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

McAlister, Kevin,< kevinmcalister@wustl.edu, Washington University in Saint Louis
Norman Schofield< schofield.norman@gmail.com,Washington University in Saint Louis

March 4, 2013

Abstract

Formal work on the electoral model often suggests that parties should locate themselves 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 an electoral valence and by allowing countries to have national and regional parties, that is, by allowing voters to face different sets of parties in different regions. We introduce the notion of a convergence coefficient, c for national and regional parties. We show that high values of c imply that there is a significant centrifugal tendency acting on parties. We used an electoral survey for the 2004 election in Canada to construct a stochastic valence model of the election with two regions Québec and the rest of Canada. Using this survey we estimate voter positions in the policy space. We use the variable choice set logit model to built a relationship between party position and vote share. We find that in the local Nash equilibrium for the election the Liberals and Conservatives locate at the national electoral mean and the Bloc Québécois at the Québec electoral mean. The NDP and the Greens, with lower valences, also diverge from the national mean. The divergence of the Bloc, the NDP and the Greens reflects the regional characteristic of the Canadian polity.

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 would certainly vote for a party. The models showed that there exists a Condorcet point at the electoral median.[1] However, when extended into spaces with more than one dimension, these two-party pure-strategy Nash equilibria generally do not exist. While attempts were made to reconcile this difference, the conditions necessary to assure that there is a pure-strategy Nash equilibrium at the electoral median were strong and unrealistic (McKelvey and Schofield;1987; Schofield 1978, 1983; Saari, 1997; Caplin and Nalebuff, 1991).

Instead of pure-strategy Nash equilibria (PNE) there often exist mixed strategy Nash equilibria, which lie in the subset of the policy space called the uncovered set (Kramer, 1978). Many times, this

*Prepared for presentation at the EPSA conference, Barcelona, June 2013, and the political economy panel at the SAET Meeting, Paris, July 2013

[1] These results can be contrasted with the citizen-candidate models of Osborne and Silvinski (1996) and Besley and Coate (1997) which did not predict convergence.

uncovered set 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 theorems and the stability theorems suggests that a model in which the individual vote is stochastic rather deterministic would be most appropriate (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 in line with multiple theories of voter behavior and under some conditions yields the property that "rational" vote maximizing parties will converge to the electoral mean.

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 quality that a candidate exhibits that is independent of the location 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. 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 partly determine how likely they are to vote for the given parties in the region.[2] Both kinds of valence can be important in determining the outcomes of elections 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 a voter's choice is determined by the voter's "utility" which depends on the valence terms and the voter's distance from the party in the policy space but are also affected 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 is a measure of the attraction the electoral mean exerts on parties in order to gain votes in an election. This coefficient is unitless and can be used to compare convergence across models. Low values of the convergence coefficient indicate 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* (Patty, 2005; Patty, 2006). High values indicate the opposite. For a simple single region model,Schofield also obtainrd the necessary and a sufficient conditions that the convergence coefficient must meet for parties to converge to the electoral mean. 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 if $c < w$, where w is the number of dimensions of the policy space of interest.

When the necessary condition fails, at least one party will have a motivation to position itself away from the electoral mean . 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. Thus $c \geq w$ implies that there cannot be a pure strategy vote maximizing Nash equilibrium (PNE) at the electoral center.

This model only answered the convergence question when the local Nash equilibrium was in the simplest case of a single national election. The problem quickly becomes more complicated when we face more complex electoral structures. For example in Canada there are four national parties and one regional party. The national parties aim to appeal to voters across Canada. The Bloc Québécois (BQ) is a regional party that runs only in Québec. Clearly the Bloc will position itself

[2] 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.

to maximize its votes in Québec and has no interest in representing the interests of voters outside Québec. In general, there may exist many reasons why parties may choose to run in some regions and not in others. They may do so for example because of deep political and economic differences with other regions or in response to too much centralization. (Riker (1964, 1987),. Parties may also choose not to run in regions where they anticipate they will do very poorly, specially if they are financially constrained.[3]

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. One would think that 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 as a regional party would converge to the electoral mean in Québec.

Moreover, when there are national and regional parties, we can no longer assume that voter choices have a multinomial logit (MNL) specification because one of its underlying assumptions, the independence of irrelevant alternatives (IIA), no longer holds (Train, 2003).[4] Using Schofield's (2007) model and the associated MNL estimations we can only analyze convergence, valence, and spatial adherence within specific regions with the analysis for each region done independently of other regions. Thus in order to deal with these more complex electorates, such as the one in Canada, we need a more general formal model of the election and a more general empirical model.

In Section 2 we develop a formal stochastic electoral model in which voters in different regions face different party bundles. We extend Schofield's (2007) formal model by allowing for national and regional parties to compete in the election. We derive the first and second order conditions necessary for a vector of party policy positions to be a local Nash equilibrium (LNE). We then define the more general convergence coefficient for each national and regional party and use the model to determine whether a party converges or nor to the a particular candidate position.

In Section 3 we introduce the varying choice set logit (VCL, see Yamamoto, 2011) model and use it to estimate the parameters necessary to find equilibria in the model. This empirical model, an extension of the mixed logit model, assumes that the error terms in voters' choices have a Type-I extreme value distribution, but that do not satisfy the independence of irrelevant alternatives (IIA) assumption so critical in the multinomial logit model. This VCL estimates the parameters at the national level while also allowing regional parameters to be estimated.

Finally, to illustrate the new formal model and the new empirical methods, in Section 4 we analyze 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.[5] 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 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

[3] We intend to apply the model presented here to the case of Britain, where there are of course at least three regions and regional parties, as well as even greater complexity in N. 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.

[4] To see this note that 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. Thus, violating IIA outside Québec.

[5] 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).

gained the support of almost half of Québécers giving it 54 seats in 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.

The VCL model allows us to estimate the parameters at the regional and national levels. We find that in Québec, the BQ is the party with the highest competence valence followed by the Liberals once policy differences are taken into account. Outside Québec, the Liberal and Conservatives were considered by voters to be equally competent at governing. The New Democrats 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 of the anticipated election, we then examined 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, 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. On the other hand, the two smaller national parties, the New Democratic Party (NDP, a left leaning party) and the Greens, diverge from the national mean to increase their vote shares.

The novelty of our analysis is that our formal and empirical models provides insights into how national and regional parties position themselves to maximize vote shares.

2 The Formal Stochastic Model

Countries are usually composed of different regions.[6] In some countries, there are vast political and/or economic differences across regions. Political differences may be due, for example, to cultural differences across regions or to the desire of a particular region for more independence from the national government. (Riker, 1964, 1987). Economic differences can arise, for example, from endowments of natural resources or from previous regional economic development such as that in manufacturing or high tech industries. When substantial political and/or economic differences exist between regions in a country, interest groups and voters in these regions coalesce to create parties that will better represent their interests at the national level.[7] Clearly, 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 recognize that there may exist parties that may have no national scope but that represent the interest of groups and voters across various states or provinces.[8] In the model developed in this paper, regional parties operate only in a single jurisdiction, a province or state. 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 a stochastic electoral model for a country that has at least one national and one regional party. Parties and voters at the national and regional levels are concerned with issues on the same policy space. That is, the policy space is defined broadly enough to include all relevant policy dimensions in the country. We allow voters' and parties' preferences to vary across regions.[9]

[6] Canada is a country with vast regional differences. Québec is by the nature of its culture and laws different from other provinces. Alberta has vast natural resources, such as the oil sands. Ontario is Canada's manufacturing base.

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

[8] Parties with support across various regions may strive to become national players as they grow. Given that we only examine 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 natural resources and the oil sands; some Québécers about preserving their

2.1 The general model

We model elections in a country where voters in different regions face different combinations of national and regional agents,[10] so that voters in different regions 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 national or a regional party and the region in which the parties compete.

Prior to the election, all parties publicly 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 the jurisdiction they seek to represent.

Let z_j represent the position of any party in X. 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 $k \in \Re = 1, .., r$ so that the set of regional parties may vary by region. On occasion we will use z_{jk} to denote the position of party j in region k. In general, we use z_j to denote the position of any party when the context clearly identifies whether we are dealing with a national or a regional party. 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,[11] let $\mathbf{z}_k$ denote the vector of policy positions of the parties 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

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

Let n_k represents 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 set of voters at the national level by $N = \bigcup_{k=1}^{r} N_k$.

