such a situation where each of the factors of interest vary, observational studies are likely to fall short in identifying these effects. This is because there are a number of campaigning strategies that we might commonly expect to find that would

confound any attempt to isolate the causal effects of leadership. Consider as examples: that some leaders may be associated with certain types of message; that both leader and message may be chosen by the campaign to address certain types of audience; and that, perhaps most importantly, leaders may adapt their messages to suit the audience. These problems are related to the issue of correlation, selection-bias and endogeneity, respectively. Thus the presence of any of these in isolation, or in combination with each other, will be enough to confound any analysis of leadership effects:

The method of randomized evaluation, however, provides some leeway in identifying the parameters of interest. In this paper we describe an empirical strategy and report the results of its implementation in the field, analyzing the impact of leaders and messages on a referendum on electoral reform in British Columbia. Working directly with the British Columbians for Single Transferable Vote political campaign we randomly assign both leaders to different localities, and randomly assign messages within those localities. This creates a hierarchical structure that allows us to directly isolate the effect of each on political outcomes of interest.

The campaign devised two political messages that they believed would enhance the prospects of a yes vote on the proposal to change the electoral system of British Columbia from the current first-past-the post system to STV, in the upcoming referendum on electoral reform to be held on May 12th 2009. The first of these arguments highlighted the fairness aspects of change in the electoral system, whilst the second emphasized the increase in voter choice and accountability corresponding with a switch to STV. The campaign trained canvassers at specially arranged training sessions. These canvassers were then randomly assigned to different local voting areas. Within each area a (random) subset of households canvassers would deliver a ”placebo” message in which voters were told that a referendum will take place on May 12th, whereas a random subset of households received, in addition, one from a list of ”treatments” in the form of an official communication from the campaign. These treatments, described more fully in section below, always involved of the two messages communicated orally by the canvassers and backed up by a campaign flyer outlining the key points of the message. Some of these messages were endorsed by leading public figures associated with the campaign. Researchers then tracked the houses visited by the campaign and implemented a simple survey in which the main question asked respondents about their intended vote in the referendum.

This approach allows us to get at leadership effects in two ways: we can separate the effect of the arguments used by the campaign from the impact of the person delivering the message; by comparing across localities we can ascertain how much of the differences in attitudes toward STV can be accounted for by the fact that the campaign randomly deployed different canvassers in different areas. Moreover, at the individual level, we can separate the effect of the message from the person endorsing it. The approach also also allowed us to consider interaction effects: for example, would a message be more effective when delivered by a leading public figure associated with the campaign.

Our results are surprising. We had designed our study in such a way so as to isolate leadership effects from other confounding factors. Contrary to what we expected to find, however, we were unable to reject the hypothesis that leaders have no impact on public opinion. Although we were clearly able to discern an effect of the campaign- households visited by the yes campaign are far more likely to vote yes- little to none of this effect appears to associated with variation in who visited these households, or with endorsement effects. That is we found no evidence that local variation is due to different canvassers being deployed, and no evidence that variation at the individual level can be explained by endorsements by leading public figures. By contrast, all of the campaign effect- an average 12 percentage point increase in the likelihood of a yes vote- is due to the messages the campaign deploys. Our findings suggest that the messages stand on their own merits, with the accountability message having a larger effect on differences in public opinion.

In the remainder of the paper we set out our research question and empirical strategy more fully, explain the details of our field experiment and implementation, and present our results.

2. Leaders or Messages

There are many situations where leaders may affect public opinion, and different ways in which the input of a leader can be related to political outcomes. Our interest is in distinguishing between the effect of innate leadership characteristics and the arguments that leaders deploy. There are good theoretical reasons to believe that each of these, in isolation, will impact upon political behavior.

There are different individual characteristics which may make an individual persuasive. Amongst attributes that might enhance a leaders persuasive ability are her seniority and/or experience; a leader may appear convincing because she seems to have the interests of

her listener at heart; and or followers may place trust in leaders because they believe that leaders have better access to information and more knowledge on the issues. The focal theory of leadership (Dewan and Myatt, 2008, 2007; Myerson, 2004; Calvert, 1995). suggests followers will be attracted to leaders because they believe others are attracted to them also this may be related to the observables described above, or to other features not always discernible by the analyst. Each of these features may contribute to a "leadership" effect, whilst nevertheless being separate from the arguments a leader deploys. Thus, in such cases, the effect of the leadership attribute, whether her trustworthiness, experience, or knowledge, is in principle distinguishable from the message that a leader conveys.

However an individual may be persuasive not due to her particular leadership abilities but due to the message that she conveys. A conflation of a leadership and message effect occurs, for example, when voters are more open to some types of arguments than others, and thus their willingness to follow a leader reflects the fact that she made an argument they were persuaded by. This suggests that arguments may stand on their own merits irrespective of the person delivering or associated with that view, and that what we might commonly think of as a leadership effect is due to the arguments she a leader has at her disposal. Indeed recent research finds that some arguments can be more effectively deployed than others, and that this can depend on the initial information followers have available to them. A leader might be more persuasive when communicating a message that conforms to latent dispositions held by audience members (Hafer and Landa, 2007). Moreover, this feature of the strategic environment, might lead to some arguments not being made at all (Hafer and Landa, 2008).

There are good reasons then to believe that both leaders and the messages matter. Moreover, in principle, these effects are distinguishable from each other: certainly one can conceptually separate the impact of a leader that is due to some innate characteristic, whether observable or not, from the impact of the arguments she makes. And yet we know of no study to date that can distinguish individual effects from argument effects in order to assess the role that leaders play in shaping public opinion. However our study does build on two others that provide partial answers to our questions of interest.

In one study, designed simply to identify leader effects, (Humphreys, Masters, and Sandbu, 2006) randomly assigned leaders to discussion groups during a national level public deliberation in the island state of Sao Tome e Principe. The random assignment was undertaken to estimate the extent to which leaders influenced discussion outcomes. These

effects were found to be substantial: leader characteristics often predicted up to 50 percent of the variation in discussion outcomes.

