

# Modeling Elections with the Variable Choice Set Logit Model : Application to the 2004 Canadian Election\*

Norman Schofield<sup>†</sup>      Kevin McAlister<sup>‡</sup>, Jee Seon Jeon<sup>§</sup>  
Washington University in Saint Louis, MO 63130      Maria Gallego.<sup>¶</sup>

November 6, 2012.

## Abstract

Previous models of elections have emphasized the convergence of parties to the center of the electorate in order to maximize votes received. More recent models of elections demonstrate that this need not be the case if asymmetry of party valences is assumed and a stochastic model of voting within elections is also assumed. This model seems able to reconcile the widely accepted median voter theorem and the instability theorems that apply when considering multidimensional policy spaces. However, these models have relied on there being a singular party bundle offered to all voters in the electorate. In this paper, we seek to extend these ideas to more complex electorates, particularly those where there are regional parties which run for office in a fraction of the electorate. We derive a convergence coefficient and set forth necessary and sufficient conditions for a generalized vector of party positions to be a local Nash equilibrium; when the necessary condition fails, parties have incentive to move away from these positions. For practical applications, we pair this finding with a microeconomic method for estimating parameters from an electorate with multiple regions which does not rely on independence of irrelevant alternatives but allows estimation of parameters at both aggregate and regional levels. We demonstrate

---

\*Prepared for presentation at the Sertel workshop on Economic Design, Paris, November 2012.

<sup>†</sup>corresponding author: (schofield.norman@gmail.com)

<sup>‡</sup>email: Kevin McAlister@wustl.edu

<sup>§</sup>jsjeon@wustl.edu

<sup>¶</sup>Department of Economics, Wilfrid Laurier University, 75 University Avenue West, Waterloo, Canada N2L 3C5 (mgallego@wlu.ca).Tel: 519.884.0710 Ext 2635

the effectiveness of this model by analyzing the 2004 Canadian election. The work we present here can be seen to complement Riker's discussion of Federalism (Riker, 1964, 1987), as it suggests that regional parties in a Federation will adopt very different policies from National parties.

## 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. These models showed that there exists a Condorcet point at the electoral median. 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 with regards to actual electoral systems (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 uncovered set includes the electoral mean, thus giving some credence to the median 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 chaos theorems which apply to multidimensional policy spaces.

The contrast between the instability theorems and the stability theorems suggest that a model in which the individual vote is not deterministic is most appropriate (Schofield, Martin, Quinn and Whitford, 1998; Quinn, Martin and Whitford, 1999). This kind of stochastic model states that the voter has a vector of probabilities corresponding to the choices available in the election. This insinuates that if the voter went to the polls for the same election multiple times, he might not make the same vote every time. This model

is in line with multiple theories of voter behavior and still yields the desirable property of showing that rational parties will converge to the electoral mean given the simple spatial framework.

Using this framework, Schofield(2007) shows that convergence to the mean need not occur given that valence asymmetries are accounted for. In this context, valence is taken to mean any sorts of quality that a candidates has that is independent of his location within a policy space. In general, valence is linked to the revealed ability of a party to govern in the past or the predicted ability of a party to govern well in the future. In recent years, models with a valence measure have been developed and utilized in studies of this sort. Schofield extends upon these models and demonstrates a necessary and sufficient condition for convergence to the mean, meaning that the joint electoral mean is a local pure-strategy Nash equilibrium (LNE) in the stochastic model with valence.

Valence can generally be divided into two types of valence: aggregate valence (or character valence) and individual valence (or sociodemographic valence). Both types of valence are exogenous to the position that a party takes in an election, meaning that these valence measures rely on some other underlying characteristic. Aggregate valence is a measure of valence which is common to all members in an electorate, and can be interpreted as the average perceived governing ability of a party for all members of an electorate (Penn, 2003). Individual valence is a bit more specific, where this kind of valence depends upon the characteristics of a voter. This kind of valence differs from individual to individual. 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 valence among African-American voters than the Republican candidate is. 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 of Type-I extreme value distributed errors (Dow and Endersby, 2004). These errors, commonly associated with microeconomic models, are typical of models that deal with individual choice, where individual utility is determined by the valence terms and the

individual's distance from the party in the policy space. This distance is weighted by  $\beta$ , a constant that is determined by the average weight that individuals give to their respective distances from the parties. The workhorse of individual choice models is the multinomial logit distribution, which is an extension of the dichotomous response logit distribution. This distribution assumes that the probability that an individual votes for a party follows the Type-I extreme value distribution, thus matching the assumed distribution of the stochastic voting model. This creates a natural empirical partner for the stochastic vote model.

Using this statistical framework and the assumption that individual choice follows this distribution, Schofield(2007) introduced the idea of the convergence coefficient,  $c$ , which is a measure of attraction to the electoral mean in an electoral system. This coefficient is unitless, thus it can be compared across models. Low values of this value indicate strong attraction to the electoral mean, meaning that the electoral mean is a local pure-strategy Nash equilibrium (Patty, 2005; Patty, 2006). High values indicate the opposite. He also lays out a necessary and a sufficient condition for convergence to the electoral mean with regards to the convergence coefficient:

1. When the dimension of the policy space is 2, then the sufficient condition for convergence to the electoral mean is  $c < 1$ .
2. 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 adopt a position away from the electoral mean in equilibrium, meaning that a LNE does not exist at the electoral mean. As a LNE must exist for the point to be a pure strategy equilibrium, this implies non-existence of a PNE at the center. Given the definition of the convergence coefficient, the general conclusion is that the smaller  $\beta$  is, the smaller the valence differences are among candidates, and the lower the variance of the electoral distribution is, the more likely there is to be a LNE at the electoral center.

However, this only answers the question where the local Nash equilibria are in the simplest case of having one electoral mean that parties are responding to. This problem can

quickly become more complicated. Imagine a country with five parties and two different regions. Four of the parties run in both regions, and are thus attempting to appeal to voters in both regions. However, one of these parties only runs in one of the regions and is only trying to appeal to the voters of this region. Thus, it would be unreasonable for it to position itself with regards to the electoral mean for the entire electorate. Rather, it wants to maximize its vote share within in the region in which it runs. Parties can choose to run in select regions for a variety of reasons. They may run for historical reasons or responsive reasons or even choose not to run in regions where they know they will not do well at all. As parties have limited resources, sometimes this kind of decision must be made.

In order to assess convergence to the electoral mean in this case, one must take into account the electoral centers that parties are responding to. In the above example, convergence to the electoral mean would mean that the first four parties converge to the overall electoral mean, or the mean of all voters in the electorate, while the fifth party would converge to the electoral mean of those individuals in its respective region. Thus, the convergence coefficient would no longer be appropriate, as it is proven only when the position for all parties is equal to zero on all dimensions. Similarly, when there are parties which run in different combinations of regions, the typical multinomial logit model is no longer appropriate because the underlying assumption of "independence of irrelevant alternatives"(IIA) is no longer met (Train, 2003). Given that there are problems with estimation of parameters from the currently utilized empirical methodology and problems with the underlying theoretical mechanism that drives the reasoning behind the convergence coefficient, we are left without the useful information gained about party tendencies in the stochastic model. Under the current framework, researchers can only analyze convergence, valence, and spatial adherence within specific regions. However, in this paper we propose a method for handling more structurally complex electorates.

In this chapter, we introduce methods for analyzing the stochastic vote model in electorates where individuals do not all vote for the same party bundle. First, this chapter will demonstrate that the convergence coefficient first defined by Schofield can be adjusted to handle any vector of party positions. We will determine the first and second order condi-

tions necessary to show that a vector of policy positions is a local Nash equilibrium (LNE). From this, we will show that the convergence coefficient for a more complex electorate can be derived in a similar manner to that used previously. We will also show the necessary and sufficient conditions for convergence. Secondly, we will introduce a method that can be used to estimate the parameters necessary to find equilibria in the model. This empirical model, an extension of the mixed logit model, will utilize the same Type-I extreme value distribution assumptions used previously, but will not rely upon the IIA assumption necessary to use the basic multinomial logit model. This varying choice set logit (VCL: see Yamamoto, 2011) will allow for aggregate estimation of parameters to occur while also allowing regional parameters to be estimated. This method of estimation along with the notions of convergence that will allow analysis of the stochastic voting model in more complex situations.

Finally, to illustrate these methods, we will analyze the Canadian elections in 2004. Canada has a regional party which only runs in one region of the country, however, in 2004, the regional party gained seats in the Parliament. As this election is an ideal testing point for these new methods, they can tell us whether or not these new methods give logical results. From this analysis, some insight can be gained as to the way in which parties can organize themselves to maximize the number of votes received.

## **2 The Formal Stochastic Model**

Countries are usually composed of different regions. 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. 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.<sup>1</sup> In national elections, while regional parties are concerned only with voters living in their jurisdiction, national parties cater to nation wide interest and seek to represent voters across all regions of the country.

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.<sup>2</sup> In our model, however, 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 of a country that has at least one national and one regional party. We assume that 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.<sup>3</sup>

## 2.1 The general model

We model elections in a country where voters in different regions face different combinations of national and regional agents,<sup>4</sup> so that voters in different regions face different bundles of parties. We study the parties choice of position by taking into account whether the agent is a national or a regional party as well as taking into account 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 corresponds to the number of dimensions of the policy space. National parties run on the same platform in all regions

---

<sup>1</sup>In Canada, 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 the manufacturing base of Canada. In the United Kingdom, regional differences are mainly due to the existence of national identities in Wales, Scotland and Northern Ireland. **MAY BE ADD SOMETHING ELSE HERE.**

<sup>2</sup>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 possibility that regional parties can grow to become national parties in the model.

