

Direction des bibliothèques

AVIS

Ce document a été numérisé par la Division de la gestion des documents et des archives de l'Université de Montréal.

L'auteur a autorisé l'Université de Montréal à reproduire et diffuser, en totalité ou en partie, par quelque moyen que ce soit et sur quelque support que ce soit, et exclusivement à des fins non lucratives d'enseignement et de recherche, des copies de ce mémoire ou de cette thèse.

L'auteur et les coauteurs le cas échéant conservent la propriété du droit d'auteur et des droits moraux qui protègent ce document. Ni la thèse ou le mémoire, ni des extraits substantiels de ce document, ne doivent être imprimés ou autrement reproduits sans l'autorisation de l'auteur.

Afin de se conformer à la Loi canadienne sur la protection des renseignements personnels, quelques formulaires secondaires, coordonnées ou signatures intégrées au texte ont pu être enlevés de ce document. Bien que cela ait pu affecter la pagination, il n'y a aucun contenu manquant.

NOTICE

This document was digitized by the Records Management & Archives Division of Université de Montréal.

The author of this thesis or dissertation has granted a nonexclusive license allowing Université de Montréal to reproduce and publish the document, in part or in whole, and in any format, solely for noncommercial educational and research purposes.

The author and co-authors if applicable retain copyright ownership and moral rights in this document. Neither the whole thesis or dissertation, nor substantial extracts from it, may be printed or otherwise reproduced without the author's permission.

In compliance with the Canadian Privacy Act some supporting forms, contact information or signatures may have been removed from the document. While this may affect the document page count, it does not represent any loss of content from the document.

Université de Montréal

**Experimentation and Political Science:
Six Applications**

par

Peter John Loewen

Département de science politique

Faculté des arts et des sciences

Thèse présentée à la Faculté des études supérieures
en vue de l'obtention du grade de Philosophiae Doctor (Ph.D.)
en science politique

Avril 2008

©Peter John Loewen, 2008



Université de Montréal
Faculté des arts et des sciences

Cette thèse intitulée:
Experimentation and Political Science:
Six Applications

par

Peter John Loewen (LOEP13027903)

a été évaluée par un jury composé des personnes suivantes:

Denis Saint-Martin
président-rapporteur

André Blais
directeur de recherche

Patrick Fournier
codirecteur

Richard Nadeau
membre du jury

James Druckman
examineur externe

Chantal Benoit Barre
représentante du doyen de la FES

Résumé

Cette thèse démontre l'utilité de l'expérimentation en science politique à l'aide de six articles. Bien que disparates quant à leurs questions et sujets, ils sont tous liés au comportement et à la psychologie politique.

Le premier article examine le rôle des considérations qui ne sont pas de l'ordre de l'intérêt personnel dans la formation des préférences pour les dépenses publiques. Je mesure l'altruisme des répondants par un jeu du dictateur. Je démontre qu'un niveau d'altruisme élevé prédit un plus fort appui pour des dépenses publiques, malgré le coût pour le répondant.

Le deuxième article s'intéresse à un paradoxe central de la participation politique: s'il est peu probable que le vote d'un simple citoyen décide du résultat d'une élection, pourquoi y a-t-il autant d'électeurs? Je démontre à l'aide d'un modèle formel que certains votent pour le bénéfice des autres. À l'aide d'une série de jeux du dictateur, je montre ensuite que les préférences variées pour certains partisans prédisent la décision de voter ou non.

Le troisième article présente les résultats d'une expérience de terrain sur les capacités de persuasion du courrier publicitaire. Durant la campagne de Michael Ignatieff pour la direction du Parti libéral du Canada en 2006, nous avons assigné de façon aléatoire du courrier aux délégués qui s'étaient engagés auprès d'autres candidats, puis nous les avons sondés. Ceux ayant reçu du courrier ont ajusté leur évaluation des autres candidats à la hausse et ont moins bien classé M. Ignatieff.

Le quatrième article montre comment le modèle de Bradley-Terry peut être utilisé pour analyser le pouvoir persuasif de l'argumentation. Au cours du référendum d'octobre 2007 sur la réforme électorale en Ontario, nous avons assigné à chacun des 520 répondants de notre sondage un argument en faveur et un argument contre la réforme, avant de leur demander leur opinion. Les arguments pour le système existant profitent d'un avantage général, tout comme les arguments qui font appel à la justice et à la représentation locale. Les arguments qui mentionnent les partis politiques sont moins persuasifs.

Le cinquième article pose la question des divers niveaux d'altruisme chez ceux qui s'identifient à un parti politique au Canada. Nous observons les différences entre les partisans dans leur allocation aux co-partisans, à d'autres partisans, et à des individus anonymes dans quatre jeux du dictateur lors d'un sondage en ligne. Tous donnent plus aux co-partisans et moins aux autres partisans. Les Néo-démocrates sont plus altruistes en moyenne que les Conservateurs et les Libéraux.

Dans le sixième article, nous nous demandons si le vote obligatoire mène aux effets de second ordre d'augmentation des connaissances et d'engagement des citoyens. Nous avons conduit une expérience de terrain avec des étudiants en âge de voter d'un cégep de Montréal. Certains étaient payés pour compléter deux sondages, d'autres étaient aussi payés pour voter lors de l'élection provinciale. Nous avons trouvé peu d'indications d'effets de second ordre.

Mots clés: expérimentation; comportement politique; psychologie politique; vote; économie comportementale; campagnes électorales

Summary

This dissertation demonstrates the usefulness of experimentation in political science through six articles. While eclectic in their questions and subjects, the articles all fall under the rubric of political psychology and behaviour.

The first article examines the role of non-self-interested considerations in the formation of preferences for public spending. I measure the altruism of respondents through the use of a dictator game experiment in a large online survey. I demonstrate that greater levels of altruism predict greater support for public spending, even when it comes at a cost to the respondent.

The second article addresses the “paradox of participation”: if the probability of a single vote deciding an election is so low, why do we still observe large numbers of voters? I demonstrate in a formal model that some individuals will vote because of the benefits accrued to others. Empirically, preferences for partisans which differ in a dictator game significantly predict the decision to vote.

The third article presents the results from a field experiment into the persuasive capacities of direct mail. The experiment was conducted in 2006 with the Michael Ignatieff campaign for the leadership of the Liberal Party of Canada. The experiment randomly assigned a direct mailing to delegates pledged to other candidates. We then surveyed these delegates. We find that those who received mail appeared to adjust their evaluations of other candidates upwards and to move Ignatieff down in their preference rankings.

The fourth article uses a Bradley-Terry model to analyze the persuasive power of arguments in a survey experiment conducted during the October 2007 Ontario referendum on electoral reform. We assigned each of 520 respondents one of six arguments for and one of six arguments against electoral reform and then measured their preference for reform. We show that arguments for the existing system enjoyed an advantage, arguments which appeal to fairness and local representation were significantly more persuasive, and arguments which mention political parties were less persuasive.

The fifth article asks if partisan identifiers in Canada differ in their levels of altruism. We examine the behaviour of partisans in four dictator games in an online survey. We compare differences between partisans in their allocations to co-partisans, other partisans, and anonymous individuals. All partisans consistently allocate the most to co-partisans and the least to other partisans. Anonymous individuals are in the middle. We also find that New Democrats are more altruistic than Conservatives or Liberals.

The sixth article asks whether compulsory voting leads to the “second-order” effects of increased citizen knowledge and engagement. We conducted a field experiment among voting-aged students at a Montreal CEGEP. Our intervention involved paying some students to complete two surveys while paying another group to also vote in a provincial election. We find little evidence of “second-order” effects.

Keywords: Experimentation; Political Behaviour; Political Psychology; Voting; Behavioural Economics; Campaigns

Table de matières

Résumé	i
Summary	iii
Liste des tableaux	xii
Liste des figures	xvi
Dédicace	xvii
Remerciements	xviii
1 Introduction	1
1.1 The Underside of Warren Miller’s Cowboy Boots	2
1.2 Defining Experimentation and Experimental Types	4
1.2.1 Experimental Types	7
1.3 An Eye Out for Inferential Monsters: The Case for Experi- mentation in Political Science	16
1.4 The Quantity of Experimentation in Political Science	20
1.4.1 The Nature of Experimentation in Political Science	23
1.5 The Qualities of Experimentation in Political Science	26
1.5.1 Statistical conclusion validity	29

1.5.2	Internal validity	31
1.5.3	Construct Validity	32
1.5.4	External Validity	35
1.5.5	Comparing the types	37
1.6	The Articles	38
1.6.1	Dictators and Purses: Altruism and Support for Public Spending	39
1.6.2	Antipathy, Affinity, and Political Participation: How Our Concern for Other Partisans Makes Us Vote	40
1.6.3	For Want of a Nail: Direct Mail and Negative Persua- sion in a Leadership Race (with Daniel Rubenson) . . .	41
1.6.4	Testing the Power of Arguments with a Bradley-Terry Model (with Daniel Rubenson and Arthur Spirling) . .	42
1.6.5	Partisanship and Altruism: Results from a Dictator Game Experiment (with Angelo Elias)	44
1.6.6	Does Compulsory Voting Lead to More Informed and Engaged Citizens: An experimental test (with Henry Milner and Bruce M. Hicks)	45
1.7	A Closing Word	46
2	Dictators and Purses: Altruism and Support for Public Spend- ing	47
2.1	Introduction	48
2.2	Altruism and Public Spending	51

2.2.1	Defining Altruism	51
2.2.2	Theoretical Link	53
2.2.3	Previous Findings	55
2.2.4	Innovation	58
2.3	Altruism and Dictator Games	59
2.4	Survey Design and Participants	62
2.4.1	Participants	66
2.5	Results	68
2.5.1	Support for Public Spending	69
2.6	Discussion and Conclusion	83
3	Affinity, Antipathy and Political Participation: How Our Concern For Other Partisans Makes Us Vote	85
3.1	Introduction	86
3.2	Group Politics	88
3.3	A Different Calculus of Voting	91
3.4	Survey and Research Design	97
3.4.1	Survey	97
3.4.2	Subject Profiles	98
3.4.3	Antipathy, Affinity, and other variables	101
3.5	Antipathy, Affinity, and Party Identification	104
3.5.1	Why Antipathy and Affinity are not just Party Identifi- cation	107
3.6	Antipathy, Affinity, and Turnout	110

3.7	Discussion and Conclusion	118
4	For Want of a Nail: Direct Mail and Negative Persuasion in a Leadership Race (with Daniel Rubenson)	121
4.1	Introduction	122
4.2	Direct Mail and Persuasion	124
4.3	The Race	135
4.4	The Experimental Study	138
4.4.1	The Experiment	139
4.4.2	The Survey	140
4.5	Results and Discussion	143
4.6	Conclusion	149
5	Testing the Power of Arguments with a Bradley-Terry Model (with Daniel Rubenson and Arthur Spirling)	152
5.1	Introduction	153
5.2	Data, Context and Connections	158
5.2.1	Experimental Design	161
5.2.2	The ‘Framing’ Connection	164
5.3	Measuring the Power to Persuade: A Statistical Model	166
5.4	The Power of Arguments: Conventional Results	170
5.4.1	Testing the Power of Arguments	170
5.4.2	Determining the Sources of Power	172
5.5	The Power of Arguments: The Bradley-Terry Method	175
5.5.1	Unstructured Results	175

5.5.2	The Sources of an Argument's Power	180
5.6	Further Applications and Conclusion	185
6	Partisanship and Altruism: Results from a Dictator Game Experiment (with Angelo Elias)	187
6.1	Introduction	188
6.2	Altruism and the Dictator Game	189
6.3	The Study	191
6.3.1	Subjects	191
6.3.2	The Survey	193
6.3.3	Dependent Variable: Dictator Game Allocations	193
6.3.4	Independent Variables	194
6.4	Results	196
6.5	Discussion and Conclusion	202
7	Does Compulsory Voting Lead to More Informed and En- gaged Citizens: An experimental test (with Henry Milner and Bruce M. Hicks)	204
7.1	Introduction	205
7.2	Existing Knowledge and the Case for Experimentation	206
7.3	Hypotheses and Experimental Design	209
7.3.1	Hypotheses	209
7.3.2	Subject Recruitment and Survey Administration	211
7.3.3	Survey and Dependent Variables	214
7.3.4	Sample Profile	216

7.4	Results	218
7.5	Conclusion	222
8	Conclusion	225
8.1	Introduction	226
8.2	Summary	226
8.3	Implications	231
8.3.1	Methodological Implications	231
8.3.2	Substantive Implications	233
8.4	Future Research	235
8.5	The Sound Down the Hall	238
	Références bibliographiques	239
	Annexe A: Criterion Validity of Dictator Game as a Measure of Altruism	xxvi
	Annexe B: Dictator Game Instructions	xxxii
	Annexe C: Question Wording and Variables for Dictators and Purses	xxxv
	Annexe D: Question Wording and Variables for Affinity, Antipathy, and Political Participation	xxxviii
	Annexe E: Treatment Assignment for For Want of a Nail	xli
	Annexe F: Sample and Subject Profile for Bradley-Terry Experiment	xliv

Annexe G: Additional Logit Results for Bradley-Terry Experiment	xlvi
Annexe H: Predicted Probabilities of FPTP dominance in Structured versus Unstructured Bradley-Terry Models	xlviii
Annexe I: Question Wording and Variables for Partisanship and Altruism	1
Annexe J: Supplementary Tables for Partisanship and Altruism	liv
Annexe K: Treatment Assignment for Compulsory Voting Experiment	lix
Annexe L: Question Wording and Variables for Compulsory Voting Experiment	lxii
Annexe M: Accord des coauteurs et permission de l'éditeur	lxviii
Curriculum Vitae	lxxvi

Liste des tableaux

1.1	Frequency of Statistical Techniques in Observational and Experimental Articles	24
2.1	Sample demographic and political characteristics	67
2.2	Willingness to pay for a carbon tax (Ordered Logistic Regression)	75
2.3	Willingness to pay for free university tuition (Ordered Logistic Regression)	76
2.4	Willingness to pay more taxes to reduce waiting times for medical services (Ordered Logistic Regression)	77
2.5	Willingness to wait longer for medical care if cancer patients can receive drug coverage (Ordered Logistic Regression)	78
2.6	Main and conditional effects of altruism when interacted with education and income (Ordered Logistic Regression)	82
3.1	Sample demographic and political characteristics	100
3.2	Partisanship and Average Allocations in the Dictator Game	107

3.3	Within-Subject Differences in Dictator Game Allocations (Wilcoxon Sign-Rank Differences)	109
3.4	Antipathy, Affinity and Turnout (Logistic Regression, Odds Ratio)	112
3.5	Separating Antipathy and Affinity (Logistic Regression)	115
4.1	2006 Liberal Party leadership election results	137
4.2	Effects of Ignatieff mail on average leadership candidate ratings (mean differences)	145
4.3	Effects of Ignatieff mail on delegates' preference ordering (Ordered Logistic Regression)	146
5.1	Argument Characteristics	165
5.2	Agreement with MMP by FPTP Arguments (mean differences)	172
5.3	Agreement with MMP by MMP Arguments (mean differences)	172
5.4	Logistic Regression of Argument Power ^a	173
5.5	Logistic Regression of Sources of Argument Power	176
5.6	Bradley-Terry Model of Argument Power	178
5.7	Power of FPTP Arguments	180
5.8	Power of MMP Arguments	180
5.9	Model of Sources of Argument Power	181
6.1	Dictator Game Allocations to Anonymous Individuals (OLS)	197
6.2	Dictator Game Allocations to Other Partisans (OLS)	198
6.3	Dictator Game Allocations to Co-Partisans (OLS)	199

6.4	Pooled Dictator Game Allocations (OLS)	200
7.1	Sample Profile	217
7.2	Effects of compulsory voting treatment on political knowledge, political discussion, and media usage (mean differences)	219
7.3	Effect of treatment on knowledge, news consumption and discussion of politics for voters and non-voters (OLS)	222
A.1	Altruism and Self-Reported Charitable Giving (Ordered Logistic Regression)	xxix
A.2	Altruism and Response to those in Distress (Ordered Logistic Regression)	xxx
A.3	Altruism and Willingness to Help Those Who Don't Help Themselves First (Ordered Logistic Regression)	xxxi
F.1	Sample Demographics and Political Characteristics	xliv
G.1	Logistic Regression of Argument Power with Argument Matchings	xlvii
H.1	Absolute difference in predicted probabilities of FPTP dominance by structured and unstructured models ^a	xliv
J.1	Dictator Game Allocations to Anonymous Individuals – Low Education (OLS)	liv

J.2 Dictator Game Allocations to Other Partisans – Low Education (OLS)	lv
J.3 Dictator Game Allocations to Co-Partisans – Low Education (OLS)	lv
J.4 Dictator Game Allocations to Anonymous Individuals – High Education (OLS)	lvi
J.5 Dictator Game Allocations to Other Partisans – High Education (OLS)	lvi
J.6 Dictator Game Allocations to Co-Partisans – High Education (OLS)	lvii
J.7 Pooled Dictator Game Allocations – Low Education (OLS) .	lvii
J.8 Pooled Dictator Game Allocations – High Education (OLS) .	lviii

Liste des figures

1.1	Share of Experimental Articles in <i>AJPS</i> , <i>APSR</i> , and <i>JOP</i> . . .	22
2.1	Dictator Game Allocations	69
2.2	Demand Curves for Public Spending	72
3.1	Affinity, Antipathy, and Turnout	117
5.1	Power of arguments: unstructured results	179
5.2	Power of arguments: structured results	184

For A.K.

*I am a lazy cat
She is as pure as the cold driven snow*

and

*For my mother and my father
To whom I owe more than I can pay*

Remerciements

I push on open doors. That is to say, I am at a point which is supposed to be the result of great effort and sacrifice, but for me it has all seemed, if not easy, then at least certain. What keeps me from pride, then, is that I am quite certain that others have done most of the work for me. Thus my debts are large. I shall try to be as sufficient in addressing them as I have been adept in amassing them.

I am indebted to Brian Thomas and James Baxter. From the time I was a child, I never lacked in intellectual confidence (whether I lacked in ability is another question!). But Thomas and Baxter were instrumental in affirming a life of intellectual pursuit and in convincing a young man in a small city that with some work his confidence could be justified, and that it was worth a shot. Somewhere a ledger exists and their names are written there in red ink.

This red ink continued a steady flow at Mount Allison University. To decamp to Sackville, NB was the best decision I ever made and I attribute it still to luck, whimsy and my parents' confidence and insistence that I leave my home town for my education. It had nothing to do with rational calculation. I owe a great debt to my first professors, especially Wayne Hunt, for teaching me that politics was ultimately about ideas, Janine Rogers, for teaching me about poetry and for that harrowing drive through Algonquin Park, and Frank Strain, for teaching me about market failures and for always

offering a place to stay. I thank, too, Rev. John Perkin and Charlie Hunter, who continue to share their wisdom with me. I should also like to thank the presidents of that institution, Ian Newbould, A. Wayne MacKay, and Robert Campbell. That I receive regular offers of assistance from all of them is a testament to how lucky I was to end up at such a personal and exceptional place. And I thank finally and most sincerely Bill Cross. It was Bill who supervised my final paper, who taught me something about being a research assistant, who introduced me to so many of Canada's political scientists, and who arranged for me to work with André Blais. He opened many doors and always made me think it was my doing.

My time in graduate school has been no less indebted. I've had the great privilege to be taught by Stuart Soroka, Elizabeth Gidengil, Louis Massicotte and André J. Bélanger, whose imprint I should think is indelible. During a sojourn in America, I was taught and greatly influenced by Jim Johnson, Kevin Clarke, Dick Niemi, and Bing and Linda Powell. They too are in my debts, but they'll have to cross the border to collect. Outside of my formal studies, I've also been lucky to form connections and friendships with Jamie Druckman, Skip Lupia, John Aldrich, James Fowler, and Charles Blattberg. Their influence is thorough-going, and for this I am richer.

Of special note are the constant and unmatched intellectual critiques I have received from Patrick Fournier, my co-director. Let him move towards the front of the line of those owed.

For financial support, I thank SSHRC, the Canada Research Chair in

Electoral Studies, and the graduate faculty at Montreal. For his unflagging willingness to fund basic research, I thank Greg Lyle and his firm, Innovative Research Group. My dissertation owes much to him.

I must thank, as well, the great colleagues I have had in Montreal. The Chair has been a wonderful place to work, and I feel lucky to count and have counted as colleagues and sometimes collaborators Pascal Doray-Demers, Henry Milner, Silvina Danesi, Sébastien Dubé, Angelo Elias, Anne-Marie Grenier, Bruce Hicks, Simon Labbé St-Vincent, Aina Gallego Dobon, François Gélinau, Indridi Indriason, Simon McDougall, André Perrella, Éric Bélanger, and Christophe Chowanietz. I owe special thanks to those with whom I shared an office at various times, namely Gianluca Cairo, Kaat Smets, Romain Lachat, Matthew Gibbons, and Jiyeon Kim. I only hope, on balance, that my stories and jokes were enough to outweigh my less redeeming qualities.

I have deliberately left out four colleagues, because their influence on me has always been more than merely professional. Marc-André Bodet was one of the first people I met in Montreal. We had offices across the hall from one another, in which we both worked alone. When he came across the hall one day with an article on behavioural economics he had clipped for me from the *Economist* I knew it was, if not true love, then at least the start of something beautiful. We decided to share an office. What has followed is a friendship I should hope will continue for at least as long as I am in need of help with statistics and Marc is in need of some half-fabricated story of my adventures.

Agnieszka Dobrzynska (known officially in our Parliament as Mr. Loewen's colleague) has been a constant during my time in Montreal. I am not sure she is aware of the value of her counsel at important times over the last five years, but I am aware enough for the both of us.

Daniel Rubenson has been a great colleague and even better friend. I don't know how two morose people can cheer one another up, but our daily phone calls rarely fail to set me right. What is more, they never fail to keep the ideas coming, which is the most important thing in this business. Daniel, may we always be upgraded to business class, may we always get a suite, and may we never decline a second shish taouk in the early morning hours.

Finally, Arthur Spirling has been a friend and colleague of the first order. We lived together during my time in Rochester, where Arthur was also my TA. If given a choice between being my TA again or being imposed upon by another visit from my Canadian friends, I am not sure which he would choose. But I am entirely confident he knows a clever method from some other discipline by which to determine this. His intelligence is surpassed only by his decency.

I have been incredibly lucky in life to be surrounded by people who seem to exceed me in every virtue but who are still solicitous of my company. I owe much to my friends.

Leslie Seidle, Danistan Saverimuthu, Mike Wodzicki, Reynolds Mastin, Will Paterson, and Robert Asselin have been great dinner companions, witty correspondents, sources of encouragement at difficult times, and sources of

great intellectual stimulation at other times. Geeta Yadav has been a welcome source of humour, conversation, and company. I am thankful for the friendship of Lorena Ruci, for nicknames if nothing else. Chris Tucker and Kate Hutchison and Matt Gauthier and Beth Hickey have been great sources of entertainment and hospitality for all my time in this city. Jason Whiting and Jeanne Bowser have been fine landlords and great friends. Jenn Nachshen has been a great source of invitations and advice and Katherine Pickering has been a constant source of great dessert. I thank Stefanie Von Hlatky for all those coffees and that squash game, and Kat Porowski for walks on the mountain.

I thank Nick Gallus for leaving keys with the concierge and for always insisting I could pay for the next dinner. I thank David Myles for listening to my songs and for sharing his own with me. I thank Nina Corfu for putting up with this on so many Saturday nights. I thank Andrew Black, both for setting me straight and settling tabs. I thank Ted Rutland for uncommon goodness. And I thank Loren McGinnis. If ever I feel badly about stretching a story, I think of Loren. And if ever I need inspiration to create new stories, I think of him, too.

My debt to those with whom I have lived is large. Tyler Ball, Shawn Walton, and Jeremie Gomand made Westmount a nice place to live. My great housemates from Mount Allison are still constants in my life. Corey Arsenaault taught me everything I ever needed to learn about dedication and focus. Maskull Lasserre and Holly Tingley taught me, and continue to teach

me, about what it means to live by one's own standards. I remain in wonder of their kindness and excellence. Ryan Merritt taught me how to play the guitar properly. If he'd only teach me how to dance properly, I would be complete. Anthony Kingsley and Christian Friis have been better friends than I've ever deserved. May we have one more weekend at Voltaire's place in Rockcliffe and one more run in the Volvo.

With debts to friends have come debts to their families. For their constant hospitality, I thank the Kingsley family, the Friis clan, the Merritts and the Lasseres. For years of kindness and concern, I thank Moshe and Ina Dobner. I thank Peter and Delia Kermack for so many fine dinners in Senneville. I thank Bill and Lorraine Black for putting so much trust in me and welcoming me back despite the outcome. Finally, a special thanks is owed to Gord and Deb Valliant for so many years of love.

Noel Baldwin has been a firm friend for ten years and my housemate for the last three. I can only assume his patience has not run its course because he travels so much and is thus out of the apartment. I am terribly lucky to count him among my closest friends.

I owe much to Elizabeth and Hart Shouldice. Liz, for being like a sister. And Hart, for being a younger brother with whom I could commiserate this fact. I thank, too, Marilyn and Rod McIvor for their warm welcomes. I should like to thank Drew Shouldice, whom I never met but whom I know through the values and lives of his children.

If ever I thought that the stars were somehow aligned in my favour, it's

because women as intelligent and beautiful as Alicia Johnston and Gallit Dobner gave me the time of day. That both are still important parts of my life speaks much more to their credit than mine.

I am deep in the red to Sam Millar and Anamitra Deb. One should be so lucky to meet one person so impressively well-rounded in their life. That I count two among my friends stretches all likelihoods. Deb, may the goats roam free, the birds pick at bottle labels, and the patrons of the world's diners live vicariously through our Sunday morning takings of stock. Sam, may you always ride lead, may I always keep your tail lights in my sights, and may we always make the ferries on time.

Brett Valliant is my oldest and dearest friend. I cannot remember a time in my life prior to him and should not wish to imagine a time after. He never fails to restore my confidence.

For André Blais, there is little I can say to account for my thankfulness. I shall leave it to the last six lines of Keats' "Dedication to Leigh Hunt, ESQ.",

But there are left delights as high as these,
And I shall ever bless my destiny,
That in a time, when under pleasant trees
Pan is no longer sought, I feel a free,
A leafy luxury, seeing I could please
With these poor offerings, a man like thee.

I owe a great debt to Andrea Kermack. No one I know lives a more examined life. Thinking of what my life would be like had I not met her is not a counterfactual I ever wish to estimate. Wherever I may find her, I will

try to square the debt.

To extended family, especially my Grandma and Grandpa and Nana and Papa, I offer my deepest thanks. Their pride and support is as unceasingly as their bewilderment at what I actually do for a living.

To James, Candace, Jacob and Aria, and to Shelley, Dan, Anna, Josh, and Eva. The only time when I question the life I have chosen is when I contemplate the joy my siblings derive from their families. I should hope to be so lucky in my life to one day have children so beautiful and unique. And they should be so lucky to have a brother and sister so patient and indulgent.

And, finally, to my parents. I have plumbed the depths of their love and my line has no slack.

1

Introduction

1.1 The Underside of Warren Miller's Cowboy Boots

In 1962, Richard Niemi graduated from Lawrence College, in his home state of Wisconsin. He studied under William Riker who would depart for the University of Rochester the same year Niemi would leave for the University of Michigan. Those were heady days in Ann Arbor. The American Voter had been released just two years earlier, and Angus Campbell, Philip Converse, Warren Miller and Donald Stokes were still in the process of revolutionizing political science. (That Niemi would participate in a second revolution at the University of Rochester speaks either to his foresight or his luck, or both).

At the end of his second year, Niemi faced his preliminary exams. He cannot now even remember whether they had a written component. But he remembers clear as day the oral component to be administered by Kent Jennings and Warren Miller. As the story goes, Niemi entered the examination room all nervous and kinetic and filled to the brim with knowledge. But rather than test his knowledge of the literature, Miller, the laconic Westerner, reposed with his cowboy boots upon his desk and asked Niemi a single question. "If we gave you \$150,000 to run an election study, " Miller inquired, "what would you do?"

This dissertation is my answer to that question. Were I ever to see the underside of Warren Miller's cowboy boots and have the question put to me, I know my answer. I would conduct as many experiments as I could imagine, as

creatively as I could muster, to better understand why individuals participate in politics and how they form their preferences.

In the pages that follow, I give my definition of experimentation and provide examples of four types of experiments. I then make a theoretical case for why an experimental framework is superior to a strictly observational one. Experimentation, I argue, gives us potentially unadulterated views into causal processes and preferences in a way not possible from a strictly observational approach. After making this argument, I survey the frequency and nature of experimentation in political science.

I find that while experimentation is slowly increasing in political science generally, it is experiencing a much more pronounced growth in political behaviour and political psychology. It does so, I argue, with greater statistical clarity and with more collaboration. It is also having a greater intellectual impact than equivalent observational research.

I then turn to questions of validity and argue that all types of experiments, given proper design, can possess high validity. Experimentation, then, does not require a retreat to the laboratory. I then conclude with a description of my experiments. I outline the contributions they make to their respective questions while articulating why experimentation was fundamental to these contributions.

1.2 Defining Experimentation and Experimental Types

Any study of experimentation requires a definition of the concept. I adopt three conditions. An experiment is a research *process* with the following three conditions:

1. Treatment.
2. Assignment to treatment or control/comparison groups is random and the process of randomization is known.
3. Ex-post measurement of results.

Each condition deserves some elaboration. Treatment refers to some intervention by a researcher whereby the units of analysis receive some stimulus. This stimulus can then be identified and expressed in a single variable or in several variables of interest. Examples include school vouchers, information about a politician, water and sunlight, placement in a political system, or participation in one game or another. The substance of treatments is quite close to infinite. Moreover, they are not constrained to one variable; they can be quite complex (Morton and Williams, 2006).

Random assignment to treatment or control groups requires that every unit have an equal probability of being assigned to a control or a treatment group (which is sometimes conditioned on some characteristic of the unit (Gibson, Caldeira and Spence, 2002)). The corollary of this claim is that

assignment to treatment must be independent of a subject's characteristics, unless specified by the treatment. An important and logical extension of this randomization requirement, especially as it relates to behavioural economics games, is that if the experiment involves placing subjects into some institution or scenario in which they are observed, then placement must be independent of subject characteristics.

It is important to note that an experiment will not always have a control group. Consider, for example, an experiment in which some survey respondents are asked to respond to a question which gives response options in the order A, B, C, while other survey respondents are asked to respond in the order C, A, B. Which of these groups is the control group is really a matter of taste. Accordingly, control comes not from having respondents in some control group in which no treatment occurs. Rather, it comes from dividing respondents into treatment groups randomly and with equal probability across subjects. In some cases, this will mean assigning some respondents no treatment (which corresponds to a more classical view of control groups), while in other instances, it will involve assigning subjects to different treatments. It is for this reason that Shadish, Cook and Campbell (2000) use the phrase "comparison group" rather than "control group."

Knowledge of the randomization procedure is the second half of the second condition. This is a logically necessary condition, because it is not possible to ensure that some inquiry involves random assignment to conditions without knowing the assignment procedure. As such, a researcher could take

data from the assignment of criminal cases to judges (which is known to be random) and then analyze the data as an experiment. Or, knowing that lottery wins are random, a researcher could examine the effects of increases in affluence (through lottery wins) on attitudes towards redistribution (e.g. Doherty, Gerber and Green, 2006). Because any effects would be a result of a random process and not a result of the unobserved actions of or events related to a respondent, then changes in attitudes can be said to be caused by an increase in wealth. But, a researcher could not conduct an experiment in the case where the distribution of the treatment merely appears random *ex-post* unless the researcher can demonstrate that there is not an unknown confounding variable which affected the distribution of treatment. In most cases, this is theoretically impossible given that the variable is unknown.

Finally, the insistence of ex-post measurement implies that an experiment is not complete until differences between treatments have been measured. An experiment is more than a research design; it is a complete inquiry which concludes with ex-post measurement of difference. This draws out the distinction between an experimental design and an experiment.

As a whole, this definition differs from Druckman et al. (2006)'s definition of an experiment by not insisting that the manipulation or application of the treatment under the control of a researcher. If we can be certain that a process has been randomized, as in the assignment of judges to cases, then this is theoretically identical to manipulation by a researcher. Accordingly, this insistence on randomization also differentiates it from Morton and Williams

(2006), who require only that a researcher intervene in a data generating process. My definition is thus closest to the definition of Gerber, Green and Kaplan (2004), though they do not make an explicit requirement for post-hoc measurement.

1.2.1 Experimental Types

There are many variants of experimentation present in our discipline. I identify four ideal types. In order of (estimated) frequency these are: survey-embedded experiments, laboratory experiments, games-based experiments, and field experiments. In the case of each, I describe their general parameters, and then outline a study which I think represents a particularly good application. In a later section, I draw out the comparative benefits and drawbacks of each type.

My classification of experiments by type rather than subject contrasts with Bositis and Steinel (1987) and McDermott (2002). However, it does reflect a more recent survey of experimentation by Druckman et al. (2006) who identify three experimental types (laboratory, field, and survey). There are three compelling motivations for examining experimentation by type. First, types are easier to define than subjects. Second, we are far from exhausting the subjects of study, but we are close to exhausting types, so it is likely more efficient to examine types. Third, studying types lets us trace out the common and differing factors of the approaches, so we can better understand the trade-offs of each. By studying subjects, we may mistake the

theoretical imprecision of a subject for the shortcoming of an experimental type.

I diverge from Druckman et al., however, by separating laboratory and game-based experiments. My reasoning is entirely inductive. Games-based experiments have focussed almost exclusively on understanding basic economic behaviour within different institutional set-ups. These games almost always feature interaction between subjects. By contrast, laboratory experiments are more likely to deal with isolated subjects and are more likely to address a broader range of questions. Sufficient differences exist to suggest that these are different types of experiments.

The consideration of experiments by type also differs from Morton and Williams (2006). They classify experiments by location: laboratory, field, or internet. They fold game-based experiments into laboratory experiments, and classify survey experiments as either field experiments (if they are done on the phone) or internet experiments. Location seems more arbitrary than type, and certainly less precise. Indeed, because of the way it collapses together experiment types which are quite different in their history and applications it invites more difficulty in exploring comparative merits. Accordingly, I rest with a distinction based on type. I present below a prototypical experiment of each type.

Survey-embedded experiments

Survey-embedded experiments feature random assignment of some subjects to some treatment within a survey, and they measure opinions ex-post. For example, the random rotation of party names on a vote preference question is a survey experiment. Question-wording experiments are also survey-embedded experiments. These experiments are of increasing importance in political science, especially public opinion and voting research. They are part of a “new look in public opinion research” (Sniderman, 1992, 219). The rise of CATI (Computer Assisted Telephone Interviewing) has allowed researchers to randomize across populations, changing the content, order, nature, or form of questions (Bartels and Brady, 1992, 122-123). The ability to record response times has afforded scholars insight into the properties of many questions and responses (Gidengil, 2002).

Burden and Klobstad (2005)’s study of partisan identification provides a clear example of a well-conducted survey-embedded experiment. There has been a longstanding incongruence between the theoretical description and the measurement of partisan identification. To wit, partisan identification is described as an affective attachment or commitment to a party (e.g. Miller and Shanks, 1996; Green, Palmquist and Schickler, 2002), but has been measured as a cognitive construct. Burden and Klobstad thus want to “reconcile the standard conceptualization of party identification with its measurement”.

To tackle this problem, they design a simple survey experiment where half of respondents are asked whether they *feel* they are Republican or Democrat.

The other half are asked the standard question of whether they *think* of themselves as Democrat or Republican. Assignment is random, demonstrated by a lack difference in their two subsamples across various baseline covariates. All other parts of the question are held constant. To ascertain whether the questions invoke different considerations (Zaller, 1992) among the subjects, they time responses. By their logic, subjects who are surveying a wider range of considerations will take longer to respond.

Their results are striking. The *feel* responses suggest a measurably more Republican electorate than the *think* responses, which helps resolve a long-standing empirical quandary in American voting literature. Namely, Republican identification seems consistently appears much lower in surveys than Republican vote totals would suggest. The increased incidence of Republican identifiers among those who received the *feel* prompt suggests that existing survey questions have been undercounting Republicans. Second, they find that respondents take longer to give *feel* responses, which they take to be evidence of a more affective basis for partisan identification. Since affective judgements are often unconscious, they may take longer for subjects to process after they are invoked. From these two findings, Burden and Klofstad draw the conclusion that a proper measure of partisan identification would ask subjects how they feel rather than how they think. By conducting a simple survey experiment, they are able to provoke an important rethinking of a concept central to electoral studies.

Laboratory experiments

Laboratory experiments are rare in political science. They are only slightly less rare if one relaxes a laboratory to include classrooms and controlled environments (like a living room in the case of Ansolabehere and Iyengar (1997)). Nonetheless, laboratory experiments are playing a role in political science which rivals the importance of survey-embedded efforts. Lodge and Taber (2005)'s recent tests of the hot cognition theory of motivated reasoning (Lodge and Taber, 2000) provides a good example of a laboratory effort.

The motivation of their experiment is to test the "hot cognition" hypothesis, which contends that individuals' assessments of other individuals, groups, and issues are affectively-charged. As a result, present evaluations are coloured by past assessments, information seeking may be biased, and the integration of new, contradictory information may be difficult. Clearly, this has important consequences for those who study opinion formation and political cognition.

Lodge and Taber work from a clear theory which provides for both a hypothesis - hot cognition - and a testable mechanism - spreading activation - where one affective node activates nearby connected or associated nodes. With this theory and mechanism they design an experimental set-up with clear manipulations of stimulus which are independent of the characteristics of the individuals.

The experiment is quite simple. Seated in front of a computer screen, subjects are shown a randomly-selected word (the prime) for a short period

of time (200 ms). The prime is either a politician, a group/party, or an issue. The negativity or positivity of the word for each respective subject is determined in a survey following the experiment (this also gives them insight into the comparative valence of the target). There is a pause - which is manipulated within the experiment - and then another word (the target) is presented on the screen, and left there until subjects identify by keystroke that the target is positive or negative. Targets are words which are generally recognized as positive (like rainbow) or negative (like cockroach). This allows another treatment: the congruence of the target and the prime. Accordingly, Lodge and Taber put different subjects through three different versions of the experiment. Afterwards, they record the general political sophistication of the subjects, which allows them to derive and measure hypotheses about the interaction of sophistication with hot cognition.

Their measurement is mean response times, and most comparison occurs within subjects. However, because of cross-sectional variation in sophistication, they can also measure between subjects. The amount of variance they have at hand - both naturally and by treatment - allows for greater inference. Moreover, assigning treatment for word-order randomly across sophistication allows for even more variance.

In support of their contentions, Lodge and Taber find that individuals have a harder time evaluating words as positive or negative when they are incongruent with the prime. Perhaps most importantly, they find that response times increase with sophistication. The implication is clear and im-

portant: more sophisticated subjects are more likely to be biased information seekers and processors. While the comparative qualities of this approach are left to a later section, it is clear that experimental control and randomization undoubtedly allow for a large degree of insight into fundamental cognitive exercises of interest to political scientists.

Game-based experiments

Games-based experiments are multi-subject experiments in which subjects interact with one another¹, for material reward. As an example of an excellent game-based experiment, I take McKelvey and Palfrey (1992)'s "An Experimental Study of the Centipede Game". Roth (1995) has identified this as a hallmark in behavioural economics and Morton (2007) has argued for its importance to political science.

The experiment is based on the centipede game, a two-player interaction in which Player 1 can chose to take a pot of money, ending the game, or pass her turn to Player 2, who can similarly take a sum of money or pass her turn back to Player 1. The catch is that the sum of money a player receives from each turn grows (but declines slightly if the other player chooses to take). The game has a set number of iterations, and as such has a backwards induction induced equilibrium in which the first player takes on the first move. The problem, and the reason for the experiment, is that this action rarely holds in the laboratory. Rather, players often play well down the game tree, and

¹Or appear to interact with one another.

some play until the end of the game. As such, they violate fundamental assumptions of non-cooperative game theory and rationality. The purpose of McKelvey and Palfrey, then, is to explain this behaviour.

This desire to explain observed behaviour is important for at least two reasons. First, it demonstrates the role of game-based experiments in getting the larger assumptions right. Whereas many economists are happy to continue using models of full rationality, behavioural economists want to refine their models to better reflect fundamental behavioural realities. Second, the “cause” they are seeking to explore is less direct than that of Lodge and Taber, for example. Whereas Lodge and Tabor want to test a theory about the way the mind works, McKelvey and Palfrey want to work backwards from an observed regularity, and see if they can explain its structure.

To do so, they rethink the nature of the game, recasting it as one of incomplete information where voters have uncertainty over the utility functions of other players. In other words, they think there is some small chance that the person against whom they are playing is altruistic. From this conclusion they specify a new estimator of subject behaviour, and then apply it to their observed data. As it fits the data well, they come to the conclusion that they have better explained the process. It is a computationally intensive process, but it represents an unrivalled integration of theory, estimators, and results (Morton, 2007). As Morton has also observed, the econometric estimation of quantal-responses has had significant payoff in the fields of legislative bargaining (Morelli, 1999) and international conflict (Signorino, 1999). It was

derived from the observation of laboratory games aimed at understanding the structure of preferences and the nature of basic behaviour.

Field Experiments

Field experiments have experienced a marked growth in political science in recent years. Their infrequent use until now is not surprising. The requirements of a field experiment are significant: the researcher has to guarantee random assignment and completed treatment in a totally uncontrolled environment. Nonetheless, concerted institutional effort has been devoted to these experiments, and their growth indicates the effectiveness of these efforts.

It is fitting, perhaps, that the most notable early field experiment in political science was Gosnell (1930)'s study of voter turnout and mobilization, and that the most notable reintroduction of field experimentation came with Gerber and Green (2000)'s study of the same question. I review their experiment as a good example of a well-executed field experiment and well-measured results.

Gerber and Green seek to measure the comparative effects of three different voter mobilization techniques: canvassing, phone calls, and direct mail. In doing so, they put to test a hypothesis that declining turnout is not a function of declining levels of contact with political parties during elections, but with a shift in the methods parties use to contact voters. In the place of a field experiment, they could use a multivariate regression set-up where

voting is regressed on different forms of reported contact and some schedule of controls. Contact, however, is not random. Rather, it is a function of both the likelihood of receiving the contact, i.e. answering your phone or door or opening your mail, and the likelihood of a party mobilizing a given voter. This is the unobservable bias which Gerber, Green and Kaplan (2004) so strongly lament.

To overcome this bias they conducted a field experiment with three different treatment types (canvass, phone, and direct mail). Working in a large city, they were able to vary the number and timing of mailers, and what the mailers said. After controlling for the likelihood of receiving the treatment (a subject of much debate between Imai (2005) and Gerber and Green (2005)), they calculated average treatment effects. Then, using data on whether individuals were contacted and official voting records, they tested the comparative effects of different treatments in a multivariate set-up. They find that face-to-face contact and mailings both increase turnout. Telephone contact has a slightly negative effect. These results leave them with strong evidence about the comparative efficacy of different treatment regimes.

1.3 An Eye Out for Inferential Monsters: The Case for Experimentation in Political Science

Suppose two political scientists wish to properly answer an empirical question and they choose to take two different approaches. The first conducts

an experiment in which she intervenes in a data-generating process, say by sending letters to voters informing them of candidates' positions. The second decides to survey voters during and after the election and ask them, among other things, about their knowledge of the candidates' positions and for whom they intend to vote. Suppose further that the two approaches differ only in this regard. Prior to data collection, the political scientists would think up a research question, perhaps review prior work in the field, and adopt or develop theory. And subsequent to these approaches, the two political scientists would likely go through the same steps. They would analyze their data, compare them to expectations, draw conclusions, and write up their results. With so much shared between these two approaches, what makes them differ so fundamentally? Put differently, why would advocates of experimentation see the first as superior to the second?

The difference, an experimentalist would reply, is that "Experiments facilitate causal inference through the transparency and content of experimental procedures, most notably the random assignment of observations to control and treatment groups (Druckman et al., 2006, 627)" This is an inference which is, in the majority of cases, probably unavailable to those who merely observe the world.

The reason for this distinction is found in my definition of experimentation. When a researcher lets a random process assign one subject to a treatment and another to control, then the researcher knows (with measured uncertainty) both the size of an effect and its cause. By contrast, the

observational political scientist must go to much greater lengths to make sound inference, relying on modeling assumptions and statistics. And, as Blalock (1979) puts it, “statistics is a poor person’s substitute for contrived laboratory experiments in which all important relevant variables have been controlled” (6). The problem for the observationalist is this: while she can know which respondents knew candidates positions and which did not, she will be hard-pressed to know why they know those positions, and to know that the same thing which caused them to know those positions is not the same thing which causes them to vote as they do. This problem of *unobserved heterogeneity* is fundamental.

Edward Leamer (1983) thoroughly captures this distinction, and he is worth quoting at length:

The truly sharp distinction between inference from experimental and inference from non-experimental data is that experimental inference sensibly admits a conventional horizon in a critical dimension, namely the choice of explanatory variables. If fertilizer is randomly assigned to plots of land, it is conventional to restrict attention to the relationship between yield and fertilizer, and to proceed as if the model were perfectly specified... In contrast, it would be foolhardy to adopt such a limited horizon with non-experimental data. But if you decide to include light level in your horizon, then why not rainfall; and if rainfall, then why not temperature; and if temperature, then why not soil depth, and if not

soil depth, then why not the soil grade; ad infinitum. Though this list is never ending, it can be made so long that a nonexperimental researcher can feel as comfortable as an experimental researcher that the risk of having his findings upset by an extension of the horizon is very low... *Still, the horizon within which we all do our statistical analyses has to be ultimately troublesome, since there is no formal way to know what inferential monsters lurk beyond our immediate field of vision* (39).

Leamer, in this eloquent argument, explicitly resists going as far as saying we can learn nothing from strictly observational research. Gerber, Green and Kaplan (2004), who make a more forceful case, are not so willing. But their point is much the same: Statistical uncertainty is generated by modeling assumptions, such that it is extremely difficult to know which modeling assumptions are correct. Observational research is plagued by our inability to identify sources of bias, and thus estimate true causal effects. This devastating shortcoming of observational as opposed to experimental research is not easily remedied, even with the addition of new cases. The principal solution, then, is not in larger samples or better estimators. Rather, it is in using true random assignment, or in the case of quasi-experiments choosing cases where other plausible explanations can be ruled out. Another manner is to identify “instances where the processes by which the independent variable (is) generated have no plausible link to unobserved factors that affect the dependent variable” (Gerber, Green and Kaplan, 2004, 23-24). Clearly,

these cases are few. The limits of observational research are thus apparent.

In theoretical terms, then, there can be little basis for the argument that experiments are inferior to observational research. And, to be fair, few political scientists would argue that they are. The reason for a dearth of experiments is probably thus a result of questions of practicality and taste. Experiments may be perceived as too costly or too impractical and difficult to implement (e.g. Green and Gerber, 2002). And in some cases, such as the effects of war, gender quotas, or material deprivation, it would be unethical (if not practically impossible) to conduct experiments. However, as I demonstrate through my own examples, experiments can be easily implemented for many questions of interest. Moreover, this can be done at a cost which does not exceed that attached to observational approaches. Experiments may also be perceived as lacking in external validity, particularly if political scientists believe they necessitate a retreat to the laboratory. In a subsequent section, however, I show that this is not the case. We can conduct experiments in a number of non-laboratory venues without fundamentally jeopardizing validity.