A second study examines the differential effects of arguments (Wantchekon, 2003). In this study presidential candidates in the 2001 national elections in Benin randomly employed different types of political messages in different communities: in some they deployed messages emphasizing the local benefits of politicians strategies; in others they used messages emphasizing national benefits. In a follow up study Wantchekon (2008) studied the effect of the latter message when devised by a group of policy "experts", comparing the voting intentions of those who received this treatment with that of those who received a standard party message.

3. Empirical Strategy

Our aim is to distinguish the effect on public opinion of innate characteristics or attributes of leaders, from other confounding factors such as the impact of the messages that they convey. Both of these elements vary systematically in a political campaign and so, in principle, when analyzing such processes we can ask how would the opinions of listeners have changed if different arguments had been employed? and how would the opinions of the population have changed if different leaders had been dispatched. responses might tell us that argument A is more effective than argument B, these results may be reversed if we switched the location of the meeting, or switched the person delivering each argument.

In practice, however, it is hard to discern whether people are convinced by the argument being deployed, or by the person who is delivering the message. An argument or leader that is effective with one audience may not work as well with another: different groups of individuals may be more or less responsive to different leaders and so we may falsely attribute effectiveness to a leader whose appeal is limited to a small sample of the population. In addition there are three related problems that we are likely to encounter in our attempt to distinguish between the effects of leaders and the arguments they make.

Firstly, certain leaders may be prone to making certain types of arguments. Even when accounting for ideological differences systematic differences between a leaders type and the message they deliver are likely to arise. For example, experienced politicians may be more cautious in the types of policies they promote, and younger politicians may promote policies that reflect their longer time horizons. Thus there may be a correlation between a

leaders characteristics and the arguments she uses; this correlation hinders any empirical assessment of the causal effect of each on political outcomes of interest.

Secondly, since a leader 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. For example, suppose that a leader uses argument A but senses that her audience might be more swayed if she deployed argument B instead; she will then be tempted to switch the emphasis of her message toward B. In this sense the audience is affected by the arguments of a leader *and* vice-versa. This strategic component of political communication introduces a classic problem of endogeneity that confounds analysis of leadership effects.

Thirdly, both arguments and leaders are typically selected because they are deemed likely to be effective in a given setting. Some leaders may be more effective when campaigning in particular areas, and this may be due to their having local roots or connections. Similarly certain types of argument may work better when deployed amongst some communities rather than others. A political campaign is likely to match leaders and arguments to local areas to maximize the impact of the campaign. This strategic component of campaigning introduces a selection bias that can confound attempts to isolate the causal effect of leaders arguments.

Each of these problems is likely to arise as part of the natural process of political campaigning. Thus it may appear impossible to distinguish leader effects in such an environment. However analyzing political behavior in a controlled environment provides some leeway to addressing our key questions. Randomizing on the variables of interest gets around the problem of selection bias and correlation described above. If, for example, a campaign were to randomly assign opinion leaders to different localities, and randomly assign different messages to different addresses within localities, then the estimates of these effects would be unbiased. One way to implement this is via a door-to-door campaign in which the opinion leaders are individual canvassers randomly assigned to localities but who deliver pre-assigned messages at randomly assigned households. That is the design we implement in our study, but the general model we describe below could also be implemented in other campaign scenarios.

3.1. **A Hierarchical Model.** During the course of a campaign different individuals, or groups of individuals, are exposed to different leaders who make the case for the campaign. Some leaders may be more persuasive in making campaign arguments and this may be due to their innate traits and communicative abilities. Moreover a leader may be

more persuasive when communicating with different groups. For example, a leadership effect may be specific to, or magnified, in certain localities.

To identify each of these these effects we analyze a hierarchical model in which individual responses reflect both idiosyncratic and group level features. More specifically, we think of individuals as clustered in local areas, where these areas may form part of a leader's communication network. Equivalently, we think of a leaders network of consisting of a number of different local neighborhoods within which individuals are clustered. The individuals response then includes a component T_k that is common to others who have come into contact with the k_{th} leader, and we think of these as the group level treatments.

At the individual level we think of a response as a function of the type of political communication that an individual is exposed to within her locality. We think of a vector T_{ai} consisting of treatments that an individual i may or may not have received. For example, for a given set of arguments that a political campaign might use then each element of T_{ai} takes on a value zero or one indicating whether the individual i was exposed to the a_{th} argument. Alternatively we might consider that a treatment consists of an argument in combination with some other form of political communication. Generally, we think of the arguments as being the individual level treatments.

In addition to both individual and group specific treatments administered by the campaign, the outcome can reflect individual specific features ε_{ij} as well as other group-specific features, τ_j. Bringing these elements together we have:

$$y_{ijk} = T_{ai}^{'}\beta_a + T_{k}^{'}\beta_k + \tau_j + \varepsilon_{ij}$$

where y_{ijk} is the political outcome of interest, k indexes leaders, j communities and I individuals.

This design allows us to draw comparisons between of a treatment group (those for whom $T_{a=1}$) to a control group ($T_{a=0}$), and allows us to provide answers to a key question: conditional on listening to leader k, how would the opinions of listeners have changed if argument A rather than argument B. For each of the comparisons, that is for each of the messages deployed, one group of voters acts as a treatment group whilst another acts as a control. Comparing outcomes across the groups then allows one to estimate the different effects of different arguments.

The empirical design we implement is a special case of this set-up in which each individual is exposed to at most one argument and one leader. Conceptually, within our

framework, it is hard to imagine a situation in which a respondent receives a message without making contact with a leader. This implies that we have no natural control group to compare the impact of the k_{th} leader. An alternative approach, similar to that used by Humphreys, Masters, and Sandbu (2006)is to exploit the fact that, as all communities receive some leader, the k parameters are determined relative to each other. If leaders have an effect then we should expect that the post-treatment differences in opinion amongst the population responds in some way to the different leaders they were exposed to. We can then deploy a standard F test to test the hypothesis that k is constant across leaders.

An alternative way of getting a handle on leadership effects directly, is by allowing some of the individual arguments to be endorsed by publicly known figures associated with the campaign. This allows us to estimate leadership effects as part of the individual level treatments. We can compare the responses of those exposed to to the $j_t h$ endorser with those who were not; or we can compare the response of those exposed to any endorser, to that of those individuals whose message was not endorsed by any individual. We discuss these elements more fully below.