<sup>3</sup>For example, in Canada, Albertans care about natural resources and the oil sands; Québécois mainly about preserving their French culture and their laws; and Ontarians about manufacturing policies.

<sup>4</sup>We will use agent and party interchangeably throughout the paper.

of the country. Regional parties compete only in the region they seek to represent.

Let  $z_j$  represent the position of any party in the policy space  $X$ . The set of national parties is denoted by  $P_{Nat} = 1, \dots, p$  and the set of regional parties in region  $k$  by  $P_k = 1, \dots, q_k$  for  $k \in \mathfrak{R} = 1, \dots, r$  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 it is clear from the context whether we are dealing with national or regional party. Regions may have more than one regional party. When region  $k$  has no regional parties,  $P_k = \emptyset$ .

Since not all parties compete in all regions, and since voters in region  $k$  do not care about parties not competing in their region, let  $\mathbf{z}_k$  be the vector of policy positions of the parties competing in region  $k$ , i.e.,

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

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

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

Let  $n_k$  represents the number of voters in region  $k$ . The total number of votes 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  $k \in \mathfrak{R}$ , let voter  $i$ 's ideal policy be  $x_i \in X$  and her utility is given by  $u_{ik}(x_i, \mathbf{z}_k) = (u_{i1k}(x_i, z_1), \dots, u_{ij_k}(x_i, z_j))$  for  $j \in P_{Nat} \cup P_k$  where voter  $i$ 's utility from party  $j$  in region  $k$  is given by

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

Here,  $u_{ijk}^*(x_i, z_j)$  is the observable component of voter  $i$ 's associated with party  $j$  in region  $k$ .  $\lambda_{jk}$  is an exogenous valence term for agent  $j$  in region  $k$ . This valence term is common for all voters in region  $k$  and gives an estimate of the perceived "quality" of party  $j$ , that is,  $j$ 's ability to govern. 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 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  $\theta_k$ , 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  $\eta_i$  is a  $s$ -vector denoting voter  $i^{\text{th}}$  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  $\alpha_{ijk} = \alpha_{jk} + \mu_{ijk}$  where  $\mu_{ijk}$  is a random shock specific to region  $k$ . Thus, the sociodemographic valence  $\alpha_{jk}$  is the ‘‘average’’ sociodemographic valence of voters in region  $k$  for party  $j$ . These regional sociodemographic valences are independent of party positions. Note that the exogenous valence  $\lambda_{jk}$  measures an average assessment of party  $j$ 's ability to govern by voters in region  $k$ , net of their sociodemographic identity.

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

The error term,  $\epsilon_{ij}$  is assumed to be commonly distributed among all voters in the country.  $\epsilon_{ijk}$ . In this model, we assume that the cumulative distribution of the errors follows a Type-I extreme value distribution. This is not only the norm in individual choice models, it also allows the theoretical model to line up with the corresponding empirical model, making the transition to testing the theoretical model easier.

Recall that we are interested in finding parties' policy positions and that different bundles of parties compete in different regions. The analysis must then be carried out first at the regional level before moving to the national level. We begin by looking at the parties' positioning game and the election for a generic region  $k \in \mathfrak{R}$ .

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  $\epsilon$ . 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  $k$ .

Given that we have defined the errors as cumulatively coming from a Type-I extreme value distribution, the probability  $\rho_{ijk}(\mathbf{z}_k)$  has a multinomial logit specification and can be estimated. Thus, given  $\mathbf{z}_k$ , the probability that  $i$  votes for  $j$  in region  $k$  is given by

$$\rho_{ijk} \equiv \rho_{ijk}(\mathbf{z}_k) = \frac{\exp[u_{ijk}^*(x_i, z_j)]}{\sum_{h=1}^{p+q_k} \exp[u_{ihk}^*(x_i, z_h)]} = \frac{1}{\sum_{k=1}^{p+q_k} \exp[u_{ihk}^*(x_i, z_h) - u_{ijk}^*(x_i, z_j)]} \text{ for all } j \in P_{Nat} \cup P_k \quad (2)$$

where to simply notation we take the dependence of  $\rho_{ijk}$  on  $\mathbf{z}_k$  as understood.

Since voters decisions are stochastic in this framework, parties cannot perfectly anticipate how voters will vote. We assume however that they can estimate their expected vote shares. With different bundles of parties competing in different regions, agents can estimate their expected regional vote share in each region  $k \in \mathfrak{R}$  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.,

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

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

In addition, national parties must also take into account that their vote share depends

on all voters in the country. However, due to the presence of regional parties in some regions of the country 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 of voters in region  $k$ ,  $\frac{n_k}{n}$ , i.e.,

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

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 does not add to 1.

We assume that agents choose their position to maximize their expected vote share given the positions of all other parties. We seek equilibria of the model where each party maximizes its votes. For the purposes of this paper, when we talk about an equilibrium, we are referring to a local Nash equilibria (*LNE*). This definition of equilibrium relies on each party maximizing its expected vote share given the position of the other parties taking into account that national parties compete in all regions and regional parties in only a single region.

A vector of positions,  $\mathbf{z}_{Nat}^* \equiv \bigcup_{k=1}^r \mathbf{z}_k^*$ , is said to be a LNE if  $\forall j \in P_{Nat} \cup \bigcup_{k=1}^r P_k$ ,  $\mathbf{z}_j^*$  is a critical point of the vote share function of party  $j$ . Moreover, 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 MNL coefficients we simulate these models and then relate any vector of party positions,  $\mathbf{z}_{Nat}$ , to a vector of vote share functions  $V(\mathbf{z}_{Nat}) = (V_1(\mathbf{z}_{Nat}), \dots, V_p(\mathbf{z}_{Nat}), V_{11}(\mathbf{z}_1), \dots, V_{q_1 1}(\mathbf{z}_1), \dots, V_{r1}(\mathbf{z}_r), \dots, V_{rq_r}(\mathbf{z}_r))$ , predicted by the particular model with  $p$  national parties and  $q_1, \dots, q_r$  regional parties where  $\mathbf{z}_k$  for  $k \in \mathfrak{R}$  is the vector of party positions restricted to those competing in region  $k$ .

In order to show that parties' positions are LNE, we need to show that given  $\mathbf{z}_{Nat} \equiv \bigcup_{k=1}^r \mathbf{z}_k$ , all agents are maximizing their vote functions. Thus, we need to show that the first derivative of the vote function of party at  $\mathbf{z}_{Nat}$  is equal to zero and then examine whether the Hessian matrix of second derivatives of the vote share function of each party at  $\mathbf{z}_{Nat}$  are at a local maximum. To do so we assess the eigenvalues of the Hessian of these vote share functions.

Given that national and regional parties face different electorates, we must study their decisions separately. For regional parties we must take into account the region in which the party competes in the election.

We begin our analysis for a typical region. In region  $k$ , given the vector of policy position  $\mathbf{z}_{Nat} \equiv \bigcup_{k=1}^r \mathbf{z}_k$ , since the probability that voter  $i$  votes for party  $j$ , either a national or a regional party, is given by (2), the impact of a *marginal* change in  $j$ 's position on the probability that  $i$  votes for  $j$  in region  $k$  is then

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

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,  $\beta_k$ ; on how likely is  $i$  to vote

in region  $k$  for  $j$ ,  $\rho_{ijk}$ , 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 only concerned 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

$$\frac{dV_{jk}(\mathbf{z}_{Nat})}{dz_j}\Big|_{z_j} = \frac{1}{n_k} \sum_{i \in N_k} \frac{d\rho_{ijk}}{dz_j} = \frac{1}{n_k} \sum_{i \in N_k} 2\beta_k \rho_{ijk}(1 - \rho_{ijk})(x_i - z_j) = 0 \quad (6)$$

where the third term follows after substituting in (5).

The FOC for party  $j$  in (6) is satisfied when

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

so that the *candidate* for party  $j$ 's vote maximizing policy in region  $k$  is

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

Here  $\alpha_{ijk}$  represents the weight that in region  $k$  party  $j$  gives to voter  $i$ . Thus, the importance that  $j$  gives to  $i$  when choosing its candidate vote maximizing policy depends on how likely is  $i$  to vote for  $j$  and for any other party in region  $k$  relative to all the other voters in region  $k$ .<sup>5</sup> Note that  $\alpha_{ijk}$  may be non-monotonic in  $\rho_{ijk}$ . To see this note that if we exclude voter  $i$  from the denominator of  $\alpha_{ijk}$  and  $\sum_{a \in N_k - i} \rho_{ajk}(1 - \rho_{ajk}) < \frac{2}{3}$  then  $\alpha_{ijk}(\rho_{ijk} = 0) < \alpha_{ijk}(\rho_{ijk} = 1) < \alpha_{ijk}(\rho_{ijk} = \frac{1}{2})$ . In this case, a voter who will vote for sure for  $j$  will receive a lower weight in  $j$ 's candidate position than a voter who will only vote for  $j$  with probability  $\frac{1}{2}$ . On the other hand,  $\alpha_{ijk}$  increases in  $\rho_{ijk}$  when  $\sum_{a \in N_k - i} \rho_{ajk}(1 - \rho_{ajk}) > \frac{2}{3}$ , since the voter who votes for sure for  $j$  receives a greater weight than one voting for  $j$  with probability  $\frac{1}{2}$ .

