Application of the VCLmodel to the British general elections of 2005 and 2010

Elena Labzina and Norman Schofield

February 21, 2014

The situation when across the electorate individual voters may choose from different sets of parties is not unusual in many countries. However, this makes the canonical version of Schofield's approach (e.g [?]) unapplicable to the estimation of the party valences and subsequent analysis of the convergence of the party positions in the straight-forward way. The formal reason for that is the violation of the crucial assumption for the major tool, the logistical regression, the independence of the irrelevant alternatives (IIA). Substantively, the concern is that it is impossible to claim in this situation that voters would give their voices to the same party were they all given the same set of alternatives as it is in the original model.

Luckily, theoretical solutions to this obstacle already exist. [?] in their paper propose a modification of the logistic regression - the varying choice logit (VCL) and showed the example of its application to the elections in Japan. The idea of this improvement is, while calculating the estimated probabilities, to consider only the options from which an individual could actually choose. This naturally splits the sample based on their choice sets. Meanwhile, the estimates for the alternatives are aggregated over all bundles in which they presents, which produces the estimates for the whole electoral system.

However, for the case of this paper, UK, similar to the existing similar work of [?] for Canada, an additional concern is that unobserved effects (possibly with a fixed component) should be accounted for as well. As [?] note, a natural solution to this problem is the hierarchical Bayesian models, which, as the technologies develop and the computers become faster, provide a handy means to control for the random unobserved effects in the model. In this work, this tool is employed on the stage of the computation of the aggregated party valences. Unfortunately, this is still a very time-consuming task, which is understandable, given that, for instance, the simplest pure spatial model of the presented analysis makes JAGS create a graph with 413049 nodes for the following estimation procedure.

Methodologically, this work can be considered a more complex continuation of [?] paper. In that paper, the authors look at Canada, where in Quebec an additional influential party, in relation to the rest of the country, runs for the elections. Meanwhile, in UK both Wales and Scotland have their specific parties that collected a significant share of votes in the regions during the General

Elections of 2010. Scottish National Party obtained 19.9 % in Scotland, coming the second after the Labour Party, that makes up to 2.6 % at the national scale. Compared to that, the share of Playd Cymru does not look impressive on the national level, 0.6 % at all, since other small parties managed to get even more [1], however, in Wales they have their 11.14 %, which, because of the IIA, makes it necessary to account for this party while calculating the probabilities for the observations from Wales.

Substantively, our paper is the next step after the work [**?**], where the General Election of 2005 and 2010 in UK in England, Wales and England are analysed separately. In their core logistic regression they include only observations from England. Their results for 2010 show the unexpected insignificance of the valence of the Labour Party, which is easily explained, given the results of this our work. The valence on the national level becomes significant, when Scotland, where the Labour Party has a tremendous support of 42 % in 2010, is added to the analysis. It is natural to assume that the Scottish electorate significantly influences the electoral strategy of the Labour party.

Another complication in relation to [**?**] is the sample size employed, which is almost 8 times larger. The last development is the amount of the models tested: besides the spatial models with and without socio-demographical valences (that turn out to be insignificant as in the paper of [**?**]), the models with traits are estimated as well.

Important to note, that obtaining the controls for the traits themselves for UK requires certain modifications to the usual component factor analysis. This is because only individuals from Scotland and Wales report[2] their attitudes towards Salmond and Jones relatively, while all individuals report the attitudes towards Brown, Cameron, and Clegg. Since the traits for Salmond and Jones may be considered crucial, this survey design creates a need to run CFA with varying number of factors for the individuals from different regions, and, then merging the sub-samples to apply in the following hierarchical Bayeasian analysis.

One of the major limitation of the paper is that the shares of the individuals voted for the parties in the sample (Table 1 and Table 2) differ from the electoral results , especially, in the number of the supporters of Liberal Democrats. This is a consequence of the voting shares in the original data set available from British Electoral Study's webpage [3]. However, this should not be a concern, because the methodology of the evaluation of the convergence of the party positions requires only the supporters of the parties in the employed sample to be representative of out-of-sample supports, which can be assumed. What is really important for the subsequent analysis is that the model predicts the probabilities in the sample well, in other words, by the design, the concern is more about the internal than external validity, since we are considering the sample as the population. Hence the fact that the estimated models predicts the in-sample voting probabilities well is sufficient not to worry about the discrepancy between the voting results

[1]

[2] BES

[3] https://www.essex.ac.uk/bes/

and the shares of the supporters in our sample.