1.4 The Quantity of Experimentation in Political Science

To measure the quantity of experimentation in political science, I use data collected by Krueger and Lewis-Beck (2008, hereafter KLB) . They analyzed

all 1756 “scientific research papers”² published in the *American Political Science Review*, the *American Journal of Political Science*, and the *Journal of Politics* in the years 1990-2005. These journals may not be a representative sample of all political science – indeed, on measures of quality they certainly are not. However, they do represent the three flagship journals of the discipline, and thus seem a reasonable choice. What is the quantity and nature of experimentation within these journals? I first consider the overall quantity and growth of experimentation, and then turn to more precise questions about how experimental articles appear to differ from their non-experimental counterparts.

Considering all articles published in the three leading journals of the discipline between 1990 and 2005, just 5.5% (97) are classified as experimental. Moreover, using KLB’s classification of 17 subjects, we find this is heavily concentrated in three subjects: political behaviour (51.5% of experimental articles), political psychology (17.5%), and minority politics (16.5%). By these lights, experimentation is neither common overall or widespread. What of its growth?

Figure 1.1 suggests that the use of experimentation in political science has increased modestly overall but much less modestly within the subfields of political behaviour and political psychology. Considering the mean share of experimental articles before and after 1999, we find that 4.8% articles are

²KLB exclude “articles dealing with methods, theory, up-dates, exchanges, communications, workshops, or symposia.”

classified as experimental including and before 1999. This rises to 6.8% after 1999 ($t = 1.87, p = .04$ (one-tailed)). The growth is more apparent if we consider the subfields of this dissertation, political behaviour and political psychology: while 11.7% of articles were experimental prior to and including 1999, this share almost doubles to 21.6% after 1999. This is a clear increase ($t = 3.10, p = .00$ (one-tailed)). By the numbers, then, the reclamation of an experimental tradition in political science is happening, if slowly and within some subfields more than others.

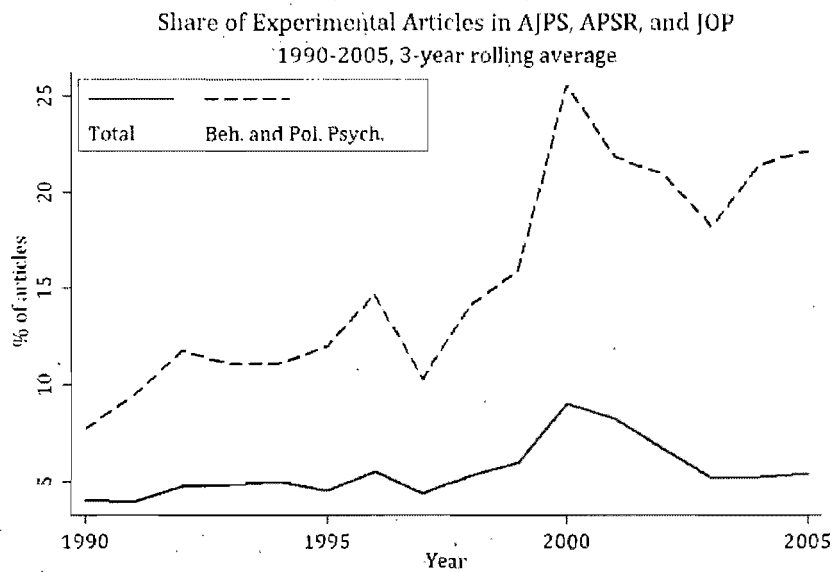


Figure 1.1: Share of Experimental Articles in *AJPS*, *APSR*, and *JOP*

1.4.1 The Nature of Experimentation in Political Science

Not only is experimentation growing generally, and within political behaviour and psychology specifically, but it is doing so with a different approach and influence than observational approaches. This difference accords to statistical approach and collaboration. First, experimental articles require less complex statistical analysis than observational articles. This is as we would expect, as the random assignment of treatment necessitates less statistical control and correction than observational research.

KLB identify a number of different statistical approaches for the analysis of data. Moving from less to more advanced, they are: tests of differences, correlations, analysis of variance (ANOVA), ordinary least squares, logit/probit, time series, maximum likelihood estimation, scaling techniques, and latent variables. Table 1.1 demonstrates the frequency of these techniques for observational and experimental papers³

On average, experimental articles appear to use less complex statistical methods than their observational counterparts. Indeed, experimental articles are more likely to use tests of difference, analysis of variance, and ordinary least squares. By contrast, they use logit and probit and other maximum likelihood techniques less frequently, and never use time-series or latent variable techniques. If we assume that the requirement of peer-review ensures

³As a single paper can have multiple techniques, these frequencies do not sum to 100% in any column.

that appropriate techniques are used in each article, then we can conclude that experimental articles require less complex statistical techniques to draw their conclusions. This arguably recommends experimentation, as neither the reader nor the analyst relies on complicated statistical techniques to draw conclusions. Whatever validity concerns are presented by statistical inference theoretically, in practice experimentation appears statistically simpler than observational research.

Table 1.1: Frequency of Statistical Techniques in Observational and Experimental Articles

Technique	Observational (%)	Experimental (%)	Difference (%)
Tests of Difference	5.1	15.5	10.4
Correlations	13.8	7.2	-6.6
ANOVA	1.0	24.7	23.8
OLS	41.4	46.4	5.0
Logit\Probit	32.5	21.6	-10.8
Time-series	5.2	0.0	-5.2
MLE	4.7	3.1	-1.6
Scaling techniques	1.8	2.1	0.3
Latent variables	0.5	0.0	-0.5

Second, experimentation appears more collaborative than observational research. Considering all the articles in KLB's dataset, the average number of authors per observational article is 1.86. The corresponding average for experimental articles is 2.30, a significant difference ($t = 5.12, p = .00$ (one-tailed)). Within political science, at least, experimentation appears to be a more collaborative approach than observational research. This could be a result of the increased operational requirements of running an experiment (e.g. using students as supervisors during an experiment and then including

them as co-authors). Alternately, it could be a result of experimental work requiring collaboration across disciplines and thus increasing the number of authors. Or, more likely, it could be a combination of both. Discerning the cause is beyond the scope of this introduction, but it should be said that those who enjoy the process and camaraderie of collaboration would be well-served to engage in experimentation. Indeed, the works that follow suggest that such was the case for me.

Third, experimentation appears more influential than observational research. While KLB do not include citation counts for each article, Druckman et al. (2006) find that experimental articles in the *American Political Science Review* are more frequently cited than their observational counterparts. Depending on the unit of comparison, they find an increase in the likelihood of an article being cited if it is experimental between 26% and 74%. What generates this advantage is unclear, but we can tentatively conclude that experimental articles are, all else equal, given more weight than their observational equivalents. As I argued in the previous section, this is just as it ought to be.

Taken together, these facts lead to the conclusion that experimental political science is appreciably different than the observational research. Experimentation is likely to be clustered within a smaller number of subfields, particularly political behaviour and political psychology; it is likely to be collaborative and statistically straightforward; and it is likely to have a greater scholarly impact. The fundamental values which guide research in politi-

cal science are not easily agreed upon, and far be it from me to suggest what another should esteem. As a question of taste, then, I believe these differences recommend experimentation to researchers in political behaviour and political psychology, particularly those who prize clear causal inference and who shun toiling alone. As will be seen in the papers that follow, my work generally conforms to these standards. (I should hope it also has a disproportionate scholarly impact).

1.5 The Qualities of Experimentation in Political Science

Thus far, I have demonstrated that experiments of several types are put to use in political science; and I have argued that experiments of all types have generated valuable empirical insights. But two questions remain. First, how exactly do we evaluate the quality of experiments? Second, are these qualities equally shared across experimental types, or is it the case that laboratory experiments are theoretically superior to other types of experiments? The implication of this second question is important for the argument of increased experimentation in political science: if laboratory experiments are the gold standard (as McDermott, 2002, would argue) and other types are theoretically inferior, then there is less of an argument for their increased use. However, if it can be demonstrated that there is no theoretical reason why other types of experiments are inferior, then an argument for their increased use remains.

It seems uncontroversial that the best criteria on which to evaluate individual experiments is their validity or their ability to make claims which other scientists regard as true or correct. It stands to reason, then, that we should evaluate experimental types with some measure of average validity. This is impossible short of doing a representative survey of experiments within types. However, it is possible to consider comparative validity between stylized versions of each experimental type. I do this using the four-fold conception of experimental validity put forward by Shadish, Cook and Campbell (2000). As MacCoun (2003) has argued, this work represents the definitive statement of a Campbell's account of validity. For the generations of social scientists who have looked to his wisdom, this posthumous work represents its pinnacle.

Political scientists tend to think of validity in terms of internal and external validity, both in the context of experiments, but also in the context of political science generally (Ferejohn, 2004). However, as Morton and Williams (2006) argue, this is not a view which has kept pace with theoretical and conceptual developments. *Au contraire*, Campbell set aside his own account of internal and external validity (Campbell and Stanley, 1966) for the refinement of Shadish, Cook and Campbell (2000).

Shadish et al's conception of validity takes four parts:

1. Statistical conclusion validity, which they define as "the validity of inferences about the correlation (covariation) between treatment and outcome."

2. Internal validity, which they take to be the “validity of inferences about whether observed covariation between A (the presumed treatment) and B (the presumed outcome) reflect a causal relationship from A to B as those variables were manipulated or measured.”
3. Construct validity, described as “the validity of inferences about the higher order constructs that represent the sampling particulars”; and,
4. External validity, set out as the “validity of inferences about whether the cause-effect relationship holds over variation in persons, settings, treatment variables, and measurement variables.” (38).

These definitions represent a substantial clarification of the internal and external validity approach, and allow for a more fulsome consideration of not only experimental designs, but also experimental types. For each of these validity concepts, they propose a series of threats to validity. In what follows, I briefly expand on each concept and make an argument about why each experimental type is beset by each set of considerations. On balance, it is not clear that one experimental type is consistently superior to others. The consequence of this is that political scientists are left with a variety of types when choosing to answer an empirical question through experimentation. They are not constrained to the laboratory or forced into the field. Rather, the importance of a particular result in any experiment will be a function of the skill of the researcher in assessing the various types of experimentation and choosing appropriately.

1.5.1 Statistical conclusion validity

While Shadish, Cook and Campbell articulate nine different threats to statistical conclusion validity, these can be categorized into two broad sets. First, those concerned with the proper application of statistical techniques. Second, those concerned with statistical power. On the criteria of proper statistical technique, it is not clear to me that any experimental type is free of the threat of a researcher choosing the wrong tests, engaging in a “star search”, or not fully accounting for the properties of her estimator. As such, these concerns do not help us adjudicate between experimental types.

The problem of statistical power is one of research design, and can be guarded against by carefully measuring outcomes, by controlling for heterogeneity in a sample, and by considering the balance between the proportions of control and treatment (Nickerson, 2005). For the purposes of considering experimental types, what matters is whether a type is flexible to adjustment. This, in turn, is a function of two things: the number of iterations of the study (i.e. how many times will the researcher repeat the experiment, with sufficient time in between to make adjustments); and the ability of the researcher to adjust the treatment/control balance midstream. On the first consideration, laboratory and games-based examples seem to have an advantage, as an experiment typically involves several iterations. Conversely, a survey experiment has more difficulty (especially if it is in the form of a poll), as it is often prohibitively costly to rerun a poll on a proper general population sample. Field experiments are more likely to involve one large-scale iteration.

Accordingly, they can do little to guard against the post-hoc discovery of low statistical power. However, careful planning in the implementation stage can avoid this problem.

On the second measure - the ability to change treatment/control balance midstream - survey experiments probably have an advantage, given the large number of respondents on average. The ability of games-based and laboratory experiments to change the balance is more a function of the complexity and the temporal administration of the treatment. Depending on the design of the experiment, the treatment/control balance could be adjusted midstream. Field experiments, on the other hand, probably face more difficulty of mid-stream adjustment, as they typically rely on large physical undertakings (such as canvassing in the case of the Gerber and Green field experiment) or the use of confederates (such as cooperation with political parties in Wantchekon (2003)'s field experiments in Benin). As such, timely adjustment of treatment/control may be, on average, beyond the abilities of the researcher.

None of the preceding threats to statistical validity are impossible to overcome in any experimental type, especially given a good research design. However, some types are more likely to face problems than others, specifically field experiments. This does not confine them to the dustbin, but it does highlight the potential challenges they face.

1.5.2 Internal validity

To begin, it should be noted that Shadish, Cook and Campbell (2000) do not see internal validity on the whole being any more or less likely in different types of studies. As they argue,

enough experience (with the list of threats) has accumulated to suggest that it applies to any descriptive molar causal inference, whether generated from experiments, correlational studies, observational studies or case studies. After all, validity is not a property of a method, it is a characteristic about knowledge claims... about causal knowledge (54).

The challenge of internal validity is to ensure that the experiment at hand produces causal effects as a result of the treatment, and not as a result of its administration or its measurement. It should be plain that this is a threat to each experimental type. That said, we can probably identify special cases where some experiment type will have a better average ability to avoid a given threat. It is unclear that any experimental type would have an advantage over another in avoiding ambiguous temporal precedence, as this plagues all research which is not carefully designed. It is possible, however, to believe that field experiments could have a higher probability of facing selection effects than other types, but only to the degree that randomization procedures are not followed and to the degree that treatment may coincide with different population groups under study. However, when properly implemented, “ran-

dom assignment definitionally eliminates selection bias” (Shadish, Cook and Campbell, 2000, 56). Field experiments face similar problems with history, as they cannot isolate respondents from outside events, as can be done in a laboratory, a games experiment, or a survey.

Conversely, all experiment types can suffer from maturation or attrition as long as they undertake measurement over some extended period of time (with the definition of extended depending on the question and instrument at hand). Similarly, every experiment type can be threatened by regression artefacts if selection is conditioned on some extreme characteristics of respondents. It is no more or less likely in any experimental type. Likewise with instrumentation, which can apply to all experimental types. In sum, as internal validity demands equal precision in each experimental type, each is susceptible to its threats, though field experiments may experience a slight disadvantage.

1.5.3 Construct Validity

Construct validity is defined as the “validity of inferences about the higher order constructs that represent the sampling particulars” (Shadish, Cook and Campbell, 2000, 64). Put slightly differently, construct validity is about the concurrence between the theoretical properties which we wish to assess and the operationalization of these measures. The authors provide an example of highly valid constructs: in measuring childhood health, growth in height and lung capacity are clearly valid constructs, because they are direct measures

of health. However, developing constructs for psychological development is more difficult, because there are no natural units of measurement, as with inches in measuring height. Rather, the researcher has to rely on constructs such as vocabulary or quantitative reasoning, which are subject to debate (Shadish, Cook and Campbell, 2000, 64–65).

There are three points of particular interest to be made. First, construct validity is important not only for questions of measurement, but even more so for questions of causal generalization, particularly for the representation of a causal process. Thus, construct validity answers the question of whether we can “generalize from a sample of instances and the data patterns associated with them to the particular target constructs they represent” (Shadish, Cook and Campbell, 2000, 21). Second, we are interested not only in the constructs of our treatments, but also of our units, settings, and outcomes. For example, if we want to study the effects of some job training program on the unemployed, then we need a construct for the unemployed. This could be anyone who is out of work, or it could be those who have been out of work for a period of time. The quality of this construct is determined by the kind of unemployed we are seeking to study. Similarly, while we are theoretically interested in what “helps” the unemployed, there are several measurable outcomes which could stand in for this. Defining which are best is the challenge of construct validity. Third, overarching all discussion of construct validity is the strong claim that research cannot be conducted without constructs. The periodic table is a construct, as is an atom (Shadish, Cook and Campbell,

2000, 65). Because there is rarely if ever a direct translation from theory to some empirical unit, then we have to be constantly concerned with construct validity.

The degree of difficulty in establishing construct validity is not the same for each experimental type. In each case, researchers have to justify their constructs given their theory and their experiment, and then have to guard against no less than thirteen threats to construct validity. The degree to which they can be guarded against varies by experimental type.

Shadish et al. identify two different types of threats. The first relates to the definition of constructs. It should be obvious that this applies with equal felicity to each type of experimentation. The second set of threats relates to reactions to treatments, by both experimenters and subjects. In this regard, threats are most likely when subjects are aware of their participation in an experiment specifically or a scientific study more broadly. In this regard, field experiments hold a clear advantage as they are the only type of experimentation in which subjects can be made completely unaware that they are participating in study.

As a whole, these threats to construct validity potentially plague all types of experiments, but perhaps they threaten field experimentation marginally less. Most importantly, these problems threaten pure observational research at least as much, as the ability to address them through design and control is often absent. As such, observational research faces these threats as or more directly than any type of experiment.

1.5.4 External Validity

Shadish et al's definition of external validity is the least revised of the original Campbell and Stanley (1966) concepts. They describe external validity as "the extent to which a causal relationship holds over variations in persons, settings, treatments, and outcomes" (Shadish, Cook and Campbell, 2000, 83). There are two important clarifications to be made to this definition. First, external validity does not refer exclusively to whether an experiment applies in the "real world". Rather, it also refers to the degree to which a causal claim made in an experiment will hold in another experiment, given some change in subjects or setting. Externality, then, is to the experiment, not the laboratory. Second, the questions of generalization inherent in establishing external validity are not only about to how broad a population an inference can be applied. As they note, generalization can occur in at least five directions:

1. Narrow to broad. By their example, using an experiment in income maintenance in one state to guide national policy.
2. Broad to narrow. For example, an individual who wishes to lose weight may ask if successful diet results from a study would apply to him individually, given his body type and exercise habits.
3. At a similar level. Applying the income maintenance experiment in a state of the same size as the experimental state would be an example.

4. To a similar or different kind. Applications to a different kind would occur if the experimental income maintenance subjects were males, and the program was adopted for females.
5. Random sample to population members. This is inference from an experiment with a random sample selection procedure to other members of the sampled population (Shadish, Cook and Campbell, 2000, 83–84).

Taking these two observations together gives us a much more fulsome view of external validity; one which draws it closely to concerns with fecundity. External validity thus becomes a question of how robustly a causal conclusion can take root and bear fruit in other contextual soils and environments.

Whereas threats to other validity types are quite diverse, the threats to external validity are common in their focus on the conditionality of results. Given that an experimental finding is the result of a proper statistical estimation, a good construct mapping, and an internally valid research design, external validity then asks if the findings are conditioned on the subjects and objects of the study, as well as other confounding factors. As such, these threats can apply equally to each type of experiment. The solution, they suggest, is random selection of subjects from the population of interest. Though they argue that this is rarely feasible in experiments, they do recommend it on the basis that “random sampling eliminates possible interactions between the causal relationship and the class of persons who are studied versus the class of persons who are not studied within the same population” (Shadish,

Cook and Campbell, 2000, 91). Given this admonishment, the superiority of survey experiments becomes apparent, provided they are performed on randomly drawn samples. Similarly, the ability of field experiments to randomly assign treatments over a randomly drawn population (or an entire population) similarly sets them apart from the typical laboratory or games-based experiment, which most typically occurs on self-selected populations. It is on these grounds, and not on the grounds that they occur “in the real world”, that survey and field experiments are superior to their counterparts on the metric of external validity.

1.5.5 Comparing the types

As the above survey of validity threats shows, laboratory experiments are by no means a clear front-runner in the race for experimental supremacy. While they are probably best for ensuring internal validity and statistical validity, they are probably not better than survey and field experiments at ensuring external validity (and not just for the conventional reason that these experiments occur in the “real world”.) Moreover, each can be equally beset by poor construct validity. The conclusion to be drawn is that the various types of proper experimentation are probably on equal footing overall in the face of validity concerns. Which is best for answering a causal question is then a function of the question, the research design possibilities at hand, and prior approaches. In the next section, I review the works which follow as individual articles. In the case of each, I articulate the question which they

seek to answer, and briefly justify the use of an experiment.

1.6 The Articles

The six articles that follow all fall into the broad categories of political behaviour and political psychology. That is, each asks questions about why and how individuals participate in politics and how they form their preferences. Perhaps the product of eclecticism, the papers otherwise vary widely in their content. I ask, for example, what role altruism plays in preferences for greater public spending. I fashion a model of voter turnout which relies on other-regarding preferences and social identification with political parties and then test this model empirically. In both the first and second paper, games-based experiments are used to reveal the preferences of actors. These revealed preferences are then shown to drive their opinions and actions. In a third paper, a field experiment examines whether direct mail is effective in changing the opinions of elite voters. In my fourth paper, I demonstrate a statistical model for measuring the power of arguments in a survey experiment. I next ask whether partisans in Canada differ in their levels of altruism, as revealed in a games-based experiment. And in a final paper, I use a field experiment to examine whether compulsory voting increases voter knowledge and engagement. This is a diverse lot. But good inference comes from variance, and the inference which is (hopefully) drawn from these papers is that a wide variety of experiments can be applied to a diverse range

of questions.

1.6.1 Dictators and Purses: Altruism and Support for Public Spending

How do individuals determine their support for new public spending programs? Is it purely a matter of self-interest, or do some citizens prefer greater public spending because they are concerned for others and wish to see their hardship relieved? In other words, is there an element of altruism in support for greater public spending? Previous studies have demonstrated a connection between non-self-interested measures and preferences on spending and taxation (e.g. Rasinski and Rosenbaum, 1987; Shiell and Seymour, 2002; Hudson and Jones, 1994; Fong, 2001). However, these studies have all relied either on verbal expressions of concern for others or an inferred degree of altruism through a series of equations.

My own contribution is to measure an individual's inclination towards altruism through their behaviour in a "dictator game" (Camerer, 2003). In this game, individuals in a large online survey were given a chance to win money and were asked how much, were they to win, they would share with a completely anonymous individual. I argue that those who give more can be said to be more altruistic. I then ask subjects about their preferences for a number of new public spending programs, in each case increasing or decreasing the cost of the program to the respondent. I find a strong and robust connection between a subject's revealed altruism and their willingness

to support a new public spending program, even at a cost to themselves.

1.6.2 Antipathy, Affinity, and Political Participation: How Our Concern for Other Partisans Makes Us Vote

Political scientists have expended significant effort to explain the paradox of voting (for the definitive review, see Blais, 2000). Why, they ask, if the probability of a single citizen's vote deciding an election is so low do we still observe such large numbers of voters? This paper proffers an answer at this question by combining two arguments. First, citizens view elections as a contest between different groups of people in which scarce resources are divided after the election. Second, some citizens care about some groups of citizens more than others, that is, they have affinity and antipathy towards various partisans. Considered together, these two arguments furnish a logic for why some citizens vote.

I demonstrate this logic through a formal, decision theoretic model. Then, using a series of dictator game experiments embedded in a large online survey, I induce respondents to reveal their preferences for some partisans over others. I then show that these revealed preferences significantly predict the decision to vote and do so independently of other well-known predictors, including partisan identification, education, age, gender, and media usage.

As with the previous paper, this study employs games-based experiments as a measurement tool to assess the preferences of individuals. By observing behaviour in a controlled environment where assignment to each condition is

under my control, I am able to directly observe individuals' preferences for some partisans over others in a more valid and reliable manner than with traditional survey questions.

1.6.3 For Want of a Nail: Direct Mail and Negative Persuasion in a Leadership Race (with Daniel Rubenson)

Direct mail is a pervasive feature of modern political campaigns. This is perhaps especially true in leadership races in which candidates for the head of a party compete to win the support of the party's elites and active members. In this paper, Rubenson and I ask two questions. First, does direct mail work at persuading these voters to increase their evaluations of and preferences for the candidate sending the mail? Second, can a leadership candidate who is outside the mainstream of his party persuade voters to come to his side by highlighting the policies which place him outside the mainstream.

To answer these two questions, we conducted a field experiment in the fall of 2006 with the Michael Ignatieff campaign for the leadership of the Liberal Party of Canada. Our experiment consisted of randomly assigning a direct mailing to delegates who were pledged to other candidates. We then surveyed these delegates using a survey instrument under the cover of a university. Our findings are rather remarkable. Not only was Ignatieff unable to persuade delegates to increase their positive evaluations of him and to move him up in their preference rankings. Quite to the contrary, those who received mail appeared to adjust their evaluations of other candidates

upwards and to move Ignatieff down in their preference rankings.

The use of a field experiment to answer this question marks two important innovations. To the best of my knowledge, this was the first field experiment conducted in cooperation with a political campaign in Canadian political science. Second, it allowed for much clearer inference than if we had only conducted a survey. Indeed, we find that survey recall of mail reception is remarkably poor. Fifteen percent of those who did not receive mail from the campaign claimed that they had, while only two-thirds of those who received mail recalled this correctly. Inferences based on survey estimates alone would be grossly incorrect. By randomly assigning mail, we can, in Leamer's terms, effectively limit the horizon of possible explanations for the negative effect we observe.

1.6.4 Testing the Power of Arguments with a Bradley-Terry Model (with Daniel Rubenson and Arthur Spirling)

Public opinion scholars and practitioners are often interested in knowing which arguments individuals find convincing and which they find unpersuasive. This is especially true in trying to explain why citizens take the positions they do on referendum questions. In this paper, we demonstrate how a statistical model typically used in biology and other sciences, the Bradley-Terry model, can be used to analyze the persuasive power of arguments in a survey experiment. The benefits of the Bradley-Terry model are that it can tell us not only whether one argument is more convincing than another, but

what components or features of an argument make it more compelling than another.

To demonstrate this, we conducted a survey experiment during the October 2007 referendum on electoral reform in the province of Ontario. Using a subject group of 520 individuals, we assigned each individual one of six arguments for electoral reform and one of six arguments against. We then asked them whether or not they supported electoral reform. Using the model, we show that arguments for the existing system enjoyed a general advantage. We then measure which components made arguments more or less persuasive. We find that arguments which appeal to fairness and local representation are significantly more persuasive, while arguments which mention political parties are less persuasive. Finally, we find a constant advantage for all arguments in favour of the status quo. We use these results to interpret the outcome of the Ontario referendum. We conclude by highlighting further possible applications of the Bradley-Terry model.

The principal contribution of this paper is not insight into electoral reform, but the introduction of a model which makes survey experimentation potentially much more efficient. With just 520 respondents and thirty-six treatment groups, we are still able to uncover information about the persuasive power of arguments. The Bradley-Terry set-up, then, has a unique value proposition. It allows survey researchers to learn much with a small number of subjects and/or large number of treatment groups.

1.6.5 Partisanship and Altruism: Results from a Dictator Game Experiment (with Angelo Elias)

Partisan identifiers differ from non-identifiers across several dimensions. We know, for example, that they are more likely to vote, to pay attention to politics, and to hold consistent positions and opinions. We also know that differences exist between partisans, for example, in their issue positions. Elias and I ask if partisans differ in a more fundamental and basic way. To wit: do partisans in Canada differ in their levels of altruism?

To test this proposition, we examine the behaviour of partisans in the four dictator game experiments described in my first two papers. We compare the differences between partisans in their allocations to co-partisans, other partisans, and anonymous individuals. We find that all partisans consistently allocate the most to co-partisans and the least to other partisans. Allocations to anonymous individuals are in the middle. We also find that New Democratic partisans are more altruistic on average than Conservative or Liberal partisans.

Our findings give us an important insight into a basic difference between Canada's partisans. In doing so, it raises a very interesting question about whether New Democratic partisanship leads individuals to be more altruistic, or whether the reverse is true. These findings were made possible by having subjects play a dictator games in a controlled environment in which they were induced to reveal their preferences. Arguably, there is no observational equivalent to this experiment, because partisans are never asked to demon-

strate (and to pay to demonstrate) their preferences for some partisans over others in the real world. Only through a games-based experiment are we able to reveal this important and fundamental difference in Canadian partisans.

1.6.6 Does Compulsory Voting Lead to More Informed and Engaged Citizens: An experimental test (with Henry Milner and Bruce M. Hicks)

The “first-order effects” of compulsory voting laws are clear. From both cross-sectional (Jackman, 1987; Blais and Carty, 1990; Blais and Dobrzynska, 1998; Franklin, 1996, 2004) and quasi-experimental (Hirczy, 1994) accounts, political scientists conclude that compulsory voting laws increase voter turnout. However, whether they generate “second-order effects” such as greater citizen knowledge and engagement in politics remains unanswered, despite the claims of advocates.

In this paper, we argue that we cannot definitively identify second-order effects of compulsory voting using existing cross-national surveys. To address this gap in our knowledge, we conducted a field experiment among voting-aged students at a Montreal CEGEP (or junior college). Our intervention involved paying some students to complete two surveys while paying another group of students to complete two surveys and vote in a provincial election. The key difference between the two groups, then, is that one faces a financial incentive to vote. Any differences in knowledge, news consumption, and political discussion are thus the result of the compulsory voting treatment. We find little evidence of such effects.

As in prior papers, the value of experimentation is the isolation of the effects of compulsory voting with measured uncertainty. Had we taken the approach of measuring differences in political knowledge and engagement between countries with and without compulsory voting we would come up against two fundamental challenges. First, we would have the thorny task of establishing the equivalence between knowledge measures in the countries. Second, we would need to account for any other differences between the countries which may explain both the decision to adopt compulsory voting and increased levels of knowledge. As we are without strong theory and as these differences are largely unknown, then we could make little progress with an observational approach. In the face of this, we take an experimental approach and draw much clearer inferences.

1.7 A Closing Word

I turn now to the presentation of these articles. After this, I conclude the dissertation with a discussion of the results presented and their implications for political behaviour and political psychology. I also discuss their implications for my own research in the future.

Chapitre 2

Dictators and Purses: Altruism and Support for Public Spending

Under review at Journal of Politics.

2.1 Introduction

How selfish soever man may be supposed, there are evidently some principles in his nature, which interest him in the fortunes of others, and render their happiness necessary to him, though he derives nothing from it, except the pleasure of seeing it. Of this kind is pity or compassion, the emotion we feel for the misery of others, when we either see it, or are made to conceive it in a very lively manner. That we often derive sorrow from the sorrows of others, is a matter of fact too obvious to require any instances to prove it; for this sentiment, like all the other original passions of human nature, is by no means confined to the virtuous or the humane, though they perhaps may feel it with the most exquisite sensibility. The greatest ruffian, the most hardened violator of the laws of society, is not altogether without it. - Adam Smith, The Theory of Moral Sentiments (1790 4)

Modern human society is characterized by a level of cooperation and coordination unmatched in history and in the rest of the animal world (e.g. Stevens and Hauser, 2004). We live within social units which, in a long view of history, are unimaginable for their size and peacefulness (e.g. Pinker, 2007). Undertakings which require an immense degree of coordination go on unnoticed for their complexity and difficulty. For scholars of many disciplines, the question of how such complexity and cooperation is initiated and maintained is central.

The provision of government services and the creation of new programs is one such example of complexity and coordination. Such programs see the monetary contributions of millions of individuals entrusted into the hands of government officials who then disburse money as instructed by politicians,

themselves elected and accountable to individual citizens in another act of great coordination. Millions of individuals rely on such programs for their sustenance and comfort, either regularly or in times of need. Indeed, government expenditures regularly make up a third or more of all economic activity in modern democracies (OECD, 2007). These programs, at least initially, rely on the support and willingness of individuals to pay for them (Pierson, 1996).

Despite the importance of these programs, we do not know enough about why citizens lend them the support so critical to their initiation and sustenance (Fong, 2001, 225). We know, for example, that spending policy responds to the preferences of citizens, moving up and down as citizens indicate a taste for more or less spending (e.g. Miller et al., 1999; McCrone and Kuklinski, 1979; Soroka and Wlezien, 2005). We also know that spending has a path-dependency such that once a program is in place it becomes difficult to abolish (e.g. Pierson, 1996). But what we do not much about is why some citizens prefer more public spending and others prefer less, especially when it relates to new public spending. Is it purely a matter of self-interest, or do some citizens support greater public spending because they are concerned for others and wish to see their pain and hardship alleviated through government action? In other words, does support for public spending depend on the benevolence of the butcher and the baker (Smith, 2007)?

I argue that there is a non-self-interested element which is crucial to support for greater public spending. In specific terms, I argue that there is a

clear altruistic motivation for supporting greater public spending on public programs. Some citizens have a greater regard for other individuals and this regard motivates them to do good for other people even when this choice comes at a cost to themselves. In the case of public policy programs, this means that individuals are willing to pay higher taxes for (new) programs even when they may not be expected to benefit, and they are willing to impose these taxes on other individuals. Accordingly, an explanation which relies solely on egoism is insufficient. Rather, a more fulsome explanation relies on the notion of *strong reciprocity* (Bowles and Gintis 2002). To demonstrate this, I show that individuals' underlying inclinations towards altruism – measured through a dictator game (Camerer, 2003) – are linked to their stated willingness to pay greater taxes for the provision of *new* public spending programs. While the importance of non-self-interested factors has been demonstrated in previous studies, this is the first study to rely on a behavioral measure of altruism rather than a stated preference or attitude. I thus establish a more conclusive link between altruism and support for greater public spending.

My demonstration proceeds in five stages. First, I give a definition and description of altruism and explain why we should expect it to be related to support for greater public spending. Second, I describe and justify the dictator game as a measure of altruism. Third, I describe the survey and subjects which are used to demonstrate the link between altruism and support for greater public spending. Fourth, I present my results. I then discuss

the implications of these findings and conclude.

2.2 Altruism and Public Spending

The importance of other-regarding orientations for the explanation of social and political behavior is becoming increasingly clear (e.g. Piliavin and Charng, 1990; Fowler, 2006; Fowler and Kam, 2007). Contra Mueller, it is not clear that “the only assumption essential to a descriptive and predictive science of human behavior is egoism” (Mueller (1986, 4) in Mansbridge (1990, 254)). Instead, economists, political scientists and sociologists alike are recognizing the analytical gains to be made from incorporating other-regarding orientations into our theoretical (e.g. Margolis, 1990) and empirical models (e.g. Frohlich and Oppenheimer, 1990; Fowler, 2006; Fowler and Kam, 2007; Fong, 2001). In this section, I define altruism. I then outline a theoretical argument as to why altruism should be related to support for greater public spending. Then, by reviewing previous work on the importance of other-regarding orientations in the formation of preferences over public spending and taxation, I identify my contribution, namely the linking of a *behavioral* measure of altruism to stated support for greater public spending.

2.2.1 Defining Altruism

In their review of altruism theory and research, Piliavin and Charng (1990) note that although definitions of altruism generally differ by discipline, this

substantial debate can be broken down into two divisions: sociobiological and psychological. The minimalist or sociobiological position (Rushton, 1982; Ridley and Dawkins, 2003; Aronfreed, 1980) defines altruism according to a behavior. As Sorrentino (1991) puts it “insofar as an act to benefit another organism is at the expense of a donor, it is altruism” (147). Or, as put by Margolis (1983) “what defines altruistic behavior is that the actor could have done better for himself had he chosen to ignore the effect of his choice on others” (quoted in Piliavin and Charng, 1990, 28). Or, as put by Piliavin and Charng (1990), we can “call behavior altruistic if it benefits the actor less than the recipient” (35). By these definitions, observing altruism requires only observing the costs and benefits of an action to an actor and a recipient.

A more strict definition, favoured by psychologists, involves intention. An altruistic act is one which the actor believes to be to the benefit of another and which does not confer extrinsic or intrinsic benefit to herself (Batson et al., 1978, 1979). The act requires a recognition of another’s suffering or need and an intentional act to reduce that suffering or fulfill that need. There is an immense amount of evidence for this psychological definition of altruism (Piliavin and Charng, 1990, 36-37), and thus my objective is not to contest this conception. Rather, my goal is to choose a conception of altruism which can be observed behaviorally and to then demonstrate its relationship to support for greater public spending. Accordingly, throughout this article I rest with a minimalist definition in which *altruism occurs when an individual undertakes an action which is to the material benefit of another*

at a material cost to herself. Such a definition has the benefit of not relying on the inference of intent. Rather, it relies only on observing an individual's actions at a given point in time. If we can demonstrate that this action did benefit someone else and that this came at a cost and could not be expected to incur a later net benefit, then we can say that the action is altruistic. Later, I show that such altruism can be measured in a "dictator game" from experimental economics.

2.2.2 Theoretical Link

Why should altruists support greater public spending? A convincing explanation relies on a theory of *strong reciprocity* (Bowles and Gintis, 2002; Fong, Bowles and Gintis, 2006). Following Fong, Bowles and Gintis (2006), strong reciprocators possess:

a propensity to cooperate and share with others similarly disposed, even at a personal cost, and a willingness to punish those who violate cooperative and other social norms, even when punishing is personally costly and cannot be expected to entail net personal gains in the future (1439).

Supporting greater public spending, even when it involves higher personal taxes, is directly analogous to strong reciprocity. First, strong reciprocators are willing to cooperate (through the paying of taxes) with those similarly disposed (other taxpayers). Moreover, they are willing to punish those who

do not contribute to the common pool resources (i.e. public goods and government spending) by imposing taxes on these individuals. They are willing to pay a cost to impose this punishment, namely a greater personal tax burden. As a preview of my results, I find exactly that. Those who display more altruistic behavior in dictator games are more willing to support new public spending programs, even when these programs impose a cost on others and themselves.

In arguing that strong reciprocity is fundamental for the provision of government spending, I add to the work of Bowles and Gintis (2002) and Fong, Bowles and Gintis (2006). Bowles and Gintis argue and demonstrate that such reciprocators are fundamental to the development of cooperation in non-kin based societies such as our own. Fong et. al. present evidence from behavioral economics experiments and survey research to show that behaviors and opinions are consistent with their theory. I go a step farther and link these two findings by embedding an economics game in a survey and using behavior in this game to predict preferences. And I do so with a clear theoretical prediction: individuals who display greater altruistic orientations in a dictator game should be willing to pay greater costs for the provision of public spending programs than those with less altruistic orientations. Even after recognizing that they may not benefit from a public policy and/or that their own benefit may be outstripped by the cost, altruists are still inclined to consider the benefits of public policies in terms of others. They are more likely to recognize how the needs of others will be served by a new policy and

to respond to those needs, even when it imposes a net cost on themselves. Accordingly, those with more altruistic orientations should be more likely to support new public spending, even when the cost of the policy increases for themselves and others.

2.2.3 Previous Findings

This is not the first study to demonstrate link between altruism or a concern for others and support for public spending, but it is the first to explicitly link game behavior with stated preferences. A review of previous studies on the role of non-self-interest factors makes this contribution more clear. For example, Shiell and Seymour (2002) examine preferences for public health insurance reform in Australia. Following Hudson and Jones (1994), they distinguish between self-interested and altruistic determinants of opinion by asking respondents which health care reforms they think would be best for them, and then which would be best for the whole population. Through a series of equations, they then estimate a coefficient for altruism. However, this requires using age, education level, and gender as proxies for altruism. They finally estimate that concern for others is about twice as influential as self-interest in determining opinion on health reform (Shiell and Seymour, 2002, 364).

Fong (2001) examines the role of self-interest and “social preferences” in generating demands for redistribution in the 1998 Gallup Poll Social Audit. She argues that some individuals have beliefs about justice and that they

care deeply about how other people are affected by redistributive policies. To test for the effects of self-interest, she uses a number of objective controls (income, education, marital status, etc) and subjective measures, such as concern over meeting family expenses and perceptions of future mobility. She then measures respondents' perceptions about whether individuals are in need as a result of their own actions or for some other reason. She finds that when people believe that poverty is not an individual's fault they are more likely to support redistributive programs. As with Shiell and Seymour, the marginal effects of self-interest do not outweigh those of redistributive preferences.

Rasinski and Rosenbaum (1987) examine the determinants of opinion on an increase in property taxes for public education in Illinois. They find that non-self-interest measures provide better explanations of citizens' opinions than self-interest measures. However, as with previous studies they rely on proxies both for self-interest (i.e. home ownership) and non-self-interest, such as views of teachers and the school board.

Also in an American context, Feldman and Steenbergen (2001) show that the prosocial orientation of humanitarianism is better than egalitarianism as an explanation of support for public spending in the US (659). Following (Staub, 1989, 50), they define prosocial orientations as consisting of "(a) a positive evaluation of human beings, (b) concern about their welfare and (c) feelings of personal responsibility for people's welfare." Like the previous studies, they use subjects' agreement or disagreement with various state-

ments to ascertain their prosocial orientations (663-666). However, contrary to Fong (2001), they do not find that the effects of humanitarianism are dependent on feelings about the deservingness of the poor. The larger point is that they show that prosocial orientations, of which altruism is clearly one, are effective for explaining support for public spending.

Corneo and Grüner (2002) examine preferences for government redistribution in twelve middle- and high-income countries. While they find that a self-interested explanation does explain significant variance in preferences for redistribution, “public values” – which is closest to my conception of altruism – and a “social rivalry effect” also explain significant variance. Moreover, these effects are consistently significant across all twelve countries, despite their different histories (post-communist and non-communist) and economic conjectures.

Finally, Sears and Funk (1990) review a large amount of evidence for the power of self-interest explanations over symbolic explanations, such as party identification, ideology, and racial tolerance for policy preferences, candidate choice, and opinions on matters of race. They cannot find a dominant place for self-interest: it plays little role in the assessment of racial issues, has a generally minimal and exceptional impact on economic issues and there is at best a modest relationship between vote choice and *personal* economic considerations¹ While Sears and Funk are not emphasizing altruism, their

¹This is, of course, in contrast to the more common phenomenon of sociotropic voting (Kinder and Kiewiet, 1981; Kiewiet, 1983; Lewis-Beck, 1988).

explanation for the dominance of symbolic politics is instructive for a more general account of why other-regarding behavior will matter in public opinion: “people may be socialized to respond to public issues in a principled and public-regarding manner. Perhaps political socialization teaches people to weigh most heavily the collective good when they don their political hats.” I agree entirely, except that I do not assume that other-regarding behavior is necessarily socialized. Rather, it could have much deeper roots (Rushton, 2004; Scourfield et al., 2004; de Quervain et al., 2004; Knafo et al., 2008; Cesarini et al., 2008).

2.2.4 Innovation

While these examples are not exhaustive of the literature on the role of non-self-interest factors in the formation of preferences for public spending, they are an informative survey for two reasons. First, they consistently demonstrate that non-self-interest factors matter for public opinion formation. Self-interest is never the whole or even the principal story. Second, they all rely on either proxies or verbal responses to questions as indicators of individuals’ altruism or concern for others. I take both of these observations to heart: first, I present below models of support for public spending which include a measure of altruism. Second, this measure does not rely on indirect measures or verbal responses. Rather, we *infer* a respondent’s level of altruism based on her *behavior* in a dictator game. In the next section, I describe these games and justify their use as an indicator of altruism. In the following sec-

tions I describe my data and then test a model of public spending preferences which includes this behavioral measure of altruism.

2.3 Altruism and Dictator Games

The use of dictator games (Forsythe et al., 1994; Camerer, 2003) is growing in political science. After first appearing in economics in the early 1980s, dictator games have been recently utilized in political science to measure such concepts as altruism, fairness, and social identification. Their results have been used to explain the decision to vote (Fowler, 2006), to participate in politics (Fowler and Kam, 2007), and to measure inter-ethnic trust (Whitt and Wilson, 2007).

The basic dictator game is as follows: Player 1 is given a sum of money or an opportunity to win money in a lottery. Player 1 is then given the opportunity to share none, some, or all of that money with Player 2. Player 2 has no opportunity to accept or reject the offer. In other words, Player 1 *dictates* the amount Player 2 receives. The game thus differs fundamentally from the more well-known “ultimatum game” in which Player 2 can accept or reject the offer (leading to a payout of zero for both players).²

Dictator game allocations to anonymous individuals provide a valid and reliable indicator of a subject’s level of altruism for at least two reasons. First, dictator games measure a subject’s willingness to improve the mate-

²In fact, in the strictest terms the dictator game is not a game at all, as it involves only one player’s choice and outcomes are not interdependent on the choices of other players.

rial welfare of another individual at a cost to self. Results thus conform to the definition of altruism specified above. Second, they measure *revealed* rather than *stated* altruism. Rather than attempting to infer the altruistic or pro-social orientations of a respondent through a series of survey questions about their behavior and invite the threat of costless deception and social desirability bias, the dictator game allows for an actual observation of their behavior.

As these measures are unconventional in political science and in public opinion research, one can anticipate several reasonable objections (see, e.g. Benz and Meier, 2008). For example, one could argue that subjects do not understand the game and are making allocations randomly. Or, one could argue that as subjects typically are playing for small stakes they are not making the decisions they would make were the stakes more substantial. Or, even if subjects are taking the game seriously, a laboratory environment does not replicate real life sufficiently to produce results which are consistent with a subject's day-to-day behavior and orientations.

Experimental economists have addressed these concerns. First, through a manipulation of the translation of allocations into payoffs, Andreoni and Miller (2002) have tested whether allocations are consistent with subjects' preferences. Their own research suggests that nearly all subjects (98%) behave consistently. In a similar manipulation across five different dictator games, Dawes and Fowler (2007) found 94% of their subjects behaved consistently. Subjects thus appear to understand the dictator game and respond

consistently and rationally when playing.

Second, while there has been some research (Cherry, Frykblom and Shogren, 2002) suggesting that subject behavior is different when playing with earned wealth rather than wealth granted by the experimenter, much research shows that as long as the stakes are real and not fictional, there is little difference in the distribution of allocations as stakes increase. In other words, subjects tend to give away the same share of a prize, regardless of its size (Camerer and Hogarth, 1999; Carpenter, Verhoogen and Burks, 2005). More importantly, the share of respondents which give away any money does not change markedly as stakes increase (Carpenter, Verhoogen and Burks, 2005). Accordingly, the degree to which a dictator game captures “other-regarding” behavior is not dependent on the size of the stakes.

Third, Benz and Meier (2008) have shown a correlation between behavior in a dictator game and altruistic behavior in real life, namely past and future giving to a university charity. Benz and Meier also highlight other noteworthy examples of the correlation between behavior in experimental settings and real world environments. For example, Karlan (2005) finds a positive relationship between trusting behavior in a trust game and repayment of microcredit loans among Peruvian subjects. In Annexe A, I provide further evidence of the criterion validity of dictator game behavior. As with any instrument in a survey, dictator games are neither perfect nor a cure-all. However, they are a more direct and robust measure of an individual’s altruistic orientation than a conventional battery of costless reports of behavior

or opinion.

My results section presents dictator game allocations and then estimates a relationship between these results and support for public spending. To anticipate my findings, I find that the dictator game allocations are similar in their distribution to earlier studies. And while I find a strong relationship between gender and allocations and a weak relationship between age and allocations, I find that allocations are independent of partisan identification, wealth, education, and other sociodemographics. This is in keeping with a view of altruism as prior to social demographics and is in keeping with a broader review of the dictator game literature (Camerer, 2003). Finally, when I move to model the relationship between behavior in the dictator game and support for public spending, I find a consistently robust relationship which suggests that more altruistic individuals support more public spending than their less altruistic counterparts.

2.4 Survey Design and Participants

The survey was conducted online. Subjects were required to login to the survey using a unique identification. This allows us to call up demographic information for those who have previously completed surveys. Those completing the survey for the first time were first asked a series of screening questions, including whether they voted in the most recent federal election and their partisan identification. Subjects then answered several questions

about recent news exposure, their attention to federal and provincial politics, and their views on federal and provincial politicians. Subjects then completed an unrelated eight-item module on empathy. They were then presented with the dictator game. Following this, they were presented with questions on their support for public spending. The survey concluded with questions related to charitable giving, the public service, and recent political events.

Altruism

Altruism was measured through a series of dictator games. Respondents were told that they were eligible to win up to four prizes of \$100 at the end of the survey in addition to a regular \$500 draw for survey participants. In the case of the first prize, they would be asked how much of it they would like to share with an anonymous individual, about whom they knew nothing.³ Complete instructions for the game are found in Annexe B.