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

---

<sup>5</sup>Suppose for example that all voters in region  $k$  are equally likely to vote for  $j$ , say with probability  $v$ , then the weight party  $j$  in region  $k$  gives to voter  $i$  in its vote maximizing policy is  $\frac{1}{n_k}$ . In this case, the weight  $j$  gives each voter is just the inverse of  $k$ 's population size.

average of the ideal policies of the voters in region  $k$  where voter  $i$ 's ideal is weighted by  $\alpha_{ijk}$ .

National parties face 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,

$$\frac{dV_j(\mathbf{z}_{Nat})}{dz_j}\Big|_{z_j} = \sum_{k \in \mathfrak{R}} \frac{n_k}{n} \frac{dV_{jk}(\mathbf{z}_k)}{dz_j}$$

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

$$\begin{aligned} \frac{dV_j(\mathbf{z}_{Nat})}{dz_j}\Big|_{z_j} &= \sum_{k \in \mathfrak{R}} \frac{n_k}{n} \frac{dV_{jk}(\mathbf{z}_k)}{dz_j} = \frac{1}{n} \sum_{k \in \mathfrak{R}} \sum_{i \in N_k} \frac{d\rho_{ijk}}{dz_j} \\ &= \frac{1}{n} \sum_{k \in \mathfrak{R}} \sum_{i \in N_k} 2\beta_k \rho_{ijk} (1 - \rho_{ijk})(x_i - z_j) = 0 \end{aligned} \quad (8)$$

The FOC for national party  $j$  in (6) is satisfied when

$$\sum_{k \in \mathfrak{R}} \sum_{i \in N_k} \beta_k \rho_{ijk} (1 - \rho_{ijk})(x_i - z_j) = 0$$

so that national party  $j$ 's vote maximizing policy is

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

Note that,  $\delta_{ijk}$  can be expressed as a function of  $\alpha_{ijk}$  in (7) as follows. Multiply the numerator and denominator of  $\alpha_{ijk}$  by  $\beta_k$  and by  $\sum_{k \in \mathfrak{R}} \sum_{i \in N_k} \beta_k \rho_{ijk} (1 - \rho_{ijk})$ , and rearrange terms to get

$$\alpha_{ijk} \equiv \frac{\sum_{k \in \mathfrak{R}} \sum_{i \in N_k} \beta_k \rho_{ijk} (1 - \rho_{ijk})}{\sum_{i \in N_k} \beta_k \rho_{ijk} (1 - \rho_{ijk})} \frac{\beta_k \rho_{ijk} (1 - \rho_{ijk})}{\sum_{k \in \mathfrak{R}} \sum_{i \in N_k} \beta_k \rho_{ijk} (1 - \rho_{ijk})}.$$

So that using  $\delta_{ijk}$  in (9), we have

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

We can re-write this expression as

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

Here  $\theta_{jk}$  is the weight given by national party  $j$  to region  $k$  relative to all regions in the country. This weight is the policy weighted aggregate probability that voters in region  $k$  vote for  $j$  or for any other party relative to all the regions in the country. Note that  $\theta_{jk}$  may be non-monotonic in  $\sum_{i \in N_k} \beta_k \rho_{ijk} (1 - \rho_{ijk})$ . To see this, let  $A_{jk} = \sum_{i \in N_k} \beta_k \rho_{ijk} (1 - \rho_{ijk})$  and note that if we exclude region  $k$  from the denominator of  $\theta_{jk}$  and  $\sum_{a \in \mathfrak{R}-k} A_{ja} < \frac{2}{3}$  then  $\theta_{jk}(A_{jk} = 0) < \alpha_{ijk}(A_{jk} = 1) < \alpha_{ijk}(A_{jk} = \frac{1}{2})$ . In this case, a region who will vote for sure for  $j$  will receive a lower  $\theta_{jk}$  weight in in  $j$ 's candidate position than a region who will only vote for  $j$  with probability  $\frac{1}{2}$ . On the other hand,  $\theta_{jk}$  is increasing in  $A_{jk}$  when  $\sum_{a \in \mathfrak{R}-k} A_{ja} > \frac{2}{3}$ , then the region who votes for sure for  $j$  receives a greater  $\theta_{jk}$  weight than one voting for  $j$  with probability  $\frac{1}{2}$ .

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

$$z_j^C = \sum_{k \in \mathfrak{R}} \sum_{i \in N_k} \theta_{jk} \alpha_{ijk} x_i = \sum_{k \in \mathfrak{R}} \theta_{jk} \sum_{i \in N_k} \alpha_{ijk} x_i$$

After substituting  $z_{jk}^C$  from (7), we obtain that 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.,

$$z_j^C = \sum_{k \in \mathfrak{R}} \theta_{jk} z_{jk}^C.$$

Let  $\mathbf{z}^C = (z_1^C, \dots, z_p^C, z_{11}^C, \dots, z_{1q_1}^C, \dots, z_{1q_r}^C, \dots, z_{rqr}^C)$  denote the *candidate* for the LNE of the game, so that  $\mathbf{z}^C$  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  $\mathbf{z}^C$  restricted to region  $k$ .

To test if  $z_{jk}^C$  (correspondingly,  $z_j^C$ ) is a local maximum of regional (correspondingly national) party  $j$ 's vote function, so that  $\mathbf{z}^C$  is a LNE of the game, we need to find the corresponding Hessian of the second derivatives of the vote share of party  $j$ . Again, 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 those competing in region  $k$ , we need the second derivative of the probability that  $i$  votes for  $j$  in region  $k$  given in (2), which we find taking the derivative of (5), i.e.,

$$\frac{d^2 \rho_{ijk}}{dz_j^2} = 2\rho_{ijk}(1 - \rho_{ijk}) [2(1 - 2\rho_{ijk})\nabla_{ijk}(z_j^C) - \beta_k I]. \quad (11)$$

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 us an ‘‘overall’’ measure of how far voter  $i$  is from party  $j$  in the policy space where each dimension is weighed by  $\beta_k$ . 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,  $\nabla_{jk}^C(z_j^C)$ ,

$$\nabla_{jk}^C(z_j^C) \equiv \frac{1}{n_k} \sum_{i \in N_k} \nabla_{ijk}(z_j^C). \quad (12)$$

From (4) and taking the derivative of (6) with respect to  $z_j$ , 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

$$\begin{aligned} \mathcal{H}_{jk} &\equiv \mathcal{H}_{jk}(\mathbf{z}_k^C) = \frac{d^2 V_{jk}(\mathbf{z}_k^C)}{dz_j^2} \Big|_{z_j^C} = \frac{1}{n_k} \sum_{i \in N_k} \frac{d^2 \rho_{ijk}}{dz_j^2} \\ &= \sum_{i \in N_k} 2\rho_{ijk}(1 - \rho_{ijk}) \left[ 2(1 - 2\rho_{ijk})\nabla_{ijk}^C(z_j^C) - \frac{1}{n_k} \beta_k I \right] \end{aligned} \quad (13)$$

where the last line follows after substituting in (11) and (12).

When all  $w$  eigenvalues of the Hessian  $\mathcal{H}_{jk}$  of regional party  $j$  in region  $k$  are negative at  $z_j^C$ , then  $j$  is at a maximum of its vote share at  $z_j^C$ . If, however, all  $w$  eigenvalues of the

Hessian  $\mathcal{H}_{jk}$  are positive then at  $z_j^C$ ,  $j$  is minimizing its vote share giving  $j$  an incentive to move away from  $z_j^C$  in order to increase its vote share. If the Hessian  $\mathcal{H}_{jk}$  has both positive and negative eigenvalues then at  $z_j^C$ ,  $j$  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  $\mathcal{H}_{jk}$ . The trace of  $\mathcal{H}_{jk}$  is also given by the sum of the main diagonal elements of  $\mathcal{H}_{jk}$ . In order to be a local maximum, and thus a LNE, the eigenvalues of party  $j$  have to all be negative implying that the trace of the Hessian must be negative as well in order for the point to be a local maximum.

For parties and voters in region  $k$ , let  $\omega$  represent a dimension in the policy space  $X$ . Denote by  $z_j^C(\omega)$  and  $x_i(\omega)$  the position of party  $j$  and the ideal policy of voter  $i$  in dimension  $\omega$  respectively. The second order condition for party  $j$  along the  $\omega$  dimension is given by the  $(\omega, \omega)$  element of the Hessian of party  $j$  in region  $k$ ,  $\mathcal{H}_{jk}$ . The  $(\omega, \omega)$  element has the following form,

$$\begin{aligned} & \sum_{i \in N_k} 2\rho_{ijk}(1 - \rho_{ijk}) \left\{ 2(1 - 2\rho_{ijk}) \frac{1}{n_k} \beta_k [x_i(\omega) - z_j^C(\omega)]^2 \beta_k - \frac{1}{n_k} \beta_k \right\} \\ = & \frac{1}{n_k} \sum_{i \in N_k} 2\beta_k \rho_{ijk}(1 - \rho_{ijk}) \left\{ 2(1 - 2\rho_{ijk}) \beta_k [x_i(\omega) - z_j^C(\omega)]^2 - 1 \right\} \end{aligned} \quad (14)$$

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

To obtain the trace of the Hessian of party  $j$  in region  $k$ ,  $trace[\mathcal{H}_{jk}]$ , just add the diagonal elements given in (14) over the  $w$  dimensions of the policy space to obtain

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

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

$$\begin{aligned} \sum_{i \in N_k} \sum_{\omega=1}^w \beta_k \rho_{ijk} (1 - \rho_{ijk}) \left\{ 2(1 - 2\rho_{ijk}) \beta_k [x_i(\omega) - z_j^C(\omega)]^2 - 1 \right\} &< 0 \\ \sum_{i \in N_k} \sum_{\omega=1}^w \alpha_{ijk} 2(1 - 2\rho_{ijk}) \beta_k [x_i(\omega) - z_j^C(\omega)]^2 &< w \end{aligned} \quad (15)$$

where after some manipulation the last line follows after substituting in (7).