Problems might arise in the estimation if the k leader effects are correlated with the local fixed effects, τ_j, as then the model would be unidentified. Consequently the model is identified when leaders are randomly assigned to local areas. This should in principle mean that our estimates of the $k-1$ leadership intercepts are uncorrelated with locality or indeed anything else we might include on the right-hand-side of the equation being estimated. However when the number of observations is not large, or when there is an imbalance on the number of households exposed to different leaders then we may need to include a wider array of fixed effects. For example, we can go one step further in our estimation by absorbing the local fixed effects, and estimating the model with a different fixed component for leader k in area j; this allows for the fact that the effect of some leaders may vary across different localities and constituencies.

To get unbiased estimates of β_a we need to ensure that, conditional on the local and leader fixed effects, the group receiving a political message is in all ways, identical to the control group except for the treatment employed. This can be achieved by random assignment within each of the j localities. For example if the people that receive the treatment (for example, to hear argument A delivered by leaders k) are a subset of those eligible to receive the same treatment, and those that receive the treatment are chosen randomly

from the group of eligible people, then, in expectation, the control and treatment group are identical on all aspects other than the treatment.

The hierarchical model described above is a suitable way of getting estimates of leadership effects with random assignment of leaders to localities, and random assignment of messages within localities. Conditional on local and leader fixed effects, in expectation the treatment and control group are identical in all regards other than the message received. However as usual this model makes some assumptions of the data that may in fact not be appropriate. In particular it assumes a particular form of leadership effect that is constant across leaders and that does not vary with the treatment. Moreover, is assumes that conditional on leaders, messages, and local area fixed effects, the remaining errors are identical and independently distributed. Moreover, even with random assignment some problems may emerge. When leaders visit multiple local districts then individual leader effects are,in fact, aggregates of each of the different local effects. And as analysts we can not be certain that these aggregates do not have a separate effect- that is the aggregate effect may work via a channel that is distinct to the leadership channel, even though this is unlikely. When we observe multiple leaders visit each district then these effects can be disentangled, but this may not always be possible.

Luckily, alternative ways of estimating the parameters from the data also exist. One such alternative is to implement a matching procedure. This works by using the data to construct a control group that is similar to the treatment group in all ways other than the treatment being administered. Both the hierarchical model described here, and the matching model, can be estimated using straightforward techniques; for example they can be estimated using the xtreg and nnmatch commands for the hierarchical and matching models in stata.

We turn now to a full discussion of an implementation of this research design using data from a unique field experiment in British Columbia.

4. The Political Campaign

4.1. **British Columbians for STV.** To implement our research design we worked directly with the British Columbians for Single Transferable Vote in their build up to the referendum on electoral reform on May 12th 2009. We initially made contact with the Citizens Assembly in British Columbia which was created by the Government of British Columbia as an independent, non-partisan assembly of citizens to examine the provinces electoral

system and to make recommendations for reform. The Assembly of 160 members, spent an 11 month period of consultations, deliberations, and hearings before advocating the Single Transferable Vote in their final report to the people of British Columbia in December 2004. Members of the assembly campaigned across the province for the proposal, with over 800 meetings being lead by assembly campaigners. Despite winning a majority on 72 of the 76 districts, the reform narrowly failed to meet the 60 percent Province wide threshold in the referendum of May 17th, 2005. A second referendum on the same proposal will be held in May 2009 and our work involved analyzing the effects of the campaign in the run in 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 them with valuable information about the effects of different elements of their campaign. As part of the implementation of this design, the campaign identified a set of mid-level activists who would serve as opinion leaders, a set of messages that the campaign believed would support their position on STV, and a set of sites where campaigning would take place. It was important that these elements were chosen by the campaign with our involvement limited to guidance in the randomization protocols we describe below.

To identify campaign effects with confidence we used the first best method of randomized assignment to ensure that observed differences are truly due to treatment effects and not due to selection or other confounding effects. In a (random) subset of households canvassers delivered a "placebo" message in which voters were simply told that a referendum will take place on May 12^{th}. The placebo contained no information that could be construed as favourable to the Yes side. In addition to these placebos a random subset of households received one from a set of "treatments". These treatments, described more fully below, came in the form of a campaign document and a presentation by the canvasser. 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 favor 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 summarizing the main points of the argument. Some of these flyers were also endorsed by leading public STV campaigners. Some treatments involved a simple endorsement with no supporting argument. We describe these treatments in more detail below.

4.2. **The Treatments.** The campaign devised a set of treatments that would be delivered by their specially trained canvassers at addresses in the Province. These treatments consisted of two messages that the campaign believed to be favorable to its position. In addition, some treatments involved both a message and an endorsement from a leading public figure associated with the campaign.

4.2.1. *The Messages.* The two messages constructed by the campaign to made the case that STV is the right electoral system for Canada, and these are shown in Figure 1. One set of arguments focused on the general claim made in favor of proportional systems that such systems are fairer. This message emphasized that STV translates votes into seat-shares more proportionally and reduces the phenomenon on wasted votes. It emphasized that plurality rule elections often lead to governments being elected on less than 50 percent of the vote, and that in the past BC governments have consisted of 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 amongst different party representatives, and a comparison of performance between local districts. The campaign argued that STV would lead to better local representation, as with more local MP's citizens would be able to choose amongst representative to go to with their concerns, including the choice to visit all local representatives.

4.2.2. *The Endorsers.* Some of these messages were also endorsed by leading public figures who supported the campaign. BC-STV settled on 4 main endorsers for the purpose of the study: David Suzuki a well known media figure and environmentalist seen in figure 2; Preston Manning politician and former leader of the Reform Party; Andrew Coyne, the editor of Macleans magazine and well known journalist; Lorne Nystrom, former National Democratic Party Member of Parliament. The list gives some balance on politician and non-politician, age and gender, ultimately these were the endorsers the campaign believed would have the biggest campaign impact.

FAIRNESS

Fairer Results

No more distorted outcomes. It will take roughly the same number of votes to elect an MLA, no matter what riding, no matter what party. Across the province, a party's share of the votes will be very similar to their share of the seats. Your vote will count.

Under the present system the Legislature we get is not the one we voted for. The current system — single-member plurality, or "first-past-the-post" — distorts people's votes, producing unrepresentative legislatures.

