University Of Groningen

Master Thesis

Mathematics

Heterogeneity for the Stochastic Blockmodel and the Issue of Separation in Political Networks

Author
B.Sc. Frank Lefeber

Supervisors
M.Sc. Mirko Signorelli
Prof. Dr. Ernst Wit

March 22, 2017

Abstract

The stochastic blockmodel is a useful tool for modeling networks in which group structures are present. However, this model assumes homogeneity among all individuals within the same group. In a recent study, Signorelli and Wit (2016) tried a stochastic blockmodel approach with added individual attributes of deputies, which allowed for heterogeneity. We explore a different approach in this study, by extending the stochastic blockmodel with individual effects, actually stepping away from its blockmodel property, keeping block interactions intact. We compare its results to the ones from the basic stochastic blockmodel. We compare two inferential procedures, one being the maximum likelihood estimation and another being penalized likelihood estimation with the adaptive lasso. We explain our methodology and test the models on some toy examples of networks. Then we use them on real data from the Finnish and Italian parliaments. During our analyses we encounter strengths and weaknesses of the various models, including a vulnerability to separation in the data.

1 Introduction

1.1 Network Model History

The essence of modeling a social network is looking at interactions between people i and j. These interactions can be regarded as realizations from a random variable from some distribution and often assumed to be independent from certain other interactions. A network can be represented by a graph by defining people in the network as the nodes of the graph and the interactions between them as edges. This network has certain properties, like edge direction, edge weights, overal density of the graph and reciprocity of directed edges (the fraction of edges for which an edge in the opposite direction exists). The nodes themselves also have certain properties, like their outgoing and ingoing degrees, which can be viewed as productivity (or expansiveness) and popularity respectively. There are possible other properties associated to the nodes like age, sex or membership of a certain group. We use these properties to determine some quantification μ_{ij} of collaboration between people i and j. There are various models that take different approaches to this problem. Holland and Leinhardt (1981) proposed a model wherein collaboration was based on two parameters from the whole network; the density of the graph θ and the reciprocity of directed edges ρ, and two nodal parameters; node expansiveness α_i and node popularity β_j. Besides individual effects there are group effects which can be considered.

Fienberg and Wasserman (1981) considered a partition of the nodes into p groups (blocks). A partition is mutually exclusive and exhaustive, each individual node i belongs to a single group r (denoted by $i \in r$). They proposed a model in which individual expansiveness α_i and popularity β_j were replaced by group expansiveness α_r and popularity β_s (for $i \in r$ and $j \in s$). Two years later Holland et al. (1983) proposed the stochastic blockmodel which also uses a partition into p blocks. This model introduces new block parameters ϕ_{rs} for interactions between groups. Any individual differences between people in the same block are meaningless in this model. Under the definition of the stochastic blockmodel it only matters how much interaction there is between groups, not which person is responsible for which interaction. Wang and Wong (1987) used a combination of individual effects α_i, β_j and group effects ϕ_{rs} for directed networks. Recently, Signorelli and Wit (2016) used the density θ, group parameters α_r, α_s, ϕ_{rs} and some individual attributes (such as sex and age) to work on political models. The stochastic blockmodel is a sensible first choice for a basic model, since the parties of a parliament form a partition on the nodes. In this thesis we start from this model and extend it with individual effects, much like the model of Wang and Wong (1987) and compare the results with the basic model.

1.2 Country Selection

On the website of Briatte (2016) there are bill-cosponsorship data for various countries in Europe. For every legislature there is an associated network graph showing all members of parliament as nodes and edges where cosponsorship occurred. Looking at the various graphs gives an idea of how much cosponsoring is going on and which parties collaborate together. However, because there is a large amount of nodes it is not obvious to recognize all patterns straight away and the edges are not weighted. With our models we will produce graphs based on the political parties in such a way that it is apparent straight away which parties collaborate. The chosen countries are Finland, as it has separated data and Italy mainly for its large size, but also to allow comparison with a graph from Signorelli and Wit (2016). What the data does not give us is information on the bills, so there is no hope of creating a bipartite graph, showing bills on the left and members of parliament or groups on the right. We can also not use the ideological placement of the bills in the quadrants created by left-right and progressive-conservative axes. Since cosponsorships are mutual we are working with an undirected network. Also, if individual i cosponsors with two other individuals j and k on the same bill, then j and k are automatically cosponsoring as well. This is an interesting property of cosponsorship networks. In the next section we go into the detail of our models.

