

Norman Schofield · Gonzalo Caballero · Daniel Kselman *Editors*

Advances in Political Economy

Institutions, Modelling and Empirical Analysis

This book presents latest research in the field of Political Economy, dealing with the integration of economics and politics and the way institutions affect social decisions. The focus is on innovative topics such as an institutional analysis based on case studies; the influence of activists on political decisions; new techniques for analyzing elections, involving game theory and empirical methods.

Schofield · Caballero
Kselman *Eds.*



Advances in Political Economy

Norman Schofield
Gonzalo Caballero
Daniel Kselman *Editors*

Advances in Political Economy

Institutions, Modelling and
Empirical Analysis

Social Sciences / Political Science

ISBN 978-3-642-35238-6



9 783642 352386

► springer.com

 Springer

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:

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

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 ρ_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:

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

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

1. Define \mathbf{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:

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

- 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

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

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

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

The convergence 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 $c_j(\mathbf{z})$ 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 $\mathbf{z} = 0$. Then ρ_{ij} is independent of i . We let Δ_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 $\text{trace}(C_j^*(\mathbf{z})) \leq w$, which we can write as

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

It can also be shown that the sufficient condition for convergence, in two dimensions, is given by $\mathbf{c} = 2\beta(1 - 2\rho_j) \text{trace}(\Delta_0) < 1$.

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 β 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 β 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 λ_j is the aggregate estimate of the exogenous valence of party j and β and Euclidian distance between voter and party has the same interpretation as within the formal model. μ_{jr} is the added utility over the aggregate valence that the average individual from region r get for voting for party j and ξ_{jrs} is the added utility over μ_{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 β 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

553 of random effects, this model assumes that each of the regional and sociodemo-
554 graphic group random effects are orthogonal to the other covariates in the model.
555 Simply put, we assume that these random effects for each person are independent
556 of one's position within the policy space. Third, by virtue of our usage of the VCL
557 model, we assume that a party's decision to run in a specific region is exogenous
558 of its perceived success within that region. This assumption can be troublesome in
559 some electoral systems where parties frequently do not remain on the same ballots
560 from year to year. However, many electoral systems with regional parties have parties
561 which are historically bound to one region or another. Thus, when we assume
562 that parties historically choose to run in a region, this model is appropriate. When all
563 three of these assumptions are met by the electorate of interest the VCL is a flexible
564 choice of estimation procedure.

565 The reason that the varying choice set logit (VCL) is the superior method when
566 handling electorates with multiple regions is that it relaxes the IIA assumption while
567 also providing us with the most information from the model. VCL relaxes IIA by al-
568 lowing each of the parameters to be estimated within each group and allowing these
569 parameters to derive the aggregate estimation of parameters through the notion of
570 partial pooling. Partial pooling is best achieved through hierarchical modeling and
571 through the use of random effects. VCL can be viewed as a specific kind of mixed
572 logit model, meaning that the mixed logit model can be used to achieve the same
573 aggregate results. However, given the structure of VCL, parameter estimates can
574 be achieved for each choice set type (i.e. region) rather than for each individual,
575 demonstrating a significant efficiency gain over the standard mixed logit model.
576 Similarly, mixed logit does not allow the researcher to estimate choice set specific
577 values of parameters, thus VCL is more efficient and informative. Another alterna-
578 tive is the multinomial probit model, which does not rely on the IIA assumption
579 either. However, the multinomial probit model does not allow the researcher to es-
580 timate parameters at the level of the individual choice set, as the errors are absorbed
581 in the error matrix and, thus, the IIA itself is absorbed. However, as with the mixed
582 logit, the individual regional values are often of as much interest as the parameter
583 values, so the mixed probit is essentially discarding information that the researcher
584 may find useful. Thus, we opt to use the VCL method when examining the behavior
585 of parties in an electorate with party choice sets that vary over the electorate.

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

594 For applications to this model, we make a few assumptions about the underly-
595 ing distributions of the parameters of interest. We assume that β , λ_j , and the ran-
596 dom effects all have underlying normal distributions. Further, we assume that all of
597 these distributions are independent of one another. This assumption follows from
598

599 our assumptions that the variables, and thus the draws in the Gibbs sampler, are all
600 orthogonal. We could easily assume that each level of the hierarchy (aggregate, re-
601 gion, sociodemographic) comes from a multivariate normal within itself. However,
602 time spent with this model has shown that this assumption is taxing computationally,
603 adding to the amount of time it takes the Gibbs sampler to converge and yielding
604 results that are virtually indiscernible from those garnered when independence is
605 assumed. However, it is unreasonable to assume that the orthogonality assumption
606 is perfectly met. For example, in some cases, region and location within the policy
607 space are correlated (as in Canada). This assumption violation will lead to biased
608 estimators. While the bias is not large, it is certainly a cause for some concern.
609 However, this problem is easily fixed.

610 Gelman et al. (2008) utilize a method to rid random effects of the collinearity
611 which causes the estimates to be biased. They propose that the problem is solved
612 very simply by adding the mean of the covariate of interest as a predictor a level
613 lower in the hierarchy than the random effect of interest. In this case, given a spec-
614 ific party, the mean of its regional level random effects and the mean of its sociode-
615 mographic level random effects are indeed situated at the respective mean of the
616 difference of Euclidian differences between the party of interest and the base party.
617 Given that this is the covariate that will theoretically be correlated with sociodemo-
618 graphic group and region, this is the mean that we need to include as a predictor in
619 the random effects. In doing this, the researcher controls for the discrepancy as if it
620 is an omitted variable and allows the random effect to take care of its own correla-
621 tion. The normal priors in this case can still be diffuse, but the mean needs to be at
622 the specified value to fix the problem.