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

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

Here, $u^*_{ijk}(x_i, z_j) = \lambda_{jk} + \alpha_{jk} - \beta_k \|x_i - z_j\|^2$ 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 j in region k.

For voters in region k, the sociodemographic aspects of voting 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.) on voting for party j in region k while η_i

French culture and their laws; and Ontarians about policies that affect the manufacturing and service sectors.

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

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

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 exogenous 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, the exogenous valence 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 policy preferences to differ across regions. Differences that in some regions in the past were deep enough to have lead to the emergence of regional parties.

The error term ϵ_{ij} is assumed to be commonly distributed among all voters in the country with the cumulative distribution of the errors following a Type-I extreme value distribution. In the empirical models below the errors also follow a Type-I extreme value distribution, thus making the transition to applying this theoretical 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 parties' policy positions in region k, $\mathbf{z}_k$, the probability that voter i votes for party j in region k is

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

Here 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

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

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. At first glance this looks like the typical multinomial logit (MNL) specification. However, this is not the case in our model since the set of parties varies across regions implying that our model violates the independence of irrelevant alternatives (IIA) assumption of the MNL models (see Section 3). We call this the variable choice set logit (VCL) model. Note that if we only have one region the VCL model has a MNL specification.

Since voters decisions are stochastic in this 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.,

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

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 $k \in \Re$.

In addition, national parties must also 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[12] 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 (4) follows after substituting in (3). Note that due to the presence of regional parties, the sum of the vote share of national parties do not add to 1.

Agents choose their position to maximize their expected vote share given the positions of all other parties. We seek the local Nash equilibria (*LNE*) of the model where each party maximizes its expected vote share taking the position of all the other national and regional parties as given.

A vector of positions, $\mathbf{z}^*_{Nat} \equiv \bigcup_{k=1}^{r} \mathbf{z}^*_k$, is said the be a LNE if $\forall j \in P_{Nat} \cup \bigcup_{k=1}^{r} P_k$, z^*_j is a critical point of the vote share function of party j. In order for these critical values to be a maximum, the Hessian matrix of second derivatives of these vote share functions must be definitely negative, meaning that the corresponding eigenvalues of these matrices must be all negative.

More simply put, a vector, $\mathbf{z}^*_{Nat} \equiv \bigcup_{k=1}^{r} \mathbf{z}^*_k$, is a 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 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 $k \in R$, 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* locally maximizes $V_j(\mathbf{z}_{Nat})$ and $V_{hk}(\mathbf{z}_k)$. Using the estimated coefficients of the VCL 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.

2.2 Equilibrium positions

Parties' positions are LNE at $\mathbf{z}_{Nat} \equiv \bigcup_{k=1}^{r} \mathbf{z}_k$, iff all agents are maximizing their vote share functions at $\mathbf{z}_{Nat}$. Thus, for parties to be at a local maximum at $\mathbf{z}_{Nat}$ two conditions must be satisfied. The

[12] 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.

first derivative of the vote function of party at $\mathbf{z}_{Nat}$ must be zero and the Hessian matrix of second derivatives of the vote share function of each party at $\mathbf{z}_{Nat}$ must be negative definite which happens only when the eigenvalues of the Hessian are all negative at $\mathbf{z}_{Nat}$.

Given that national and regional parties face different electorates, their decisions must be studied separately. In addition, the region in which the parties competes 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 (2), the impact of a *marginal* change in j's position on this probability is then

$$\frac{d\rho_{ijk}(\mathbf{z}_{Nat})}{dz_j}|_{\mathbf{z}_{-j}} = 2\beta_k \rho_{ijk}(1-\rho_{ijk})(x_i - z_j) \tag{5}$$

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 in region k for j, ρ_{ijk}, and 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 (3), 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

$$\frac{dV_{jk}(\mathbf{z}_{Nat})}{dz_j}|_{z_{-j}} = \frac{1}{n_k}\sum\nolimits_{i\in N_k}\frac{d\rho_{ijk}}{dz_j} = \frac{1}{n_k}\sum\nolimits_{i\in N_k} 2\beta_k\rho_{ijk}(1-\rho_{ijk})(x_i - z_j) = 0 \tag{6}$$

where the third term follows after substituting in (5). The FOC for party j in (6) is satisfied when

$$\sum\nolimits_{i\in N_k} \rho_{ijk}(1-\rho_{ijk})(x_i - z_j) = 0$$

so that the *candidate* for party j's vote maximizing policy in region k regardless of whether j is a national or a regional party is

$$z_{jk}^C = \sum\nolimits_{i\in N_k} \mu_{ijk} x_i \quad \text{where} \quad \mu_{ijk} \equiv \frac{\rho_{ijk}(1-\rho_{ijk})}{\sum_{i\in N_k}\rho_{ijk}(1-\rho_{ijk})} \tag{7}$$

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 for any other party, $(1-\rho_{ijk})$, in region k relative to all other voters in region k.[13] Note 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 $\mu_{ijk}(\rho_{ijk}=0) < \mu_{ijk}(\rho_{ijk}=1) < \mu_{ijk}(\rho_{ijk}=\frac{1}{2})$. So that when j's vote share in region k is low enough (excluding voter i), when i votes for sure for j, i receives a lower weight in j's candidate position than a voter who will only vote for j with probability $\frac{1}{2}$ (an "undecided" voter). In this case, j caters to "undecided" voters in region k by giving them a higher weight in j's policy in region k. When $\sum_{v\in N_k - i}\rho_{vjk}(1-\rho_{vjk}) > \dfrac{2}{3}$, then μ_{ijk} increases in ρ_{ijk}. Thus, when party j in region k expects a large enough vote share (excluding voter i), it gives a core supporter (a voter who votes for sure for j) a higher weight than other voters in its position as there is no risk of doing so.

[13] 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 $\mu_{ijk} = \frac{1}{n_k}$.

Thus, (7) says that if the only concern of a party, national or regional, 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 (4) national party j adjusts its position so as to maximize its expected *national* vote share, so that,

$$\frac{dV_j(\mathbf{z}_{Nat})}{dz_j}|_{z_{-j}} = \sum\nolimits_{k\in\Re} \frac{n_k}{n}\frac{dV_{jk}(\mathbf{z}_k)}{dz_j}$$

Thus, when maximizing its expected vote share, national party j takes into account how changes its position affects its votes in each region of the country. After substituting (6) into the above equation, the first order condition for national party j gives

$$\begin{aligned} \frac{dV_j(\mathbf{z}_{Nat})}{dz_j}|_{z_{-j}} &= \sum\nolimits_{k\in\Re} \frac{n_k}{n}\frac{dV_{jk}(\mathbf{z}_k)}{dz_j} = \frac{1}{n}\sum\nolimits_{k\in\Re}\sum\nolimits_{i\in N_k}\frac{d\rho_{ijk}}{dz_j} \\ &= \frac{1}{n}\sum\nolimits_{k\in\Re}\sum\nolimits_{i\in N_k} 2\beta_k\rho_{ijk}(1-\rho_{ijk})(x_i - z_j) = 0 \end{aligned} \tag{8}$$

The FOC for national party j in (8) is satisfied when

$$\sum\nolimits_{k\in\Re}\sum\nolimits_{i\in N_k} \beta_k\rho_{ijk}(1-\rho_{ijk})(x_i - z_j) = 0$$

so that national party j's candidate for vote maximizing policy is

$$z_j^C = \sum\nolimits_{k\in\Re}\sum\nolimits_{i\in N_k}\delta_{ijk}x_i \quad \text{where} \quad \delta_{ijk} \equiv \frac{\beta_k\rho_{ijk}(1-\rho_{ijk})}{\sum_{k\in\Re}\sum_{i\in N_k}\beta_k\rho_{ijk}(1-\rho_{ijk})} \tag{9}$$