The measure of altruism used in this paper is the amount of money which a respondent indicated she would give away to an entirely anonymous individual should she win the prize. This amount can range from 0 to 100, though it is rescaled for interpretive ease. In order to ensure that my results are robust to the specification of this independent variable, I operationalize altruism in

³For the other three prizes, they were similarly asked how much they would be willing to share with an anonymous individual about whom they knew nothing except which political party the respondent typically supported (Conservative, Liberal or NDP; in the case of Quebec residents, Conservative, Liberal or Bloc Quebecois). The presentation order of the anonymous recipients was randomized.

three ways. First, as a continuous variable from 0 to 1 (\$0 to \$100). Second, as a categorical variable reading 0 for those who give away nothing, 1 for those who give away an amount less than or equal to the median (\$20), and 2 for those who give away something greater than the median. Third, as a dummy variable which reads 1 for all individuals who give away any money in the dictator game and 0 for those who gave away nothing.

Willingness to pay for public goods

Respondents were also asked four questions about their willingness to pay for public services. In the case of each question, respondents were presented with a public policy proposal which was to the benefit of others but imposed a direct cost on the respondent. For two questions, the initial cost was high. If a respondent indicated an unwillingness to pay, they were then presented with lower costs. The question continued until they agreed to pay or the last category was reached. For the other two questions, the costs began low but increased with each indication of a willingness to pay. The questions were as follows:

- One proposed solution to fight climate change and decrease air pollution is to impose carbon taxes. Supporters of these environmental policies say such taxes would result in cleaner air and better health for everyone. Would you support carbon taxes if you knew it would cost you \$2000 (\$1500, \$1000, \$500, \$100) more per year to heat your home, ride the bus, and drive a car?

- Some politicians and policy groups propose making the first four years of university free for all qualified students, just like high school. This will result in greater accessibility to university education. Would you support the elimination of tuition fees if it cost you \$100 (\$250, \$500, \$1000, \$2000) more per year in taxes?
- Provincial health care programs often do not cover the cost of drugs for those with cancer. This can make fighting cancer financially taxing for cancer sufferers and their families. Would you support covering the cost of cancer drugs if you knew it would increase average emergency wait times for non-critical injuries (such as ear infections, the flu, or small cuts) by one hour (90 minutes, two hours, three hours, five hours, ten hours)?
- Wait times for many medical procedures (such as cataract surgery and hip and knee replacements) are currently longer than recommended by doctors. If tax dollars were guaranteed to go to these priority areas and to reduce wait times, would you be willing to pay \$2000 (\$1500, \$1000, \$500, \$100) more per year in taxes?

Other variables

Other standard survey variables were used to capture respondent demographics as well as party identification. Question wording for all variables is available in Annexe C.

2.4.1 Participants

Participants were drawn from a commercial online survey administered monthly by a national public opinion research firm in Canada. The firm uses a panel comprised of respondents recruited by the firm through email solicitations and respondents who self-select into the panel at the firm's website. Participants are invited by email to participate in one of four survey waves. The study was conducted in the final two weeks of May 2007. The effective sample size is 2648 respondents.

Compared to a university-based convenience sample, the online survey methods affords a large number of respondents and a relatively representative sample, particularly in regards to age, education and income. Table 2.1 shows the characteristics of the sample and their bivariate relationship to dictator game allocations, measured from 0-1. The average age of the panel was 50 years (*s.d.* = 19.6, *range* = 18,87). Males were slightly overrepresented in the panel (52.4%). The sample was also diverse on measures of income and education. Finally, the sample included a measurable number of partisans from each of Canada's four political parties. I address objections to my non-random sample in end of the results section, while noting that respondents in online samples have been demonstrated to mimic the behavior of those in more conventional telephone surveys (Best et al., 2005) while being less susceptible to social desirability biases (Taylor and Thomas, 2005).

Table 2.1: Sample demographic and political characteristics

Variable	% or Mean	Bivariate regression on Altruism	<i>p</i>	
Age	50.0	0.0424	.08	
French	15.3%	-0.0329	.01	
Female	47.6%	0.0278	.04	
Employed	62.1%	0.0110	.27	
Unemployed	3.5%	0.0102	.70	
Student/Homemaker/Retired	34.4%	-0.0129	.20	
Household Income	<\$40000	21.9%	-0.0136	.30
	\$40000 to \$60000	21.1%		
	\$60000 to \$80000	18.7%		
	>\$80000	38.2%		
Education	High School or less	13.1%	-0.0189	.16
	Some College	26.4%		
	Some University	60.5%		
Conservative ID	Weak	11.3%	0.0096	.57
	Strong	6.6%		
Liberal ID	Weak	16.9%	0.01084	.31
	Strong	4.7%		
NDP ID	Weak	7.1%	0.0242	.39
	Strong	2.9%		
BQ ID	Weak	2.5%	-0.0212	.43
	Strong	2.8%		
N=2648				

2.5 Results

Dictator game allocations in this survey resemble those in previous studies with “pay” (versus “non-pay”) conditions (see Fowler, 2006, 677). As Figure 1 shows, the distribution is tri-modal, with a plurality of respondents (approximately 40%) giving away no money, 30% giving away half of their potential winnings, and the next highest frequency (10%) giving away a quarter of their winnings. In comparison to findings in smaller university-based convenience samples, I do find some significant bivariate relationships between demographic variables and altruism, though not many (Table 2.1). Respondents who completed the survey in French give three dollars less on average to the anonymous respondent. Females give three dollars (or about 15.5%) more on average. This is consistent with a repeated findings that females give more in dictator games (see e.g. Eckel and Grossman, 1996). As important as these differences is the lack of differences on other sociodemographic measures. While most previous studies have used smaller samples (see e.g. Camerer, 2003), even with a large sample I cannot find significant differences according to employment status, household income, education, or partisan identification. This has important implications for the study of altruism. If dictator games are an analogue to altruism and if previous empirical studies have used sociodemographic variables as proxies for altruism (e.g. Hudson and Jones, 1994; Shiell and Seymour, 2002), then this suggests that they may have been off the mark. If altruism is a deeply-ingrained, prosocial ori-

entation, then it should correlate with them weakly, at best. Indeed, this is just as it appears, even with a large sample.

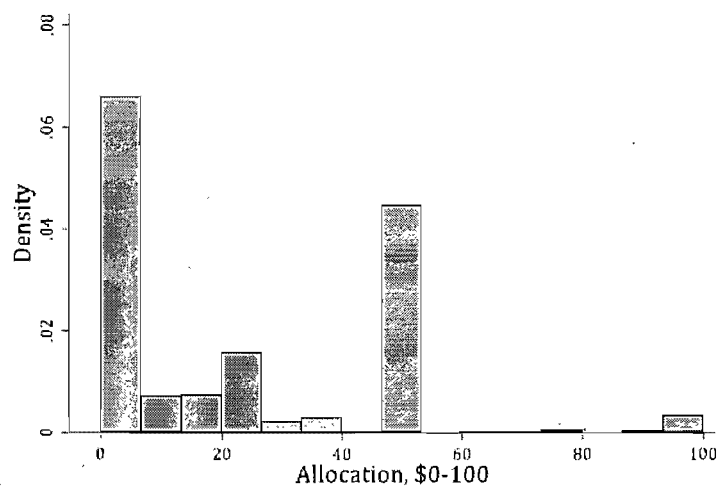


Figure 2.1: Dictator Game Allocations

2.5.1 Support for Public Spending

The results suggest that individuals with more altruistic orientations are more likely to support higher public spending. I begin by estimating demand curves for four public services in question according to dictator game allocations (Figure 2.2). In each case, I divide the sample into those who allocated nothing in the dictator game, those who allocated something less than or equal to the median (\$20), and those who allocated more than the median. Each graph then plots the percentage of individuals willing to pay for a public service at each price. Taking the top left panel in Figure 2 as an

example, we see that the percentage of respondents supporting a carbon tax policy decreases as the price of the policy increases. But, as importantly, a consistent difference between those who allocate nothing in the dictator game and those who allocate something remains. This difference ranges from about 10 percentage points at a cost of \$100 to 3 percentage points at the highest cost (\$2000). At the same time, the differences between those who allocate below and above the median are not apparent. This suggests that the most important difference is between those who allocate something and those who allocate nothing in the dictator game.

The top right and bottom left graphs in Figure 2 also demonstrate clear differences between those who allocate nothing and those who allocate something. Altruists appear more willing to pay for free tuition for university students and they appear more willing to spend more time in a hospital waiting room for non-threatening ailments if they know that this will ensure better drug coverage for those with cancer. Even when I change the cost terms from money to time, then, I find a consistent influence of altruism. However, as with a carbon tax there is not a clear difference between those who give a small amount of money away in a dictator game and those who give a lot.

Finally, the bottom right panel demonstrates willingness to pay higher taxes in exchange for shorter wait times in hospitals. As with the previous examples, a clear pattern emerges where the altruistic are more willing to pay for this public good than the non-altruistic. However, in contrast to the

previous three examples, there is now not only a clear difference between those who give nothing and those who give something but also a difference between those who give below the median and those who give above the median.

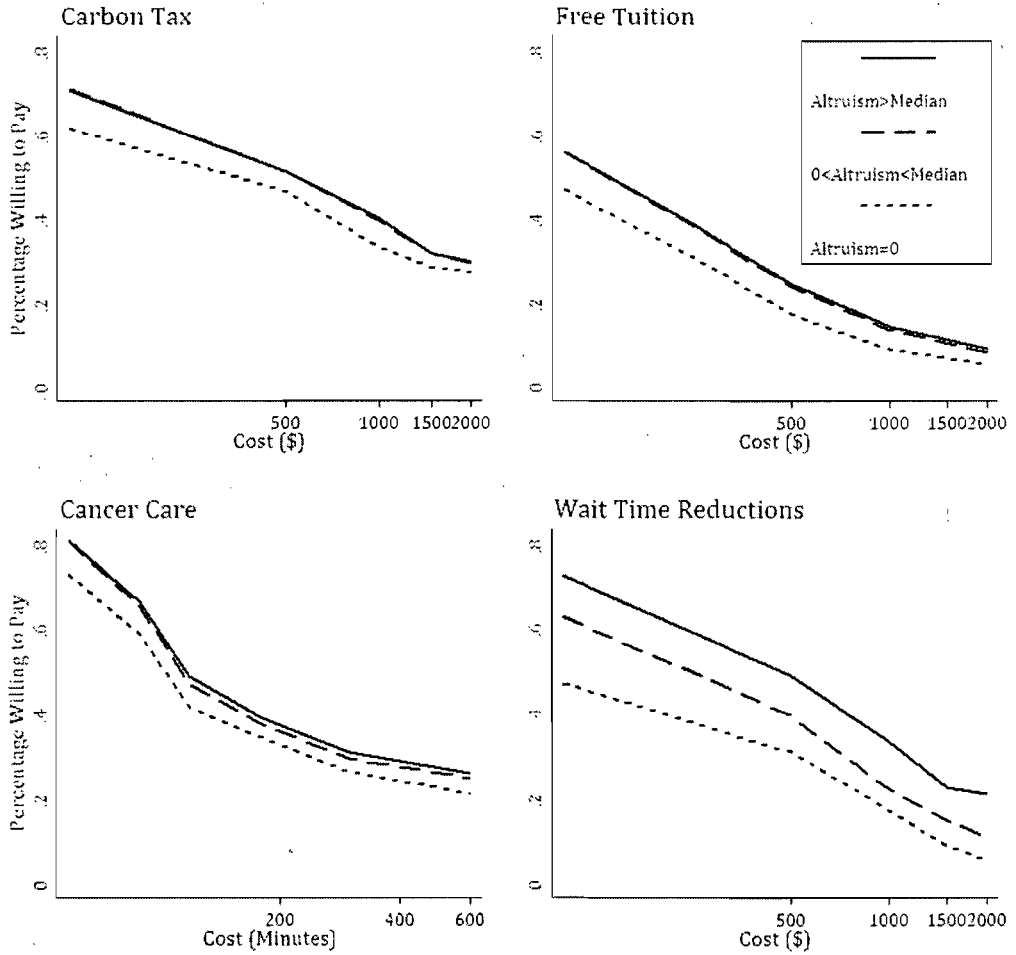


Figure 2.2: Demand Curves for Public Spending

All of these figures have demonstrated a difference in the willingness to incur personal costs for greater public services consistent with a theory of altruistic motivation and strong reciprocity. However, these differences could just as easily be a function not of altruism but of some unobserved factor in the population. I address this through a series of regressions (Tables 2.2-2.5). Each regression is an ordered logit model. The dependent variable is the highest cost category a respondent was willing to enter for the public service in question before answering no or don't know.⁴ This corresponds to 7 categories for the willingness to wait longer in return for better cancer coverage (Table 2.5) and 6 categories for the other three questions.

Each model includes a number of relevant covariates. Dummy variables indicate if a respondent is female or a francophone. They also measure whether the individual works in the public sector, is employed in the private sector, unemployed, a homemaker, or self-employed, or is a student. Retired respondents act as the employment reference category. Categorical variables measure age category, income category, education category, strength of partisan identification, and two variables measure the current and prospective economic uncertainty of respondents. All are scaled from 0 to 1. Following

⁴In the case of questions in which the price was declining with each refusal, the category is the cost at which the respondent first indicated a willingness to pay. Either way, the variables are constructed so that a higher category indicates a willingness to pay a greater amount. In the cases where questions have begun with a high price, respondents who answer Don't Know are discarded, as I cannot ascertain the price at which they would agree to pay. In the cases where questions have begun with a low price (Cancer Care and Free Tuition) I treat Don't Know responses as an unwillingness to pay as respondents who do not know if they would pay for a good at some price can be assumed to be unwilling to pay for it at a higher price.

previous work (see, e.g. Corneo and Grüner, 2002), I expect females and the better educated to be more supportive of public spending. I expect the support of the old to be contingent on the nature of the public good, as the older are more likely to benefit from better medical services but are likely less concerned about environmental degradation and free university tuition. I expect the wealthier to be less opposed to greater spending as the indicated price increases represent a smaller share of their income. And I expect those who work in the public sector to be more favourable to increased public spending (e.g. Blais, Blake and Dion, 1997). I do not have strong expectations based on prior work for the employed and unemployed or for those who face economic uncertainty. On the one hand, those who anticipate future economic trouble may also anticipate not paying taxes; despite a mention of increased cost in the question, they may be willing to indicate willingness to pay for a public good which they feel they could use in the future. Accordingly, I have no expectations for these variables. Similarly, I have no strong expectations for preferences among homemakers and the self-employed.

Among partisans, I expect New Democrats, Liberals, and supporters of the Bloc Québécois all to support greater public spending. I expect partisans of the Conservative party to support less public spending (for an organization of spending preferences along partisan lines, see Lewis and Jackson, 1985; Blais, Gidengil, Nadeau and Neviite, 2002).

As outlined above, I operationalize altruism in three ways. These operationalizations are Models 1 to 3 in Tables 2.2-2.5. In reviewing my results,

Even in the face of several controls thought to influence support or opposition to greater public spending, indicators of altruism are associated with a greater willingness to pay for public spending. According to the estimates provided in the tables, the effect of altruism is always significantly different than zero, and the effects are often larger than those of other well-known correlates. Taking a carbon tax as an example (Table 2.2), those who give away something in a dictator game have 1.37 times higher odds of paying the maximum stated price for a carbon tax than those who give away nothing, according to Model 3. While this effect is smaller than the effects of age, education, student status, public sector employment and various partisan identifications, it is greater than the negative effects of current or anticipated economic hardship, age, income, gender, and employment status. Altruism is not the whole story, but it is explaining important variance in willingness to pay for public spending.⁵

Faced with the choice of paying higher taxes for free university tuition (Table 2.3), altruists are 1.47 times more likely to be in the highest cost category than non-altruists, according to Model 3. This a greater effect than income, education, gender, employment or unemployment, public sector employment, and current or future economic uncertainty. It is also a stronger effect than moving from no partisanship to a strong Liberal partisanship. As with the Carbon Tax, the model fit is also significantly improved with the

⁵A likelihood ratio test suggests that it improves the fit of the model significantly over a model without altruism. This is the case for every specification of altruism in each model.

addition of altruism.

The largest effect of altruism is in the willingness to pay for reduced waiting times for medical services (Table 2.4). Model 3 suggests that those who give something away in a dictator game are 1.63 times more likely to be in the highest cost category of willingness to pay than non-altruists. This effect is larger than the effects for income, education, gender, employment status, current and future economic difficulty. It is also larger than the effect of each partisan identification except one. It is smaller than the effect of being a student. Most importantly, it is nearly half the size of the difference between the oldest and youngest respondent. This is thus a rather remarkable magnitude, as the question makes explicit mention of medical surgeries which are most commonly for elderly patients (i.e. hip replacement and cataract removal).

Finally, altruists are willing to wait longer than non-altruists for standard medical procedures if those who have cancer can have better drug coverage. Model 3 suggests that this effect is greater than every other predictor except anticipated economic hardship, BQ partisanship, and age. The willingness of altruists to bear costs for others, then, extends to non-monetary costs such as time.

Taken together, these results suggest a consistent pattern. Those who exhibit more altruism in a dictator game state a greater willingness to pay for greater public spending. This willingness persists in the face of other well-known correlates of the decision to support or oppose greater public spending.

Moreover, altruism often exhibits a stronger effect than these other factors, consistent with earlier findings (Rasinski and Rosenbaum, 1987; Hudson and Jones, 1994; Fong, 2001; Shiell and Seymour, 2002).

These results are open to an obvious criticism: they are generated from a non-random sample in which the educated and wealthy are overrepresented. If the effect of altruism differs systematically between the wealthy and the non-wealthy, or between the highly educated and others, then these results may not obtain in the entire population. This endangers the generalizability or external validity (Shadish, Cook and Campbell, 2000) of my findings.

To examine this possibility, I have repeated the regressions and added interaction terms between altruism and education category and altruism and income category. If the main effect for altruism remains significant and is not washed out by the marginal effects in the interactions, this suggests that the role of altruism is present across the population. My results, presented in Table 2.6, suggest that this is the case for three of four goods. Only in the case of cancer care does the main effect of altruism fail to reach statistical significance, despite being in the right direction. Moreover, I note that this is the only question for which costs are expressed in non-monetary terms, perhaps making the inducement to altruism less strong. Additionally, only two interactions of eight are significant, suggesting that the effect of altruism is direct and consistent across the population. In sum, these findings on the importance of altruism for support for public spending likely extend to a wider (and less educated and wealthy) population.

Table 2.6: Main and conditional effects of altruism when interacted with education and income (Ordered Logistic Regression)

	Carbon Tax	<i>S.E</i>	<i>p</i>	Free Tuition	<i>S.E</i>	<i>p</i>	Wait Times	<i>S.E</i>	<i>p</i>	Cancer Drugs	<i>S.E</i>	<i>p</i>
Altruism (Model 3)	2.65	0.60	.00	1.41	0.27	.07	2.52	0.55	.00	1.31	0.25	.16
Education*Altruism	0.60	0.16	.06	0.98	0.22	.91	0.53	0.14	.01	1.08	0.24	.74
Income*Altruism	0.62	0.15	.04	1.12	0.23	.59	1.05	0.24	.01	0.76	0.16	.18
Age	0.73	0.23	.32	0.50	0.13	.01	2.38	0.74	.01	1.73	0.46	.04
Income	1.53	0.30	.03	1.11	0.90	.54	1.45	0.27	.85	1.30	0.22	.12
Education	3.06	0.63	.00	1.24	0.21	.22	2.08	0.41	.00	1.06	0.18	.75
French	0.78	0.11	.07	0.28	0.04	.00	0.75	0.10	.04	0.80	0.18	.30
Female	1.07	0.10	.41	1.14	0.09	.10	0.93	0.09	.49	1.16	0.09	.07
Employed	1.49	0.38	.11	1.05	0.23	.81	1.37	0.34	.21	1.39	0.30	.13
Unemployed	1.18	0.41	.64	1.23	0.37	.50	1.61	0.55	.17	1.41	0.43	.26
Self-Employed	1.17	0.22	.39	1.03	0.17	.86	1.12	0.20	.55	1.23	0.19	.18
Student	2.31	0.70	.01	1.92	0.48	.01	2.20	0.66	.01	1.64	0.42	.05
Homemaker	1.19	0.29	.47	1.19	0.24	.38	1.34	0.31	.21	0.95	0.19	.80
Public Sector	0.65	0.14	.04	1.03	0.19	.86	0.77	0.16	.20	0.92	0.17	.64
Making Ends Meet	0.82	0.05	.00	0.99	0.05	.86	0.84	0.05	.00	1.12	0.06	.03
Job Loss Worry	0.89	0.05	.00	1.01	0.05	.82	0.95	0.05	.28	0.85	0.04	.00
NDP ID	2.87	0.66	.00	2.53	0.46	.00	2.93	0.65	.00	1.30	0.24	.15
Conservative ID	0.31	0.05	.00	0.61	0.09	.00	0.70	0.11	.02	0.85	0.12	.27
Liberal ID	1.69	0.29	.00	1.34	0.20	.05	1.26	0.22	.17	1.27	0.19	.11
BQ ID	2.19	0.57	.00	5.38	1.39	.00	1.31	0.36	.33	2.06	0.50	.00
Cut 1	-0.11			-0.44			0.30			-0.99		
Cut 2	0.70			0.42			1.36			0.04		
Cut 3	1.24			1.34			2.11			1.01		
Cut 4	1.58			2.23			2.70			1.52		
Cut 5	1.67			3.05			2.78			2.09		
Cut 6										2.51		
LR χ^2	256.81			227.57			142.21			45.96		
LR χ^2 >~ Altruism Model	0.00			0.00			0.00			0.09		
N	1789			2287			1728			2199		

2.6 Discussion and Conclusion

Explanations of support for public spending which rely on self-interest alone are, however theoretically elegant, empirically implausible. When asked about spending programs from which they cannot be expected to benefit or from which a net gain will not be realized, many citizens still indicate support. I argue that this is due to non-self-interested considerations, particularly altruism. To support this claim, I have demonstrated a consistent and strong link between allocations in a dictator game and support for greater public spending on a variety of programs with a variety of costs. By linking a behavioral indicator of altruism with a willingness to pay higher taxes for public programs, I have established a stronger empirical link than previous studies and added support to arguments for strong reciprocity.

This work is not without objections, particularly that it has occurred over a non-random sample. While my sample is more representative than a typical university convenience sample, it is still open to claims of limited generalizability or external validity (Shadish, Cook and Campbell, 2000). Despite this, I have shown that allocations in dictator games are generally unrelated to sociodemographic variables (see Table 2.1). Moreover, when I control for the marginal effects of education and income, I find that the link still exists for goods with monetary costs. Accordingly, in the absence of an argument about why those who would refuse participation in the study would not behave similarly or hold similar preferences, I argue that my results

generalize broadly.

These results have important implications both methodologically and substantively. From a methodological standpoint, these results demonstrate that we can embed games from behavioral economics into large-sample surveys and learn from the results. In doing so we can arguably introduce more valid and reliable indicators of concepts such as altruism.

From a substantive viewpoint, these results tell us that individuals who exhibit more altruism hold different preferences for public spending than those citizens who do not. However, such individuals do not likely represent a majority of the population. Accordingly, any appeals for public spending which rely on altruism alone are not as likely to garner support as broadly as those calls which also incorporate an element of self-interest. This likely has important implications for those who wish to explain the rise and fall of preferences for more public spending (e.g. Soroka and Wlezien, 2005). More generally, these findings reinforce the call to incorporate more than just self-interest into our explanations of spending preferences.

It is not our claim that these findings call into question the importance of self-interest in the explanation of political action. But they do demonstrate that other-regarding behavior can have an equally and sometimes stronger impact. Our accounts of political behavior, then, should be open to explanations which move beyond simple self-interest. They should, to put it differently, take regard of other considerations.

Chapitre 3

Affinity, Antipathy and Political Participation: How Our Concern For Other Partisans Makes Us Vote

To be submitted at Canadian Journal of Political Science.

3.1 Introduction

"...voting is essentially a group experience" (Lazarsfeld et al 1968, 137)

"Gypsies, tramps, and thieves, these are the people who will vote for McGovern..." As sung at the 1972 Republican National Convention.

Elections are not the simple aggregation of millions of individual and independent decisions whether and for whom to vote. Nor are they simply about self-interested decisions. Elections are instead a competition between groups of people who rely on more than self-interest when deciding when and how to participate in politics. These individuals engage in other-regarding behaviour in which they consider the benefits of an election outcome for whole groups of people. If we wish to understand the decision to participate in politics, we need to take account of this fact.

In this article, I provide an interpretation of other-regarding behaviour and electoral participation in which I argue that antipathy and affinity towards others - specifically, other partisans - can be used to explain the decision to vote or not to vote. Using a game from behavioural economics - the dictator game - I demonstrate empirically that citizens who have stronger preferences or greater concern for some partisans than others are more likely to vote. This suggests that models of voter turnout which rely only on self-

regarding considerations - including even duty and social obligation - are incomplete. A more fulsome account of the decision to vote takes account of this variation in individuals' concern for others.

The demonstration of this argument proceeds as follows. In the next section, I justify a conception of politics as a contest between groups of citizens. In this view, politics is not just a contest between parties fighting for the support of individual citizens. Rather, it is a fight over scarce resources between groups of citizens and their respective representatives in political parties. As such, how individuals feel about various partisan groups matters for their understanding of politics. In section 3, I formalize a calculus for voting which is consistent with this view of politics, particularly by incorporating a regard for others. The model demonstrates how affinity for co-partisans and antipathy towards other partisans can drive the decision to vote. It thus departs from the conventional rational choice model of voting by incorporating concern for others and not relying on a duty term to explain the paradox of participation. It is also thus similar to those models presented by Fowler (2006), Fowler and Kam (2007) and Edlin, Gelman and Kaplan (2007). In presenting the model, I argue that such an account provides a more satisfactory theoretical explanation of the decision to vote than a model which depends on duty, resources, or partisan identification. In section four, I describe a large online survey experiment which uses dictator games to measure antipathy and affinity. Dictator games involve giving a subject a sum of money and then observing how much of that money they

are willing to share with a recipient. I describe the properties of dictator games and their suitability for the measurement of affinity and antipathy. I show in the fifth section that behaviour in these games is consistent with what we should expect according to partisan identification. Partisans give more to their fellow partisans and less to the partisans of other parties and this difference increases with strength of partisanship. For example, those individuals who identify with the Conservative party give more money to fellow Conservatives than to Liberals. And this difference increases with strength of partisanship. The sixth section presents multiple regression models of the decision to vote in the 2006 Canadian federal election which incorporate measures of antipathy and affinity. Closely resembling conventional models of turnout, these models demonstrate that antipathy and affinity matter independent of other well-known correlates of the decision to vote, such as media attention, party identification, education, income, and election competitiveness. Moreover, the models suggests that both affinity and antipathy independently predict turnout, but that affinity has a slightly stronger effect. I discuss these findings and conclude in the seventh section.

3.2 Group Politics

Politics can be understood as a contest between groups of people. Three sets of evidence support this view. First, we generally understand parties as having different bases of support; bases which can generally be described in

terms of social groups.¹ The Liberal party, for example, is the party of visible minorities and Catholics, of Quebec federalists, of francophones outside of Quebec. The Conservative party is traditionally the party of protestants, rural Canadians, and Westerners. And the New Democratic Party is a party of union members, women, and increasingly urban dwellers (Blais, 2005; Blais, Gidengill, Nadeau and Nevitte, 2002; Bibby, 1990). While there is some debate over the importance of social groupings for vote choice (see, e.g., Clarke et al., 1979; LeDuc, 1984), it remains true that parties often think of their support in terms of groups and pursue votes accordingly (for a popular account, see Wells, 2006). And, as an empirical matter, we can explain vote choice as a function of group membership (Blais, Gidengill, Nadeau and Nevitte (2002), for a non-Canadian example, see Abramson, Aldrich and Rhode (2006)).

Second, the rhetoric of parties frames politics as a competition between different groups. In doing so, parties attempt to paint a positive picture of the individuals who support them and paint a negative picture of the individuals supporting other parties. While the invocation of gypsies, tramps, and thieves is perhaps too strong, parties do draw caricatures of their supporters and their opponents. Take, for example, the leaders' debate during the 2006 Canadian federal election.² Stephen Harper, the leader of the Con-

¹This is especially true if we conceive of class membership as a specific instance of a group membership (see, e.g. Evans, 1999; Hout, Brooks and Manza, 1993, for accounts of the enduring importance of class).

²All the following quotes are drawn from the 2006 Canadian federal leaders' debate, according to the transcript of the Canadian Press (2006).

servative Party, characterized his party as “on the side of the people who work hard, pay their taxes, and play by the rules.” Similarly, the leader of the New Democratic Party, Jack Layton, cast his party as the one which would “make politicians in Parliament accountable to you, and we’ll work day in and day out, not for the well-connected, but for working families... We’ll ensure dignity and respect for seniors. And we’ll make sure there’s opportunities for young people...” Prime Minister Paul Martin accused the Conservative party of being on the side of “richer Canadians” and opposed to the interests of their “working class” counterparts. In all of these rhetorical appeals, leaders are framing their policy offerings in terms of the groups which they benefit. Their intention is to draw a picture of the type of people who support their party. If a favourable picture can be drawn, then voters are more likely to be convinced that casting a ballot for a party is going to benefit people whom they like and people who are like them.³

Third, we know that voters think about their membership in parties in the same way they think about their membership in other groups. Indeed, Campbell et al. (1960)’s original conception of party identification was that it was similar to affiliation with other groups, whether religious, ethnic or racial (see also Greene, 2004, 136-137). Recent research has confirmed this view and argued that individuals identify with parties the same way they iden-

³In this respect, politicians also have an incentive to overstate the size of the group of voters they represent. The larger the group to which a voter belongs, the more individuals who will benefit from an election win. This perhaps explains the tendency of voters to overestimate the chances (and thus size) of the parties which they support (see, e.g. Bartels, 1988; Blais and Bodet, 2006).

tify with other social groups (Green, Palmquist and Schickler, 2002; Greene, 2004). In doing so, they adopt positive images not only of the parties, but of the people who support the parties. And they sometimes adopt more negative views of those who support other parties (for more general work on social identity theory, see Tajfel, 1978). Arguably, this suggests that our conceptions of party identification are incompletely served by questions which ask only about attachment to a party and not about feelings towards a party's supporters.

Taken together, these arguments suggest that people vote as groups, parties conceive of elections as contests between groups of voters, and voters think of parties and partisans in the same terms in which they think of other social groups. When we combine this with the fact that some individuals are motivated by a concern for others, then it becomes puzzling that existing models of the decision to vote would not take account of affinity towards others. What is needed, then, is an account of voting in which (some) individuals take account of their feelings of the various groups who stand to benefit from an election's outcome. In the next section, I formalize such a conception.

3.3 A Different Calculus of Voting

Riker and Ordeshook (1968) provide perhaps the definitive self-interested account of turnout. Indeed, in his extensive review of the decision to vote or

not to vote, Blais (2000) takes this as *the* rational choice model. A “paradox of participation” emerges from this model, namely in that it predicts no or very low turnout. To review, the original model posits three components: B , the benefits an individual receives from an election outcome, C , the costs an individual incurs in voting, and P , the probability that an individual’s vote will be decisive. An individual decides to vote if $PB > C$. The problem with the model is immediately apparent. In only the rarest circumstances is P ever anything but infinitesimally small. Indeed, as Fowler (2006, 675) observes, numerous scholars have demonstrated formally (Edlin, Gelman and Kaplan, 2007; Chamberlain and Rothchild, 1981) and empirically (e.g. Gelman, Katz and Bafumi, 2004; Mulligan and Hunter, 2003) that in any election, P is about equal to $1/N$. It thus does not make rational sense for a voter to go to the polls.

To resolve this paradox of participation, Riker and Ordershook proposed adding a duty term, D , resulting in $D + PB > C$. Thus, if a citizen’s sense of duty plus the discounted benefits of winning were greater than the cost of voting, then they would cast a ballot. In his extensive review of the literature, Blais (2000, 2–11) outlines six additional amendments to the model by rational choice scholars (Downs, 1957; Ferejohn and Fiorina, 1974; Mueller, 1989; Uhlaner, 1986, 1989*a,b*, 1999; Niemi, 1976; Barry, 1978; Aldrich, 1993) and four non-rational choice-based explanations (e.g. Brady, Verba and Schlozman, 1995; Rosenstone and Hansen, 1994; Blais, 2000, 13-14). After this, he too comes down on an explanation which “assumes that citizens are con-

cerned with the well-being of their community as much as with their own self-interest” and that encapsulates such a concern in a *sense of duty*.

In my view, an explanation which relies on duty is only half-right. That many citizens have a sense of duty seems uncontroversial. That it would take the form of a concern for others seems equally uncontroversial. But, this is still a static explanation (Fowler, 2006, 675) because it does not condition this concern for others on the importance of the election. In other words, it does not explain why a sense of duty would be greater for some elections than others. As such, it does little to explain the variation we see in levels of turnout between national and local elections, for instance.⁴ Finally, it does not clearly specify whether this obligation to the group is oriented towards others in the group, i.e. individuals want to help others, or whether it is self-oriented, i.e. an individual wants to feel as though she is a member of the group.

We can find a way out of this paradox, I and others would argue, if we develop a model which allows for a concern for others, and which effectively takes into account the outcome of elections. Such a model is also more consistent with a view of politics in which groups of people fight over power and resources, rather than a view in which parties simply play out a competition in front of unconnected and solitary citizens who think only of benefits to themselves.

I present a model of turnout in which the decision to vote depends on

⁴These variations likewise cannot be explained by discounted benefits.

the difference in regard which an individual has for the supporters of various political parties and for the benefits which will accrue to them given some election outcome (for similar models, see Fowler, 2006; Fowler and Kam, 2007; Edlin, Gelman and Kaplan, 2007). My model differs from these in a few subtle respects. First, in contrast to Fowler and Kam, I allow for the presence of more than two parties and do not demand that voters be evenly divided between them. Second, in contrast to Edlin et al., I do not include a feedback mechanism to explain habitual voting. That said, the model clearly owes its fundamental intuition and implications to these prior models.

As with the classical model of voter turnout, the model assumes that there are costs to voting which individuals consider in the decision to vote. While these costs are often small, they are not nil. Individuals face costs, for example, in determining where and how to vote and in learning about parties and issue positions. Additionally, as with the classical model, voters consider the benefits to themselves. But they discount these benefits by the probability of their vote being decisive for their preferred party. However, unlike the classical model, this model assumes that voters also care about benefits to others, specifically the benefits which are realized by supporters of the winning party. The more they care about those supporters in contrast to supporters of other parties, the more likely they are to vote in an election.

Formally, the model assumes that two sets of benefits exist. First, benefits to self: B_S . Second, benefits to the supporters of the winning party: B_0 . In real terms, both sets of benefits could include changes in tax laws which

favour the supporters of the winning party more than the losing party, on average. They could similarly include new spending measures (or cuts) which disproportionately favour (harm) the supporters (opponents) of the winning party. The important point to be noted here is that these benefits need not be construed in terms of patronage, but rather about the larger programmatic differences in spending which occur between parties of various stripes (e.g. Blais, Blake and Dion, 1993). By omission, the model thus assumes that no generalized benefits exist for all members of society given the election of some party over another (contra Fowler and Kam, 2007).

The model further assumes that voters do consider P , the probability of casting a deciding or tying vote. As Fowler notes (Fowler, 2006), in any election in which the outcome is uncertain this term generally equals $\frac{1}{N}$. The model also incorporates the concern or *affinity* of an individual for supporters of their own party, α_{aff} , and *antipathy* for supporters of other parties, α_{ant} . Finally, the model assumes that voters conceive of the election in terms of a competition between citizens who support their party, who make up some share of the population (n_1), and those who support other parties, who make up another share of the population (n_2). Accordingly, I assume that $n_1 + n_2 = 1$ and that voters adopt a mean level of antipathy towards all other partisans in their calculus.⁵ Whereas the decision to vote in a classical

⁵This corresponds to our Max-Mean specification in the empirical tests. Alternately, we can assume that voters conceive of elections as a contest between their own group of supporters and the supporters of the least preferred group. This corresponds to the Max-Min specification in the empirical tests. The two specifications produce very similar results.

model depends on $D + PB > C$, in this model an individual votes if :

$$P(B_S + \alpha_{aff}B_ONn_1 + \alpha_{ant}B_ONn_2) > C,$$

which rearranges as:

$$PN\left(\frac{B_S}{N} + \alpha_{aff}B_On_1 + \alpha_{ant}B_On_2\right) > C$$

As P is approximately equal to $1/N$, then the PN term drops away.⁶ Benefits to self, B_S , are similarly discounted by N and drop away from the model. However, as the benefits to others, B_O , are not discounted by P or N , then this can be a sufficient motivator to vote in the face of costs. Intuitively, voters care about the benefits incurred by others, and these benefits add up to a non-trivial sum when a group is sufficiently large. The model thus reduces to:

$$(\alpha_{aff}B_On_1 + \alpha_{ant}B_On_2) > C$$

Since one individual's vote can confer a non-trivial benefit on a whole group of people and deny the benefit to other groups, those voters who care about the utility of others can now be motivated to vote. This is true even in

⁶ $N * \frac{1}{N} = \frac{N}{N} = 1.$

the face of C .⁷ As the difference in their concern for others increases, i.e. as α_{aff} increases and/or α_{ant} increases, then voters should become more likely to vote. Substantively, if politics is a competition between groups than those who see one group as more deserving of an election's spoils than others will be more likely to vote. Antipathy and affinity, then, should explain some of the decision to vote. I next describe a survey experiment designed to test this proposition.

3.4 Survey and Research Design

My study relies on an online survey of some 2035 respondents conducted by a commercial public opinion research firm in Canada in May 2007. The respondents are broadly but certainly not perfectly representative of the population. The survey contained conventional questions about political participation and political preferences, but also included a series of dictator games from behavioural economics aimed at revealing affinity and antipathy towards other partisans. Below, I describe the survey, the survey participants, and the variables drawn from the survey.

⁷We could add into the left hand side of the model terms for duty, party identification, or some individual utility derived from voting regardless of the outcome. We could similarly add a negative cost term to reflect the cost of shame for a group member who does not vote. All of these factors could explain variation in the baseline probability of any individual voting. This is a case for including relevant control variables in an empirical model. However, as these variables are not central to the model presented here, I leave them out of the formal discussion.

3.4.1 Survey

The survey was conducted online. Subjects were required to login to the survey using a unique identification. This allows me to call up previously entered demographic information from those who have completed prior surveys. Those completing the survey for the first time were first asked a series of screening questions, including whether they voted in the most recent federal election and their partisan identification. Subjects answered several questions about recent news exposure, their attention to federal and provincial politics, and their views on federal and provincial politicians. Subjects then completed an unrelated eight-item module on empathy. They were next presented with the dictator game battery. Following this, they were presented with questions concerning their support for public spending, their past charitable giving, their views of the public service, and their views of recent political events. The final effective sample was 2035 respondents.⁸

3.4.2 Subject Profiles

Compared to a university-based convenience sample, the online survey methods affords a large number of respondents and comparatively representative population, particularly in regards to age, education and income. Compared to a telephone survey, it allows us to present subjects with more complex or

⁸The sample is limited by three factors. First, I eliminate those who have not indicated whether they voted in the 2006 federal election. Second, I eliminate those whose constituency is not identified. Finally, I eliminate those for whom values are missing on income and education.

complicated questions, such as the dictator game questions used to measure affinity and antipathy, while not sacrificing the advantages of a broadly representative sample. It should be noted, however, that the sample does not perfectly resemble one which is randomly drawn.

Table 3.1 shows the characteristics of the sample and compares them to the 2006 Canadian Election Study, a RDD telephone survey. The average respondent in the internet sample is slightly younger ($t = 2.60, p < .00$). There is also a lower incidence of French respondents ($\chi^2 = 13.93, p < .00$) and female respondents ($\chi^2 = 7.06, p < .00$) in the internet sample. Finally, the internet sample appears both wealthier ($\chi^2 = 123.55, p < .00$) and more educated ($\chi^2 = 349.76, p < .00$).

In addition to the socio-demographic differences, the internet sample also appears to be more politically engaged. It exhibits a higher incidence of both weak and strong Conservative identifiers, weak Liberal identifiers, weak and strong NDP identifiers, and strong Bloc Quebecois identifiers. Overall, the internet sample has a higher incidence of those who identify with a party than the CES sample ($\chi^2 = 24.99, p < .00$).

The most glaring difference between the sample and the general population is the incidence of turnout (it is 91.4% in the internet sample, 90.5% in the CES post-election survey, but only 64.7% in the population). The panel is quite clearly overpopulated by those who claim to have voted in the last federal election. Because vote is our dependent variable, this imbalance cannot be ameliorated by a control variable. To address this, I weight the

data according to the actual rate of turnout in the 2006 federal election. As a result, my regressions rely on control variables to account for differences on sociodemographic variables and party identification, and a weighting to address the over-reporting of voting (for a similar approach, see Blais et al., 2004).⁹

3.4.3 Antipathy, Affinity, and other variables

To begin with standard variables, survey questions were used to capture respondent demographics as well as party identification. Furthermore, respondents were asked how many days a week they read the newspaper and watch television news. They were equally asked how many hours a week they spend reading internet news. For interpretive ease, all variables are recoded from 0 to 1. Question wordings are available in Annexe D.

Antipathy and affinity are measured through a series of dictator games (see Camerer, 2003), the properties of which are discussed in more detail below. In addition to a regular \$500 draw for survey participation, respondents were told that they were eligible to win up to four prizes of \$100 at the end of the survey. In the case of one prize, they were asked how much of it they

⁹To address the higher frequency of voters, I weight my sample to reflect national (rather than provincial or regional) rates of turnout. I make no corrections within demographic groups. Accordingly, voters receive a weight of 0.708 and non-voters a weight of 4.08. I have also estimated a rare events logit (King and Zeng, 2001) for each model with corrections for the frequency of the dependent variable, but no McCullagh and Nelder correction or variance cluster correction. In the case of each model in Tables 3.4 and 3.5, the substantive results remain the same.

Table 3.1: Sample demographic and political characteristics

Variable	Current % or Mean (SD)	CES % or Mean (SD)
Age	49.7 (13.5)	50.8 (16.5)
French	18.2%	22.3%
Female	49.1%	52.7%
Household Income	<\$40000	36.7%
	\$40000 to \$60000	20.5%
	\$60000 to \$80000	16.4%
	>\$80000	26.4%
Education	High School or less	37.6%
	Some College	25.6%
	Some University	36.9%
Conservative ID	Weak	11.5%
	Strong	6.3%
Liberal ID	Weak	15.5%
	Strong	6.5%
NDP ID	Weak	5.4%
	Strong	2.9%
BQ ID	Weak	5.6%
	Strong	2.1%
Total Party ID	62.5%	55.8%
Voted	91.3%	90.5%
N	2035	4057

would share, should they win, with an anonymous individual about whom they knew nothing. For the other three prizes, they were similarly asked how much they would be willing to share with an anonymous individual about whom they knew nothing except which political party the respondent typically supported (Conservative, Liberal or NDP; in the case of Quebec residents, Conservative, Liberal or Bloc Quebecois). The presentation or-

der of the anonymous recipients was randomized. Question wording for the dictator games is available in Annexe B.

The comparison of these amounts can give us important information about the level of antipathy and affinity individuals have for supporters of other parties. For example, if an individual indicated she would give \$50 to a Conservative but nothing to a Liberal, then we may say she has more antipathy for Liberals than a respondent who gave the same amount to partisans of both parties. Alternately, we could say she has more affinity for Conservatives. I leave a discussion of the operationalization of these specific variables to Section 6.

In the past, dictator games have been used to measure other-regarding behaviour, whether altruism, social identification or fairness (e.g. Fowler, 2006; Fowler and Kam, 2007; Whitt and Wilson, 2007). The question remains as to whether they can be used to measure antipathy and affinity for other partisans. I argue that differences in partisan allocations in a dictator game are just such evidence of antipathy and affinity towards other partisans. That is, if an individual is willing to allocate \$50 to a supporter of the Conservative party and \$0 to a supporter of the Liberal party, then they can be said to have antipathy towards Liberals and/or affinity for the Conservatives, particularly because they are displaying a real difference in their concern for others, and the display of this concern comes at a real cost. Behaviour in the dictator game thus closely resembles that which we would expect according to social identity theory and the preference for some groups over

others (see Tajfel, 1978; Greene, 2004). It can similarly be construed as a measurement of “social distance” between respondents and other partisans (Hoffman, McCabe and Smith, 1996)

The use of behavioural economics games in political science and especially in voting behaviour research is rather unconventional. As a result, several reasonable objections can be anticipated (see Benz and Meier, 2008, 2-3). First, one could contend that subjects do not understand the play of the game and instead make allocations more or less randomly. However, Andreoni and Miller (2002) and Dawes and Fowler (2007) have shown through the manipulation of payoff functions that individuals do understand the game and do not simply make up allocations as they go along. Instead, their allocations consistently and rationally match their stated preferences. Second, it could be argued that the small stakes of the games mean that individuals would play differently if the stakes were higher. Most research, however, suggests that subjects play consistently provided the stakes are real (as they are in our game) (e.g. Camerer and Hogarth (1999); Carpenter, Verhoogen and Burks (2005), but see also Cherry, Frykblom and Shogren (2002)). Third, it can be argued that despite consistent play, behaviour in dictator games does not correspond to the real world equivalents we wish to measure. Benz and Meier (2008), however, review strong evidence of the correlation between dictator game allocations to anonymous individuals (taken as a measure of altruism) and charitable giving, among other actions (for a longer review, see Loewen (2008*b*)). Accordingly, I am confident behaviour in these games

reveals respondents' concern for others.

These objections aside, the dictator game has substantial advantages over, for example, asking subjects to state how they feel about other partisans (e.g. Greene, 2004). First, stated opinions are arguably more subject to social desirability than revealed preferences in a dictator game, precisely because the former are virtually costless to make. Dictator game allocations ask subjects to put their dollars where their hearts are. Second, while asking subjects to allocate money may seem like an abstract task, it is likely one which subjects can undertake with more consistency and meaning than, for example, trying to translate their preferences for some partisans over others onto a 7-point Likhart scale or onto a 0-100 thermometer. Indeed, the dictator game is desirable because it asks subjects to demonstrate their affection for some groups over others at a cost to themselves, and it does so in quantities which a subject can readily understand.

3.5 Antipathy, Affinity, and Party Identification

Table 3.2 demonstrates the different allotments of partisans in the dictator games. The differences in these allocations suggest that the dictator game does uncover affinity and antipathy between political supporters, as subjects give more to co-partisans than they do to rival partisans. Moreover, they are likely to give more to those who are not identified with a party than those

who are identified with another party. For example, the first row in Table 3.2 shows that weak Conservative identifiers give \$21.20 to other Conservatives on average, while they give only \$12.50 and \$12.70 to Liberal and New Democratic respondents, respectively. These differences become more stark when we consider the allocations of strong Conservative identifiers. These individuals on average give other Conservatives \$26.70, while they give Liberal and New Democratic respondents only \$10.60 and \$10.20 respectively. A similar pattern obtains for New Democratic, Bloc Quebecois and Liberal partisans. They allot more money to their co-partisans than to other partisans, and these differences are larger for strong partisans than for weak. This pattern only fails to obtain with regards to the allocations of strong Liberals and strong Conservatives to Bloc partisans.

Those who do not identify with a political party - more than a third of our sample - appear to give less to partisans compared to anonymous individuals. On average, non-partisans give \$22.40 to anonymous recipients (partisans give about the same on average). They conversely give between \$15.30 and \$16.70 to partisans. Taken together, all of these results suggest that the allocations in the dictator game are consistent with respondent partisanship or non-partisanship.

