The variable choice set logit model applied to the 2004 Canadian election

Maria Gallego[*], Norman Schofield[†]
Kevin McAlister[‡], Jee Seon Jeon[§]
Center in Political Economy, Washington University,
1 Brookings Drive, Saint Louis, MO 63130.
Tel (1)314 935 5630, Fax (1) 314 935 5688.

August 13, 2013

Abstract

Formal work on the electoral model often suggests that parties should locate at the electoral mean. Recent research has found no evidence of such convergence. In order to explain non-convergence, the stochastic electoral model is extended by including a competence and sociodemographic valance in a country where regional and national parties compete in the election. That is, the model allows voters to face different sets of parties in different regions. We introduce the notion of a convergence coefficient, c for regional and national parties and show that when c is high there is a significant centrifugal tendency acting on parties. An electoral survey of the 2004 election in Canada is used to construct a stochastic electoral model of the election with two regions: Québec and the rest of Canada. The survey allows us to estimate voter positions in the policy space. The variable choice set logit model is used to built a relationship between party position and vote share. We find that in the local Nash equilibrium for the election the two main parties with high competence valence, the Liberals and Conservatives, locate at the national electoral mean and the Bloc Québécois, with the highest competence valence, locates at the Québec electoral mean. The New Democratic Party has a low competence valence but remains at the national mean. The Greens, with lowest competence valence, locate away from the national mean to increase its vote share.

Key words: stochastic vote model, valence, local Nash equilibrium, convergence coefficient.

1 Introduction

Early work in formal political theory focused on the relationship between constituencies and parties in two-party systems. It generally showed that in these cases, parties had strong incentive to converge to the electoral median (Hotelling, 1929; Downs 1957; Riker and Ordeshook 1973). These models assumed a one-dimensional policy space and non-stochastic policy choice, meaning that voters voted with certainly for a party. The models tended to show that there exists a Condorcet point, or pure strategy Nash equilibrium (PNE), at the electoral median. However, when extended

[*]Department of Economics, Wilfrid Laurier University, 75 University Avenue West, Waterloo, Canada N2L 3C5 (mgallego@wlu.ca). Tel: 519.884.0710 Ext 2635
[†]corresponding author: (schofield.norman@gmail.com)
[‡]email: kevinmcalister@wustl.edu
[§]jsjeon@wustl.edu

into multidimensional spaces, later work showed that these two-party pure-strategy Nash equilibria generally do not exist (Schofield, 1978, 1983; Saari 1997; Caplin and Nalebuff 1991).

Instead of a PNE, there often exist mixed strategy Nash equilibria, which lie in the subset of the policy space called the uncovered set (Kramer 1978). This uncovered set often includes the electoral mean, thus giving some credence to a "central" voter theorem in multiple dimensions (Poole and Rosenthal 1984; Adams and Merrill 1999; Merrill and Grofman 1999; Adams 2001). However, this seems at odds with the instability theorems in multidimensional policy spaces.

The contrast between the instability and the stability theorems suggests that we should model an individual's vote as stochastic rather deterministic (Schofield *et al.* 1998; Quinn *et al.* 1999). A stochastic model assumes that the voter has a vector of probabilities corresponding to the choices available in the election. This model is compatible with many theories of voter behavior. Under some conditions this model yields the property that "rational", or vote maximizing, parties will converge to the mean of the electoral distribution. .

Schofield (2007) shows however that convergence to the mean need not occur when valence asymmetries are incorporated. "Valence" is taken to mean any sort of characteristic "perceived" quality that a political candidate exhibits that is independent of the candidates position within the policy space. We consider two types of valence: one associated with the *competence* of the candidate, the other derived from the sociodemographic characteristics of voters.[1] The competence valence is assumed to be common to all members of the electorate and can be interpreted as the average perceived governing ability of a party for all voters in the electorate (Penn, 2003). In this paper, the competence valence is assumed to be common to voters in a given region. The sociodemographic valence depends on the voter's characteristics and thus differs from individual to individual. Due to regional differences, in this paper we assume that the common sociodemographic characteristics of voters in a region partly determine how likely they are to vote for the parties competing in that region. Both kinds of valence can be important in determining electoral outcomes and are necessary to consider when building models of this sort.

Recent empirical work on the stochastic vote model has relied upon the assumption that voter's choices are determined by the voter's "utility" which depends on the valence terms and the voter's policy distances from the various parties in the policy space Voter choice is stochastic and modeled by Type-I extreme value distributed errors (Dow and Endersby 2004). Using this framework, Schofield (2007) introduced the idea of the *convergence coefficient*, c, which can be regarded as measure of the attraction the electoral mean exerts on parties in order to gain votes in an election. Being unitless this coefficient can be used to compare convergence across different models. A low value of the coefficient indicates a strong attraction for parties to locate at the electoral mean. In other words, the joint electoral mean will be a *local pure-strategy Nash equilibrium* (*LNE)*, (Patty 2005; Patty 2006). High values indicate the opposite. For a single region model, Schofield also obtained the necessary and a sufficient conditions that the convergence coefficient must meet for there to be a LNE at the the electoral mean to which all parties converge. When the dimension of the policy space is 2, then the *sufficient* condition for convergence is that $c < 1$. The *necessary* condition for convergence is $c < w$, where w is the number of dimensions of the policy space.

When the necessary condition fails, at least one party will position itself away from the electoral mean in order to increase its vote share. Thus, a LNE does not exist at the electoral mean. Clearly a vector of positions must be an LNE for it to be a pure strategy equilibrium. So that when $c \geq w$, there cannot be a pure strategy vote maximizing Nash equilibrium at the electoral center.

[1]For example, in United States elections, African-American voters are very much more likely to vote for the Democratic candidate than they are to vote for the Republican candidate. Thus, it can be said that the Democratic candidate is of higher average valence among African-American voters than the Republican candidate is.

Using Schofield (2007), we can only examine whether the local Nash equilibrium is given by the electoral mean in the simplest case where there is only one region in the country. The problem quickly becomes more complicated when we face more complex electoral structures. For example, Canada has four national parties and one regional party. While national parties aim to appeal to voters across Canada, the Bloc Québécois (BQ) represents the voice of Québec separatists in the Canadian Parliament and ignores voters outside Québec. In the general regional model that we consider, parties may only run in some regions. This may be due to deep political and economic differences with other regions, in response to too much centralization (Riker 1964, 1987), or because they anticipate doing very poorly in a region, specially if they are financially constrained.[2]

To assess convergence to the electoral mean when there are national and regional parties, one must take into account the electoral centers that parties respond to. Convergence to the electoral mean in Canada would mean that the four national parties converge to the national electoral mean, or the mean of all Canadian voters, while the Bloc would converge to the Québec electoral mean.

In trying to maximize votes, parties respond to the anticipated electoral outcome as well as to the positions of their competitors. When there are regional parties, the set of parties varies by region. National parties must then also take regional differences into account when setting their national policy positions, that is, the position of national parties at the national level should depend on the positions they would like to adopt in each region. For example, the BQ's policy position will affect the position that national parties would want to adopt in Québec and this will in turn influence their positions at the national level. This implies that the *independence of irrelevant alternatives* (IIA) assumption made in *multinomial logit* (MNL) models cannot be used to study elections in countries with regional parties. That is, we cannot assume that voter choices have a multinomial logit (MNL) specification in the formal model as done in Schofield (2007). Moreover, using Schofield's (2007) model, we can only analyze convergence, valence, and spatial adherence within specific regions with the analysis for each region done independently of other regions but we cannot study the election for the entire country. To study electorates like the Canadian one, we need a more general formal model of the election and a more general empirical framework.

One of the main contributions of this paper is the development of a stochastic electoral model for a country with regional and national parties in which the set of parties competing in the national election differs by region. In the formal stochastic model developed in the Technical Appendix voters in different regions face different party bundles. Assuming that parties maximize their vote shares, we derive the first (FOC) and second (SOC) order conditions necessary for a candidate vector of party policy positions to be a LNE. The FOC identifies the first order conditions for a possible equilibrium. From the SOC we derive the convergence coefficient for each regional and national party and then use each party's coefficient to determine the conditions under which the party remains at or diverges from the possible euilibrium position.

We show that each party's policy position at the regional level is given by a weighted average of the positions of voters in this region. The weight that party j gives voter i in region k in its regional position depends on the likelihood that i votes for j in region k relative to the aggregate probability that all other voters in region k vote for j. Interestingly, the possible positions for national parties at the national level are a weighted average of their regional positions. The weight that national party j gives to each region depends on the likelihood that region k votes for j relative to the aggregate probability that all other regions vote for j. That is, national parties take regional differences into account when setting their national positions. We believe that this model is the first to showi how national parties choose their national positions when regional and national parties

[2]We are working on applying the model to the case of Britain, where there are of course at least three regions and regional parties, as well as even greater complexity in Northern Ireland. The first version of the model for Britain (Schofield *et al.* 2011) made it clear that it was necessary to develop a regional model.

compete in a national election in a situation where regional differences matter.

We also show in the technical Appendix that Schofield's (2007) model is a special case of the one developed in this paper. That is, if the country has only one region, the formal model presented in this paper reduces to Schofield's (2007) model. The advantage on this new more general model is that we can study elections in countries where regional difference are crucial in determining the electoral outcome at the national level.

Our second major contribution is the development of an empirical methodology that allows us to study an election in which voters in different regions face different bundles of parties. The *varying choice set logit* model (VCL), based on Yamamoto (2011) presented in Section 3 estimates the parameters necessary to find the equilibria of the model. The VCL, an extension of the mixed logit model, assumes that the error terms in voters' choices have a Type-I extreme value distribution. The VCL allows for say, party ℓ to influence voter's decisions between, say, parties j and h, which the multinomial logit (MNL) model rules out. The basic idea is that we infer a set of possible choices for the parties in each region, we then use the Bayesian framework to guess how the the vote shares of the parties depend on their guesses about the positions of the other parties. Since these guesses involve competition in different regions, the IIA is assumption is violated implying that we cannot study an election in a country with regional and national parties using a MNL framework. Instead we use the VCL model to estimate the parameters first at the regional level; then assuming that the estimated regional parameters come from their own distribution, estimate those at the national level. The way we fix the model is based on Roemer's notion of the core of the party (Roemer, 2011). That is we use the weighted average of the voters who choose each party as a way of estimating how the party can guess the degree of support that it has. Using these guesses we can then model the reasoning of the opportunists who wish to change the part position in order to gain further votes. We give detail of the formal model in a technical appendix.

The third mayor contribution is the application of the new formal model and the new empirical methodology to the study of the 2004 Canadian election. This election is of particular interest because it was the first election since the early eighties were the governing Liberals faced a united right under the newly merged Conservative Party (CP) of Canada.[3] In addition, the issue of Québec separation remained prominent after the failure of two agreements that were to bring "Québec back into the Constitution," which raised the prominence of the Bloc Québécois (BQ) in the election. Moreover, the ongoing infighting within the Liberal Party culminated in Paul Martin replacing Jean Chrétien as prime minister on 12 December 2003. In early 2004, a major scandal on Liberal sponsorship during the 1995 Québec referendum broke forcing Martin to call an early election for June 2004. In spite of running only in Québec and facing only a quarter of the Canadian electorate, the Bloc gained the support of almost half of Québécers giving it 54 out of the 75 seats Québec has in the Canadian Parliament. The prominence of this regional party in the 2004 Canadian election prompted us to develop a formal model in which national and regional parties compete in the election as well as using the variable choice set logit model to show that in order to maximize votes the Bloc positioned itself at the Québec rather than the Canadian electoral mean.

In Section 4, we use the formal model and the VCL methodology to study the 2004 Canadian election in two regions: Québec and the rest of Canada. We find that in Québec, the BQ has the highest competence valence followed by the Liberals once policy differences and the sociodemographic valences are taken into account. Outside Québec, the Liberal and Conservatives were considered by voters to be equally competent at governing. The New Democratic Party (NDP, a

[3] Unable to make a break through in Eastern Canada, the western based Reform Party rebranded itself as the Canadian Reform Alliance Party. Alliance was also unable to appeal to Eastern Canadians. After long deliberations Alliance and the Progressive Conservatives merged in December 2003 to form the Conservative Party of Canada. These types problems in federal systems are not unusual in first-past-the-post plurality systems (Riker 1982).

left leaning party) and the Greens had the lowest competence valences in both Québec and the rest of Canada. Assuming that parties use the VCL model as a heuristic model of the anticipated election, we then examine whether parties would locate at their corresponding electoral means. The analysis of the Hessian of second derivatives of the party's vote share functions together with the convergence coefficient for each party shows that if the two major national parties, the Liberals and Conservatives, locate at the national electoral mean they would be maximizing their vote shares as would the BQ if it locates at the Québec mean. Even though the NDP has a low competence valence, this does not entice it to move from the national mean to increase votes. The Greens with the lowest competence valence diverge from the national mean to increase its vote share.

Using Schofield's (2007) model we have studied the party's positioning strategies in countries that operate under different political systems assuming a single region to examine (using MNL estimates of the election) whether there is convergence to the electoral mean in different countries. The necessary condition for convergence in Schofield (2007) is that the convergence coefficient be less than the dimension of the policy space. The convergence coefficient is dimensionless and thus can be used to compare convergence across election, countries and political systems. We compare convergence across political systems in Gallego and Schofield (2013).[4] The convergence coefficients for the 2005 and 2010 UK elections were not significantly different from 1, meeting the necessary condition for convergence to the mean. For the 2000, 2004 and 2008 US presidential elections, the convergence coefficient is significantly below 1 in 2000 and 2004 thus meeting the sufficient and necessary conditions for convergence; and not significantly different from 1 in 2008, only meeting the necessary condition for convergence. This suggests that the centrifugal tendency in the majoritarian polities like the United States and the United Kingdom is very low. In contrast, the convergence coefficient gives an indication that the centrifugal tendency in Israel, Poland and Turkey is very high. In these proportional representation systems with highly fragmented polities the convergence coefficients are significantly greater than 2 (the dimension of the policy space) failing to meet the necessary condition for convergence to the mean. In the anocracies[5] of Georgia, Russia and Azerbaijan, the convergence coefficient is not significantly different from the dimension of the policy space (2 for Georgia and Russia and 1 for Azerbaijan), thus failing the necessary condition for convergence. While the analysis for Georgia and Azerbaijan shows that not all parties converge to the mean, in Russia it is likely that they did. Thus, in Russia opposition parties found it difficult to diverge from the position adopted by Putin's party, the electoral mean.

Using the new formal model and the new empirical methodology that allow us to take regional differences into account, we find, in this paper, that the Liberals, Conservatives and NDP converged to the Canadian mean and the BQ to the Québec mean. The Greens, a small national party, locates away from the national mean to increase its vote share. We show that popular Bloc Québécois, the party with the highest competence valence, affected the Canadian election. We decompose the analysis between Québec and the rest of Canada, and show that given that Québec has a quarter of the Canadian population and controls a quarter of the parliamentary seats in the House of Commons, the Bloc not only affected the positions of the national parties in Québec, but also their positions in the rest of Canada and thus also affecting the electoral outcome in the rest of Canada.

[4] The work described in this paragraph can also be seen in Schofield *et al.* (2011d)for the UK; in Schofield *et al.* (2011c) for the US; in Schofield *et al.* (2011b) for Israel; in Schofield *et al.* (2011e) for Poland; in Schofield *et al.* (2011a) Turkey; in Schofield *et al.* (2012) for Azerbaijan and Georgia; and in Schofield and Zakharov (2010) for Russia.

[5] Anocracies are countries in which the president/autocrat governs along an elected legislature. The President, however, exerts undue influence on legislative elections.

2 Modelling Elections in Multi-Regional Countries

We model elections in countries where there are vast political and/or economic differences across regions.[6] Political differences may originate from cultural differences across regions or from the desire of a particular region for more independence from the national government (Riker, 1964, 1987). Economic differences arise from endowments of natural resources or from previous regional economic development. When substantial political and/or economic regional differences exist, interest groups and voters in these regions coalesce to create parties that better represent their interests at the national level.[7] National parties cater to nationwide interest and seek to represent voters across all regions of the country. Regional parties, on the other hand, are concerned only with representing the interests of voters in their jurisdiction.

We assume that regional parties operate only in a single jurisdiction, a province or state.[8] Moreover, there may be regions with no regional parties as the political and economic actors as well as voters in these regions feel that their interests are well represented by national parties.

We develop an electoral model for a country with at least one national and one regional party. The preferences of parties and voters are defined over the same space at both regional and national levels. That is, the policy space is defined broadly enough to include all relevant policy dimensions in the country.[9] We allow voters' and parties' preferences vary across regions and study the policy positioning game of regional and national parties in response to the anticipated electoral outcome.

$$u_{ijk}(x_i, z_j) = \lambda_{jk} + \alpha_{jk} - \beta_k \|x_i - z_j\|^2 + \epsilon_{ij} = u^*_{ijk}(x_i, z_j) + \epsilon_{ijk} \quad (1)$$

Here, $u^*_{ijk}(x_i, z_j)$ is the observable component of voter i's utility associated with party j in region k. The term λ_{jk} is the *competence valence* for agent j in region k. This valence is common across all voters in region k and gives an estimate of the perceived "quality" of party j or of j's ability to govern. We model voters' common belief on j's quality by assuming that an individual voter's perception is distributed around the mean perception in region k, i.e., $\lambda_{ijk} = \lambda_{jk} + \xi_{ijk}$ where ξ_{ijk} is a random iid shock specific to region k. This regional valence is independent of party positions. Moreover, since regional party j in region k never runs in other regions of the country, the model says nothing about the belief that voters in other regions have on j's ability to govern. This is not a problem as voters outside of region k cannot vote for regional party j in region k.