- We usually have a majority government that has been elected by a minority of voters - a **'false majority'.**
- Sometimes the winning party does not even have the most votes – a **'wrong winner',** as happened in 1996.
- A majority of our votes are usually **'wasted'** – in 2005, 64% of voters had no effect on the outcome. Both major parties, and all smaller parties and independent candidates, have suffered from our antiquated voting system.

Many countries around the world, including Germany, Ireland, Sweden, Australia and New Zealand, use modern systems that avoid these problems.

BC-STV will give British Columbians the governments they vote for — and that's the way it should be.

ACCOUNTABILITY

Greater Voter Choice

Under our present voting system voters only have one candidate per party to choose from and are only represented by one candidate.

But under STV, parties nominate multiple candidates in each district and the voters have the final say on which ones are elected.

This means:

- **No more 'safe seats'** - the voter's choice is all that counts.
- Since voters are represented by several candidates from the same, or different, parties at election time, they can **compare how all of their district MLAs measure up** between elections.
- More voter choice means **MLAs have a stronger incentive** to represent their community's best interests in the legislature.
- If you need help after Election Day **you can approach** the MLA you voted for, the one who lives closest to you, **or all of your MLAs**.

BC-STV puts the power in the hands of the voters — and that's the way it should be.

Accountable Local MLAs

FIGURE 1. Messages

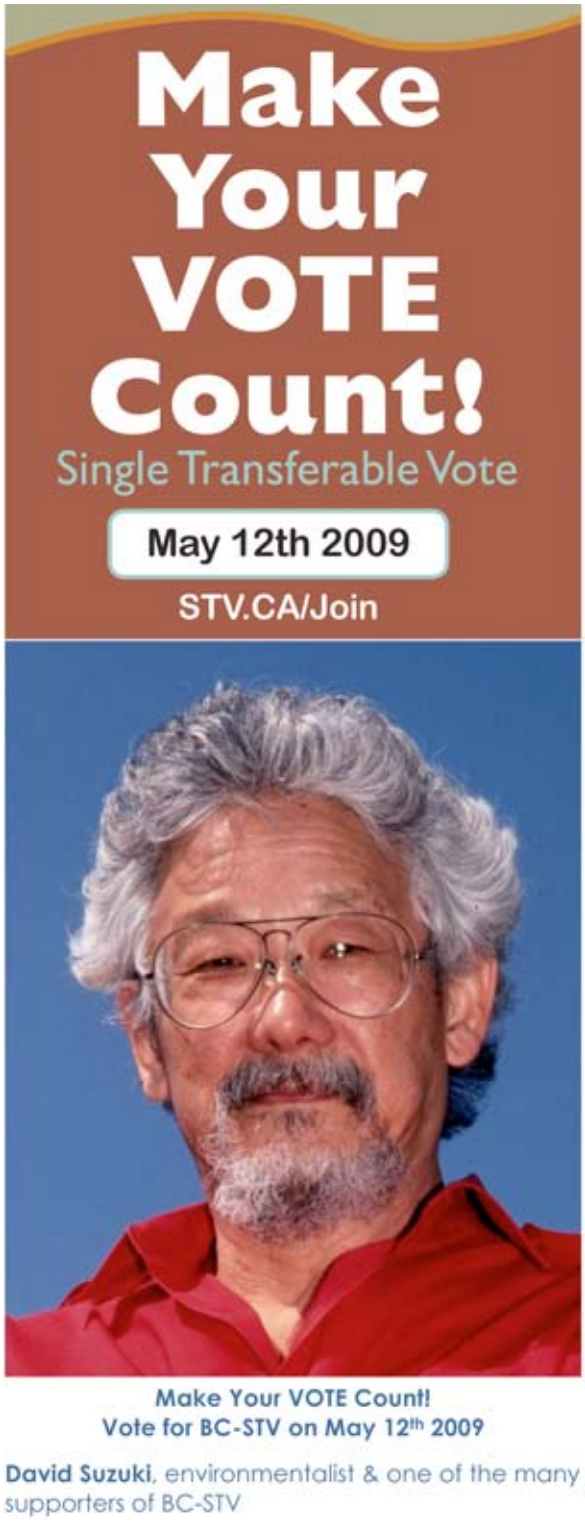

FIGURE 2. Messages with Endorsement

4.3. **Voting Areas.** Canadian provinces are broken down into ridings of which there are 308 in all of Canada, which return representatives to the Federal Parliament. These are areas are broken down further into political areas called voting areas. The design involved 215 voting areas in the Greater Vancouver area that are shown in Figure 4. Canvassers were randomly assigned to voting areas and treatments administered within voting areas.