Note that δ_{ijk} can be expressed as a function of μ_{ijk} in (7) as follows. Multiply the numerator and denominator of μ_{ijk} by β_k and by $\sum_{k\in\Re}\sum_{i\in N_k}\beta_k\rho_{ijk}(1-\rho_{ijk})$, and re-arrange terms to get

$$\mu_{ijk} \equiv \frac{\sum_{k\in\Re}\sum_{i\in N_k}\beta_k\rho_{ijk}(1-\rho_{ijk})}{\sum_{i\in N_k}\beta_k\rho_{ijk}(1-\rho_{ijk})}\frac{\beta_k\rho_{ijk}(1-\rho_{ijk})}{\sum_{k\in\Re}\sum_{i\in N_k}\beta_k\rho_{ijk}(1-\rho_{ijk})}.$$

So that using δ_{ijk} in (9), we have

$$\mu_{ijk} \equiv \frac{\sum_{k\in\Re}\sum_{i\in N_k}\beta_k\rho_{ijk}(1-\rho_{ijk})}{\sum_{i\in N_k}\beta_k\rho_{ijk}(1-\rho_{ijk})}\delta_{ijk}.$$

We can re-write this expression as

$$\delta_{ijk} \equiv \theta_{jk}\mu_{ijk} \quad \text{where} \quad \theta_{jk} \equiv \frac{\sum_{i\in N_k}\beta_k\rho_{ijk}(1-\rho_{ijk})}{\sum_{k\in\Re}\sum_{i\in N_k}\beta_k\rho_{ijk}(1-\rho_{ijk})} \tag{10}$$

Here θ_{jk} is the weight given by national party j to region k relative to all regions in the country. The term θ_{jk} is the weighted (β_k) aggregate probability that voters in region k vote for j or for any other party relative to all regions. Similar to μ_{ijk}, θ_{jk} may be non-monotonic in $\sum_{i\in N_k}\beta_k\rho_{ijk}(1-\rho_{ijk})$. To see this, let $C_{jk} = \sum_{i\in N_k}\beta_k\rho_{ijk}(1-\rho_{ijk})$ and exclude region k from the denominator of θ_{jk}. When $\sum_{q\in\Re-k} C_{jq} < \frac{2}{3}$ then $\theta_{jk}(C_{jk}=0) < \theta_{jk}(C_{jk}=1) < \theta_{jk}(C_{jk}=\frac{1}{2})$. In this case, a region who will vote for sure for j will receive a lower θ_{jk} weight in j's candidate position than a region who will only vote for j with probability $\frac{1}{2}$. That is, suppose j's expected vote share (excluding region k) is low enough. Then if region k is "undecided" ($\theta_{jk} = \frac{1}{2}$) national party j caters to voters in region k by giving region k a higher weight in its policy position than if voters in region k were to vote for sure for j. On the other hand, when $\sum_{q \in \Re - k} C_{jq} > \dfrac{2}{3}$ then θ_{jk} increases in C_{jk}. That is, if national party j's expected vote share (excluding region k) is high enough then j can give region k a greater θ_{jk} weight as there is no risk to doing so.

Using (9) and (10), the candidate for the vote maximizing position of national party j, z_j^C is

$$z_j^C = \sum_{k \in \Re} \sum_{i \in N_k} \theta_{jk} \mu_{ijk} x_i = \sum_{k \in \Re} \theta_{jk} \sum_{i \in N_k} \mu_{ijk} x_i \tag{11}$$

After substituting z_{jk}^C from (7), national party j's candidate vote maximizing position is a weighted average of the candidate positions it would adopt in each region where it only competing in that region, i.e.,

$$z_j^C = \sum_{k \in \Re} \theta_{jk} z_{jk}^C. \tag{12}$$

Let $\mathbf{z}_{Nat}^C = (z_1^C, ..., z_p^C, z_{11}^C, ..., z_{1q_1}^C, ..., z_{1q_r}^C, .., z_{rq_r}^C)$ denote the *candidate* for the LNE of the game, so that $\mathbf{z}_{Nat}^C$ satisfies the first order conditions for national and regional parties given in (8) and (6) in all regions of the country. Let $\mathbf{z}_k^C$ denote the vector $\mathbf{z}_{Nat}^C$ restricted to region k.

To find whether 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 find whether the corresponding Hessian of the second derivatives of the vote share of party j is negative definate. To do so the analysis must consider the region in which the parties compete.

To find the Hessian of party j's vote share function in (3) for parties competing in region k, we need the second derivative of the probability that i votes for j in region k in (2), which we find taking the derivative of (5), i.e.,

$$\frac{d^2 \rho_{ijk}}{dz_j^2} = 2\rho_{ijk}(1 - \rho_{ijk}) \left[2(1 - 2\rho_{ijk}) \nabla_{ijk}(z_j^C) - \beta_k I \right]. \tag{13}$$

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 of region k's electorate around j's candidate policy position, $\nabla_{jk}^C(z_j^C)$,

$$\nabla_{jk}^C(z_j^C) \equiv \frac{1}{n_k} \sum_{i \in N_k} \nabla_{ijk}(z_j^C). \tag{14}$$

From (4) and taking the derivative of (6) with respect to z_{jk}, the $w \times w$ Hessian of second order derivatives for party j in region k located at z_j^C, $\mathcal{H}_{jk}$, is then

$$\begin{aligned} \mathcal{H}_{jk} &\equiv \mathcal{H}_{jk}(\mathbf{z}_k^C) = \frac{d^2 V_{jk}(\mathbf{z}_k^C)}{dz_j^2}|_{z_{-j}^C} = \frac{1}{n_k} \sum_{i \in N_k} \frac{d^2 \rho_{ijk}}{dz_j^2} \\ &= \sum_{i \in N_k} 2\rho_{ijk}(1 - \rho_{ijk}) \left[2(1 - 2\rho_{ijk}) \nabla_{ijk}^C(z_j^C) - \frac{1}{n_k} \beta_k I \right] \end{aligned} \tag{15}$$

where the last line follows after substituting in (13) and (14).

When all w eigenvalues of the Hessian $\mathcal{H}_{jk}$ of regional party j in region k are negative at z_{jk}^C, then j is at a maximum of its vote share at z_{jk}^C in region k. If, however, all w eigenvalues of the Hessian $\mathcal{H}_{jk}$ are positive then at z_{jk}^C, j is minimizing its vote share giving j an incentive to move away from z_{jk}^C in order to increase its vote share. If the Hessian $\mathcal{H}_{jk}$ has both positive and negative eigenvalues then at z_{jk}^C, j is at a saddle point of its vote share and should move to increase votes.

To find the eigenvalues of $\mathcal{H}_{jk}$, recall that the trace of the Hessian is equal to the sum of the eigenvalues associated with $\mathcal{H}_{jk}$ and is also given by the sum of the main diagonal elements of $\mathcal{H}_{jk}$. In order for z_{jk}^C 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 so that the trace of $\mathcal{H}_{jk}$ must also be negative.

Let ω represent a dimension in the policy space X. For parties and voters in region k, denote by $z_j^C(\omega)$ and $x_i(\omega)$ the position of party j and the ideal policy of voter i in dimension ω respectively. Using (15), 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

$$\begin{aligned} &\sum\nolimits_{i \in N_k} 2\rho_{ijk}(1-\rho_{ijk}) \left\{ 2(1-2\rho_{ijk}) \frac{1}{n_k} \beta_k \left[x_i(\omega) - z_{jk}^C(\omega) \right]^2 \beta_k - \frac{1}{n_k} \beta_k \right\} \\ = \; & \frac{1}{n_k} \sum\nolimits_{i \in N_k} 2\beta_k \rho_{ijk}(1-\rho_{ijk}) \left\{ 2(1-2\rho_{ijk}) \beta_k \left[x_i(\omega) - z_{jk}^C(\omega) \right]^2 - 1 \right\} \qquad (16) \end{aligned}$$