Table 3.3 demonstrates the within-subject differences in allocations. Each cell presents the average within-subject differences by donor. For example, the cell in the upper left demonstrates the average difference in allotments to Conservatives and Liberals by Conservative donors. I then use a Wilcoxon

sign-rank test to determine the significance of this finding. As a consequence of being non-parametric, the Wilcoxon does not assume that the quantities being compared are normally distributed as with a conventional t-test. The test reports a probability that the direction of the real difference in scores is the opposite of that observed. Accordingly, the p-values in parentheses represent the probability that the difference exhibited is in fact in the other direction (Wilcoxon, 1945).

An examination of these results finds that, as with the observations in Table 3.2, within-subject allocations are consistent with partisanship. Conservatives allocate significantly more to Conservatives than Liberals, more to Conservatives than New Democrats, more to Conservatives than to Bloquistes, and more to Conservatives than anonymous individuals. Moreover, they give more to anonymous individuals than to any other partisans. As importantly, they make no distinction between Liberal and New Democratic recipients. Liberal, New Democratic and Bloc identifiers make similarly consistent allocations, allocating their co-partisans significantly more money than other partisans and non-partisans.

A final observation is warranted. Fowler and Kam (2007) find that participants in their experiments exhibit a bias against Republicans. A similar bias against Conservatives is exhibited in these data. Non-partisan recipients give significantly less to Conservatives than to New Democrats or Liberals. New Democrats similarly give less to Conservatives than to Liberals, and Liberals give less to Conservatives than to New Democrats. Two possible sources of

this Conservative bias both support a view of politics as a struggle between different groups of partisans. Non-Conservatives may exhibit less concern for Conservative partisans because they believe they are unfairly enriched by the current Conservative government. They could also exhibit less concern because they believe that, as a group, Conservatives are less in need of the support of others, consistent with a view of Conservative supporters as well-off financially. These explanations are not exclusive, and both support the view that citizens approach politics with clear distinctions between groups of partisans, differences which translate into varying levels of concern.

Table 3.2: Partisanship and Average Allocations in the Dictator Game

Donor/Recipient	Anon.	Cons.	Liberal	New Dem.	Bloc Que.	N
Weak Conservative	20.7	21.2	12.5	12.7	2.9	311
Strong Conservative	22.3	26.7	10.6	10.2	19.3	168
Weak Liberal	23.6	16.3	23.4	19.5	8.8	414
Strong Liberal	24.0	11.3	28.5	15.7	15.9	143
Weak NDP	23.1	11.7	18.3	28.8	14.5	188
Strong NDP	23.4	13.2	19.1	33.8	13.5	70
Weak BQ	20.6	12.7	12.2		20.7	80
Strong BQ	19.5	12.1	11.4		29.1	82
Non-Partisan	22.4	15.3	16.7	16.6	16.7	896

3.5.1 Why Antipathy and Affinity are not just Party Identification

Given the preceding the results, it can be objected that rather than measuring the affinity and antipathy towards other partisans, the dictator game allocations are simply a different measure of partisan identification. Three pieces of evidence militate against this contention. First, many partisan identifiers give nothing to their co-partisans: 30.1% of Conservative identifiers, 36.6% of Liberal identifiers 27.7% of NDP identifiers and 36.8% of Bloc identifiers give nothing to their fellow partisans. Second, most non-partisans allocate money to one or more co-partisans. Indeed, only 41.3% of non-identifiers allocate nothing to all other partisans. Together, these findings suggest that partisan identification is neither sufficient nor necessary to display differing levels of concern for the partisans of other parties. Third, as the models presented below in Tables 4 and 5 show, when measures of affinity and antipathy are added to a turnout model with party identification, all variables remain significant and the marginal effects of partisan identification remain unchanged. Indeed, as the models below demonstrate, I obtain stronger results when I model the decision to turnout as a function of concern for others, as suggested by our theoretical model. Taken together, this evidence suggests that I am tapping into feelings which, while related to the traditional measure of partisan identification, are not one and the same. Instead, they reach into another element of partisanship, particularly that which involves the feelings

of partisans towards other partisans and not just formal parties (see also Green, Palmquist and Schickler, 2002; Greene, 2004).

Table 3.3: Within-Subject Differences in Dictator Game Allocations (Wilcoxon Sign-Rank Differences)

Donor	Con-Lib Mean (ρ)	Con-NDP Mean (ρ)	Con-BQ Mean (ρ)	Lib-NDP Mean (ρ)	Lib-BQ Mean (ρ)	Con-Anon Mean (ρ)	Lib-Anon Mean (ρ)	NDP-Anon Mean (ρ)	BQ-Anon Mean (ρ)
Conservative	11.4 (.00)	11.4 (.00)	13.3 (.00)	-0.4 (.63)	4.6 (.02)	2.0 (.00)	-9.4 (.00)	-9.5 (.00)	-10.1 (.00)
Liberal	-9.7 (.00)	-4.0 (.00)	5.4 (.00)	6.6 (.00)	11.1 (.00)	-8.7 (.00)	1.0 (.02)	-5.5 (.00)	-10.8 (.00)
New Democrat	-6.4 (.00)	-18.0 (.00)	-1.4 (.95)	-11.1 (.00)	0.9 (.68)	-11.0 (.00)	-4.6 (.00)	6.3 (.00)	-3.0 (.97)
BQ	0.6 (.27)		-12.9 (.00)		-13.5 (.00)	-7.7 (.00)	-8.2 (.00)		5.2 (.00)
Non-Partisan	-1.4 (.02)	-1.5 (.01)	0.0 (.80)	-0.0 (.37)	1.2 (.62)	-7.1 (.00)	-5.7 (.00)	-6.0 (.00)	-4.6 (.01)

3.6 Antipathy, Affinity, and Turnout

My contention is that those who display higher amounts of antipathy towards the supporters of other parties and higher amounts of affinity for supporters of their party should be more likely to vote than those who do not make a distinction between the supporters of various parties. Moreover, this effect should be independent of other predictors of the decision to vote, such as education, income, gender, political interest, and partisan identification.

Table 3.4 presents results from three logistic regressions. The first presents a standard model in which the decision to vote is regressed on party identification, sociodemographic factors, three measures of news consumption, and the closeness of the race in the respondent's constituency. In keeping with many prior research findings (e.g. Leighley and Nagler, 1992*b,a*; Strate et al., 1989) older, more educated and wealthier citizens are all more likely to vote. Likewise, those who identify with a political party are more likely to have reported casting a ballot (Huckfeldt and Sprague, 1992). While females appear less likely to vote, and internet news consumption has no effect on the turnout decision, all other variables conform to a standard account of turnout (for a similar turnout model using Canadian data, see Blais, Gidengill, Nadeau and Nevitte, 2002).

The second and third models add two combined measures of affinity and antipathy. The first, Max-Min, is the difference between the maximum allocation to a partisan less the minimum allocation to a partisan, rescaled 0-1.

The second, Max-Mean, is the difference between the maximum allocation to a partisan less the mean allocation to all other partisans, rescaled 0-1. So, if a respondent gave \$50 to a Conservative, \$40 to a Liberal and \$30 to a New Democrat, then Max-Min would read 0.2 $((\$50-\$30)/100)$. Max-Mean would read 0.15 $((\$50-(\$40+\$30)/2)/100)$. Obviously, these two measures are closely related ($r = .95, p = .00$).

Table 3.4: Antipathy, Affinity and Turnout (Logistic Regression)

	Empty			Max-Min			Max-Mean		
	Coef	S.E.	ρ	Coef	S.E.	ρ	Coef	S.E.	ρ
Affinity, Antipathy				0.71	0.27	0.01	0.82	0.33	0.01
Margin	0.55	0.39	0.16	0.63	0.39	0.12	0.60	0.39	0.12
Party ID	1.19	0.15	0.00	1.14	0.15	0.00	1.13	0.15	0.00
Age	2.24	0.29	0.00	2.21	0.29	0.00	2.21	0.29	0.00
Education	0.68	0.14	0.00	0.66	0.14	0.00	0.66	0.14	0.00
Income	0.74	0.14	0.00	0.66	0.14	0.00	0.66	0.14	0.00
French	0.14	0.15	0.33	0.13	0.15	0.35	0.13	0.15	0.36
Female	-0.49	0.11	0.00	-0.50	0.11	0.00	-0.49	0.11	0.00
Newspaper	0.55	0.14	0.00	0.56	0.14	0.00	0.57	0.14	0.00
TV News	0.62	0.16	0.00	0.60	0.16	0.00	0.60	0.16	0.03
Internet News	-0.56	0.28	0.04	-0.61	0.22	0.00			
Constant	-2.02	0.21	0.00	-2.07	0.22	0.00	-2.06	0.22	0.00
LR χ^2	380.98			388.17			387.47		
LR 2 > 1, $\rho = 0.01$									
LR 3 > 1, $\rho = 0.01$									
N=2035									

Both of these variables are significant and positive, suggesting that increased affinity and/or antipathy increases the probability of voting, even in the face of standard controls. Moreover, the magnitude of this effect is not small and is only eclipsed by the effects of age, partisan identification and income. According to column three, the difference in average probability of voting between our youngest and oldest respondent is some 43 percentage points. Moving from no party identification to a strong party identification and holding all other variables at their mean increases the average probability of voting by 23 percentage points. The effect of moving from the lowest to the highest income category is 16 percentage points, on average. By contrast, the effect for affinity and/or antipathy is 16 percentage points. This is greater than the effect of gender (11 percentage points), television news consumption (13 percentage points), newspaper consumption (12 percentage points), and internet news consumption (14 percentage points). The decision to vote clearly depends on more than just sociodemographic factors, media attention, and even an identification with one party or another. It also depends on how much an individual is concerned with the well-being of co-partisans versus those who support other parties.¹⁰

My formal model suggests that any increase in antipathy or affinity is likely to increase turnout. The combined measures provided above suggest this is just the case. However, the question remains as to which element has a stronger effect if observed separately.

¹⁰These quantities are all based on Clarify estimates with 1000 simulations.

In Table 3.5, I make a distinction between Antipathy and Affinity. Because I am not now just measuring a distance between two allocations, I need a common reference point against which to calculate affinity and antipathy. Dictator game allocations to a completely anonymous individual provide just such a reference point. Accordingly, I define affinity as the difference between the maximum allocation to a partisan and the allocation to the anonymous individual. Affinity thus demands that an individual have more concern for at least one group of partisans than completely anonymous individuals. I censor negative values at zero. Antipathy is defined as the completely anonymous allocation less the minimum partisan allocation. Antipathy thus demands that a respondent like some group of partisans less than individuals about whom they know nothing. As with affinity, I censor negative values at zero. As with the previous two variables, these are rescaled from 0 to 1. Importantly, these two variables are significantly but not strongly correlated ($r = -0.10, p = .00$).¹¹

The results in Table 3.5 suggest little change in the effect of the control variables. More importantly, they suggest that affinity plays a slightly stronger role than antipathy. I use Figure 3.1 to show the comparative effects of these two measures on the probability of voting when considered together. The Z-axis (the vertical axis on the left) measures the probability of voting.

¹¹It is possible that some respondents make allocations inconsistent with their preferences, namely by allocating the most to supporters of a party with which they do not identify. I find that 4.1% of respondents meet this condition. They are retained in the analysis

Table 3.5: Separating Antipathy and Affinity (Logistic Regression)

	1			2		
	Coef	S.E.	ρ	Coef	S.E.	ρ
Antipathy				0.77	0.28	0.01
Affinity				0.92	0.42	0.03
Margin	0.55	0.39	0.16	0.61	0.39	0.12
Party ID	1.19	0.15	0.00	1.14	0.15	0.00
Age	2.24	0.29	0.00	2.19	0.29	0.00
Education	0.68	0.14	0.00	0.66	0.14	0.00
Income	0.74	0.14	0.00	0.76	0.14	0.00
French	0.14	0.15	0.33	0.14	0.15	0.34
Female	-0.49	0.11	0.00	-0.50	0.11	0.00
Newspaper	0.55	0.15	0.00	0.56	0.15	0.00
TV News	0.62	0.16	0.00	0.59	0.16	0.00
Internet News	-0.56	0.28	0.04	-0.64	0.28	0.02
Constant	-2.08	0.22	0.00	-2.08	0.22	0.00
LR χ^2	380.98			392.60		
LR 2 > 1, $\rho = 0.00$						
N=2035						

The X axis measures antipathy and the Y axis measures affinity. The baseline probability of voting (65%) is represented by the ‘floor’ of the graph. The plane shows that both antipathy and affinity increase the probability of voting from 65% to around 80%. The slightly stronger effect for affinity can be seen in the top right hand corner of the plane: when antipathy is equal to 0 the probability of voting at maximum affinity is 82% (as shown by the change in the shade on the plane). By contrast, when affinity is 0, maximum antipathy leads to a probability of voting of just below 80%. On balance, however, both effects appear substantively important, even if one is slightly smaller than the other. Moreover, they do so in the face of traditional con-

trols for party identification. I can thus better explain the decision to vote by incorporating measures of individuals' concern for other partisans.

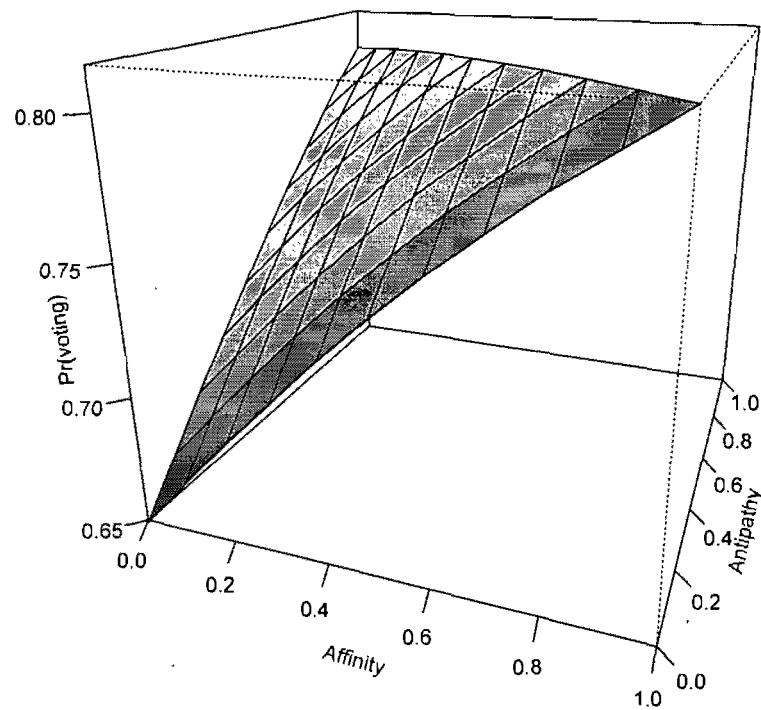


Figure 3.1: The figure demonstrates the increase in the probability of voting at each level of antipathy contingent on the level of affinity, and vice-versa. As the two variables cannot sum to more than one, the plane shows the range of all possible predicted values. Predicted probabilities are based on 1000 Clarify simulations.

3.7 Discussion and Conclusion

I have presented an explanation of turnout which was captured in a simple decision theoretic model and demonstrated empirically using a series of dictator games embedded in an online survey. Moreover, these results were shown to be robust to a number of conventional controls. These results lend support not only to my account, but the similar accounts of Fowler (2006), Fowler and Kam (2007), and Edlin, Gelman and Kaplan (2007).

This model of turnout is more fulsome than one which relies on traditional conceptions of party identification or duty. It recognizes that party identification is about more than a preference for one party over another. Instead, it is membership in a social group. Moreover, it recognizes that senses of obligation to others are likely to drive decisions to participate, especially when the stakes of participation increase. The fulsomeness of this model, then, comes from taking a broader view of partisanship and a view of duty which is not deaf to instrumental outcomes.

These findings have important implications for our study of politics and the decision to participate in politics. The results support the view of politics as a competition between groups in which individuals are concerned not only with their own well-being, but also the well-being of others. On the positive side of the ledger, this is an encouraging result for those who desire a politics which is typified by civic concern and not just by pure self-interest. Indeed, these results suggest that many people participate in elections because they

care about others.

These findings are not entirely positive, however. Looking at these results from the perspective of antipathy, we see that as a preference for some citizens over others increases, some individuals are more likely to vote. By extension, this means that if groups of partisans can be made more polarized and more distrusting of one another, then their likelihood of participating should increase. In short, if politics can be made more negative, then voter turnout could be expected to increase. This does not necessarily recommend negative politics, but it does call into question the view that greater voter participation is necessarily a virtue. Indeed, it supports a much older view that high turnout is not necessarily indicative of civic engagement but of conflict (Berelson, Lazarsfeld and McPhee, 1954).

Whether one takes these results as positive or negative, the sum result is that a better model of turnout incorporates individuals' concern for others. Elections are not contested by parties which appeal to a disaggregated collection of atomized individuals. Rather, they are contested by parties who compete for the support of groups of voters. In doing so, they often portray the supporters of other parties in an unfavourable light. As the story goes, an election win for an opposing party is thus likely not only to perhaps make aggregate welfare worse off, but especially to comparatively enrich those who gave the party their support. According to this conception of elections, the views individuals hold of those supporting other parties matter. As their concern for those who support other parties differs from their concern for

those in their own party, they become more likely to vote. Of all the explanations for turnout, then, we should have increased affection for those which incorporate other-regarding preferences.

Chapitre 4

For Want of a Nail: Direct Mail and Negative Persuasion in a Leadership Race (with Daniel Rubenson)

Under review at Quarterly Journal of Political Science. The experiment was conceived by Rubenson and Loewen. Loewen negotiated an agreement with the campaign. Rubenson and Loewen oversaw the execution of the experiment. The analysis was performed collectively. Loewen wrote the first draft of the paper. Subsequent drafts have been performed equally between Rubenson and Loewen.

4.1 Introduction

Does direct mail work? Political campaign managers certainly believe it does. In nearly every type of political campaign at every level of competition, some form of mail is used. Sometimes this mail serves the purpose of outlining a candidate's position, or casting an opponent's position in an unfavorable light. At other times it is used for fundraising. It sometimes serves a mobilizing function, encouraging potential voters to participate in an election. Most often it takes up several of these tasks at once. Whatever its purpose, there seems little question that direct mail is a frequently used tool in politics generally.

The ubiquity of direct mail is easily explained. It is a relatively cheap manner in which to reach a large number of voters. Moreover, when its design incorporates individual level data on a voter's preferences or concerns (or even their consumer habits and financial status) it promises still greater potential effectiveness. Most importantly, direct mail allows parties or candidates to personally connect with voters through potentially highly targeted messages. This combination of low cost and tailored messaging should only increase the importance of direct mail in the future. Despite this, the persuasive effects of political direct mail have not undergone systematic academic study. The question remains: Is direct mail an effective tool for persuading voters? More precisely, is direct mail an effective tool to persuade elites to support a party leadership candidate who holds controversial positions?

This paper presents evidence from a field experiment into the effectiveness of direct mail in changing elite vote intentions and increasing a candidate's likeability. As such it speaks to two literatures. First, a growing literature of field experimental research into the effectiveness of modern campaign techniques. Second, to an important literature on the strategic communication of leaders. The experiment we present was conducted in cooperation with the Michael Ignatieff campaign in the 2006 Liberal Party of Canada leadership race. To our knowledge, this represents the first field experiment in Canadian politics and the first within the context of leadership elections. It marks, then, an extension of both geography (Canada) and, more importantly, domain (elite politics). As we discuss in more detail below, the Ignatieff campaign provides us with an interesting and unique case for testing the persuasive power of direct mail. Ignatieff was a candidate who sought to change the direction of the Liberal Party on several important and controversial issues. Other candidates who were closer to the Liberal consensus had the job of convincing delegates that they were the best person to manage and implement that consensus. Ignatieff had to not only win over delegates to the view that he was the best person to lead the party; he also took up the more difficult task of persuading delegates to adopt new, non-mainstream positions on core policies. He did so boldly and unambiguously.

The findings are striking. Contrary to campaigns' beliefs about the benefits of their strategy, for at least one frontrunning candidate, there was no positive effect from communicating controversial campaign positions di-

rectly. To the contrary, we find evidence of a negative persuasion effect. These findings correspond with other recent studies demonstrating *contrast* (Chong and Druckman, 2007a) or *boomerang* effects (Peffley and Hurwitz, 2007; Haider-Markel and Joslyn, 2001; Johnson et al., 2003). Taken together, these findings raise a warning about leaders' persuasion efforts : Ineffective or weak arguments do not risk merely falling on deaf ears. Rather, they carry the risk of increasing opposition to a candidate or policy among those who are initially opposed or ambivalent. While these findings do not conclusively demonstrate the disutility of direct mail or the inability of leaders to persuade, they do raise important questions about the conditions under which leaders can change the minds of elites.

The paper is organized as follows. We begin by situating our research in existing literature on leadership communication and literature on the persuasive capacities of direct mail. We then briefly discuss the race in which the experiment occurred. In section four we outline our field experiment and justify its use in comparison to other inferential techniques. Section five presents our model and results. After discussing our findings, we conclude.

4.2 Direct Mail and Persuasion

As in other jurisdictions, direct mail is ubiquitous in Canadian political campaigns. Older evidence suggesting the importance of printed materials, such as that presented by Paltiel (1974), has been confirmed by recent analyses

of modern campaigns. Carty and Eagles (2005), in particular, have documented the importance of printed advertising for modern local campaigns. Using data from the 2000 Canadian federal election, they observe that print advertising was the largest expense of candidates in all parties. While this material encompasses much more than just direct mail, our own conversations with local and national campaign managers suggest that direct mail makes up a large portion of this spending and often the largest. Clearly, it is a tool frequently drawn from a campaign manager's toolbox. This trend promises to continue as parties become increasingly adept at collecting individual level data and mining it for insights which can then be leveraged through direct contact with individual voters (see Carty, Cross and Young, 2000; Gibson and Rommele, 2001; Norris, 2003, and for a more popular account Wells, 2006).

The importance of direct mail in general elections is probably surpassed by its importance in party leadership races. Whether conventions or direct elections, leadership races seem especially amenable to this campaign tool. These races are often paid little sustained attention by the media, especially for less competitive candidates. They tend to feature candidates who are often difficult to distinguish on ideological or policy grounds (Vavreck, Spiliotes and Fowler, 2002). Moreover, party leadership campaigns are increasingly large scaled affairs in which it is difficult for candidates to personally reach every member in the electorate through face-to-face meetings (Cross, 1996, 312). At the same time, the number of eligible voters (i.e. party members)

relative to the typical budget does make it possible to reach each voter by mail, often multiple times. Mail thus allows a candidate to speak directly to each party member or delegate. In races with many candidates, persuasion becomes a principal activity as campaigns seek to build coalitions which can deliver a majority of delegates or voters over a series of ballots. Direct mail plays an important role in this persuasion. Wearing's (1988) accounts of the 1976 and 1983 Progressive Conservative and 1984 Liberal leadership convention campaigns and Flanagan's (2003) account of the 2002 Harper campaign for the leadership of the Canadian Alliance provide convincing evidence of the importance that campaign managers assigned to direct mail in these races—an importance that we think generalizes fairly easily to all leadership races and elite politics more generally. Whether direct mail actually works, however, remains unclear.

In contrast to the political science literature, marketing is one field in which direct mail has been extensively studied. As a result, a substantial and broad literature exists. Among its findings, the marketing literature includes theory and knowledge about the elements of direct mail which make for success (eg Nash, 1984; Elsner, Krafft and Huchzermeier, 2004), how direct mail campaigns (especially coupons) affect purchasing (Bawa and Shoemaker, 1989; Bult and Wansbeek, 1995), how they affect incremental sales and how direct mail campaigns can be optimized based on past purchasing information (Allenby, Leone and Jen, 1999; Neslin, Henderson and Quelch, 1985). Moreover, much of this literature includes an experimental element.

For example, Irons, Little and Klein (1983) present a meta analysis of *sixty* field experiments on the effects of coupons on purchasing habits.

We can learn clear methodological lessons from this literature, particularly about the analytical power of field experiments. But despite this, it is unclear how much we can apply the lessons of consumer behavior to electoral politics. The decision to consume more goods or change the mix of goods that an individual consumes does not accurately reflect the nature of political choice in which a decision is forced (you have to vote at a certain time), zero sum (you have to vote for one candidate and not others) and essentially civic (in that one is likely, in making one's choice, to think about more than self interest or the meeting of a need). In short, individuals may bring a substantially different calculus to vote choice, one which is responsive in a different way—or not at all—to direct mail efforts. What is more, direct mail may vary systematically in its design from that in the commercial world. Accordingly, we look principally to evidence within politics and political science.

Whether in general elections or leadership contests, there is a lack of systematic evidence on the effectiveness of political direct mail. Examining direct mail effects using existing data is problematic for two reasons. First, even if we can assume that party and campaign spending is measured consistently and correctly (Ansolabehere and Gerber, 1994), its accounting is often not precise enough to identify direct mail outlays specifically (see Loewen, 2005, for a Canadian account). Second, even if we could observe the

different types of spending, we could not easily extract strong causal statements from these observations. While we discuss this at greater length in the next section, the basic problem is easily stated: Because spending decisions and communications strategies are not developed randomly, we cannot determine if the effects of campaign practices are a function of the types and extent of the method or the unobserved factors which influence campaigns to choose some methods over others. This problem is far from unique to Canada or Canadian political science. Indeed, there are many examples of observational research on campaign effectiveness that are confronted with this empirical problem (for British examples see Johnston and Pattie, 1998; Pattie, Johnston and Fieldhouse, 1995; Whiteley and Seyd, 1994, for American examples see Holbrook and McClurg, 2005; Vavreck, Spiliotes and Fowler, 2002). What is required is some form of inquiry not subject to Leamer's "inferential monsters lurking beyond our immediate field of vision" (1983, 83). That is, some form of inquiry where we can reasonably limit the number of possible explanatory variables and focus on one in particular—i.e. direct mail.

A growing line of research has sought to confront this problem of unobserved heterogeneity in campaign effects by engaging in field experiments. This research program has been both wide and deep. It covers several different campaign methods including direct mail, door-to-door canvassing, various telephone techniques and leafleting; and it reaches down into several types of elections, several different types of campaigns and several different

locales. The most important feature of these experiments is the random assignment of a treatment of interest to a well defined population, followed by a statistical analysis of the effects of the treatment (Green and Gerber, 2004, 11–22). While these studies have not been without criticism—particularly in terms of execution and estimation (eg Imai, 2005)—they have allowed for strong conclusions to be drawn on the effects of direct mail, especially as it relates to mobilization. Following Green and Gerber’s (2004) summary, while non-partisan direct mail seems to increase turnout, mail which expresses opposition to a candidate does not seem to have an effect. Partisan mail is effective in mobilizing partisans but not in bringing “swing voters” to the polls. On balance, the mobilizing effects of direct mail appear highly conditional and modest.

Less work has been undertaken on the persuasive effects of partisan direct mail. One early study examines the effects of a single candidate mailing in a weakly contested Democratic congressional primary (Miller and Robyn, 1975). It found no effect, though it was conducted over a rather small sample. Bositis, Baer and Miller (1985) conducted a unique experiment on timing and order effects in a Committeeman endorsement letter. Following up an election with a survey, they found persuasion to vary across message timing and order. Gerber (2004) single handedly expanded the field, conducting field experiments with five different campaigns during the 1999–2000 election cycle. These experiments—conducted during a mayoral race, a New Jersey state assembly election, a state legislative race in Connecticut, a Congress-

sional primary and a Congressional general election—examined the effect of campaign mailings on vote totals, which we take to be a test of the persuasive capacities of direct mail. In some cases, post-election surveys were used to estimate effects while in others they were measured by ward-level differences in vote totals. The results generally show that while incumbent mailings had little effect (except in primaries), challenger mailings were effective in some cases.

Taken together with the mobilization literature, it is difficult to arrive at a firm conclusion on the effectiveness of direct mail. Its utility is contingent both on the type of race and the type of candidate. As a consequence, these results do not directly inform our expectations of the persuasiveness of direct mail in a leadership race. However, they do demonstrate two things. First, we can effectively ascertain the causal properties of campaign methods through field experiments. Second, at least some of the claims of those who advocate direct mail appear to be false. The mobilizing capacity of direct mail has, at best, been overestimated by its advocates. Might it be the same for its persuasive properties?

While proponents of direct mail maintain that it serves to persuade voters to support the candidate sending the mail, there is evidence that attitudes and opinions can be resistant to such attempts, under certain circumstances (Knowles and Linn, 2003). As Peffley and Hurwitz (2007) point out, this is particularly the case when it comes to contentious issues that people hold intense attitudes about. In these instances it can be difficult to move opinions,

as attitudes can be resistant to attempts at persuasion.

There is also the possibility that arguments aimed at swaying individuals' opinions can have the effect of moving attitudes in the opposite direction to that intended by the argument. That is, receiving more information about, say, a candidate can make that candidate *less* attractive to certain voters. Chong and Druckman study the impact of competing "frames", or arguments, on opinion formation and find evidence of such *contrast* effects whereby, "weak frames will backfire in the face of strong competition by pushing the recipient further in the direction of the stronger frame than if he or she had been exposed only to the strong frame" (2007*a*, 7). Similarly, the results uncovered by Peffley and Hurwitz in their study of attitudes towards the death penalty among blacks and whites in the United States, are illustrative of similar *reactance* or *boomerang* effects (2007, 13). Both Chong and Druckman (2007*a*) and Peffley and Hurwitz (2007), as well as others (eg Johnson et al., 2003), make the point that such negative effects of persuasion attempts are most likely to occur among engaged, knowledgeable citizens; what Lodge and Taber refer to as *motivated reasoners* (2000). These people are those whom we would expect to latch onto confirmatory information while subjecting contradictory information to increased scrutiny in a manner which confirms their predispositions or increases their stock of negative considerations. We argue that delegates to a leadership convention are prime candidates for such a label. In many cases, they are long-time party members. They have likely invested significant time and money in securing their

spots as delegates. And they are likely to feel strongly about the candidates involved in the election.

Given delegates' levels of sophistication and commitment, the stakes attached to communicating information become all the higher for candidates. This is even more so if the candidate happens to also be a polarizing one. While Ignatieff was considered by most to be the clear front runner in the race, on many salient issues he adopted positions apart from the median of the activists of the Liberal Party. In his campaign material, including his direct mail, Ignatieff called for the eventual constitutional recognition of Quebec as a "nation", for the righting of the "fiscal imbalance" and for continued Canadian involvement in a war in Afghanistan. In addition, his support for the 2003 US led invasion of Iraq was well publicized in Canada. These were all positions outside the historical and recent mainstream of the Liberal Party. For example, while Ignatieff supported the extension of Canada's military mission in Afghanistan to at least 2009, this view was shared by only 36 percent of Liberal delegates from outside Quebec. Fifty-seven percent of delegates believed Canada's involvement should either end immediately or in 2007. Similarly, only 37 percent of delegates outside of Quebec supported a parliamentary resolution to recognize Quebec as a nation within Canada. Ignatieff's position for constitutional recognition was much stronger than a simple parliamentary resolution and thus likely even less supported (Strategic Counsel, 2006).

Aside from his support for the Iraq war, the policy positions articulated

by Ignatieff in the leadership race were not well known before the start of the campaign. Indeed, Ignatieff was better known for his record as an international human rights scholar and activist and an advocate of centrist social policy—positions much more in line with the Liberal Party mainstream. Thus, Ignatieff sought not merely to lead the party, but to move it in a certain direction. His candidacy needed to not only inform delegates of his views; he needed to persuade people to change their views. Thus, clearly communicating a policy direction outside the party mainstream was a risky strategy. It is in this respect that Ignatieff's gamble provides interesting and unique insight into questions of strategic communication.

A growing theoretical literature focuses on the strategies of communication available to leaders (Zeckhauser, 1969; Shepsle, 1970, 1972; Aragonés and Neeman, 2000; Meirowitz, 2005; Dewan and Myatt, 2008). There are a number of dimensions along which such communication can be characterized. For instance, one might think of communication as varying between degrees of clarity and obfuscation. That is, a leader can deliver her message with varying levels of precision. Moreover, a leader's judgment about policies—the substance of her communication—can reflect more or less “sense of direction” (Dewan and Myatt, 2008, 2).

It might seem obvious that a leader is better off when she communicates clearly. However, numerous scholars have noted that ambiguity and obfuscation have strategic advantages (eg Zeckhauser, 1969; Shepsle, 1970, 1972). Shepsle points out that, “. . . observed ambiguity often typically involves pre-

cisely those issues on which the election hinges” (1972, 555) and underlines “the politician’s advantage in speaking ‘half truths’ and in varying his appeals with variations in audience and political climate” (1972, 559). Aragonés and Neeman (2000) provide a model in which candidates opt for ambiguity in order to remain flexible and because it allows them to broaden their appeal. Dewan and Myatt follow in this vein and argue that “attention-seeking leaders will intentionally obfuscate” (2008, 9).

As we have argued, Ignatieff presented a set of policies that deviated from the Liberal mainstream; and he did so with some measure of clarity. There was no mistaking where he stood on controversial issues such as Afghanistan or national unity. And, as it turned out, delegates were not convinced by his “judgment” or “sense of direction” (Dewan and Myatt, 2008). The theoretical literature on leadership communication and strategic ambiguity suggests that leaders such as Ignatieff ought to equivocate and obfuscate in delivering their message. Dewan and Myatt’s (2008) model illustrates that when communication skills are endogenous (leaders can manipulate the clarity of their message), the relative influence amongst the best communicators is greater for those with lower variance in their judgment. In other words, in the case of the Liberal leadership candidates, while Ignatieff was certainly a good communicator, his appeal was circumscribed by the perception among delegates that he would take the party in the wrong direction. From this arises our empirical question: when elite voters are confronted with a controversial position via direct mail, do they become more likely to support a candidate

or less likely? That is, does direct mail persuade, or does it merely inform and lead to possibly negative effects?

After describing the race for the Liberal leadership, we turn to an examination of whether Ignatieff's gamble paid off. That is, whether direct mail had the desired effect of persuading voters or if voters in this election were resistant to the information presented in campaign mail—or worse, susceptible to reactance.

4.3 The Race

After losing the January 2006 federal election, Prime Minister Paul Martin resigned as parliamentary leader of the Liberal Party of Canada. In the subsequent weeks the party outlined the conditions of its leadership selection process—much of which was predetermined by the party's constitution. A leadership convention was held in Montreal on December 3, 2006. Delegates to the convention were elected from among party members. In addition to ex-officio delegates who were guaranteed a place at the convention, the regular delegates from each federal electoral district were allotted to leadership candidates according to the total preferences of all members in that electoral district. On the first ballot non-ex-officio delegates were thus obliged to vote for the candidate to whom they were pledged. Indeed, they received marked ballots upon their arrival at the convention. This apportionment process occurred at a "Super Weekend" at the end of September. Only those party

members who were of good standing as of July 1, 2006 were allowed to vote in the Super Weekend.

The race was nothing if not exciting. More than twenty names were identified as potential candidates and eleven officially entered. By the time of the delegate selection meetings the field had narrowed to eight candidates. Michael Ignatieff was the clear front runner, obtaining the support of about 30 percent of pledged delegates, as well as many ex-officios (see Table 4.1). Ignatieff, recently returned from more than twenty years outside the country as an academic and journalist, was generally seen as being on the right of the party. He was a polarizing candidate. Bob Rae, a former Premier of Ontario (as the leader of the social democratic New Democratic Party), was the clear second place candidate. He could also be regarded as polarizing. Rounding out the top four were Gerard Kennedy, a former Ontario provincial cabinet minister, and Stéphane Dion, a former federal cabinet minister (and political scientist) known far more for intellectual battles with sovereigntist/separatist leaders in Quebec than for his political *panache*. The bottom four comprised Ken Dryden, Joe Volpe, Scott Brison and Martha Hall Findlay.

To the surprise of many, Dion would eventually win the leadership. Results from the pre-convention delegate selection and the four ballots at the convention are presented in Table 4.1. Dion finished in third place on the first ballot, just two delegates ahead of Kennedy. He would receive Kennedy's endorsement after widening his lead on the second ballot. On the strength

Table 4.1: 2006 Liberal Party leadership election results (%)

Candidate	Pre convention delegates		1 st ballot		2 nd ballot		3 rd ballot		4 th ballot	
Ignatieff	1,377	(29.3)	1,412	(29.3)	1,481	(31.6)	1,660	(34.5)	2,084	(45.3)
Rae	943	(20.1)	977	(20.3)	1,132	(24.1)	1,375	(28.5)		
Kennedy	820	(17.5)	854	(17.7)	884	(18.8)				
Dion	754	(16.1)	856	(17.8)	974	(20.8)	1,782	(37.0)	2,521	(54.7)
Dryden	238	(5.1)	238	(4.9)	219	(4.7)				
Volpe	226	(4.8)	156	(3.2)						
Brison	181	(3.5)	192	(4.0)						
Hall Findlay	46	(1.0)	130	(2.7)						
Undeclared	112	(2.4)								
Total votes	4,697	(100.0)	4,815	(100.0)	4,690	(100.0)	4,817	(100.0)	4,605	(100.0)

of that hand-tipping he would vault past both Ignatieff and Rae on the third ballot, thus eliminating Rae. He defeated Ignatieff on the fourth and final ballot, receiving 54.7 percent of the votes to Ignatieff's 45.3 percent. Rather than polarizing delegates, as the two front-runners had, Dion was successful in portraying himself as a safe second-choice. Whether by luck or design, he appeared a master of convention politics.

Our experiment was situated within the period between the election of delegates and the convention in Montreal, what Wearing calls the "second stage" of delegated conventions (1988). This period provided a crucial test of the persuasive ability of campaigns. Rather than selling memberships and encouraging supporters to stand as delegates, campaigns in this period of the process were dedicated to ensuring delegates attended the convention; and, crucially, to persuading delegates for other candidates to select their candidate as their next choice should their preferred leadership candidate fall off the ballot or withdraw. Among many tactics, direct mail played an

important role in this critical period. For example, the Dion campaign sent a DVD featuring a series of short interviews with their candidate. The Ignatieff campaign sent a 40-page bilingual policy book entitled “Agenda for Nation Building: Liberal Leadership for the 21st Century.”¹ The book outlined in unusual detail Ignatieff’s policy on the economy, the environment, the constitution, national unity and foreign affairs.² The Ignatieff campaign also sent out a simple color brochure summarizing Ignatieff’s positions. As we outlined above, Ignatieff’s stand on many of these issues was in contrast to the prevailing opinion within the Liberal Party.

4.4 The Experimental Study

Our experiment consisted of two components: First, a randomized program of direct mail from a front-running campaign conducted over a subset of elected delegates in the last week of October and first week of November, 2006. Second, an academic mail-back survey of the same delegates which measured, among other things, their likeability evaluations of each candidate as well as their preferences between the various leadership candidates. We describe each in more detail below.

¹The title on the French side of the book was “Bâtir notre nation: le leadership libéral pour le 21e siècle.”

²Electronic versions of both documents can be seen at http://www.politics.ryerson.ca/rubenson/downloads/Ignatieff_book.pdf and <http://www.politics.ryerson.ca/rubenson/downloads/english.pdf>.

4.4.1 The Experiment

Our experiment relied on a partnership with the Michael Ignatieff campaign. After the selection of delegates at the end of September, we randomly selected a subset of 800 delegates from those who had a current address on the official party list of delegates.³ In addition to restricting our sample to those delegates who had addresses, we also excluded Quebec, Manitoba and British Columbia.⁴ Among these 800 delegates, we identified those who had not pledged to support Ignatieff at delegate selection meetings, reducing our sample to 567. Among these remaining delegates, we randomly assigned 100 to receive two pieces of mail from the Ignatieff campaign and 200 to receive one piece of mail.⁵ All those who received mail received a copy of Ignatieff's 40-page policy book in the last week of October. Those who were assigned to receive a second piece of mail also received a copy of a color brochure in the first week of November. This material was developed by the campaign and was identical to that sent to all delegates not included in the experiment. By randomly assigning mail we (theoretically) ensured that the reception of mail was not a function of a respondent's personal characteristics or preferences. As with conventional random assignment in a laboratory, this affords us much analytical leverage.

³This represents approximately 16 percent of delegates.

⁴Delegates from Quebec were excluded as they were subject to a different ad campaign by the Ignatieff campaign. Manitoba and British Columbia were excluded from the party's delegate list at the time because of incomplete delegate lists or disputes between several campaigns over the status of various delegates.

⁵We describe our treatment assignment procedure in more detail in Annexe E.

4.4.2 The Survey

One week after the second wave of mailing, we mailed each delegate within our subset an academic survey from the Department of Politics at Ryerson University. By sending the survey under the cover of the University, we concealed any connection between the survey and the experiment. Moreover, we excluded any mention of the field experiment on our respective academic websites. The survey included a postage-paid return envelope, as well as an ethics disclaimer and short introduction. The survey obviously made no mention of the experiment, though it did include recall questions on the reception of direct mail from campaigns since the selection of delegates. Most pertinent to our study, the survey included questions about preferences for, and evaluations of candidates, which allowed us to test the persuasive effects of direct mail.

The advantages of combining a survey and an experiment become clear when we consider the typical alternative approach to studying the impact of campaigns on individual voters, i.e. a survey which may or may not include contextual information about the campaign (for leadership campaign examples see Perlin, 1988; Stewart, 1997; Vavreck, Spiliotes and Fowler, 2002; Bartels, 1987). As Gerber and Green (2000); Green and Gerber (2004) have argued, relying on a survey alone to gage the effects of direct mail—and other campaign contacts more generally—suffers from two problems. First, individual respondents are demonstrably poor at recalling whether or not they have received mail from a campaign. For example, our survey included

a recall question which asked delegates to identify from which campaigns they had received mail since the conclusion of delegate selection meetings. Because we know which delegates received mail from the Ignatieff campaign we were able to measure the level of error in delegate recall. Of those who did not receive mail from the campaign, 85 percent correctly recalled that they received no mail. However, 15 percent did report receiving mail. The case is more grave with those who did receive mail, with less than two-thirds (64 percent) correctly recalling receiving mail. Moreover, based on a question-wording experiment embedded in our survey, we found that recall was not improved by giving some delegates a further prompt identifying the types of mail they may have received.⁶ Accordingly, even with a carefully designed survey we would risk serious measurement error in identifying who received direct mail from a campaign. Our study avoids this pitfall because we know to whom the campaign sent mail.⁷

Second, political campaigns are often strategic in their targeting of direct mail. Mailings are targeted and tailored to reflect a campaign's beliefs about the recipient. For example, campaigns may be more likely to send mail

⁶The question asked "Do you recall receiving any postal mail (snail mail) from any of the campaigns since the Delegate Election Meetings at the end of September?" The question wording experiment then added "For example, have any campaigns sent you mail soliciting support for later ballots, or telling you about events that their candidate is holding in your area?"

⁷We also know that the mail at least landed in their mailbox. We used the same addresses for the surveys as for the mail, so it is not possible that a delegate received and responded to our survey without receiving the mail. While delegates may very well be selective about what they choose to read (see Barlett et al., 1974), we can be certain that we are at least dealing with cases in which they had the opportunity to read the mail sent to their address.

to those whom they believe are at least open to supporting the campaign. By merely observing the relationship between direct mail and behaviors, we cannot know whether any direct mail effect is the result of the mail itself or the individual in question being predisposed to support the party or candidate. Even with a bevy of control variables this problem cannot be easily solved statistically, if at all (Gerber, Green and Kaplan, 2004). However, in the case of our experiment we know that the assignment of mail was random and thus uncorrelated with individual characteristics. Any observed effect of mail on leader ratings or preference orderings is likely the result of mail.

In comparison to the experiments outlined above, one caveat is in order. Those experiments typically test the effects of a treatment—direct mail for example—on a directly observable behavior such as voting as determined by an official record. Our experiment, by contrast, still relies on estimates of an effect drawn from a survey. As we could not peer inside the ballot boxes at the party convention, we are left to ascertain the effects of direct mail through our survey questions. We are thus left open to many of the problems associated with survey responses. However, we are not confronted with the more fundamental problems of respondent recall or the strategic allotment of a treatment. As a result of this, the effects we observe are “real” to the extent that surveys capture “real” aspects of delegates’ considerations and evaluations in the run up to the convention.

Our final sample includes 161 respondents, a response rate of 28 percent.

This sample is evenly balanced between those who did receive mail (81) and those who did not (80). Treatment is unrelated to the pledged support of delegates ($\chi^2 = 7.78$, $p = 0.35$), province of residence ($\chi^2 = 3.73$, $p = 0.81$), or delegate type⁸ ($\chi^2 = 8.95$, $p = 0.26$). Most importantly, survey response is unrelated to our 3-category treatment assignment ($\chi^2 = 0.61$, $p = 0.74$).

4.5 Results and Discussion

In a leadership race such as the one we study here, direct mail has two principal aims. First, campaigns want to make their candidate more likeable while at the same time decreasing voters' positive evaluations of rival candidates. Second, and more important, in multi-ballot elections, campaigns want to persuade voters to shift their support to the campaign's candidate. In other words, the aim is to convince voters to change their preference rankings of candidates. We examined the effects of direct mail across two different measures in order to assess its effectiveness in achieving each of these aims. In each case, we compare those who did and did not receive mail using relatively simple models (Achen, 2002).⁹ Rather than specifying complicated models, we rely on the power of random assignment.

We first measure whether those who received direct mail evaluate the like-

⁸Delegates are classified by the party according to gender, age and aboriginal status.

⁹Our treatment regime specified that some individuals receive two pieces and others one piece. Because of our relatively small n , we have collapsed these two treatments into one in the analysis. Our substantive results do not change when we consider those who received two pieces of mail separately. We note again that all those in treatment received the detailed policy book.

ability of the eight candidates differently than those who did not. The expectation of those sending direct mail—at least for the campaign in question—was that mailers would increase positive evaluations of their own candidate and reduce positive evaluations of other candidates. Table 4.2 reports the results of *t*-tests on differences in the mean rating of candidates with and without mail. We use a conventional 0-100 rating scale. Initially we only considered one-sided hypotheses in the direction expected by the campaign. That is, Ignatieff mail should make Ignatieff more likeable and other candidates less likeable. As is clear from the results in Table 4.2, we find little evidence of such positive effects for direct mail. Only in the case of Ken Dryden is the test statistically significant and in the expected direction. Receiving mail from the Ignatieff campaign appears to have caused delegates to reduce their positive evaluations of Dryden. However, those who received mail did not give higher ratings to Ignatieff, on average. Moreover, some of the results are statistically significant in the opposite direction to that anticipated by the campaign. In the case of Dion, Brison and Volpe, it appears that direct mail from the Ignatieff campaign *increased* the likeability of these candidates. On the whole, receiving mail did not move the opinions of those who were not already pledged to support Ignatieff in the expected and desired direction.¹⁰

¹⁰We have also estimated these effects with separate OLS regressions for each candidate with leader rating on the lefthand side and mail and a small number of control variables on the righthand side. Our results do not change.

Table 4.2: Effects of Ignatieff mail on average leadership candidate ratings^a

Candidate	Mean rating with no mail (s.d.)		Mean rating with mail (s.d.)		<i>p</i> -value ^b	N Mail (No mail)
Michael Ignatieff	46.3	(32.6)	47.8	(33.7)	0.38	75 (80)
Bob Rae	61.4	(33.6)	63.6	(33.9)	0.66	75 (81)
Gerard Kennedy	73.6	(26.1)	73.9	(26.7)	0.53	74 (80)
Stéphane Dion	72.6	(23.7)	77.3	(21.4)	0.90	75 (79)
Ken Dryden	60.7	(26.0)	54.7	(26.9)	0.08	72 (79)
Joe Volpe	15.5	(22.6)	20.8	(25.5)	0.91	70 (79)
Scott Brison	42.4	(27.5)	54.3	(26.1)	0.99	67 (76)
Martha Hall Findlay	49.5	(27.0)	49.7	(27.9)	0.51	65 (78)

^a Note: Calculations of difference rely on unpaired *t*-tests with an assumption of unequal variance.