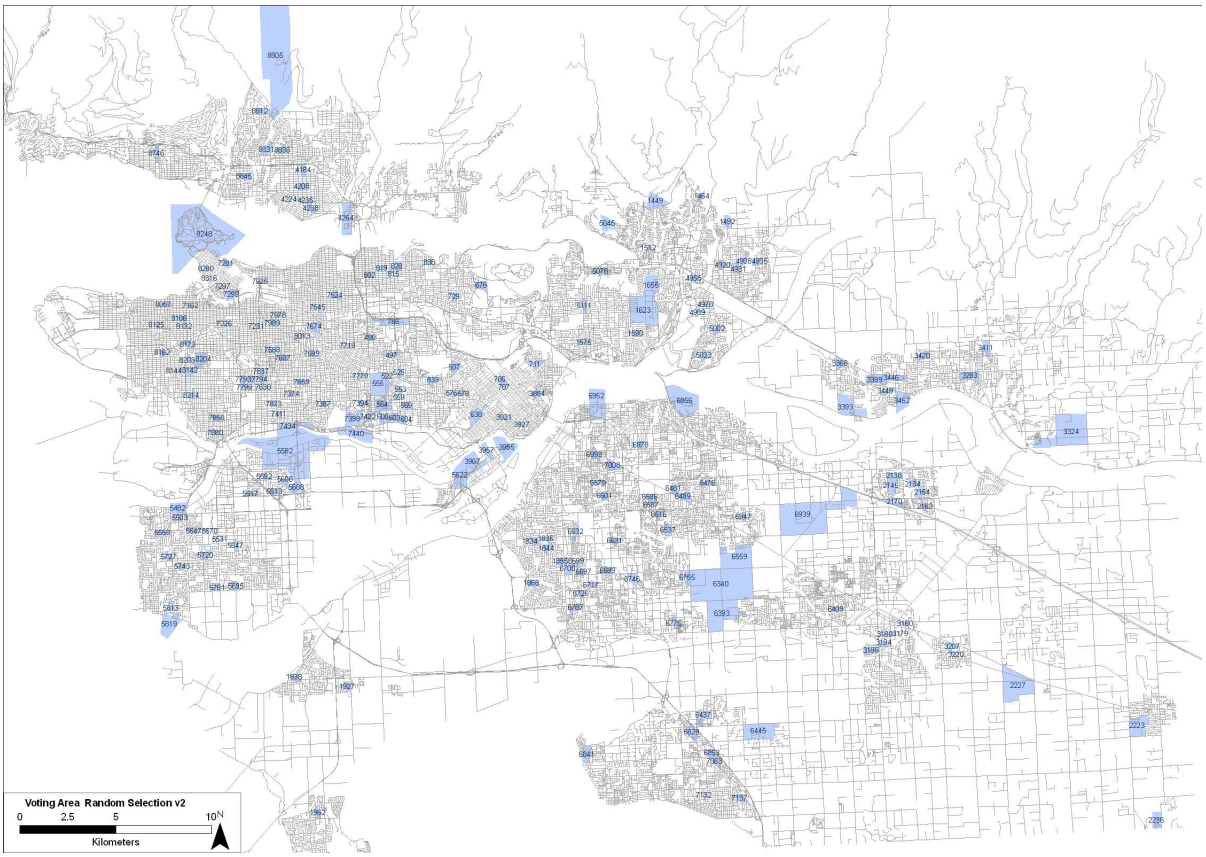

FIGURE 3. Location of Sampled Canvassing Sites

4.4. **Randomization Procedure.** A group of canvassers were trained to take part in the evaluation programme. Training took place over several days with the first canvassers trained by the research team, and later canvassers trained by members of the campaign team. Canvassers were told to follow their usual campaigning style, but were given instructions as to which areas to visit, and which households to visit in each area, with the former being randomly assigned on a daily basis.

In each voting area canvassers were provided with a map of the area on which a randomly generated starting point was indicated, along with a route that they would follow.

Canvassers 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 10th residential structure, and if not they would go to the next residential structure.

In addition canvassers were given instructions as to which order they should deliver their messages. A typical daily canvasser sheet is shown in Figure 4.

BC STV MESSAGE TESTING CAMPAIGNER FORM **Campaigner NAME**_____ TOTALS

Date: ___ March 2009
Date: ___ April 2009

	Number of Unsuccessful calls (no one at home):	_____
	Number of Unsuccessful calls (no common language):	_____
	Number of Unsuccessful calls (no eligible voters):	_____
	Number of Unsuccessful calls (not interested):	_____

You	Side of Street	CASE ID	Riding	Voting Area	Apt #	House #	Street Name	Type	Message	Endorser	Start Time	Visit Length	Stayed on message?	Team contact?	Apprx age	Gender	Knowledge of STV/referend
1	L	**10101**	NEW	128				○Apt/Mult ○Hse/Single	ACCOUNT	SUZ	__:__ HH MM	___ MINS	○YES ○NO	○YES ○NO	__YRS	○M ○F	____ [0-10]
1	L	**10102**	NEW	128				○Apt/Mult ○Hse/Single	Endorser Only	MNG	__:__ HH MM	___ MINS	○YES ○NO	○YES ○NO	__YRS	○M ○F	____ [0-10]
1	L	**10103**	NEW	128				○Apt/Mult ○Hse/Single	FAIRNESS	MNG	__:__ HH MM	___ MINS	○YES ○NO	○YES ○NO	__YRS	○M ○F	____ [0-10]
1	L	**10104**	NEW	128				○Apt/Mult ○Hse/Single	Placebo Only	.	__:__ HH MM	___ MINS	○YES ○NO	○YES ○NO	__YRS	○M ○F	____ [0-10]
1	L	**10105**	NEW	128				○Apt/Mult ○Hse/Single	FAIRNESS	SUZ	__:__ HH MM	___ MINS	○YES ○NO	○YES ○NO	__YRS	○M ○F	____ [0-10]
1	L	**10106**	NEW	128				○Apt/Mult ○Hse/Single	ACCOUNT	.	__:__ HH MM	___ MINS	○YES ○NO	○YES ○NO	__YRS	○M ○F	____ [0-10]
1	L	**10107**	NEW	126				○Apt/Mult ○Hse/Single	FAIRNESS	SUZ	__:__ HH MM	___ MINS	○YES ○NO	○YES ○NO	__YRS	○M ○F	____ [0-10]
1	L	**10108**	NEW	126				○Apt/Mult ○Hse/Single	ACCOUNT	MNG	__:__ HH MM	___ MINS	○YES ○NO	○YES ○NO	__YRS	○M ○F	____ [0-10]
1	L	**10109**	NEW	126				○Apt/Mult ○Hse/Single	Placebo Only	.	__:__ HH MM	___ MINS	○YES ○NO	○YES ○NO	__YRS	○M ○F	____ [0-10]
1	L	**10110**	NEW	126				○Apt/Mult ○Hse/Single	Placebo Only	.	__:__ HH MM	___ MINS	○YES ○NO	○YES ○NO	__YRS	○M ○F	____ [0-10]
1	L	**10111**	NEW	126				○Apt/Mult ○Hse/Single	FAIRNESS	.	__:__ HH MM	___ MINS	○YES ○NO	○YES ○NO	__YRS	○M ○F	____ [0-10]
1	L	**10112**	NEW	126				○Apt/Mult ○Hse/Single	Endorser Only	SUZ	__:__ HH MM	___ MINS	○YES ○NO	○YES ○NO	__YRS	○M ○F	____ [0-10]

NOTES: 1/216

FIGURE 4. Example of Daily Canvassing Sheet

4.5. **Matching of Treatments to Voters.** In our design, the types of message delivered to households varied in the content of the message delivered and the identity of the endorser associated with the message according to the distribution given in Table 1. The planned overall matching of treatments to individuals aimed to give overall balance over each of the treatments

In practice, data collection fell considerably short of what had been planned, and so the sample of respondents is smaller than anticipated. However sufficient data was still collected to allow core questions to be answered.