This justification is in a way valid as an answer to another possible critique - the extent to which the employed Bayesian models managed to converge. Despite the opinion that a partial convergence (e.g a confident convergence only in relation to a subset of the parameters and a somewhat convergence of the whole model) can not be considered a convergence at all, in the given case we should not worry too much about this if the models is good in the predictions. This remark may be considered even more legitimate if to remember that the whole Bayesian model here is a form of the adjustment and refinement of the logistic model that has been proven to converge by [**?**].

1 Formal model in the application to UK

This is a modification of the canonical Schofield's analysis for the case with varying choice sets of the parties. Because of the violation of IIA, the application of the classical logistic regression is impossible.

In UK, three major parties - the Conservative, Labour and Liberal Democrats - gained 88.2% of the votes in the General Elections of 2010. The rest of the votes are split among minor parties (at the national scale), which could have been neglected in the following analysis had all the voters the same alternatives in their electoral choice bundles, since our major interest is the investigation of the valences and convergence of the major parties. The problem appears from the two parties that run for elections only in one region, Scottish National Party in Scotland and Playd Cymru in Wales, hence in such a setting IIA is violated: for instance, we cannot assume that an individual in Wales does not vote for SNP, because he does not want, since he simply does not have such an option at all.

In the further analysis, we assume that our full set of parties consists of the three major parties (Labour(1), Conservative(2), Liberal Democrats(3)) and two parties which cause the violation of IIA (SNP(4) and Plaid Cymru (5)). Then, denoting r(i) the region of individual i[4] and $u_{i,\,j}$ the utility the individual gets from voting for party j, x_i the political position of the individual, z the position of the parties, the three possible sets of the alternatives are:

$$r(i)1 \implies \bar{u}_{i\,m(i)} = \{1,2,3\} \implies u_i(x_i, z|r(i) = 1) = \{u_{i\,1}, u_{i\,2}, u_{i\,3}\}$$

$$r(i) = 2 \implies \bar{u}_{i\,m(i)} = \{1,2,3,4\} \implies u_i(x_i, z|r(i)) = 2) = \{u_{i\,1}, u_{i\,2}, u_{i\,3}, u_{i\,4}\}$$

$$r(i) = 3 \implies \bar{u}_{i\,m(i)} = \{1,2,3,5\} \implies u_i(x_i, z|r(i) = 3) = \{u_{i\,1}, u_{i\,2}, u_{i\,3}, u_{i\,5}\}$$

The Bayesian approach enables to include the random regional effects for the three major parties ($\mu_{i\,j}$, $i = 1,2,3$, $j = 2,3$). Following [**?**], their priors are assumed normal with the zero mean and the random variance from the gamma distribution.

[4] 1 for England, 2 for Scotland, 3 for Wales

The general form of the most straight-forward version of the individual utility is:

$$u(x_i, z_j) = \lambda_j - \beta||x_i - z_j|| + \sum_{k \in m(i)} \alpha_k \tau + \mu_{jr(i)} + (\theta * \nu_i) + \epsilon_i \tag{1}$$

Here λ_j is the valence of party j, $\sum_{k \in m(i)} \alpha_k \tau$ - the traits (see for more details the section on the models with traits), $(\theta * \nu_i)$ - socio-demographic valences. Naturally, in the applied settings various time of the aggregation of the random and fixed are to be tried, possibly, by "trial and error" checking which models fit and converge the best.

Assuming that for some combinations of j and r(i) the regional-party effect is significantly different from zero, e.g $E(\mu_{jr(i)}) \neq 0$, then the *mixed valence* of party j is:

$$\lambda;\prime_j = \lambda_j + \frac{1}{n} \sum_{r(i): E(\mu_{jr}) \neq 0} n_r \mu_{jr} \tag{2}$$

Logically, the way to calculate the aggregated probabilities of the voting for the parties differs from the situation with the common choice bundle across the sample. In a way, these probabilities, considering the set of 5 parties, are tentative, since no individual in the sample chooses from all of them. However, these probabilities can be understood as a general characteristic of the each party, a predicted vote share, that incorporates two probabilities: a randomly selected observation is from a region, and, given the region, he or she chooses the party.

For the estimated probability that individual i votes for party j, $p_{i\,j} = \frac{e^{u_{i\,j}}}{\sum\limits_{k \in m(i)} e^{u_{i\,k}}}$, where $m(i)$ is the set of parties from which i chooses, then the probability of voting for j in the region of i, $r(i)$, is

$$p_{j\,r} = \frac{\sum\limits_{i \in r} p_{i\,j}}{n_r}$$

and the aggregated probabilities are

$$p_j = \frac{1}{n} \sum_{r \in m(i)} n_r \, p_{j\,r}$$

Here n_r is the number of the observations for region r and n is the total size of the sample. It is easy to be that n_j's sum up to one.