Define regional party  $j$ 's convergence coefficient as the left hand side of (15), i.e.,

$$c_{jk}(\mathbf{z}_{Nat}^C) \equiv \sum_{i \in N_k} \sum_{\omega=1}^w \alpha_{ijk} 2(1 - 2\rho_{ijk}) [\beta_k [x_i(\omega) - z_j^C(\omega)]]^2 \quad (16)$$

Regional party  $j$ 's convergence coefficient in region  $k$  depends on the weight given by party  $j$  to each voter in region  $k$ ,  $\alpha_{ijk}$ , on the aggregate probability that voters in region  $k$  vote for  $j$ , and on how dispersed voters in region  $k$  are in aggregate around  $j$ 's candidate position,  $[\beta_k [x_i(\omega) - z_j^C(\omega)]]^2$  which takes into account the weight that voters give to differences with  $j$ 's policies  $\beta_k$ .

We have the following result: if when competing in region  $k$ , party  $j$  locates at its critical point  $z_j^C$ , the trace of the Hessian of its vote share function,  $\text{trace}(\mathcal{H}_{jk})$ , will be negative if  $j$ 's convergence coefficient is less than the dimension of the policy space,  $w$ , i.e.,

$$c_{jk}(\mathbf{z}_{Nat}) < w \quad (17)$$

Let us now examine the necessary conditions for national party  $j$  located at its critical point  $z_j^C$  to be at a maximum of its vote share function. To do so we look at the Hessian of second derivatives of  $j$ 's vote share function in (4) by taking the derivative of (8) to obtain

$$\begin{aligned} \mathcal{H}_j &\equiv \mathcal{H}_j(\mathbf{z}_{Nat}^C) = \frac{d^2 V_j(\mathbf{z}_{Nat}^C)}{dz_j^2} \Big|_{z_j^C} = \sum_{k \in \mathfrak{R}} \frac{n_k}{n} \frac{d^2 V_{jk}(\mathbf{z}_k^C)}{dz_j^2} = \sum_{k \in \mathfrak{R}} \frac{n_k}{n} \mathcal{H}_{jk} \quad (18) \\ &= \sum_{k \in \mathfrak{R}} \frac{n_k}{n} \sum_{i \in N_k} 2\rho_{ijk} (1 - \rho_{ijk}) \left[ 2(1 - 2\rho_{ijk}) \nabla_{ijk}^C(z_j^C) - \frac{1}{n_k} \beta_k I \right] \end{aligned}$$

where the last term follows from (13) after substituting in (11) and (12).

Like in the regional case, if all  $w$  eigenvalues of the Hessian  $\mathcal{H}_j$  of national party  $j$

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

The main diagonal element of the Hessian for national party  $j$  along the  $\omega$  dimension is given by the  $(\omega, \omega)$  element of the Hessian of national party  $j$ ,  $\mathcal{H}_j$ . From (18), the  $(\omega, \omega)$  element of  $\mathcal{H}_j$  is given by

$$\begin{aligned} & \sum_{k \in \mathfrak{R}} \frac{n_k}{n} \sum_{i \in N_k} 2\rho_{ijk}(1 - \rho_{ijk}) \left\{ 2(1 - 2\rho_{ijk}) \frac{1}{n_k} \beta_k [x_i(\omega) - z_j^C(\omega)]^2 \beta_k - \frac{1}{n_k} \beta_k \right\} \\ = & \sum_{k \in \mathfrak{R}} \frac{n_k}{n} \frac{1}{n_k} \sum_{i \in N_k} 2\beta_k \rho_{ijk}(1 - \rho_{ijk}) \left\{ 2(1 - 2\rho_{ijk}) \beta_k [x_i(\omega) - z_j^C(\omega)]^2 - 1 \right\} \quad (19) \end{aligned}$$

To obtain the trace of the Hessian of national party  $j$ ,  $trace[\mathcal{H}_j]$ , just add the diagonal elements given in (19) over the  $w$  dimensions of the policy space to obtain

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

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

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

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,  $trace(\mathcal{H}_j) < 0$ , iff

$$\sum_{\omega=1}^w \sum_{k \in \mathfrak{R}} \sum_{i \in N_k} \theta_{jk} \alpha_{ijk} 2(1 - 2\rho_{ijk}) \beta_k [x_i(\omega) - z_j^C(\omega)]^2 < w \quad (20)$$

Define national party  $j$ 's convergence coefficient as the left hand side of (20), i.e.,

$$c_j(\mathbf{z}_{Nat}^C) \equiv \sum_{\omega=1}^w \sum_{k \in \mathfrak{R}} \sum_{i \in N_k} \theta_{jk} \alpha_{ijk} 2(1 - 2\rho_{ijk}) \beta_k [x_i(\omega) - z_j^C(\omega)]^2 \quad (21)$$

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

National party  $j$  will want to stay at its critical point  $z_j^C$  if it is maximizing its vote share, that is, if

$$c_j(\mathbf{z}_{Nat}) < w \quad (22)$$

This says that national party  $j$  will remain at its critical position if its convergence coefficient is less than 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 (16), we can re-write (21) as

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

In a model in which parties compete at the national and regional levels, 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 the convergence coefficient  $j$  faces 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 than 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 were only national parties, and considered conditions under which the national mean vector  $\mathbf{z}_{Nat} = (0, \dots)$  could be a LNE. In this earlier model there is a single  $\beta$ , and we note that the term  $\frac{1}{n} \sum_{i \in N} [x_i(\omega) - \mathbf{0}]^2$  is in fact the total electoral variance about the origin, Writing this as  $\sigma^2$  gives the necessary condition for convergence as

$c(\mathbf{z}_{Nat}) = 2\beta(1 - 2\rho_1)\sigma^2 < w$ , where  $\rho_1$  is the probability that a generic voter chooses the lowest valence party,1, when all parties are at the origin.

In this paper, we make two key departures from previous papers that have used this stochastic vote model. First, and certainly the most important departure, we intend to assess convergence in a model where the position vector of interest does not have all of the parties at the joint aggregate electoral origin. As explained before, in cases where there are regional parties that do not run in all parts of an electorate, there is no incentive for these agents to locate at the overall electoral mean. Rather, in line with other median voter resultss, these parties have incentives to locate at their respective electoral means, meaning that they position themselves on the ideal point of the average voter that actually has the choice to vote for that party. Thus, should we find that parties in an electoral system converge to the electoral mean in equilibrium, we should find that parties that run in all regions of an electorate converge to the joint electoral mean and regional parties converge to their respective regional electoral means. Previous papers have adjusted the scale of the policy space such that the electoral mean corresponds to the origin of the policy space and this allowed for some convenient cancelation to occur in proofs. For the purposes of this paper, though, we cannot make those cancelations and, thus, we are assessing convergence for a general vector of party positions rather than a zero vector. Second, we assume a second kind of valence, an individual valence, that was not previously included in utility equation. We intend to assess convergence to the mean given these individual valence measures as well, showing proofs including these variables.

The first derivative of  $V_j(z)$  with respect to one dimension of the policy space is:

$$\frac{dV_j(\mathbf{z})}{dz_j} = \frac{2\beta}{n} \sum_{i=1}^n (z_j - x_i) \rho_{ij} (1 - \rho_{ij})$$

Of course, a LNE has to be at a critical point, so all the set of possible LNE can be obtained by setting this equation to 0. Note that this derivative is somewhat different than that from earlier works as we do not assume that  $\rho_{ij}$  equals  $\rho_j$  (being independent of  $i$ ). This is due to the fact that we do not assume that all parties are located at the electoral mean.

This result is important in a couple of ways. First, we see that the first derivative does not rely on  $\lambda_j$  or  $\alpha_{ij}$  in any way aside from the calculation of the probability,  $\rho_{ij}$ , that an individual  $i$  votes for party  $j$ . This is an encouraging result because any resulting measures that assess convergence (i.e. the convergence coefficient) will not depend on the individual level valences. Previously, Schofield(2007) only showed that the convergence coefficient could be calculated when we assume a common valence for agent  $j$  across all members of an electorate. This finding allows us to expand the convergence coefficient notion to include these individual level valences as long as they are exogenous of a voter's ideal point. Second, after doing some simple algebra, it is easy to see that when a party locates at its respective electoral mean, the equation always equals zero, meaning that it is always at a critical point. This is also a good result, because it gives further support to the idea that the electoral mean is always a possible LNE.