The data shortfall arose due to a combination of factors. First although 2592 households were targeted for canvassing, engaging sufficient numbers of canvassers for the duration required proved very difficult. In part this was due to the fact that the randomization strategy required greater time spent traveling than normally needed for canvassing. The effect was that 1044 rather than 2592 houses were canvassed.

Enumeration teams were able to locate and visit 968 (93%) of these households. Often enumerators would have to visit a household several times before making contact. Enumeration was successful, in that the enumerators were able to get a full response to the survey, in 520 cases. Of the remainder, in 255 cases enumerators were unable to find anybody at home after multiple visits (at least 3 per house, often 6); in 192 cases respondents refused to respond to the survey (although in 32 of these cases the respondents did agree to answer a single question on voting intentions).

TABLE 1. Research Design Distribution of Messages and Endorsers

		Endorser					
		None	**Coyne**	**Manning**	**Nystrom**	**Suzuki**	**Total**
Message	**Account.**	216	126	126	126	126	720
		(8%)	(5%)	(5%)	(5%)	(5%)	(28%)
	Fairness	216	126	126	126	126	720
		(8%)	(5%)	(5%)	(5%)	(5%)	(28%)
	Endorser	648	126	126	126	126	1152
	only	(25%)	(5%)	(5%)	(5%)	(5%)	(44%)
	Total	1080	378	378	378	378	2592
		(42%)	(15%)	(15%)	(15%)	(15%)	(100%)

4.6. **Voter profiles.** Our aim was not to reach a population representative of all voters in the BC area but to construct a sample of households who were *likely to be reached by the campaign* during door-to door canvassing. Thus the aim of the study was to discern leadership and message effects amongst a representative sample of those contacted by the campaign. For this reason the campaign canvassers were encouraged to use their normal canvassing techniques with the exception of strict implementation of the randomization

protocols described above. Canvassers were asked to provide basic descriptive information on the respondents. This basic information allowed us to get an overview of the sample of respondents, and of the type of people the campaign is likely to make contact with in door-to-door canvassing.

Our sample is composed of 48% women and 52% men. The median age is 48 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 last referendum. Of those reporting, 65% report not having voted in the last referendum; and of those that voted, 58.6% reported having voted Yes and 41.4% reported having voted no.

Enumerators followed up by making contact at the door and administering a short survey. The survey revealed a number of interesting findings related to our sample and we list these below. Perhaps the key finding of interest to the campaign is listed as Finding 5 below. Amongst those indicating a preference between STV and the status quo, support is high for STV. These responses are provided in more detail in Table 3.

5. Results

5.1. **Canvasser Effects.** The aim of our study is to separate the effects of leadership and message in a campaigning environment. An interesting issue arises due to the way in which messages are delivered in such a setting, namely it is hard to conceptualize a message without a leader to deliver it. This presents a problem in identifying a leadership

Table 2. Sample Distribution of Messages and Endorsers in Surveyed Areas

		Endorser					
		None	**Coyne**	**Manning**	**Nystrom**	**Suzuki**	**Total**
Message	**Account.**	50 (10%)	20 (4%)	18 (3%)	24 (5%)	26 (5%)	138 (27%)
	Fairness	48 (9%)	26 (5%)	27 (5%)	26 (5%)	28 (5%)	155 (30%)
	Endorser only	126 (24%)	26 (5%)	21 (4%)	22 (4%)	32 (6%)	227 (44%)
	Total	224 (43%)	72 (14%)	66 (13%)	72 (14%)	86 (17%)	520 (100%)

TABLE 3. Reasons to support or Oppose STV

How strongly do you agree or disagree with the following statements:	Agree strongly	Agree	Disagree	Disagree strongly	All
A Our present voting system works well and does not need to be changed	8	41	38	13	100%
B STV would lead to a more diverse representation in the legislative assembly	23	63	11	2	100%
C Under STV the results of elections would better reflect the way people vote	22	61	13	4	100%
D STV would lead to weak political parties	12	29	50	9	100%
E STV would make politicians more responsive to the needs of voters	14	61	18	7	100%
F STV is too complicated a system	10	28	50	13	100%

effect that is due to different canvassers being assigned to different areas. A useful approach is to analyze the relative differences across canvassers. A priori one should expect that, if leadership effects exist, then variance in the opinions of respondents should reflect differential exposure to leaders. Put another way, a necessary condition for an individual leadership effect is that some leaders are more persuasive than others.

This approach was adopted by Humphreys, Masters, and Sandbu (2006) in their analysis of leadership effects in deliberative exercises. In their study individuals were assigned to record deliberations on the use of natural resources in Sao Tome e Principe. As no such meetings took place without an official recorder of the debate, their study, as ours, lacked a natural control group consisting of leaderless citizens. To detect leadership effects they estimated a series of models on deliberative outcomes controlling for leader specific fixed

effects, using an F test to test the hypothesis that that the joint difference between these individual leader fixed effects was zero. Indeed they were unable to reject this hypothesis, concluding that a large part of the variance in post-treatment outcomes reflected the influence of leaders.

Here we perform a similar analysis for different regression specifications reported in Table 4 below. The analysis is based on different specifications of the linear probability model with tow different dichotomous endogenous variables: the first (Yes) is the likelihood of a yes vote with no voters and likely abstainers is the reference group; the second (Vote) is the likelihood of a yes vote with only no voters as the reference group. The final model estimated is on a continuous variable: the number of yes voters by voting area (weighted by the population of the voting area).

The first column reports the regression being estimated with the first term being the endogenous variable, and the remaining terms the regressors that include various combinations of the relevant treatments: campaign, message, endorser, as well as interactions between message and endorser. For example the inclusion of campaign estimates the overall campaign effect relative to the group receiving only the placebo, the inclusion of the message estimates estimates the effect of the message relative to receiving either a placebo or an endorsement only. Each of the regressions includes an estimated intercept for $k-1$ canvassers.