Then, the conservative estimates (assuming the zero correlation across the regions) of the standard errors for the probabilities on the national level are:

$$p_j = \frac{1}{n^2} \sum_{r \in m(i)} n_r^2 Var(\,p_{j\,r})$$

.

Meanwhile, to calculate the hessian matrix and the convergence coefficients still the individual level probabilities must be used.

The Hessian for the varying choice sets and the 2-dimensional policy space[5]:

$$H_j(z_j^*) = \begin{pmatrix} \frac{2\beta}{n'} \sum_{i=1}^{n'} p_{i\,j}(1-p_{i\,j})(2\beta(x_{i1}-z_{j1})^2(1-2p_{ij})-1) & \frac{4\beta^2}{n'} \sum_{i=1}^{n'} (x_{i2}-z_{j2})(x_{i1}-z_{j1})p_{i\,j}(1-p_{i\,j})(1-2 \\ \frac{4\beta^2}{n'} \sum_{i=1}^{n'} (x_{i2}-z_{j2})(x_{i1}-z_{j1})p_{i\,j}(1-p_{i\,j})(1-2p_{ij}) & \frac{2\beta}{n'} \sum_{i=1}^{n'} p_{i\,j}(1-p_{i\,j})(2\beta(x_{i2}-z_{j2})^2(1-2p_{ij}) - \end{pmatrix} \quad (3)$$

Important to note, $n^{'}$ is not the total number of the observations in the sample, but the number of the individuals having party j in their choice bundle, since a party is expected to maximize across those who can vote for it. Hence, for our case $n^{'}_{LAB} = n^{'}_{CON} = n^{'}_{LAB} = N$, $n^{'}_{SNP} = n_2$ and $n^{'}_{PC} = n_3$.

If the eigenvalues for each Hessian are negative, then the vector of the positions is LNE.

The convergence coefficient for each party:

$$c_j(z) = c_j^1(z) + c_j^2(z) = \frac{2\beta}{n'} \sum_{i=1}^{n'} p_{i\,j}(1-2\,p_{i\,j})(x_{i1}-z_{j1})^2 + \frac{2\beta}{n'} \sum_{i=1}^{n'} p_{i\,j}(1-2\,p_{i\,j})(x_{i2}-z_{j2})^2 \quad (4)$$

$c_j^1(z)$ and $c_j^2(z)$ are the components related to each dimension of the policy.

Each party's position may evaluated according to the following rules (e.g [?]):

- If $c_j(z) > 2$ then there is no convergence.
- If $c_j(z) < 2$, $c_j^1(z) < 1$ and $c_j^2(z) < 1$ then the system converges.

Meanwhile, the convergence coefficient for the whole electoral system:

$$c(z) = max(\{c_j(z)\}) \quad (5)$$

To check the core characteristic of the system - the general convergence - the same above criteria should be applied.

2 Data methodology

The data for this work comes from British Electoral Survey 2009-2010. All variables might be divided into 4 groups:

1. The core individual variables: the party voted and the region of the respondent;

[5] Based on [?], pp 13-14

2. Survey questions (see Table *) used to construct the individual score along 2 dimensions along which the distances are to be calculated further: the nationalism (anti-EU) and economic (anti-taxes);

3. Socio-economic valences: sex, income and education

4. Survey questions to calculate the traits (see Table *)

In relation to the full sample downloaded from the page of the survey, only the observations having 'Don't know' in the questions used for CFA were dropped. The major intent was to keep as many observations as possible to avoid unintended bias and to have non-equal samples employed for testing different models, that is why for (2), (3), and (4) the multiple imputation via the R package MICE was performed[6] (no values were missing among the variables from (1)).

On the next computational stage of the analysis, various modifications of the equation (1) were tried. Based on the results for the pure spatial model (e.g $\alpha_k = 0$ and $\theta = 0$), as a result of the unequal subsamples with the same choice sets and that j=4 and j=5 belong each only to a single choice set, the best solution was found to assume non-zero $\mu_{jr(i)}$'s for $(j, r(i)) = \{(1,2),(2,2),(1,3),(2,3)\}$ only (j=3, Liberal Democrats, is the reference party). These way of accounting for the regional effects is used in more sophisticated models as well.

3 The General Election of 2010

The Table present the comparison of the electoral results for the General Elections of 2010 in UK. It is seem that among the parties of our interest the Liberal Democrats and SNP are overrepresented in the sample, while the Conservatives and, especially, PC are under-represented. This is the feature of the sample and, for example, [?], who used the same survey, have very similar summary statistics.