^b For Ignatieff ratings, the *p*-value is from the onesided alternative hypothesis that $\Delta > 0$; for all other candidate ratings the alternative hypothesis is $\Delta < 0$.

As a consequence of the multiballot nature of a competitive delegated convention, moving a candidate up in delegates' preference rankings is a principal objective for campaigns. Indeed, a particular feature of a contest such as this one is that candidates have little choice but to communicate with delegates supporting rivals in order to entice them to change their mind on later ballots, contrary to general elections where a candidate can choose to not speak to a large portion of the electorate. Given that the final ballot pairing in this race was far from obvious, campaigns were compelled to send mail to all delegates. For a potentially polarizing candidate such as Ignatieff this presents a dilemma. He would want to get his message out but that message may in fact be damaging to him among delegates who have an antipathy toward him.

Table 4.3: Effects of Ignatieff mail on delegates' preference ordering^a

Variable	Model 1		Model 2	
	Odds ratio	<i>p</i> -value	Odds ratio	<i>p</i> -value
Ignatieff mail	0.642 (0.156)	0.07	0.590 (0.160)	0.05
Attention to the race			0.843 (0.046)	0.00
Interest in the race			1.169 (0.124)	0.14
Senior			1.018 (0.262)	0.95
Youth			1.586 (0.791)	0.36
Female			1.043 (0.321)	0.89
Cut 1	-0.524 (0.157)		-0.467 (0.688)	
Cut 2	0.583 (0.143)		0.659 (0.733)	
Cut 3	1.186 (0.159)		1.275 (0.734)	
Wald	3.26		34.90	
Prob > χ^2	0.071		0.000	
N	160		160	

^a Note: Odds ratios are from ordered logit models, clustering on province; robust standard errors in parentheses.

Our results suggest that receiving direct mail from the Ignatieff campaign seems to have done little to achieve the objective of moving Ignatieff up in delegates' preference orderings. Quite the opposite, receiving mail appears to have moved Ignatieff down in the preference rankings of some delegates. Table 4.3 presents results from two ordered logit models, both of which take Ignatieff's position in a delegate's preference ranking as the dependent variable. We constructed the variable from three questions. The first asked delegates to identify their second choice. The second asked delegates to identify their third choice. The final question asked delegates to identify any candidates for whom they would never vote. We are thus left with four categories: Never Choose \rightarrow Ambivalent \rightarrow Third Choice \rightarrow Second Choice. The first model includes only a dummy variable indicating whether the delegate received mail from the Ignatieff campaign.¹¹ The second model adds a 0 to 10 measure of a respondent's interest in the campaign, a 0 to 10 measure of the respondent's attention paid to the campaign and dummy variables indicating whether the respondent was a youth delegate, a senior or a female delegate to improve the precision of our estimates.¹²

The results in our first model suggest a negative effect of mail on preference orderings. The odds of making Ignatieff second choice versus all the other options are 35.8 percent lower for those who received mail than for

¹¹We present robust standard errors calculated over provincial clusters. As campaigns were organized provincially, we want to control for unobserved differences across provinces.

¹²Interest and attention appear unrelated to the reception of mail. Interest: $\beta = -0.35, p = .20$; Attention: $\beta = -0.24, p = .26$.

those who did not receive mail. After controlling for attention and interest paid to the race and delegate demographics, the reception of mail continues to have a strong negative effect on the vote choice: For delegates who received mail, the odds of making Ignatieff their second choice over all other options are 41 percent lower than for those who did not receive the mail. Our second model provides a better fit of the data, a more accurate classification of cases and a less ambiguously significant effect for direct mail. It is an effect, however, quite contrary to the campaign's expectations.

Taken together, these results lead us to a clear conclusion. In the face of crystallized preferences, receiving one or two mailings from a campaign was not enough to positively alter delegates' assessments or intentions. Rather, if it had any systematic effect it was in making delegates more negative towards Ignatieff's candidacy. This finding is consistent with recent work highlighting *contrast* (Chong and Druckman, 2007a) or *boomerang* effects (Peffley and Hurwitz, 2007; Haider-Markel and Joslyn, 2001; Johnson et al., 2003). These studies argue that when individuals are motivated, engaged and hold intense prior views on issues and candidates, attempts at persuasion can backfire. By merely informing voters, Ignatieff gave them more reasons to vote against him.

In the case of Ignatieff's mail, there was much on which motivated reasoners could take hold. His positions on foreign policy, the constitution and fiscal federalism were well outside of the mainstream of the party he was seeking to lead. Presented with clear evidence of this, delegates who harbored

prior neutral or negative dispositions about Ignatieff may have become even less disposed to his candidacy. This finding should give campaign managers pause. Political direct mail is a communication of a message which can have three effects. It can increase the appeal of a candidate; it can have no effect; or it can decrease the appeal of a candidate. If direct mail makes clear positions or attributes which voters find objectionable, it may have such a negative effect. For a candidate as polarizing as Michael Ignatieff, this final outcome appears to have been very real.

4.6 Conclusion

Given the mixed evidence on the mobilizing effects of direct mail and given the lack of evidence of positive persuasion effects, why do we observe campaigns devoting substantial resources to this tool? We have three explanations. First, campaign operatives are certain that these tools work. This message is rather consistently delivered in trade publications such as *Campaigns and Elections* and in operative training sessions such as the “universities” which Canadian parties hold prior to elections. It only makes sense to use these tools, given the received wisdom. Second, it is not difficult to talk oneself into believing that a chosen campaign tool is working despite a lack of evidence of positive effects or evidence to the contrary. In the hubbub and stress of a campaign an operative will look for any affirmation that things are on the right track. A positive comment about direct mail can

quickly become enough to convince one of larger effects. Similarly, it is easy to become convinced of the importance of direct mail, when one knows it is being used by other campaigns. A third possibility exists—one which is less pessimistic about the analytical abilities of campaign managers. Even if direct mail were known to have very small effects, it may still be the most efficient use of resources. Volunteers cannot be bought, professional call centers and automated calls are demonstrably inefficient, a candidate can only work phones or shake hands a certain number of hours each day and time cannot be stretched. The implication is that a campaign which did not spend its remaining money on direct mail may not be able to spend it at all. Moreover, direct mail can be sent at a relatively low cost and can often be easily scaled up into repeated or more substantial mailings. Indeed, once a campaign has settled on a message and obtained a list of voters, the marginal cost of mailing consists only of the cost of producing materials and postage. Knowing this, why would a campaign not spend whatever extra resources it had on printed material? Perceiving that direct mail has some effect, knowing that it is widely used in other campaigns and being able to send it rather economically, what campaign manager could be expected to take the risk of not sending the mail?

We think a similar logic holds when explaining why Ignatieff would communicate such controversial positions. While a *post-hoc* analysis suggests that his positions were controversial and costly, the campaign may not have been able to conclude this during the course of the election. Having achieved

a front-runner status on the strength of his organization and the appeal of his personality, Ignatieff's campaign may have wrongly concluded that he was capable also of moving delegates' on a core set of issues. Or, they may have misread the appeal and popularity of these issues to begin with. Either way, only a more strategically adroit candidate, faced with the possibility of remaining silent and losing, could be convinced not to communicate positions which he believed to be both correct and compelling. Like the fall of a kingdom for the want of a nail, who would risk the loss of a campaign for the want of mail? Our results would suggest that even if direct mail is the most sensible expenditure given resource and time constraints, it may be a message better left unsent.

Chapitre 5

Testing the Power of Arguments with a Bradley-Terry Model (with Daniel Rubenson and Arthur Spirling)

Under review at Public Opinion Quarterly. Loewen conceived of the experiment. The experiment was designed by Loewen and Daniel Rubenson. Arthur Spirling developed the statistical model. Spirling and Loewen performed the Bradley-Terry analysis. Loewen performed the conventional analysis. Loewen wrote the first draft of the article and Spirling wrote a second draft. Spirling, Rubenson and Loewen all completed the final drafts.

5.1 Introduction

Public opinion research and political science have long been centrally concerned with assessing the persuasive power of arguments (see, e.g., Aristotle (322BC/1991) for an early treatment, and e.g. Mutz, Sniderman and Brody (1996) for a more recent discussion). For example, does a call for greater social welfare spending which appeals to a respondent's sense of equality elicit greater support than an appeal based on self-interest? Or, in a more oppositional manner, if a liberal argument for higher taxes is pitted against a conservative argument for lower taxation, which garners more support? By exposing respondents to various arguments in a controlled setting, public opinion research can tell us a lot about the foundations of preferences, as well as rhetorical and political success in real world arenas.

Elections are perhaps the most obvious arena in which issues and arguments compete (see, e.g., Page, 1978; Erikson and Wright, 1997; Ansolabehere, Snyder and Stewart, 2001). However, in this case, issues and arguments are often 'bundled' together such that it becomes difficult to determine the effects of issues and arguments independent of candidates. A more systematic discussion of positions is seen in *referendums*—formal plebiscites in which voters directly convey their preferences on a binary choice that may concern anything from a new policy commitment to potentially profound constitutional changes. Since so much is at stake in such national reform votes, political scientists and public opinion researchers have an ob-

vious interest in attempting to understand *why* citizens choose as they do. In some senses, this exercise ought to be empirically easier for referendums since, relative to elections, such votes are depersonalized: the candidate is an inanimate *proposition*, rather than an individual (or collection of individuals). Voters thus receive more focused and explicit discussions of the *issue* at stake, with party and other (partisan) loyalties sometimes at cross-cutting angles.¹ As a result, scholars have direct access to the explicit arguments and types of reasoning used to buttress a position, rather than having to piece together the more rhetorical, nebulous and eclectic sources of appeal that a particular party or politician uses. Using such political arguments and voter responses—observationally or experimentally—analysts ought to be able to teach us much about the outcomes of these events.

Unfortunately, analysts of public opinion face pronounced methodological problems when attempting to assess the ‘power’ of different political arguments. First, in observational studies, it is not always apparent precisely *which* arguments voters have received. Objectively measuring probable exposure is difficult and asking voters to self-report their treatment is problematic. Another solution is to take an *experimental* approach within a survey, and have subjects receive one of k total contrasting arguments, followed by a report on which argument they agreed with. For increasing values of k ,

¹Consider, for example, the UK’s EEC referendum vote of 1975: Prime Minister Wilson and most of his cabinet united with the majority of Margaret Thatcher’s Conservatives to recommend a ‘yes’ vote. Meanwhile, a left-wing contingent of Labour MPs—and some right-wing Tories—campaigning for a ‘no’.

however, the size of the relevant sample on which inferences are based is quickly reduced. Alternatively, exposing each subject to a large number of arguments and then soliciting opinions on which is ‘most convincing’, risks a learning or cumulative effect where responses are a function of previous and current treatments in the series, rather than the content of the argument *per se*.

These two problems are exacerbated if we wish to pit particular arguments against one another in order to mimic the way that debates take place in practice. Consider the case where we have m arguments in favour of a position, and the same number against. Randomly assigning a total of T respondents to receive one pro argument and one con argument yields just $n = \frac{T}{m^2}$ subjects per treatment. For a sample of, say, 500 individuals picking between, say, 6 pairs of arguments, $n < 15$: that is, there are less than 15 respondents per condition! Clearly then, whether using conventional comparison of means tests or logistic regression, statistical power is an issue in this set-up. And, of course, assigning multiple treatments risks the learning/fatigue effect noted above.

We suggest an innovative solution to this problem: a novel application of the Bradley-Terry model of pairwise comparisons (Bradley and Terry, 1952). This model has a unique value proposition. It can give survey researchers a comparatively large amount of information in a small amount of survey space. By way of example, we use a survey experiment of just 520 respondents in which respondents are assigned one of six arguments for and six against

electoral reform (creating 36 treatment groups): We show that both means-test and traditional logit methods fail to give us clear insights into the power of arguments for and against reform. However, by conceptualizing arguments as contestants engaged in bouts with other arguments, we can use the model to estimate the probability that an argument wins. We can then rank-order arguments according to their power to persuade. Moreover, we can also determine the sources of an argument's power, i.e. the components which make it more persuasive than another argument. We will show that the utility of this method exceeds more traditional approaches. Our principal motivation, then, is to introduce these models to public opinion researchers.

Substantively, we are motivated by the question of which arguments for and against *electoral reform* are most powerful. To this end, we conducted a survey experiment in the midst of a province-wide referendum on whether to change electoral systems in Ontario, Canada. This new data consists of randomly assigned respondents receiving one of six arguments in favor of the existing First Past the Post (FPTP) electoral system and one of six arguments in favor of the proposed Mixed Member Proportional (MMP) system. By knowing which arguments influence a voter to choose one electoral system over another and by knowing the sources of their power we can gain important insight into support or opposition to electoral reform. By understanding *why* certain arguments are more powerful, we also contribute to the literature on 'framing' in public opinion.

Ontario is one of many jurisdictions that have put the question of re-

forming the First Past the Post system to a referendum. Citizens in Ireland (1958), New Zealand (1993), Italy (1991, 1993) and other provinces of Canada (British Columbia in 2004, and Prince Edward Island in 2005) have been asked for their preferences on the rules that turn votes into seats. Outcomes in referendums such as these have the potential to fundamentally alter the structure of a polity, the number of parties, the power of its elites and the representation of minority groups (see, e.g., Gallagher, 1998; Norris, 1997; Vowles et al., 1998; Horiuchi and Saito, 2003). More generally, referendums and initiatives play an important and often major role in several modern democracies. As such, analysts have devoted considerable attention to the desirability and effects of this device (see, e.g., Hamilton, 1970; Vanderleeuw and Engstrom, 1987; Johnston et al., 1996; Smith, 2001; Lupia and Matsusaka, 2004; Matsusaka, 2005) and it is not difficult to identify controversial, newsworthy and wide-ranging examples of this application of direct democracy.² Our secondary motivation, then, is to use this model to gain useful insight into the substantively meaningful question of support for electoral reform.

We proceed as follows. In the next section we describe our research design

²Instances include Britain's *European Communities membership referendum* of 1975 that determined the UK would remain in the ECC and similar recent referendums in several other European countries; California's *Proposition 13* of 1978 that severely limited property taxes in that state; the Charlottetown Accord of 1992 that would have divided federal and provincial powers in Canada but was ultimately defeated; the *1999 Australian referendum* that retained the monarchy; various referendums in the Republic of Ireland that dealt with laws concerning abortion (1992, 2002), divorce (1995) and the death penalty (2001); and the 2005 vote that approved a new constitution for a democratic Iraq.

and the substantive context in which it was conducted. We also draw some connections to the large ‘framing’ literature in public opinion. In Section 3 we introduce our novel econometric approach for pairwise comparison, the Bradley-Terry model. In Sections 4 and 5 we describe our results and compare and contrast the utility of traditional analysis with our method. We conclude by suggesting other questions to which this model might be applied in the analysis of public opinion.

5.2 Data, Context and Connections

Our experiment occurred within the context of a referendum on electoral reform conducted in the province of Ontario, Canada, in October 2007. The referendum was conducted concurrently with a provincial election. In the previous provincial election, the Liberal Party of Ontario had pledged to hold a ‘Citizens’ Assembly’ to consider the question of electoral reform (Cross, 2005, 77). Making good on their promise, the government called together 104 citizens—one randomly-selected from each electoral district plus a chair—in the summer of 2006 and tasked them with deliberating about electoral systems and possibly recommending a shift from Ontario’s First Past the Post (FPTP) system. Any recommendation would be put to a referendum and a change in systems would require a 60% majority and a majority of support in 60% of electoral districts. The Assembly eventually recommended the move to a Mixed Member Proportional (MMP) system, similar to that

in place in Germany and New Zealand.

The First Past the Post system is perhaps the most simple of all electoral systems, and some commentators cite this feature as one of its best qualities. Citizens vote for candidates in local constituencies and the candidate with the most votes (the plurality) wins the constituency's seat. To win election, then, candidates must have some degree of local appeal in a clearly defined constituency—again a distinct advantage of this arrangement according to some pundits. The party with the most seats (almost always) forms the government. As FPTP systems most often result in single-party majority governments, the party in power can exercise an almost exclusive influence over legislation until the time of the next election (Blais, Forthcoming). Some argue that this allows for a more 'decisive' polity with clear lines of electoral responsibility: large parties are over-represented relative to their vote shares and small parties may be entirely excluded from legislative and executive representation (see Powell, 2000, for a discussion). Ontario has operated under a FPTP system since 1792.

Mixed Member Proportional systems incorporate elements of the FPTP system and proportional representation systems, which had previously been the principal reform alternative to FPTP systems (Blais, Forthcoming). MMP systems elect some share of legislators from local constituencies in a plurality fashion identical to FPTP. However, the remaining share of representatives are allotted to parties according to a party vote such that the overall composition of the legislature is proportional to each party's support. These

legislators are drawn from lists composed by the parties. In the proposed system, voters would thus cast two ballots, one for a local representative and one to indicate their party preference. The result is a mix of some local representation and greater proportionality, with parties playing a more central role. Some have argued that MMP represents a compelling ‘middle way’ in this regard (Dunleavy and Margetts, 1999). According to the Ontario proposal, ninety representatives would be elected locally and thirty-nine would be elected from party lists.

The Ontario MMP proposal lost the ensuing referendum, garnering just 38% of the vote and a majority in only five districts. It was not a loss attributable to the proposing government, as they won the concurrent election rather overwhelmingly. Thus, this loss was much to the consternation of electoral reform advocates who believed that the arguments for electoral reform were clearly superior to those in support of the current system. Indeed, in the days following the referendum, advocates claimed that had the public been better educated on the proposed system, especially its underlying values, then support would have been higher, perhaps surpassing the 60% threshold required for reform (Fenlon, 2007).

To put to the test claims that the arguments for the MMP system are more powerful upon exposure than those for the FPTP system, we developed six arguments in favor of each position, drawn from campaign materials produced by both sides of the campaign, conversations with advocates on both sides and academic literature on electoral systems (e.g. Blais and Massicotte,

2002; Powell, 2000). We describe these arguments in the next section.

5.2.1 Experimental Design

To assess the comparative power of arguments, we conducted an online survey experiment with 520 voting-aged Ontarians in the last week of the referendum campaign. Our experimental module was contained at the end of the survey, after three sections related to federal politics, provincial politics and environmental issues. No questions prior to the experimental module were related to electoral reform. A profile of our subjects is found in Annexe F.

In the experiment, subjects read: “As you may know, there will also be a referendum during the October 10th election. The purpose of the referendum is to determine whether Ontario should change its electoral system from the current first past the post system to a mixed member proportional system. We’d like to present you with an argument for each system and then get your view on which system you prefer.” Respondents were then presented with one of six arguments in favor of the existing system and one of six arguments in favor of the proposed system. The order of the arguments was also randomized. After receiving the arguments, respondents were then asked to indicate their preference between MMP and FPTP; they were not able to give a ‘Don’t know’ response and had to indicate a preference before proceeding to the end of the survey.

The six arguments for FPTP and MMP were:

- A first past the post system is better because it creates strong majority governments that can implement their policies. (FPTP1)
- A first past the post system is better because it makes sure that every Member of Provincial Parliament is elected from a constituency. (FPTP2)
- A first past the post system is better because one ballot is less confusing than two. (FPTP3)
- A first past the post system is better because it allows local party members to choose all of a party's candidates. (FPTP4)
- A mixed member proportional system is worse because it will lead to unstable coalition governments. (FPTP5)
- A mixed member proportional system is worse because it puts too much control in the hands of parties and political elites. (FPTP6)

And,

- A mixed member proportional system is better because it makes sure that parties that have support of some of the population still get some representation. (MMP1)
- A mixed member proportional system is better because it lets voters indicate their preference for both a local representative and a party. (MMP2)

- A mixed member proportional system is better because parties should get the same share of seats as their share of the vote. (MMP3)
- A mixed member proportional system is better because it will lead to more diversity in the legislature. (MMP4)
- A first past the post system is worse because it gives some parties a share of the seats much larger than their share of the vote. (MMP5)
- A first past the post system is worse because it shuts out small parties. (MMP6)

As the pro and con arguments were assigned randomly and independently, we have 36 approximately equal sized groups.³

We are not only interested in the power of arguments, but also their sources of power. In other words, what is it about the content of arguments that makes some more powerful than others? We note that several arguments share common components. For example, some arguments reference local candidates, others appeal to notions of proportionality or fairness, others reference government stability and effectiveness, others highlight the strengths or weaknesses of the ballot structure in FPTP and MMP. Many make reference to political parties, while another refers to increased diversity in the legislature as a result of MMP. Some arguments are phrased in positive terms (i.e. a first past the post system is *better* because...), while others

³A χ^2 test indicates that the assignments are independent of one another ($\chi^2 = 26.41$ $p = .39$).

are phrased negatively (i.e. a mixed member proportional system is *worse* because...). And some arguments combine many of these features. Table 5.1 summarizes the characteristics of the arguments. We are thus not only interested in which arguments are more powerful, but also what components make an argument more persuasive than another. Accordingly, we use both conventional methods (in Section 5.4.2) and a structured Bradley-Terry model (in Section 5.5.2) to attempt to identify the most persuasive arguments and the components which make them persuasive.

5.2.2 The ‘Framing’ Connection

We see a connection between our endeavors here and the large literature on ‘framing’ in public opinion. In that work, the central conceit is that “an issue can be viewed from a variety of perspectives and be construed as having implications for multiple values or considerations” (Chong and Druckman, 2007*b*, 104). In our experiment, different arguments for the *same* system are phrased in different ways, highlighting particular features of the voting arrangements under debate. In this sense, we are firmly within the ‘emphasis’ framing paradigm—wherein the frames are not simply equivalent ways to convey the exact same information (see Druckman, 2004, for discussion of this clarification). Since each of these characteristics potentially appeals differently to voters via a distinct “frame in communication” (Chong and Druckman, 2007*b*, 104), we might expect some frames to do better than

Table 5.1: Argument Characteristics

Argument	Fairness/ Proportionality	Local Members/ Control	Government Formation/ Stability/Effectiveness	Ballot Structure	Diversity	Political Parties	Negative
FPTP1			X				
FPTP2		X					
FPTP3				X			
FPTP4		X				X	
FPTP5			X				X
FPTP6						X	X
MMP1	X					X	
MMP2		X		X		X	
MMP3	X					X	
MMP4					X		
MMP5	X					X	X
MMP6	X					X	X

others in convincing voters. Indeed, we will demonstrate this to be the case.

As an aside, we candidly note that a fundamental issue that is difficult for us to explore directly in our study is that of ‘moderators’. These variables essentially condition the effect of the frame, perhaps amplifying or damping-down its consequences for the subject (see Druckman, 2001; Haider-Markel and Joslyn, 2001, for example). Predispositions such as ‘values’ or ‘experience’ are generally considered important, though this is less of a concern for us because the referendum in Ontario was a *new* issue for voters whose views may not have settled *a priori* of the framing information they received. Moreover, as we will demonstrate when we describe the Bradley-Terry model below, it is the *relative* performance of *randomly assigned* arguments that we wish to examine here, rather than their performance in some absolute sense.

5.3 Measuring the Power to Persuade: A Statistical Model

We contend that arguments vary in their ‘power to persuade’ citizens who receive them. The central econometric concern is to *estimate* this persuasion ‘power’, possibly as a function of the characteristics of the arguments themselves. We begin by supposing that in any two-way comparison in which an argument pertaining to FPTP is paired with one pertaining to MMP, one argument makes a more convincing case than the other for its respective electoral system; it is thus more likely to persuade the respondent. For any

given comparison or ‘contest’ between arguments—or ‘players’— i and j let $\pi_{i,j} \in (0, 1)$ be the probability that the respondent finds i more compelling than j and thus prefers the electoral system (implicitly or explicitly) promoted by i . Write the odds that argument i ‘beats’ argument j as a function of their ‘powers’ α_i, α_j which are latent (and thus unobserved) but positive valued

$$\frac{\pi_{i,j}}{\pi_{j,i}} = \frac{\pi_{i,j}}{1 - \pi_{i,j}} = \frac{\alpha_i}{\alpha_j} \quad \forall i, \forall j. \quad (5.1)$$

Implicit in Equation (5.1) is the fact that $\pi_{j,i} \equiv 1 - \pi_{i,j}$. In other words, there can be no tied contests: either i wins and j loses or j wins and i loses.⁴ The task is to estimate α_i and α_j . Suppose that i and j compete against one another a total of $N_{i,j}$ times and let $n_{i,j}$ be the number of times i beats j . If all these contests are independent, then it is natural to assume $n_{i,j}$ is distributed as a binomial $(N_{i,j}, \pi_{i,j})$. With t total arguments competing—i.e. all the specific i s and j s that are actually compared in the survey—the likelihood is

$$L(\boldsymbol{\alpha}|\mathbf{n}) = \frac{\prod_i \alpha_i^{n_{i,j}}}{\prod_{i < j} (\alpha_i + \alpha_j)^{N_{i,j}}}. \quad (5.2)$$

Maximization of (5.2) yields estimates of the elements of $\boldsymbol{\alpha}$ subject to the identification restriction that some α_i set equal to one.⁵ Notice that, via an assumption of transitivity, the $\boldsymbol{\alpha}$ may be estimated even if $n_{i,j}$ is zero for some pairings, so long as there does *not* exist some subset of arguments that

⁴Effectively, this means not allowing ‘don’t know’ responses.

⁵Alternatively, the researcher could set $\sum_i \alpha_i = 1$.

never meet with others in competition.⁶ A logit linear form of the problem proceeds by defining $\exp(\lambda_i) \equiv \alpha_i$ and executing the maximization above. In this case, with rearrangement of Equation (5.1),

$$\pi_{i,j} = \frac{\exp(\lambda_i)}{\exp(\lambda_i) + \exp(\lambda_j)}, \quad (5.3)$$

and, in terms of log-odds,

$$\log \left[\frac{\Pr(i \text{ beats } j)}{\Pr(j \text{ beats } i)} \right] = \lambda_i - \lambda_j. \quad (5.4)$$

The intuitive message from Equations (5.3) and (5.4) is clear: the larger the value of λ_i *relative* to λ_j , the more likely it is that argument i beats argument j in a pairwise contest.

The approach here may appear unusual to public opinion researchers, but is well known to statisticians as the 'Bradley-Terry model' for pairwise comparison (Bradley and Terry, 1952) and has been used by psychologists interested in subjects selecting items from choice sets for some time (see, for example, Luce, 1959; Thurstone, 1959).⁷ The model has seen use in other fields such as biology (e.g. Stuart-Fox et al., 2006), genetics (e.g. Sham and Curtis, 1995), the investigation of journal citation patterns (e.g. Carter and Spirling, Forthcoming; Stigler, 1994) and sports science (e.g. Agresti, 2002).

⁶More technically, the design must be 'connected' though it need not be 'complete'.

⁷In practice, Thurstone (1959), opts for a slightly different form of the model: in selecting a function, f for the link $\pi_{i,j} = f^{-1}(\lambda_i - \lambda_j)$, he uses the inverse probit; as seen in Equation (5.4), Bradley and Terry (1952) use the logit.

Spirling (2007) considers an application to the United States Senate, however we can find no other application of this model in political science.⁸

An important extension of the Bradley-Terry model is provided by Springall (1973), in which ‘player’-specific (here, argument-specific) variables x_{i1}, \dots, x_{ip} are used to predict the ‘power’ of the players (the λ) directly. These independent variables enter the model via the linear predictor in the sense that

$$\lambda_i = \sum_{r=1}^p \beta_r x_{ir}, \quad (5.5)$$

and interest focuses on the estimated coefficients β_1, \dots, β_p . In the current context, these β inform the analyst as to the source, or ‘cause’, of the arguments’ power. Firth (2005) devotes considerable effort to designing software for the fitting of both ‘unstructured’ models of the form given in Equation (5.4), and ‘structured’ versions as noted in Equation (5.5). Our work below utilizes his implementation in conjunction with the R language and environment (R Development Core Team, 2006). In the next section though, we review the use of more conventional techniques before demonstrating the utility of our model.

⁸Indeed, a search of all major political science journals on JSTOR resulted in only one hit; an article mentioning the Bradley-Terry model in an end note (Monroe, 1995).

5.4 The Power of Arguments: Conventional Results

5.4.1 Testing the Power of Arguments

We consider two conventional methods of testing the power of arguments. The first is to observe the mean level of support for MMP according to the assigned FPTP and MMP arguments and then identify which arguments elicit significantly higher mean levels of support. The results in Tables 5.2 and 5.3 do just this, comparing support for each argument. The arguments are arranged in order from most to least powerful.⁹ Of the six arguments for the existing system, FPTP6 appears to be the weakest of all arguments. Indeed, four arguments elicit significantly higher support. However, according to convention *t*-tests of means, all other arguments appear equally powerful.

We can find even less difference in the persuasive power of arguments for a mixed member system (Table 5.3). When we consider all respondents, we find all arguments are equally persuasive, with the exception of MMP6, which is significantly more persuasive than MMP4.

These findings are instructive as far as they go. We can create rank orderings of arguments but without much certainty about differences in their levels of power. We may also be able to infer the reasons for differences in persuasiveness, or power. For example, we could infer that the difference

⁹As the quantity is percentage agreement with MMP, FPTP arguments with lower support are more persuasive.

between FFTP1 and FFTP6 exists because FFTP1 references government effectiveness while FFTP6 does not. However, our inference would end there. We would not be able to make statements with measured uncertainty about why FFTP1 is more powerful than FFTP6. At best, we are left with an uncertain rank ordering and little insight into the sources of an argument's power.

Rather than a means-test, we could also perform a logistic regression in which we set support for MMP as the dependent variable and dummy out the arguments a respondent received. We do this in Table 5.4. The logit tests clearly suffer from low statistical power. Using conventional significance levels, the coefficients tell us nothing about which arguments are more or less powerful. We do note, however, that both *age* and *gender* have significant effects. The results suggest that MMP is more preferred by women while it is less supported by older citizens. The results also demonstrate that even with 520 respondents, the logit model is able to pick up the larger effects due to gender and age. If the effects of arguments for and against electoral reform are small—and we would expect them to be on a sophisticated panel—then relying on a traditional logit analysis arguably significantly increases the risk of Type II errors. This difficulty only becomes greater if we dummy out each argument pairing, as shown in Table G.1 in Annexe G. Here only one argument pairing variable out of 36 reaches conventional levels of statistical significance, less than we would expect even by chance.

Table 5.2: Agreement with MMP by FFTP Arguments (mean differences)

Argument	Mean % (sd)	Dominates (p -value, one-tailed)
FFTP2	41.4 (49.5)	FFTP6 (0.02)
FFTP1	44.4 (49.9)	FFTP6 (0.06)
FFTP3	45.2 (50.1)	FFTP6 (0.08)
FFTP4	45.6 (50.1)	FFTP6 (0.08)
FFTP5	47.8 (50.2)	—
FFTP6	56.7 (49.8)	—

Table 5.3: Agreement with MMP by MMP Arguments (mean differences)

Argument	Mean % (sd)	Dominates (p -value, one-tailed)
MMP6	51.2 (50.3)	MMP4 (0.08)
MMP3	49.3 (50.3)	—
MMP5	48.2 (50.3)	—
MMP1	46.3 (50.1)	—
MMP2	45.1 (50.0)	—
MMP4	40.6 (49.3)	—

5.4.2 Determining the Sources of Power

In the previous section, we demonstrated that with a large number of treatments we could learn little about the power of arguments using a traditional logit set-up. We did succeed in constructing rank orderings of the power of arguments according to their mean levels of support, but we could not infer from this the sources of their power, i.e. the elements of some argument

Table 5.4: Logistic Regression of Argument Power^a

Variable	Model 1			Model 2		
	Coefficient	se	<i>p</i> -value	Coefficient	se	<i>p</i> -value
FFTP2	-0.13	0.30	0.64	-0.13	0.30	0.66
FFTP3	0.01	0.32	0.96	-0.05	0.32	0.87
FFTP4	0.07	0.30	0.83	0.00	0.31	1.00
FFTP5	0.14	0.29	0.64	0.04	0.31	0.90
FFTP6	0.47	0.31	0.14	0.46	0.32	0.15
MMP1	-0.21	0.30	0.50	0.17	0.57	0.76
MMP2	-0.08	0.30	0.79	-0.12	0.31	0.70
MMP3	0.12	0.33	0.71	0.12	0.33	0.72
MMP4	-0.25	0.31	0.42	-0.27	0.13	0.40
MMP5	0.04	0.32	0.90	0.02	0.32	0.95
MMP6	0.14	0.32	0.64	0.07	0.32	0.81
Female				0.41	0.18	0.03
Age				-0.22	0.08	0.00
N	520			520		
Pseudo R^2	0.01			0.03		

^a FFTP1 is the reference category.

which made it more convincing than other arguments for a similar policy position. These results should be of little surprise to seasoned analysts of public opinion. We have designed an experiment with a very large number of treatment groups and a small number of respondents. This lack of results is not a criticism of these methods *per se*. Rather, it is a demonstration of a limit of these methods. A similar limit exists if we attempt to locate the sources of an argument's power.

In Table 5.1, we identified the characteristics of the various arguments for

and against electoral reform. For example, did the argument appeal to the value of local control, did it reference fairness, etc? We create eight such variables: Fairness, Local, Government Stability/Effectiveness, Ballot Structure, Diversity, Parties, Negative Frame, FPTP. *Fairness* indicates an argument which makes appeals based on proportionality, while *Local* is a variable indicating that an argument makes reference to local representation or control. *Stability* essentially refers to the notion that governments are more likely to survive longer periods and be able to make policy in some systems rather than others. *Ballot Structure* refers to an argument about how simple or confusing the actual voting papers are in practice, while *Diversity* is a reference to arguments making claims of variegation or heterogeneity of representation in the legislature. *Parties* indicates that an argument mentions political parties and *Negative* occurs when an argument is in favor of a system only insofar as it highlights a *detracting* feature of the opposing position. Finally, not recorded in Table 5.1, *FPTP* indicates that the argument is in favor of the existing (status quo) system: of course, this takes a value of '1' for FPTP1 through FPTP6. Because some arguments share some characteristics but not others, the question of which characteristics really make an argument persuasive is an interesting one, but not necessarily easily answered.

One way of locating the sources of an argument's power is to create variables that indicate whether an argument (and its opposing position) contain certain characteristics. Each of the seven variables is coded as 1 if a respondent received only an MMP argument which had this characteristic, 0 if both

or neither variables had this characteristic and -1 if only the FPTP argument had this characteristic. For example, both FPTP3 and MMP2 refer to ballot structure in a manner favorable to their side. So, if a respondent received both FPTP3 and MMP2, then the Ballot Format variable would read 0. If they received MMP2 but not FPTP3, then Ballot Format would read 1. Vice-versa and the variable would read -1.

Using a logit setup, we regress the choice for MMP on these seven argument variables, as well as controls for age and gender. As the results reported in Table 5.5 indicate, we find no significant differences in the argument characteristics, though we do again find significant demographic effects. This suggests that there are no differences in the ability of various argument components to persuade. Moreover, we still fail to find significant predictors if we perform a series of bivariate regressions between each component and support for MMP.

5.5 The Power of Arguments: The Bradley-Terry Method

5.5.1 Unstructured Results

We have previously presented a rank-ordering of arguments by mean support. We were able to uncover some differences using tests of means,

Table 5.5: Logistic Regression of Sources of Argument Power

Variable	Coefficient	se	<i>p</i> -value
Fairness	0.55	0.51	0.28
Local Control	0.44	0.35	0.21
Stability/Effectiveness	0.30	0.42	0.48
Ballot Format	0.37	0.48	0.44
Diversity	0.08	0.77	0.92
Parties	-0.14	0.30	0.65
Negative Frame	0.02	0.18	0.92
Female	0.41	0.18	0.03
Age	-0.23	0.07	0.00
Constant	0.14	0.60	0.82
N	520		
Pseudo R^2	0.03		

but we were unable to compute from these the probability that one argument would beat another. We have also attempted to ascertain comparative power through a logit analysis to little avail. Moreover, we have (unsuccessfully) tried to determine the sources of arguments' power using a second logit analysis. In place of these more conventional approaches, we now present Bradley-Terry results. Our first 'unstructured' results are in Table 5.6. Unstructured results consist of a power coefficient for each argument, λ_i , which can be used to compare its power with that of other arguments. The reported coefficient for FPTP2, the most powerful argument, is set to 0. All other coefficients (and their *p*-values), are reported in comparison to the power of this FPTP argument. The arguments are ordered from most

to least powerful. By examining the coefficients we can determine which arguments are estimated as less powerful than others. Moreover, by examining the p -values, we can determine which arguments are significantly less powerful than FPTP2 (i.e. MMP4 and FPTP6). With such a rank ordering, the Bradley-Terry model does not differ much from a mean-support type approach outlined above and the order of the arguments appears the same. However, this approach does offer the distinct advantage that we can easily compute the estimated probability that one argument dominates another. Recall Equation (5.3); if we wanted to know the probability that FPTP2 defeated MMP2, we enter the power estimates into the equation and find that:

$$\pi_{2,2} = \frac{\exp(0.00)}{\exp(0.00) + \exp(-0.422)} = 0.60. \quad (5.6)$$

Similarly, we could calculate the probability that FPTP2 beats the more evenly matched MMP6 at 0.55. We cannot make similar probability statements from merely observing means. Nor can we derive them from a logistic regression which returns insignificant coefficients.

Tables 5.7 and 5.8 break up the results by FPTP and MMP and in the case of each argument includes an indication of which other arguments it dominates. The arguments do not now appear as well-matched as in Table 5.6. Rather, we see that FPTP2 dominates one argument for MMP, and FPTP1 also weakly dominates a MMP argument. No argument for MMP systematically dominates a FPTP argument. The advocates of a FPTP system had a clear

Table 5.6: Bradley-Terry Model of Argument Power

Argument	Power	<i>p</i> -value
FFTP2	0.000	—
FFTP1	-0.137	0.63
FFTP3	-0.151	0.63
MMP6	-0.197	0.50
FFTP4	-0.203	0.49
MMP3	-0.224	0.45
FFTP5	-0.278	0.34
MMP5	-0.302	0.30
MMP1	-0.343	0.23
MMP2	-0.422	0.12
MMP4	-0.593	0.04
FFTP6	-0.605	0.05

advantage.

In Figure 5.1 we present our findings in a slightly different way. Here, each ‘fan’ represents an argument. A fan’s size is proportional to the argument’s ‘power’ as estimated in the Bradley-Terry model. Each *section* of each fan is sized proportionally to the probability that the argument would defeat each of the eleven others in a contest. As with Table 5.6, we see that the top three most powerful arguments all concern FFTP, while three of the four least powerful arguments concern MMP. This further suggests an overall advantage for FFTP advocates. In the next section we examine whether there in fact was a systematic advantage for FFTP arguments.

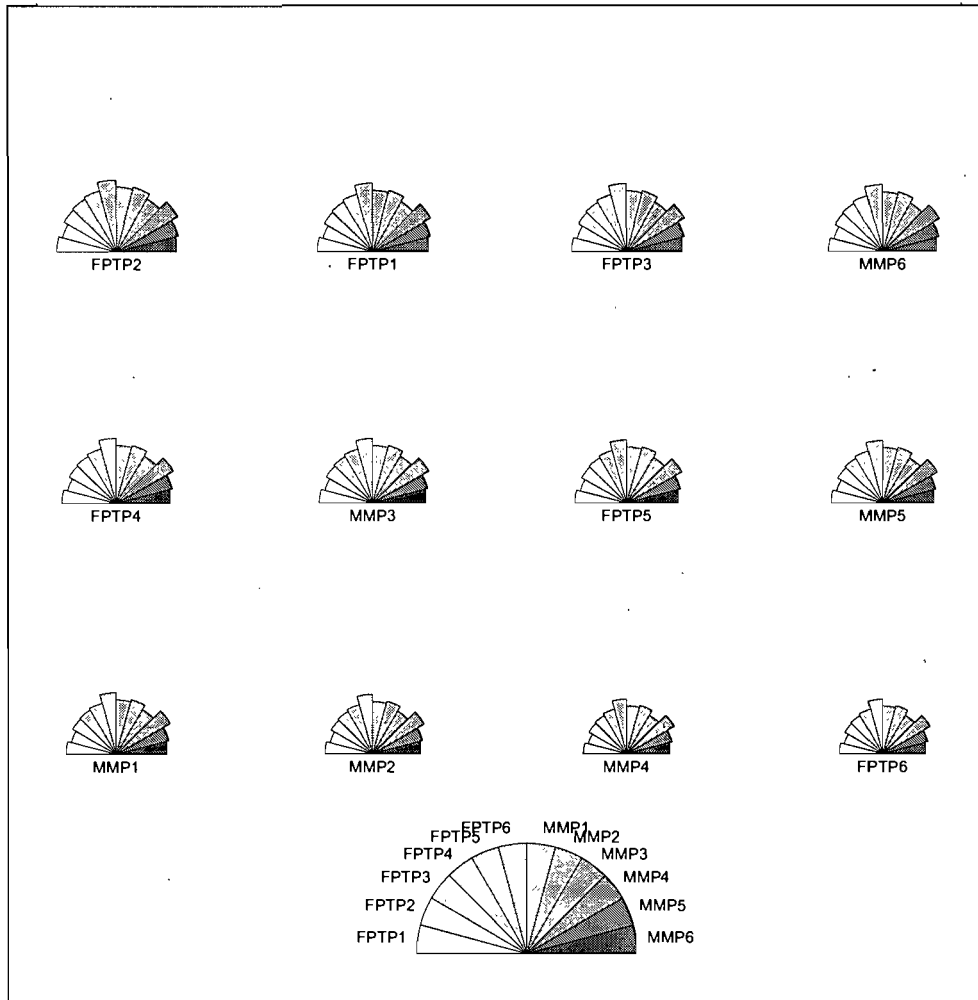


Figure 5.1: Power of arguments: unstructured results. Each 'fan' represents an argument and its size is proportional to its 'power' (thus FPTP2 is the largest and most powerful). Each section of each fan is sized proportionally to the probability that the argument beats the others in a direct contest. Note the key at the bottom of the diagram.

Table 5.7: Power of FPTP Arguments

Argument	Power	Dominates (<i>p</i> -value)
FPTP2	0.000	MMP4 (0.04); FPTP6 (0.05)
FPTP1	-0.137	MMP4 (0.10)
FPTP3	-0.151	—
FPTP4	-0.203	—
FPTP5	-0.278	—
FPTP6	-0.605	—

Table 5.8: Power of MMP Arguments

Argument	Power	Dominates (<i>p</i> -value)
MMP6	-0.197	—
MMP3	-0.224	—
MMP5	-0.302	—
MMP1	-0.343	—
MMP2	-0.422	—
MMP4	-0.593	—

5.5.2 The Sources of an Argument's Power

In the previous section, we estimated the power of twelve arguments for and against electoral reform. Our task now is to estimate the sources of this power. In other words, what are the characteristics of an argument that make it able to overcome other arguments—in our case, for or against electoral reform? In Table 5.1 we summarized the attributes of each of the FPTP and MMP arguments. By including these traits as covariates in a structured model we are able to estimate the sources of an argument's power—or

‘ability’—to persuade.¹⁰

Since the model is estimated via maximum likelihood, standard procedures for finding the ‘best-fit’ are available: for example, the analyst may use the Akaike or Bayes Information Criterion. We use the AIC to determine a ‘best-fit’ model in Table 5.9. As a result, we finish with four predictors. We take this systemic advantage for some components over others as evidence that at least some of our sample was responding to arguments systematically, rather than merely stating prior opinions. With the exception of *Parties*, all the variables are significant at the 10% level. When we repeat the logit analysis in Table 5.5 using these variables, only Fairness achieves significance at the 10% level.

Table 5.9: Model of Sources of Argument Power

Variable	Coefficient	se	<i>p</i> -value
FPTP	0.282	0.147	0.055
Fairness	0.534	0.240	0.026
Local	0.318	0.177	0.071
Parties	-0.275	0.180	0.127
AIC	141.88		

All else equal then, an argument in favor of the status quo appears systematically advantaged. *Local* is a characteristic of arguments for both FPTP and MMP, as an MMP system does retain local representation. However,

¹⁰This is Equation (5.5) in Section 5.3.

MMP does not allow the local selection of *all* candidates, suggesting that in actual rhetoric the scope of local arguments favors FPTP. *Fairness* arguments clearly advantage MMP, as proportionality is a central component of this system and is never present in arguments for FPTP. Finally, the mention of political parties makes an argument less appealing. This is likely to the disadvantage of MMP, as the logic of the system (and thus at least some of the arguments in its favor) rely on matching the votes and seats of political parties. Parties are thus arguably more central to an understanding of MMP than FPTP.¹¹

Using this model and the formula in Equation (5.3) we can then estimate the probability that one argument dominates another according to its characteristics. For example, we can estimate the probability that FPTP2 (which mentions local control) dominates MMP6 (which makes a fairness argument):

$$\pi_{2,6} = \frac{\exp(0.282 + 0.318)}{\exp(0.282 + 0.318) + \exp(0.534)} = 0.52. \quad (5.7)$$

By contrast, if we pit FPTP2 against MMP4, an argument which has no significantly powerful attributes, then the probability of FPTP dominance rises to 0.65.

Figure 5.2 presents our findings in a slightly different way. The figure

¹¹As an aside, we note that the issue of ‘negative’ framing is a popular research topic in political science (see, for example, Ansolabehere and Iyengar, 1997). We found that ‘Negative’ was not a significant predictor (on its own) for the structured model, though it had a negative sign implying that voters are (all else equal) less convinced by such phrasing.

compares the probability of an argument with various covariate profiles winning against the modal argument (which does not mention ‘fairness’, ‘local’ or ‘parties’, but does make reference to ‘FPTP’). The arrows show the effect of moving from one covariate profile to another. For example, in a contest with the modal argument, an argument that is identical but which *mentions* fairness would do better: the probability of a win for this second argument increases from 0.5 to around 0.64. By contrast, an argument that is identical but which does not *mention* FPTP would do worse: the probability of a win for this second argument decreases from 0.5 to around 0.45. On the left side of the plot, we report the predicted probabilities for the arguments in the study *based on their covariate profiles*. Reading from top to bottom, we see that FPTP2 is once again the most powerful argument in keeping with our unstructured results.¹²

¹²In Annexe H, we consider the differences between the predicted probabilities generated from structured and unstructured models. We find the models generate similar predictions, on average.