Since all examined voting areas receive some leader the k parameters are determined relative to each other. The second column of the table then presents the test statistic for each of the model specifications for the null hypothesis that the leadership intercepts are constant across the class of canvassers. The final column gives a *p* value for the likelihood of observing a test score of that magnitude given the null hypothesis.

The use of an estimated leadership intercept is, on the one hand, a very restrictive way of estimating the leadership effect, and is based on the strong assumption that the effect is constant across the class of canvassers. But it also has advantages. It can account for the constant effect of differences between canvassers not immediately discernible to the analyst, though perhaps constituting part of a leadership effect via its impact on public opinion.

For the first three estimations shown, all using the comparison between yes and no voters, would lead us to reject the null hypothesis of no significant differences across canvassers at the 10% significance level. The next two reported specifications include abstainers

in the comparison group and each has a test statistic of similar magnitude, though not statistically significant at conventional levels.

The next specification looks at canvasser effects only amongst the sample of respondents who had explicit contact with the campaign ie. they did not receive a placebo. It seems natural that, if leadership effects are present, then those exposed to the campaign most susceptible to such effects; but here we find no evidence here of the said effect.

However, and despite the fact that canvassers were randomly assigned to voting areas on a daily basis, each of these estimates may conflate leadership with geographic effects. The simplest way to isolate these effects is then to consider how much of the variance on the intended vote between voting areas is due to difference between canvassers. This is the estimation in the last row of the table: in deploying a standard F test to test the hypothesis that k i constant across leaders we fail to reject the null, based on a test statistic of 1.39.

Overall the evidence in this section presents at the very best, weak and circumstantial evidence of leadership effects. However based on the cleanest analysis separating leadership from local specific effects, we would reject the hypothesis that differences across voting areas are due to exposure to different leaders.

TABLE 4. Test Hypothesis of Constant Leader Effect

Model	Test statistic	Prob>F
Vote, Campaign, Leader Fixed Effects	F(16,279)=1.521	0.091
Vote, Message, Campaign, Leader Fixed Effects	F(16,278)=1.517	0.093
Vote, Message, Endorser, Interactions, Leader Fixed Effects	F(16, 277) = 1.510	0.095
Yes, Campaign, Leadership Fixed Effects	F(16,455)=1.348	0.164
Yes, Message, Endorser, Leader Fixed Effects	F(16,454)=1.136	0.155
Yes, Message, Endorser, Leader Fixed Effects	F(16,454)=0.136	0.155
Yes, [Message, Endorser,if Campaign=1], Leader Fixed Effects	F(73,280)=1.09	0.306
Average Yes by Voting Area, Campaign, Leader Fixed Effects	F(14,77)=1.390	0.179

5.2. **Message Effects.** In this section we report the results of estimating the effects of the individual level treatments in a linear probability model that conditions on canvasser-voting area fixed effects. Table 5 reports results when comparing Yes voters with no voters and abstainers combined (YES). Whereas Table 6 presents the same results when comparing yes voters with no voters only (VOTE). In each of these regressions the treatment

group consists of all those receiving a message - either emphasizing the accountability or the fairness aspects- and the control group consists of those receiving either the informative placebo or the endorsement only.

These models suggest that, conditioning on all geographic data, canvassing by the campaign has a significant and substantial overall effect on the likelihood of a yes in the referendum on the introduction of STV in British Columbia. The estimated magnitude of this effect is in the region of 12-13 percentage points, depending on which model specification we adopt: thus the effect is of a similar size whether yes voters are compared with no voters or with no voters and abstainers combined. Model 2 in each of these tables reports the effect of receiving a message from the campaign with the control group being those that received either the placebo or the endorsement only. The estimated effects vary, a marginal effect of 8.6 percentage points when analyzing YES, and a marginal effect of 11.3 when analyzing VOTE. When analyzing the former columns 3 and 4 of Table 5 suggest that the accountability message absorbs most of the message effect. Whereas analysis of the latter suggests the effects of the messages are evenly balanced.

TABLE 5. Impact of Campaign and Message on YES vote with Fixed Effects for Canvasser by Voting Area

	(1)	(2)	(3)	(4)
Campaign	0.118**			
	(0.0546)			
Message		0.0861*		
		(0.0471)		
Fairness			0.0320	
			(0.0561)	
Accountable				0.125**
				(0.0598)
Constant	0.325***	0.365***	0.371***	0.370***
	(0.0461)	(0.0339)	(0.0333)	(0.0352)
Observations	478	478	357	297
Number of voting areas	77	77	76	76
R^2	0.011	0.008	0.001	0.015

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

5.3. **Endorser Effects.** We next consider the impact of public endorsements on the likelihood of a yes vote. In this analysis the treatment group varies to includes all those

TABLE 6. Impact of Campaign and Message on VOTE with Fixed Effects for Canvasser by Voting Area

	(1)	(2)	(3)	(4)
Campaign	0.118**			
	(0.0566)			
Message		0.113**		
		(0.0474)		
Fairness			0.121**	
			(0.0585)	
Accountable				0.113*
				(0.0598)
Constant	0.561***	0.587***	0.586***	0.601***
	(0.0482)	(0.0343)	(0.0340)	(0.0343)
Observations	301	301	301	301
Number of voting areas	71	71	65	63
R^2	0.003	0.009	0.005	0.013

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

whose message contained an endorsement from a leading public figure, or who received the endorsement only. Table 7 estimates these effects on the likelihood of a yes vote using a linear probability model and conditioning on all geographic data. The first model estimates the effect of receiving the endorser treatment where the control group in this analysis includes those receiving the placebo only, and those who received a message with no endorsement. The second model estimates the effect of receiving the endorser treatment conditioning on receiving a message, and so the control group in this analysis includes only those receiving the placebo.

The analysis does not suggest that endorsements have an effect on public opinion, and indeed, when combined with results reported earlier, they suggest that the largest campaign effect is due to the arguments used in favor of STV. However, the message treatment conflates both a message and an endorser effect, as some respondents received a message that was endorsed by a leading public figure. The final model estimated separates these effects by including an interaction term which allows the effect of the message to vary with endorsement. The interaction term reports the marginal effect of receiving a message endorsed by a leading public figure, over and above that obtained upon receiving either of the treatments -message or endorser- in isolation. Analyzing the coefficient we are unable to reject the hypothesis that the interaction effect is zero and thus that the main

effects are in fact constant. Similar results are obtained when, instead of comparing yes voters with no voters and those who abstain, the comparison group is no voters only.