[1] Based on 2010 British Election Survey campaign Internet panel data

Table 1. 2010 Election in Great Britain

	Population			Sample[1]	
Party	Vote %	Seats	Seat %	Observations	Observations %
Conservative	36.1	306	47.1	3,097	35.43
Labour	29.0	258	39.7	2,350	26.89
Liberal Democrats	23.0	57	8.8	2,384	27.28
Scottish National Party	1.7	6	0.9	210	2.40
Plaid Cymru	0.6	3	0.9	43	0.49
Others	9.6	20	2.6	656	7.51
Total	100	650	100	8,740	100

[6] The results of the CFAs and the summary statistics for (3) are almost the same with and without imputation, hence this may be considered a legitimate solution.

Being interested in the internal mechanism driving the party self-positioning, our major only requirement to the sample, given that we are going to test the convergence at the electoral mean in terms of the policy dimensions, is that the average across the full sample equals to zero, which is satisfied (see Table 3). Table 2 presents the vote shares split by the regions of the analysis.

Table 2. 2010 Election in Great Britain by Region: Voting

	All			Scotland			Wales		
Party[1]	election (%)	obs[2]	obs (%)	election (%)	obs	obs (%)	election (%)	obs	obs (%)
Con	36.1	3,097	38.31	16.7	134	16.75	26.1	110	28.50
Lab	29.0	2,350	29.07	42.0	283	35.38	36.2	137	35.49
LibDem	23.0	2,384	29.49	18.9	173	21.63	20.1	96	24.87
SNP	1.7	210	2.60	19.9	210	26.25	-	-	-
PC	0.6	43	0.56	-	-	-	11.3	43	11.14
Total	90.4	8,084	100	97.5	777	100	93.7	370	100

[1] Only major parties and region-specific parties: Con: Conservative Party; Lab: Labor Party; LibDem: Liberal Democrat Party; SNP: Scottish National Party; PC: Plaid Cymru

[2] Sample based on BES 2010 containing only observations of those voted for the 5 parties

Table 3. 2010 Factor Analysis

	Nationalism	Economy
1. EU membership	0.894	
2. EU cooperation	0.845	0.174
3. Nuclear plan	0.281	0.395
4. Tax-spend	-0.325	-0.388
5. Tax exemption		0.373
6. Mansion tax	0.118	0.632
7. Tax relief		0.294
8. Ecotax	0.266	0.392
n	8084	
% variance	0.223	0.140
Cumulative % Variance	0.223	0.363

The averages of the positions of those voted for the parties are given by the vector:

$$z^* = \begin{pmatrix} & Lib & Con & LibDem & SNP & PC \\ nat & -0.31 & 0.50 & -0.34 & -0.11 & -0.09 \\ econ & -0.18 & 0.40 & -0.31 & -0.31 & -0.10 \\ \end{pmatrix}$$

3.1 Spatial models

The classical model must be modified in a way to include the assumption of varying party bundles:

$$u(x_i, z_j) = \lambda_j - \beta||x_i - z_j|| + \mu_{jr(i)} + \epsilon_i$$

Here for individual i the utilities are defined only for the parties from which he or she could actually choose.

Because of the structure of the set, namely, that significantly more observations come from England than from other two regions, the best convergence is obtained for the model, in which $\mu_{jr(i)} \neq 0$ for $(jr(i)) \in \{(1,2)(2,2)(1,3)(2,3)\}$.

The substantively chosen point priors for the valences are estimated with the classical approach via logistic regression. The definite numbers are not necessary for the estimation itself, since the model confidently converges to non-zero values without these point priors, but they, along with the *counterfactual estimates*, may serve as the reference points to see which benefits in terms of the estimates the adjusted model brings. They are to be compared with the *mixed valences* from the Bayesian model.

To estimate the priors for the valence for Conservative and Labour Party (that of Liberal Democrats is assumed to be the reference level with the 0 valence), the subsample of the observations from England is taken (to avoid the violation of IIA). To estimate the prior for the valence of Scottish National Party, the logistic model was run only on the observations from Scotland; for the valence of Plaid Cymru - only on those from Wales.

Priors: $\hat{\lambda_1} = -0.099$, $\hat{\lambda_2} = 0.276$, $\hat{\lambda_3} = 0$, $\hat{\lambda_4} = 0.233$, $\hat{\lambda_5} = -0.827$

$$\lambda^{prior} = (\lambda^{prior}_{PC}, \lambda^{prior}_{Lab}, \lambda^{prior}_{LibDem}, \lambda^{prior}_{SNP}, \lambda^{prior}_{Con}) = (-0.827, -0.099, 0, 0.233, 0.276)$$

After running 25000 iterations of 3 chains in JAGS, the model confidently converges according to the Gelman-Rubin diagnostic, all partial potential scale reduction factors and multivariate potential scale reduction factor equal to 1, and to the Heidelberger-Welch diagnostic, where both tests have passed in all chains and for all coefficients [7] (see the plots on Gibbs in the Appendix).