To test if a critical point is a local maximum in the vote function, thus a LNE, we need a second order condition. The Hessian matrix of second derivatives is a  $w \times w$  matrix defined as follows:

- Let  $v_t = (x_{1t}, x_{2t}, \dots, x_{nt})$  be the vector of the  $t^{th}$  coordinates of the positions of the  $n$  voters and let  $z_j = (z_{1j}, z_{2j}, \dots, z_{nj})$  and  $\langle v_t - z_{tj}, v_s - z_{sj} \rangle$  be the scalar product, with  $\Delta_0 = [\langle v_t - 0, v_s - 0 \rangle]$  the electoral covariance matrix about the origin. Then

diagonal entries of the Hessian for candidate  $j$  have the following form:

$$\frac{1}{n} \sum_{i=1}^n 2\beta(\rho_{ij})(1 - \rho_{ij})(2\beta(x_{it} - z_{tj})^2(1 - 2\rho_{ij}) - 1)$$

- The off diagonal elements have the following form:

$$\frac{1}{n} \sum_{i=1}^n 4\beta^2(x_{is} - z_{sj})(x_{is} - z_{tj})\rho_{ij}(1 - \rho_{ij})(1 - 2\rho_{ij})$$

- where  $s \neq t$ , and  $s = 1, \dots, w$ , and  $t = 1, \dots, w$ ,

Given this matrix, if all  $w$  eigenvalues of the Hessian are negative given  $\mathbf{z}$ , then we can say that the position of interest is a LNE.

Unlike previous models of this sort, there is no characteristic matrix that the Hessian can be reduced to in order to assess whether or not a point is a local Nash equilibria. Thus, for the proper second order test, the eigenvalues of the Hessian must be found. However, as in earlier works, a reduced equation can be used to find a convergence coefficient, a unitless measure of how quickly the second derivative is changing at a given point. This convergence coefficient can be viewed substantively as a measure of how much a rational, vote-optimizing party is attracted to a certain position. As the coefficient becomes large, the party is repelled from the position.

We know that the trace of the Hessian is equal to the sum of the eigenvalues associated with the matrix. In order to be a local maximum, and thus a LNE, the eigenvalues have to all be negative. Thus, the trace of the Hessian must be negative as well in order for the point to be a local maximum. Given the equation for the main diagonal elements, we can see that it relies on  $\beta$ ,  $\rho_{ij}$ , and the squared distance between the individual's ideal point on one dimension and the party's position on the same dimension. As  $\beta$  and  $\rho_{ij}$  are necessarily positive, the only way in which the second derivative can be negative is if  $2\beta(x_i - z_i)^2(1 - 2\rho_{ij})$  is greater than 1. Thus, this is the value of interest when trying to assess whether or not a point is a local maximum. This value can be viewed as the measure of how fast the probability that voter  $i$  votes for party  $j$  changes as the party makes small moves. We reason that the mean of  $2\beta(x_i - z_i)^2(1 - 2\rho_{ij})$  over all voters is

an equivalent concept to the convergence coefficient that does not rely on parties being positioned at the electoral origin. However, this is only for one dimension, so the full definition of the convergence coefficient is:

$$c(\mathbf{z}) = \frac{1}{n} \sum_{i=1}^w \sum_{j=1}^n 2\beta(x_{it} - z_{tj})^2(1 - 2\rho_{ij})$$

In words, the convergence coefficient is equal to the sum of mean values of

$$2\beta(x_i - z_i)^2(1 - 2\rho_{ij})$$

over all individuals in the electorate for each dimension of the policy space. This notion is supported by the fact that when all parties do locate at the electoral origin, this definition of the convergence coefficient is equivalent to the definition provided in Schofield (2007).

Given this definition of the convergence coefficient, we can derive necessary and sufficient conditions for convergence to a given vector of party positions. Given a vector of party positions, a sufficient condition for the vector being a local Nash equilibrium is that  $c(\mathbf{z}) < 1$ . If  $c(\mathbf{z})$  is less than 1, then we can guarantee that the second derivatives with respect to each dimension are less than 0. This eliminates the possibility that the party is located at a saddle point. A *necessary* condition for convergence to the vector of interest is that  $c(\mathbf{z}) < w$ . However, for the position to be a LNE, each second derivative has to be negative. Thus, each constituent part of  $c(\mathbf{z})$  must be less than 1.

It is important to note that a convergence coefficient can be calculated for each party in the electoral system. Previously, given that all of the parties have been attempting to optimize over the same population, an assumption could be made that the highest convergence coefficient would belong to the party which had the lowest exogenous valence. However, with the slight restructuring of the model to include individual level valences and parties which run in singular regions, as  $\rho_j$  can no longer be reduced down to a difference of valences, we can no longer make the assumption that the lowest valence party will be the first to move away from the mean should that be equilibrium behavior. In fact, given that there are multiple definitions of valence in the equation and multiple values of

these valences for each region, a notion of lowest valence party becomes very difficult to define. Thus, the convergence coefficient should be calculated for each party to ensure a complete analysis of convergence behavior. Then the party with the highest convergence coefficient represents the electoral behavior of the system. Thus, for an electoral system, the convergence coefficient is:

$$c(\mathbf{z}) = \arg \underset{p}{c_p}(z)$$

In summary, the method for assessing whether or not a vector of party positions is a LNE is as follows:

1. Define  $z^*$ , or the vector of party positions in the policy space.
2. Check that each party position meets the first order condition given the other party positions:

$$\frac{dV_j(z)}{dz_j} = \frac{2\beta}{n} \sum_{t=1}^w \sum_{i=1}^n (x_i - z_j) \rho_{ij} (1 - \rho_{ij}) = 0$$

- Note that each party's respective electoral mean is a position that is always a critical point in the vote function.

3. Define the Hessian,  $C_j(\mathbf{z})$  for each party position as follows:

- diagonal entries are

$$\frac{1}{n} \sum_{i=1}^n 2\beta(\rho_{ij})(1 - \rho_{ij})(2\beta(x_{it} - z_{tj})^2(1 - 2\rho_{ij}) - 1)$$

.where  $t = 1, \dots, w$ .

- The off diagonal elements have the following form

$$\frac{1}{n} \sum_{i=1}^n 4\beta^2(x_{is} - z_{js})(x_{it} - z_{jt})\rho_{ij}(1 - \rho_{ij})(1 - 2\rho_{ij})$$

4. Check the eigenvalues for each Hessian. If all of the eigenvalues are negative, the vector of positions is a local Nash equilibrium.

5. The necessary condition that the eigenvalues all be negative is that  $trace(C_j(\mathbf{z})) < \mathbf{0}$ . Since  $\beta(\rho_{ij})(1 - \rho_{ij}) > 0$  this reduces to:  $\sum_{t=1}^w \sum_{i=1}^n 2\beta(\rho_{ij})(1 - 2\rho_{ij})(x_{itw} - z_{tj})^2 < w$ .
6. In two dimensions. the further sufficient condition is that  $\det(C_j(\mathbf{z})) > \mathbf{0}$ , which is equivalent to the condition that  $\sum_{t=1}^w \sum_{i=1}^n 2\beta(\rho_{ij})(1 - 2\rho_{ij})(x_{itw} - z_{iw})^2 < 1$ .
7. Calculate the convergence coefficient for each party,

$$c_j(\mathbf{z}) = \frac{1}{n} \sum_{i=1}^w \sum_{i=1}^n 2\beta(\rho_{ij})(1 - 2\rho_{ij})(x_{itw} - z_{iw})^2$$

The highence coefficient, labelled  $c(\mathbf{z})$ , represents the electoral system.

- If  $c(\mathbf{z}) > w$ , then we cannot have convergence. If , however  $c(\mathbf{z}) < 1$ , then the sufficient condition is satisfied, and the system converges to the vector of interest. If  $c(\mathbf{z}) \leq w$ , check the components of  $c_j(z)$  in dimension  $w$ , if all are less than 1, then the system converges to  $\mathbf{z}$ .
- To compare this general model with the one presented in Schofield (2007), suppose that all parties adopt the same position at the electoral mean  $\mathbf{z} = \mathbf{0}$ . Then  $\rho_{ij}$  is independent of  $i$ . We let  $\Delta_0$  be the  $w$  by  $w$  electoral covariance matrix about the origin. Then
- 

$$C_j(\mathbf{z}) = (\rho_j)(1 - \rho_j)4\beta^2(1 - 2\rho_j)\Delta_0(1 - 2\beta I)$$

( where  $I$  is the  $w$  by  $w$  identity matrix. Since  $(\rho_j)(1 - \rho_j)(2\beta) > 0$ , we can identify the Hessian with the matrix

$$C_j^*(\mathbf{z}) = [2\beta(1 - 2\rho_j)\Delta_0 - I]$$

Thus the eigenvalues are determined by the necessary condition  $trace(C_j^*(\mathbf{z})) \leq w$ , which we can write as

$$\mathbf{c} = 2\beta(1 - 2\rho_j)trace(\Delta_0) \leq w$$

It can also be shown that the sufficient condition for convergence, in two di-

mensions, is given by  $\mathbf{c} = 2\beta(1 - 2\rho_j)\text{trace}(\Delta_0) < \mathbf{1}$ .

### 3 Estimation Strategies Given Varying Party Bundles

In order to utilize the stochastic election model proposed above, we need to have measures of valence, both aggregate and individual, for each party in the system, and an estimation of  $\beta$  along with the data in order to analyze equilibrium positions within the system. Typically, given the assumptions of the model, it is an easy translation of data to conditional logit model to equilibrium analysis. However, this is only true when all of the voters exist in one region. In other words, this only works when all voters vote with the same bundle of alternatives on the ballot. However, as shown in the beginning, when there are regional parties in a country which only run in one region, and are thus on the ballot for only a fraction of members of an electorate, the situation quickly becomes more complicated.