TABLE 7. The impact of Message and Endorser on YES with fixed effects by voting area

	(1)	(2)	(3)
Message		0.0781	0.126*
		(0.0479)	(0.0710)
Some Endorser	0.0572	0.0426	0.0906
	(0.0470)	(0.0477)	(0.0709)
Message with Endorser			-0.0891
			(0.0479)
Constant	0.381***	0.346***	0.325***
	(0.0339)	(0.0401)	(0.0462)
Observations	478	478	478
Number of Voting Areas	77	77	77
R^2	0.004	0.010	0.012

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 8 uses the estimates derived in model 3 of Table 7, and reports the substantive effect of each of the treatments. Overall, whereas receiving a message from the campaign has a large effect on the likelihood that a respondent intends to vote yes, we find no evidence of direct or indirect effects of receiving an endorsement.

5.4. **Accountability or Fairness.** The preceding results suggest that it is the message that the campaign conveys, rather than individual opinion leaders who have an impact on public opinion. Our analysis also allows us to compare the impact of different messages used by the campaign when controlling for geographic and canvasser fixed effects.

5.5. **Matching Estimates.** Our main result is that that, whereas the campaign has an overall effect of around 12 percentage points, this effect is entirely due to the message that the campaign uses. We find no evidence of endorser effects in the data. However these results are based on fixed effects estimations that make arguably strong assumptions about

TABLE 8. Marginal Effects of Message and Endorser Treatments[a]

	Message Effects	
1	Marginal Effect of Receiving a Message when there is no Endorser	+13.6 percentage points [Standard error = 7.1 points]*
2	Marginal Effect of Receiving a Message when there is an Endorser	+8.6 percentage points [Standard error = 6.6 points]
3	Overall Average Marginal Effect of Receiving a Message	+11.1 percentage points [Standard error = 6.0 points]*
	Endorser Effects	
4	Marginal Effect of Seeing an Endorser when there is no Message	+5.6 percentage points [Standard error = 7.7 points]
5	Marginal Effect of Seeing an Endorser when there is a Message	+0.6 percentage points [Standard error = 6.4 points]
6	Average Marginal Effect of Endorsements	+4.3 percentage points [Standard error = 4.8 points]
	Overall Effects	
7	Overall Marginal Effect of Canvassing (Endorser or Message)	+11.8 percentage points [Standard error = 5.7 points]**
8	Observations	301

[a] * Significant at the 90% level , ** Significant at the 95% level. Estimates are from OLS models with voting area fixed effects. Similar estimates obtain from model with matching at the voting area level.

TABLE 9. Impact of Accountability and Fairness Messages

	(1) VOTE	(2) VOTE	(3) YES	(4) YES
Fairness	0.121**		0.0320	
	(0.0585)		(0.0561)	
Accountable		0.113*		0.144**
		(0.0593)		(0.0590)
Constant	0.586***	0.602***	0.371***	0.365***
	(0.0340)	(0.0343)	(0.0333)	(0.0335)
Observations	216	214	357	336
Number of voting areaa	65	63	76	74
R^2	0.028	0.023	0.001	0.022

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

the data. An alternative is to estimate a matching model, the reports of which we report below.

TABLE 10. Impact of Campaign and Message with Matching

	(1)	(2)	(3)	(4)
Campaign	0.106*			
	(0.0538)			
Message		0.0772		
		(0.0485)		
Fairness			0.0349	
			(0.0572)	
Accountable				0.125**
				(0.0598)
Constant	0.325***	0.365***	0.371***	0.370***
	(0.0461)	(0.0339)	(0.0333)	(0.0352)
Observations	413	413	310	297
Number of voting areas	77	77	77	77
R^2	0.009	0.006	0.001	0.015

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

6. CONCLUSION

Our study has analyzed whether leaders are effective because of some innate qualities—for example, clarity, trustworthiness or focality—or because of the particular arguments they employ? To analyze these effects systematically we analyzed a political campaign, working together with the BC-STV campaign in the May 2009 British Columbia referendum on electoral reform. By randomly assigning canvassers to voting areas, who then randomly assigned messages to households, we were able to obtain estimates of two leadership effects: (i) the between voting area differences in opinion that are due to the assignment of different canvassers; and (ii) the impact on the campaign of endorsements by leading public figures. Neither of these strategies yields any evidence of leadership effects: a fairly substantial overall campaign effect is entirely due to the arguments used by the campaign, not to individuals making the case for reform. Thus, although our research design was developed in order to detect leadership effects that may be masked by standard approaches using observational data, we were unable to detect leadership effects even with this improved design.

REFERENCES

CALVERT, R. L. (1995): "The Rational Choice of Social Institutions: Cooperation, Coordination, and Communication," in *Modern Political Economy: Old Topics, New Directions*, ed. by J. S. Banks, and E. A. Hanushek. Cambridge University Press, Cambridge, UK.

DEWAN, T., AND D. P. MYATT (2007): "Leading the Party: Coordination, Direction amd Communication," *American Political Science Review*, 101(4), 825–843.

——— (2008): "The Qualities of Leadership: Communication, Direction and Obfuscation," *American Political Science Review*, 102(3).

HAFER, C., AND D. LANDA (2007): "Deliberation as Self-Discovery and Institutions for Political Speech," *Journal of Theoretical Politics*, 19(3), 329–360.

——— (2008): "Majoritarian Debate," unpublished paper, New York University.

HUMPHREYS, M., W. MASTERS, AND M. E. SANDBU (2006): "The Role of Leadership in Democratic Deliberations: Results from a Field Experiment in São Tomé and Principe," *World Politics*, forthcoming.

MYERSON, R. B. (2004): "Justice, Institutions, and Multiple Equilibria," *Chicago Journal of International Law*, 5(1), 91–107.

WANTCHEKON, L. (2003): "Clientelism and Voting Behavior: Evidence from a Field Experiment in Benin," *World Politics*, 55.