Suppose that in the above equation we only had $\frac{1}{n_k} \sum_{i \in N_k} \beta_k \left[x_i(\omega) - z_{jk}^C(\omega) \right]^2$, 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 (16) over the w dimensions of the policy space to obtain

$$\begin{aligned} trace[\mathcal{H}_{jk}] \; &\equiv \; \sum\nolimits_{\omega=1}^{w} \frac{1}{n_k} \sum\nolimits_{i \in N_k} 2\beta_k \rho_{ijk}(1-\rho_{ijk}) \left\{ 2(1-2\rho_{ijk}) \beta_k \left[x_i(\omega) - z_{jk}^C(\omega) \right]^2 - 1 \right\} \\ &= \; \frac{1}{n_k} \sum\nolimits_{i \in N_k} \sum\nolimits_{\omega=1}^{w} 2\beta_k \rho_{ijk}(1-\rho_{ijk}) \left\{ 2(1-2\rho_{ijk}) \beta_k \left[x_i(\omega) - z_{jk}^C(\omega) \right]^2 - 1 \right\} (17) \end{aligned}$$

Therefore, regional party j will be at a maximum at z_j^C iff all eigenvalues are negative. The necessary condition for this is that $trace[\mathcal{H}_{jk}] < 0$, which gives

$$\begin{aligned} \sum\nolimits_{i \in N_k} \sum\nolimits_{\omega=1}^{w} \beta_k \rho_{ijk}(1-\rho_{ijk}) \left\{ 2(1-2\rho_{ijk}) \beta_k \left[x_i(\omega) - z_{jk}^C(\omega) \right]^2 - 1 \right\} \; &< \; 0 \\ \text{or} \; \sum\nolimits_{i \in N_k} \sum\nolimits_{\omega=1}^{w} \mu_{ijk} 2(1-2\rho_{ijk}) \beta_k [x_i(\omega) - z_{jk}^C(\omega)]^2 \; &< \; w \qquad (18) \end{aligned}$$

Now let

$$v_s = \frac{1}{n_k} \sum\nolimits_{i \in N_k} \left(\left[x_i(s) - z_{jk}^C(s) \right] \right) : v_t = \frac{1}{n_k} \sum\nolimits_{i \in N_k} \left(\left[x_i(t) - z_{jk}^C(t) \right] \right)$$

so the variances on the sth and tthaxes are (v_s, v_s), (v_t, v_t). Then after some manipulation the last line follows after substituting in (7).

Define regional party j's *convergence coefficient* as the left hand side of (18), i.e.,

$$c_{jk}(\mathbf{z}_k^C) \equiv \sum\nolimits_{i \in N_k} \sum\nolimits_{\omega=1}^{w} \mu_{ijk} 2(1-2\rho_{ijk}) \beta_k](v_\omega, v_\omega)^2 \qquad (19)$$

Here $\frac{1}{n_k} \sum_{\omega=1}^{w} (v_\omega, v_\omega)^2$ is the total electoral variance in region k. Party j's convergence coefficient in region k depends on the weight given by party j to each voter in region k, μ_{ijk} in (7), on the probability that voters in region k vote for j, and on how dispersed voters in region k are around j's candidate position, $\beta_k[x_i(\omega) - z^C_{jk}(\omega)]^2$ which takes into account the weight that voters give to differences with j's policies β_k.

We have the following result: if when competing in region k, party j locates at its critical point z^C_{jk}, then the trace of the Hessian of its vote share function, $trace[\mathcal{H}_{jk}]$, will be negative only if 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^C_{jk} to maximize its vote share is that

$$c_{jk}(\mathbf{z}^C_{Nat}) < w \tag{20}$$

Let us now examine the necessary conditions for *national* party j located at its critical point z^C_j 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 (4) by taking the derivative of (8) to obtain

$$\begin{aligned}\mathcal{H}_j &\equiv \mathcal{H}_j(\mathbf{z}^C_{Nat}) = \frac{d^2V_j(\mathbf{z}^C_{Nat})}{dz_j^2}|_{z^C_{-j}} = \sum_{k\in\Re}\frac{n_k}{n}\frac{d^2V_{jk}(\mathbf{z}^C_k)}{dz_j^2} = \sum_{k\in\Re}\frac{n_k}{n}\mathcal{H}_{jk} \\ &= \sum_{k\in\Re}\frac{n_k}{n}\sum_{i\in N_k}2\rho_{ijk}(1-\rho_{ijk})\left[2(1-2\rho_{ijk})\nabla^C_{ijk}(z^C_j) - \frac{1}{n_k}\beta_k I\right]\end{aligned} \tag{21}$$

where the last term follows from (15) after substituting in (13) and (14).

Like in the regional case, if all w eigenvalues of the Hessian $\mathcal{H}_j$ of national party j are negative at z^C_j, then j is at a maximum of its vote share at z^C_j. If, however, all w eigenvalues of the Hessian $\mathcal{H}_j$ are positive then at z^C_j, j is minimizing its vote share and j will move away from z^C_j to increase its votes. If the Hessian $\mathcal{H}_j$ has both positive and negative eigenvalues then at z^C_j, j is at a saddle point of its vote share and should move to increase votes.

The main diagonal element of the Hessian of national party j along the ω dimension is given by the (ω,ω) element of the Hessian of national party j, $\mathcal{H}_j$. From (21), the (ω,ω) element of $\mathcal{H}_j$ is

$$\begin{aligned}&\sum_{k\in\Re}\frac{n_k}{n}\sum_{i\in N_k}2\rho_{ijk}(1-\rho_{ijk})\left\{2(1-2\rho_{ijk})\frac{1}{n_k}\beta_k\left[x_i(\omega)-z^C_j(\omega)\right]^2\beta_k - \frac{1}{n_k}\beta_k\right\} \\ =\ &\frac{1}{n}\sum_{k\in\Re}\sum_{i\in N_k}2\beta_k\rho_{ijk}(1-\rho_{ijk})\left\{2(1-2\rho_{ijk})\beta_k\left[x_i(\omega)-z^C_j(\omega)\right]^2 - 1\right\}\end{aligned} \tag{22}$$

To obtain the trace of the Hessian of national party j, $trace[\mathcal{H}_j]$, just add the diagonal elements given in (22) over the w dimensions of the policy space to obtain

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

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

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

where after some manipulation the last line follows from (9).

Then using (10), we get that the trace of $\mathcal{H}_j$ will be negative, $trace(\mathcal{H}_j) < 0$, only if

$$\sum_{\omega=1}^{w}\sum_{k\in\Re}\sum_{i\in N_k}\theta_{jk}\mu_{ijk}2(1-2\rho_{ijk})\beta_k\left[x_i(\omega)-z^C_j(\omega)\right]^2 < w \tag{23}$$

Define national party j's *convergence coefficient* as the left hand side of (23), i.e.,

$$c_j(\mathbf{z}_{Nat}^C) \equiv \sum_{\omega=1}^{w} \sum_{k\in\Re} \sum_{i\in N_k} \theta_{jk}\mu_{ijk} 2(1-2\rho_{ijk})\beta_k \left[x_i(\omega) - z_j^C(\omega)\right]^2 \tag{24}$$

National party j's convergence coefficient depends on the weight given by party j to region k, θ_{jk}; on the weight given by party j to each voter in region k, μ_{ijk}; on the aggregate probability that voters in region k vote for j; and on how dispersed voters in the country are around j's candidate position in the ω dimension, $\beta_k \left[x_i(\omega) - z_j^C(\omega)\right]^2$ where account is taken of the weight given by voters to difference with j's policies, β_k.

National party j stays at its critical point z_j^C iff it is maximizing its vote share, that is, only if

$$c_j^C(\mathbf{z}_{Nat}) < w \tag{25}$$

This says that the *necessary* condition for national party j to remain at the candidate position is that its convergence coefficient be less that the dimension of the policy space, w.

Note that national party j's convergence coefficient can be expressed as a function of the convergence coefficients it faces in each region since using (19), we can re-write (24) as$c_j(\mathbf{z}_{Nat}^C) \equiv \sum_{\omega=1}^{w} \sum_{k\in\Re} \sum_{i\in N_k} \theta_{jk}\mu_{ijk} 2(1-2\rho_{ijk})\beta_k \left[x_i(\omega) - z_j^C(\omega)\right]^2 = \sum_{k\in\Re} \theta_{jk} c_{jk}(\mathbf{z}_{jk}^C)$

Summarizing our results: In a model in which there are national and regional parties, our main result states that parties will locate or converge to the critical position satisfying the first order condition if the convergence coefficient for national and regional parties is less than the dimension of the policy space, w. In this case, if all parties locate at their critical point they will also 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. If all the convergence coefficients are less than w, then the vector of candidate positions $\mathbf{z}_{Nat}^C$ is an LNE of the game.