The reason that a new method is necessary is that multinomial logit models are reliant upon the assumption of independence of irrelevant alternatives. Simply put, IIA is a statement that requires that all odds ratios be preserved from group to group, even if the choice sets are different.

1. When IIA is violated, the multinomial logit specification is incorrect if we want to do any estimation procedures with this data.

Yamamoto(2011) proposed an appropriate model, called the *varying choice set logit model*.(VCL) This model, which follows the same specification as the typical multinomial logit model when Type-I extreme value errors are assumed, is the same as used above to derive the convergence coefficient: that is:

$$\rho_{ij}(\mathbf{z}) = \frac{\exp(u_{ij}^*(x_i, z_j))}{\sum_{k=1}^p \exp(u_{ik}^*(x_i, z_k))}$$

Thus the framework of the formal model and the empirical model still match, allowing easy transition from empirical estimations of parameters to analyzing the equilibria of the system given the parameters.

The VCL differs from typical logistic regression models, though, by not relying on the IIA assumption. This is done by allowing there to be individual logistic regression

models for each choice set type then aggregating these estimates to make an aggregate estimate of valence for the entire electorate. In this case, each choice set type is seen as a region, as each region has a different bundle of parties offered to voters. In these models, we can assume that parameters are common to all regions in an electorate or that the parameters have values that are region specific. For example, in our model, we assume that  $\beta$  is common to all members of the electorate regardless of region. On the other hand, we assume that both types of valence 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.

The VCL model uses random effects for the individual choice set types, meaning that for each individual type of choice set in an electorate, we estimate the parameters of interest for the individuals within that choice set. Then, using these estimates, we assume that these individual estimates come from their own distribution, and we use that to determine the best aggregate estimate for a parameter within the model. For our model, we assume the following specification for the observed utility gained by voter  $i$  from voting for party  $j$ :

$$u_{ij}^*(x_i, z_j) = \lambda_j + \beta \|z_j - x_i\| + \mu_{jr} + \xi_{jrs}$$

where  $\lambda_j$  is the aggregate estimate of the exogenous valence of party  $j$  and  $\beta$  and Euclidian distance between voter and party has the same interpretation as within the formal model.  $\mu_{jr}$  is the added utility over the aggregate valence that the average individual from region  $r$  get for voting for party  $j$  and  $\xi_{jrs}$  is the added utility over  $\mu_{jr}$  that the average member from sociodemographic group  $s$  gets from voting for party  $j$ . This clearly hierarchical specification of valence lends itself very well to the VCL model. As with typical logit models, the probability that voter  $i$  votes for party  $j$  follows the typical logit specification, which states that the probability that the voter votes for party  $j$  is the ratio of the exponentiated utility of voting for  $j$  to the sum of the utility gained for voting for each party. This model clearly lines up with the formal model specified before and makes the

VCL a very attractive choice when attempting to estimate parameters from an electorate with a clear regional structure.

Using the VCL, however, places a few light assumptions on the model, as any estimation procedure does. First, given the structure of the utility equation, we assume that  $\beta$  is common over all members of the electorate, regardless of region or sociodemographic group. This is not a departure from previous papers which have utilized this assumption. This simply means that individuals only differ in how they view each of the parties and not how much weight they apply to the differences between their ideal points and the parties' ideal points. Second, by virtue of the usage of random effects, this model assumes that each of the regional and sociodemographic group random effects are orthogonal to the other covariates in the model. Simply put, we assume that these random effects for each person are independent of one's position within the policy space. Third, by virtue of our usage of 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 can be troublesome in some electoral systems where parties frequently do not remain on the same ballots from year to year. However, many electoral systems with regional parties have parties which are historically bound to one region or another. Thus, when we assume that parties historically choose to run in a region, this model is appropriate. When all three of these assumptions are met by the electorate of interest the VCL is a flexible choice of 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 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. Similarly, mixed logit does not allow the researcher to estimate choice set specific values of parameters, thus VCL is more efficient and informative. 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, as the errors are absorbed in the error matrix and, thus, the IIA itself is absorbed. However, as with the mixed logit, the individual regional values are often of as much interest as the parameter values, so the mixed probit is essentially discarding information that the researcher may find useful. Thus, we opt to use the VCL method when examining the behavior of parties in an electorate with party choice sets that vary over the electorate.

The structure of the VCL 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 effects 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  $\beta$ ,  $\lambda_j$ , 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 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 (as 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 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.

## **4 Application to Canadian Elections**

In recent history, 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. The Liberal Party gained the most seats (135 seats) and the largest percentage of the vote (36.7 percent), however it failed to gain a majority of the seats in Parliament and needed to form a coali-

tion government in order to control the legislature. Paul Martin and the Liberals initially formed a coalition with the New Democratic Party (NDP), a liberal party whose support increased from the 2000 elections, in order to control government (19 seats, 15.7 percent). The Liberal Party's main opponent was the newly formed Conservative Party of Canada, the party formed by the merger of the Alliance Party and the Progressive Conservative party, which significantly chipped into the Liberal's vote share. After splitting support in the 2000 elections, the merger of the two parties gave the Conservative Party hope of controlling the Canadian government. Given exposure of scandal within the Liberal Party, the Conservative Party and the Liberal Party were neck and neck in the weeks leading up to the elections. However, the relative inexperience of the new party led to key mistakes prior to the elections and the Conservative Party was not able to garner a seat majority and was not able to form a coalition to control government.

Perhaps the most interesting aspect of the 2004 Canadian elections was Quebec's regional party, Bloc Quebecois (BQ). The BQ only ran in Quebec and, thus, was only on the ballot for approximately twenty percent of Canadians. However, their support within the region was overwhelming, with nearly fifty percent of Quebec voters voting for the party. This strong showing put quite a dent in the Liberal Party's showing within the region and made the BQ a significant player in the Canadian parliament (54 seats, 12.4 percent). Similarly, while not quite on the scale of the BQ, the Green Party was another small party which undoubtedly played a part in reducing the vote share of the Liberal Party. Though support for the party increased in the 2004 elections, its small initial voter base kept it from receiving any seats within parliament. However, it did gain a significant portion of votes in the election (0 seats, 4.3 percent).

To study the 2004 Canadian election we used the survey data for Canada collected by Blais et al. (2006). Table 1 shows vote shares within the sample and the overall vote shares. The similarity between these two sets of shares suggests that the sample is fairly representative of the Canadian electorate. Table 1 also has columns for those voters within Quebec, as Bloc Quebecois only ran within Quebec.

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

	Actual	Sample - All	Sample - Quebec
Liberal	36.71	34.34	25.13
NDP	15.65	18.45	8.02
Conservative	29.66	31.55	9.01
Green	4.29	3.71	2.68
BQ	12.42	11.95	55.08

Table 1: Actual and Sample Vote Percentages

"decentralization." The 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 Quebec 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 given in Table 3, we computed the value for 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 electoral 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 axis (xaxis), 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 axis (y axis), and could be regarded as opposed to decentralization.

The survey asked voters which party they would be voting for, so we estimated party positions as the mean of voters for that party. The party positions in the policy space are

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 2: Survey Items

given by the vector:

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

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 to the left on the social access while the Conservative party lies in the upper right quadrant, as shown in Figure 1. Figure 1 also shows the distribution of voters in Canada. From this, we see that most voters have a moderately leftist view on social issues and are fairly

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

Table 3: Weighting Coefficients for Canada

evenly split on decentralization issues, with most voters lying right in the middle. In Figure 1, the "Q" represents the electoral mean within Quebec, which is noticeably left of the overall electoral mean. Figure 2 shows the voter distribution for Quebec only. The majority of voters in Quebec advocate more liberal social policies than the average voter in Canada. Similarly, voters in Quebec tend to want more decentralization of government, as Quebec has a strong regional identity and wants to maintain its somewhat independent state. This, along with the differences that are easily seen from the two plots, are evidence that the two regions have strong regional identities.[Insert Figures 1 and 2 here]

The survey also collected sociodemographic data. For each respondent, sex, age, and education level were recorded. Age was divided into four categories: 18-29, 30-49, 50-65, 65 and older. Education was divided into three categories: No High School Diploma, High School Diploma but No Bachelors, Bachelors or Higher. 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 in 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 it is always a good idea to examine the relationships between the variables and see if it makes sense to 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 varying choice set logit proposed earlier, we estimate  $\beta$  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. We ran each Gibbs sampler for around 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 shown in Table 4. We show the VCL estimates of the parameter values and the corresponding 95 percent credible intervals. In this example, we use the Liberal Party as the base group, thus their valence is always restricted at 0. For the model, we report  $\beta$  and the aggregate valences first. We then report the regional effect for each party. 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 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.