The sociodemographic aspects of voting for voters in region k are modelled by θ_k, a set of s -vectors $\{\theta_{jk} : j \in P_{Nat} \cup P_k\}$ representing the effect of the s different sociodemographic parameters (gender, age, class, education, financial situation, etc.) have on voting for party j in region k while η_i is an s-vector denoting voter i^{th} individual's sociodemographic characteristics. The composition $\alpha_{ijk} = \{(\theta_{jk} \cdot \eta_i)\}$ is a scalar product representing voter i's *sociodemographic valence* for party j in region k. We assume that voters with common sociodemographic characteristics share a common evaluation or bias for party j that is captured by their sociodemographic characteristics. We model this by assuming that an individual voter's sociodemographic valence varies around the mean sociodemographic valence in region k, $\alpha_{ijk} = \alpha_{jk} + \nu_{ijk}$ where ν_{ijk} is a random iid shock specific

[6] For example, in Canada, Québec is by the nature of its history, culture and laws different from other provinces; Alberta has vast natural resources (the oil sands); and Ontario has large manufacturing, high tech and service sectors.

[7] The Bloc Québécois was created after a failed attempt to bring Québec back into the Canadian Constitution.

[8] There may exist parties that may have no national scope but that represent the interest of groups and voters across various provinces or states. Parties with support across various regions may strive to become national players as they grow. Since we examine only one election in the model, we rule out the existence of multi-regional parties as well as the the possibility that regional parties can grow to become national parties in the model.

[9] In Canada, Albertans care about the oil sands; some Québécers about preserving their French culture and their laws; and Ontarians about policies that affect the manufacturing, high tech and service sectors.

to region k. Thus, the sociodemographic valence α_{jk} is the "average" sociodemographic valence of voters in region k for party j. These regional sociodemographic valences are independent of party positions. The competence valence λ_{jk} measures an average assessment of party j's ability to govern by voters in region k and since we control for voters' sociodemographic biases, λ_{jk} measures j's ability to govern *net* of any sociodemographic bias these votes may have.

The term $\|x_i - z_j\|$ is the Euclidean distance between voter i's ideal policies x_i and party j's position z_j. The coefficient β_k is the *weight* given to policy differences with party j by all voters in region k. This weight varies by region to allow preferences to differ across regions. Differences that in some regions were deep enough in the past to have lead to the emergence of regional parties.

The error term ϵ_{ijk}, commonly distributed among all voters in region k, come from a Type-I extreme value distribution. Assumption also made in empirical models below which makes the transition to applying this theoretical model to the 2004 Canadian election easier.

To find parties' policy positions in a model where varying sets of parties compete in different regions, the analysis must be first carried out at the regional level before moving to the national level. We begin by examining the parties' positioning game in region $k \in \Re$.

Given the stochastic assumption of the model and the parties' policy positions in region k, $\mathbf{z}_k$, the probability that voter i votes for party j in region k is

$$\begin{aligned}\rho_{ijk}(\mathbf{z}_k) &= \Pr[u_{ijk}(x_i, z_j) > u_{ihk}(x_i, z_h), \text{ for all } h \neq j \in P_{Nat} \cup P_k,] \\ &= \Pr[\epsilon_{hk} - \epsilon_{jk} < u^*_{ijk}(x_i, z_j) - u^*_{ihk}(x_i, z_h), \text{ for all } h \neq j \in P_{Nat} \cup P_k]\end{aligned}$$

where the last line follows after substituting in (18) and Pr stands for the probability operator generated by the distribution assumption on ϵ. Thus, the probability that i votes for j in region k is given by the probability that $u_{ijk}(x_i, z_j) > u_{ihk}(x_i, z_h)$, for all j and h in $P_{Nat} \cup P_k$, i.e., that i gets a higher utility from j than from any other party competing in region k.

With the errors coming from a Type-I extreme value distribution and given the vector of party policy positions $\mathbf{z}_k$, the probability that i votes for j in region k has a logit specification, i.e.,

$$\rho_{ijk} \equiv \rho_{ijk}(\mathbf{z}_k) = \frac{\exp[u^*_{ijk}(x_i, z_j)]}{\sum_{h=1}^{p+q_k} \exp[u^*_{ihk}(x_i, z_h)]} = \frac{1}{\sum_{k=1}^{p+q_k} \exp[u^*_{ihk}(x_i, z_h) - u^*_{ijk}(x_i, z_j)]} \tag{2}$$

for all $j \in P_{Nat} \cup P_k$ where to simply notation we take the dependence of ρ_{ijk} on $\mathbf{z}_k$ as understood.

This *stochastic multi-regional* (SMR) model does not rely on the *independence of irrelevant alternatives* (IIA) assumption made in Multinomial Logit (MNL) models since we allow the presence of, say, party ℓ to affect voter choices between, say, parties j and h. This is particularly important in our model since voters in different regions face different bundles of parties in the election. Note, that when there is only one region, our SMR model reduces to that developed in Schofield (2007).

Since voters' decisions are stochastic in our framework, parties cannot perfectly anticipate how voters will vote but can estimate their *expected* vote shares. With varying sets of parties competing in different regions, agents can estimate their expected regional vote share in each region and given these regional vote shares, national parties can estimate their expected national vote shares.

For party $j \in P_{Nat} \cup P_k$ competing in region k, its *expected vote share in region* k is the average of the probabilities over voters in region k, i.e.,

$$V_{jk}(\mathbf{z}_k) = \frac{1}{n_k} \sum\nolimits_{i \in N_k} \rho_{ijk} \quad \text{for } j \in P_{Nat} \cup P_k, \tag{3}$$

with the sum of vote shares in each region adding up to 1, $\sum_{j \in P_{Nat} \cup P_k} V_{jk}(\mathbf{z}_k) = 1$ for all $k \in \Re$.

National parties must, in addition, take into account that their *expected* vote share depends on all voters in the country. However, due to the presence of regional parties and since the number of voters varies across regions, the expected national vote share of party j *cannot* be estimated as the average of the probabilities of voters across the country. Rather, j's expected national vote share depends on the vote share j expects to obtain in each region in the country. We assume that the *expected national vote share of party* j is the weighted average of its expected vote share in each region, where the weight of region k is given by the proportion[10] of voters in region k, $\frac{n_k}{n}$, i.e.,

$$V_j(\mathbf{z}_{Nat}) = \sum_{k\in\Re} \frac{n_k}{n} V_{jk}(\mathbf{z}_k) = \sum_{k\in\Re} \frac{n_k}{n}\frac{1}{n_k} \sum_{i\in N_k} \rho_{ijk} = \frac{1}{n} \sum_{k\in\Re} \sum_{i\in N_k} \rho_{ijk}. \tag{4}$$

The third term in (21) follows after substituting in (20). Note that due to the presence of regional parties, the sum of the vote share of national parties do not add to 1.

The objective is to find the *local Nash equilibria* (LNE) of party positions where each party takes the position of all the other national and regional parties as well as that of voters as given.

2.0.1 Is there convergence at the regional or national levels?

The technical Appendix gives the vector of possible vote maximizing positions for regional and national parties. We now need to determine whether the parties are maximizing their vote shares at these critical points.. To find whether the possible choice for position z_{jk}^C (correspondingly, z_j^C) is a local maximum of regional (correspondingly national) party j's vote share function, so that z_{Nat}^C is a LNE of the game, we need to examine whether the second order condition determines whether the regional (correspondingly national) vote share function of party j is at a maximum, minimum or a saddle point. To do so we need to check whether the Hessian of the second derivatives of regional (correspondingly national) party j's vote share function is negative definite. Since regional and national parties face different electorates the analysis must consider whether the party is a regional or national party and the region in which the parties compete. These second order conditions for regional and national parties are derived in the Technical Appendix and give the conditions under which regional (correspondingly national) party j is maximizing its vote share when located at its critical point z_{jk}^C (correspondingly, z_j^C).

The TechnicalAppendix shows that the *necessary* condition for party j in region k to converge to or remain at z_{jk}^C in order to maximize its vote share is that

$$c_{jk}(\mathbf{z}_{Nat}^C) < \sum_{\omega=1}^{w} 1 = w \tag{5}$$

where $c_{jk}(\mathbf{z}_k^C)$ is party j's *convergence coefficient in region* k given by

$$c_{jk}(\mathbf{z}_k^C) \equiv \sum_{\omega=1}^{w} \sum_{i\in N_k} \mu_{ijk} 2(1-2\rho_{ijk})\beta_k[x_i(\omega) - z_{jk}^C(\omega)]^2. \tag{6}$$

Note that $c_{jk}(\mathbf{z}_k^C)$ depends on the weight j gives each voter in region k, μ_{ijk} in (24) and on the probability that each voter in region k votes for j, ρ_{ijk} in (19). It also depends on how dispersed voters in region k are around j's possible choice for a position in the ω dimension, $\beta_k[x_i(\omega) - z_{jk}^C(\omega)]^2$, which takes into account the weight that voters give to differences with j's policies β_k. By aggregating over all policy dimensions, j's convergence coefficient in region k takes into account the dispersion over voters' positions across all the w dimensions of the policy space.

[10] We could have assumed instead that the weight of each region depends on the share of seats each region gets in the national parliament. The results presented below would then depend on seat rather than vote shares but would remain substantially unchanged. Note that the number of parliamentary seats that each region gets is, in general, based on the proportion of the population living in that region.

We can also write j's convergence convergence condition in region k in (5) as the sum of the convergence coefficients along each dimension, i.e.,

$$c_{jk}(\mathbf{z}_k^C) \equiv \sum\nolimits_{\omega=1}^{w} c_{jk}(\mathbf{z}_{jk}^C(\omega)) < \sum\nolimits_{\omega=1}^{w} 1 = w \tag{7}$$

where $c_{jk}(\mathbf{z}_k^C(\omega))$, the convergence coefficient of party j in region k along dimension ω is given by

$$c_{jk}(\mathbf{z}_k^C(\omega)) \equiv \sum\nolimits_{i\in N_k} \mu_{ijk} 2(1-2\rho_{ijk})\beta_k[x_i(\omega) - z_{jk}^C(\omega)]^2 \tag{8}$$

So that, if convergence is met dimension by dimension, i.e., if $c_{jk}(\mathbf{z}_j^C(\omega)) < 1$ for all ω, then party j in region k will be maximizing its vote share at $\mathbf{z}_{jk}^C$. If convergence is not met in at least one dimension, that is, if $c_{jk}(\mathbf{z}_{jk}^C(\omega)) > 1$ for some ω, then party j in region k will have an incentive to move from $\mathbf{z}_{jk}^C$ in at least dimension ω.

Result 1: Convergence at the regional level

For any party j competing in region k,

- j will remain at its critical point, z_{jk}^C, only if j's convergence coefficient in region k is less than the dimension of the policy space, w. That is, when $c_{jk}(\mathbf{z}_{Nat}^C) < w$, j is maximizing its vote share at z_{jk}^C and has no incentive to move as doing so would decrease its vote share.
- When $c_{jk}(\mathbf{z}_{Nat}^C) \geq w$, then at z_{jk}^C, j is at a minimum or at a saddle point of its vote share function and moves away from z_{jk}^C to increase its votes. In the two dimensional case, $w = 2$, if $c_{jk}(\mathbf{z}_k^C(\omega))$ in (8) is less than one in one dimension and greater than one in the other, party j in region k is at a saddle point and will not locate at z_{jk}^C.

This result depends on all other regional and national parties locating at their corresponding critical points. If to increase votes j moves away from z_{jk}^C to another position, other regional or national parties may then also find it in their interest to move from their critical points.

The *necessary* condition for *national* party j to remain at z_j^C to maximize its vote share given in Appendix A is that

$$c_j^C(\mathbf{z}_{Nat}) < \sum\nolimits_{\omega=1}^{w} 1 = w \tag{9}$$

where $c_j^C(\mathbf{z}_{Nat})$ is national party j's *national convergence coefficient* given by

$$c_j(\mathbf{z}_{Nat}^C) \equiv \sum\nolimits_{k\in\Re} \theta_{jk} \sum\nolimits_{\omega=1}^{w} \sum\nolimits_{i\in N_k} \mu_{ijk} 2(1-2\rho_{ijk})\beta_k \left[x_i(\omega) - z_j^C(\omega)\right]^2. \tag{10}$$

Note that $c_j(\mathbf{z}_{Nat}^C)$ depends on the weight given by j to each region, θ_{jk} in (**??**); on the weight that j gives to each voter in region k, μ_{ijk} in (24); on the probability that each voter in region k votes for j ρ_{ijk} in (19); and on how dispersed voters are around j's possible choice for a position in the ω dimension in region k, $\beta_k \left[x_i(\omega) - z_j^C(\omega)\right]^2$. By aggregating over all dimensions and all regions, $c_j(\mathbf{z}_{Nat}^C)$ also accounts for the dispersion over voters across all the w dimensions of the policy space in all regions.

National party j's convergence coefficient can also be expressed as a function of the convergence coefficients it faces in each region since using (6), we can re-write (10) as

$$c_j(\mathbf{z}_{Nat}^C) = \sum\nolimits_{k\in\Re} \theta_{jk} c_{jk}(\mathbf{z}_{jk}^C). \tag{11}$$

Thus, $c_j(\mathbf{z}^C_{Nat})$ is a weighted average of the convergence coefficient national party j faces in each region where the weight θ_{jk} is the weight party j gives each region in its policy function given in (**??**). When national party j's convergence coefficient in each region satisfies the convergence condition, i.e., if $c_{jk}(\mathbf{z}^C_{jk}) < w$ for all $k \in \Re$, then j's national position also satisfies the convergence condition at the national level.

National party j's convergence condition in (9) can also be re-written as the sum of the convergence coefficients along each dimension, i.e.,

$$c_j(\mathbf{z}^C_{Nat}) \equiv \sum\nolimits_{\omega=1}^{w} c_j(\mathbf{z}^C_j(\omega)) < \sum\nolimits_{\omega=1}^{w} 1 = w \tag{12}$$

where $c_j(\mathbf{z}^C_j(\omega))$, the convergence coefficient of national party j along dimension ω is given by

$$c_j(\mathbf{z}^C_j(\omega)) \equiv \sum\nolimits_{k\in\Re} \theta_{jk} \sum\nolimits_{i\in N_k} \mu_{ijk} 2(1-2\rho_{ijk})\beta_k \left[x_i(\omega) - z^C_j(\omega)\right]^2 \tag{13}$$

When convergence is met dimension by dimension, i.e., $c_j(\mathbf{z}^C_j(\omega)) < 1$ for all ω, then national party j is maximizing its vote share at $\mathbf{z}^C_j$. Convergence is not met when $c_j(\mathbf{z}^C_j(\omega)) > 1$ for some ω, which gives national party j an incentive to locate away from $\mathbf{z}^C_j$ in at least in dimension ω.

Result 2: Convergence at the national level

For any national party j,

- j will converge to its critical point, z^C_j, only if the value of j's national convergence coefficient in (10) is less than the dimension of the policy space, w. So that if $c_j(\mathbf{z}^C_{Nat}) < w$, then j is maximizing its vote share at z^C_{jk} and has no incentive to move.
- When $c_j(\mathbf{z}^C_{Nat}) \geq w$, then j is either at a minimum or at a saddle point of its vote share function and moves away from z^C_j to increase its votes. In the two dimensional case, $w = 2$, if $c_j(\mathbf{z}^C_j(\omega))$ in (13) is less than one in one dimension and greater than one in the other, national party j is at a saddle point and will want to move from z^C_{jk}.

This result depends on all other regional and national parties locating at their corresponding critical points. If national party j moves away from z^C_j to increase its votes, other regional and national parties may also find it in their interest to move away from their critical point.

Clearly, if the party with the highest convergence coefficient does not want to move from the possible choice for its location (because it is maximizing its votes) then no other party will want to move from its possible position either. Thus, the party with the highest convergence coefficient determines whether parties converge to their possible positions as whole.

Define the convergence coefficient of the election as the highest convergence coefficients of all national and regional parties as

$$c(\mathbf{z}^C_{Nat}) = \max\left\{\max c^C_j(\mathbf{z}^C_{Nat}), j \in P_{Nat}; \max c_{jk}(\mathbf{z}^C_k), j \in P_1; ...; \max c_{jk}(\mathbf{z}^C_k), j \in P_r\right\} \tag{14}$$

Result 3: Electoral convergence

The vector of possible positions $\mathbf{z}^C_{Nat}$ is a LNE of the election when

- all parties, regional and national, want to remain at their position. This happens only when the convergence coefficient of the election $c(\mathbf{z}^C_{Nat})$ in (14) is less that the dimension of the policy space, w. That is, if $c(\mathbf{z}^C_{Nat}) < w$ then the convergence coefficient of any regional or national party will also be less than w. In this case, the position announced by parties prior to the election coincide with $\mathbf{z}^C_{Nat}$ so that no party wants move from its possible position.

- the dimensional components of $c_{jk}(\mathbf{z}_k^C(\omega))$ and $c_j(\mathbf{z}_j^C(\omega))$ in ω dimension given in (8) and (13) are all less than 1.
- If $c(\mathbf{z}_{Nat}^C) \geq w$, at least one party will want to diverge from its position and $\mathbf{z}_{Nat}^C$ is not a LNE of the election.