[7] with a minor exception in one chain for μ_{Lab3}, which is expected since the confidence interval contains 0

	estimate	credible interval
β	0.873	[0.913,0.833]
λ_{Lab}	-0.102	$[-0.163, -0.039]$
λ_{Con}	0.259	[0.193, 0.325]
λ_{SNP}	0.227	[0.024,0.432]
λ_{PC}	-0.762	[-1.127,-0.409]
μ_{Lab2}	0.589	[0.391,0.792]
μ_{Con2}	-0.466	[-0.729,-0.208]
μ_{Lab3}	0.458	[0.193, 0.726]
μ_{Con3}	-0.056	[-0.363,0.247]
DIC	00	

Table 1: Pure spatial Bayesian model with regional effects

Based on the table above, the valences (the order of the parties is the same as in the aggregated estimates):

In England:

$$(\lambda_{PC}^{mixed}, \lambda_{LibDem}^{mixed}, \lambda_{Lab}^{mixed}, \lambda_{Con}^{mixed}, \lambda_{SNP}^{mixed}) = (--, 0, -0.102, 0.259, --)$$

In Scotland:

$$(\lambda_{PC}^{mixed}, \lambda_{LibDem}^{mixed}, \lambda_{Lab}^{mixed}, \lambda_{Con}^{mixed}, \lambda_{SNP}^{mixed}) = (--, 0, 0.487, -0.470, 0.227)$$

In Wales:

$$(\lambda_{PC}^{mixed}, \lambda_{LibDem}^{mixed}, \lambda_{Lab}^{mixed}, \lambda_{Con}^{mixed}, \lambda_{SNP}^{mixed}) = (-0.762, 0, 0.091, 0.622, --)$$

the mixed valences are:

$$(\lambda_{PC}^{mixed}, \lambda_{LibDem}^{mixed}, \lambda_{Lab}^{mixed}, \lambda_{Con}^{mixed}, \lambda_{SNP}^{mixed}) = (-0.762, 0, 0.002, 0.213, 0.227)$$

[Model with demographics is to be added (I need to run again more iterations). So far what I saw - they are non-significant when added to this model, the valences do not change. Significance looks slightly worse. However, this model already fits well, so this is a technical step, page on which I will add before the last draft.]

3.2 Spatial models with traits

Following the usual sequence of the investigation, models capturing the personal attitudes towards the leaders of the parties are presented in this section.

Important to note that the usual task of the addition of the traits is not as straight-forward as it would be in a country where voters in all regions choose

	Endland			Scotland			Wales				
	d	e	ci	d	e	ci	d	e	ci	d	e
p_1	0.280	0.280	0.277,0.283	0.354	0.352	0.346, 0.359	0.355	0.352	0.339, 0.365	0.291	0.29
p_2	0.414	0.413	0.407, 0.420	0.167	0.169	0.156, 0.182	0.285	0.286	0.261, 0.311	0.383	0.38
p_3	0.307	0.307	0.303, 0.310	0.216	0.216	0.211, 0.221	0.250	0.251	0.239, 0.262	0.295	0.29
p_4	-	-	-	0.262	0.263	0.258, 0.267	-	-	-	0.026	0.02
p_5	-	-	-	-	-	-	0.111	0.112	0.109, 0.114	0.005	0.00

Table 2: Table 5: Predicted and sample voting probabilies (s - sample, e - estimates, ci - confidence interval)

from the same set of parties. In the case of Britain, for the estimation of the loadings for the traits the common factor analysis cannot be applied to the full sample, since in that case one of two (bad) choices would have to be made. The source of the problem is that the respondents in Scotland and Wales in addition to the questions about Brown, Cameron, and Clegg were asked about Salmon and Jones relatively as well.

Given that, the first fast option is to include only the attitudes towards the leader of the common set of parties into analysis. This is a poor solution since then the influence of the other two regional leaders on the factors for Brown, Cameron, and Clegg would be completely 'lost'. Furthermore, the traits for Salmon and Clegg are not estimated at all. Another solution might be assigning 'fake' zero attitudes towards Salmon and Jones to those individuals who had not been asked about them. This approach is misleading as well: in that case these 'fake' attitudes would affect the estimation of the real loadings for the traits of Salmon and Jones.