623 One practical note is necessary regarding the time necessary to achieve conver-
624 gence within the model. Convergence of the VCL can be quite slow given a large
625 number of choice set types and individual observations. Similarly, as random effects
626 are estimated for each party, the number of parties and the number of sociodemo-
627 graphic groups can slow down the rate at which samples are derived from the Gibbs
628 sampler. Though it is a time consuming method, the sheer amount of information
629 gained from the VCL is, thus, the best choice when it is necessary to use a discrete
630 choice model which does not rely on IIA.

631 632 633 634 **4 Application to Canadian Elections**

635
636 In recent history, Canadians have elected at least three different parties to the Fed-
637 eral legislature and 2004 was no different. However, the 2004 election in Canada
638 was significant because it yielded the first minority government for Canada since
639 1979. The Liberal Party gained the most seats (135 seats) and the largest percentage
640 of the vote (36.7 percent), however it failed to gain a majority of the seats in Parli-
641 ament and needed to form a coalition government in order to control the legislature.
642 Paul Martin and the Liberals initially formed a coalition with the New Democratic
643 Party (NDP), a liberal party whose support increased from the 2000 elections, in
644

Table 1 Actual and sample vote percentages

	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

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 "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

Table 2 Survey items

691		
692	Inequality	How much to you think should be done to reduce the gap
693		between the rich and the poor in Canada?
694		(1) much more—(5) much less
695	Women	How much do you think should be done for women?
696		(1) much more—(5) much less
697	Gun Only police/military	Only the police and the military should be allowed to have guns.
698		(1) strongly agree—(4) strongly disagree
699	Iraq War	As you may know, Canada decided not to participate in the war
700		against Iraq.
701		Do you think this was a good decision or a bad decision?
702		(1) good decision (2) bad decision
703	Left-Right	In politics, people sometimes talk of left and right.
704		Where would you place yourself on the scale below?
705		(0) left—(11) right
706	Welfare	The welfare state makes people less willing to look after
707		themselves.
708		(1) strongly disagree—(4) strongly agree
709	Standard of Living	The government should see to it that everyone has a decent
710		standard of living.
711		(1) leave people behind (2) Don't leave people
712	Quebec	How much do you think should be done for Quebec?
713		(1) much more—(5) much less
714	Moving Cross Region	If people can't find work in the region where they live, they
715		should move to where the jobs are?
716		(1) strongly disagree—(4) strongly agree
717	Federal-provincial	In general, which government looks after your interests better?
718		(1) provincial (2) no difference (3) federal
719		

720
721
722 Federal versus Provincial governments in their lives. A greater desire for decentral-
723 ization implies higher values on this axis. The questions used in the factor analysis
724 can be found in Table 2.

725 Using the factor loadings given in Table 3, we computed the value for each voter
726 along the social and decentralization dimensions. The mean and median values of
727 voters' positions along these two dimensions in Canada are at the electoral origin,
728 (0; 0). To illustrate, a voter who thinks that more should be done to reduce the gap
729 between rich and poor would tend to be on the left of the Social axis (x axis), while
730 a voter who believes that the federal government does a better job of looking after
731 peoples' interests would have a negative position on the D axis (y axis), and could
732 be regarded as opposed to decentralization.

733 The survey asked voters which party they would be voting for, so we estimated
734 party positions as the mean of voters for that party. The party positions in the policy
735 space are given by the vector:
736

Table 3 Weighting coefficients for Canada

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

$$z^* = \begin{bmatrix} S & Lib. & Con. & NDP & Grn. & BQU \\ D & -0.17 & 1.27 & -0.78 & -0.63 & -1.48 \\ & -0.38 & 0.32 & 0.05 & -0.13 & 0.23 \end{bmatrix}$$

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 axis while the Conservative party lies in the upper right quadrant, as shown in Fig. 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 evenly split on decentralization issues, with most voters lying right in the middle. In Fig. 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.

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

Fig. 1 Distribution of voters and party positions for Canada in 2004

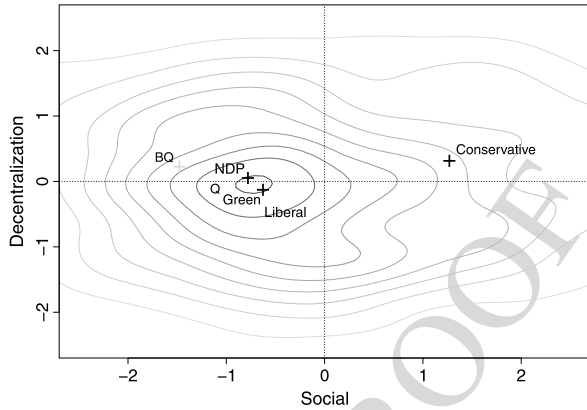
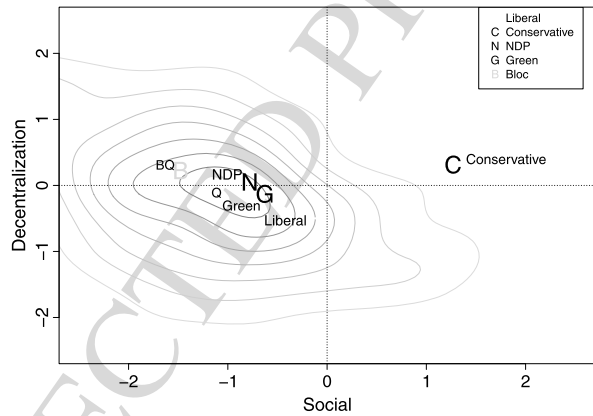


Fig. 2 Distribution of voters and party positions for Quebec in 2004



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 β and the variances 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 β 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 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