2.1 Summarizing our results

In a model in which regional and national compete in the national election, our main result states that parties will locate or converge to the critical position satisfying the first order condition if the convergence coefficient for both national and regional parties is less than the dimension of the policy space, w as all parties will be maximizing their vote shares. Moreover, the convergence coefficient of national party j must be the weighted average of j's convergence coefficient in each region where the weight θ_{jk} in (**??**) is the weight j gives to each region in equilibrium. If the convergence coefficient of all regional and national parties are less than w, then there is electoral convergence and the vector of possible positions $\mathbf{z}_{Nat}^C$ is an LNE of the election.

If, on the other hand, the convergence coefficient of at least one regional or national party, say j, is greater than w then j will have an incentive to deviate from its critical position in order to increase its votes. Other parties may then also find it in their interests to move from their possible positions. In this case, the vector of possible positions $\mathbf{z}_{Nat}^C$ will *not* be a LNE of the election.

The above theoretical analysis generalizes the earlier work of Schofield (2007). In that work it was implicitly assumed that there was only a single region. Schofield's (2007) model considered conditions under which the electoral mean vector (normalized to be at the origin) $\mathbf{z}_{Nat} = \mathbf{0} \equiv (0, .., 0)$ could be a LNE. Note that if we assume that there is only one region in the present model, then the weight given to policy differences in the voter's utility in (18) is the same for all voters in the country, i.e., there is a single β as assumed in Schofield (2007). From (26) we can see that in this case the electoral variance around the electoral mean vector reduces to $\frac{1}{n}\sum_{i\in N}[x_i(\omega) - \mathbf{0})]^2$. Writing this as σ^2 and imposing the assumptions on (10) gives that the necessary condition for party 1 to converge to the electoral origin as $c_1(\mathbf{z}_{Nat}) = 2\beta(1-2\rho_1)\sigma^2 < w$, where ρ_1 is the probability that a generic voter chooses the lowest valence party 1, when all parties locate at the origin. Since the incentive to converge to or diverge from the origin is greatest for party 1, the result presented in Schofield (2007) is that convergence in the election is determined by the incentives of party 1.

The theoretical model presented in this paper gives a method to assess whether a vector of party positions is a LNE in a model with national and regional parties summarized as follows:

1. Define the vector of possible party positions in the policy space $\mathbf{z}_{Nat}^C \equiv \bigcup_{k=1}^{r} \mathbf{z}_k^C$.
2. Check that each party's possible position meets the FOC given in (23) for regional parties and in (**??**) for national parties while holding the position of other parties constant.
   - Note that if party j weights each voter in region k equally, so that from (24) $\mu_{ijk} = \dfrac{1}{n_k}$ then j locates at
     $$z_{jk} = \frac{1}{n_k}\sum\nolimits_{N_k} x_i.$$
     In this case, j locates at the mean of the ideal points of voters in region k, i.e., locates at *region k's electoral mean.* Under this assumption, the regional electoral mean is always a critical point of the vote share function of the parties competing in region k.

- Note also that if national party j weights each region according to their population share so that from (**??**) $\theta_{jk} = \frac{n_k}{n}$, and also weights all voters in region k equally, so that from (24) $\mu_{ijk} = \frac{1}{n_k}$ then j locates at

$$z_j = \sum_{k\in\Re} \theta_{jk} \sum_{i\in N_k} \mu_{ijk} x_i = \sum_{k\in\Re} \frac{n_k}{n} \sum_{i\in N_k} \frac{1}{n_k} x_i = \frac{1}{n} \sum_{k\in\Re} \sum_{i\in N_k} x_i.$$

  In this case, j locates at the mean of the ideal points of all voters in the country, i.e., locates at the *national electoral mean.* Under these assumptions, the national electoral mean is always a critical point in the vote function of national party j.

- Note also that if there is only one region, so that all parties are national parties, then the national electoral mean is always a critical point in vote share function of all parties. This is the reason that Schofield (2007) examined whether the national electoral mean, normalized to be at the origin, $z_{Nat} = (0, .., 0)$, was a LNE of the election.

3. To find if at the possible position, each party is maximizing its vote share we need to look at the SOC on the regional and national party's vote shares given in Appendix A. We look at the Hessian of second order derivatives of each party's regional vote share function, H_{jk} in (27) and of each national party's vote share function, H_j in (33), to examine whether at the corresponding critical point the party's vote share is at a maximum, a minimum or a saddle point. From Appendix A, we also know that regional (respectively national) party j is at a maximum if H_{jk} (respectively H_j) is negative definite. Since the trace of any Hessian is equal to the sum of the eigenvalues associated the Hessian and is also given by the sum of the main diagonal elements of the Hessian, we know that for z^C_{jk} (respectively z^C_j) to be a local maximum of regional (respectively national) party j's vote share function, the eigenvalues of H_{jk} (respectively H_j) have to be all negative thus implying that the trace of H_{jk} (respectively H_j) must then also be nbleegative. If trace of the Hessian of all regional parties in (29) and of all national parties in (35) is negative, then each party is maximizing its vote at the possible position. The vector of candidate positions is then a LNE of the election.

4. Recall from Section 2.0.1 that whether a party converges to the possible position depends on the value of its convergence coefficient which are derived from the SOCs. To determine convergence, calculate the convergence coefficient for each party at the regional level using (6) and for each national party using (10). Using the convergence coefficients of each party, calculate the convergence coefficient of the election $c(\mathbf{z})$ using (14). If $c(\mathbf{z}) \leq w$, check the convergence condition in each dimension, i.e., check $c_{jk}(\mathbf{z}^C_k(\omega))$ and $c_j(\mathbf{z}^C_j(\omega))$ in ω dimension given in (8) for party j in region k and in (13) for national party j. If all are less than 1, then the system converges to $\mathbf{z}^C_{Nat}$, the LNE of the election. If $c(\mathbf{z}) > w$, at least one party will not converge to the candidate position and $\mathbf{z}^C_{Nat}$ will not be a LNE of the election.

We now describe the empirical methodology that we use to apply the formal stochastic multi-regional model to the 2004 Canadian election.

3 Estimation Strategies Given Varying Party Bundles

We are interested in applying the formal model developed in Section **??** to study convergence to the national and Québec electoral means in the 2004 Canadian election. To study the election, we need to estimate the probability that each voter votes for each party in each region, ρ_{ijk} using (19) which means estimating the observable component of the voter's utility, $u^*_{ijk}(x_i, z_j)$ in (18). To estimate $u^*_{ijk}(x_i, z_j)$ we need estimates of the competence and sociodemographic valences for all parties, λ_{jk} and α_{jk} for $j \in P_{Nat} \cup P_k$ and $k \in \Re$ and of the weights given by voters to the policy differences with parties in each region, β_k for $k \in \Re$. We also need the exogenously given positions of voters as well as the party's positions. Note that in equilibrium each party's position depends on the weights parties give voters and regions in their policy positions, meaning that we also need estimates of μ_{ijk} in (24) and of θ_{ij} in (**??**). Recall that these weights are endogenously determined since they depend on the probability that all voters in each region vote for the party and these in turn depend on the party positions. Using the estimates of λ_{jk}, α_{jk} and β_k we can estimate $u^*_{ijk}(x_i, z_j), \rho_{ijk}$, μ_{ijk} and θ_{ij}. We can then calculate the convergence coefficient for regional parties using (6) and for national parties using (10). The convergence conditions derived in Results 1, 2 and 3 then determine whether the parties converge or not to the possible position.

Note that to estimate λ_{jk} and α_{jk} and β_k we cannot use multinomial logit (MNL) model since it relies on the *independence of irrelevant alternatives* (IIA) assumption. IIA requires that all odds ratios between the probabilities that each voter in each region votes for a pair of parties j and h, $\frac{\rho_{ijk}}{\rho_{ihk}}$ be independent of party ℓ where ρ_{ijk} and ρ_{ihk} are given by (19) and that this odds ratio be preserved from region to region. Since set of paries varies across regions, IIA is violated in the formal and empirical model. MNL models can then not be used in the estimation procedures.

Yamamoto (2011) proposes a model that overcomes these problems: the *varying choice set logit* (VCL) model, a variant on the typical hierarchical multinomial logistic regression model. We adapt Yamamoto's VCL model to our regional setting as unlike the MNL models it does not rely on the IIA assumption. The VCL estimates individual logistic regression models for each region then assuming that the regional parameters come from their own distribution, aggregates these parameters to estimate the valences at the national level.

The VCL model allows for voter's utility to be region specific, i.e., in the *empirical* estimation we assume that the utility i derives from voting for party j in region k, given in (18) in Section **??**, is

$$u_{ijk}(x_i, z_j) = \lambda_{jk} + \alpha_{jk} - \beta_k \|x_i - z_j\|^2 + \epsilon_{ij} = u^*_{ijk}(x_i, z_j) + \epsilon_{ijk} \tag{15}$$

where λ_{jk} is the average competence valence of party j in region k; α_{jk} is the average sociodemographic valence and represents the utility that an average voter, with given sociodemographic characteristics, gets from voting for party j in region k. The weight given by voters to policy differences with parties in region k is measured by β_k. This hierarchical specification of the valence terms lends itself very well to the VCL model.

The error terms ϵ_{ijk} come from a Type-I extreme value distribution, as assumed in the Stochastic Multi-Regional (SMR) model developed in Section **??** used to derive the convergence coefficient. The *empirical* probability that voter i votes for party j in region k has then a logit specification

$$\rho_{ijk} \equiv \rho_{ijk}(\mathbf{z}_k) = \frac{\exp[u^*_{ijk}(x_i, z_j)]}{\sum_{h=1}^{p+q_k} \exp[u^*_{ihk}(x_i, z_h)]} = \frac{1}{\sum_{k=1}^{p+q_k} \exp[u^*_{ihk}(x_i, z_h) - u^*_{ijk}(x_i, z_j)]}. \tag{16}$$

Note that ρ_{ijk} is the same as in (19). In Section 4.3 we will use this probability to estimate the parameters of voter i's utility in (15) in two regions, Québec and the rest of Canada for the 2004 Canadian election. Under these assumptions, the framework of the formal and the empirical models match, making the transition to the estimation of the parameters of the formal model easy. We can then analyze the equilibria of the system using these parameter estimates and framework of the SMR model of Section **??**. Because VCL model does not rely on the IIA assumption, it is the proper model to use when estimating the parameters for an electorate with a regional structure.

The VCL model uses random effects for each region. This means that for each region we estimate the parameters of interest for voters in that region.[11] Then, using these estimates, we assume that these individual estimates come from their own distribution, and use their distribution to determine the best national estimate for a parameter within the model.

Using the VCL, however, places a few light assumptions on the model, as any estimation procedure does. First, as already specified in the utility function in (15), we allow for voters' policy preferences to differ by region, different β_k. Second, by using random effects, this model assumes that each of the regional and sociodemographic group random effects are orthogonal to each other and to other covariates in the model; in particular, are independent of voter's position in the policy space as assumed in voter's utility in (15) and as assumed in (18) in the formal model. Third, by using the VCL model we assume that a party's decision to run in a specific region is exogenous of its perceived success in that region. This assumption is inconsequential when studying a single election but would be problematic if we were studying a sequence of elections in a country with an unstable party system that changes from election to election as is the case in recent Polish elections. However, many electoral systems with regional parties have parties which are historically bound to one region or another and this is independent of their success.[12] Thus, this model is appropriate when there are regional parties representing specific regions. When these three assumptions are met by the electorate of interest the VCL is the proper estimation procedure.

The reason that the varying choice set logit (VCL) is the superior method when handling electorates with multiple regions is that it relaxes the IIA assumption while also providing us with the most information from the model. VCL relaxes IIA by allowing each of the parameters to be estimated within each group (i.e., region) and by allowing these parameters to derive the aggregate (i.e., national) estimation of parameters through the notion of partial pooling. Partial pooling is best achieved through hierarchical modeling and through the use of random effects. VCL can be viewed as a specific kind of mixed logit model, meaning that the mixed logit model can be used to achieve the same aggregate results. However, given the structure of VCL, parameter estimates can be achieved for each choice set type (i.e., region) rather than for each voter, demonstrating a significant efficiency gain over the standard mixed logit model. Moreover, the mixed logit does not allow us to estimate region specific parameters, thus VCL is more efficient and informative.[13]

The structure of the VCL model lends itself to Bayesian estimation methods very easily. While random effects can be estimated in a frequentist manner, as is demonstrated with Yamamoto's (2011) expectation-maximization algorithm for estimation using the VCL, the implementation of the estimation procedure is much easier in a Bayesian hierarchical setting. Assuming that each of the parameters of interest (both random and fixed effects) come from commonly used statistical distributions, generally those within the Gamma family, a Gibbs sampler is easily set up and can be utilized to garner estimates of the parameters of interest.

For applications to this model, we make a few assumptions about the underlying distributions of the parameters of interest. We assume that the Euclidean distance parameter β_k, the competence

[11] If the competence and sociodemographic valences are individual specific, the VCL is able to accommodate parameters of both types by using a random effects hierarchical structure, meaning that the parameters estimated for each region are assumed to come from some probability distribution, generally a normal distribution, as assumed in the SMR model of Section **??**. This method of estimation is best done utilizing random effects.

[12] This is the case for the Bloc Québécois in Canada as the main reason it came into existence was to promote and negotiate the secession of Québec from Canada.

[13] Another alternative is the multinomial probit model, which does not rely on the IIA assumption either. However, the multinomial probit model does not allow the researcher to estimate parameters at the level of the individual choice set, i.e., at the regional level, as the errors are absorbed in the error matrix and, thus, the IIA itself is absorbed. However, as with the mixed logit, the regional values are often of as much interest as those at the national level, so the mixed probit is essentially discarding information that the researcher may find useful.

valence λ_{jk}, the sociodemographic valence α_{jk} and the random effects all have underlying normal distributions.[14] Further, we assume that all of these distributions are independent of one another. This assumption follows from our assumptions that the variables, and thus the draws in the Gibbs sampler, are all orthogonal. We could easily assume that each level of the hierarchy (aggregate, region, sociodemographic) comes from a multivariate normal within itself. Time spent with this model has shown that this assumption is taxing computationally, adding to the amount of time it takes the Gibbs sampler to converge and yielding results that are virtually indiscernible from those garnered when independence is assumed. It is unreasonable, however, to assume that the orthogonality assumption is perfectly met. For example, in some cases, region and location within the policy space are correlated (e.g., the Bloc Québécois in Canada). This assumption violation will lead to biased estimators. While the bias is not large, it is certainly a cause for some concern. Nevertheless, this problem is easily fixed.

Gelman *et al.* (2005) utilize a method to rid random effects of the collinearity which causes the estimates to be biased. They propose that the problem is solved very simply by adding the mean of the covariate of interest as a predictor a level lower in the hierarchy than the random effect of interest. In this case, given a specific party, the mean of its regional level random effects and the mean of its sociodemographic level random effects are indeed situated at the respective mean of the difference of Euclidian differences between the party of interest and the base party. Given that this is the covariate that will theoretically be correlated with sociodemographic group and region, this is the mean that we need to include it as a predictor in the random effects as also assumed in the SMR model in Section **??**. In doing this, we control for the discrepancy as if it is an omitted variable and allows the random effect to take care of its own correlation. The *normal priors* in this case can still be diffuse, but the mean needs to be at the specified value to fix the problem.

With regards to *prior specification* for the parameters, we choose to use the conjugate priors for each of the parameters of interest. This is to say that we choose to utilize *normal priors* on the policy weights β_k and on the competence and sociodemographic valences, λ_{jk} and α_{jk} for all $k \in \Re$ and utilize inverse-Gamma priors on any variance terms. This is the typical prior specification. Though some models have trouble achieving convergence to a stationary distribution when given diffuse priors, this model is normal enough in its specification that diffuse priors do not cause problems. Therefore, for the VCL model, we tend to utilize very diffuse priors where we define the priors as having a mean of zero and a very high variance. While no proper prior can be completely uninformative, these high variance priors allow the priors to provide very little information to the model. Thus, these priors allow for the estimates to be almost completely driven by the data. For the purposes of this model, having priors centered at zero can be seen as a more stringent test of the estimates, as the priors will very slightly drag estimates towards zero, or the fixed estimate of the intercept for the base group. However, this model is flexible enough to incorporate any priors.

One practical note is necessary regarding the time necessary to achieve convergence within the model. Convergence of the VCL can be quite slow given a large number of choice set types (i.e., regions) and voters. Similarly, as random effects are estimated for each party, the number of parties and the number of sociodemographic groups can slow down the rate at which samples are derived from the Gibbs sampler. Though it is a time consuming method, the sheer amount of information gained from the VCL is, thus, the best choice when it is necessary to use a discrete choice model which does not rely on IIA.

Using the formal model and the VCL methodology we now study the 2004 Canadian election.

[14] In the formal model in Section **??**, we assume that λ_{jk} and α_{jk} are the mean of the voter's competence and sociodemographic valences in region k. The assumptions of the formal and empirical models then match, thus making the transition to applying the formal model easier.

4 Application to the 2004 Canadian Election

Since 1921, Canadians have elected at least three different parties to the Federal legislature and 2004 was no different. However, the 2004 election in Canada was significant because it yielded the first minority government for Canada since 1979.