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

The above theoretical analysis generalizes the earlier work of Schofield (2007). In that work it was assumed that there was only a single region, and considered conditions under which the electoral mean vector (normalized to be at the origin) $\mathbf{z}_{Nat} = (0,..,0)$ could be a LNE. Under these assumptions in the current more general model the weight given to policy differences in the voter's utility in (1) is the same for all voters in the country, i.e., there is a single β. From (14) we can see that this variance term reduces to $\frac{1}{n}\sum_{i\in N}\left[x_i(\omega) - \mathbf{0}\right]^2$ which measures the total electoral variance about the origin. Writing this as σ^2 and imposing the assumptions on (24) 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, convergence in the election is determined by the incentives of party 1.

The general theoretical model gives a method to assess whether or not a vector of party positions is a LNE in a model with national and regional parties. The method is summarized as follows:

1. Define the vector of candidate 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 position meets the first order condition given in (6) for regional parties and in (8) for national parties while holding the position of the other parties constant.

- Note that if regional party j weights each voter in region k equally, so that from (7) $\mu_{ijk} = \frac{1}{n_k}$ 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 regional party j.

- Note also that if national party j weights each region according to their population share so that from (9) $\theta_{jk} = \frac{n_k}{n}$, and also weights all voters in region k equally, so that from (7) $\mu_{ijk} = \frac{1}{n_k}$ then j locates at

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

Party j then locates at the mean of the ideal points of 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.

3. The Hessian for a regional party at the candidate position is given by (15) and that of a national party at the candidate position is given by (21). Find the eigenvalues of the Hessian for national and regional parties. If all of the eigenvalues are negative, the vector of positions is a local Nash equilibrium. The necessary condition is that the eigenvalues of the Hessian of the national and regional parties be all negative, that is, the trace of the national parties given in (23) and the trace of regional parties in (17) are both negative.

4. Calculate the convergence coefficient for each for each regional party using (19) and for each national party using (24). The convergence coefficient of each party determines whether or not the party wants to move from the candidate position. Clearly, if the party with the highest convergence coefficient does not want to move from the candidate location (because it is maximizing its votes) then no other party will want to move from its location either. Thus, the party with the highest convergence coefficient determines whether parties converge to their candidate positions as whole. The convergence coefficient of the election is then the highest of the convergence coefficients of the national and regional parties. i.e.,

$$c(\mathbf{z}) = \max \left\{ \max c_j^C(\mathbf{z}_{Nat}), j \in P_{Nat}; \max c_{jk}(\mathbf{z}_k^C), j \in P_1; ...; \max c_{jk}(\mathbf{z}_k^C), j \in P_r \right\}$$

If $c(\mathbf{z}) > w$, then at least one party will not converge to the candidate position.

We now describe the empirical implementation of the methodology identified by the formal stochastic model before applying it to the 2004 Canadian election.

3 Estimation Strategies Given Varying Party Bundles

In order to utilize the stochastic election model proposed above, we need to have measures of competence and socio-demographic valence for all national and regional parties. In addition we need to estimate the regional weights given by voters to the policy differences they have with the parties, β_k for $k \in \Re$. We also need to estimate the party positions and well as the weights parties give regions and voters in their policy space, i.e., we need to estimate μ_{ijk} in (7) and θ_{ij} in (10).

When there is only one region, the assumption that the random errors in voters stochastic decisions have a Type-I extreme value distribution implies that voters' probabilistic choices have a multinomial logit (MNL) specification that can be easily estimated. However, as shown in Section 2, when there is at least one regional and one national party competing in the election the situation is far more complex. The fact that voters in different regions of the country face different sets of parties, implies that we can no longer use use the MNL model to estimate the parameters of the model. The reason is simple: MNL models rely on the assumption of independence of irrelevant alternatives (IIA) which requires that all *odds ratios* between the probability that each voter i in each region $k \in \Re$ votes for a pair of parties j and h, ρ_{ijk}/ρ_{ihk} be independent of party ℓ in region k where ρ_{ijk} and ρ_{ihk} are given by (2). If we use the MNL model, the model would require that this odds ratio be preserved from region to region, even though the choice sets varies across regions. Since IIA is violated in the regional model, we cannot use the MNL model as the estimation procedure.

Yamamoto (2011) proposes a model that overcomes these problems: the *varying choice set logit* (VCL) model. Like the MNL model, the VCL model assumes that the errors in voters choices follow at Type-I extreme value distribution, the same assumption used in Section 2 to derive the convergence coefficient. That is, the VCL model assumes that the probability that voter i votes for party j in region k is given by (2). Thus, the framework of the formal and the empirical models match, making the transition to empirical estimations of the parameters of the model easy. We can then analyze the equilibria of the system given the parameter estimates.

The VCL differs from MNL regression models in that it avoids the IIA assumption. This is done by allowing there to be individual logistic regression models for each choice set type, meaning each region, then aggregating these estimates to make an estimate of valences at the national level. In this case, each region has a different bundle of parties offered to voters.

The VCL model allows, just as assumed in Section 2, that the parameters of the voter's utility be region specific. Thus, the empirical model assumes that the utility voter i derives from voting for party j in region k, given (1), is

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

Here λ_{jk}, the competence valence of party j is region specific as is the weight given by voters to policy differences with parties, β_k. The term α_{jk} is the added utility that voter i gets over the competence valence in region k from voting for party j and represents the utility that average member from sociodemographic group s gets from voting for party j in region k. This hierarchical specification of the valence terms lends itself very well to the VCL model. With the error terms coming from a Type-I extreme value distribution, the empirical probability that voter i votes for party j has a logit specification as given in (2). So that the empirical and formal models clearly line up and makes using the VCL the proper choice 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.[14] Then, using these estimates, we assume that these individual estimates come from their own distribution, and we use that to determine the best

[14] If we assume that 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. This method of estimation is best done utilizing random effects.

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, 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. 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 within 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.[15] Thus, this model is appropriate when there are regional parties representing the specific interests of that region. 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 and 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 individual, demonstrating a significant efficiency gain over the standard mixed logit model. The mixed logit does not allow the researcher to estimate choice set specific values of parameters, thus VCL is more efficient and informative. [16]

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 valence λ_{jk}, the sociodemographic valence α_{jk} and the random effects all have underlying normal distributions. 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. However, 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

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

[16] 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.

from those garnered when independence is assumed. However, it is unreasonable 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. However, 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 as a predictor in the random effects. In doing this, the researcher controls 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.

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 individual observations. 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.

We now study the 2004 Canadian election using the VCL model.

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 them in a "neck-and-neck" race making the "the election too close to call."[17] 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.[18] As the campaign advanced, the Conservatives made two major mistakes. A Conservative MP accused Martin of being soft on child pornography. 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 rightwing 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.

[17] 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).

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

[**Insert Tables 1A, 1B and 1C here**]

Regional differences were of primal importance in this election.[19] The Liberals rating plummeted when Sponsorship scandal broke, specially in Québec with Québécers massively turning to the Bloc. The Liberals only 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[20] in Table 1A). 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 and the Prairie Provinces. The Conservatives who clearly 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 that 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. Givent the difference in political environments we study the election in these two regions: Québec (Q) and Canada outside Québec (C/Q).

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 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.

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, did not gain any seats in the Parliament.

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 1B shows vote shares within the sample and the actual vote shares. The similarity between these two sets of shares suggests that the sample is fairly representative of the Canadian electorate. Table 1B also has the data for Québec, as Bloc Québécois only ran in Québec.

The factor analysis performed on the voters' responses in the survey questions led us to conclude that there were two factors or policy dimensions: one "social," the other "decentralization." The

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

[20] 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).

social dimension is 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 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 their lives. A greater desire for decentralization implies higher values on this axis. The questions used in the factor analysis can be found in Table 2.

Using the factor loadings from the factor analysis given in Table 3 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 the origin, $(0, 0)$. 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.