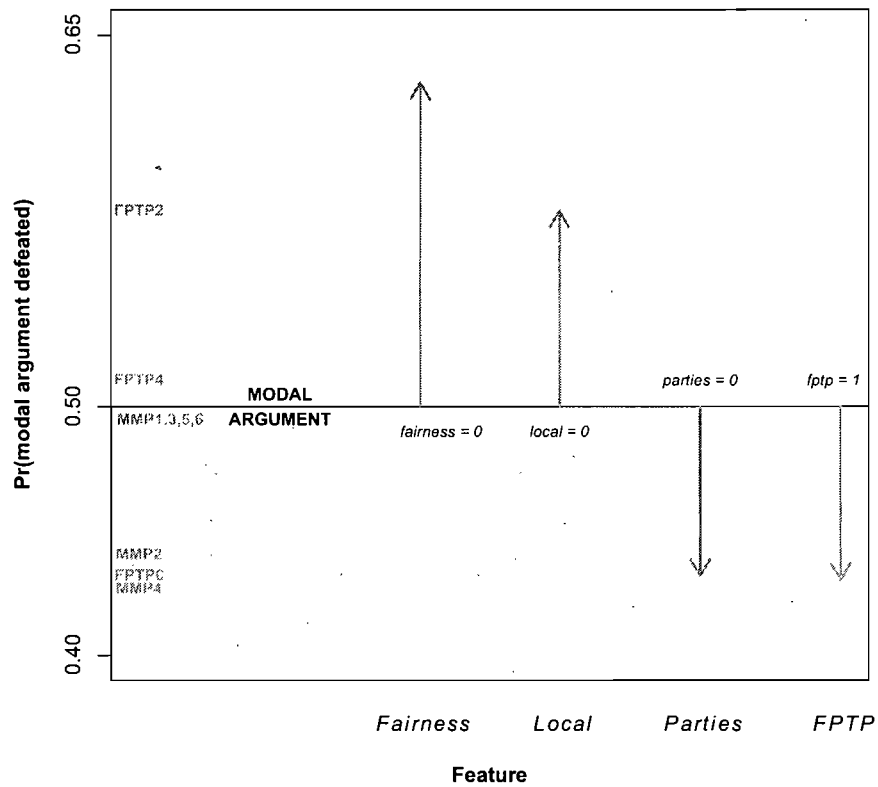


Figure 5.2: Power of arguments: structured results. The arrows show the effect of moving from one covariate profile to another. On the left side of the plot, we report the predicted probabilities for the arguments in the study based on their covariate profiles. The modal arguments are FPTP1, FPTP3, and FPTP5.

5.6 Further Applications and Conclusion

We have demonstrated that Bradley-Terry models can generate meaningful results from pairwise matchings of arguments, even when the number of treatment groups is large and the sample size small. Similar results were not uncovered by conventional means-tests or logit analysis. Moreover, structured models can be used to identify the source of an argument's persuasive power. These models are thus well-suited for a situation in which a researcher wishes to test a large number of arguments on a small sample size.

We have also gained substantive insight into the comparative power of arguments for and against electoral reform. While the arguments were generally well-matched, we have found a small advantage in status-quo arguments for a FPTP system, in arguments which reference local candidates and in arguments which appeal to fairness norms. Overall, this advantaged arguments for the existing system in Ontario, and goes some way in explaining the failure of electoral reform in the province. They also conform with a more general status quo bias in referendums and initiatives (e.g. see Bowler and Donovan, 1998). These insights were not apparent from conventional analyses.

Just as these models have found application in a large number of other fields, we can find further applications for them in public opinion research. For example, we could measure the comparative power of frames over a small number of respondents without the need to repeat frames (e.g., Chong and

Druckman, 2007a). The models could also be used to identify sources of persuasion in studies using a ‘jury method’ (e.g., London, Meldman and Lanckton, 1970). They could be used in experimental studies determining the characteristics of candidate attractiveness (e.g., Sigelman, 1990). Indeed, any research design with dichotomous outcomes and direct confrontation between units or actors is a candidate for these models.

Our suggested approach is not without drawbacks. It does require a minimum degree of fluency in the R environment (though some of its components can be undertaken in STATA). Similarly, it requires pairwise comparison and dichotomous outcomes. Accordingly, it cannot be used if researchers are interested in measuring degrees of support for a position and/or if they are not interested in testing arguments in an oppositional manner. Similarly, a researcher cannot easily estimate tournament-specific (or respondent-specific) effects with a Bradley-Terry model. It is not a panacea. However, if researchers wish to make pairwise comparisons with dichotomous outcomes, this approach provides a powerful solution when a number of arguments need to be tested and a small number of respondents are at hand. Indeed, we are convinced that when matched against more conventional methods, these models are up to the test.

Chapitre 6

Partisanship and Altruism: Results from a Dictator Game Experiment (with Angelo Elias)

Under review at Political Behavior. Loewen designed, executed and analysed the experiment. Loewen drafted the first version of the article. Elias helped draft a second. Loewen completed the final draft.

6.1 Introduction

Partisanship matters. Partisans are more likely to vote than non-partisans (Verba, Schlozman and Brady, 1995; Franklin, 1996, 2002; Norris, 2002; Dalton, 2006; Blais, Gidengill, Nadeau and Nevitte, 2002) and to do so consistently (Campbell et al., 1960; Dalton, 2006; Blais, Gidengill, Nadeau and Nevitte, 2002; Miller and Shanks, 1996). Partisans are more likely to follow politics (Campbell et al., 1960) and they are more likely to have structured opinions (Hamill, Lodge and Blake, 1985; Sharp and Lodge, 1985; Lodge and Hamill, 1986). As important as differences between partisans and non-partisans are those among partisans. We know, for example, that partisanship can lead to systematically different views on issues Blais, Gidengill, Nadeau and Nevitte (2002). We also know that different partisans attend to and receive news and political arguments differently (e.g. Zaller, 1992; Bartels, 2002; Johnston, 1992). These basic differences between partisans are well-known and generally travel across several countries (e.g. Clarke et al., 2004).

This note asks if partisans differ in another way, specifically if partisans of some parties are more altruistic than others. That is, are some partisan identifiers systematically more generous when given the opportunity to improve the lot of others at a cost to themselves (Rushton, 1982; Ridley and Dawkins, 2003; Aronfreed, 1980; Piliavin and Charng, 1990; Margolis, 1983)? In keeping with a growing trend in political science, we answer this question

experimentally (see Druckman et al., 2006). The altruism of partisan identifiers in Canada is measured through a series of dictator games (Camerer, 2003) - a tool from behavioural or experimental economics - embedded in an online survey. By measuring the impact of different partisan identifications on allocations to completely anonymous individuals, other partisans, and co-partisans, we can identify whether partisans differ in their altruism. We find that New Democrats exhibit more altruism than Conservatives and Liberals in two of three scenarios, and that Liberals exhibit more altruism than Conservatives in one scenario. When we consider all scenarios together, we find that New Democrats exhibit more altruism on average. We also find that all partisans exhibit more altruism towards co-partisans than towards anonymous individuals and more altruism towards anonymous individuals than towards the supporters of other parties.

This note proceeds as follows. We first describe the dictator game and defend it as a measure of altruism. We also briefly review the properties of the game. We then describe the survey in which the games were embedded. We finally present and discuss our results.

6.2 Altruism and the Dictator Game

We use a minimalist definition of altruism: *an altruistic act occurs when an individual undertakes action which is to the material benefit of another at a material cost to herself*. To satisfy this definition of altruism we do not

need to observe or infer intent (contra Batson et al., 1978, 1979). We merely need to observe an action which confers a benefit and comes at a cost. In this manner, our definition is consistent with other well-known minimalist accounts (e.g. Sorrentino, 1991; Margolis, 1983).

We measure the altruism of partisans through a series of dictator games (see Camerer, 2003) embedded within an online survey. The dictator game is played as follows: Player A is given some sum of money (or the chance to win this money). She is then asked how much of this money (should she win) that she would like to share with Player B. The distribution is made and the game ends. In short, Player A ‘dictates’ a share of the prize to Player B. Any non-zero sum represents an altruistic action as it improves the material lot of Player B at a cost to Player A. Larger allocations thus represent more altruistic actions.

While the use of dictator games is rather new in political science (e.g. Fowler, 2006; Fowler and Kam, 2007; Whitt and Wilson, 2007), it is common in economics. In that context, economists have demonstrated that allocations in dictator games are consistent and rational (Andreoni and Miller, 2002), they differ little according to the size of the stakes provided that the stakes are not fictional (though they can be merely expected) (Camerer and Hogarth, 1999; Carpenter, Verhoogen and Burks, 2005; Fowler, 2006), and they correlate as expected with real world behavior (Benz and Meier, 2008).

The few studies which have considered the effects of partisanship on dictator game allocations to anonymous individuals (Fowler, 2006; Fowler and

Kam, 2007) have typically failed to find significant relationships between partisanship and allocations. This is consistent with findings that sociodemographic variables, with the exception of gender (Eckel and Grossman, 1996, 1998), are generally unrelated to dictator game allocations (see also Camerer, 2003). These findings, however, should be viewed with caution as dictator games typically occur within convenience samples which are both small and unrepresentative of the general population. Fowler (2006), for example, fails to find significant relationships, but his data suggest that they *may* exist between allocations and age and partisanship and go undetected because of a small sample.

6.3 The Study

6.3.1 Subjects

Our survey was conducted in May 2007 by a Canadian commercial public opinion research firm. As the survey was conducted online, respondents were required to login to the survey using a unique identification. While the original survey included 5399 respondents, we include only those who identify with a political party, live outside of the province of Quebec¹, and completed the relevant questions. We are left with 1942 respondents in our effective sample. Our sample is probably more politically sophisticated than

¹The exclusion of Quebec from analysis of political behavior in Canada is commonplace (see, e.g. Blais, Gidengill, Nadeau and Neviite, 2002).

the general population of partisans in Canada. For example, while 59% of our effective sample indicates having attended at least some university, the equivalent quantity in the most recent Canadian Election Study is just 37%.² To address this, we include controls for education in our models. At the end of the note, we report simple robustness checks for the effects of sophistication.

This objection aside, we are confident that an online survey is an appropriate method by which to evaluate the altruism of partisans through a dictator game. First, while our sample may not be properly representative, we know no reason why self-selection into our panel would affect the direction or nature of the relationship between altruism and partisanship. We are, in other words, confident that the partisans in our sample use the same considerations in deciding whether to be altruistic as partisans in the general population (Best et al., 2005). Second, giving money in a dictator game as a function of social desirability is likely lower in an online survey than in a laboratory or a telephone survey (Taylor and Thomas, 2005) as respondents are not interacting with a human experimenter or caller. Accordingly, an online approach may allow for a more accurate measure of behavior in the dictator game.

²While it would be ideal to compare our sample to census data, this is not possible given that our sample is limited to partisans. Accordingly, we have compared it against the randomly-sampled Canadian Election Study.

6.3.2 The Survey

People completing the survey for the first time responded to a number of screening questions, including a query about their partisan identification. All respondents were presented with a series of questions about recent news exposure, attention to federal and provincial politics, and views on various politicians. Subjects then completed an eight-item measure of empathy based on a larger instrument (see Loewen, Lyle and Nachshen, 2007; Wakabayashi et al., 2006) as well as four dictator games which are described in the next paragraph. Following this, subjects answered questions concerning their support for public spending, past charitable giving, views of the public service, and views of recent political events.

6.3.3 Dependent Variable: Dictator Game Allocations

The dictator game experiment consisted of four iterations. Subjects read an introduction to the game and were then presented with instructions. Upon being informed that they would have four chances to win one of four \$100 prizes at the end of the survey, subjects were asked how much they would like to share with a completely anonymous individual about whom they know nothing and who would never know their identity. Subjects entered their preferred split. The game was then repeated three times with subjects being informed that the anonymous recipient supports the Conservative, Liberal, or New Democratic parties. The order of the recipients was randomized. Re-

spondents could give away any amount between \$0 and \$100. The complete instructions for the game are available in Annexe B. For the purposes of our model, we create three different dependent variables. Allocations to *Anonymous Individuals*, to *Other Partisans*, i.e. partisans of party other than the respondent's, and to *Co-Partisans*, i.e. partisans of the same party as the respondent. We also create a pooled variable, *Pooled Allocations*, which pools all allocations by partisan respondents.

6.3.4 Independent Variables

The principal independent variables, *Conservative*, *Liberal*, and *New Democrat*, are based on the standard question "Thinking about federal politics in Canada, generally speaking, do you usually think of yourself as Liberal, Conservative, N.D.P, or none of these?" Those who indicated a partisanship received the standard follow-up question: "And, generally speaking, how strongly do you think of yourself as a (party)?" Respondents who indicated a strong or fairly strong identification with any party are included. All others are excluded. (Blais et al., 2001) We make this exclusion because we are interested exclusively in differences between partisans, rather than between non-partisans and partisans.

We include a number of other independent variables, including income, education, age, gender, and a dummy for unemployment. The question wording for all variables is available in Annexe I. For interpretive ease, all independent variables are rescaled from 0 to 1.

We also include another variable measuring the nature of respondents' concern for others. *Empathy* measures the capacity of respondents to vicariously experience the distress of others and to feel a motivation to alleviate that distress. Prior research has demonstrated a strong link between empathy and prosocial, helping behaviours (Vitaglione and Barnett, 2003; Staub, 1978, 1980). More pointedly, the Empathy-Altruism hypothesis argues that those who feel greater empathy exhibit greater altruism (Batson et al., 1981; Kruger, Hicks and McGue, 2001; Frey and Meier, 2004). Accordingly, more empathic respondents should make more generous allocations (Batson et al., 1979; Batson, 1994). The variable is a scale measuring average responses to eight questions drawn from a larger scale (see Loewen, Lyle and Nachshen, 2007; Wakabayashi et al., 2006). Having controlled for the effects of altruism due to empathy, any results due to partisanship are arguably a result of the different behavioural norms towards helping associated with each partisan identification. This is precisely the quantity we wish to identify: the independent of contribution of partisanship to altruism holding constant another well-known and basic determinant of altruism.³

In the case of each model, we specify an ordinary least squares linear regression. As the values of our dependent variable are left-censored, a tobit model may be more appropriate. However, our tobit results do not differ substantively, i.e. we find no differences in the significance of our partisanship

³We note that the correlations between empathy and each partisanship - New Democrat ($r = .08, p = .00$), Liberal ($r = .08, p = .00$), Conservative ($r = -.14, p = .00$) - are weak but significant.

variables and no differences in the ordering or direction of effects. Accordingly, we report more easily interpreted OLS coefficients (Roncek, 1992).

6.4 Results

Table 6.1 presents the results when the dependent variable is dictator game allocations to completely anonymous individuals. We observe the following means on this measure: New Democrats give \$25.52 ($sd = \23.81), Liberals give \$23.25 ($sd = \25.35), and Conservatives give \$21.72 ($sd = \25.96). These suggest that New Democrats appear more altruistic than Liberals ($t = 1.49, p = .07$) and more altruistic than Conservatives ($t = 2.44, p = .01$). Liberals in turn appear more altruistic than Conservatives ($t = 1.17, p = .12$).⁴ However, controlling for all other factors we find only one significant difference on account of partisanship. Being a New Democrat makes one weakly more altruistic towards completely anonymous individuals than being a Conservative. However, there is no difference between New Democrats and Liberals, and no difference between Liberals and Conservatives. In keeping with previous findings, women donate \$3.01 more than men. We do not find other significant sociodemographics effects. As expected, empathic respondents give significantly more. For the most basic measure of altruism, then, New Democrats appear to more altruistic than one other group of partisans.

When altruism is measured as allocations to partisans of another party

⁴Reported p-values for bivariate means tests are one-tailed.

Table 6.1: Dictator Game Allocations to Anonymous Individuals (OLS)

Variable	Coefficient	(Std. Err.)
Liberal	-1.940	(1.562)
Conservative	-2.856†	(1.607)
New Democrat	0.000	(0.000)
Income	-0.290	(1.544)
Education	-2.340	(1.666)
Female	3.099*	(1.213)
Age Group	2.468	(2.199)
Unemployed	-4.691	(3.291)
Empathy	9.159**	(2.716)
Intercept	18.485**	(2.842)
<hr/>		
N	1958	
R ²	0.017	
F (8,1949)	4.270	
<hr/>		
Significance levels : † : 10% * : 5% ** : 1%		

(Table 6.2), we observe the following means: New Democrats give \$22.90 ($sd = \20.26), Liberals give \$16.13 ($sd = \20.47), and Conservatives give \$12.19 ($sd = \19.94). As with giving to anonymous individuals, New Democrats appear more altruistic than Liberals ($t = 5.44, p = .00$) and Conservatives ($t = 8.66, p = .00$). Liberals also appear more altruistic than Conservatives ($t = 3.85, p = .00$). Controlling for all other factors, the marginal effect of being a Conservative is to give \$9.33 (95%*c.i.* = \$6.81, \$11.84) less than New Democrats, while the marginal effect of being a Liberal is to give \$5.91 (95%*c.i.* = \$3.46, \$8.35) less than a New Democrat. Liberals demonstrate significantly more altruism than Conservatives ($F = 10.76, p = .01$) but significantly less altruism than New Democrats. As in the case of anonymous

recipients, we find women are again more generous, giving \$2.53 more than men. We also find that more educated respondents exhibit more altruism and older respondents demonstrate less. In keeping with the previous finding, the more empathic are also more generous.

Table 6.2: Dictator Game Allocations to Other Partisans (OLS)

Variable	Coefficient	(Std. Err.)
Liberal	-5.908**	(1.247)
Conservative	-9.326**	(1.282)
New Democrat	0.000	(0.000)
Income	-1.890	(1.232)
Education	-2.927*	(1.329)
Female	2.531**	(0.968)
Age Group	-3.068†	(1.755)
Unemployed	-3.264	(2.626)
Empathy	7.356**	(2.167)
Intercept	21.622**	(2.268)
<hr/>		
N	1958	
R ²	0.055	
F (8,1949)	14.164	
<hr/>		
Significance levels : † : 10% * : 5% ** : 1%		

When altruism is measured as an allocation to co-partisans (Table 6.3), we observe the following means: New Democrats give \$31.92 ($sd = \25.07), Liberals give \$24.35 ($sd = \24.99), and Conservatives give \$23.82 ($sd = \25.29). Once again, New Democrats appear more altruistic than Liberals ($t = 4.96, p = .00$) and Conservatives ($t = 5.21, p = .00$). However, there appears to be no difference between Liberals and Conservatives ($t = 0.41, p = .34$). Controlling for all other factors, we find that those identifying as Lib-

eral or Conservative demonstrate significantly less altruism than those who identify as New Democrats. The difference is equal to approximately \$7.00; the difference between Liberal and Conservatives does not achieve statistical significance. There are no significant sociodemographic predictors. However, empathy is very powerful.

Table 6.3: Dictator Game Allocations to Co-Partisans (OLS)

Variable	Coefficient	(Std. Err.)
Liberal	-7.060**	(1.551)
Conservative	-6.945**	(1.596)
New Democrat	0.000	(0.000)
Income	-2.195	(1.533)
Education	-0.810	(1.654)
Female	0.885	(1.205)
Age Group	2.308	(2.183)
Unemployed	-2.548	(3.268)
Empathy	13.228**	(2.697)
Intercept	23.390**	(2.822)
<hr/>		
N	1958	
R ²	0.031	
F (8,1949)	7.893	
<hr/>		
Significance levels : † : 10% * : 5% ** : 1%		

As a final test, we have pooled our observations across allocations and run a fourth model. In this case the dependent variable is *Pooled Allocations*. We include the same demographic controls as before, as well as variables indicating whether the allocation was to a *Co-Partisan* or to an *Other Partisan*. Allocations to anonymous individuals by New Democrats serve as the reference group. As we have multiple observations for each respondent, we specify

robust standard errors.

The results from the pooled model confirm the results found above. Liberal and Conservative partisans exhibit significantly less altruism than New Democrats, on average. However, they do not differ significantly from one another ($F = 1.47, p = .23$). Second, we find that all partisans allocate significantly less to other partisans than to anonymous individuals and to co-partisans. Third, we find a significant difference between allocations to anonymous individuals and (larger) allocations to co-partisans. Finally, we find that the more empathic are more generous, as are women. We fail to find significant effects for age, income, or education.

Table 6.4: Pooled Dictator Game Allocations (OLS)

Variable	Coefficient	(Rob. Std. Err.)
Liberal	-4.936**	(1.303)
Conservative	-6.334**	(1.323)
Other Partisan	-7.120**	(0.415)
Co-Partisan	2.574**	(0.450)
Female	2.139*	(1.002)
Income	-1.493	(1.275)
Education	-2.045	(1.320)
Unemployed	-3.565	(2.592)
Empathy	9.897**	(2.254)
Intercept	23.068**	(2.012)
<hr/>		
N	5874	
R ²	0.056	
F (9,1957)	79.387	
<hr/>		
Significance levels :	† : 10%	* : 5% ** : 1%

Taken together, these results suggest that partisanship matters for altru-

ism. The differences we find – that partisans discriminate between partisans and among partisans and non-partisans – conform to a more positive and less suspicious view of fellow partisans. They also conform to the view that partisans feel socially closer to their co-partisans than to other partisans (Hoffman, McCabe and Smith, 1996; Green, Palmquist and Schickler, 2002; Greene, 2004; Goeree et al., 2008; Eckel and Grossman, 1996)

As discussed above, the fact that our sample may overrepresent the politically sophisticated could lead to inferences which do not generalize to the entire population of partisans. To check against this, we ran each regression on two different subgroups: the low educated and the highly educated, defined as those with at least some university education. In each case, the relative positioning of the parties did not change (i.e. New Democrats revealed more altruism than Liberals, and Liberals sometimes revealed more altruism than Conservatives). The differences of note are that among low education partisans, Liberals do not behave more altruistically than Conservatives; and among high education partisans, Conservative display significantly less altruism than New Democrats. In the pooled analysis among the highly-educated, Liberals appear more generous than Conservatives. These differences in significance aside, we are confident that the relationship between partisanship and altruism generalizes to the larger population of partisans in Canada.⁵

⁵These supplementary tables are available in Annexe J.

6.5 Discussion and Conclusion

Partisans differ from non-partisans and they differ from one another in their views on issues and in their political behavior. We have examined whether partisans differ in their levels of basic altruism as revealed in a series of dictator games. Our results suggest that in two of three cases New Democrats are significantly more altruistic than Conservatives and Liberals. In the case of other partisans, Liberals are significantly more altruistic than Conservatives.

These findings have two important implications. Methodologically, we have shown how we can use an experimental framework and tools borrowed from a cognate discipline – in this case, experimental or behavioural economics – to learn about partisanship. Similar techniques have been used to explore altruism and voting (Fowler, 2006; Fowler and Kam, 2007; Loewen, 2008*a*), interethnic trust (Whitt and Wilson, 2007), and preferences for public spending (Loewen, 2008*b*). The experimental turn in political science (Druckman et al., 2006) still has much to teach us.

The larger substantive implication of these findings is that partisans maintain distinct views of other partisans and that these views are important enough to effect their level of altruism. This conforms to a view of partisanship as not just a perceptual screen for the filtering of information (Campbell et al., 1960; Bartels, 2002) and not just a standing opinion or running tally (Fiorina, 1981). Instead, partisanship involves viewing partisans as distinct social groupings (Green, Palmquist and Schickler, 2002; Greene, 2004), some

of whom are more deserving of help than others. When these findings are combined with more recent findings about the deep-rooted nature of partisanship (Dawes and Fowler, 2008; Settle, Dawes and Fowler, 2008) an important puzzle emerges: do altruistic individuals choose to identify with the New Democratic Party with greater frequency or does identification with the party lead to a more altruistic orientation? Understanding how individuals sort themselves into different parties and then have their views shaped by this identification is an open and important question. That such stark differences in a basic behavior present along partisan lines only increases the importance of this question.

Chapitre 7

Does Compulsory Voting Lead to More Informed and Engaged Citizens: An experimental test (with Henry Milner and Bruce M. Hicks)

Forthcoming at The Canadian Journal of Political Science. Henry Milner initiated the project and secured the agreement of the participating institution and the Director General of Elections. Milner, Hicks and Loewen designed the experiment. Loewen performed the analysis and drafted the journal version of the article (an earlier version appeared as a working paper drafted in equal parts by all three authors). Loewen saw the article through the submission and revision process.

7.1 Introduction

In his well-known Presidential Address to the *American Political Science Association*, Arend Lijphart (1997) called for compulsory voting as a solution to unequal electoral participation in the United States. In doing so, he restated the main arguments of the advocates of compulsory voting. Most importantly, compulsory voting would increase turnout in elections. Second, compulsory voting would lead to a more politically knowledgeable and engaged electorate.¹

There can be no quibble with Lijphart's first assertion, which we regard as a first-order effect of compulsory voting. The cross-sectional (Jackman, 1987; Blais and Carty, 1990; Blais and Dobrzynska, 1998; Franklin, 1996, 2004) and quasi-experimental (Hirczy, 1994) evidence for this claim is clear. Compulsory voting increases turnout in national elections on average by some 10 to 15 percentage points - and even more in regional and local elections. However, the evidence for the second-order effects of compulsory voting is much less clear, at least partly because of the difficulty of making causal claims about cross-national differences in more subjective variables like political knowledge and engagement.

We argue that an experimental approach is an appropriate way in which to address this gap in our knowledge. To this end, we conducted an experiment in the winter of 2007 in the midst of the Quebec provincial election.

¹Lijphart also claimed that compulsory voting could reduce the incentive for attack ads and reduce the influence of money in politics. We do not test these claims.

Our experiment required that one group of (first-time) voters complete two surveys to receive a monetary reward, while another group was also required to vote in the provincial election. That is, they faced a financial penalty if they chose to abstain from voting. We take between-group differences in knowledge, news consumption, and political discussion as measures of the second-order effects of compulsory voting. To anticipate our findings, we find little evidence of the second-order effects of compulsory voting.

In section 2, we briefly review existing knowledge on the second-order effects of compulsory voting. In doing so, we draw a connection between the lack of current evidence and the value of an experimental approach, an approach which is gaining currency in political science (see Druckman et al., 2006). In section 3, we first operationalize Lijphart's second-order claim in the form of three hypotheses and then describe our experimental design and procedure. In section 4, we present our results. We then conclude.

7.2 Existing Knowledge and the Case for Experimentation

We lack a body of systematic empirical knowledge about the second-order effects of compulsory voting. For example, Bilodeau and Blais (2005) could uncover no empirical studies to support Lijphart's claim. To fill the gap, they attempted to substantiate his claim in three ways. They first examined whether citizens in Western European countries with compulsory voting re-

port that they discussed politics more than those in non-compulsory countries. Second, they examined the behaviour of immigrants to New Zealand from compulsory-voting Australia. Third, they examined the behaviour of immigrants to Australia from compulsory voting countries. In each case they sought differences in reported levels of political discussion, interest in politics and attitudes toward voting, but were unable to find evidence of second-order effects due to compulsory voting.

A recent analysis of Belgian survey data by Engelen and Hooghe (2007) could not find evidence of knowledge effects from compulsory voting. They used the hypothetical question “what if voting were not compulsory” to isolate those who vote to avoid sanction. They find evidence that those who vote to avoid sanction are less knowledgeable about and engaged in politics, suggesting that while compulsory voting is effective at bringing the otherwise less engaged to the polls, it is not necessarily effective at increasing their knowledge levels. Another recent study using data from the Polish Election Survey used the same method in reverse, asking non-voters what they would do if voting were compulsory (Czesnik, 2007). Not surprisingly, those who reported voting to avoid sanction were the least interested and knowledgeable. As with the Belgian study, this merely demonstrates that compulsory voting would bring the otherwise less knowledgeable to the polls. Finally, Ballinger (2007) looked at the British and Australian evidence, concluding that Australian respondents are no better-informed about political systems than British respondents.

While all of these studies are informative, they illustrate two methodological obstacles to testing the second-order effects of compulsory voting. First, in contrast to an objective measure like turnout, there is a major problem with cross-national comparability in survey questions tapping political knowledge. It is very difficult to establish that two national scales are measuring the same type and amount of political knowledge (King et al., 2004). Moreover, even if our scales are measuring exactly the same quantities, we cannot be certain that each country requires the same amount of knowledge for effective democratic citizenship. Second, even if one can come up with directly comparable measures of survey knowledge, the analyst will still be confronted with a problem of unobserved heterogeneity. It is entirely plausible that countries which adopt compulsory voting are also those which have a more engaged citizenry than those countries which do not require compulsory voting. Hence, we cannot assume that any observed differences are a function of compulsory voting and not some unobserved variable(s) in the populations (Gerber, Green and Kaplan, 2004).

In the absence of a change in electoral law within a country allowing for a before-and-after quasi-experiment, there is no unambiguous empirical basis for determining the second-order effects of compulsory voting. What is needed, therefore, is a method which decouples the presence of compulsory voting from pre-existing levels of citizen engagement and knowledge. One such method is an experiment which randomly assigns some voters to a treatment which resembles one context (i.e. compulsory voting), while as-

signing others to a control condition. This is an analytical strategy in keeping with the experimental turn in political science (Druckman et al., 2006; McDermott, 2002; Lupia, 2002; Druckman and Lupia, 2007). We now describe and report results from one such experiment.

7.3 Hypotheses and Experimental Design

7.3.1 Hypotheses

Following Lijphart, compulsory could lead to a more informed and engaged electorate (Lijphart, 1997, 10). We operationalize these second-order effects in the form of three hypotheses:

- H1: Those who face a financial incentive against abstention should learn more about politics than those who do not face a similar incentive.
- H2: Those who face a financial incentive against abstention should discuss politics more frequently than those who do not face a similar incentive.
- H3: Those who face a financial incentive against abstention should follow the news more frequently than those who do not face a similar incentive.

To each of these three hypotheses we add this common extension: the second-order effects should be greatest among those who would not otherwise

go to the polls.

To test these hypotheses, we conducted an experiment among eligible-to-vote students at a Montréal CEGEP during the March 2007 provincial election. The logic of our experimental design is quite simple. We recruited a group of students to participate in a study about ‘youth attitudes’, consisting of two surveys administered approximately one-month apart, at either end of a provincial election campaign. All students who completed these surveys were eligible to receive \$25 (CDN).² However, to receive this money a randomly selected subset of the students were also required to vote in the provincial election.³ Accordingly, we were left with two groups, one of which faced a financial disincentive if they chose not to vote, the other of which faced no such disincentive. By comparing differences between these two groups in political knowledge, media news consumption and reported discussion about politics, we are able to draw inferences about the effects of compulsory voting-like incentives on voters, especially first-time voters. We note that those in our treatment condition faced a financial incentive to vote, which is not theoretically identical to the prospect of losing money through a fine (Kahneman and Tversky, 1979*b*; Kahneman, 2003; Cohen and Blum, 2002). However, we feel confident that, for the purposes of our exper-

²This compares with compulsory voting fines of \$20 (AUS) in Australia, and far exceeds fines in countries such as Argentina (approximately \$3.25 to 6.50 (CDN)) or some Swiss cantons (approximately \$3 (CDN)).

³The requirement that all participants be entitled to be paid was a requirement of the Director General of Elections, so that in a formal sense it was not a matter of people being paid to vote.

iment, this sufficiently approximates compulsion. Moreover, it is the closest we could come within reasonable ethical limitations.

The office of the Director General of Elections in Quebec is responsible for the administration of elections in the province, including the registration of voters and the administration of polling stations. Their cooperation made it possible to verify voting by our subjects.⁴ The survey was conducted at Vanier College, a Montreal English-language CEGEP with over 5000 students from a variety of socio-economic groups and a wide number of ethnicities, the majority of whom are in pre-university programs.

7.3.2 Subject Recruitment and Survey Administration

Recruitment occurred in over 60 Vanier College classes, specifically targeting students in pre-university social science and commerce general education courses (i.e. those with minimal admission requirements). The targeted classes were those most likely to contain students who would be at least 18 years of age on Election Day, the voting age in Quebec (as in the rest of Canada). Interested students were asked to fill out a registration form. This form contained ten unrelated questions,⁵ one of which asked if the students expected to vote in the upcoming Quebec election.

Our subject recruitment occurred in two waves. First, once the election was formally announced, 205 students who filled out the forms and who

⁴We should like to note that this required no small effort on the part of the DGE.

⁵This included questions such as “Do you play sports on campus?”, “Do you own a cellphone?”, and “Do you plan to go on to university?”

were eligible to vote were invited by email or telephone to complete the questionnaire in a class-room at the college on a given date and time. Our initial sample included all students who indicated on the recruitment form that they did not expect to vote. The balance of participants were drawn randomly from those who indicated they intended to vote. Half of the 205 were randomly assigned to two treatment rooms and the other half were randomly assigned to two control rooms.⁶ In total, 55 students showed up as instructed. All subjects were administered instructions, a research consent form and a questionnaire, with the only difference being the subjects who attended the two treatment rooms were informed of the future obligation to vote. Subjects in the control group rooms were not informed that any subjects were being asked to vote. The subjects were not told that the survey was associated in any way with the election, only that there would be a second questionnaire in approximately one month's time.

To expand our sample, we then sent out an email or telephoned those who did not turn up at the first invitation and to 255 of the remaining students who had filled out the forms (and stated that it was likely they would vote). We offered the option of either completing the attached survey by email or completing it in a secretary's office on the college campus at a time of their convenience (within a five-day window). Once again, assignment to treat-

⁶Vanier College has two closely-situated campuses. To ensure maximum ease of participation, students were given a choice of coming to a room on either of the campuses, and the time coincided with the weekly universal break when no classes are supposed to be scheduled. This break occurs in the middle of day.

ment was randomly determined (for details concerning the randomization process, see Annexe K). At the end of the first round we had 82 subjects in the control condition and 101 in the treatment condition. Overall, 52 percent of subjects completed the first survey online, while the remainder filled out a paper copy version. Results of this first survey displayed no significant differences between our control and treatment group in political knowledge, political discussion, or media usage. Moreover, we could find no significant bivariate differences in demographics. We take this as evidence of proper randomization and balance (see Annexe K for more details).

The second round of the survey was administered in the five days prior to the election. All subjects from the first round of the survey were emailed the second survey and asked to complete it online or to complete it in on paper at the same secretary's office within the five day window. The email text differed for those in the treatment and control groups only in regards to the obligation to vote. The deadline to complete the second questionnaire coincided with the close of polls on Election Day (March 26, 2007). One-hundred and forty-three (143) students completed the second questionnaire (all but six completed it electronically).

All subjects had to complete and sign a research consent form in person to give permission to the college to provide the DGE with their name and address in order to verify that they had voted. Hence, excluding those who failed to fill out consent forms as well as those who we could not officially confirm had voted, we had 55 subjects in the control group and 66 in the

treatment group at the end of the study.⁷

7.3.3 Survey and Dependent Variables

Those subjects who chose to participate in the experiment (either on paper or online) were all given the same survey. The first survey asked them a number of questions about media usage, political discussion, and attitudes toward politics and political involvement, followed by 11 political knowledge questions.

As the overall purpose of the experiment is to determine whether those who have a financial incentive to vote (or a financial disincentive not to vote) engage more in and learn more about politics, we carefully selected a variety of different knowledge questions. These ranged from questions about the positions of the parties on the issues (e.g. on raising university tuition), to relevant political facts (e.g. which party was in power when the election was called), to knowledge about the elections, (date, and eligibility to vote). In sum, we included a variety of knowledge questions which should distinguish those with a rudimentary knowledge about politics generally and current Quebec politics specifically. We did much the same with the second questionnaire. However, we added several political knowledge questions, bringing the total to 20. Nine repeated the previous questions verbatim, two repeated

⁷The attrition rate between the first and second round surveys was slightly higher among those in the treatment condition than the control condition (32.9% and 34.7% respectively). We have excluded those whom the DGE could not find on a voters list so as to verify their having voted.

them in altered form, and 9 new questions were added, almost all of which were closely linked to developments in the campaign (These questions are available in Annexe L). We are confident that the full battery of questions provides an appropriate instrument for uncovering any significant knowledge differences between our two groups relevant to electoral participation in this time and place. A variable called *Knowledge* measures respondents' political knowledge as a percentage of questions correctly answered.

We measure political discussion using four questions. The first two questions ask respondents how often they follow what is occurring in "government and public affairs" and how closely they have followed the Quebec election. Each question allows four response categories ranging from "Never" to "Most of the time." The next two questions query how often respondents discuss current events with friends and family, with response categories ranging from "Never" to "Very often". We scale each of these responses from 0 to 1, and then create a variable, *Discussion*, which averages these scores. Accordingly, a subject with a high score follows current events and discusses them with friends and family. A subject with a lower score engages in less such discussion.

Our final dependent variable is *Media Usage*. We queried subjects on how many days a week they read the newspaper, watch the national news on TV, listen to news on the radio, or read news on the internet. Our final variable measures the average number of days a week an individual consumes all of these media. Accordingly, a subject with the maximum possible score

(7) would consume all of these media every day, whereas a subject with the lowest possible score (0) would consume none of these media on any day of the week.

7.3.4 Sample Profile

Table 7.1 presents a profile of our final subjects and their scores on relevant variables. Our subject pool certainly reflects what we would expect from a convenience sample at an English CEGEP. Our subjects are young and principally English. While they are likely more interested in politics than their peers who declined to participate in the survey, they cannot easily be described as politically sophisticated. In the first round of the survey, subjects answered less than one in three knowledge questions correctly (28.4%). In the second survey, the percentage of correctly answered questions rose to just 43.1%, and this despite the majority of the questions being repetitions of first round questions. Similarly, our subjects cannot be easily described as “news junkies” or political conversationalists. Indeed, subjects report consuming news on the radio, TV, internet and newspaper less than two days per week. The average subject would only report discussing news with family and friends somewhere between rarely and sometimes. Finally, when we examine the other political activities of our subjects, we do not find strong evidence of political engagement. Just one in twenty subjects have ever written a newspaper or contacted a television or radio program regarding a political issue. Only half of subjects report ever having signed a written or email petition.

We do not have general population statistics with which to compare these scores. However, we are inclined to believe that our sample does not grossly over-represent political sophisticates. Indeed, the growth in knowledge between the first and second rounds of the survey suggests that subjects were capable of learning more over the course of an election. And subjects could certainly increase their media consumption and political discussion if so inclined. In sum, this is a reasonable sample on which to test the proposition that compulsory voting encourages greater political engagement, as growth in these measures over the course of the campaign was possible.

Table 7.1: Sample Profile

	Mean	(S.D.)
Age	18.9	(1.36)
Female	75.2%	(43.3%)
First Language French	19.0%	(39.3%)
First Language English	60.3%	(49.1%)
First Language Other	20.7%	(40.7%)
First Round Knowledge Score	28.4%	(18.5%)
Second Round Knowledge Score	41.6%	(22.7%)
First Round News Consumption	2.72	(1.45)
Second Round News Consumption	2.27	(1.33)
First Round Discussion Score	0.60	(0.16)
Second Round Discussion Score	0.52	(0.16)
Contacted a Newspaper	5.0%	(21.8%)
Called into TV or Radio	5.7%	(23.4%)
Taken part in a protest	22.5%	(41.9%)
Signed a petition	52.1%	(50.1%)

7.4 Results

We find little support in our data for the above hypotheses.⁸ Table 7.2 presents differences in *Knowledge*, *Discussion*, and *Media Usage* according to treatment. The cells under control and treatment present a mean and a standard deviation for each group. The final row provides the results or a t-test of mean differences between those who received the treatment and those who did not.

As can be seen, the overall difference in knowledge scores in the second round between groups under treatment and control conditions is not significant. On average, both groups appear to be able to answer approximately four of ten political knowledge questions correctly.

We next consider the possibility that the treatment students did try to learn more about politics but were unable to do so. We find no evidence that they increased their general engagement with politics through discussion, which could have signalled greater effort at learning. Rather, by the end of the campaign those in control and treatment both appeared to engage in conversation with friends and family somewhere between the “Rarely” and “Sometimes” response categories.

When it comes to media usage, however, there is some indication that subjects in the treatment condition consumed more news by the end of the

⁸As we found no significant differences between treatment and control conditions in our first round scores, we limit the analysis to second round scores. We have done similar analysis using differences between first and second round scores as dependent variables. Substantive results do not change.

campaign than those in the control condition. Media usage does seem to increase with treatment, though not at a 95% significance level. The estimate suggests that those in the control condition on average reported consuming all forms of news an average of 2.05 days out of 7, while those in the treatment condition reported consuming all forms of news 2.43 days out of 7. It is hard to know how much significance to attribute to this as we do not know at which point greater media consumption begins to bestow knowledge benefits, or at which point it signals a more engaged electorate. In itself, it is consonant with the claims of compulsory voting advocates, but it leaves a new puzzle in that it does not manifest itself in any measurable increase in knowledge.

Table 7.2: Effects of compulsory voting treatment on political knowledge, political discussion, and media usage (mean differences)

	Control (S.D)	Treatment (S.D.)	<i>p, Trt>Ctrl</i>
Knowledge Score	0.40 (0.21)	0.43 (0.24)	0.25
Discussion Score	0.52 (0.16)	0.52 (0.17)	0.48
Media Usage	2.05 (1.15)	2.43 (1.45)	0.07
N	55	66	

Aside from our media usage finding, we have not found support for the hypotheses that financially compelling individuals to vote causes them to become more politically attentive and knowledgeable citizens. It is possible that this is because our treatment was simply not strong enough. Indeed, in the case of some subjects, our monetary incentive was not enough to compel them to vote. This reasonably leads to the question of whether we can expect

to find second-order effects where no first-order effects are present. A more fair test of Lijphart's hypotheses would be to exclude those in the treatment condition who did not vote, and to look for effects particularly among those who did not intend to vote at the outset of the study but were assigned the treatment and voted. We test this proposition in Table 7.3. We limit our analysis to those in the control condition who completed both surveys and those in the treatment condition who completed both surveys and voted.⁹ Our approach is to use an OLS regression with the following form:

$$Y(\textit{Knowledge}) = a + \beta_1 * \textit{Treatment} + \beta_2 * \textit{ExpVote} + \beta_3 * \textit{ExpVote} * \textit{Treatment} + \beta_4 * \textit{Allo} + \beta_5 * \textit{French} + \beta_6 * \textit{Female} + \epsilon$$

where:

Treatment indicates the subject was in the treatment condition;

ExpVote indicates the subject initially reported that they expected to vote;

*ExpVote * Treatment* is an interaction between *Treatment* and *ExpVote*;

Allo indicates an Allophone respondent;

French indicates a French respondent; and,

⁹This regression does not include those in the treatment group who we have identified as non-voting. But, it does include non-voters in the control group. The reason for the exclusion is that we want to isolate effects among those for whom the experiment worked (i.e. those who voted) and then compare them to what our "electorate" would look like without compulsory voting (i.e. one which included voters and non-voters).

Female indicates a female respondent;

If we wish to isolate the effect of treatment (treatment=1) on initial non-voters (voter=0), we are left with the following equation:

$$Y(\textit{Knowledge}) = a + \beta_1 * \textit{Treatment} + \beta_4 * \textit{Allo} + \beta_5 * \textit{French} + \beta_6 * \textit{Female} + \epsilon$$

Accordingly, the specific effect of compulsory voting on the knowledge acquisition (or levels of discussion or media usage) among non-voters is captured by the coefficient on Treatment.¹⁰

We should note that we do not include several other variables which we know are related to political knowledge and engagement (see Fournier, 2002). Because we are using a randomly assigned experiment, we can assume that these factors are equally present in both our control and treatment conditions. Including them theoretically should not change the estimated effects of the compulsory voting treatment. Accordingly, we exclude them and rest with a more simple model.

As Table 7.3 demonstrates, while we find a treatment effect on news consumption for those who intended to vote in the first place, we can find no effect of the treatment for those who would otherwise be non-voters. We are unable to reject the null hypothesis that compulsory voting does not increase the news consumption of non-voters. Moreover, on both our knowledge vari-

¹⁰The treatment effect for those who intended to vote and did vote is captured by the addition of the Treatment coefficient and the Treatment*ExpVoter interaction coefficient. Finally, the effect of expecting to vote in the first place is captured by Expected to Vote.

Table 7.3: Effect of treatment on knowledge, news consumption and discussion of politics for voters and non-voters (OLS)

	Knowledge	<i>t</i>	News	<i>t</i>	Discussion	<i>t</i>
Treatment	-0.06	-0.54	-0.70	-1.03	-0.07	-1.03
Expected to Vote	0.11	1.47	-0.03	-0.07	0.12**	2.61
Treatment*Exp to Vote	0.11	0.94	1.40*	1.91	0.06	0.77
Allophone	-0.03	-0.58	-0.10	-0.31	-0.03	-0.81
Francophone	0.10*	1.76	-0.30	-0.90	0.01	0.45
Female	0.00	-0.05	-0.43	-1.41	-0.04	-1.63
Constant	0.30**	3.74	2.52**	5.35	0.47**	9.25
Adjusted R^2	0.08		0.08		0.10	
N	107		103		107	

* = $p < 0.10$, ** = $p < 0.05$

able and discussion variable, we cannot find a significant effect of treatment either among those who intended to vote or those who did not. In sum, our data do not give us any good basis for rejecting the null hypotheses: to the extent that our experiment reproduces a compulsory voting environment, we do not find that compulsory voting boosts political knowledge or discussion about politics. All that is left is a small effect on media usage among those who originally intended to vote.

7.5 Conclusion

There is little question about the first-order effects of compulsory voting. Countries which have compulsory voting exhibit significantly higher levels

of voter turnout. This alone may be enough to recommend its implementation. Its second-order effects, however, are much less established. We have attempted to test one mechanism by which second-order effects may be generated, namely financial compulsion.

If a relationship between compulsory voting and greater political engagement exists, it is likely so for a number of reasons beyond mere financial compulsion. Political parties in compulsory voting environments may expend more effort educating voters. Or, countries with compulsory voting may also possess or develop a political culture which encourages greater engagement in politics. Or, compulsory voting may compel the media to place a greater effort on educating voters. There are, in other words, many plausible mechanisms by which compulsory voting may be associated with increased political engagement. However, as we have argued, we cannot easily adjudicate between these by cross-sectional research alone. An experimental approach can fill some of this gap.

We have used such an approach to answer a very specific question: do the financial incentives of a compulsory voting environment increase citizen knowledge, discussion, and media consumption. Our results suggest that though a sufficient motivator for getting an uninformed voter to the polls, avoiding forgoing money cannot be assumed to be a sufficient motivator for getting him or her to learn more about politics. Our results thus place the ball back in the court of the advocates of compulsory voting; especially those who suggest that individuals will seek out more information so as to make

correct decisions when compelled to vote. This is hardly the end of the story. But advocates of compulsory voting will need to provide a more compelling, empirically-based micro-story about how it makes for better - or at least more informed - citizens.

8

Conclusion

8.1 Introduction

This dissertation set out to demonstrate the application of experimentation in political science, especially in political behaviour and political psychology. It took up a deliberately eclectic set of questions in an attempt to show that several different types of experiments can be deployed to answer several outstanding questions. Taken together, the results suggest that there is much to learn from an experimental approach. By extension, the results also suggest that the relative paucity of experimentation in political science is likely unwarranted. Indeed, it is more likely the result of a non-experimental culture within empirical political science than of the theoretical or empirical unsuitability of experimentation. There is little which theoretically or practically rules out the use of (more) experimentation in most domains of political science. Following a summary of my work, I draw out five more implications of this work for our discipline. The first three are methodological while the last two are substantive. I then suggest future research questions and directions. I finish with a final word.

8.2 Summary

Chapter 2 presented results from a large online survey exploring the relationship between altruism and preferences for public spending. Using a game from behavioural economics (the dictator game), I induced respondents to reveal their levels of altruism or concern for others. Then, using a series of

questions about new public spending programs, I determined the maximum amount of money that each respondent was willing to pay for new programs. In a series of models, I demonstrated that those individuals who reveal a higher degree of altruism also express a greater willingness to pay for new public spending programs. Whereas previous works have demonstrated a link between altruism and support for public spending, this is the first to *observe* altruism in a dictator game experiment and then show its ability to predict public spending preferences. As such, it provides clear evidence of a non-self-interested basis for public spending preferences.

Chapter 3 similarly used a dictator game experiment to explain an important political puzzle. Political scientists have long-wondered why we see such high levels of voter participation when any single voter has infinitesimally low odds of affecting the outcome of an election. To explain this paradox of participation, I present a formal explanation closely modeled on those of Fowler and Kam (2007) and Edlin, Gelman and Kaplan (2007). My model draws attention to two different feelings which partisans can hold. The first is affinity for those who support the same party as them. The second is antipathy towards those who support another party. Because election outcomes matter for the well-being of the supporters of each party, the model demonstrates that those who feel affinity and/or antipathy should be more likely to vote. I confirm this empirically by inducing subjects to reveal their levels of affinity and antipathy in a dictator game experiment and then demonstrating that these two variables both positively predict voter participation. Moreover,

they do so in the face of several conventional explanations, such as partisan identification, education, age, income, gender, election competitiveness, and media attentiveness. This chapter thus uses an experimental game to demonstrate the importance of other-regarding orientations for the decision to participate in politics.