Facing the political fall from the Sponsorship scandal, on May 22, Paul Martin, the newly minted un-elected prime minister, was forced to call an early election for June 28, 2004. The 2004 campaign did not run smoothly for the two major parties, the Liberals and the new Conservatives. The pre-election polls consistently showed both in a "neck-and-neck" race making "the election too close to call."[15] By mid-campaign the Conservatives were slightly ahead of the Liberals. However, the polls consistently showed that, regardless of who was ahead, the winning party would only form a minority government.[16] As the campaign advanced, the Conservatives made two major mistakes. A Conservative MP accused Martin of being soft on child pornography and Ralph Kline, the Progressive Conservative premier of Alberta, announced that his government was considering a two-tier health care system that would include a substantial private sector component. The Liberals and many Canadians reacted strongly against both issues. Changing gears, the Liberals portrayed Harper as an extreme right-wing Conservative, encouraging New Democratic Party (NDP)-supporters to vote strategically. By the last week of the campaign the Liberals were ahead of the Conservatives with polls indicating that the Liberals would only win a minority government.The Empirical Appendix gives details of the 2004 election results.

Regional differences were of primal importance in this election.[17] The Liberals rating plummeted when Sponsorship scandal broke, specially in Québec with Québécers massively turning to the Bloc Québécois. The Liberals partially recovered from this blow as indicated by the late campaign poll (see e.g., provincial results of the Ekos June 21-24, 2004 poll[18] in Table **A3**). This coupled with the resurgence of support for Québec sovereignty meant that, in contrast to their situation in the rest of Canada, the Liberals main competitor in Québec was the Bloc Québécois (polling at 51%) not the Conservatives (polling at 11%). The Liberals could then not ignore the effect the Bloc would have on its electoral prospects in Québec. Moreover, the Liberals while ahead in Ontario, were polling poorly in Alberta and less so in British Columbia (BC) and the Prairie Provinces. The Conservatives who dominated in the Western provinces (BC, Alberta and the Prairie provinces) were slightly behind the Liberals in Ontario, and were polling low in Québec. The NDP understood it was polling low everywhere but especially in Québec. When choosing their policies, the parties who understood they faced different political environments in Québec and the rest of Canada must have adjusted their policies to account for these differences. Given the difference in political environments we study the election in these two regions: Québec and Canada outside Québec.

The 2004 election results are given in Tables 1B (national) and 1C (by province). The Liberals under Martin won the 2004 election with 135 (44%) seats out of 308, down 37 from the 2000 election. This was the first minority government since 1979 (informally supported by the NDP). Relative to the 2000 election, the Liberals lost votes in Ontario and in Québec winning 75 out of 106 Ontario seats in 2004 (down from 100 out of 103 in 2000) and 21 out of 75 Québec seats in 2004 (down from 36 out of 75 in 2000). They held onto the 14 seats they had in the Western provinces since 2000, gaining in British Columbia and losing in Manitoba.

[15] Canadians were polled almost on a daily basis throughout the campaign with no coverage in the first week or the last last three days of the campaign (Pickup and Johnston, 2007).

[16] The last time a party won more than fifty percent of the vote in Canada was in 1984.

[17] As happens in federations where regional differences are accentuated by various political events (Riker 1987).

[18] The 5,254 sample reflects the regional, gender and age composition of the Canadian population in the Census (see http://www.ekos.com/admin/articles/26June2004BackgroundDoc.pdf).

The Conservatives won the second largest number of seats, wining more seats (99) than both of its two predecessors in 2000 (Alliance 66 and Progressive Conservatives, PC, 12). Its vote share (30%) was, however, lower than that of its predecessors combined (Alliance 26% and PC 12%). Support for the CP came mainly from Western Canada and in spite of making some progress in Ontario, gaining 24 seats, they failed to make in roads in the Atlantic Provinces. They won no seats in Québec. The Conservatives were still seen by many as mainly representing western interests.

Support for the Bloc Québécois soared in 2004 as almost half (49%) of Québécers voted for them, thus winning 54 out of 75 Québec seats with 12.4% of the national vote. The NDP, the other major winner in this election, almost doubled its vote share relative to 2000 and managed to add 6 members to its caucus mostly in Ontario and British Columbia. The Greens' support increased relative to 2000 but starting from a very low base, won no seats.

4.1 Policy Dimensions and Sociodemographic Data

To study the 2004 Canadian election we used the survey data collected by Blais *et al.* (2006). Table **A4** of the Empirical Appendix shows the actual and sample vote shares. The similarity between these two sets of vote shares suggests that the sample is fairly representative of the Canadian electorate. Table **A4** also has the data for Québec, as the Bloc Québécois only ran in Québec.

We used voters' responses to the survey questions listed in Table **1** taken from Blais *et al.* (2006) to estimate their position in the latent policy space. The factor analysis first finds the correlation between these questions, then determines a lower number of unobserved variables or factors. The factor analysis led us to conclude that there were two latent factors or policy dimensions: one "social," the other "decentralization."[19] Table **2** gives the factor loadings or weights that the factor analysis assigns to each question. Using these weights we identified that the social dimension as a weighted combination of voters' attitudes towards (1) the gap between poor and rich, (2) helping women, (3) gun control, (4) the war in Iraq and (5) their position the left-right scale. We coded the social dimension such that *lower* values along this dimension imply *higher* interest in social programs so as to have a left-right scale along this axis. The decentralization dimension included voters' attitudes towards (1) the welfare state, (2) their standard of living, (3) inter-jurisdictional job mobility, (4) helping Québec and (5) the influence of Federal versus Provincial governments in respondents lives. A *greater* desire for decentralization implies *higher* values along this axis.

[**Insert Tables 1 and 2 here**]

Using the factor loadings from the factor analysis given in Table **2** we computed the position of each voter along the social and decentralization dimensions. The mean and median values of voters' positions along these two dimensions in Canada are at $(0, 0)$, the origin (see Table **A1**). To illustrate, a voter who thinks that more should be done to reduce the gap between rich and poor would tend to be on the left of the Social (S) axis ($x - axis$), while a voter who believes that the federal government does a better job of looking after peoples' interests would have a negative position on the D ($= y - axis$), and could be regarded as opposed to decentralization.

The survey asked voters which party they would be voting for, so we estimated party positions as the mean of voters for that party.[20] From Table **A2**, the party positions in the policy space are

[19] The factor analysis performed on these questions showed evidence of only two factors or dimensions. Given no evidence of a third factor, the analysis below is carried out using a two dimensional space.

[20] While using the mean is a crude measure of party position, other methods, more computationally intensive, provide similar estimates. For example, we used Aldrich and McKelvey **(1977)** scores to place the parties in the latent policy space. The positions found were not very different from the mean estimates. To check the robustness of estimates from the VCL model with regards to party position we jittered the positions for each party taking 100

then given by the vector:

$$z^*_{Nat} = \begin{bmatrix} & Lib & Con & NDP & Grn & BQ \\ S & -0.17 & 1.27 & -.78 & -0.63 & -1.48 \\ D & -0.38 & 0.32 & 0.05 & -0.13 & 0.23 \end{bmatrix}$$

These party positions correspond closely with those estimated by Benoit and Laver (2006), obtained using expert opinions in 2000. As with these estimates, the Liberal Party locates on the lower left quadrant while the Conservatives lie opposite in the upper right quadrant. Figure 1 and Table B1 show the distribution of voters and the party position for all of Canada with "Q" representing the electoral mean in Québec, which differs from the national electoral mean. While the mean Québécer wants more social programs than the average Canadian; the average Québec and non-Québec voters and thus average Canadian voter seem neutral with regards to decentralization. This contrasts with the average Bloc Québecois supporter who wants greater decentralization (see the above vector of party positions).[21] Figure 2 shows the voter distribution for Québec only.

[**Insert Figures 1 and 2 here**]

The *electoral covariance matrix* for the entire sample of 862 respondents ∇^{Canada} (Table **A1**) is

$$\nabla^{Canada} = \begin{bmatrix} & S & D \\ S & \sigma_S^2 = 2.78 & \sigma_{SD}^C = 0.0 \\ D & \sigma_{SD}^C = 0.0 & \sigma_D^2 = 1.14 \end{bmatrix}.$$

While at the national level there is no covariance between the two dimensions, $\sigma_{SD}^C = 0.0$, the variances on these two orthogonal axes differ. The "total"variance is $\sigma_C^2 \equiv \sigma_S^2 + \sigma_D^2 = 2.78 + 1.14 = 3.92$ with an *electoral standard deviation* (*esd*) $\sigma_C = 1.98$. We also have that for the sample outside Québec (C/Q) of 675 respondents the electoral covariance matrix is

$$\nabla^{C/Q} = \begin{bmatrix} 2.70 & 0.12 \\ 0.12 & 1.18 \end{bmatrix}$$

The "total" variance is $\sigma_{C/Q}^2 \equiv \sigma_S^2 + \sigma_D^2 = 3.88$ with an esd $\sigma_{C/Q} = 1.97$. For C/Q, the variance along the social dimension is smaller and along the decentralization dimension higher than in the national sample. While the two dimensions seem orthogonal to each other for Canada, the covariance between them is positive in the C/Q sample. Québec, with a sample of 187 respondents, has an electoral covariance matrix

$$\nabla^{Q} = \begin{bmatrix} 1.48 & -0.57 \\ -0.57 & 0.98 \end{bmatrix}$$

random samples from a bivariate normal distribution centered at the mean party positions with a variance of 1 on each axis and no covariance and ran the VCL model. The results show that differences in the estimates only occurred when the draws were far away from the mean positions, meaning that small changes on party positions had little influence the outcome. Given the strong prior information on where parties should be and since these positions match closely with estimates from other papers, we feel confident that using the mean of those voting for the party is a reasonable method for estimating the positions of parties within the created latent policy space.

[21] Supporters of the Bloc Québecois are mainly French Québécers who want the cessation of Québec from Canada. Note that not all French Québécers support the Bloc or want greater decentralization. Moreover, according to the 2006 Census, 40% of the Québec population is none French speaking. Polls suggest that non-French speaking Québécers want Québec to remain in Canada and support greater centralization. It is then not surprising to find that the mean Québécer is neutrally located along the decentralization dimension.

whose "total" variance is $\sigma_Q^2 = 2.46$ with esd $\sigma_Q = 1.57$. The variances along the two dimensions in Québec are smaller than in all of Canada. Moreover, while in all of Canada and in C/Q the covariance between the two dimensions is zero, or close to zero, for the Québec sample it is negative.

The differences in the electoral covariance matrices between the C/Q and Q samples show that the electoral distributions in the two regions, as illustrated in Figures 1 and 2, differ. In addition, non-Québécers prefer fewer social programs and more centralization than Québécers (Table B1). The median Québécer is to the left of the mean Québécer in the decentralization dimension.

Figures 1 and 2 together with Tables 1B and 1C and the electoral covariance matrices suggest that there are significant regional differences across the electorate in these two regions. These differences are driven by Québec and Alberta whose residents wanted greater decentralization but for different reasons. Québécers wanted to ensure the survival their culture, language, laws, and to control the composition of its population by managing its immigration policy. Thus, due to its distinct nature, Québécers wanted decentralization for cultural reasons. Albertans wanted control over the regions vast natural resources, mainly its oil sands, and did not want to share its oil revenues with the rest of Canada. Thus, Alberta wanted economic decentralization.

For each respondent, the survey collected their sex, age, and education level. We coded age into three categories (18-29, 30-65, and 65+) and education into two categories (college and no college degree). While there are no major sociodemographic differences between non-Québec and Québec respondents (Table B1), there are differences in the characteristics of party supporters (Table B2). The mean Liberal supporter is older than that of other parties with the youngest mean supporter voting Green. More than half of those voting Liberal, NDP and BQ were women with more than half of those voting Conservatives or Green being men.

4.2 Modelling the 2004 Canadian election

Clarke *et al.*(2005) point out that in the last stages of the campaign there was no clear winner as the two front runners, the Liberals or Conservatives, were running neck-and-neck. Polls indicated that neither party would win a majority of votes. Since support for both parties hovering around 33−35% (Table **A3**), neither party was expected to win a parliamentary majority either. Parties, voters and political commentators were speculating on which party would form a minority government.

With Canada having a first-past-the-post system and the election too close to call, parties were targeting marginal ridings. Spending more resources (e.g., on canvassing or advertising) on marginal ridings is a totally different aspect of the campaign than finding the policies that lead to electoral victories. Parties may cater their message when their leader visits a particular riding, but they don't design election policies to target individual ridings. For if they did, the party risks alienating not only its core supporters but also voters in another riding and would also be vulnerable to changes in voters' mood. To avoid being seen as wavering in their position parties make only small changes to their policy during the campaign.

The election had two different battle fronts: in Québec between the Liberals and the Bloc and in the rest of Canada, basically in Ontario, between the Liberals and Conservatives.[22] Table **A3** shows that the Conservatives had given up in Québec but were trying to break through in Eastern Canada mainly in Ontario. The Liberals had given up in the West and, as in previous elections, knew they needed to win Ontario and not perform too badly in Québec to stay in office. Even though marginal ridings were important, the main battles were at the Ontario and Québec levels. Under these conditions, current accurate information at the provincial level is important to the parties. It is then not surprising that this marked the first election in which Canadians were polled

[22] This is not uncommon in Federal system with vast regional differences (Riker, 1982).

on a daily basis. These public polls were closely followed by all parties. They reported national and provincial vote shares. Even though poll sizes were large for a national poll, they were too small to give riding level shares. We can assume that parties used the vote shares given by these public polls and by their own internal polls as a heuristic measure of the electoral outcome when making decisions about changes in positions. Thus, with the election too close to call and no party expected to win a majority, it seems reasonable to assume that the parties choose policies to maximize their vote shares.

4.3 VCL Model of the 2004 Canadian Election

We now examine whether Canada's national parties[23] converged to the national mean and whether the Bloc Québécois (BQ) converged to the Québec mean in the 2004 election. To do so we apply the formal *stochastic multi-regional* (SMR) model presented in the Technical Appendix to the election in the two regions (Canada outside Québec and Québec) using the estimates provided by the *varying choice set logit* (VCL) model presented in Section 3.

We use the voter's utility function in (15) to estimate the probability that voter i votes for party j in region k in (16) for all national parties and the BQ in Québec and in the rest of Canada. Using the VCL model we estimate the competence and sociodemographic valences for all national parties and for the BQ, $(\lambda_{jk}, \alpha_{jk})$ for $j \in LP, CP, NDP, GPC, BQ$ and the regional weights given by voters to the policy differences with parties, β_k in the two regions $k \in C/Q, Q$. In the results presented below, we report the effect that education by age group has on the probability that i votes for j in each region rather than reporting the average regional sociodemographic valence α_{jk}. The competence valence, λ_{jk}, is given by the intercept in the VCL regression once voters' sociodemographic characteristics and their differences with party policies are taken into account.[24]. We use the Liberal Party as the base party in the VCL estimations, so that the coefficients of the models are measured relative to that of the Liberals whose coefficients are standardized to be zero.

Note that due to the structure of the VCL and the underlying random effects model, sociodemographics are viewed as categorical so that groups can be constructed. As noted previously, parsimony is very important when estimating the VCL model as the time to convergence and the time necessary to run the Gibbs sampler can be long (each sociodemographic group has a random effect for each region being considered), thus we examine the relationships between the variables to see if we should keep them all in the model. After experimentong with the model for some time, we noticed that the relationship between sex and vote was yielded spurious by age and education. Thus, to preserve time and allow the Gibbs sampler to run efficiently, our model does not include sex as a variable.

Given some correlation between the random effects of interest and the independent variable of Euclidian difference, we use the random effects correction procedure proposed in Section 3. We include the mean difference for each party in each region's respective random effects by setting the mean of the normal priors to the random effects at this value as assumed in the SMR model of Section **??**. To assist in convergence of the VCL, we create a diffuse gamma hyperprior for the variance of each prior. As stated before, this model takes a while to converge, so we let the Gibbs sampler run some time. We ran each Gibbs sampler for 100,000 iterations and received nice normal distributions for each of the parameters of interest. Allowing the Gibbs sampler to run this long

[23] The Liberals (LPC), the Conservatives (CPC), the New Democratic Party (NDP) and the Green Party (GPC).

[24] As in related work, we assume that the intercept term of the spatial model for each party can be used as an estimate of the party's competence variance. In our models of US and British politics, we used voter perceptions of candidate traits as estimates of competence valence. However these more refined estimates essentially matched the intercept estimates.

reduces the effects of the inherent autocorrelation that occurs in the sampler.

The VCL parameter estimates and their corresponding 95 percent credible confidence intervals given in Table 4 are derived from these Markov Chain Monte Carlo simulations. We also report the deviance information criterion (DIC), which is a hierarchical model analogue to the Akaike information criterion (AIC) or the Bayesian information criterion (BIC). When the posterior distribution is assumed to be multivariate normal (as it is in this case), the DIC functions as a measure of model quality rewarding a model with a small number of parameters, but penalizing a model that does not fit the data well. The DIC can be seen as a measure of the log-likelihood of the posterior density. Lower values of DIC are preferred.

[**Insert Table 3 about here**]

Table **3** suggest that for Canada outside Québec, the Liberals and Conservatives were considered equally able to govern as the competence valence of the Conservatives is not statistically different from zero and that of the Liberals (the baseline party) is standardized to zero. By adding the competence valence to the non-Quebec regional random effect, we see that Liberals and Conservatives are considered almost equivalent in terms of competence outside Québec. The NDP was considered of lower competence than the Liberal or Conservatives. However, its positioning in the policy space allowed it to be a significant competitor outside of Québec and to win additional seats (see Table B2 and Figures 1 and 2). The Greens have the lowest valence outside Québec, mainly because it is a one-issue party and even though its votes increased relative to 2000, starting from a very low base it won no seats. Their competence valence is also significantly lower than that of the NDP.