From this model, we can see a number of things. First, as would have been predicted before running the model, the Liberal Party is the highest valence party in Canada outside of Quebec. However, the Conservative Party is almost equivalent in valence level. By

	<b>Canada(Aggregate)</b>	<b>Outside Quebec</b>	<b>Quebec</b>
$2*\beta$	.256 (.220,.294)	.267 (.227,.310)	.231 (.152,316)
$2*\lambda_{NDP}$	-.593 (-1.96,659)	-.556 (-.767,-.348)	-1.200 (-1.79,-.651)
$2*\lambda_{CP}$	-.139 (-1.180,.823)	-.024 (-.225,.175)	-.316 (-1.796,-.651)
$2*\lambda_{GPC}$	-1.775 (-3.294,.264)	-2.233 (-2.644,-1.855)	-2.310 (-3.262,-1.499)
$2*\lambda_{BQ}$	.278 (-1.356,1.772)	— —	.649 (.284,1.014)
Sample Size	862	675	187
DIC	2029.291		
Gelman-Rubin	1.02		

Table 4: VCL Model for 2004 Canadian Elections

simply adding the aggregate valence to the Non-Quebec regional random effect, we can see that the two are almost equivalent in valence outside of Quebec. However, this model shows that the BQ is, in fact, the highest valence party in Canada. This makes sense, given that of the people that could actually vote for the party, nearly 50 percent of them did. This exemplifies one of the strengths of this model, which is that it accurately specifies this party as the highest valence party, even though it is only available to around 25 percent of the electorate. Thus, if we view parties as entities that look down and see a uniform electorate of members without specific regional affiliation or sociodemographic groups, then they would estimate that BQ is the highest valence party.

Outside of Quebec, as mentioned before, the Conservative Party and the Liberal Party are the highest valence parties, with almost equivalent valence. The NDP is of somewhat lower valence as the party simply does not have the same presence as its larger Liberal counterpart. However, its valence and positioning in the preference space of Canada allows it to be a significant competitor outside of Quebec. The lowest valence party outside of Quebec is the Green Party, which makes plenty of sense as it is was (and is still) more of a one-issue dimension party and fails to have mass appeal to the electorate.

Inside Quebec, BQ is the highest valence party, with an even larger valence than that estimated by the aggregate valence measure. The Liberal Party also has a strong presence

in Quebec; however, given that BQ and the Liberal Party are in similar areas of the preference space, they compete for many of the same voters and BQ simply has a stronger presence in Quebec. The Conservative Party is of somewhat lower valence within Quebec, as it fails to draw voters that instead choose to vote for BQ. The lowest valence party in Quebec is also the Green Party.

Recall that we are interested in finding where the parties will locate in the policy space in order 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 election outcome and then use this information to find their most preferred position taking into account their estimates of where other parties will locate.

One possibility is that all parties will locate at their respective electoral means, meaning that  $\mathbf{z}^*$  is as follows:

$$\mathbf{z}^* = \begin{bmatrix} & Lib. & Con. & NDP & Grn. & BQU \\ S & 0 & 0 & 0 & 0 & -1.11 \\ D & 0 & 0 & 0 & 0 & -0.08 \end{bmatrix}$$

Notice that this means that BQ will not locate at the same position as the other parties as it only runs in Quebec, so its regional mean is at the mean of voters in Quebec. Given this vector of party positions and the information about the voter ideal points, we can calculate the Hessian of the vote function for each party as well as the convergence coefficient,  $c(\mathbf{z}^*)$  for each party. For the Hessians, we are interested in the eigenvalues associated with the Hessians for each party; if they are both negative, then the Hessian is negative definite and the party location is at a local maximum. Given  $\mathbf{z}^*$ , if any of the Hessians are not negative definite, then one of the parties will not choose to locate at this position in equilibrium. Similarly, we can check the convergence coefficients to see if they meet the necessary condition for convergence. Given that any of these conditions fail, the party for which they fail will choose to move elsewhere in the policy space at equilibrium and. Given that the Green Party is the lowest valence party in both regions, as well as at

the aggregate level, we can assume that if a party is going to move, it will be the Green Party. We now examine the Hessians and  $c(\mathbf{z}^*)$  for each party.

$$\mathcal{H}_{Lib} = \begin{bmatrix} -.0365 & -.0004 \\ -.0004 & -.0705 \end{bmatrix}; \mathcal{H}_{NDP} = \begin{bmatrix} .0021 & .0012 \\ .0012 & -.0362 \end{bmatrix}; \mathcal{H}_{Con} = \begin{bmatrix} -.0326 & -.0002 \\ -.0002 & -.0676 \end{bmatrix}$$

$$\mathcal{H}_{GPC} = \begin{bmatrix} .0085 & .0085 \\ .0085 & -.0091 \end{bmatrix}; \mathcal{H}_{BQ} = \begin{bmatrix} -.1194 & .0034 \\ .0034 & -.1286 \end{bmatrix}$$

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

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

From the Hessian's and their corresponding eigenvalues, we can see that two parties will diverge from the vector of electoral means. The NDP and the Green Party both have positive eigenvalues, meaning that  $\mathbf{z}^*$  is not a vote maximizing position for them and, thus, not a LNE. It is interesting to note that both of these parties  $\mathbf{z}^*$  is a saddle point . Thus, when they choose a better position, it will still be on the mean of the decentralization axis as the second eigenvalue represents that axis.

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)<sup>6</sup>, thus we need to check the

---

<sup>6</sup>It is interesting to note that the convergence coefficient need not be positive, as is the case with  $c_{BQ}(z^*)$ . This simple indicates a particularly strong desire to stay in the given position. A negative convergence coefficient indicates a quickly changing local maximum, meaning that a small departure from

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:

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

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 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 equilibrium set of positions where no party can do any better by moving given the positions of the other party. This method is necessary given that each party can potentially be optimizing over a different portion of the electorate. In this case, while the other four parties are attempting to optimize their respective vote shares over all of Canada, BQ is only trying to optimize its vote share among those voters in Quebec. Thus, this style of optimizer is necessary for finding the optimizing positions in Canada.[ Insert Figure 3 here]

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

$$z_{opt}^* = \begin{bmatrix} & Lib. & Con. & NDP & Grn. & BQ \\ S & .0524 & .0649 & 1.099 & 2.337 & -1.069 \\ D & -.0259 & -.0264 & .0266 & .2281 & -.1290 \end{bmatrix}$$

Fortunately for our measures, the vote optimizing positions echo what we were told by 

---

this position would result in a large decrease in vote share.

	Current	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

Table 5: Vote Shares given Various  $z^*$ s

the convergence coefficients: the NDP and the Green Party have incentive to move away from the electoral mean while the other parties want to stay there. Given that these two parties are of relatively low valence, their relocation has little effect on the maximizing positions for the largest three parties. However, in accordance with the equilibrium theory of proposed by Schofield (2007), the parties locate along the same axis, with distances away from their electoral means proportional to their respective perceived valence differences.

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 current positions, at the electoral mean, 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 aggregate valences plus the regional random effects).

This table strengthens our notion that the vector of means is not a LNE as the Green Party, the BQ, and the Liberals all do better when the Green Party and the NDP locate away from the mean. As the Green Party is one of the parties that is dissatisfied with the electoral mean, it can choose to move to a more extreme position and do better. The NDP is forced to adapt and do worse than it would if the parties all located at their respective electoral means.

## Conclusion

In this paper, we proposed a method for examining the vote maximizing positions of parties in electoral systems with parties that do not run in every region. When parties do not run in every region, different voters have different party bundles at the polls and

existing theories of valence and empirical methods for estimating valence are no longer appropriate. We proposed a more generalized notion of the convergence coefficient which is able to handle any generalized vector of party positions and tell us whether or not these positions are a local Nash equilibrium for the given electoral system. We also proposed a new method for estimating the parameters necessary to utilize the convergence coefficient that does not rely on the IIA assumption. Though methods of doing so already exist, the sheer amount of information gained from the Varying Choice Set Logit makes it the ideal model to run when examining voting tendencies within complex electorates that have clear hierarchical structures.

Using these methods, we examined the 2004 Canadian elections. Using the new empirical methods, we found that even though it only ran in Quebec, a region that makes up around 25 percent of Canada's population, the Bloc Quebecois was the highest valence party in Canada in the 2004 elections. Using these empirical findings, we found that parties were not able to maximize their respective vote shares by locating at the joint electoral mean, which included BQ locating at the mean of voters in Quebec and not at the joint electoral mean. Rather, the lower valence parties were able to maximize vote shares by taking more extreme positions within the policy space. This finding is in direct contrast of widely accepted theories that political actors can always maximize their vote shares by taking positions at the electoral center.

Given the accurate outcomes of these methods, there are a number of more complex situations in which these methods can be used. First, this type of model is not limited to the two region case and can be applied to cases where there are numerous "party bundles" which arise in a nation's electorate. A region, in this case, is equivalent to a party bundle; thus, a region can be a combination of many regions (the case when a party runs in two out of three regions, for example). Similarly, in further uses of this model, it is possible to examine equilibria where parties have perfect information about each of the voters, meaning that parties know each voter's region, sociodemographic group, and ideal point. Given this information, new equilibria can be computed and differences can be examined. This further demonstrates the general nature of the new definition of the convergence coefficient and its ability to handle an even wider variety of electorate types

than previously .

## 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] BLAIS, A., FOURNIER, P., GIDENGIL, E., NEVITTE, N., EVERITT, J. (2006) "Election 2006: How big were the changes. . . really?", Working paper, Universite de Montreal.
- [5] CAPLIN, A. and NALEBUFF, B. (1991), "Aggregation and Social Choice: A Mean Voter Theorem", *Econometrica*, 59, 1–23.
- [6] 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.
- [7] DOWNS, A. (1957) *An Economic Theory of Democracy* (New York: Harper and Row).
- [8] 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 University Press.
- [9] HOTELLING, H. (1929), "Stability in Competition", *Economic Journal*, 39, 41–57.
- [10] KRAMER, G. (1978), "Existence of Electoral Equilibrium", in P. Ordeshook (ed.) *Game Theory and Political Science* (New York: New York University Press) 375–389.