Chapter 4 presented results from a field experimental test of the persuasive capacities of direct mail in a party leadership race. Working with a front-running campaign, we tested whether a candidate who communicated controversial policy positions could be successful in changing the minds of party elites. As such, the work represents several important extensions. It pushes field experimentation into Canada and into elite politics. It also extends this method into the study of the strategic communication of political leaders, an important and growing field of research in political science. Our findings are equally important. Contrary to the campaign's expectations, direct mail failed to persuade delegates. What is more, it appears to have made their evaluations of the candidate and his positions more negative. As such, our results raise an important caveat for those campaigning among elites. The communication of controversial positions risks not only falling on deaf ears but also making elites less favourable towards a candidate.

Chapter 5 presented a statistical method for the analysis of survey experiments in which two arguments are pitted against one another. The Bradley-Terry model, typically used in biology and other natural sciences, was shown to be more efficient at uncovering the differences in power between argu-

ments. It was also shown to be more efficient at uncovering the cause of differences in power between arguments. We demonstrated this using a survey experiment conducted during the 2007 referendum on electoral reform in the province of Ontario. As such, we learned important facts about why the existing electoral system was retained in that province. To wit, there was a general bias for the status quo, but there was also an unambiguous advantage for arguments which appealed to local representation. This substantially advantaged the existing system. More generally, this work shows that matching the theoretical properties of statistical models with the experiments they are analyzing can lead to clearer and more informative results.

Chapter 6 asked whether some partisans in Canada are more altruistic than other partisans. This question has not yet been answered in Canada, though one can imagine that many opinions abound among partisans as to who is the most virtuous! This is likely because altruism is difficult to observe unambiguously in real world settings. We overcome this problem by having partisans reveal their altruism towards others through a series of dictator games. In this controlled environment, we were able to vary the recipient of a respondent's altruism. Comparing differences by recipient revealed two facts about partisanship and altruism unknown until now. First, all partisans in Canada are most altruistic towards their co-partisans and least generous towards other partisans. Their generosity to those whose partisanship they do not know falls between these two extremes. Second, New Democratic partisans are more altruistic, on average, than Liberal and Conservative par-

tisans. Taken together, these results raise important questions about the origins and effects of partisanship.

Chapter 7 asked whether, as claimed by advocates, compulsory voting is effective at increasing the political knowledge, attention, and discussion of voters. The answer to this question has thus far remained unclear, due in no small part to the difficulty of making causal inferences from cross-national survey data. To proffer an answer to this question, we randomly assigned a group of voting-age college students to complete surveys at either end of a provincial election in exchange for a monetary reward. We randomly assigned a second group to similarly complete two surveys, but also to vote in the election, in exchange for the same monetary reward. As the only difference between these two groups is the monetary compulsion to vote among the second group, any differences in political knowledge, attention, or discussion can be attributed to being in a compulsory voting condition. We are unable to reject the hypothesis that compulsory voting has no effect on these variables, with the exception of a weak effect on political attention. While this is hardly the end of the debate, it does represent the most clear extant empirical evidence against the claims of compulsory voting advocates. The strength of this claim is, in no small part, a result of its experimental basis.

8.3 Implications

8.3.1 Methodological Implications

It may not be obvious that such an eclectic lot of chapters would generate clear implications for our discipline. But I suggest that there are, for me at least, three clear methodological lessons to be learned from these chapters. First, collaboration matters. The efficiencies in collaboration do not come from the division of labour, particularly in writing. Indeed, my own experience is that any such efficiency gains are lost in the conversations which precede and follow the writing! But collaboration is clearly beneficial for the generation of ideas, for the refining of arguments, and for the execution of experiments. As I was not randomly assigned to these papers and these collaborators, I cannot say what the resulting work would have been like had I toiled alone. However, it is my own experience that in the case of this work, it was made better by collaboration. The larger data from our discipline would suggest that experimental work is more collaborative and also exerts a larger influence on our discipline. I would suggest that these elements are all related, and I recommend them strongly going forward.

Second, while experiments may appear prohibitively expensive and practically difficult, they need not be in practice. The unrivalled clarity of the insights offered by experiments makes them appealing not only to academics but also to practitioners. Each experiment presented in this dissertation represented a partnership between the author(s) and some implementing or-

ganization, whether a political campaign, a commercial polling firm, or a government body and policy think tank. By asking clear questions and generating clear answers of value to practitioners, political scientists can develop experimental studies which others are willing to aid in implementing. We should not, then, be put off by the initial apparent difficulty in funding and implementing experiments.

Third, experimental political science presents many opportunities for arbitrage, which in turn allows experimentalists to have an impact on many questions of interest. Because experimental methods can offer clarity where previous observational approaches have provided conflicting accounts, experimentalists can make a contribution to existing questions at a lower cost than their observationalist counterparts. Consider the example of the second-order effects of compulsory voting. Rather than relying on potentially incomparable survey questions in a number of countries, and rather than waiting on some country to change from compulsory to non-compulsory voting, or vice-versa, we were able to quickly and efficiently provide key evidence in this debate. Similarly, taking the example of altruism and public spending, by embedding an experimental game in an online survey, I was able to provide the most clear evidence to date of the importance of altruism for public spending preferences. Rather than searching for another survey through which this could be demonstrated with some marginally different question on altruism, or searching for yet another opaque estimation technique, I simply observed the revealed altruism of respondents and measured

their preferences. In more colloquial terms, I was able to get in and get out with minimal bleeding. Because so few questions in political science have received an experimental treatment, experimentalists can have a disproportionate impact. Given the simplicity of experimentation and the ease with which experimental results can be interpreted, this should appeal to political scientists of all vintages.

Taken together, then, this work suggests that experimentation has important methodological role to play in political science. It is a role which is collaborative, practical, and efficient in the generation of knowledge.

8.3.2 Substantive Implications

As outlined in the summaries, each of these chapters has made a contribution to the questions which they set out to answer. Despite being a diverse lot, two larger substantive implications can be drawn from these findings. The first substantive implication is that citizen politics are characterized by an imbalance in power between persuasion attempts in favour of a status quo and those in favour of a change in direction. I presented field experimental evidence of the difficulty one front-running leadership campaign had in persuading party elites to adopt positions – and by extension a candidacy – outside of the party's mainstream. What is notable about this is not only did the persuasion attempt set the candidate backwards, but it arguably set the party backwards as well. That is, party elites rejected a proposed change which may have made the party better off in the long-term. One explana-

tion for this resistance to change is that party elites were biased against any change from the party's median position, despite its appeal to the broader electorate. We similarly noted in our Bradley-Terry experiment on electoral reform that a constant advantage was enjoyed by arguments for the status quo. This advantage was not a result of the virtues or features of the existing system. Instead, the advantage appeared shared by *all* arguments for the status quo. For scholars of public opinion and political psychology especially, these results add to our existing knowledge about the difficulty in persuading citizens to change their minds (see, e.g. Lodge and Taber, 2005). They also carry implications for those theorists who desire a more deliberative politics in which arguments meet on equal footing (see, e.g. Elster, 1998).

The second substantive implication of this research is that the political behaviour of individuals is insufficiently explained by egoist and self-interested considerations. Instead, I have shown in three different chapters the other-regarding orientations of citizens matter for how they choose to participate in politics and how they form their preferences. In the case of two chapters, I showed that such considerations work as effective *explanators* of preferences and behaviour. In a third chapter, I showed that such considerations, in this case altruism, act as *indicators* which in turn give us important insights into the differences between partisans in Canada. We need not abandon methodological individualism to account for non-egoist and non-self-interested considerations in political behaviour. We merely need to recognize that the well-being of others plays an important part for some citizens in their own

utility calculations. To disregard this is to risk specifying incomplete models of political behaviour. To embrace this (or at least to begrudgingly try it on for a period of time) is to potentially better explain important aspects of political behaviour and political psychology.

8.4 Future Research

I have argued that this dissertation has demonstrated that experimentation can be widely-applied in political science. I have also argued that this work has demonstrated important substantive insights for our understanding of politics, particularly in highlighting the importance of cognitive or preferential biases and the importance of other-regarding orientations.

Going forward, I envision a research program which extends the embedding of games-based experiments in surveys and takes seriously the two substantive insights highlighted above.¹ When these insights are combined with the observation that there is an increasing interest in the genetic basis of other-regarding behaviour generally (Rushton, 2004; Scourfield et al., 2004; de Quervain et al., 2004; Knafo et al., 2008; Cesarini et al., 2008) and political behaviour in particular (Dawes and Fowler, 2008; Settle, Dawes and Fowler, 2008) an exciting set of questions emerge. I outline three such examples below.

First, a series of experiments could explore whether we can explain indi-

¹Indeed, the first two questions which follow are those I proposed to take up in the most recent round of postdoctoral applications and for which I have received funding.

viduals' differences in rates of political and community participation according to their revealed levels of trust. A large body of literature on "social capital" has argued that this connection exists. However, it typically relies on expressed levels of generalized trust in others. What is lacking is a behavioural measure of trust in which an individual's propensity to cooperate and to trust others is observed. By embedding in a survey a multiple-round ultimatum game in which utility maximization depends on cooperation, we could induce survey respondents to reveal their levels of basic trust. By then correlating this trust with respondents' reported levels of participation, we can gain insight into the extent to which basic social trust is necessary for collective action.

Second, a series of experiments could be conducted to examine whether individuals' differences in vote choice can be explained by the variance observed in their willingness to punish others in ultimatum games. Ultimatum games can reveal a voter's preference for fair outcomes or egalitarianism (Dawes et al., 2007) and their willingness to bear a cost for more fair outcomes. Such preferences have been used to explain vote choice for left-wing political parties (e.g. Deth and Scarbrough, 1995). However, left-wing vote choice has have never been shown to be related to observed individual preference for egalitarianism. By correlating individual behaviour in a series of ultimatum games and reported vote choice in several elections, we can determine if preferences for egalitarianism effectively explain vote choice for left-wing parties.

Third, we could explore the degree to which preferences for the status quo and biases against change are genetically structured. For example, by combining a series of loss-aversion experiments (see, e.g. Kahneman and Tversky, 1979*a*) in an online study with a DNA sample of respondents, or alternately a twins study, we could estimate both the heritability of loss aversion and the relationship between loss aversion and preferences for the status quo in questions of public opinion and vote choice. Such a research program may seem costly and prohibitively complex. However, I would argue that the benefits of understanding how both our genes and our environments structure fundamental political behaviours necessitates ambitious research programs, even if the initial climb seems steep. Indeed, the growing field of genetics and politics has much to teach us about what we observe in the social world of politics. We should not shy away from this possibility and we should recognize that experimentation provides a useful tool for forays into this field. Finally, I note that such research likely has several interested partners who could be convinced to share in the costs.

Taken together, the combination of experimentation and a substantive interest in other-regarding preferences and preferential biases points to an interesting research program growing forward. It is one which would not have been clear to me had I not first undertaken the preceding experiments.

8.5 The Sound Down the Hall

A final word. Much of the task of writing this dissertation has happened at night. As the night gets on and the time between the buses home grows, I have found it easier to push on for a few more hours. It is in the quiet of the night that I've understood best what it is I have tried to accomplish in this dissertation.

But as I have entered the watch that ends the night, I think I've often heard a knocking around the corner. It is not the security guard with the keys who passes by around midnight. Nor is it the cleaning lady who ghosts around a few hours after him, with her Spanish music playing on her stereo. It is not just the wind and I am certain it is not Lionel Groulx. No, it is either the clacking tongue of another inferential monster ready to offer up an alternative explanation for my data or it is the sound of Warren Miller's cowboy boots coming down the hall. Either way, I welcome them in my office.

Références bibliographiques

- Abramson, Paul R., John H. Aldrich and David W. Rhode. 2006. *Change and Continuity in the 2004 Elections*. Washington: Congressional Quarterly Press.
- Achen, Christopher H. 2002. "Toward a New Political Methodology: Micro-foundations and ART." *Annual Review of Political Science* 5:423–50.
- Agresti, Alan. 2002. *Categorical Data Analysis*. 2 ed. New York: John Wiley and Sons.
- Aldrich, John H. 1993. "Rational Choice and Turnout." *American Journal of Political Science* 37:246–278.
- Allenby, Greg M, Robert P Leone and Lichung Jen. 1999. "A Dynamic Model of Purchase Timing with Application to Direct Marketing." *Journal of the American Statistical Association* 94(446):365–74.
- Andreoni, James and John Miller. 2002. "Giving According to GARP: An

Experimental Test of the Consistency of Preferences for Altruism.” *Econometrica* 70(2):737–53.

Ansolabehere, Stephen and Alan Gerber. 1994. “The Mismeasure of Campaign Spending: Evidence from the 1990 U.S. House Elections.” *Journal of Politics* 56(4):1106–18.

Ansolabehere, Stephen, James M. Snyder and Charles Stewart. 2001. “Candidate Positioning in U.S. House Elections.” *American Journal of Political Science* 45(1):136–159.

Ansolabehere, Stephen and Shanto Iyengar. 1997. *Going Negative: How Political Advertisements Shrink and Polarize the Electorate*. New York: The Free Press.

Aragonés, Enriqueta and Zvika Neeman. 2000. “Strategic Ambiguity in Electoral Competition.” *Journal of Theoretical Politics* 12(2):183–204.

Aristotle. 322BC/1991. *Rhetoric*. New York: Oxford University Press (translation by George A. Kennedy).

Aronfreed, Justin. 1980. *Conduct and Conscience*. New York: Academic Press.

Ballinger, Chris. 2007. Compulsory Voting: Palliative Care for Democracy in the UK. Presented at the ECPR Joint Sessions workshop on “Compulsory Voting: Principles and Practice”, Helsinki, May 7 - 12, 2007.

- Barlett, Dorothy L, Pamela B Drew, Eleanor G Fahle and William A Watts. 1974. "Selective Exposure to a Presidential Campaign Appeal." *Public Opinion Quarterly* 38(2):264-70.
- Barry, Brian. 1978. *Sociologists, Economists, and Democracy*. Chicago: University of Chicago Press.
- Bartels, Larry and Henry Brady. 1992. The State of Quantitative Political Methodology. In *Political Science: The State of the Discipline II*, ed. Ada W. Finifter. Washington: American Political Science Association.
- Bartels, Larry M. 1987. "Candidate Choice and the Dynamics of the Presidential Nominating Process." *American Journal of Political Science* 31(1):1-30.
- Bartels, Larry M. 1988. *Presidential Primaries and the Dynamics of Public Choice*. Princeton: Princeton University Press.
- Bartels, Larry M. 2002. "Beyond the Running Tally: Partisan Bias in Political Perception." *Political Behavior* 24:117-150.
- Batson, C. Daniel. 1994. "Why act for the public good? Four answers." *Personality and Social Psychology Bulletin* 20:603-610.
- Batson, C. Daniel, A. Christine Harris, Kevin D McCaul, Michael Davis and Timothy Schmidt. 1979. "Comparison or compliance? Alternative dispositional attributions for one's own helping behavior." *Social Psychology Quarterly* 42:405-409.

- Batson, C. Daniel, Bruce D. Duncan, Paula Ackerman, Terese Buckley and Kimberly Birch. 1981. "Is empathic emotion a source of altruistic motivation?" *Journal of Personality and Social Psychology* 40:290–302.
- Batson, C. Daniel, Jay S. Coke, M.L. Jasnosi and Michael Hanson. 1978. "Buying kindness: Effect of an extrinsic incentive for helping on perceived altruism." *Personality and Social Psychology Bulletin* 4:86–91.
- Bawa, Kapil and Robert W Shoemaker. 1989. "Analyzing Incremental Sales From a Direct Mail Coupon Promotion." *Journal of Marketing* 53:66–78.
- Benz, Matthias and Stephan Meier. 2008. "Do people behave in experiments as in the field? Evidence from donations." *Experimental Economics*.
- Berelson, Bernard, Paul Lazarsfeld and William McPhee. 1954. *Voting*. Chicago: University of Chicago Press.
- Best, Samuel J., Brian Krueger, Clark Hubbard and Andrew Smith. 2005. "An Assessment of the Generalizability of Internet Surveys." *Social Science Computer Review* 131:131–145.
- Bibby, Reginald W. 1990. *Fragmented Gods*. Toronto: Stoddart.
- Bilodeau, Antoine and André Blais. 2005. "Le vote obligatoire a-t-il un effet de socialisation politique?" Presented to Colloque int. vote obligatoire, Inst. d'études Polit. Lille, Oct. 20-21.

- Blais, André. 2000. *To Vote or Not to Vote? The Merits and Limits of Rational Choice*. Pittsburgh PA: University of Pittsburgh Press.
- Blais, André. 2005. "Accounting for the Electoral Success of the Liberal Party in Canada." *Canadian Journal of Political Research* 38:821–840.
- Blais, André. Forthcoming. Introduction. In *To Adopt or Not Adopt First Past the Post*, ed. André Blais. Cambridge: Cambridge University Press.
- Blais, André and Agnieszka Dobrzynska. 1998. "Turnout in Electoral Democracies." *European Journal of Political Research* 33(2):239–61.
- Blais, André, Donald Blake and Stéphane Dion. 1993. "Do Parties Make a Difference? Parties and the Size of Government in Liberal Democracies." *American Journal of Political Science* 37:40–62.
- Blais, André, Donald Blake and Stéphane Dion. 1997. *Governments, Parties and Public Sector Employees : Canada, United States, Britain and France*. Pittsburgh: University of Pittsburgh Press.
- Blais, Andre, Elisabeth Gidengil, Neil Nevitte and Richard Nadeau. 2004. "Where does turnout decline come from?" *European Journal of Political Research* 43(2):221–236.
- Blais, Andre, Elisabeth Gidengil, Richard Nadeau and Neil Nevitte. 2001. "Measuring Party Identification: Britain, Canada, and the United States." *Political Behavior* 23(1):5–22.

- Blais, André, Elisabeth Gidengil, Richard Nadeau and Neil Nevitte. 2002. Do Party Supporters Differ? In *Political Behaviour: Theory and Practice in a Canadian Context*, ed. Joanna Everitt and Brenda O'Neill. Toronto: Oxford University Press.
- Blais, André, Elisabeth Gidengill, Richard Nadeau and Neil Nevitte. 2002. *Anatomy of a Liberal Victory: Making Sense of the Vote in the 2000 Canadian Election*. Broadview Press.
- Blais, André and Louis Massicotte. 2002. Electoral Systems. In *Comparing Democracies 2: New Challenges in the Study of Elections and Voting*, ed. Lawrence LeDuc, Richard G Niemi and Pippa Norris. London: Sage.
- Blais, André and Marc-André Bodet. 2006. "How Do Voters Form Expectations about the Parties' Chances of Winning the Election?" *Social Science Quarterly* 87(3):477–493.
- Blais, André and R. Kenneth Carty. 1990. "Does Proportional Representation Foster Voter Turnout." *European Journal of Political Research* 18(2):167–181.
- Blalock, Herbert M. 1979. *Social Statistics*. Toronto: McGraw-Hill.
- Bositis, David A, Denise L Baer and Roy E Miller. 1985. "Cognitive Information Levels, Voter Preferences, and Local Partisan Political Activity: A Field Experimental Study on the Effects of Timing and Order of Message Presentation." *Political Behavior* 7(3):266–84.

- Bositis, David A. and Douglas Steinel. 1987. "A synoptic history and typology of experimental research in political science." *Political Behavior* 9:263–284.
- Bowler, Shaun and Todd Donovan. 1998. *Demanding Choices: Opinion, Voting and Direct Democracy*. Ann Arbor MI: University of Michigan Press.
- Bowles, Samuel and Herbert Gintis. 2002. *Origins of Human Cooperation*. In *Genetic and Cultural Evolution of Cooperation*, ed. Peter Hammerstein. Cambridge: MIT Press.
- Bradley, Ralph Allan and Milton E. Terry. 1952. "Rank analysis of incomplete block designs, I. the method of paired comparisons." *Biometrika* 39:324–345.
- Brady, Henry, Sidney Verba and Kay Lehman Schlozman. 1995. "Beyond SES: A Resource Model of Political Participation." *British Journal of Political Science* 89(2):271–94.
- Bult, Jan Roelf and Tom Wansbeek. 1995. "Optimal Selection for Direct Mail." *Marketing Science* 14(4):378–94.
- Burden, Barry C. and Casey A. Klofstad. 2005. "Affect and Cognition in Party Identification." *Political Psychology* 26(6):869–886.
- Camerer, Colin F. 2003. *Behavioral Game Theory: Experiments in Strategic Interaction*. Princeton: Princeton University Press.

- Camerer, Colin F. and Robin M. Hogarth. 1999. "The Effects of Financial Incentives in Experiments: A Review and Capital-Labor-Production Framework." *Journal of Risk and Uncertainty* 19(1-3):7-42.
- Campbell, Angus, Philip E. Converse, Warren E. Miller and Donald E. Stokes. 1960. *The American Voter*. New York: John Wiley and Sons.
- Campbell, Donald T. and Julien C. Stanley. 1966. *Experimental and quasi-experimental designs for research*. Skokie, IL: Rand McNally.
- Canadian Press. 2006. "Debate Transcripts." Transcript of 2006 Canadian federal leaders debate. Available at: www.theglobeandmail.com/generated/realtime/specialDecision2006.html.
- Carpenter, Jeffery, Erik Verhoogen and Stephen Burks. 2005. "The Effect of Stakes in Distribution Experiments." *Economics Letters* 86(2):393-98.
- Carter, David B. and Arthur Spirling. Forthcoming. "Under the Influence? Intellectual Exchange in Political Science." *PS: Political Science and Politics*.
- Carty, R. Kenneth and Munroe Eagles. 2005. *Politics Is Local: National Politics at the Grassroots*. Don Mills: Oxford University Press.
- Carty, R. Kenneth, William Cross and Lisa Young. 2000. *Rebuilding Canadian Party Politics*. Vancouver: UBC Press.

- Cesarini, David, Christopher T. Dawes, James H. Fowler, Magnus Johannesson, Paul Lichtenstein and Björn Wallace. 2008. "Heritability of cooperative behavior in the trust game." *Proceedings of the National Academy of Sciences* 105:3721–3726.
- Chamberlain, Gary and Michael Rothchild. 1981. "A Note on the Probability of Casting a Decisive Vote." *Journal of Economic Theory* 25:152–162.
- Cherry, Todd L., Peter Frykblom and Jason F. Shogren. 2002. "Hardnose the Dictator." *The American Economic Review* 92(4):1218–1221.
- Chong, Dennis and James N. Druckman. 2007a. "Framing Public Opinion in Competitive Democracies." *American Political Science Review* 101(4):637–55.
- Chong, Dennis and James N. Druckman. 2007b. "Framing Theory." *Annual Review of Political Science* 10:103–26.
- Clarke, Harold D., David Sanders, Marianne C. Stewart and Paul Whiteley. 2004. *Political Choice in Britain*. Oxford: Oxford University Press.
- Clarke, Harold D., Jane Jenson, Lawrence LeDuc and Jon Pammett. 1979. *Political Choice in Canada*. Toronto: McGraw-Hill Ryerson.
- Cohen, Jonathan D. and Kenneth I. Blum. 2002. "Reward and Decision." *Neuron* 36:193–198.

- Corneo, Giacomo and Hand Peter Grüner. 2002. "Individual Preferences for Political Redistribution." *Journal of Public Economics* 83:83–107.
- Cross, William. 1996. "Direct Election of Provincial Party Leaders in Canada, 1985-1995: The End of the Leadership Convention?" *Canadian Journal of Political Science* 29(2):295–315.
- Cross, William. 2005. "The Rush to Electoral Reform in the Canadian Provinces: Why Now?" *Representation* 41(2):75–84.
- Czesnik, Mikolaj. 2007. Is Compulsory Voting a Remedy? Evidence from the 2001 Polish Parliamentary Elections. Presented at the ECPR Joint Sessions workshop on "Compulsory Voting: Principles and Practice", Helsinki, May 7 - 12, 2007.
- Dalton, Russell J., ed. 2006. *Citizen Politics: Public Opinion and Political Parties in Advanced Industrial Democracies*. Fourth ed.
- Dawes, Christopher T. and James H. Fowler. 2007. "Social Preferences and Political Participation." University of California-San Diego. Working Paper.
- Dawes, Christopher T. and James H. Fowler. 2008. "Partisanship, Voting, and the Dopamine D2 Receptor Gene." University of California-San Diego. Working Paper.
- Dawes, Christopher T., James H. Fowler, Tim Johnson, Richard McEl-

- reath and Oleg Smirnov. 2007. "Egalitarian Motives in Humans." *Nature* 446:794–796.
- de Quervain, Dominique J.-F., Urs Fischbacher, Valerie Treyer, Melanie Schellhammer, Ulrich Schnyder, Alfred Buck and Ernst Fehr. 2004. "The Neural Basis of Altruistic Punishment." *Science* 305(5688):1254–1258.
- Deth, Jan W. Van and Elinor Scarbrough, eds. 1995. *The Impact of Values*. Oxford: Oxford University Press.
- Dewan, Torun and David P. Myatt. 2008. "The Qualities of Leadership: Direction, Communication and Obfuscation." LSE Political Science and Political Economy Group Working Paper: PSPE:WP1:08.
- Doherty, Daniel, Alan S. Gerber and Donald P. Green. 2006. "Personal Income and Attitudes toward Redistribution: A Study of Lottery Winners." *Political Psychology* 27:441–458.
- Downs, Anthony. 1957. *An Economic Theory of Democracy*. New York: Harper and Row.
- Druckman, James N. 2001. "The Implications of Framing Effects for Citizen Competence." *Political Behavior* 23:225–256.
- Druckman, James N. 2004. "Political Preference Formation: Competition, Deliberation and the (Ir)relevance of Framing Effects." *American Political Science Review* 98:671–86.

- Druckman, James N. and Arthur Lupia. 2007. Mind, Will, and Choice: Lessons from Experiments in Contextual Variation. In *The Oxford Handbook on Contextual Political Analysis*, ed. Charles Tilly and Robert E. Goodin. Oxford University Press.
- Druckman, James N., Donald P. Green, James H. Kuklinski and Arthur Lupia. 2006. "The Growth and Development of Experimental Research in Political Science." *American Political Science Review* 100(4):627–637.
- Dunleavy, Patrick and Helen Margetts. 1999. "Mixed Electoral Systems in Britain and the Jenkins Commission on Electoral Reform." *British Journal of Politics and International Relations* 1(1):12–38.
- Eckel, Catherine C. and Philip J. Grossman. 1996. "Altruism in Anonymous Dictator Games." *Games and Economic Behavior* 16(2):181–191.
- Eckel, Catherine C. and Philip J. Grossman. 1998. "Are Women Less Selfish than Men? Evidence from Dictator Experiments." *Economic Journal* 108(448):726–735.
- Edlin, Aaron, Andrew Gelman and Noah Kaplan. 2007. "Voting as a Rational Choice: Why and How People Vote To Improve the Well-Being of Others." *Rationality and Society* 19(3):293–314.
- Elsner, Ralf, Manfred Krafft and Arnd Huchzermeier. 2004. "Optimizing Rhenania's Direct Marketing Business Through Dynamic Multilevel Mod-

eling (DMLM) in a Multicatalog-Brand Environment.” *Marketing Science* 23(2):192–206.

Elster, Jon, ed. 1998. *Deliberative Democracy*. Cambridge: Cambridge University Press.

Engelen, Bart and Marc Hooghe. 2007. Compulsory Voting and its Effects on Political Participation, Interest and Efficacy. Presented at the ECPR Joint Sessions workshop on “Compulsory Voting: Principles and Practice”, Helsinki, May 7 - 12, 2007.

Erikson, Robert S. and Gerald C. Wright. 1997. Voters, Candidates, and Issues in Congressional Elections. In *Congress Reconsidered*, ed. Lawrence C. Dodd and Bruce I. Oppenheimer. Washington DC: Congressional Quarterly Press.

Evans, Geoffrey, ed. 1999. *The End of Class Politics? Class Voting in Comparative Context*. Oxford: Oxford University Press.

Feldman, Stanley and Marco Steenbergen. 2001. “The Humanitarian Foundations of Public Support for Welfare.” *American Journal of Political Science* 45:658–677.

Fenlon, Brodie. 2007. “Referendum on Electoral Reform an ‘Unmitigated Disaster’.” *The Globe and Mail*. October 11, 2007, Online edition.

URL: <http://theglobeandmail.com>

- Ferejohn, John. 2004. Internal and External Explanation. In *Problems and Methods in the Study of Politics*, ed. Ian Shapiro, Rogers Smith and Tarek Masoud. Cambridge: Cambridge University Press.
- Ferejohn, John A. and Morris P. Fiorina. 1974. "The Paradox of Not Voting: A Decision Theoretic Analysis." *American Political Science Review* 68:525–546.
- Fiorina, Morris P. 1981. *Retrospective Voting in American National Elections*. New Haven: Yale University Press.
- Firth, David. 2005. "Bradley-Terry Models in R." *Journal of Statistical Software* 12.
- Flanagan, Tom. 2003. "Database Party: The 2002 Leadership Campaign for the Canadian Alliance." *Canadian Parliamentary Review* 26.
- Fong, Christina M. 2001. "Social Preferences, Self-Interest, and the Demand for Redistribution." *Journal of Public Economics* 82:225–246.
- Fong, Christina M., Samuel Bowles and Herbert Gintis. 2006. *Strong reciprocity and the welfare state*. Vol. 1 of *Handbook on the Economics of Giving, Reciprocity and Altruism* Elsevier chapter 23, pp. 1439–1464.
- Forsythe, Robert, Joel L. Horowitz, N. E. Savin and Martin Sefton. 1994. "Fairness in Simple Bargaining Experiments." *Games and Economic Behavior* 6(3):347–369.

- Fournier, Patrick. 2002. The Uninformed Canadian Voter. In *Citizen Politics: Research and Theory in Canadian Political Behaviour*, ed. Joanna Everitt and Brenda O'Neill. Toronto: Oxford University Press.
- Fowler, James H. 2006. "Altruism and Turnout." *Journal of Politics* 68:647–683.
- Fowler, James H. and Cindy Kam. 2007. "Beyond the Self: Altruism, Social Identity, and Political Participation." *Journal of Politics* 69:813–827.
- Franklin, Mark N. 1996. Electoral Participation. In *Comparing Democracies: Elections and Voting in Global Perspective*, ed. Laurence LeDuc, Richard G Niemi and Pippa Norris. Thousand Oaks CA: Sage.
- Franklin, Mark N. 2002. The Dynamics of Electoral Participation. In *Comparing Democracies 2: New Challenges in the Study of Elections and Voting*, ed. Laurence LeDuc, Richard G, Niemi and Pippa Norris. Thousand Oaks CA: Sage.
- Franklin, Mark N. 2004. *Voter Turnout and the Dynamics of Electoral Competition in Established Democracies Since 1945*. Cambridge: Cambridge University Press.
- Frey, Bruno S. and Stephan Meier. 2004. "Pro-social behavior in a natural setting." *Journal of Economic Behavior and Organization* 54:65–88.
- Frohlich, Norman and Joe A. Oppenheimer. 1990. "Choosing Justice in

Experimental Democracies with Production.” *American Political Science Review* 84:461–477.

Gallagher, Michael. 1998. “The Political Impact of Electoral System Change in Japan and New Zealand, 1996.” *Party Politics* 4(2):203–228.

Gelman, Andrew, Jonathan N. Katz and Joseph Bafumi. 2004. “Standard Power Voting Indexes Don’t Work: An Empirical Analysis.” *British Journal of Political Science* 34(4):657–674.

Gerber, Alan S. 2004. “Does Campaign Spending Work?: Field Experiments Provide Evidence and Suggest New Theory.” *American Behavioral Scientist* 47(5):541–74.

Gerber, Alan S. and Donald P. Green. 2000. “The Effects of Canvassing, Direct Mail, and Telephone Contact on Voter Turnout: A Field Experiment.” *American Political Science Review* 94(3):653–63.

Gerber, Alan S. and Donald P. Green. 2005. “Correction to Gerber and Green (2000), Replication of Disputed Findings, and Reply to Imai (2005).” *The American Political Science Review* 99(2):301–313.

Gerber, Alan S, Donald P Green and Edward H Kaplan. 2004. The Illusion of Learning from Observational Research. *in* Shapiro, Smith and Massoud (2004) pp. 251–73.

Gibson, James L., Gregory A. Caldeira and Lester Kenyatta Spence. 2002.

"The Role of Theory in Experimental Design: Experiments Without Randomization." *Political Analysis* 10(4):362–375.

Gibson, Rachel and Andrea Rommele. 2001. "Changing Campaign Communications: A Party-Centered Theory of Professionalized Campaigning." *The Harvard International Journal of Press Politics* 4:31–43.

Gidengil, Elisabeth. 2002. Bringing Politics Back In: The Study of Public Opinion in Canada. In *Citizen Politics: Research and Theory in Canadian Political Behavior*, ed. Joanna Everitt and Brenda O'Neill. Toronto: Oxford University Press.

Goeree, Jacob K., Margaret A. McConnell, Tiffany Mitchell, Tracey Tromp and Leeat Yariv. 2008. "Linking and Giving Among Teenage Girls." California Institute of Technology Working Paper.

Gosnell, Harold F. 1930. *Why Europe Votes*. Chicago: University of Chicago Press.

Green, Donald P. and Alan S. Gerber. 2002. Reclaiming the Experimental Tradition in Political Science. In *Political Science: The State of the Discipline, 3rd Edition*, ed. Helen V. Milner and Ira Katznelson. New York: W.W. Norton & Co.

Green, Donald P. and Alan S. Gerber. 2004. *Get Out the Vote*. Washington: Brookings Institution Press.

- Green, Donald P., Bradley Palmquist and Eric Schickler. 2002. *Partisan Hearts and Minds: Political Parties and the Social Identities of Voters*. New Haven: Yale University Press.
- Greene, Steven. 2004. "Social Identity Theory and Party Identification." *Social Science Quarterly* 85(1):136–153.
- Haider-Markel, D and M Joslyn. 2001. "Gun Policy, Opinion, Tragedy, and Blame Attribution: the Conditional Influence of Issue Frames." *Journal of Politics* 63:520–43.
- Hamill, Ruth C., Milton Lodge and Frederick Blake. 1985. "The Breadth, Depth, and Utility of Class, Partisan, and Ideological Schemata." *American Journal of Political Science* 29:850–870.
- Hamilton, Howard D. 1970. "Direct Legislation: Some Implications of Open Housing Referenda." *American Political Science Review* 64(1):124–137.
- Hirczy, Wolfgang. 1994. "The Impact of Mandatory Voting Laws on Turnout: A Quasi-Experimental Approach." *Electoral Studies* 13:64–76.
- Hoffman, Eizabeth, Kevin McCabe and Vernon L. Smith. 1996. "Social Distance and Other-Regarding Behavior in Dictator Games." *The American Economic Review* 86(3):653–660.
- Holbrook, Thomas M and Scott D McClurg. 2005. "The Mobilization of Core Supporters: Campaigns, Turnout, and Electoral Composition in

- United States Presidential Elections." *American Journal of Political Science* 49(4):689–703.
- Horiuchi, Yusaku and Jun Saito. 2003. "Reapportionment and Redistribution: Consequences of Electoral Reform in Japan." *American Journal of Political Science* 47(4):669–682.
- Hout, Mike, Clem Brooks and Jeff Manza. 1993. "The Persistence of Classes in Post-Industrial Societies." *International Sociology* 8(3):259–277.
- Huckfeldt, Robert and John Sprague. 1992. "Political Parties and Electoral Mobilization: Political Structure, Social Structure, and the Party Canvass." *The American Political Science Review* 86(1):70–86.
- Hudson, John and Philip R. Jones. 1994. "The importance of the 'ethical voter': An estimate of altruism." *Journal of Political Economy* 10:499–509.
- Imai, Kosuke. 2005. "Do get-out-the-vote calls reduce turnout? The importance of statistical methods for field experiments." *American Political Science Review* 99(2):283–300.
- Irons, Karl W, John DC Litle and Robert L Klein. 1983. "Determinants of Coupon Effectiveness." *Advances and Practices of Marketing Science*, Proceedings of the 1983 ORSA/TIMS Marketing Science Conference (March).
- Jackman, Robert W. 1987. "Political Institutions and Voter Turnout in the Industrial Democracies." *American Political Science Review* 81(2):405–24.

- Johnson, Blair T, Aaron Smith-McLallan, Ley A Killeya and Kenneth D Levin. 2003. Truth or Consequences: Overcoming Resistance to Persuasion with Positive Thinking. *in* Knowles and Linn (2003) pp. 215–33.
- Johnston, Richard. 1992. “Party Identification and Campaign Dynamics.” *Political Behavior* 14:311–331.
- Johnston, Richard, Andre Blais, Elisabeth Gidengil and Neil Nevitte. 1996. *The Challenge of Direct Democracy: The 1992 Canadian Referendum*. Montreal: McGill-Queen’s University Press.
- Johnston, Ronald J and Charles J Pattie. 1998. “Campaigning and Advertising: An Evaluation of the Components of Constituency Activism at Recent British General Elections.” *British Journal of Political Science* 28(4):677–85.
- Kahneman, Daniel. 2003. “A Psychological Perspective of Economics.” *The American Economic Review* 93:162–168.
- Kahneman, Daniel and Amos Tversky. 1979a. “Prospect Theory: An Analysis of Decision under Risk.” *Econometrica* 157:263–291.
- Kahneman, Daniel and Amos Tversky. 1979b. “Prospect theory: An analysis of decisions under risk.” *Econometrica* 47:313–327.
- Karlan, Dean S. 2005. “Using Experimental Economics to Measure Social Capital And Predict Financial Decisions”. Working Papers 909 Economic Growth Center, Yale University.

- Kiewiet, D. Roderick. 1983. *Macroeconomics and Micropolitics*. Chicago: University of Chicago Press.
- Kinder, Donald R. and D. Roderick Kiewiet. 1981. "Sociotropic Politics: The American Case." *British Journal of Political Science* 11:129–141.
- King, Gary, Christopher J.L. Murray, Joshua A. Saloman and Ajay Tandon. 2004. "Enhancing the Validity and Cross-Cultural Comparability of Measurement in Survey Research." *American Political Science Review* 98(1):191–207.
- King, Gary and Langche Zeng. 2001. "Logistic Regression in Rare Events Data." *Political Analysis* 9(2):137–163.
- Knafo, Ariel, S. Leon Israel, Ariel Darvasi, Rachel Bachner-Melman, Florina Uzevovsky, L. Jonathan Cohen, Elliot J. Feldman, Ehud Lerer, Efrat Laiba, Yifat Raz, Lubov Nemanov, Inga Gritsenko, Christian Dina, Galila Agam, Brian Dean, George Bornstein and Richard P. Ebstein. 2008. "Individual differences in allocation of funds in the dictator game associated with length of the arginine vasopressin 1a receptor RS3 promoter region and correlation between RS3 length and hippocampal mRNA." *Genes and Brain Behavior* 7:266–275.
- Knowles, Eric S and Jay Linn, eds. 2003. *Resistance and Persuasion*. Mahwah NJ: Lawrence Erlbaum Associates.

- Krueger, James S. and Michael S. Lewis-Beck. 2008. "Is OLS Dead?" *The Political Methodologist* 15(2):2-5.
- Kruger, Ribert F., Brian M. Hicks and Matt McGue. 2001. "Altruism and antisocial behaviour: Independent tendencies, unique personality correlates, distinct etiologies." *Psychological Science* 12:397-402.
- Leamer, Edward E. 1983. "Let's Take the Con Out of Econometrics." *American Economic Review* 73(1):31-43.
- LeDuc, Lawrence. 1984. Canada: The Politics of Stable Dealignment. In *Electoral Change in Advanced Industrial Democracies: Realignment or Dealignment*, ed. Russell J. Dalton, Paul Beck and Scott Flanagan. Princeton: Princeton University Press.
- Leighley, Jan E. and Jonathan Nagler. 1992a. "Individual and Systemic Influences on Turnout: Who Votes? 1984." *The Journal of Politics* 54(3):718-740.
- Leighley, Jan E. and Jonathan Nagler. 1992b. "Socioeconomic Class Bias in Turnout, 1964-1988: The Voters Remain the Same." *The American Political Science Review* 86(3):725-736.
- Lewis, Alan and Diane Jackson. 1985. "Voting Preferences and Attitudes to Public Expenditure." *Political Studies* 33(3):457-466.
- Lewis-Beck, Michael. 1988. *Economics and Elections*. Ann Arbor, MI: University of Michigan Press.

- Lijphart, Arend. 1997. "Unequal Participation: Democracy's Unresolved Dilemma." *American Political Science Review* 91(1):1-14.
- Lodge, Milton and Charles S Taber. 2000. Three Steps Toward a Theory of Motivated Reasoning. In *Elements of Reason: Cognition, Choice, and the Bounds of Rationality*, ed. Arthur Lupia, Matthew D McCubbins and Samuel L Popkin. Cambridge: Cambridge University Press pp. 183-213.
- Lodge, Milton and Charles S. Taber. 2005. "The Automaticity of Affect for Political Leaders, Groups, and Issues: An Experimental Test of the Hot Cognition Hypothesis." *Political Psychology* 26(3):455-482.
- Lodge, Milton and Ruth Hamill. 1986. "A Partisan Schema for Political Information Processing." *American Political Science Review* 80:505-520.
- Loewen, Peter John. 2005. "How Do Local Candidates Spend Their Money? And Does it Matter?" Paper presented at the Canadian Political Science Association Annual Meeting.
- Loewen, Peter John. 2008a. "Antipathy, Affinity, and Political Participation: How Our Concern for Others Makes Us Vote." Working paper. Université de Montréal.
- Loewen, Peter John. 2008b. "Dictators and Purses: Altruism and Support for Public Spending." Working paper. Université de Montréal.
- Loewen, Peter John, Greg Lyle and Jennifer S. Nachshen. 2007. "An eight-item form of the Empathy Quotient (EQ) and an application to charitable

giving.” Working paper. Université de Montréal. Available at:http://www.crcee.umontreal.ca/pdf/Eight_Question_20ES_final.pdf.

London, Harvey, Philip J. Meldman and A. Van C. Lanckton. 1970. “The Jury Method: How the Persuader Persuades.” *Public Opinion Quarterly* 34:171–83.

Luce, R. Duncan. 1959. *Individual Choice Behaviour: A Theoretical Analysis*. New York: J. Wiley.

Lupia, Arthur. 2002. “New Ideas in Experimental Political Science.” *Political Analysis* 10:319–324.

Lupia, Arthur and John G. Matsusaka. 2004. “Direct Democracy: New Approaches to Old Questions.” *Annual Review of Political Science* 7:463–82.

MacCoun, Robert J. 2003. “Review of Shadish, Cook, and Campbells *Experimental and quasi-experimental designs for generalized causal inference*.” *Journal of Policy Analysis and Management* 22:330–332.

Mansbridge, Jane. 1990. Expanding the Range of Formal Modeling. In *Beyond Self Interest*, ed. Jane Mansbridge. Chicago: University of Chicago Press.

Margolis, Howard. 1983. *Selfishness, Altruism, and Rationality*. Cambridge: Cambridge University Press.

- Margolis, Howard. 1990. Dual Utilities and Rational Choice. In *Beyond Self Interest*, ed. Jane Mansbridge. Chicago: University of Chicago Press.
- Matsusaka, John G. 2005. "Direct Democracy Works." *Journal of Economic Perspectives* 19(2):185–206.
- McCrone, Donald J. and James H. Kuklinski. 1979. "The Delegate Theory of Representation." *American Journal of Political Science* 23:278–300.
- McDermott, Rose. 2002. "Experimental Methodology in Political Science." *Political Analysis* 10(4):325–42.
- McKelvey, Richard D. and Thomas R. Palfrey. 1992. "An Experimental Study of the Centipede Game." *Econometrica* 60(4):803–836.
- Meirowitz, Adam. 2005. "Informational Party Primaries and Strategic Ambiguity." *Journal of Theoretical Politics* 17(1):107–36.
- Miller, R.E. and K.L. Robyn. 1975. "A Field Experimental Study of Direct Mail in a Congressional Primary Campaign: What Effects Last Until Election Day?" *Experimental Study of Politics* 5:1–37.
- Miller, Warren E. and Merrill Shanks. 1996. *The New American Voters*. Cambridge: Harvard University Press.
- Miller, Warren E., Roy Pierce, Jacques Thomassen, Richard Herrera, Soren Holmberg, Peter Esaiasson and Bernhard Wessels. 1999. *Policy Representation in Western Democracies*. Oxford: Oxford University Press.

- Monroe, Burt L. 1995. "Fully Proportional Representation." *American Political Science Review* 89(4):925–40.
- Morelli, Massimo. 1999. "Demand Competition and Policy Compromise in Legislative Bargaining." *The American Political Science Review* 93(4):809–820.
- Morton, Rebecca. 2007. Why the Centipede Game Experiment Is Important for Political Science. In *Positive Changes in Political Science: The Legacy of Richard D. McKelvey's Most Influential Writings*, ed. John H. Aldrich, James E. Alt and Arthur Lupia. Ann Arbor: University of Michigan Press.
- Morton, Rebecca and Kenneth Williams. 2006. Experimentation in Political Science. In *Oxford Handbook of Political Methodology*, ed. Janice Box-Steffensmeiser, David Collier and Henry Brady. New York: Oxford University Press.
- Mueller, Dennis C. 1986. "Rational egoism versus adaptive egoism as fundamental postulate for a descriptive theory of human behavior." *Public Choice* 51:3–23.
- Mueller, Dennis C. 1989. *Public Choice II*. Cambridge: Cambridge University Press.
- Mulligan, Casey B. and Charles G. Hunter. 2003. "The Empirical Frequency of a Pivotal Vote." *Public Choice* 116(1-2):31–54.

- Mutz, Diana C., Paul M. Sniderman and Richard Brody, eds. 1996. *Political Persuasion and Attitude Change*. Ann Arbor: University of Michigan Press.
- Nash, Edward L. 1984. *The Direct Marketing Handbook*. New York: McGraw-Hill.
- Neslin, Scott A, Carline Henderson and John Quelch. 1985. "Consumer Promotions and the Acceleration of Product Purchases." *Marketing Science* 4(2):147–65.
- Nickerson, David W. 2005. "Scalable Protocols Offer Efficient Design for Field Experiments." *Political Analysis* 13(3):233–252.
- Niemi, Richard G. 1976. "Costs of Voting and Nonvoting." *Public Choice* 27:115–119.
- Norris, Pippa. 1997. "Choosing Electoral Systems." *International Political Science Review* 18(3):297–312.
- Norris, Pippa. 2002. *Democratic Phoenix: Reinventing Political Activism*. Cambridge: Cambridge University Press.
- Norris, Pippa. 2003. *A Virtuous Circle: Political Communications in Post-Industrial Democracies*. Cambridge: Cambridge University Press.
- OECD. 2007. *OECD Factbook 2007: Economic, Environmental and Social Statistics*. Organization for Economic Development and Cooperation.