Electoral differences between Québec and the rest of Canada come across clearly in Table 4. The failed attempts to bring Québec back into the Constitution and the Liberal Sponsorship Scandal led to the prominence of the BQ in the 2004 election (Table 1A). It is then no surprise that the BQ had the highest competence valence in Québec in 2004. The Liberals, who by the end of the campaign had partially recovered in the pre-election polls (Table A1), came in second in terms of competence. Essentially, with the BQ and the Liberals similarly positioned in the policy space, they compete for many of the same voters in Québec. However, what the BQ's competence valence shows is that the political environment was such that Québécers believed the BQ to be simply better at representing their interests in Ottawa independent of the BQ's position.[25] Recall that the Conservatives had given up in Québec as suggested by the pre-elections polls (Table 1A), it is then no surprise that their valence was significantly lower than that of the Liberals (the base party) and the BQ's. The significantly negative competence valence of the NDP signals that Québécers considered them less able to govern than either the BQ and the Liberals and at par with the Conservatives. Like in the rest of Canada, the Greens had the lowest competence valence in Québec.

We estimated the effect that education has by age group on the probability of voting for each party and report these coefficients for Québec and the rest of Canada (Table **3**).[26] The sociodemographic valences of these groups varied by party and by region once the policy differences with parties and the competence valence are taken into account. For example, voters outside Québec between 30 and 65 who have at least a college education were significantly less likely to vote for the Greens (-2.4) than for the NDP (-1.02) than for the Conservatives (-0.55) relative to the Liberals (0.0). For Québec, voters with less than college education in the 30 to 65 age group were more likely to vote for the BQ than for any other party.

[25] Clearly, with 75 out of 308 seats, the BQ leader can never become prime minister in Canada.

[26] Note that to avoid having too many rows in Table 4, we included the coefficients by education-age group in the form of a sub-matrix in the corresponding regional regression column.

The VCL model also allows us to examine the results at the *national* level. The results are however less clear than at the regional level as none of the competence valences at the national level are statistically different from zero. Thus showing the advantage of using the VCL model to estimate the results by regions within a single model. The estimation at the national level can mask vast regional differences in the estimated parameters. For example, while the BQ's competence valence is significantly higher than that of the Liberals in Québec, at the national level the results suggest that these two parties are considered equally able to govern as the Bloc's competence valence is not statistically different that of the Liberals which is standardized to zero. The insignificance of the Bloc's national competence valence is due to the fact that even though nearly half of Québécers voted for the BQ (Table 1B), its vote share represents only 12.4% of the national vote (Table 1B) as Québec represents only a quarter of the Canadian population.

4.4 Is there convergence to the electoral means in Canada?

We are interested in finding whether parties converge to the their corresponding mean and whether the means represent a local Nash equilibrium (LNE) of the election. We assume that parties use polls and other information at their disposal to form an idea of the anticipated electoral outcome. Parties then use their expectation of the electoral outcome to find a policy position that allows them to maximize their vote shares taking into account their estimates of where other parties locate.

Using voters' positions and the VCL estimates of the parameters $(\beta_k, \lambda_{jk}, \alpha_{jk})$ for $k = C/Q, Q$ given in Table 4, we estimate the vote share function of each party at different policy positions. The assumption here is that parties use the the VCL parameters and the framework of the formal SMR model to form a heuristic empirical model of the anticipated electoral outcome and then use this heuristic model to position themselves so as to maximize their expected vote share. Given a vector of possible positions, we estimate the Hessians for regional and national parties given respectively by (27) and (33), to examine whether these parties are maximizing their expected vote shares at their corresponding means.

Assuming that parties locate at their respective electoral means in Canada translates into national parties locating at the national mean and the BQ at Québec mean. From Table B1, the vector of possible positions $\mathbf{z}^{\mathbf{C}}_{Nat}$ is given by

$$\mathbf{z}^{\mathbf{C}}_{Nat} = \begin{bmatrix} & Lib & Con & NDP & Grn & BQ \\ S & 0 & 0 & 0 & 0 & -1.11 \\ D & 0 & 0 & 0 & 0 & -0.08 \end{bmatrix} \tag{17}$$

The necessary condition for convergence to the possible positions were presented in Section 2.0.1. Remember that if any party fails to meet the necessary condition for convergence at the possible position, this party has an incentive to move away from this position in order to increase its vote share. The analysis in Section 4.3 showed that the Greens have the lowest competence valence in both regions. Thus, if any party has an incentive to move to increase its vote share, it is the Greens.

Taking the VCL parameters estimates given in Table 4, we calculate the observable component of voter's utility $u^*_{ijk}(x_i, z_j)$ using (15) and then estimate the probability that each voter i in Québec and in the rest of Canada votes for each party, ρ_{ijk} for $j \in LP, CP, NDP, GPC, BQ$ using (16). Taking the estimates of ρ_{ijk} for each voter and the weight voters give policy differences with parties β_k for $k = C/Q, Q$ we calculate the weight that each party gives each voter i in each region, μ_{ijk} using (24) and calculate the weight that each national party gives each region θ_{jk} using (**??**).

Taking the estimates of ρ_{ijk}, μ_{ijk} and β_k and voter i's and party j's position in the each dimension, we then calculate party j's convergence coefficient in each region, c_{jk}, for $j \in LP, CP, NDP$,

GPC, BQ and $k = C/Q, Q$ using (6). Then using c_{jk} and θ_{jk}, we calculate the *national* convergence coefficient for each national party c_j for $j \in LP, CP, NDP, GPC$ using (11). Note that being a regional party, we already have the Bloc's convergence coefficient.

We can now test if each party meets the necessary condition for convergence when it locates at its corresponding electoral mean. From Section 2.0.1, we know that the necessary condition is satisfied if the national convergence coefficient of national parties and the BQ's convergence coefficient are each less than the dimension of the policy space, which in the Canadian case is $w = 2$. When we performed the simulations to find the VCL estimates given in Table **4**, we also estimated ρ_{ijk} for all voters, the weights μ_{ijk} and θ_{jk} and the national and regional convergence coefficients c_j for $j \in LP, CP, NDP, GPC$ and c_{BQ} in each simulation. From the simulations we constructed the 95% credible intervals around these estimates.They are also given in Table **4**.

[**Insert Table 4 here**]

We now use the convergence tests given in Results 1 and 2 in Section 2.0.1 to assess convergence to the vector of possible positions, the mean vector given in (17). From Table 5, we see that the national convergence coefficient of the Liberals and Conservatives and the convergence coefficient of the Bloc Québécois are significantly less than 1 and thus significantly less than the dimension of the policy space, $w = 2$. Thus, the necessary condition for convergence to the possible position, the national electoral mean for the national parties and the Québec mean for the BQ given in (17), has been met. Consequently, none of these parties want to move from their corresponding mean.

The convergence coefficient for the NDP is not significantly different from 1 and thus is significantly below $w = 2$. The NDP then meets the convergence criterion. Given that the Greens' convergence coefficient is not significantly different from $w = 2$, it is not clear whether the Greens have an incentive to stay at the national mean.

Since the NDP and the Greens are the lowest valence parties in both regions, using the results in Section 2.0.1 we can also test for convergence to the national mean by decomposing their convergence coefficients into their two constituent portions along the social and the decentralization dimensions using (13). We can then check to see if either or both of these dimensional convergence coefficients are greater than 1. When the convergence coefficient along a particular dimension is greater than one, the eigenvalue along this dimension must be positive. Under this condition, the party at the possible position, the national mean, is not maximizing its vote share. Table **5** gives the estimates for the convergence coefficient for each dimension and their 95% credible intervals.

[**Insert Tables 5 and 6 here**]

The NDP is maximizing its vote share at the national mean when all other parties locate at their respective means, as the convergence coefficients along both dimensions are significantly less than 1 (Table **5**). a The Green Party, with the highest convergence coefficient, has an incentive to move from the national mean when all other parties locate at their respective means. The reason is simple: the Greens are at a saddle point at the national mean as evidenced by the fact that while the convergence coefficient along the social dimension is significantly greater than 1 that along the decentralization dimension it is significantly less than 1. The Greens will then be the first to diverge from the mean vector in (17). Once the Greens locate elsewhere, other parties may also want to move away from their corresponding means. Note that the convergence coefficient of the election is given by that of the party with the highest convergence coefficient, in this case that of the Greens, as given in (14). Since the Greens have the highest convergence coefficient and want to diverge from the national mean, then the mean vector is not a LNE of the election.

4.5 Vote maximizing positions

Taken as they are, we do not know if the information given by the convergence coefficients actually matches the vote maximizing tendencies of the parties. In order to give validity to the proposed tests, we need to use optimization methods to show that the vote maximizing positions for parties are not located on the mean vector. Using computational techniques, we can allow a party to optimize over their voter base given the other parties' positions and allow each party to do this in rotation until no party wants to move anymore.

Given the global optimization procedure used, we can assume that parties' locations upon convergence in the optimizer correspond to the LNE on which they are maximizing their vote shares. Keeping with the theory presented in Section 2, this optimizer also allows for the weights that each national party places on each region θ_{jk} in (**??**) and the weights that each party place on each voter a the regional level μ_{ijk} in (24) to be endogenous to their positions. Given voter's utility in (15), we simply optimize over the weighted average probability that a person votes for a party in the voter's region. This method is necessary given that each party can potentially be optimizing over a different portion of the electorate. In this case, while the four national parties are attempting to optimize their respective vote shares over all of Canada, BQ is only trying to optimize among voters in Québec. Thus, this slightly augmented optimizer is necessary for finding the optimizing positions in Canada.

Figure 3 shows the vote optimizing positions for each party in Canada, which are as follows:

$$z^*_{opt} = \begin{bmatrix} & Lib & Con & NDP & Grn & BQ \\ S & -0.151 & 0.085 & 3.736 & -3.886 & -1.140 \\ D & -0.113 & -0.073 & 0.823 & 1.808 & -0.073 \end{bmatrix}$$

We can also get the vote optimizing weights that the national parties apply to each region, though they really are not all that different from the weights applied at the mean vector:

$$\begin{bmatrix} & C/Quebec & Quebec \\ Lib & .838 & .161 \\ Con & .870 & .130 \\ NDP & .918 & .081 \\ Grn & .880 & .119 \end{bmatrix}$$

[**Insert Figure 3 here**]

Fortunately for our measures, the vote optimizing positions echo what we were told by the convergence coefficients: the Greens have incentive to move away from the national electoral mean while the Liberal and Conservatives want to stay close to the national mean and the BQ wants to stay at the Québec mean. Similarly, the NDP must adapt to the adjustments by other parties and move away from the electoral mean, as well.

Given that the NDP and the Greens have a relatively low competence valence, their relocation has little effect on the maximizing positions for the largest three parties: the Liberal, the Conservatives and the Bloc Québécois. However, it is important to note that the vote optimizing positions for the Liberals and the Conservatives are slightly different from their respective electoral means. This is due to the fact that they can both do slightly better when the Green Party moves away from the electoral mean by taking positions on opposite sides of the electoral mean. Given that these two parties make these small changes, the smaller valence parties are forced to move to the fringes of the electorate to try to get the votes of the extremists.

This begs the question, though, how much better can the parties do at these positions than they did at their current positions? Table **6** shows the vote shares in the sample for each party at their actual positions, at the national and regional electoral means, and at the vote maximizing positions determined by the optimization routine. These vote shares are calculated as the mean probability that a voter votes for each party. Note that these means are unweighted, as the weights are simply perceptions of the parties as to which voters are more likely to be swayed to vote for their party. This is to say that a perceived increase in votes may not actually occur when parties use these weights to calculate optimal positions. However, given their beliefs about voters, they are maximizing their perceived vote shares.

Table 7 demonstrates this point as the Green Party and the NDP actually do worse when they locate at the vote optimizing positions. This is likely due to both the fact that their perceptions did not result in a vote share increase and that when the Green Party moves away from the mean, they initially do better, but the larger national parties can take advantage of this move with small moves of their own causing the Green Party to move to the fringe of the electorate to gain what votes they can. This is to say at the vote maximizing LNE, the Green Party does worse than it would have had the mean vector been an LNE. The NDP suffers from a similar problem and takes a small decrease in votes at the vote maximizing LNE. This is corroborated by the fact that the Liberals and the Conservatives both do better moving slightly away from the mean.

Conclusion

In this paper makes three contributions to the literature. First we develop a formal stochastic multi-regional (SMR) election model, then we present the Variable Choice Logit (VCL) methodology that allows us to estimate the parameters of the voter's utility when the sets of parties varies by region. We then use the SMR model and the VCL methodology to study the 2004 Canadian election.

The formal SMR model allows voters in different regions face different bundles of national and regional parties. Using the first order conditions we find the party's candidate positions. The second order conditions (derived in Appendix A) show that the necessary conditions for parties to converge to or remain at their candidate positions is that the convergence coefficients for national and regional parties must be less than the dimension of the policy space, w. This SMR model generalizes the model presented in Schofield (2007). The SMR model gives the conditions under which there is a LNE for any number of national and regional parties.

The VCL methodology that we propose adapts that developed by Yamamoto (2011) to our multi-regional setting. This methodology relaxes the independence of irrelevant alternatives (IIA) assumed in the multinomial logit (MNL) models. Using the VCL model we estimate the parameters of the voter's utility function in each region, something that cannot be done using the MNL model as it relies on the IIA assumption being satisfied. We show that the VCL model allows us to take advantage of more information than the MNL models do. Thus the VCL model is ideal when examining voting tendencies within complex electorates that have clear hierarchical structures in countries where the sets of parties varies by region.

We use the formal SMR model and the VCL methodology to study the 2004 Canadian election. First we estimate the parameters of the voter's utility using VCL methodology, then applying the framework of the SMR model we examine whether parties converge or not to their respective electoral means using the parameters estimates produced by the VCL model. Our results show that the Bloc Québécois has the highest competence valence in Québec with the Liberal having the second highest, the Conservatives and the NDP coming in third and the greens last. In the rest of Canada, the Liberals and Conservatives are considered as equally able to govern in the 2004 election. Our new methodology for estimating the regional valences allow us to show that while at the national level the competence valence of all parties seem indistinguishable from each other, this masks important regional differences. Québécers and voters in the rest of Canada had very different beliefs on who was more capable of representing their interest in Ottawa. Our analysis also showed that if national parties located at the national electoral mean and the BQ located at the Québec mean then not all parties would be maximizing their respective expected vote shares. Rather, to maximize their vote share the Greens with the lowest competence valence in both regions would prefer to adopt a more extreme position in the policy space to increase its vote shares. Thus, our analysis showed that two parties, the BQ and the Greens would not locate at the national electoral mean. This finding is in direct contrast to widely accepted theories (Downs 1957, and Hinich 1977) that political actors can always maximize their vote shares by taking positions at the electoral center.

5 References

ADAMS, J. (2001). *Party competition and responsible party Government* Ann Arbor, MI, University of Michigan Press.

ADAMS, J. & MERRILL III S. (1999). "Modelin party strategies and policy representation in multiparty elections: Why are strategies so extreme?", *American Journal of Political Science*, 43, 765–781.

ALDRICH, J.H. & MCKELVEY, R.D. (1977). A method of scaling with applications to the 1968 and 1972 presidential elections. *The American Political Science Review* , 71(1): 111–130.

BENOIT, K. & LAVER, M. (2006). *Party policy in modern democracies*, Routledge, London.

BLAIS A., FOURNIER P., GIDENGIL E., NEVITTE N. & EVERITT J. (2006). Election 2006: How big were the changes. . . really?, Working paper, Universite de Montreal.

CAPLIN, A. & NALEBUFF, B. (1991). Aggregation and social choice: A Mean Voter Theorem, *Econometrica*, 59, 1–23.

CLARKE, H.D., KORNBERG, A., MACLEOD, J. & SCOTTO, T.J. (2005). Too close to call: Political choice in Canada, 2004. *PS: Political Sci & Pol.* 38, 247–253.

DOW, J.K. & ENDERSBY, J. (2004). Multinomial Logit and Multinomial Probit: A comparison of choice models for voting research, *Electoral Studies*, 23, 107–122.

DOWNS, A. (1957). *An Economic Theory of Democracy.* New York: Harper and Row.

GALLEGO, M. & SCHOFIELD, N. (2013). The convergence coefficient across political regimes. Unpublished Manuscript. Washington University in St. Louis.

GELMAN, A., PARK, D., SHOR, B., BAFUMI, J. & CORTINA, J. (2008). *Red State, blue State, rich state, poor state: Why Americans Vote the Way They Do,*Princeton, NJ: Princeton University Press.

HINICH, M.J. (1977). Equilibrium in Spatial Voting: The median voter theorem is an artifact. *Journal of Economic Theory* 16:208–219.

HOTELLING, H. (1929). Stability in competition, *Economic Journal*, 39, 41–57.

KRAMER G. 1978. Existence of electoral equilibrium, in P. Ordeshook (Ed.) *Game theory and political science.*New York: New York University Press 375–389.

MERRILL, S. & GROFMAN, B. (1999). *A unified theory of voting.* Cambridge, U.K., Cambridge University Press.