2 Methods

2.1 Basic Blockmodel

The files from Briatte (2016) are in .gexf format, representing the networks as graphs. Using Gephi, we can retrieve edge lists and node lists from these files. These edges have weights and are undirected. Since any deputy in a parliament has to be part of a single party, the parties in the parliament form a partition on the deputies. This allows us to make a stochastic blockmodel. For a parliament consisting of n deputies which form p parties we model y_{ij}, the number of cosponsorships between deputy i from party r and deputy j from party s. The number of cosponsorships is a count and we assume it to be a draw from a Poisson distribution with mean μ_{ij}. A network model is a stochastic blockmodel if the random vectors X_{ij} and X_{kl} are identically distributed when $i, k \in r$ and $j, l \in s$. So also μ_{ij} should be the same as μ_{kl} under the same conditions. In other words: we assume homogeneity among all members of the same party, they are assumed to behave identically. This is not very realistic and that is exactly what we try to improve with the extended model. In the first model μ_{ij} depends on several parameters. The overall weighted density of the graph θ_0 is a measure of how frequently cosponsorships occur, so it should affect cosponsorships between deputies i and j. In a sense it sets some baseline for how many cosponsorships happen between any two deputies. We also consider the productivity of i and j: α_i and α_j. However, since this is a stochastic blockmodel we only consider α_r and α_s, the group productivity of the parties they belong to respectively. The last piece of the model is the block interactions parameter ϕ_{rs} which represents the tendency for parties to cosponsor with each other or avoid certain parties. Using the logarithm as a link function, the model becomes

$$log(\mu_{ij}) = \theta_0 + \alpha_r + \alpha_s + \phi_{rs} = X\beta, \tag{1}$$

where β is a vector of coefficients to be estimated and X is the design matrix. The vector y contains the observed amount of cosponsorships between all deputy pairs (i, j), its entries are $(y_{12}, y_{13}, \ldots, y_{p-1\ p})$. Note that we write y_{ij} to specify some k for which y_k is the entry of y which contains the number of cosponsorships between deputies i and j. It is a vector containing the entries of an upper triangular weighted adjacency matrix of all deputies, sorted by row first and column second. For the design matrix X we need to define dummy variables $D_r(ij)$ and $D_{rs}(ij)$ such that we can rewrite (1) as

$$log(\mu_{ij}) = \theta_0 + \sum_{r=1}^{p} \alpha_r D_r(i) + \sum_{s=1}^{p} \alpha_s D_s(j) + \sum_{r \leqslant s}^{p} \phi_{rs} D_{rs}(ij). \tag{2}$$

We define $D_r(k)$ first. For k in $1 \dots n$, for r in $1 \dots p$:

$$D_r(k) = \begin{cases} 1 \text{ if } k \in r \\ 0 \text{ if } k \notin r \end{cases}$$

Now we define $D_{rs}(ij)$ for the structure of ϕ_{rs}. For all pairs (i, j) with i, j in $1 \dots n$, for all r, s in $1 \dots p$:

$$D_{rs}(ij) = \begin{cases} 1 \text{ if } (i \in r \wedge j \in s) \vee (i \in s \wedge j \in r) \\ 0 \text{ otherwise} \end{cases}$$

In θ_0 we already captured the overall productivity of the whole network. So these α_r that we introduced can not for example all be positive. In such a case, a constant can be removed from all α_r and added to θ_0. This presents an identifiability issue if we leave α_r unconstrained. Similarly, we need constraints for ϕ_{rs}. We take the identifiability conditions from Signorelli and Wit (2016),

$$\sum_{r=1}^{p} \alpha_r = 0 \qquad \text{and} \qquad \sum_{s=1}^{p} \phi_{rs} = 0 \quad \forall r.$$

These conditions are then used to compute the new dummy variables

$$T_r = D_r - D_1 \quad \forall r > 1 \qquad \text{and} \qquad T_{rs} = D_{rs} - D_{rr} - D_{ss} \quad \forall s > r \geqslant 1.$$

Now we can specify the columns of the structure matrix X. They consist of

$$T_2 \dots T_p, T_{12} \dots T_{1p}, T_{23} \dots T_{2p}, \ \dots \ , T_{p-1\ p}.$$

We do not need a column for the intercept in X, as it is added automatically by the functions we use. Now we can rewrite (2) as

$$log(\mu_{ij}) = \theta_0 + \sum_{r=2}^{p} \alpha_r T_r(i) + \sum_{s=2}^{p} \alpha_s T_s(j) + \sum_{r<s}^{p} \phi_{rs} T_{rs}(ij) = X\beta. \tag{3}$$

We can use y and X for certain functions in R, like *glm* and *glmnet*. We first look at an extension of the model, before going into detail about using these functions.