- [11] MERRILL, S. and GROFMAN, B. (1999) *A United Theory of Voting* (Cambridge, U.K.: Cambridge University Press).
- [12] PATTY, J. W. (2005), "Local Equilibrium Equivalence in Probabilistic Voting Models", *Games and Economic Behavior*, 51, 523–536.
- [13] 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* (forthcoming).
- [14] PENN, E. (2003), "A Model of Far-sighted Voting" (Unpublished Paper, Harvard University).
- [15] POOLE, K. and ROSENTHAL, H. (1984), "US Presidential Elections 1968–1980", *American Journal of Political Science*, 28, 283–312.
- [16] QUINN, K., MARTIN, A. and WHITFORD, A. (1999), "Voter Choice in Multiparty Democracies", *American Journal of Political Science*, 43, 1231–1247.
- [17] RIKER, W. H. (1953). *Democracy in the United States*. New York: Macmillan
- [18] RIKER, W. H. (1964). *Federalism: Origin, Operation, Maintenance*. Boston, MA: Little Brown.
- [19] RIKER, W. H. (1987). *The Development of American Federalism*. Boston, MA: Kluwer.
- [20] RIKER, W. H. and ORDESHOOK, P. C. (1973) *An Introduction to Positive Political Theory* (Englewood Cliffs, NJ: Prentice-Hall).
- [21] SCHOFIELD, N. (2007), "The mean voter theorem: Necessary and sufficient conditions for convergent equilibrium", *Rev Econ Stud* 74, 965-980.
- [22] SCHOFIELD, N., MARTIN, A., QUINN, K. and WHITFORD, A. (1998), "Multi-party Electoral Competition in the Netherlands and Germany: A Model Based on Multinomial Probit", *Public Choice*, 97, 257–293.

- [23] TRAIN, K. (2003) Discrete Choice Methods for Simulation (Cambridge, U.K.: Cambridge University Press).
- [24] YAMAMOTO, T. (2011), "A Multinomial Response Model for Varying Choice Sets, with Application to Partially Contested Multiparty Elections." Working Paper.

## 4.1 AppendixA

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]
}}}
}
\newpage
\FRAME{ftbpFU
}{3.961in
}{3.022in
}{Opt
}{\Qcb{Distribution of voters and party positions for Canada in 2004}
}{14.newcanadaall.eps
}{\special{language "Scientific Word";type "GRAPHIC";maintain-aspect-ratio
TRUE;display "USEDEF";valid_file "F";width 3.961in;height 3.022in;depth
Opt;original-width 10.083in;original-height 7.6754in;cropleft "0";croptop
"1";cropright "1";cropbottom "0";filename 'E:/1.1.1.1.me.canada/graphics/14.newCanadaAll.ep
\FRAME{ftbpFU
}{3.961in
}{3.022in
}{Opt
}{\Qcb{Distribution of voters and party positions for Quebec in 2004}
}{14.newcanadaquebec.eps
}{\special{language "Scientific Word";type "GRAPHIC";maintain-aspect-ratio
TRUE;display "USEDEF";valid_file "F";width 3.961in;height 3.022in;depth
Opt;original-width 10.083in;original-height 7.6754in;cropleft "0";croptop
"1";cropright "1";cropbottom "0";filename 'E:/1.1.1.1.me.canada/graphics/14.newCanadaQuebec
\FRAME{ftbpFU
}{3.961in
}{3.022in
}{Opt
}{\Qcb{Vote maximizing positions in Canada 2004}
}{14.newcanadaoptimal.eps
}{\special{language "Scientific Word";type "GRAPHIC";maintain-aspect-ratio
TRUE;display "USEDEF";valid_file "F";width 3.961in;height 3.022in;depth
Opt;original-width 10.083in;original-height 7.6754in;cropleft "0";croptop

```

## 5 APPENDIX B ELECTION STATISTICS

<b>Table A3:</b> Descriptive Statistics by Region									
	<b>Canada (<math>n = 862</math>)</b>			<b>C/Q (<math>n = 675</math>)</b>			<b>Québec (<math>n = 187</math>)</b>		
<b>Variable</b>	<b>Mean</b>	<b>Median</b>	<b>SD</b>	<b>Mean</b>	<b>Median</b>	<b>SD</b>	<b>Mean</b>	<b>Median</b>	<b>SD</b>
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

<b>Table A4:</b> : Descriptive Statistics by Party						
Variable	Mean	Median	SD	Mean	Median	SD
	<b>Liberals</b>			<b>Conservatives</b>		
	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
	<b>New Democratic</b>			<b>Greens</b>		
	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
<b>Bloc Québécois (in Québec only, <math>n = 103</math>)</b>						
	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		51 1		0.50	
Education	7.31		7		2.06	

		<b>Outside Quebec (<math>n = 675</math>)</b>				<b>Quebec (<math>n = 187</math>)</b>			
		<b>Spatial (S)</b>		<b>Spatial+Socio.</b>		<b>Spatial (S)</b>		<b>Spatial+Socio.</b>	
Party	Var	Est. <sup>a</sup>	t-value	Est. <sup>a</sup>	t-value	Est. <sup>a</sup>	t-value	Est. <sup>a</sup>	t-value
	$\beta$ Soc	0.25***	11.67	0.25***	11.43	0.19***	4.49	0.21***	4.56
	$\beta$ Dec	0.48***	6.87	0.47***	6.67	0.53**	3.15	0.48**	2.78
CP	$\lambda_{CP}$	-0.02	0.15	1.07	1.77	-0.24	0.76	-0.01	0.01
	age			0.00	0.68			-0.02	1.42
	gender(f)			0.12	0.59			-0.43	0.70
	education			-0.13*	2.44			0.17	1.10
NDP	$\lambda_{NDP}$	-0.54***	5.03	1.00	1.60	-1.51***	3.76	1.35	0.87
	age			-0.02*	2.30			-0.04*	2.11
	gender(f)			0.56	0.26			-0.71	1.15
	education			-0.10	1.82			-0.03	0.16
GPC	$\lambda_{GPC}$	-2.22***	10.88	1.11	0.98	-2.26***	4.78	0.67	0.27
	age			-0.38**	2.66			-0.03	0.87
	gender(f)			-0.29	0.69			-1.90	1.63
	education			-0.19	1.78			0.06	0.26
BQ	$\lambda_{BQ}$					0.75***	3.67	2.21*	2.23
	age							-0.02	1.74
	gender(f)							-0.65	1.64
	education							-0.01	0.15
<i>LL</i>		-678		-668		-202		-195	
AIC		1365		1365		415		426	

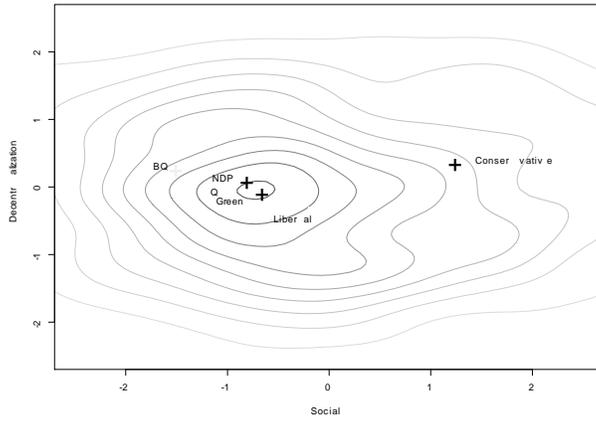


Figure 1: Distribution of voters and party positions for Canada in 2004

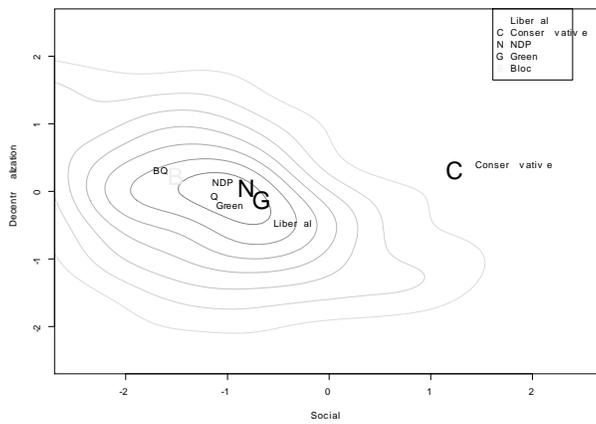


Figure 2: Distribution of voters and party positions for Quebec in 2004

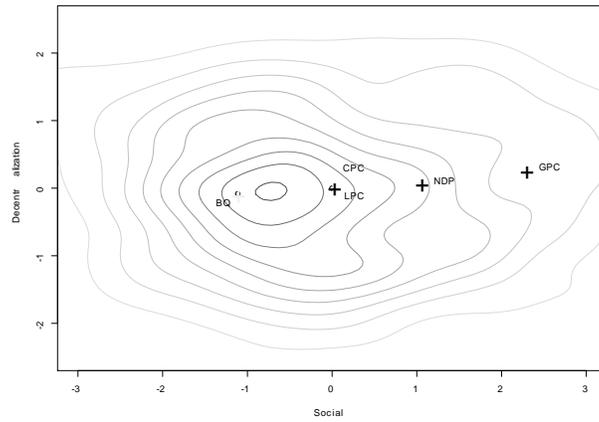


Figure 3: Vote maximizing positions in Canada 2004