PATTY, J. W. (2005), Local equilibrium equivalence in probabilistic voting models, *Games and Economic Behavior*, 51, 523–536.

PATTY, J. W. (2006), Generic difference of expected vote share and probability of victory maximization in simple plurality elections with probabilistic voters, *Social Choice and Welfare*, 28(1): 149–173.

PENN, E (2009) A model of far-sighted voting. *American Journal of Political Science*, 53:36–54.

PICKUP, M, JOHNSON, R. (2007). Campaign trial heats as electoral information: Evidence from the 2004 and 2006 Canadian federal elections. *Electoral Studies* 26, 460–476.

POOLE, K. and ROSENTHAL, H. (1984), US presidential elections 1968–1980, *American Journal of Political Science*, 28, 283–312.

QUINN, K., MARTIN, A. and WHITFORD, A. (1999), Voter choice in multiparty democracies, *American Journal of Political Science*, 43, 1231–1247.

RIKER, W. H. (1964). *Federalism: origin, operation, maintenance.* Boston. MA, Little Brown.

RIKER, W. H. (1987). *The Development of American federalism.* Boston, MA, Kluwer.

RIKER, W. H. and ORDESHOOK, P. C. (1973) *An introduction to positive political theory* Englewood Cliffs, NJ: Prentice-Hall.

RIKER, W H. (1980). Implications from the disequilibrium of majority rule for the study of institutions. *American Political Science Review* 74, 432–446.

RIKER, W H. (1982). The Two Party System and Duverger's law: An essay on the history of political science.. *American Political Science Review* , 76:753–766.

ROEMER, J. E. (2011). A theory of income taxation where politicians focus upon core and swing voters. *Social Choice and Welfare,* 36, pp. 383-422.

SAARI, D. (1997) The Generic existence of a core for q-Rules. *Economic Theory* 9:219–260

SCHOFIELD, N. (1978)Instability of simple dynamic games. *Review of Economic Studies* 45: 575–594.

SCHOFIELD, N. (1983) Generic instability of majority rule. *Review of Economic Studies.* 50: 695–705

SCHOFIELD, N. (2006) Equilibria in the spatial stochastic model of voting with party activists, *Review of Economic Design* 10(3): 183–203.

SCHOFIELD, N. (2007), The mean voter theorem: Necessary and sufficient conditions for convergent equilibrium, *Review of Economic Studies* 74: 965–980.

SCHOFIELD, N. ZAKHAROV, A. (2010). A stochastic model of the 2007 Russian Duma election, *Public Choic*e 142, 177-194.

SCHOFIELD, N. GALLEGO, M. OZDEMIR, U. and ZAKHAROV, A. (2011a)Competition for Popular Support: A valence model of elections in Turkey ." *Social Choice and Welfare* 36(3-4): 451-48

SCHOFIELD, N.,CLAASSEN, C, OZDEMIR, U., and ZAKHAROV, A. (2011b)Estimating the effects of activists in two-party and multiparty systems: A comparison of the United Atates in 2008 and Israel in 1996. *Social Choice and Welfare* 36(3-4)483-518.

SCHOFIELD, N. CLAASSEN,C. OZDEMIR, U. (2011c). Empirical and formal models of the US presidential elections in 2004 and 2008. In: Schofield N, Caballero G. (Eds.) The political economy of institutions, democracy and voting, Springer, Berlin, pp.217-258.

SCHOFIELD, N. GALLEGO, M and J.S.JEON (2011e) Leaders, Voters and activists in elections in the Great Britain 2005 and 2010. *Electoral Studies* 30(3): 484–496.

SCHOFIELD, N, J.S.JEON, MUSKHELISHVILI M, OZDEMIR, U, and TAVITS M. 2011d. Modeling elections in post-communist regimes: Voter perceptions, political leaders and activists, in: Schofield, N., Caballero G. (Eds.) The political economy of institutions, democracy and voting, Springer, Berlin, pp.259-301.

SCHOFIELD, N, J.S.JEON, MUSKHELISHVILI M (2012). Modeling elections in the Caucasus. *Journal of Elections, Public Opinion and Parties* 22(2):187-214.

SCHOFIELD, N., MARTIN, A., QUINN, K. and WHITFORD, A. (1998), Multiparty Electoral Competition in the Netherlands and Germany: A model based on mutinomial probit, *Public Choice*, 97: 257–293.

TRAIN, K. (2003) *Discrete choice methods for simulation.* Cambridge, U.K., Cambridge University Press.

YAMAMOTO, T. (2011), "A Multinomial Response Model for Varying Choice Sets, with Application to Partially Contested Multiparty Elections. Typescript.

5.1 Acknowldgement

This paper was presented at the EPSA conference, Barcelona, June 2013; at the political economy panel at the SAET Meeting, Paris, July 2013; the Public Economic Theory (PET) conference in Lisbon, July 2013. We thank seminar participants for their comments. We also thank the two anonymous referees of this paper for their very helpful suggestions.

6 Technical Appendix :The Stochastic Multi-Regional Model

We model an election in a country where there are regional differences and where voters in different regions face different combinations of national and regional agents,[27] that is, face varying sets of parties. We study the parties choice of position and voters choice of party by taking into account whether the agent is a regional or a national party and the region in which the parties compete.

Prior to the election, all regional and national parties simultaneously announce their policy position in X, an open convex subset of Euclidian space, $\mathbb{R}^w$, where w is finite and represents the number of dimensions of the policy space. Whereas national parties run on the same platform in all regions of the country, regional parties cater only to voters in their own jurisdiction.

Let z_j represent the policy position of a national party in X and z_{jk} denote the position of party j in region k, regardless of whether it is a regional or national party. The set of national parties is denoted by $P_{Nat} = 1, ..., p$ and the set of regional parties in region k by $P_k = 1, ..., q_k$ for $k \in \mathbb{R} = 1, .., r$ so that the set of regional parties may vary by region. Regions may have more than one regional party. When region k has no regional parties, $P_k = \emptyset$.

Given that voters in region k vote only for parties competing in their region, let $\mathbf{z}_k$ denote the vector of policy positions of the parties[28] competing in region k

$$\mathbf{z}_k = (z_1, ..., z_p, z_{k1}, ...z_{kq_k}) \in X^{b_k} \quad \text{where} \quad b_k = p + q_k \text{ for } k \in \Re = \{1, .., r\}.$$

The positions of all parties across all regions represented by $\mathbf{z}_{Nat}$ is given by

$$\mathbf{z}_{Nat} = \bigcup_{k=1}^{r} \mathbf{z}_k \in X^b \quad \text{where} \quad b = p + q_1 + q_2 + ... + q_r$$

Let n_k represent the number of voters in region k. The total number of voters in the country is the sum of voters across all regions, $n = \sum_{k=1}^{r} n_k$. Denote the set of voters in region k by N_k and the set of voters at the national level by $N = \bigcup_{k=1}^{r} N_k$.

For voters in region k, denote voter i's ideal policy by $x_i \in X$ and i's utility by $u_{ik}(x_i, \mathbf{z}_k) = (u_{i1k}(x_i, z_1), \ldots, u_{ijk}(x_i, z_j), ...u_{ijq_k}(x_i, z_{q_k}))$ for $j \in P_{Nat} \cup P_k$ where voter i's utility from party j in region k is

$$u_{ijk}(x_i, z_j) = \lambda_{jk} + \alpha_{jk} - \beta_k \|x_i - z_j\|^2 + \epsilon_{ij} = u^*_{ijk}(x_i, z_j) + \epsilon_{ijk} \tag{18}$$

[27] We use agent and party interchangeably throughout the paper.

[28] Voters in region k indirectly care about the position of all parties competing in all regions of the country as their position affects the location of all parties competing in region k.

Here, $u^*_{ijk}(x_i, z_j)$ is the observable component of voter i's utility associated with party j in region k. The term λ_{jk} is the *competence valence* for agent j in region k. This valence is common across all voters in region k and gives an estimate of the perceived "quality" of party j or of j's ability to govern. We model voters' common belief on j's quality by assuming that an individual voter's perception is distributed around the mean perception in region k, i.e., $\lambda_{ijk} = \lambda_{jk} + \xi_{ijk}$ where ξ_{ijk} is a random iid shock specific to region k. This regional valence is independent of party positions. Moreover, since regional party j in region k never runs in other regions of the country, the model says nothing about the belief that voters in other regions have on j's ability to govern. This is not a problem as voters outside of region k cannot vote for regional party j in region k.

The sociodemographic aspects of voting for voters in region k are modelled by θ_k, a set of s -vectors $\{\theta_{jk} : j \in P_{Nat} \cup P_k\}$ representing the effect of the s different sociodemographic parameters (gender, age, class, education, financial situation, etc.) have on voting for party j in region k while η_i is an s-vector denoting voter i^{th} individual's sociodemographic characteristics. The composition $\alpha_{ijk} = \{(\theta_{jk} \cdot \eta_i)\}$ is a scalar product representing voter i's *sociodemographic valence* for party j in region k. We assume that voters with common sociodemographic characteristics share a common evaluation or bias for party j that is captured by their sociodemographic characteristics. We model this by assuming that an individual voter's sociodemographic valence varies around the mean sociodemographic valence in region k, $\alpha_{ijk} = \alpha_{jk} + \nu_{ijk}$ where ν_{ijk} is a random iid shock specific to region k. Thus, the sociodemographic valence α_{jk} is the "average" sociodemographic valence of voters in region k for party j. These regional sociodemographic valences are independent of party positions. The competence valence λ_{jk} measures an average assessment of party j's ability to govern by voters in region k and since we control for voters' sociodemographic biases, λ_{jk} measures j's ability to govern *net* of any sociodemographic bias these votes may have.

The term $\|x_i - z_j\|$ is the Euclidean distance between voter i's ideal policies x_i and party j's position z_j. The coefficient β_k is the *weight* given to policy differences with party j by all voters in region k. This weight varies by region to allow preferences to differ across regions. Differences that in some regions were deep enough in the past to have lead to the emergence of regional parties.

The error term ϵ_{ijk}, commonly distributed among all voters in region k, come from a Type-I extreme value distribution. Assumption also made in empirical models below which makes the transition to applying this theoretical model to the 2004 Canadian election easier.

To find parties' policy positions in a model where varying sets of parties compete in different regions, the analysis must be first carried out at the regional level before moving to the national level. We begin by examining the parties' positioning game in region $k \in \Re$.

Given the stochastic assumption of the model and the parties' policy positions in region k, $\mathbf{z}_k$, the probability that voter i votes for party j in region k is

$$\begin{aligned}\rho_{ijk}(\mathbf{z}_k) &= \Pr[u_{ijk}(x_i, z_j) > u_{ihk}(x_i, z_h), \text{ for all } h \neq j \in P_{Nat} \cup P_k,] \\ &= \Pr[\epsilon_{hk} - \epsilon_{jk} < u^*_{ijk}(x_i, z_j) - u^*_{ihk}(x_i, z_h), \text{ for all } h \neq j \in P_{Nat} \cup P_k]\end{aligned}$$

where the last line follows after substituting in (18) and Pr stands for the probability operator generated by the distribution assumption on ϵ. Thus, the probability that i votes for j in region k is given by the probability that $u_{ijk}(x_i, z_j) > u_{ihk}(x_i, z_h)$, for all j and h in $P_{Nat} \cup P_k$, i.e., that i gets a higher utility from j than from any other party competing in region k.

With the errors coming from a Type-I extreme value distribution and given the vector of party policy positions $\mathbf{z}_k$, the probability that i votes for j in region k has a logit specification, i.e.,

$$\rho_{ijk} \equiv \rho_{ijk}(\mathbf{z}_k) = \frac{\exp[u^*_{ijk}(x_i, z_j)]}{\sum_{h=1}^{p+q_k} \exp[u^*_{ihk}(x_i, z_h)]} = \frac{1}{\sum_{k=1}^{p+q_k} \exp[u^*_{ihk}(x_i, z_h) - u^*_{ijk}(x_i, z_j)]} \qquad (19)$$

for all $j \in P_{Nat} \cup P_k$ where to simply notation we take the dependence of ρ_{ijk} on $\mathbf{z}_k$ as understood.

This *stochastic multi-regional* (SMR) model does not rely on the *independence of irrelevant alternatives* (IIA) assumption made in Multinomial Logit (MNL) models since we allow the presence of, say, party ℓ to affect voter choices between, say, parties j and h. This is particularly important in our model since voters in different regions face different bundles of parties in the election. Note, that when there is only one region, our SMR model reduces to that developed in Schofield (2007).

Since voters' decisions are stochastic in our framework, parties cannot perfectly anticipate how voters will vote but can estimate their *expected* vote shares. With varying sets of parties competing in different regions, agents can estimate their expected regional vote share in each region and given these regional vote shares, national parties can estimate their expected national vote shares.

For party $j \in P_{Nat} \cup P_k$ competing in region k, its *expected vote share in region* k is the average of the probabilities over voters in region k, i.e.,

$$V_{jk}(\mathbf{z}_k) = \frac{1}{n_k} \sum\nolimits_{i \in N_k} \rho_{ijk} \quad \text{for } j \in P_{Nat} \cup P_k, \tag{20}$$

with the sum of vote shares in each region adding up to 1, $\sum_{j \in P_{Nat} \cup P_k} V_{jk}(\mathbf{z}_k) = 1$ for all $k \in \Re$.

National parties must, in addition, take into account that their *expected* vote share depends on all voters in the country. However, due to the presence of regional parties and since the number of voters varies across regions, the expected national vote share of party j *cannot* be estimated as the average of the probabilities of voters across the country. Rather, j's expected national vote share depends on the vote share j expects to obtain in each region in the country. We assume that the *expected national vote share of party* j is the weighted average of its expected vote share in each region, where the weight of region k is given by the proportion[29] of voters in region k, $\frac{n_k}{n}$, i.e.,

$$V_j(\mathbf{z}_{Nat}) = \sum\nolimits_{k \in \Re} \frac{n_k}{n} V_{jk}(\mathbf{z}_k) = \sum\nolimits_{k \in \Re} \frac{n_k}{n} \frac{1}{n_k} \sum\nolimits_{i \in N_k} \rho_{ijk} = \frac{1}{n} \sum\nolimits_{k \in \Re} \sum\nolimits_{i \in N_k} \rho_{ijk}. \tag{21}$$

The third term in (21) follows after substituting in (20). Note that due to the presence of regional parties, the sum of the vote share of national parties do not add to 1.

The objective is to find the *local Nash equilibria* (LNE) of party positions where each party takes the position of all the other national and regional parties as well as that of voters as given.

6.1 Equilibrium positions

A vector of party policy positions, $\mathbf{z}^*_{Nat} \equiv \bigcup_{k=1}^{r} \mathbf{z}^*_k$, is a *local Nash equilibrium*, LNE, if each party locates itself at a local maximum in its vote share function. This means, that given the opportunity to make moves in the policy space and relocate its policy platform, no vote-maximizing party would choose to do so. We assume that parties can estimate how their vote shares would change if they *marginally* move their policy position. The local Nash equilibrium is that vector $\mathbf{z}_{Nat}$ of party positions so that no party may shift position by a small amount to increase its vote share either at the national or regional level. More formally, a LNE is a vector $\mathbf{z}_{Nat}$ such that for all national parties, their vote share functions, $V_j(\mathbf{z}_{Nat})$ for $j \in P_{Nat}$ and for all regional parties their vote share functions $V_{hk}(\mathbf{z}_k)$ for $h \in P_k$ and $k \in \Re$, are weakly locally maximized at their corresponding positions. To avoid problems with zero eigenvalues we also define a *SLNE* to be a vector that *strictly*

[29] We could have assumed instead that the weight of each region depends on the share of seats each region gets in the national parliament. The results presented below would then depend on seat rather than vote shares but would remain substantially unchanged. Note that the number of parliamentary seats that each region gets is, in general, based on the proportion of the population living in that region.

locally maximizes $V_j(\mathbf{z}_{Nat})$ and $V_{hk}(\mathbf{z}_k)$. Using the estimated coefficients of the SMR model we simulate these models and then relate any vector of party positions, $\mathbf{z}_{Nat}$, to a vector of vote share functions $V(\mathbf{z}_{Nat}) = (V_1(\mathbf{z}_{Nat}), ..., V_p(\mathbf{z}_{Nat}), V_{11}(\mathbf{z}_1), ..., V_{q_1 1}(\mathbf{z}_1), ..., V_{r1}(\mathbf{z}_r), ..., V_{rq_r}(\mathbf{z}_r))$, predicted by the model with p national parties and $q_1, ..., q_r$ regional parties where $\mathbf{z}_k$ for $k \in \Re$ is the vector of party positions restricted to those competing in region k.

Parties' positions are LNE at $z_{Nat} \equiv \bigcup_{k=1}^{r} z_k$, if and only if (iff) all agents are maximizing their vote share functions at z_{Nat}. Suppose parties position themselves at their corresponding positions in z_{Nat}, then parties will be at a local maximum if two conditions are satisfied. The first order condition (FOC) identifies the critical points of the vote share function, that is, the points at which the party's vote share function is at a maximum, minimum or a saddle point at z_{Nat}. To find these critical points take the first derivative of the vote function and set it equal to zero. To find whether a critical point is a maximum, minimum or a saddle point of the party's vote share function we need the second order condition (SOC) at z_{Nat}. This SOC is the second derivative of vote share function, called the Hessian matrix of second derivatives of the vote share function of each party. At z_{Nat}, the party is a maximum if the Hessian is negative definite which happens only when the eigenvalues of the Hessian are all negative at z_{Nat}.