[**Insert Tables 2 and 3 here**]

The survey asked voters which party they would be voting for, so we estimated party positions as the mean of voters for that party. The party positions in the policy space are 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 shows the distribution of voters and the party position for all of Canada with "Q" representing the electoral mean in Québec, which is noticeably left of the overall electoral mean. Figure 2 shows the voter distribution for Québec only. The majority of voters in Québec advocate more liberal social policies than the average voter in Canada and want more decentralization of government, as Québec has a strong regional identity and wants to maintain its somewhat independent state.

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

The *electoral covariance matrix* for the entire sample of 862 respondents ∇^C (Table B1) as

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

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 $\sigma^2_{C/Q} \equiv \sigma^2_S + \sigma^2_D = 3.88$ 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, has an electoral covariance matrix

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

whose "total" variance is $\sigma^2_Q = 2.46$ 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 illstrated 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 (immigration). 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.

The survey also collected sociodemographic data. For each respondent, sex, age, and education level were recorded. Age was divided into four categories: 18-30, 30-65, 65 and older. Education was divided into three categories: No College degree, more than college degree. While Table B1 shows that there are no major sociodemographic differences between non-Québec and Québec respondents, there are differences in the characteristics of party supporters. The mean Liberal supporter is older than that of other parties with the youngest voters tend to support the Greens. More than half those voting Liberal, NDP and BQ were women with more than half of those voting Conservatives or Green were men.

4.2 Modelling the 2004 Canadian election

Clarke *et al.*(2005) point out that in the last stages of the campaign it was not clear which of the two front runners, the Liberals or Conservatives, would win as the election was to close to call. The polls indicated that neither party would win a majority of votes. Since support for both parties was hovering around the $33-35\%$ range (Table 1A), neither party expected to win a parliamentary majority either. This promoted parties, voters and political commentators to speculate which party would form a minority government.

With Canada having a fist-past-the-post system and the election so 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 their core supporters but also voters in another riding and would also be vulnerable to changes in voters' mood. To avoid being seen as unsure of their position parties make only small changes to their policy during the campaign.

The election had two different battle fronts: the one in Québec was between the Liberals and the Bloc and in the rest of Canada, basically in Ontario, between the Liberals and Conservatives.[21] Table 1A shows that the Conservative 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 on a daily basis. These public polls were closely followed by all parties. They reported national and provincial vote shares as the poll sizes were too small to give riding level shares. Parties used then 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 apply the formal national-regional stochastic electoral model to the elections in the two regions: Canada outside Québec and Québec using the varying choice set logit (VCL) model. When estimating the VCL model we use the Liberal Party as the base party, so that the coefficients of the models are measured relative to that of the Liberals whose coefficients are standardized to zero.

Due to the structure of the VCL and the underlying random effects model, sociodemographics are viewed as categorical so that groups can be made. 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. In this case, after toying with the model for some time, it seemed 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.

Using the VCL proposed in Section 3, for each region we estimate the weight given to policy differences with each party β_k and the valences for a model with sociodemographics. For the model, given some correlation between the random effects of interest and the independent variable of Euclidian difference, we use the random effects correction procedure proposed earlier. 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. To assist in convergence of the VCL, we create a diffuse gamma hyperprior for the variance of each prior. As stated before, this model does take a while to converge, so it is necessary to let the Gibbs sampler for this model run a while (see the Gibbs sampling algorithm in Appendix A) We ran each Gibbs sampler for 100,000 iterations and received nice normal distributions for each of the parameters of interest. Similarly, allowing the Gibbs sampler to run this long reduces the effects of the inherent autocorrelation that occurs in the sampler.

The results of the VCL are given in Table 4. We show the VCL estimates of the parameter values and the corresponding 95 percent credible/confidence intervals. Note that while the sociodemographic random effect values may be of substantive interest sometimes, they are included simply as controls in this case, thus we do not report these values. We also report the *deviance information criterion* (DIC), which is a hierarchical model analogue to AIC or BIC. When the posterior

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

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 4 about here**]

The results in Table 4 suggest a number of things. 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 can see that the valence of the Conservatives and the Liberals are almost equivalent outside of 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 lowest valence party outside Québec is the Green Party, mainly because it is a one-issue party that failed to appeal to the electorate.

The political differences between Québec and the rest of Canada come across clearly in Table 4. Given the failed attempts to bring Québec back into the Constitution that led the prominence of the BQ in the 2004 election (Table 1A), it is no surprise that the BQ had the highest competence valence in Québec in 2004. The Liberals, who by the end of the campaign had recovered in the 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.[22] 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 goven than the BQ and the Liberals and at par with the Conservatives. As in the rest of Canada, the Greens had the lowest competence valence in Québec.

Table 4 also includes the coefficients of the socio-demographic variables included in our analysis: the Age and Education dummies. We report the cofficients for each age group and education level for both outside and in Québec. The socio-demographic valences of these groups varied by party and by region once the policy differnces betweeen parties are taken into account. For example, voters outside Québec who are between 30 and 65 years old and 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.

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 valences 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 these two parties seem to have the same valence as that of the BQ is not statistically different from zero. This is mainly due to the fact that although nearly half of Québécers voted for the BQ (Table 1B), Québec represents only a quarter of the Canadian population.

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

4.4 Is there convergence to the electoral means in Canada?

Recall that we are interested in finding where the parties will locate in the policy space so as to maximize their vote share. Because the outcome of the election depends on these vote shares, we assume that parties use polls and other information at their disposal to form an idea of the anticipated electoral outcome. Parties then use information on voters' responses to changes in position to find a policy position that will allow then to increase their vote shares taking into account their estimates of where other parties will 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 can calculate the vote share function of the party at different policy positions. The assumption here is that parties use the VCL model as a heuristic estimate of the anticipate electoral outcome to position themselves so as to maximize their expected vote share. Given a candidate position, we can use (21) and (15) for national and regional parties respectively to examine whether these parties are maximizing their expected vote shares at these positions.

One possible location is for each party to locate at its respective electoral mean, meaning in Canada that national parties locate at the national mean and the BQ locates at Québec mean. From Table B1 and B2 the vector of candidate positions $\mathbf{z}^C_{Nat}$ is then given by

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

We can then calculate the party's expected vote shares at this position. Using (21) and (15) to find the Hessian of second derivatives at the candidate position for each national and regional party respectively. The Hessian for each party at $\mathbf{z}^C_{Nat}$ are given by

$$\begin{aligned} \mathcal{H}_{Lib} &= \begin{bmatrix} -.0365 & -.0004 \\ -.0004 & -.0705 \end{bmatrix}; \quad \mathcal{H}_{NDP} = \begin{bmatrix} .0021 & .0012 \\ .0012 & -.0362 \end{bmatrix}; \quad \mathcal{H}_{Con} = \begin{bmatrix} -.0326 & -.0002 \\ -.0002 & -.0676 \end{bmatrix} \\ \mathcal{H}_{GPC} &= \begin{bmatrix} .0085 & .0085 \\ .0085 & -.0091 \end{bmatrix}; \quad \mathcal{H}_{BQ} = \begin{bmatrix} -.1194 & .0034 \\ .0034 & -.1286 \end{bmatrix} \end{aligned} \tag{26}$$

In order to determine whether each party is maximizing its vote share at $\mathbf{z}^C_{Nat}$ we need to find the eigenvalues associated with the Hessians for each party; if they are both negative, then the Hessian is negative definite at the candidate position implying that at $\mathbf{z}^C_{Nat}$ the party is at a local maximum. If the Hessian of any party is not negative definite at $\mathbf{z}^C_{Nat}$, then this party will have an incentive to move from the candidate position. Other parties may then also find it in their best interest to move. The vector of candidate positions is then not a LNE. The eigenvalues of the Hessian for each party given in (26) are

$$eigen(\mathcal{H}|\mathbf{z}^C_{Nat}) = \begin{bmatrix} & Lib & NDP & Con & Green & BQ \\ Eigen1 & -.0365 & .0021 & -.0326 & .0085 & -.1183 \\ Eigen2 & -.0705 & -.0361 & -.0676 & -.0092 & -.1297 \end{bmatrix} \tag{27}$$

The analysis of the eigenvalues of the Hessians in (27) and their confidence intervals allow us to conclude with a high degree of certainty that if the Liberals and Conservatives locate at the national electoral mean they are maximizing their expected vote shares. Similarly, if the BQ locates at the Québec electoral mean then it will also be maximizing its expected vote share in Québec. Thus, if all parties locate at the respective means, the Liberals, the Conservatives and the Bloc will not want to move from their candidate position. Note however that the eigenvalues indicate that the

Greens and the NDP are at a saddle point when they locate at the national electoral mean and will thus want to move away from the national mean.