We propose the following solution to this problem:

1. Split the sample into the individuals with the same set of attitudinal questions (Scotland, Wales, and the rest)

2. For each sub-sample's attitudinal variables run the multiple imputation [8]

3. For the sub-sample with just three parties in the choice bundle, run 3-factor analysis

4. For for the subsamples for Scotland and Wales, run 4-factor analysis

5. Merge the sub-samples getting 5 traits for the full sample, while zero dummies not used in the following analysis, are assigned to contractual traits to keep the table balanced

An additional advantage of this approach is the possibility to compare the factor loading across the regions even on this primary stage of analysis. The patterns are expected to be similarm but slightly differ (see tables in the Appendix).

[8] With the use of the R package mice. See the details on the imputation in Appendix.

region	party	Brown	Cameron	Clegg	Salmond	Jones
England	Labour	0.84	-0.53	-0.21	-	-
	Conservative	-0.54	0.65	-0.18	-	-
	Liberal Democrats	-0.03	-0.39	0.43	-	-
Scotland	Lab	0.7	-0.31	-0.06	-0.29	-
	Conservative	-0.66	1.1	-0.25	-0.42	-
	Liberal Democrats	-0.15	-0.1	0.48	-0.18	-
	Scottish NP	-0.4	-0.21	-0.15	0.81	-
Wales	Labour	0.7	-0.42	-0.17	-	-0.04
	Conservative	-0.68	0.81	-0.18	-	-0.13
	Liberal Democrats	-0.13	-0.3	0.59	-	0.01
	Plaid Cymru	-0.19	-0.08	-0.31	-	0.44
Overall		0	0	0	0	0

Table 3: Average traits by regions and parties

The resulting models with traits differ from the canonical in the way that the bundles of the traits are different for the individuals with different party choices:

$$E(u(x_i, z_j)) = \lambda_j + \beta||x_i - z_j|| + \sum_{k \in m(i)} \alpha_k \tau + \mu_{jr(i)}$$

Here m(i) is the set of parties from which the individual i chooses, and τ_k is the trait for the party leader of the party k. Hence, if r(i) is England then $|m(i)| = 3$, and if r(i) is Scotland or Wales: $|m(i)| = 4$.

After the addition of the socio-demographic valences:

$$E(u(x_i, z_j)) = \lambda_j + \beta||x_i - z_j|| + \sum_{k \in m(i)} \alpha_k \tau + \mu_{jr(i)} + (\theta * \nu_i)$$

[Here I will add a page as in the section on the Spatial Models. This is an extremely slow model (even without demographics).]

3.3 Conterfactual analysis

In this section I will present what happens if we forget that IIA is violated, meaning the usual logistic model is run for the whole sample. [I will add more text and a detailed table, but the key estimates are already given in Table 6, (6).)

3.4 Comparison of the results

[To be updated]

However, based only on the three models, the relative advantage of the adjustment of the model is seen. Given, that our model is correct, since the probability predictions are very close to the sample, it is seen how wrong it would be to apply the counterfactual model(6) or even the models with priors(5).

	(1)	(2)	(3)	(4)	(5)	(6)
β	-0.76***	-0.87***	-	-	-	-0.76
λ_{lab}	-0.01	0.002***	-	-	-	-0.02
λ_{con}	0.28***	0.213***	-	-	-	0.28***
λ_{snp}	0.23***	0.227***	-	-	-	-3.9***
λ_{pc}	-0.83**	-0.762**	-	-	-	-5.4***

Table 4: Table 6: Comparison of the models

Models:
(1) - Priors
(2) - Pure spatial
(3) - Spatial model with socio-demographics valences
(4) - Spatial model with traits
(5) - Spatial model with traits and socio-demographic valences
(6) - Counterfactual

3.5 Assessment of the convergence

To check the application of the median voter theorem, let's assume that the parties take median position in relation to their constituencies:

$$z^0 = (z^0_{lab}, z^0_{con}, z^0_{lib}, z^0_{snp}, z^0_{pc}) = \begin{pmatrix} 0 & 0 & 0 & -0.19 & -0.09 \\ 0 & 0 & 0 & -0.13 & -0.11 \end{pmatrix} \tag{6}$$

In the previous section, the estimates for the valences of the major parties were shown to be robust across the models with various sets of random and fixed effects. Given this, based on the considerations of the parsimony, the pure spatial model with random regional effects is employed for the assessment of the convergency.