6.1.1 Parties' critical points (FOC)

Let us now find the critical points of the vote share of party j for $j \in P_{Nat} \cup \bigcup_{k=1}^{r} P_k$. Note that since regional and national parties face different electorates, their positioning decisions must be studied separately. Thus, the region in which the parties compete must also be taken into account.

We begin our analysis in region k. Given the vector of policy positions $\mathbf{z}_{Nat}$, and since the probability that voter i votes for party j in region k, whether a national or a regional party, is given by (19), the impact of a *marginal* change in j's position has on this probability is

$$\frac{d\rho_{ijk}(\mathbf{z}_{Nat})}{dz_j}|_{\mathbf{z}_{-j}} = 2\beta_k \rho_{ijk}(1 - \rho_{ijk})(x_i - z_j) \tag{22}$$

where $\mathbf{z}_{-j}$ indicates that we are holding the positions of all parties but j fixed. The effect that j's change in position has on the probability that i votes for j in region k depends on the weight given to the policy differences with parties in region k, β_k; on how likely is i to vote for j, ρ_{ijk}, and to vote for any other party, $(1 - \rho_{ijk})$; and on how far apart i's ideal policy is from j's, $(x_i - z_j)$.

If both national and regional parties were concerned only with maximizing their votes in region k, then from (20), party j adjusts its position to maximize its expected vote share in region k, that is, party j's first order condition in region k is

$$\frac{dV_{jk}(\mathbf{z}_{Nat})}{dz_j}|_{\mathbf{z}_{-j}} = \frac{1}{n_k} \sum_{i \in N_k} \frac{d\rho_{ijk}}{dz_j} = \frac{1}{n_k} \sum_{i \in N_k} 2\beta_k \rho_{ijk}(1 - \rho_{ijk})(x_i - z_j) = 0 \tag{23}$$

where the third term follows after substituting in (22). The FOC for party j in (23) is satisfied when

$$\sum_{i \in N_k} \rho_{ijk}(1 - \rho_{ijk})(x_i - z_{jk}) = 0.$$

Solving this equation for z_j gives the *candidate (C)* for party j's vote maximizing policy in region k, regardless of whether j is a national or a regional party, as

$$z_{jk}^C = \sum_{i \in N_k} \mu_{ijk} x_i \quad \text{where} \quad \mu_{ijk} \equiv \frac{\rho_{ijk}(1 - \rho_{ijk})}{\sum_{i \in N_k} \rho_{ijk}(1 - \rho_{ijk})}. \tag{24}$$

The term μ_{ijk} represents the weight that party j gives to voter i in region k when choosing its candidate vote maximizing policy in region k. This weight depends on how likely is i to vote for j, ρ_{ijk}, and to vote for any other party, $(1-\rho_{ijk})$, in region k relative to all other voters in region k.[30]

Note that μ_{ijk} is endogenously determined in the model as this weight depends on how likely are all voters to vote for j in region k which in turn depend on the positions of all parties in region k. Note also that μ_{ijk} may be non-monotonic in ρ_{ijk}. To see this exclude voter i from the denominator of μ_{ijk}. When $\sum_{v\in N_k - i} \rho_{vjk}(1-\rho_{vjk}) < \dfrac{2}{3}$ then $\mu_{ijk}(\rho_{ijk}=0) = \mu_{ijk}(\rho_{ijk}=1) = 0 < \mu_{ijk}(\rho_{ijk} = \frac{1}{2})$. So that when j's vote share in region k is low enough (excluding voter i) and i votes for sure for j (a core supporter), i gets a lower weight (in fact zero weight) in j's candidate position than a voter who will only vote for j with probability $\frac{1}{2}$ (an "undecided" voter) and gets the same weight than a voter who will not vote for j, $\rho_{ijk} = 0$. In this case, j caters to "undecided" voters in region k by giving them a higher weight in j's policy in region k than the weight j gives a core supporter. This will be the most frequent case. When $\sum_{v\in N_k - i} \rho_{vjk}(1-\rho_{vjk}) > \dfrac{2}{3}$, μ_{ijk} increases in ρ_{ijk}. When party j in region k expects a large enough vote share (excluding voter i), it gives a core supporter a higher weight in its position than it gives other voters as there is no risk of doing so.

6.2 Second Order Conditions SOC

Thus, (24) says that if the only concern of a party, regional or national, competing in region k is the voters in region k, then its candidate equilibrium position is a weighted average of the ideal policies of the voters in region k where voter i's ideal is weighted by μ_{ijk}.

National parties face, however, a different problem as they must consider their election prospects in all regions of the country. Given the vector of national policy positions, $\mathbf{z}_{Nat}$, from (21) national party j adjusts its position so as to maximize its expected *national* vote share, so that,Section 6.1.1 gives the candidate position for national and regional parties. We now need to determine whether at these critical points parties are maximizing their vote shares.

To find whether the candidate position z_{jk}^C (correspondingly, z_j^C) is a local maximum of regional (correspondingly national) party j's vote share function, so that $\mathbf{z}_{Nat}^C$ is a LNE of the game, we need to examine whether the second order condition of the party's vote share function, that is, the corresponding Hessians of the second derivatives of the parties' vote share function is negative definite. Since regional and national parties face different electorates the analysis must consider whether the party is a regional or a national party and the region in which the parties compete.

To find the Hessian of party j's vote share function in (20) for parties competing in region k, we need the second derivative of the probability that i votes for j in region k in (19), which we find by taking the derivative of (22), i.e.,

$$\frac{d^2\rho_{ijk}}{dz_j^2}|_{\mathbf{z}_{-j}} = 2\rho_{ijk}(1-\rho_{ijk})\left[2(1-2\rho_{ijk})\nabla_{ijk}(z_j^C) - \beta_k I\right]. \tag{25}$$

Here $\nabla_{ijk}(z_j^C) = \beta_k (x_i - z_j^C)^{\mathbf{T}}(x_i - z_j^C)\beta_k$ where superscript $\mathbf{T}$ is used to denote a column vector and I is the w-identity vector. Note that $\nabla_{ijk}(z_j^C)$ is a $w \times w$ matrix of cross product terms giving an "overall" measure of how far voter i is from party j in the policy space where each dimension is weighed by β_k. In addition, note that if we average over all voters in region k, we get the variance

[30] For example if all voters in region k are equally likely to vote for j, say with probability p, then the weight party j in region k gives to voter i in its candidate vote maximizing policy is $\mu_{ijk} = \frac{1}{n_k}$, i.e., j gives all voters the same weight in its policy position.

of region k's electorate around j's candidate policy position,

$$\nabla^C_{jk}(z^C_j) \equiv \frac{1}{n_k}\sum\nolimits_{i\in N_k} \nabla_{ijk}(z^C_j). \tag{26}$$

From (21) and taking the derivative of (23) with respect to z_{jk}, the $w \times w$ Hessian of second order derivatives for party j in region k located at z^C_j, $\mathcal{H}_{jk}$, is then

$$\begin{aligned}\mathcal{H}_{jk} &\equiv \mathcal{H}_{jk}(\mathbf{z}^C_k) = \frac{d^2V_{jk}(\mathbf{z}^C_k)}{dz_j^2}|_{z^C_{-j}} = \frac{1}{n_k}\sum\nolimits_{i\in N_k}\frac{d^2\rho_{ijk}}{dz_j^2} \\ &= \sum\nolimits_{i\in N_k} 2\rho_{ijk}(1-\rho_{ijk})\left[2(1-2\rho_{ijk})\nabla^C_{ijk}(z^C_j) - \frac{1}{n_k}\beta_k I\right] \end{aligned}\tag{27}$$

where the last line follows after substituting in (25) and (26).

When all the w eigenvalues of the Hessian H_{jk} of regional party j in region k are negative at z^C_{jk}, then at this critical point j's vote share in region k is at a maximum. If, however, all the w eigenvalues of the Hessian H_{jk} are positive at z^C_{jk}, then at this critical point j's vote share is at a minimum thus giving j an incentive to move away from z^C_{jk}, i.e., to locate at a different point, to increase its vote share. If the Hessian H_{jk} has both positive and negative eigenvalues at z^C_{jk} then at this critical point j's vote share is at a saddle point and j should move to increase its votes. Therefore to find whether j's vote share function is at maximum we need to determine the conditions under which the Hessian H_{jk} will be negative definite as when these conditions are not met, j will move to increase its votes.

Recall that the trace of the Hessian is equal to the sum of the eigenvalues associated with H_{jk} and is also given by the sum of the main diagonal elements of H_{jk}. In order for z^C_{jk} to be a local maximum of j's vote share function in region k, the eigenvalues of the Hessian of party j have to be all negative, implying that the trace of H_{jk} must then also be negative. Thus, the Hessian H_{jk} at z^C_{jk} is negative definite when the trace of H_{jk} is negative.

To find the trace of H_{jk}, we need the diagonal elements of H_{jk}. Let ω represent a dimension in the policy space X. For parties and voters in region k, denote by $z^C_j(\omega)$ and $x_i(\omega)$ the position of party j and the ideal policy of voter i in dimension ω respectively. Using (27), the second order condition for party j along the ω dimension is given by the (ω,ω) element of $\mathcal{H}_{jk}$. The (ω,ω) element of $\mathcal{H}_{jk}$ has the following form

$$\begin{aligned}&\sum\nolimits_{i\in N_k} 2\rho_{ijk}(1-\rho_{ijk})\left\{2(1-2\rho_{ijk})\frac{1}{n_k}\beta_k\left[x_i(\omega)-z^C_{jk}(\omega)\right]^2\beta_k - \frac{1}{n_k}\beta_k\right\} \\ =\ & \frac{1}{n_k}\sum\nolimits_{i\in N_k} 2\beta_k\rho_{ijk}(1-\rho_{ijk})\left\{2(1-2\rho_{ijk})\beta_k\left[x_i(\omega)-z^C_{jk}(\omega)\right]^2 - 1\right\}. \end{aligned}\tag{28}$$

Suppose that in the above equation we only had $\frac{1}{n_k}\sum_{i\in N_k}\beta_k\left[x_i(\omega)-z^C_{jk}(\omega)\right]^2$, this expression gives the weighted variance of voters ideal points from j's candidate position in the ω dimension, that is, how dispersed voters in region k are from j's candidate position in the ω dimension.

To obtain the trace of the Hessian of party j in region k, $trace[\mathcal{H}_{jk}]$, just add the diagonal elements given in (28) over the w dimensions of the policy space to obtain

$$\begin{aligned} trace[\mathcal{H}_{jk}] &\equiv \sum\nolimits_{\omega=1}^{w}\frac{1}{n_k}\sum\nolimits_{i\in N_k} 2\beta_k\rho_{ijk}(1-\rho_{ijk})\left\{2(1-2\rho_{ijk})\beta_k\left[x_i(\omega)-z^C_{jk}(\omega)\right]^2-1\right\} \\ &= \frac{1}{n_k}\sum\nolimits_{\omega=1}^{w}\sum\nolimits_{i\in N_k} 2\beta_k\rho_{ijk}(1-\rho_{ijk})\left\{2(1-2\rho_{ijk})\beta_k\left[x_i(\omega)-z^C_{jk}(\omega)\right]^2-1\right\} \end{aligned}\tag{29}$$

Party j in region k will be at a maximum at z_{jk}^C if all eigenvalues are negative, i.e., if $trace[\mathcal{H}_{jk}] < 0$. The *necessary* condition for the Hessian to be negative definite at z_{jk}^C (i.e., $trace[\mathcal{H}_{jk}] < 0$) is then

$$\sum\nolimits_{\omega=1}^{w} \sum\nolimits_{i \in N_k} \beta_k \rho_{ijk}(1-\rho_{ijk}) \left\{ 2(1-2\rho_{ijk})\beta_k \left[x_i(\omega) - z_{jk}^C(\omega)\right]^2 - 1 \right\} \quad < \quad 0$$
$$\text{or} \quad \sum\nolimits_{\omega=1}^{w} \sum\nolimits_{i \in N_k} \mu_{ijk} 2(1-2\rho_{ijk})\beta_k [x_i(\omega) - z_{jk}^C(\omega)]^2 \quad < \quad w. \tag{30}$$

Define party j's *convergence coefficient in region* k as the LHS of (30), i.e.,

$$c_{jk}(\mathbf{z}_k^C) \equiv \sum\nolimits_{\omega=1}^{w} \sum\nolimits_{i \in N_k} \mu_{ijk} 2(1-2\rho_{ijk})\beta_k [x_i(\omega) - z_{jk}^C(\omega)]^2. \tag{31}$$

This is the convergence coefficient of party j in region k given in (6).

The following proof is summarized in Result 1 in Section 2.0.1. If when competing in region k, party j locates at its critical point, z_{jk}^C, then the trace of the Hessian of its vote share function, $trace[\mathcal{H}_{jk}]$, will be negative only if the value of the convergence coefficient of party j in region k is less than the dimension of the policy space, w. Thus, the *necessary* condition for party j in region k to converge to or remain at z_{jk}^C to maximize its vote share is that

$$c_{jk}(\mathbf{z}_k^C) < w \tag{32}$$

On the other hand, if $c_{jk}(\mathbf{z}_{Nat}^C) \geq w$, then the necessary condition for convergence has not been met. In this case, party j's vote share function in region k is at a minimum or at a saddle point.

Let us now examine the necessary conditions for *national* party j located at its critical point z_j^C to be at a maximum of its vote share function. To do so we look at the Hessian of second derivatives of j's vote share function in (21) by taking the derivative of (**??**) to obtain

$$\begin{aligned} \mathcal{H}_j \quad &\equiv \quad \mathcal{H}_j(\mathbf{z}_{Nat}^C) = \frac{d^2 V_j(\mathbf{z}_{Nat}^C)}{dz_j^2}|_{z_{-j}^C} = \sum\nolimits_{k \in \Re} \frac{n_k}{n} \frac{d^2 V_{jk}(\mathbf{z}_k^C)}{dz_j^2} = \sum\nolimits_{k \in \Re} \frac{n_k}{n} \mathcal{H}_{jk} \\ &= \quad \sum\nolimits_{k \in \Re} \frac{n_k}{n} \sum\nolimits_{i \in N_k} 2\rho_{ijk}(1-\rho_{ijk}) \left[2(1-2\rho_{ijk}) \nabla_{ijk}^C(z_j^C) - \frac{1}{n_k}\beta_k I \right] \end{aligned} \tag{33}$$

where the last term follows from (27) after substituting in (25) and (26).

Like in the regional case, if all the w eigenvalues of the Hessian $\mathcal{H}_j$ of national party j are negative at z_j^C, then at this critical point j is maximizing its national vote share. If, however, all the w eigenvalues of the Hessian $\mathcal{H}_j$ are positive at z_j^C, then j is minimizing of its national vote share and j will move away from z_j^C to increase its votes. If the Hessian $\mathcal{H}_j$ has both positive and negative eigenvalues at z_j^C, then j is at a saddle point of its national vote share and should move to increase votes. That is, we need the conditions on the Hessian that determine whether j's national vote share function is at maximum as when these conditions are not satisfied, j will move to increase its votes.

As in the regional case, we use the trace of the Hessian of national party j, $trace[H_j]$, to determine whether j's Hessian, H_j, is negative definite. The main diagonal element of $\mathcal{H}_j$ along the ω dimension is the (ω, ω) element of $\mathcal{H}_j$. From (33), the (ω, ω) element of $\mathcal{H}_j$ is

$$\begin{aligned} &\sum\nolimits_{k \in \Re} \frac{n_k}{n} \sum\nolimits_{i \in N_k} 2\rho_{ijk}(1-\rho_{ijk}) \left\{ 2(1-2\rho_{ijk}) \frac{1}{n_k} \beta_k \left[x_i(\omega) - z_j^C(\omega)\right]^2 \beta_k - \frac{1}{n_k}\beta_k \right\} \\ = \quad &\frac{1}{n} \sum\nolimits_{k \in \Re} \sum\nolimits_{i \in N_k} 2\beta_k \rho_{ijk}(1-\rho_{ijk}) \left\{ 2(1-2\rho_{ijk})\beta_k \left[x_i(\omega) - z_j^C(\omega)\right]^2 - 1 \right\} \end{aligned} \tag{34}$$

The $trace[\mathcal{H}_j]$ is just the sum of the diagonal elements given in (34) over the w dimensions, i.e.,

$$\begin{aligned} trace[\mathcal{H}_j] &\equiv \sum_{\omega=1}^{w}\sum_{k\in\Re}\frac{n_k}{n}\frac{1}{n_k}\sum_{i\in N_k} 2\beta_k\rho_{ijk}(1-\rho_{ijk})\left\{2(1-2\rho_{ijk})\beta_k\left[x_i(\omega)-z_j^C(\omega)\right]^2-1\right\} \\ &= \frac{1}{n}\sum_{\omega=1}^{w}\sum_{k\in\Re}\sum_{i\in N_k} 2\beta_k\rho_{ijk}(1-\rho_{ijk})\left\{2(1-2\rho_{ijk})\beta_k\left[x_i(\omega)-z_j^C(\omega)\right]^2-1\right\} \end{aligned}$$

Therefore, national party j will be at a maximum at z_j^C if $trace(\mathcal{H}_j) < 0$, implying that

$$\begin{aligned} \sum_{\omega=1}^{w}\sum_{k\in\Re}\sum_{i\in N_k} \beta_k\rho_{ijk}(1-\rho_{ijk})\left\{2(1-2\rho_{ijk})\beta_k\left[x_i(\omega)-z_j^C(\omega)\right]^2-1\right\} &< 0 \\ \text{or} \quad \sum_{\omega=1}^{w}\sum_{k\in\Re}\sum_{i\in N_k} \delta_{ijk}2(1-2\rho_{ijk})\beta_k\left[x_i(\omega)-z_j^C(\omega)\right]^2 &< w \end{aligned}$$

where after some manipulation the last line follows from (**??**).