The formal model presented in Section 2 determined the necessary and sufficient conditions for parties to converge to or remain at the candidate position. Remember that if any party fails to meet the necessary condition for convergence then this party has an incentive to move away from the candidate 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 it will be the Greens.

Using the VCL parameters estimates in Table 4 and (2) we can estimate the probability that a voter in Québec and in the rest of Canada vote for each party. Then using (24) and (19) respectively for national and regional parties, the estimated probabilities, and the VCL estimates in Table 4, we can calculate the convergence coefficients for each party and then examine whether each party meets the necessary condition for convergence, that is if the party's convergence coefficient is less than the dimension of the policy space, $w = 2$. The convergence coefficient for the four national parties and for the BQ are

$$c_j(\mathbf{z}^*) = \begin{bmatrix} & Lib. & NDP & Con. & Grn. & BQ \\ c(\mathbf{z}^*) & 1.031 & 1.518 & 1.071 & 1.945 & -.5921 \end{bmatrix} \quad (28)$$

The analysis of the eigenvalues of the Hessians in (27) and the convergence coefficients in (28) of all parties allow us to conclude that two national parties will diverge from the national electoral mean. The NDP and the Greens both have positive eigenvalues, meaning that $\mathbf{z}^C_{Nat}$ is not a vote maximizing position for them and, thus, not a LNE. . However, when they choose a better position, it will still be on the mean of the decentralization axis as the second eigenvalue represents that axis and has a negative value.

We can also utilize the test of convergence coefficients to assess convergence to the vector of interest. Here, we see that all of the convergence coefficients, except for BQ's, are greater than one but less than w (which in this case is 2) , thus we need to check the largest one to see if it indicates convergence to the mean vector. The largest convergence coefficient belongs to the Green Party and examination of the constituent portions of its $c(\mathbf{z}^*)$ shows:

$$c_{GPC}(\mathbf{z}^C_{Nat}) = 1.379 + .5657$$

where 1.379 corresponds to the social axis. This means that the Green Party is not maximizing its vote share at the mean social position. These values indicate that the Green Party is also located at a saddle point when given the mean vector, just as the Hessian test did.

However, taken as they are, we do not know if these two tests actually match the vote maximizing tendencies of the parties. Thus, 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. In a Gibbs sampling style of optimizer, we create an Nash style optimization method in which each party optimizes its vote share given the positions of the other parties. If we do this for each party in rotation beginning at some arbitrary starting values, the parties should eventually converge on the local Nash equilibrium set of positions where no party can do any better by marginally moving given the positions of the other parties. 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 style of 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:

$$z^*_{opt} = \begin{bmatrix} & Lib & Con & NDP & Grn & BQ \\ S & 0.0524 & 0.0649 & 1.099 & 2.337 & -1.069 \\ D & -0.0259 & -0.0264 & 0.0266 & 0.2281 & -0.1290 \end{bmatrix}$$

[**Insert Figure 3 here**]

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

This begs the question, though, how much better can the parties do at these positions than they did at their current positions? Table 5 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 predicted using the actual valences from each region (i.e., the competence and socio-demographic valences plus the regional random effects).

[**Insert Table 5 here**]

Table 5 strengthens our conclusions that the vector of means is not a LNE as the Greens, the BQ, and the Liberals all do better when the Greens and the NDP locate away from the national mean. As the Greens is one of the parties dissatisfied with the national mean, it can choose to move to a more extreme position and gain votes. The NDP is forced to adapt and does worse than it would if the parties all located at their respective electoral means.

Conclusion

In this paper, we develop a formal stochastic model of an election in which voters in different regions face different bundles of national and regional parties. We show that the necessary condition for a party to converge to or remain at a candidate position is that the convergence coefficient is is less than the dimension of the policy space. When the policy space is two dimensional the sufficient condition for parties to converge to the candidate position is that the convergence coefficient is less than the 1. This model is more general than that proposed by Schofield (2007) since it handles any number of national and regional parties. Usung this model we can study whether any proposed vector of candidate positions for national and regional parties is a local Nash equilibrium for the election.

Given the presence of national and regional parties we show that due to the violation of the IIA assumption we can no longer use MNL valence models to estimate voters choices and thus cannot derive the party positions from these estimates. To overcome this problem we propose a new method for estimating the parameters necessary to estimate the convergence coefficient that does not rely on the IIA assumption. We show that the VCL model allows us to take advantage of more information than other exiting empirical methods. Thus the VCL model is ideal when examining voting tendencies within complex electorates that have clear hierarchical structures.

Using the formal model and its VCL empirical implementation, we examined the 2004 Canadian election. Our results show that the Bloc Québécois has the highest competence valence in Québec with the Liberals and Conservatives being tied in the rest of Canada 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 two national but lower valence parties needed to adopt more extreme positions in the policy space. Thus, our analysis showed that two national parties, the NDP 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 parties can always maximize their vote shares by taking positions at the electoral center.

References

[1] ADAMS, J. (2001) Party Competition and Responsible Party Government. Ann Arbor, MI, University of Michigan Press.

[2] ADAMS, J. and MERRILL, S., III (1999), "Modeling Party Strategies and Policy Representation in Multiparty Elections: Why are Strategies so Extreme?", *American Journal of Political Science*, 43, 765–781.

[3] BENOIT, K., LAVER, M. (2006), *Party Policy in Modern Democracies*, Routledge, London.

[4] BESLEY T. and COATE, S. (1997) An Economic Model of Representative Democracy. *Quarterly Journal of Economics* 112: 85-114.

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

[6] CAPLIN, A. and NALEBUFF, B. (1991), Aggregation and Social Choice: A Mean Voter Theorem, *Econometrica*, 59, 1–23.

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

[8] DOW, J. K. and ENDERSBY, J. (2004), Multinomial Logit and Multinomial Probit: A Comparison of Choice Models for Voting Research, *Electoral Studies*, 23, 107–122.

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

[10] 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.

[11] HINICH, M. J. 1(977). Equilibrium in Spatial Voting: The Median Voter Theorem is an Artifact. *Journal of Economic Theory* 16:208–219.

[12] HOTELLING, H. (1929), Stability in Competition, *Economic Journal*, 39, 41–57.

[13] KRAMER, G. (1978), Existence of Electoral Equilibrium, in P. Ordeshook (Ed.) *Game Theory and Political Science.*New York: New York University Press, pp.375–389.

[14] McKELVEY R.D. and SCHOFIELD, N. (1987) Generalized Symmetry Conditions at a Core Point, *Econometrica* 55 : 923-933

[15] MERRILL, S. and GROFMAN, B. (1999) *A Unified Theory of Voting.*Cambridge, U.K., Cambridge University Press.

[16] OSBORNE M.J. and SILVINSKI, A. (1996), A Model of Political Competition with Citizen Candidates. *Quarterly Journal of Economics* 111: 65-96.

[17] PATTY, J. W. (2005), Local Equilibrium Equivalence in Probabilistic Voting Models. *Games and Economic Behavior*, 51, 523–536.

[18] 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*, 29: 149-173.

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

[20] 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.

[21] POOLE, K. and ROSENTHAL, H. (1984), US Presidential Elections 1968–1980, *American Journal of Political Science*, 28, 283–312.

[22] QUINN, K., MARTIN, A. and WHITFORD, A. (1999), Voter Choice in Multiparty Democracies, *American Journal of Political Science*, 43, 1231–1247.

[23] RIKER, W. H. (1964). *Federalism: Origin, Operation, Maintenance.* Boston. MA, Little Brown.

[24] RIKER, W. H. (1987). *The Development of American Federalism.* Boston, MA, Kluwer.

[25] RIKER, W. H. and ORDESHOOK, P. C. (1973) *An Introduction to Positive Political Theory* Englewood Cliffs, NJ: Prentice-Hall.