The Hessian matrices for the parties are:

$$H_{Lab|z^*} = \begin{pmatrix} -0.185 & -0.008 \\ -0.008 & -0.222 \end{pmatrix} \; H_{Con|z^*} = \begin{pmatrix} -0.266 & -0.0308 \\ -0.031 & -0.275 \end{pmatrix} \; H_{LibDem|z^*} = \begin{pmatrix} -0.203 & -0.034 \\ -0.034 & -0.238 \end{pmatrix}$$

$$H_{SNP|z^*} = \begin{pmatrix} -0.098 & 0.014 \\ 0.014 & -0.211 \end{pmatrix} \; H_{PC|z^*} = \begin{pmatrix} 0.024 & 0.003 \\ 0.002 & -0.060 \end{pmatrix}$$

$$eigen(H|z^*) = \begin{pmatrix} & Lib & Con & LibDem & SNP & PC \\ Nat & -0.183 & -0.240 & -0.182 & -0.096 & 0.024 \\ Econ & -0.224 & -0.302 & -0.258 & -0.213 & -0.060 \end{pmatrix}$$

$$\{c(z^*)_j\} = \begin{pmatrix} c(z_1^*) \\ c(z_2^*) \end{pmatrix} = \begin{pmatrix} & Lib & Con & LibDem & SNP & PC \\ Nat & 0.115 & -0.201 & 0.080 & 0.178 & 0.126 \\ Econ & 0.084 & -0.239 & 0.057 & 0.089 & 0.072 \\ \end{pmatrix} =$$

$$\begin{pmatrix} Lib & Con & LibDem & SNP & PC \\ 0.2 & -0.44 & 0.137 & 0.267 & 0.197 \\ \end{pmatrix}$$

$$c(z) = max(\{c(z^*)_j\}) = 0.267 < 1$$

We observe that the only party for which the median position is the saddle point is Playd Cymru, the rest of the conditions for the convergence of the electoral system are strongly satisfied. Dependent on the perception of PC, the situation can be evaluated in two different perspectives. First, formally, the system does not converge if we consider all parties equally important. The second approach is to remember that the inclusion of Plaid Cymru was rather technical to formally control for the violation of IIA, meanwhile it is not a major player, and many parties not included in the analysis got much more voters.

In terms of three major parties the electoral system confidently converges, and this does not change with the inclusion with the only other candidate to be considered a major party (SNP). An interesting observation is the negative convergence coefficient for the Conservatives. This means that they extremely benefit from their position and even a small deviation from the median position would cause a significant decrease of the voting support.

4 Conclusion

5 Appendix

		Brown Trait	Cameron Trait	Clegg Trait
Brown	Knowledge	0.810	-0.210	0.154
	Feeling	0.866	-0.330	
	Competence	0.874	-0.282	
	Interests	0.871	-0.238	0.117
	Trustworthy	0.878	-0.220	0.119
Cameron	Knowledge	-0.221	0.834	0.117
	Feeling	-0.363	0.834	
	Competence	-0.247	0.757	
	Interests	-0.239	0.868	
	Trustworthy	-0.178	0.855	
Clegg	Knowledge			0.818
	Feeling			0.827
	Competence			0.832
	Interests	0.180		0.768
	Trustworthy	0.146	0.174	0.729
	n	6898		
	% variance	0.273	0.255	0.217
	Cumulative % Variance	0.273	0.528	0.746

Table 5: Factor Loadings for Traits in England in 2010

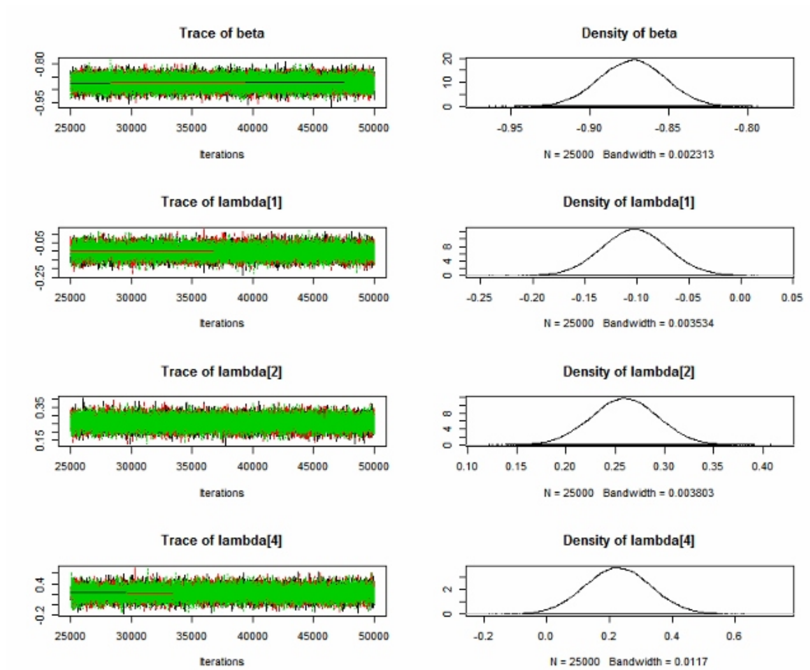
		Brown Trait	Cameron Trait	Clegg Trait	Salmond Trait
Brown	Knowledge	0.826	-0.162	0.164	
	Feeling	0.905	-0.229		
	Competence	0.895	-0.172	0.103	
	Interests	-0.137			
	Trustworthy	0.878	-0.133		
Cameron	Knowledge	-0.134	0.806	0.186	
	Feeling	-0.274	0.856		
	Competence	-0.160	0.841	0.160	
	Interests	-0.152	0.840		
	Trustworthy	-0.106	0.835	0.120	
Clegg	Knowledge		0.119	0.845	
	Feeling			0.828	
	Competence		0.116	0.848	
	Interests	0.158	0.114	0.787	
	Trustworthy	0.139	0.198	0.703	
Salmond	Feeling	-0.104			0.815
	Competence				0.996
	n	800			
	% variance	0.241	0.219	0.199	0.099
	Cumulative % Variance	0.241	0.460	0.659	0.757