- Page, Benjamin. 1978. *Choices and Echoes in Presidential Elections*. Chicago: University of Chicago Press.
- Paltiel, Khayyam Z. 1974. "Party and Candidate Expenditures in the Canadian General Election of 1972." *Canadian Journal of Political Science* 7(2):341-52.
- Pattie, Charles J, Ronald J Johnston and Edward A Fieldhouse. 1995. "Winning the Local Vote: The Effectiveness of Constituency Campaign Spending in Great Britain, 1983-1992." *American Political Science Review* 89(4):969-83.
- Peffley, Mark and Jon Hurwitz. 2007. "Resistance and Persuasion in Response to Arguments against The Death Penalty Among Whites and Blacks." *American Journal of Political Science* 51(4).
- Perlin, George, ed. 1988. *Party Democracy in Canada: The Politics of National Party Conventions*. Scarborough: Prentice-Hall Canada Inc.
- Pierson, Paul. 1996. "The New Politics of the Welfare State." *World Politics* 48(2):143-179.
- Piliavin, Jane Allyn and Hong-Wen Charng. 1990. "Altruism: A review of recent theory and research." *Annual Review of Sociology* 16(1):27-65.
- Pinker, Steven. 2007. "A History of Violence." *The New Republic* March 17, 2007.

- Powell, G Bingham. 2000. *Elections as Instruments of Democracy: Majoritarian and Proportional Visions*. New Haven CT: Yale University Press.
- R Development Core Team. 2006. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. ISBN 3-900051-07-0.
URL: <http://www.R-project.org>
- Rasinski, Kenneth A and Susan M. Rosenbaum. 1987. "Support of Tax Increases for Education: A Comparison of Two Social Psychological Perspectives." *Journal of Applied Social Psychology* 17:990–1006.
- Ridley, Mark and Richard Dawkins. 2003. The Natural Selection of Altruism. In *Altruism and Helping Behavior*. Hillsdale: Lawrence Earlbaum Associates.
- Riker, William H. and Peter C. Ordeshook. 1968. "A Theory of the Calculus of Voting." *American Political Science Review* 62:25–42.
- Roncek, Dennis W. 1992. "Learning More From Tobit Coefficients: Extending a Comparative Analysis of Political Protest." *American Sociological Review* 57(4):503–507.
- Rosenstone, Steven and John Mark Hansen. 1994. *Mobilization, Participation and Democracy in America*. New York: MacMillan.
- Roth, Alvin E. 1995. Bargaining Experiments. In *The Handbook of Exper-*

- imental Economics*, ed. John H. Kagel and Alvin E. Roth. Princeton: Princeton University Press.
- Rushton, J. Phillippe. 1982. "Altruism and Society: A Social Learning Perspective." *Ethics* 3:425–446.
- Rushton, J. Phillippe. 2004. "Genetic and environmental influences on antisocial behavior: a meta-analysis of twin and adoption studies." *Proceedings of the Royal Society* 271:2583–2585.
- Scourfield, Jane, Bethan John, Neilson Martin and Peter McGuffin. 2004. "The development of prosocial behavior in children and adolescents: a twin study." *Journal of Child Psychology and Psychiatry* 45:927–935.
- Sears, David O. and Carolyn L. Funk. 1990. The role of self-interest in social and political attitudes. In *Beyond Self-Interest*, ed. Jane J. Mansbridge. Chicago: University of Chicago Press.
- Settle, Jaime, Christopher T. Dawes and James H. Fowler. 2008. "The Heritability of Partisan Attachment." University of California-San Diego. Working Paper.
- Shadish, William R., Thomas D. Cook and Donald T. Campbell. 2000. *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*. Boston: Houghton Mifflin.
- Sham, Pak and Dave Curtis. 1995. "An Extended Transmis-

- sion/Disequilibrium Test (TDT) for Multiallele Marker Loci." *Annals of Human Genetics* 59(3):323–36.
- Shapiro, Ian, Rogers Smith and Tarek Massoud, eds. 2004. *Problems and Methods in the Study of Politics*. Cambridge: Cambridge University Press.
- Sharp, Carol and Milton Lodge. 1985. "Partisan and Ideological Belief Systems: Do They Differ?" *Political Behavior* 7:147–166.
- Shepsle, Kenneth A. 1970. "A Note On Zeckhauser's 'Majority Rule with Lotteries on Alternatives': The Case of the Paradox of Voting." *Quarterly Journal of Economics* 84(4):705–10.
- Shepsle, Kenneth A. 1972. "The Strategy of Ambiguity: Uncertainty and Electoral Competition." *American Political Science Review* 66(2):555–69.
- Shiell, Alan and Janelle M. Seymour. 2002. "Preferences for public health insurance: egotism or altruism?" *Journal of Public Economics* 29:356–369.
- Sigelman, Lee. 1990. "Hair Loss and Electability: The Bald Truth." *Journal of Nonverbal Behavior* 14(4):269–83.
- Signorino, Curtis S. 1999. "Strategic Interaction and the Statistical Analysis of International Conflict." *The American Political Science Review* 93(2):279–297.
- Smith, Adam. 1790. *The Theory of Moral Sentiments*. Sixth (2006) ed. Sao Paulo: MetaLibri.

- Smith, Adam. 2007. *An Inquiry into the Nature and Causes of The Wealth of Nations*. Sao Paulo: MetaLibri.
- Smith, Mark A. 2001. "The Contingent Effects of Ballot Initiatives and Candidate Races on Turnout." *American Journal of Political Science* 45(3):700–706.
- Sniderman, Paul M. 1992. The New Look in Public Opinion Research. In *Political Science: The State of the Discipline II*, ed. Ada W. Finifter. Washington: American Political Science Association.
- Soroka, Stuart N. and Christopher Wlezien. 2005. "Opinion Policy Dynamics: Public Preferences and Public Expenditure in the United Kingdom." *British Journal of Political Science* 35:665–689.
- Sorrentino, Richard M. 1991. "Evidence for Altruism: The Lady is Still in Waiting." *Psychological Inquiry* 2:147–150.
- Spirling, Arthur. 2007. "Power Tool: Measuring Power in Political Science—A New Methods with Application to the United States Senate." *University of Rochester Working Paper*.
- Springall, Anthony S. 1973. "Response surface fitting using a generalization of the Bradley-Terry paired comparisons model." *Applied Statistics* 22:59–68.
- Staub, Ervin. 1978. *Positive social behavior and morality*. Vol. 1 New York: Academic Press.

- Staub, Ervin. 1980. *"Personality: Basic aspects and current research"*. Englewood Cliffs: Prentice-Hall.
- Staub, Ervin. 1989. Individual and Societal (Group) Values in a Motivational Perspective and Their Role in Benevolence and Harmdoing. In *Social and Moral Values: Individual and Societal Perspectives*, ed. Janusz Reykowski Nancy Eisenberg and Ervin Staub. Hillsdale, NJ: Erlbaum.
- Stevens, Jeffrey R. and Marc D. Hauser. 2004. "Why be nice? Psychological constraints on the evolution of cooperation." *Trends in Cognitive Science* 8:60-65.
- Stewart, David K. 1997. "The Changing Leadership Electorate: An Examination of Participants in the 1992 Alberta Conservative Leadership Election." *Canadian Journal of Political Science* 30(1):107-28.
- Stigler, Stephen M. 1994. "Citation Patterns in the Journals of Statistics and Probability." *Statistical Science* 9:94-108.
- Strate, John M., Charles J. Parrish, Charles D. Elder and Coit Ford. 1989. "Life Span Civic Development and Voting Participation." *The American Political Science Review* 83(2):443-464.
- Strategic Counsel, The. 2006. "Liberal Leadership: Poll of Liberal leadership delegates conducted November 21-27, 2006. www.thestrategiccounsel.com."

- Stuart-Fox, Devi M., David Firth, Adnan Moussalli and Martin J. Whiting. 2006. "Multiple Signals in Chameleon Contests: Designing and Analysing Animal Contests as a Tournament." *Animal Behaviour* 71:1263–1271.
- Tajfel, Henri. 1978. Social Categorization, Social Identity, and Social Comparisons. In *Differentiation Between Social Groups*, ed. Henri Tajfel. London: Academic Press.
- Taylor, Humphrey. D. and Randall Thomas. 2005. "How Does Social Desirability Affect Responses ? Differences in Telephone and Online." Paper presented at the annual meeting of the American Association For Public Opinion Association, Fontainebleau Resort, Miami Beach, FL,.
- Thurstone, Louis Leon. 1959. *The Measurement of Values*. Chicago: University of Chicago Press.
- Uhlener, Carole J. 1986. "Political Participation, Rational Actors, and Rationality: A New Approach." *Political Psychology* 7:551–573.
- Uhlener, Carole J. 1989a. "Rational Turnout: The Neglected Role of Groups." *American Journal of Political Science* 33:390–422.
- Uhlener, Carole J. 1989b. "Relational Goods' and Participation: Incorporating Sociability into a Theory of Rational Action." *Public Choice* 62:253–285.
- Uhlener, Carole J. 1999. What the Downsian Voter Weighs: A Reassessment of the Costs and Benefits of Action. In *Information, Participation and*

- Choice: An Economic Theory of Democracy in Perspective*, ed. Bernard Grofman. Ann Arbor: University of Michigan Press pp. 67–79.
- Vanderleeuw, James and Richard Engstrom. 1987. "Race, Referendums, and Roll-Off." *Journal of Politics* 49(4):1081–1092.
- Vavreck, Lynn, Constantine J Spiliotes and Linda L Fowler. 2002. "The Effects of Retail Politics in the New Hampshire Primary." *American Journal of Political Science* 46(3):595–610.
- Verba, Sidney, Kay Lehman Schlozman and Henry E. Brady. 1995. *Voice and Equality: Civic Volunteerism in American Politics*. Cambridge MA: Harvard University Press.
- Vitaglione, Guy D. and Mark A. Barnett. 2003. "Assessing a new dimension of empathy: Empathic anger as a predictor of helping and punishing desires." *Motivation and Emotion* 27:301–325.
- Vowles, Jack, Peter Aimer, Susan Banducci and Jeffrey Karp. 1998. *Voters Victory? New Zealands First Election Under Proportional Representation*. Auckland: Auckland University Press.
- Wakabayashi, Akio, Simon Baron-Cohen, Sally Wheelwright, Nigel Goldfeld, Joe Delaney, Debra Fine, Richard Smith and Lenora Weil. 2006. "Development of short forms of the Empathy Quotient (EQ-Short) and the Systemizing Quotient (SQ-Short)." *Personality and Individual Differences* 41:929–940.

- Wantchekon, Leonard. 2003. "Clientalism and Voting Behavior: Evidence from a Field Experiment in Benin." *World Politics* 55:399–422.
- Wearing, Joseph. 1988. The High Cost of High Tech: Financing the Modern Leadership Campaign. in Perlin (1988) pp. 72–83.
- Wells, Paul. 2006. *Right Side Up: The Fall of Paul Martin and the Rise of Stephen Harper's New Conservatism*. Toronto: Douglas Gibson Books.
- Whiteley, Paul F. and Patrick Seyd. 1994. "Local Party Campaigning and Electoral Mobilization in Britain." *Journal of Politics* 56(1):242–52.
- Whitt, Sam and Rick K. Wilson. 2007. "The Dictator Game, Fairness, and Ethnicity in Postwar Bosnia." *American Journal of Political Science* 51:655–668.
- Wilcoxon, Frank. 1945. "Individual Comparisons by Ranking Methods." *Biometrics Bulletin* 1:80–83.
- Zaller, John. 1992. *The Nature and Origins of Mass Opinion*. Cambridge: Cambridge University Press.
- Zeckhauser, Richard. 1969. "Majority Rule with Lotteries on Alternatives." *Quarterly Journal of Economics* 83(2):696–703.

Annexe A: Criterion Validity of Dictator Game as a Measure of Altruism

In this annexe, I provide further evidence of the dictator game as an indicator of altruism. Table A.1 shows the relationship between dictator game behavior and self-reported charitable donations in the last two years. I measure altruism dichotomously, as in Model 3 in Tables 2.2–2.5, and use an ordered logit. Charitable giving was measured with the following question: “Thinking about the past 2 years, what was the total amount of all your charitable donations together?” Respondents could choose from one of eight categories:

- Did not give to charity over past 2 years
- Under \$500
- \$500 to \$999
- \$1000 to \$4999
- \$5000 to \$9999

- \$10000 to \$49999
- \$50000 to \$99999
- More than \$100000
- Prefer not to say

I exclude those who did not disclose their giving (2.93% of the total sample). I note that disclosure is unrelated to altruism in a bivariate logits using each of the three specifications of altruism.

As can be seen in Table A.1, altruistic behavior in a dictator game is positively and significantly related to self-reported charitable giving. According to the estimates, those who gave away any money in a dictator game had odds of being in the highest donation category 38% greater than those who gave away nothing in the dictator game.

Table A.2 shows the relationship between altruistic behavior and disagreement with two statements. The first statement is: "It is difficult for me to contain my feelings when I see people in distress." The second statement is "I have little compassion for people in need who are unwilling to take the first step to help themselves." Respondents were allowed to indicate that they strongly agree (1), somewhat agree (2), somewhat disagree (3), and strongly disagree (4) with the statements. Altruistic behavior in a dictator game should negatively predict disagreement with the first statement and positively predict disagreement with the second. Tables A.2 and A.3 suggests this is just the case. In the case of responding to those in distress,

those who gave away money in the dictator game have odds of answering strongly disagree over all other categories only 78% as large than those who give away nothing. Similarly, the odds of those who give away money in the dictator game have strongly disagreeing with the second statement are only 38% larger than those who give away nothing. Taken together, these three results increase the content validity of the dictator game as an indicator of an altruistic orientation.

Table A.1: Altruism and Self-Reported Charitable Giving (Ordered Logistic Regression)

	Model 3	<i>S.E.</i>	<i>p</i>
Altruism	1.39	0.11	.00
Age	9.19	2.64	.00
Income	3.14	0.39	.00
Education	2.43	0.30	.00
French	0.36	0.05	.00
Female	0.95	0.08	.52
Employed	0.82	0.19	.38
Unemployed	0.37	0.12	.00
Self-Employed	1.20	0.20	.28
Student	0.60	0.17	.07
Homemaker	0.86	0.20	.52
Public Sector	1.18	0.23	.40
Making Ends Meet	0.72	0.04	.00
Job Loss Worry	0.88	0.04	.01
NDP ID	0.72	0.14	.10
Conservative ID	1.49	0.23	.01
Liberal ID	0.86	0.13	.36
BQ ID	0.61	0.16	.07
Cut 1	1.29		
Cut 2	1.60		
Cut 3	2.69		
Cut 4	4.58		
Cut 5	5.50		
Cut 6	7.51		
Cut 7	8.06		
Cut 8	8.40		
LR χ^2	712.02		
LR $\chi^2 > \sim$ Altruism Model	0.00		
N=2160			

Table A.2: Altruism and Response to those in Distress (Ordered Logistic Regression)

	Model 3	<i>S.E.</i>	<i>p</i>
Altruism	0.78	0.07	.00
Age	0.22	0.07	.00
Income	1.19	0.15	.15
Education	1.18	0.14	.17
French	1.59	0.21	.00
Female	0.54	0.05	.00
Employed	0.62	0.15	.05
Unemployed	0.83	0.27	.56
Self-Employed	0.73	0.12	.06
Student	0.86	0.24	.57
Homemaker	0.75	0.16	.18
Public Sector	1.74	0.35	.01
Making Ends Meet	0.90	0.05	.07
Job Loss Worry	0.85	0.04	.00
NDP ID	0.62	0.12	.02
Conservative ID	1.17	0.18	.31
Liberal ID	0.65	0.10	.01
BQ ID	0.70	0.18	.17
Cut 1	-2.84		
Cut 2	-0.43		
Cut 3	1.71		
LR χ^2	215.01		
LR $\chi^2 > \sim$ Altruism Model	0.00		
N=2183			

Table A.3: Altruism and Willingness to Help Those Who Don't Help Themselves First (Ordered Logistic Regression)

	Model 3	<i>S.E.</i>	<i>p</i>
Altruism	1.38	0.11	.00
Age	2.44	0.68	.00
Income	1.03	0.12	.79
Education	1.83	0.21	.00
French	1.16	0.15	.25
Female	1.74	0.15	.00
Employed	1.90	0.44	.01
Unemployed	2.47	0.77	.00
Self-Employed	1.41	0.23	.04
Student	1.52	0.40	.12
Homemaker	2.11	0.44	.00
Public Sector	0.61	0.12	.01
Making Ends Meet	1.04	0.06	.41
Job Loss Worry	0.99	0.05	.78
NDP ID	5.08	0.98	.00
Conservative ID	0.52	0.08	.00
Liberal ID	1.56	0.24	.00
BQ ID	2.36	0.59	.00
Cut 1	0.16		
Cut 2	2.19		
Cut 3	3.63		
LR χ^2	261.80		
LR $\chi^2 > \sim$ Altruism Model	0.00		
N=2185			

Annexe B: Dictator Game Instructions

The complete text of the dictator game experiment is as follows:

In addition to our normal \$500 cash prize for completing the survey, we will be drawing four other prizes at the end of this survey. One person in this study will be randomly chosen to receive each prize.

In each draw, the prize is \$100. Should you win any of the draws, your answer to the questions below will determine the amount of each prize that you receive. Remember that your answer is completely anonymous.

(1) Below, you will see two boxes. In the first box, enter how much of a \$100 prize you would keep if you won one of the additional draws. In the other box, indicate how much you'd like to give away to an anonymous individual who will also be randomly chosen. You know nothing about this anonymous individual.

You must choose how to divide the \$100 between yourself and the anonymous individual. You may keep all, none, or some of the money - the decision is up to you and will be completely anonymous. The total of the two boxes must add up to \$100. Once you have made your decision, please hit next.

(2) Below, you will see two boxes. In the first box, enter how much of a \$100 prize you would keep if you won one of the additional draws. In the other box, indicate how much you'd like to give away to an anonymous individual who will also be randomly chosen. You know nothing about this anonymous individual except that they support the Conservative Party.

You must choose how to divide the \$100 between yourself and the anonymous individual. You may keep all, none, or some of the money - the decision is up to you and will be completely anonymous. The total of the two boxes must add up to \$100. Once you have made your decision, please hit next.

(3) Below, you will see two boxes. In the first box, enter how much of a \$100 prize you would keep if you won one of the additional draws. In the other box, indicate how much you'd like to give away to an anonymous individual who will also be randomly chosen. You know nothing about this anonymous individual except that they support the Liberal Party

You must choose how to divide the \$100 between yourself and the anonymous individual. You may keep all, none, or some of the money - the decision is up to you and will be completely anonymous. The total of the two boxes must add up to \$100. Once you have made your decision, please hit next.

(4) Below, you will see two boxes. In the first box, enter how much of a \$100 prize you would keep if you won one of the additional draws. In the other box, indicate how much you'd like to give away to an anonymous individual who will also be randomly chosen. You know nothing about this anonymous individual except that they support the New Democratic Party

You must choose how to divide the \$100 between yourself and the anonymous individual. You may keep all, none, or some of the money - the decision is up to you and will be completely anonymous. The total of the two boxes must add up to \$100. Once you have made your decision, please hit next.

The order of questions 1-4 was randomized.

Annexe C: Question Wording and Variables for Dictators and Purses

Altruism is the amount of money given away to the completely anonymous individual in the dictator game. The amount ranges from \$0 to \$100. In Model 1, Altruism is rescaled from 0 to 1. In Model 2, Altruism is transformed to 0 for those who give away nothing, 1 for those who give away something less than or equal to the median (\$20), and 2 for those who give away more than the median. In Model 3, Altruism reads 1 for those who gave away something and 0 for those who give away nothing.

Age is a six category variable measuring age group. It is rescaled to 1. Values are 18-24 (0), 25-34 (1/6), 35-44 (2/6), 45-54 (3/6), 45-54 (4/6), 55-64 (5/6), 65 and older (1).

Income is a four category variable measuring household income in the last year. It is rescaled from 0 to 1. Values are <\$40000 (0), \$40000 to \$60000

(1/3), \$60000 to \$80000 (2/3) and >\$80000 (1).

Education is a three category variable measuring highest level of education. It is rescaled from 0 to 1. Values are high school or less (0), at least some college (1/2), and at least some university (1).

French is a dummy variable reading 1 if a respondent completed the survey in French and 0 otherwise.

Female is a dummy variable reading 1 if a respondent is a female and 0 otherwise.

Employed, Unemployed, Self-Employed, Student, Homemaker, and Public Sector Employee are all dummy variable reading 1 if a respondent indicates the occupation and 0 otherwise. The question wording is "Which of the following best describes your current job status." Response categories are "working on your own business within your home," "working on your own business outside of your home," "working at an employer's business full-time/part-time," "currently unemployed," "full-time student," "full time student, working part time," "part time student, working full time," "homemaker," and "retired."

Making Ends Meet measures a respondent's agreement with the state-

ment "Thinking about your monthly bills, how difficult is it for you and your family to make ends meet." Responses are "Not difficult at all" (0), "Not very difficult" (1/3), "Somewhat difficult" (2/3), "Very difficult" (1).

Future Job Loss measures the response to the question "How concerned are you that either you or the main household earner might become unemployed in the next six months?" Responses are "Not at all concerned" (0), "Not very concerned" (1/3), "Somewhat concerned" (2/3), "Very concerned" (1).

Partisan identification – Liberal ID, Conservative ID, NDP ID, and BQ ID are all variables reading 1 when a respondent identifies as a strong identifier of a party, 1/2 when they identify as a moderate identifier, and 0 otherwise. Identification is determined with the question "Thinking about federal politics in Canada, generally speaking, do you usually think of yourself as Liberal, Conservative, N.D.P, or none of these?" Those who identified a party then received the standard follow-up: "And, generally speaking, how strongly do you think of yourself as a (party)?"

Annexe D: Question Wording and Variables for Affinity, Antipathy, and Political Participation

Max-Min is the maximum allocation to a partisan less the minimum allocation. The variable is rescaled from 0 to 1.

Max-Mean is the maximum allocation to a partisan less the mean of allocations to other partisans. The variable is rescaled from 0 to 1.

Affinity is the maximum allocation to a partisan less the allocation to a completely anonymous individual. Negative values are censored at 0. The variable is rescaled 0 to 1.

Antipathy is the allocation to a completely anonymous individual less the minimum allocation. Negative values are censored at 0. The variable is rescaled 0 to 1.

Partisan identification – is a variable reading 1 when a respondent identifies as a strong identifier of a party, 1/2 when then identify as a moderate identifier, and 0 otherwise. Identification is determined with the question “Thinking about federal politics in Canada, generally speaking, do you usually think of yourself as Liberal, Conservative, N.D.P, or none of these?” Those who identified a party then received the standard follow-up: “And, generally speaking, how strongly do you think of yourself as a (party)?”

Income is a four category variable measuring household income in the last year. It is rescaled from 0 to 1. Values are <\$40000 (0), \$40000 to \$60000 (1/3), \$60000 to \$80000 (2/3) and >\$80000 (1).

Age is a six category variable measuring age group. It is rescaled to 1. Values are 18-24 (0), 25-34 (1/6), 35-44 (2/6), 45-54 (3/6), 45-54 (4/6), 55-64 (5/6), 65 and older (1).

Education is a three category variable measuring highest level of education. It is rescaled from 0 to 1. Values are high school or less (0), at least some college (1/2), and at least some university (1).

Female is a dummy variable reading 1 if a respondent is a female and 0 otherwise.

Unemployed is a dummy variable reading 1 if a respondent is currently unemployed and 0 otherwise.

Voting is a dummy variable reading 1 when respondents indicate having voted in response to the question: "In talking to people about elections, we find that they are sometimes not able to vote because they're not registered, they don't have the time, or they have difficulty getting to the polls. Did you happen to vote in the last federal election?"

Annexe E: Treatment Assignment for For Want of a Nail

Our treatment assignment procedure occurred in three steps:

- An official list of delegates was provided to campaigns by the Liberal Party of Canada following delegate selection meetings. We first excluded all those who did not have a proper address and then excluded those from three provinces: Quebec, Manitoba and British Columbia. Delegates from Quebec were excluded as they were subject to a different ad campaign by the Ignatieff campaign. Those in Manitoba and British Columbia were excluded because delegate lists were not finalized at the time of treatment assignment due to disputes over the eligibility of several delegates. Using the random number generator function in Excel, we assigned each delegate a random number and then ranked delegates from largest to smallest number. The first 800 delegates were selected for the study. We originally included delegates pledged to support Ignatieff in our first sample because we expected a second campaign to

participate in the experiment. We included Ignatieff delegates to allow us to test the effectiveness of the second campaign's direct mail on delegates committed to other candidates. Ultimately, the second campaign did not participate, but not before we had sent a treatment schedule to the Ignatieff campaign.

- Among the 800 selected delegates, we identified and excluded all those who were not pledged to support Michael Ignatieff. The leadership selection process of the Liberal party requires those who stand as delegate candidates to formally declare their allegiance prior to delegate selection meetings. This information is retained in the official party list. This left 567 delegates.
- Among the remaining delegates, we assigned them a second random number and ranked them from largest to smallest number. The first 100 delegates were assigned to receive two pieces of mail from the Ignatieff campaign. The next 200 delegates were assigned to receive one piece of mail. The remaining delegates (267) were assigned to receive no mail for the period of the study.

In the course of receiving completed surveys we identified as many as four individuals in our control condition who may have been treated by the campaign. Because the campaign eventually mailed every delegate, those from whom we received completed surveys after November 27th may have received mail from the campaign. However, our statistical and substantive

results do not change when we rerun our analyses with these individuals excluded.

Annexe F: Sample and Subject Profile for Bradley-Terry Experiment

Our subjects were drawn from an internet panel containing a list of approximately 15000 specifically named individuals. The panel is managed by a commercial polling firm. Though the panel is national, our experiment was provincial. At the time of the survey the panel had 3575 registered panelists in the province of Ontario. Respondents either self-select into the panel through the surveying firm's website, or they are recruited through solicitation by phone and email. Upon registration in the panel respondents are assigned a number between 1 and 31. In each subsequent month they are invited to complete the survey in one of four weeks according to this number.

Our experiment occurred in the last week of the referendum campaign (October 2 to 9, 2007) and was limited to those living in the province of Ontario. Of 3575 Ontarians registered in the panel, 844 were invited to complete the survey during this week. 565 agreed to complete the survey. Of those 565, 520 agreed to participate in the Bradley-Terry experiment. This

gives us a response rate of 61.6%.

Table F.1 presents a profile of the experimental subjects.

Table F.1: Sample Demographics and Political Characteristics

Variable		%
Age	19–24	1.7%
	25–34	9.6%
	35–44	16.5%
	45–54	26.5%
	55–64	31.7%
	65+	13.9%
Female		52.6%
Household Income	<\$40000	12.89%
	\$40000 to \$60000	16.15%
	\$60000 to \$80000	17.12%
	>\$80000	38.65%
	Refused	15.19%
Education	High School or less	8.85%
	Some College	28.46%
	Some University	62.69%
<i>N</i>	520	

**Annexe G: Additional Logit
Results for Bradley-Terry
Experiment**

Table G.1: Logistic Regression of Argument Power with Argument Matchings

Variable	Model 1			Model 2		
	Coef	SE	<i>p</i> -value	Coef	SE	<i>p</i> -value
FPTP2/MMP1	-0.59	0.78	0.45	-0.69	0.79	0.39
FPTP3/MMP1	0.52	0.84	0.54	0.39	0.86	0.65
FPTP4/MMP1	-0.22	0.81	0.78	-0.31	0.82	0.71
FPTP5/MMP1	0.32	0.80	0.69	0.16	0.81	0.85
FPTP6/MMP1	0.69	0.95	0.47	0.65	0.96	0.50
FPTP1/MMP2	-0.29	0.83	0.73	-0.42	0.84	0.62
FPTP2/MMP2	0.29	0.76	0.71	0.36	0.77	0.65
FPTP3/MMP2	-0.83	0.84	0.32	-1.03	0.86	0.23
FPTP4/MMP2	-0.04	0.77	0.96	-0.35	0.79	0.66
FPTP5/MMP2	-0.00	0.74	1.00	-0.13	0.76	0.87
FPTP6/MMP2	0.59	0.80	0.46	0.45	0.82	0.59
FPTP1/MMP3	-0.51	0.86	0.55	-0.62	0.88	0.49
FPTP2/MMP3	0.43	0.79	0.58	0.33	0.81	0.68
FPTP3/MMP3	-0.04	0.90	0.96	-0.12	0.92	0.90
FPTP4/MMP3	0.06	0.78	0.93	-0.02	0.79	0.98
FPTP5/MMP3	0.36	0.86	0.67	0.20	0.87	0.82
FPTP6/MMP3	0.69	0.95	0.47	0.71	0.97	0.46
FPTP1/MMP4	-0.11	0.75	0.89	-0.23	0.76	0.76
FPTP2/MMP4	0.05	0.80	0.95	0.10	0.82	0.90
FPTP3/MMP4	-1.07	1.00	0.29	-1.15	1.01	0.26
FPTP4/MMP4	0.10	0.74	0.90	-0.07	0.75	0.92
FPTP5/MMP4	-1.00	0.83	0.23	-1.22	0.86	0.15
FPTP6/MMP4	0.36	0.86	0.67	0.31	0.87	0.72
FPTP1/MMP5	0.43	0.79	0.58	0.34	0.80	0.67
FPTP2/MMP5	-1.68	0.97	0.08	-1.82	0.98	0.06
FPTP3/MMP5	0.54	0.78	0.49	0.41	0.80	0.61
FPTP4/MMP5	0.18	0.93	0.85	0.09	0.94	0.92
FPTP5/MMP5	0.34	0.82	0.69	0.14	0.84	0.87
FPTP6/MMP5	0.43	0.79	0.58	0.41	0.80	0.61
FPTP1/MMP6	0.06	0.78	0.93	0.04	0.79	0.96
FPTP2/MMP6	0.05	0.80	0.95	-0.16	0.81	0.84
FPTP3/MMP6	0.36	0.86	0.67	0.14	0.87	0.88
FPTP4/MMP6	-0.04	0.90	0.96	-0.04	0.92	0.97
FPTP5/MMP6	0.77	0.82	0.35	0.50	0.84	0.55
FPTP6/MMP6	0.18	0.77	0.81	-0.02	0.79	0.98
Female				0.45	0.19	0.02
Age				-0.24	0.08	0.00
Constant	-0.18	0.61	0.76	0.24	0.80	0.76
N	520			520		
Pseudo <i>R</i> ²	0.04			0.07		

Annexe H: Predicted Probabilities of FPTP dominance in Structured versus Unstructured Bradley-Terry Models

Given structured and unstructured results, the question remains about which researchers should prefer. We believe this is a matter of what questions a researcher wishes to answer. If we wished to understand the power of exact arguments in the referendum, we would be well-served to consider the unstructured results, as these arguments closely match those made during the campaign. However, if we were looking forward to another campaign and wished to develop new arguments, we could learn more from the structured results. Indeed, these would allow us to design optimal arguments which combined effective components and avoided less effective ones. Fortunately, despite these models giving us different types of information on power, the predictions which result from them are very similar. Table H.1 demonstrates

the absolute difference in predicted probabilities of FPTP dominance according to our unstructured and structured models. The mean absolute difference between the model predictions is just 0.0196.² This suggests that researchers could choose a structured or unstructured model according to their own analytical needs without concern for making inferences greatly different from those they would make with the other model.

Table H.1: Absolute difference in predicted probabilities of FPTP dominance by structured and unstructured models^a

MMP Argument↓	FPTP Argument⇒					
	1	2	3	4	5	6
1	0.05	0.00	0.04	0.02	0.01	0.00
2	0.01	0.03	0.01	0.02	0.02	0.04
3	0.02	0.03	0.01	0.01	0.02	0.03
4	0.04	0.00	0.04	0.02	0.01	0.00
5	0.04	0.01	0.03	0.01	0.00	0.01
6	0.01	0.04	0.01	0.02	0.03	0.04

^a The mean absolute difference is 0.0196.

²We calculate this by taking the average of the absolute difference between the predicted probabilities from the structured and unstructured models.

Annexe I: Question Wording and Variables for Partisanship and Altruism

Allocation to Anonymous Individuals is the amount of money given away to the completely anonymous individual in the dictator game. The amount ranges from \$0 to \$100.

Allocation to Other Partisans is the average of the amount given away to partisan recipients who do not support the identifier's party, i.e. the average of allocations to Liberal and Conservative supporters for New Democratic identifiers. The amount ranges from \$0 to \$100.

Allocation to Co-Partisans is the amount of money given away to the partisan recipient who supports the identifier's party, i.e. the allocation to a New Democrat by New Democratic identifiers. The amount ranges from \$0 to \$100.

Pooled Allocation is the amount of money given away to all recipients. The amount ranges from \$0 to \$100.

Other Partisan is a dummy variable indicating that a Pooled Allocation was to the partisan of a different party than the respondent.

Co-Partisan is a dummy variable indicating that a Pooled Allocation was to a partisan of the same party as the respondent.

Partisan identification – Liberal, Conservative, and New Democrat are all dummy variables reading 1 when a respondent identifies as a moderate or strong identifier of a party and 0 otherwise. Identification is determined with the question “Thinking about federal politics in Canada, generally speaking, do you usually think of yourself as Liberal, Conservative, N.D.P, or none of these?” Those who identified a party then received the standard follow-up: “And, generally speaking, how strongly do you think of yourself as a (party)?” Those who indicated a very strongly or fairly strongly were retained.

Income is a four category variable measuring household income in the last year. It is rescaled from 0 to 1. Values are <\$40000 (0), \$40000 to \$60000 (1/3), \$60000 to \$80000 (2/3) and >\$80000 (1).

Age is a six category variable measuring age group. It is rescaled to 1. Values are 18-24 (0), 25-34 (1/6), 35-44 (2/6), 45-54 (3/6), 45-54 (4/6), 55-64 (5/6), 65 and older (1).

Education is a three category variable measuring highest level of education. It is rescaled from 0 to 1. Values are high school or less (0), at least some college (0.5), and at least some university (1).

Female is a dummy variable reading 1 if a respondent is a female and 0 otherwise.

Unemployed is a dummy variable reading 1 if a respondent is currently unemployed and 0 otherwise.

Empathy is an interval-level variable measuring the average agreement with eight questions measuring empathic capacity. Respondents who score 0 have the lowest empathic capacity; those scoring 1 have the highest capacity. For each question, respondents receive a score of 0, 1, or 2 based on their response of total, somewhat, or no agreement (or disagreement in the case of negatively-keyed questions). The total summed score is then divided by 16. The questions are as follows (*negative-keyed questions are italicized*): “I find it easy to put myself in somebody else’s shoes.” “I am good at predicting how someone will feel.” “I am quick to spot when someone in a group is

feeling awkward or uncomfortable.” “Other people tell me I am good at understanding how they are feeling and what they are thinking.” “*I find it hard to know what to do in a social situation.*” “*I often find it hard to judge if something is rude or polite.*” “*It is hard for me to see why some things upset people so much.*” “*Other people often say that I am insensitive, though I dont always see why.*”

Annexe J: Supplementary Tables for Partisanship and Altruism

Table J.1: Dictator Game Allocations to Anonymous Individuals – Low Education (OLS)

Variable	Coefficient	(Std. Err.)
Liberal	-3.356	(2.541)
Conservative	-1.186	(2.484)
Income	1.550	(2.370)
Education	5.417	(3.854)
Female	5.084**	(1.867)
Age Group	3.891	(3.722)
Unemployed	-3.858	(4.128)
Empathy	-0.119	(4.286)
Intercept	18.812**	(4.487)
N		802
R ²		0.017
F (8,793)		1.73
Significance levels : † : 10% * : 5% ** : 1%		

Table J.2: Dictator Game Allocations to Other Partisans – Low Education (OLS)

Variable	Coefficient	(Std. Err.)
Liberal	-6.342**	(2.073)
Conservative	-8.168**	(2.027)
Income	-1.330	(1.934)
Education	0.668	(3.145)
Female	3.146*	(1.523)
Age Group	-2.962	(3.037)
Unemployed	-1.486	(3.368)
Empathy	1.491	(3.498)
Intercept	22.997**	(3.661)
<hr/>		
N	802	
R ²	0.036	
F (8,793)	3.656	
<hr/>		
Significance levels : † : 10% * : 5% ** : 1%		

Table J.3: Dictator Game Allocations to Co-Partisans – Low Education (OLS)

Variable	Coefficient	(Std. Err.)
Liberal	-8.476**	(2.514)
Conservative	-7.025**	(2.457)
Income	-2.775	(2.344)
Education	0.492	(3.812)
Female	1.074	(1.847)
Age Group	2.508	(3.682)
Unemployed	-1.398	(4.083)
Empathy	7.069†	(4.240)
Intercept	27.138**	(4.439)
<hr/>		
N	802	
R ²	0.025	
F (8,793)	2.572	
<hr/>		
Significance levels : † : 10% * : 5% ** : 1%		

Table J.4: Dictator Game Allocations to Anonymous Individuals – High Education (OLS)

Variable	Coefficient	(Std. Err.)
Liberal	-1.522	(1.979)
Conservative	-5.027*	(2.120)
Income	-1.308	(2.052)
Female	1.378	(1.599)
Age Group	1.675	(2.747)
Unemployed	-4.695	(5.435)
Empathy	15.511**	(3.509)
Intercept	14.302**	(3.392)
<hr/>		
N	1156	
R ²	0.031	
F (7,1148)	5.209	
<hr/>		
Significance levels : † : 10% * : 5% ** : 1%		

Table J.5: Dictator Game Allocations to Other Partisans – High Education (OLS)

Variable	Coefficient	(Std. Err.)
Liberal	-5.799**	(1.562)
Conservative	-10.509**	(1.673)
Income	-2.271	(1.619)
Female	2.006	(1.262)
Age Group	-3.194	(2.168)
Unemployed	-5.763	(4.289)
Empathy	11.360**	(2.769)
Intercept	17.004**	(2.677)
<hr/>		
N	1156	
R ²	0.073	
F (7,1148)	12.929	
<hr/>		
Significance levels : † : 10% * : 5% ** : 1%		

Table J.6: Dictator Game Allocations to Co-Partisans – High Education (OLS)

Variable	Coefficient	(Std. Err.)
Liberal	-6.266**	(1.981)
Conservative	-7.066**	(2.123)
Income	-1.649	(2.055)
Female	0.773	(1.601)
Age Group	1.754	(2.750)
Unemployed	-4.312	(5.442)
Empathy	17.475**	(3.514)
Intercept	19.663**	(3.397)
<hr/>		
N		1156
R ²		0.04
F (7,1148)		6.787
<hr/>		
Significance levels : † : 10% * : 5% ** : 1%		

Table J.7: Pooled Dictator Game Allocations – Low Education (OLS)

Variable	Coefficient	(Rob. Std. Err.)
Liberal	-6.058**	(2.060)
Conservative	-5.460**	(2.010)
Other Partisan	-7.320**	(0.693)
Co-Partisan	1.495*	(0.723)
Female	3.101*	(1.553)
Age	1.146	(3.169)
Income	-0.852	(2.043)
Education	2.192	(3.100)
Unemployed	-2.247	(3.483)
Empathy	2.813	(3.657)
Intercept	24.924**	(3.733)
<hr/>		
N		2406
R ²		0.043
F (10,801)		26.057
<hr/>		
Significance levels : † : 10% * : 5% ** : 1%		

Table J.8: Pooled Dictator Game Allocations – High Education (OLS)

Variable	Coefficient	(Rob. Std. Err.)
Liberal	-4.529**	(1.685)
Conservative	-7.534**	(1.775)
Other Partisan	-6.982**	(0.515)
Co-Partisan	3.323**	(0.573)
Female	1.385	(1.316)
Age	0.078	(2.193)
Income	-1.743	(1.653)
Education	0.000	(0.000)
Unemployed	-4.923	(3.814)
Empathy	14.782**	(2.850)
Intercept	18.209**	(2.789)
<hr/>		
N	3468	
R ²	0.071	
F (9,1155)	52.241	
<hr/>		
Significance levels : † : 10% * : 5% ** : 1%		

Annexe K: Treatment Assignment for Compulsory Voting Experiment

The randomization of participants proceeded in three steps. First, we identified all subjects (119) who indicated on the initial recruitment form that they did not expect to vote or were unsure. Using a random number generator, we assigned each of these subjects a number and then ranked them according to this number. The top half were assigned to the treatment condition and the bottom half to the control condition. Second, we then assigned a random number to all potential participants who indicated they were likely to vote. We selected the top 86 of these participants. The top half of the selected group was assigned to the treatment condition and the bottom half was assigned to the control condition. Third, to expand our sample using an online survey we invited the remaining 255 eligible participants to take part in the study. We assigned subjects to treatment and control prior to contact using the method of random number assignment and then ranking described above. However, in this instance 70 percent were assigned to treatment and

the remaining 30 percent were assigned to control.

We have checked our randomization procedure across several key variables, and found only one significant difference between conditions in the first round, suggesting that our randomization worked. In each case, we test balance using a χ^2 test of the relationship between treatment and the variable in question. Our treatment was balanced according to gender ($\chi^2 = 0.82$, $p < .37$), with female participants making up 73% of the treatment group and 67% of the control group. Internet usage was also insignificantly related to treatment assignment ($\chi^2 = 5.84$, $p < .44$). Most importantly, there was no difference in the average knowledge scores on the first wave of the survey between the two groups ($\chi^2 = 7.06$, $p < .63$). The same is true political discussion and media news consumption.

We did encounter one possible problem in our randomization. Specifically, considering all those we invited to participate, those who were assigned to the control group chose to participate in larger numbers (66%) than those in the treatment group (54%). This is a significant difference ($\chi^2 = 6.50$, $p < .03$) and raises the possibility of a difference between those who were assigned to the treatment condition and then chose to participate compared to those in the control condition who chose to participate. Because the treatment condition requires more effort than the control condition (i.e. voting), those who chose to participate under the treatment regime may be more motivated in general. This general level of motivation may also make them more likely to seek out political information. If these groups are unbalanced,

any growth found in political knowledge among the treatment group could be attributed to their general levels of motivation (which could differ from the control group) rather than the incentive to learn imposed by mandatory voting. Nevertheless, other factors led us to lay aside this concern. In our first round of invitations, which invited potential participants to a room but did not tell them the details of the experiment, we had 31 participants in the control condition and 22 in the treatment. No participants who showed up declined to fill out the survey. Despite being randomly assigned, we had about 50% more participants in the control conditions show up than those in the treatment. But as this is due to chance, there is no unobserved effect among our first set of participants. When we consider subjects from both rounds of invitations, the possible motivation effect disappears and the difference between the two groups likelihood of participating in the experiment is no longer significant. Taken together, all of these tests suggest that our randomization procedure did not lead to any unobserved differences between the groups which could also be expected to affect knowledge acquisition or political engagement.

Annexe L: Question Wording and Variables for Compulsory Voting Experiment

First Round Knowledge Score: The first round knowledge score is the percentage of the following questions answered correctly. Response categories are given in parenthesis with the correct answer in bold.

- Between the Parti Qubcois and the Liberals, which would you say is further to the right (i.e. more conservative) than the other? (Parti Québécois, **Liberals**)
- In this country, what is the maximum number of years between elections allowed by law? (3, 4, **5**, 6, DK).*
- Which of the following best describes who is entitled to vote in Quebec elections? (Resident of Quebec, Taxpayer in Quebec, Landed Immigrant in Quebec, **Canadian citizen living in Quebec**, DK)
- Which party was in power in Quebec when the Quebec election was called? (Parti Québécois, **Liberals**, Parti conservateur, ADQ, DK).*

- When the election was called, which party had the second largest number of seats in the Assemblée Nationale? (Parti Québécois, **Liberals**, Parti conservateur, ADQ, DK).*
- Which party leader has raised questions about Quebec’s approach of “reasonable accomodation” of minorities? (André Boisclair, **Mario Dumont**, Gilles Duceppe, Stéphane Dion, DK).*
- The date of the Quebec election is the (15 March, **26 March**, 15 April, 26 April, DK).*
- Which party wants to maintain the freeze on university tuition fees? (Parti Québécois, **Liberals**, ADQ, DK).*
- Which party leader advocates paying mothers who stay home with the children? (André Boisclair, **Mario Dumont**, Francoise David, Jean Charest, DK).*
- The Charest government has proposed selling off part of a provincial park. In which region have they proposed this? (Mont Tremblant, **Orford**, St. Maurice, Charlevoix, DK).*
- Which party leader is taking credit for Quebec having made progress on eliminating the fiscal imbalance with Ottawa? (André Boisclair, Mario Dumont, Francoise David, **Jean Charest**, DK).*

Second Round Knowledge Score: The second round knowledge score is the percentage correctly answered of the following questions plus first round

questions marked with an asterisk. Response categories are given in parenthesis with the correct answer in bold.

- The leader of the Quebec Liberal Party is (please write in, **Jean Charest**).
- The leader of the Parti Québécois is (please write in, **André Boisclair**).
- The leader of the ADQ (Action démocratique) is (please write in, **Mario Dumont**).
- Of the three main parties, which is the most federalist? (Parti Québécois, **Liberals**, ADQ, DK).
- During the campaign, an important moment came with decisions announced by Jim Flaherty on March 19th. What is his position? (**Federal Finance Minister**, Quebec Finance Minister, Premier of Ontario, Premier of Alberta.)
- How many party leaders participated in the March 13th debate? (One, Two, **Three**, Four, Five, DK.)
- Which party leader appeared confused at one point about whether Quebec was divisible or indivisible? (André Boisclair, Mario Dumont, Francoise David, **Jean Charest**, DK).
- Which party leader was criticized at one point for using the term “slanted eyes”? (**André Boisclair**, Mario Dumont, Francoise David,

Jean Charest, DK).

- The polls show how many parties have the support of at least one-quarter of the voters? (One, Two, **Three**, Four, DK.)

First and Second Round Political Discussion are calculated as the average response to three questions in the first round, and four in the second round. The response category indicating the least frequency is set to 0 and the most frequency is set to 1. The second round questions were as follows, with those from the first round indicated by an asterisk:

- Some people seem to follow what's going on in government and public affairs most of the time. Others aren't that interested. Do you follow what's going on in government and public affairs most of the time, some of the time, rarely or never?
- Some people seem to follow what's going on in the Quebec election campaign most of the time. Others aren't that interested. Have you been following what's going on in the Quebec election campaign most of the time, some of the time, rarely or never?*
- How often do you talk about current events or things you have heard about in the news with your FAMILY very often, sometimes, rarely or never?*
- How often do you talk about current events or things you have heard about in the news with your FRIENDS very often, sometimes, rarely

or never?*

First and Second Round Media Usage are calculated as the average response of the following four questions. The questions were in both the first and second survey, and are preceded by the following preamble: “Here are some ways that people get news and information. Over the last 7 days, please estimate on how many days you have done each of the following. (please circle the number of days)”.

- Read a newspaper. (0-7).
- Watch the news on TV. (0-7).
- Listen to the news on the radio. (0-7).
- Read news on the internet. (0-7).

Political activities are determined by four questions in the first round survey, all preceded by the preamble “Here is a quick list of things that some people have done to express their views. For each one, please indicate whether you have ever done it or not.”:

- Contacted a newspaper or magazine to express your opinion on an issue.
- Called in to a radio or television talk show to express your opinion on a political issue, even if you did not get on the air.
- Taken part in a protest, march or demonstration.

- Signed an e-mail or a written petition about a social or political issue.

Curriculum Vitae

The author was born in Calgary, Alberta, on February 13, 1979. He was raised in North Bay, Ontario. He attended Mount Allison University in Sackville, New Brunswick, from 1998 to 2002. He graduated with a Bachelor of Arts (First Class Honours) in Political Science and was awarded a Gold A. He commenced graduate studies at l'Université de Montréal in the fall of 2002. He entered the doctoral program in the fall of 2003. He spent the 2004-2005 academic year as a visitising PhD student in the Department of Political Science at the University of Rochester. After a four-month leave to work on a political campaign in Nova Scotia, he began his dissertation in the spring of 2006.