Then using (**??**), we get that the trace of $\mathcal{H}_j$ will be negative, $trace(\mathcal{H}_j) < 0$, if

$$\sum_{\omega=1}^{w}\sum_{k\in\Re}\sum_{i\in N_k} \theta_{jk}\mu_{ijk}2(1-2\rho_{ijk})\beta_k\left[x_i(\omega)-z_j^C(\omega)\right]^2 < w \tag{35}$$

Define national party j's *national convergence coefficient* as the LHS of (35), i.e.,

$$c_j(\mathbf{z}_{Nat}^C) \equiv \sum_{\omega=1}^{w}\sum_{k\in\Re}\sum_{i\in N_k} \theta_{jk}\mu_{ijk}2(1-2\rho_{ijk})\beta_k\left[x_i(\omega)-z_j^C(\omega)\right]^2$$

This is the national convergence coefficient of national party j given in (10).

National party j will locate at z_j^C if it is maximizing its vote share, that is, only if

$$c_j(\mathbf{z}_{Nat}^C) < w \tag{36}$$

This is given in (9). Thus, the *necessary* condition for national party j to remain at the candidate position is that its national convergence coefficient be less that the dimension of the policy space, w. This is the convergence condition for national party j given in Result 2 in Section 2.0.1.

7 Empirical Appendix : Election Statistics

Table A1: Descriptive Statistics of the survey sample by Region

	Canada ($n = 862$)			**C/Q** ($n = 675$)			**Québec** ($n = 187$)		
Variable	**Mean**	**Median**	**SD**	**Mean**	**Median**	**SD**	**Mean**	**Median**	**SD**
Social	0.00	-0.22	1.67	0.31	0.12	1.64	-1.11	-1.09	1.22
Decentralization	0.00	0.00	1.07	0.02	0.03	1.09	-0.08	-0.11	0.99
Age	50.41	50	15.84	50.77	50.00	15.55	49.08	48	16.82
Female	0.51	1	0.50	0.51	1	0.50	0.51	1	0.50
Education	7.164	7	2.10	7.14	7	2.10	7.26	7	2.10

TableA2: Descriptive Statistics of the Survey Sample by Party and Region						
Variable	Mean	Median	SD	Mean	Median	SD
	Liberals			**Conservatives**		
	Canada ($n = 296$)			Canada ($n = 272$)		
Social	-0.17	-0.35	1.29	1.27	1.18	1.54
Decen	-0.38	-0.46	1.04	0.32	0.33	1.04
Age	53.07	53	15.21	50.91	50	16.20
Female	0.53	1	0.5	0.45	0	0.50
Educ	7.33	8	2.10	6.87	7	2.10
	Canada outside Québec ($n = 249$)			Canada outside Québec ($n = 255$)		
Social	-0.10	-0.18	1.323	1.36	1.38	1.52
Decen	-0.34	-0.45	1.06	0.38	0.39	1.03
Age	52.65	53	14.45	51.15	50	16.06
Female	0.52	1	0.50	0.46	0	0.5
Educ	7.40	8	2.06	6.81	7	2.09
	Québec ($n = 47$)			Québec ($n = 17$)		
Social	-0.54	-0.83	1.10	-0.06	-0.21	1.012
Decen	-0.62	-0.64	0.86	-0.61	-0.70	0.67
Age	55.76	58	18.73	47.35	44	18.43
Female	0.60	1	0.50	0.41	0	0.51
Educ	6.99	7	2.34	7.71	9	1.96
	New Democratic			**Greens**		
	Canada ($n = 159$)			Canada ($n = 32$)		
Social	-0.78	-0.84	1.34	-0.63	-0.93	1.41
Decen	0.05	0.02	1.05	-0.13	-0.19	0.93
Age	47.52	47	15.70	44.94	43	14.20
Female	0.57	1	0.50	0.44	0	0.50
Educ	7.28	7	2.09	7.06	7	2.09
	Canada outside Québec ($n = 144$)			Canada outside Québec ($n = 27$)		
Social	-0.70	-0.83	1.34	-0.48	-0.81	1.41
Decen	-0.03	-0.02	1.05	-0.04	0.03	0.83
Age	48.01	48	16.20	44.56	43	13.85
Female	0.58	1	0.50	0.48	0	0.51
Educ	7.29	7	2.12	7	7	2.09
	Québec ($n = 15$)			Québec ($n = 5$)		
Social	-1.51	-1.23	1.23	-1.40	-1.11	1.26
Decen	0.26	0.46	0.99	-0.55	-0.40	1.43
Age	42.87	43	8.91	47	44	17.64
Female	0.47	0	0.52	0.20	0	0.45
Educ	7.2	7	1.74	7.4	8	2.30

Bloc Québécois (in Québec only, $n = 103$)			
	Mean	Median	SD
Social	-1.48	-1.56	1.10
Decentralization	0.23	0.11	0.92
Age	47.32	47	15.85
Female	0.51	1	0.50
Education	7.31	7	2.06

Table A3: EKOS[a] June 21-24, 2004 Poll (Percentages out of 5,254 respondents)[b]							
Parties[c]	BC	Alberta	Praries	Ontario	Québec	Atlantic	National
LP	30	23	29	38	28	39	32.6
CP	34	58	37	35	11	33	31.8
NDP	27	12	30	21	7	28	19.0
BQ					51		11.2
GPC	7	7	5	5	3	0	4.9

[a] Reflects gender and age composition of the Canadian census population by region (http://www.ekos.com/admin/articles/26June2004BackgroundDoc.pdf).

[b] BC=British Columbia, Praries=Saskatchewan and Manitoba, Atlantic= New Brunswick, Prince Edward Island„ Nova Scotia, and Newfoundland and Labrador

[c] CP= Conservatives, LP= Liberals, NDP=New Democratic, BQ=Bloc Québécois, GPC= Greens.

Table A4: Actual and sample vote shares				
	Actual		**Sample Vote share (%)**	
	Vote %	Seat (%)	All	Québec
Liberal	36.71	135 (44)	34.34	25.13
Conservative	29.66	99 (32)	31.55	9.01
NDP	15.65	19 (6)	18.45	8.02
BQ	12.42	54 (18)	11.95	55.08
Green	4.29		3.71	2.68
Ind.	0.5	1 (0.3)		
Total	99.2	308	100	100

Table A5: Provincial Votes (%) and Seats in the 2004 Canadian Election										
Region	Western Provinces								Ontario	
Provinces[a]	BC		AB		SK		MB		ON	
Party[b]	Vote	Seats	Vote	Seats	Vote	Seats	Vote	Seats	Vote	Seats
CP	36.3	22	61.7	26	41.8	13	39.1	7	31.5	24
LP	28.6	8	22.0	2	27.2	1	33.2	3	44.7	75
BQ										
NDP	26.6	5	9.5		23.4		23.5	4	18.1	7
GPC	6.3		6.1		2.7		2.7		4.4	
Ind	0.3	1			4.6				0.3	
Total[c]	98.1	36	99.3	28	99.7	14	98.5	14	99.0	106
Region	Québec		Atlantic Provinces							
Provinces[a]	QC		NB		NS		PEI		NL	
Party[b]	Vote	Seats	Vote	Seats	Vote	Seats	Vote	Seats	Vote	Seats
CP	8.8		31.1	2	28.0	3	30.7	0	32.3	2
LP	33.9	21	44.6	7	39.7	6	52.5	4	48.0	5
BQ	48.9	54								
NDP	4.6		20.6	1	28.4	2	12.5		17.5	
GP	3.2		3.4		3.3		4.2		1.6	
Ind	0.1		0.2		0.1				0.6	
Total	99.4	75	99.9	10	99.5	11	99.9	4	100	7

[a] BC= British Columbia, AB= Alberta, SK=Saskatchewan, MB = Manitoba, ON= Ontario, QC = Québec, NB = New Brunswick, NS= Nova Scotia, PEI = Prince Edward Island, NL = Newfoundland and Labrador.

[b] CP= Conservatives, LP= Liberals, BQ=Bloc Québécois, NDP=New Democratic, GP= Greens, Ind=independent.

8 Tables and Figures to be inserted

Table 1: Survey Questions taken from Blais *et al.* (2006) for the 2004 Canadian election	
Inequality	How much to you think should be done to reduce the gap between the rich and the poor in Canada? (1) much more - (5) much less
Women	How much do you think should be done for women? (1) much more - (5) much less
Gun only police/military	Only the police and the military should be allowed to have guns. (1) strongly agree - (4) strongly disagree
Iraq War	As you may know, Canada decided not to participate in the war against Iraq. Do you think this was a good decision or a bad decision? (1) good decision (2) bad decision
Left-Right	In politics, people sometimes talk of left and right. Where would you place yourself on the scale below? (0) left - (11) right
Welfare	The welfare state makes people less willing to look after themselves. (1) strongly disagree - (4) strongly agree
Standard of Living	The government should see to it that everyone has a decent standard of living. (1) leave people behind (2) Don't leave people
Quebec	How much do you think should be done for Quebec? (1) much more - (5) much less
Moving Cross Region	If people can't find work in the region where they live, they should move to where the jobs are? (1) strongly disagree - (4) strongly agree
Federal-provincial	In general, which government looks after your interests better? (1) provincial (2) no difference (3) federal

inserted

Table 2: Factor loadings for Canada

Components	**Social**	**Decentralization**
Inequality	0.36	-0.03
Women	0.35	0.07
Gun only police/military	0.20	0.52
Iraq War	0.30	0.20
Left-Right	0.38	-0.06
Welfare	0.37	-0.17
Standard of Living	0.38	-0.05
Quebec	-0.35	0.00
Moving cross region	0.27	-0.48
Federal-provincial	-0.09	-0.65
SD ($\sqrt{var}$)	1.67	1.07
% Var	28	11
Cumulative % Var	28	39

Table 3: VCL Model for 2004 Canadian Election by Region (baseline=Lib)

	Canada	Outside Québec			Québec		
	Coeff. (conf. int.)[a]	Coeff. (conf. int.)[a]			Coeff. (conf. int)[a]		
β_k	0.256 * (0.22,0.29)	0.267 * (0.23,0.31)			0.231 * (0.15,0.32)		
	New Democratic Party						
λ_{NDP}	-0.593 (-1.96,0.66)	-0.556 * (-0.77,-0.35)			-1.200 * (-1.80,-0.65)		
Age		18-30	30-65	65+	18-30	30-65	65+
<col[c]		-0.325 (-1.13,0.46)	-0.333 * (-0.65,-0.03)	-0.638 * (-1.17,-0.13)	-1.650 * (-3.51,-0.19)	-0.526 (-1.40,0.32)	-3.153 * (-6.77,-0.13)
>col[c]		0.348 (-0.54,1.31)	-1.012 * (-1.42,-0.63)	-0.752 (-1.83,0.18)	-2.169 * (-5.58,-1.16)	-1.228 * (-2.26,-0.29)	-2.387 * (-5.76,-0.45)
	Conservative Party of Canada						
λ_{CPC}	-0.139 (-1.18,0.82)	-0.240 (-0.23,0.18)			-0.316 * (-1.80,-0.65)		
Age		18-30	30-65	65+	18-30	30-65	65+
<col[c]		0.198 (-0.57,0.99)	0.167 (-0.13,0.47)	0.084 (-0.36,0.53)	-0.298 (-1.74,1.04)	-0.314 (-1.40,0.70)	-1.120 (-2.62,0.04)
>col[c]		0.168 (-0.82,1.19)	-0.553 * (-0.96,-0.15)	0.420 (-0.42,1.31)	0.404 (-1.20;2.14)	-2.619 (-1.13,0.87)	-0.236 (-1.68,1.08)
	Green Party of Canada						
λ_{GPC}	-1.775 (-3.29,0.26)	-2.233 * (-2.64,-1.86)			-2.310 * (-3.26,-1.50)		
Age		18-30	30-65	65+	18-30	30-65	65+
<col[c]		-1.757 * (-3.02,-0.56)	-2.029 * (-2.62,-1.47)	-2.805 (-4.17,-1.79)	-2.178 * (-4.07,-0.51)	-2.618 * (-4.43,-1.23)	-2.562 * (-4.40,-1.16)
>col[c]		-3.191 * (-6.67,-1.34)	-2.401 * (-3.10,-1.79)	-3.265 * (-6.59,-1.44)	-2.831 * (-6.04,-0.69)	-2.233 * (-3.74,-1.00)	-2.978 * (-6.37,-0.99)
					Bloc Québécois		
λ_{BQ}	0.278 (-1.36,1.77)				0.649 * (0.28,1.01)		
Age					18-30	30-65	65+
<col[c]					0.975 * (0.10,1.98)	0.979 * (0.36,1.63)	0.258 (-0.54,1.02)
>col[c]					0.577 (-0.74,1.94)	0.607 (-0.03,1.26)	0.263 (-0.99,1.37)
n	862	675			187		
DIC	2029.291						
GR[b]	1.02						

[a] Numbers in brackets are 95% credible intervals
[b] Gelman-Rubin
[c] “<col” = Less than college degree and “>col” = More than college degree
* 95% Credible Interval Does Not Include 0.

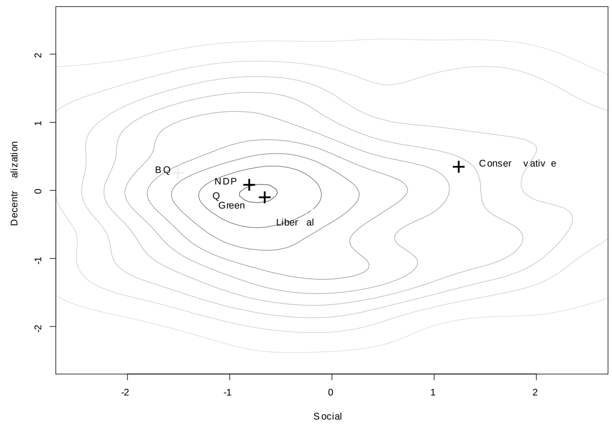

Figure 1: Distribution of voters and party positions for Canada in 2004

Table 4: Weights and Convergence Coefficients for the Mean Vector

	LPC	NDP	CPC	GPC	BQ
θ_{Quebec}	.159	.088	.134	.139	—
	(.151,.165)	(.083,.093)	(.127,.139)	(.134,.145)	—
$\theta_{C/Quebec}$	.841	.911	.866	.861	—
	(.835,.849)	(.906,.917)	(.860,.872)	(.855,.866)	—
Convergence Coefficient	.528	1.176	.627	1.870	-.043
	(.454,.603)	(.992,1.356)	(.537,.719)	(1.579,2.155)	(-.084,-.008)

Table 5: Convergence Coefficient Parts for NDP and Greens

	NDP	GPC
Con.Coef. - Social	.833	1.328
	(.703,.959)	(1.121,1.531)
Con.Coef - Decentralization	.343	.541
	(.289,.396)	(.463,.624)

Table 6: Vote Shares given various policy positions

	Actual	Mean	Optimal
LPC	36.71	34.48	39.47
CPC	29.66	31.85	36.60
NDP	15.65	18.16	10.12
GPC	4.29	3.43	1.21
BQ	12.42	12.07	12.59

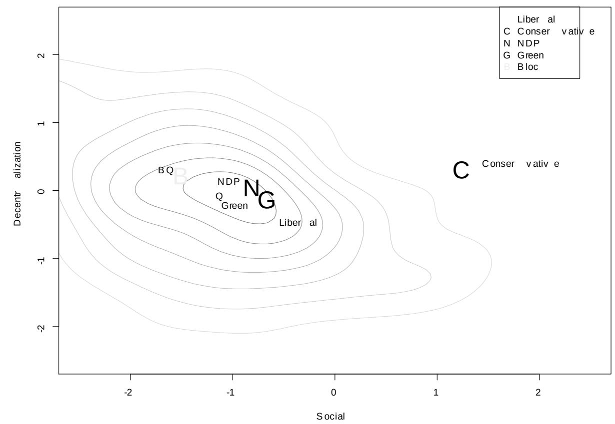

Figure 2: Distribution of voters and party positions for Quebec in 2004

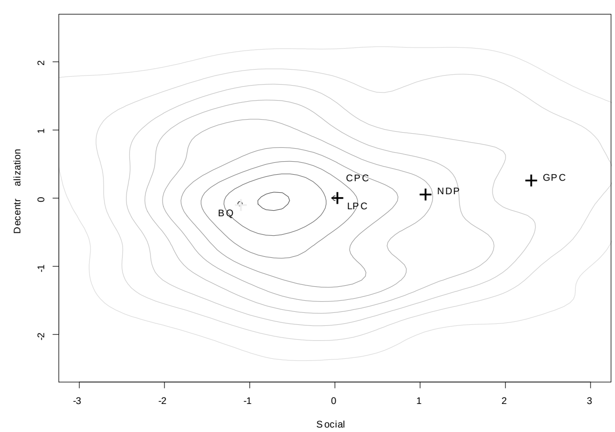

Figure 3: Vote maximizing positions in Canada 2004