Table 6: Factor Loadings for Traits in Scotland in 2010

		Brown Trait	Cameron Trait	Clegg Trait	Jones Trait
Brown	Knowledge	0.785	-0.291	0.132	
	Feeling	0.857	-0.376		
	Competence	0.864	-0.321		
	Interests	0.851	-0.278		
	Trustworthy	0.854	-0.274		
Cameron	Knowledge	-0.251	0.799		
	Feeling	-0.406	0.785		
	Competence	-0.285	0.808		
	Interests	-0.316	0.839		
	Trustworthy	-0.272	0.835		
Clegg	Knowledge	0.103		0.797	
	Feeling		-0.111	0.818	0.126
	Competence			0.815	0.121
	Interests			0.808	
	Trustworthy	0.105	0.130	0.756	
Jones	Feeling				0.777
	Competence			0.115	0.818
	n	800			
	% variance	0.241	0.219	0.199	0.099
	Cumulative % Variance	0.241	0.460	0.659	0.757

Table 7: Factor Loadings for Traits in Wales in 2010

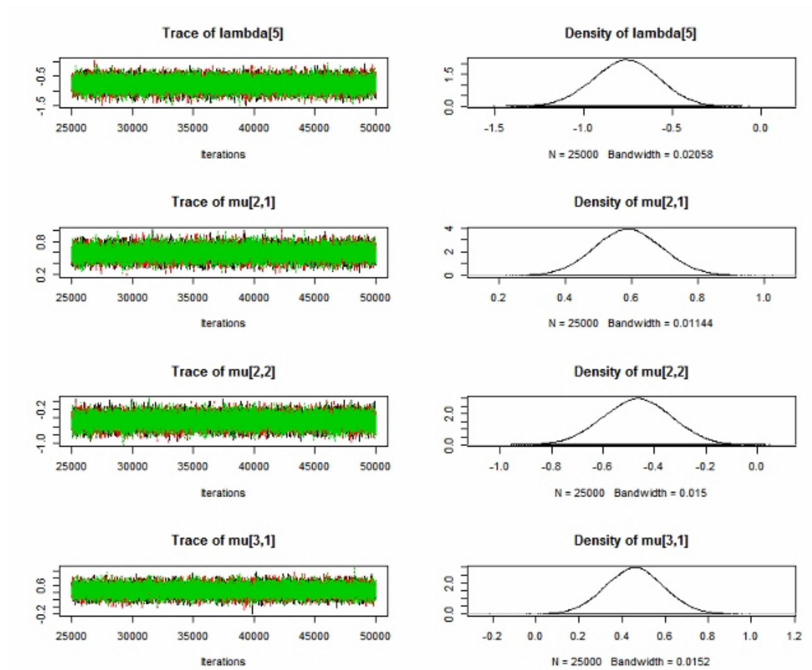
(1)



$_1.png$

Figure 1: Gibb's sampling: Pure spatial models with regional effects 1

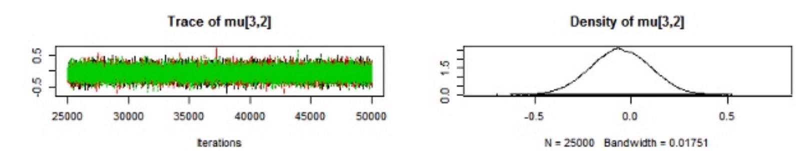
(1)



$_2.png$

Figure 2: Gibb's sampling: Pure spatial models with regional effects 2

(1)



$_3.png$

Figure 3: Gibb's sampling: Pure spatial models with regional effects 3