[26] RIKER, W H. 1980. Implications from the disequilibrium of majority rule for the study of institutions. *Amererican Political Science Review* 74, 432–446.

[27] RIKER, W H. 1982. The Two Party System and Duverger's Law: An Essay on the History of Political Science.. *Amererican Political Science Review* , 76:753–766.

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

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

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

[31] SCHOFIELD, N. (2006) Equilibria in the Spatial Stochastic Model of Voting with Party Activists, *Review of Economic Design* 10 (3): 183-203.

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

[33] SCHOFIELD, N. GALLEGO, M and J.S.JEON (2011 Leaders, Voters and Activists in Elections in the Great Britain 2005 and 2010. *Electoral Studies* 30(3) :484-496.

[34] 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.

[35] TRAIN, K. (2003) *Discrete Choice Methods for Simulation*. Cambridge, U.K., Cambridge University Press.

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

4.5 Appendix A

This appendix gives the algorithm for the Gibbs sampling.

```
model{

for(i in 1:N) {
for(k in 1:K) {
v[i,k] <- alpha[k] + beta[1]*(d[(N*(k-1))+i]-d[i]) + m[region[i],k] +
ed[region[i], education[i], k] + ag[region[i],education[i],age[i],k]

expv[i,k] <- exp(v[i,k])
pv[i,k] <- expv[i,k]/sum(expv[i,1:K])
vote[i]  ~ dcat(pv[i, 1:K])
}}

beta[1] ~ dnorm(0,taub[1])I(-5,5)

alpha[1] <- 0
alpha[2] ~ dnorm(0,taua[2])
alpha[3] ~ dnorm(0,taua[3])
alpha[4] ~ dnorm(0,taua[4])
alpha[5] ~ dnorm(0,taua[5])

m[1,1] <- 0
m[1,2] ~ dnorm(0,taum[1,2])
m[1,3] ~ dnorm(0,taum[1,3])
m[1,4] ~ dnorm(0,taum[1,4])
m[1,5] <- -100000
m[2,1] <- 0
m[2,2] ~ dnorm(0,taum[2,2])
m[2,3] ~ dnorm(0,taum[2,3])
m[2,4] ~ dnorm(0,taum[2,4])
m[2,5] ~ dnorm(0,taum[2,5])

taub[1] ~ dgamma(.1,.1)I(.1,10)
```

```
taua[2] ~ dgamma(.1,.1)I(.1,10)
taua[3] ~ dgamma(.1,.1)I(.1,10)
taua[4] ~ dgamma(.1,.1)I(.1,10)
taua[5] ~ dgamma(.1,.1)I(.1,10)
taum[1,2]~dgamma(.1,.1)I(.1,10)
taum[1,3]~dgamma(.1,.1)I(.1,10)
taum[1,4]~dgamma(.1,.1)I(.1,10)
taum[2,2]~dgamma(.1,.1)I(.1,10)
taum[2,3]~dgamma(.1,.1)I(.1,10)
taum[2,4]~dgamma(.1,.1)I(.1,10)
taum[2,5]~dgamma(.1,.1)I(.1,10)

for(f in 1:e){
ed[1,f,5] <- -10000
}

for(f in 1:e){
for(z in 1:4){
ed[1,f,z] ~ dnorm(0,taued[1,f,z])
taued[1,f,z] ~ dgamma(.01,.01)I(.01,10)
}}

for(f in 1:e){
for(z in 1:5){
ed[2,f,z] ~ dnorm(0,taued[2,f,z])
taued[2,f,z] ~ dgamma(.01,.01)I(.01,10)
}}

for(f in 1:e){
for(w in 1:a){
ag[1,f,w,5] <- -10000
}}

for(f in 1:e){
for(z in 1:4){
for(w in 1:a){
ag[1,f,w,z] ~ dnorm(0,tauag[1,f,w,z])
tauag[1,f,w,z] ~ dgamma(.01,.01)I(.01,10)
}}}

for(f in 1:e){
for(z in 1:5){
for(w in 1:a){
ag[2,f,w,z] ~ dnorm(0,tauag[2,f,w,z])
tauag[2,f,w,z] ~ dgamma(.01,.01)I(.01,10)
}}}

for(f in 1:e){
```

```
for(z in 1:4){
for(w in 1:a){
tot[1,f,w,z] <- alpha[z] + m[1,z] + ed[1,f,z] + ag[1,f,w,z]
}}}

for(f in 1:e){
for(z in 1:5){
for(w in 1:a){
tot[2,f,w,z] <- alpha[z] + m[2,z] + ed[2,f,z] + ag[2,f,w,z]
}}}
}
```

5 APPENDIX B: ELECTION STATISTICS

Table B1: 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

Table B2: : Descriptive Statistics by Party						
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 1A: 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 1B: 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 1C: Provincial Votes (%) and Seats in the 2004 Canadian Election										
Region	Western Provinces									
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			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.

Table 2: Survey Questions	
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

Table 3: 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 4: 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+	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						

(brackets give credible confidence intervals)

[a] Numbers in brackets are credible confidence intervals

[b] Gelman-Rubin

[c] “<col” = Less than college degree and “>col” = More than college degree

* significant at the 5%.level

Table 5: Vote Shares given various policy positions			
	Actual	Mean	Optimal
LPC	36.71	33.42	33.43
CPC	29.66	33.34	33.29
NDP	15.65	17.89	16.96
GPC	4.29	3.55	3.80
BQ	12.42	11.79	12.52

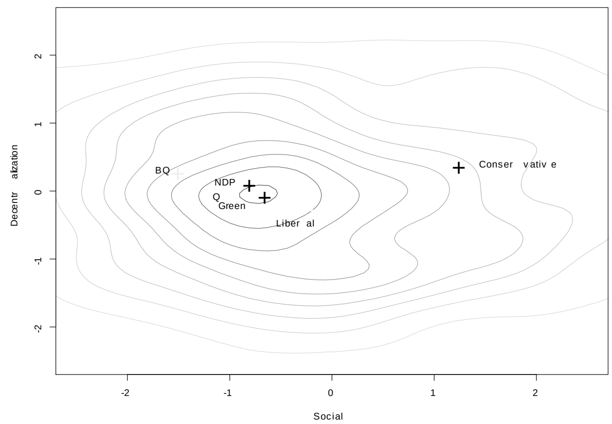

Figure 1: Distribution of voters and party positions for Canada in 2004

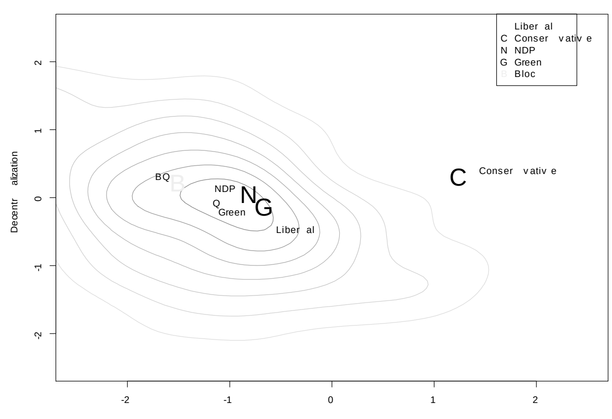

Figure 2: Distribution of voters and party positions for Quebec in 2004

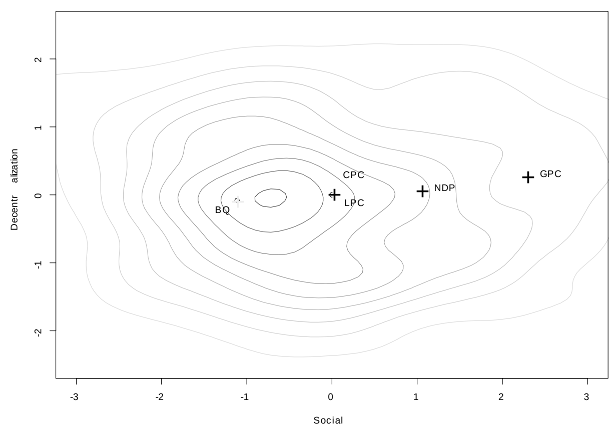

Figure 3: Vote maximizing positions in Canada 2004