2.2 Extended Blockmodel

The problem with the blockmodel given by (1) is that there is no heterogeneity. Each individual is regarded as a pawn in a group and this is not the way it is in reality. Each individual has their own unique behavior and a good model should nourish this property, instead of ignoring it. In order to include this in the model, we need to make some adjustments to the basic model. To account for differences among deputies of the same party we add fixed effects to each individual, based on their productivity. So we change from the group productivity α_r and α_s to the individual productivity α_i and α_j. To make the distinction clear, we will call them γ_i and γ_j. This is no longer a stochastic blockmodel, but we keep the collaboration preferences ϕ_{rs} in the model. The new model becomes

$$log(\mu_{ij}) = \theta_0 + \gamma_i + \gamma_j + \phi_{rs}. \tag{4}$$

With this model y_{ij} can remain the same, but we have to adjust the design matrix X in such a way that it represents the individual productivity instead of the group productivity. We do this by introducing new dummy variables for γ, while keeping the ones for ϕ the same. For k in $1 \dots n$, i in $1 \dots n$:

$$D_k(i) = \begin{cases} 1 \text{ if } k = i \\ 0 \text{ if } k \neq i \end{cases}$$

So for the extended blockmodel equation (2) becomes

$$log(\mu_{ij}) = \theta_0 + \sum_{r=1}^{p} \gamma_i D_k(i) + \sum_{s=1}^{p} \gamma_j D_k(j) + \sum_{r \leqslant s}^{p} \phi_{rs} D_{rs}(ij). \tag{5}$$

For identifiability the constraint for γ has to change

$$\sum_{i=1}^{n} \gamma_i = 0 \qquad \text{and} \qquad \sum_{s=1}^{p} \phi_{rs} = 0 \quad \forall r,$$

leading to the new dummy variable

$$T_k = D_k - D_1 \;\; \forall k > 1.$$

Now, the columns of the design matrix X consist of

$$T_2 \dots T_n, T_{12} \dots T_{1p}, T_{23} \dots T_{2p}, \;\dots\;, T_{p-1\;p}.$$

So equation (5) becomes

$$log(\mu_{ij}) = \theta_0 + \sum_{r=2}^{n} \gamma_i T_k(i) + \sum_{s=2}^{n} \gamma_j T_k(j) + \sum_{r<s}^{p} \phi_{rs} T_{rs}(ij) = X\beta. \tag{6}$$

With this new X we have a structure based on individual fixed effects and group tendencies.

2.3 Inference and Graphing

The y and X developed in the previous sections are ready for usage with some functions in R. With *glm* we calculate the maximum likelihood estimates (MLE). In maximum likelihood estimation the best fit is given by the model with the most parameters (it tends to overfit), this means that all parameter estimates will be nonzero. In the parliament some pairs of parties will have preferences to work together ($\phi_{rs} > 0$) or avoid ($\phi_{rs} < 0$) each other, but they might also be impartial ($\phi_{rs} = 0$). The MLE can not provide an exact zero for an estimate. Another method of parameter estimation is penalized likelihood estimation, which we can do with *glmnet*. With penalized likelihood we maximize the log-likelihood minus a penalty term. By introducing this penalty term we introduce a little bias but we can shrink and select parameters for the model, improving model interpretation. The penalty term increases with the number of parameters in the model, so unless parameters have enough impact on the model they will be set to zero. The lasso (Least Absolute Shrinkage and Selection Operator) introduced by Tibshirani (1996) is a tool that allows for shrinkage of variables as well as selection. The adaptive lasso from Zou (2006) adds weight terms associated to each parameter so we can penalize certain parameters more than others. It also provides consistent estimators. In the adaptive lasso we use the following penalty term:

$$\lambda \sum_j w_j |\beta_j| \qquad \text{with weights} \qquad w_j = \left(\frac{1}{\hat{\beta}_j}\right)^{\gamma},$$

using the MLE estimates $\hat{\beta}_j$. For consistent estimators γ needs to be larger than 1, we choose to use $\gamma = 2$. In this penalty term λ is the tuning parameter that determines how much penalization happens. It is a decreasing sequence of values, allowing more parameters in the model as it becomes smaller. At $\lambda = 0$ the model is equal to the MLE, because the penalty is zero in that case. For each λ in the sequence we get parameter estimates maximizing the penalized likelihood. From these we select the best model using the Bayesian information criterion (BIC). We use BIC because it prefers smaller models compared to for example AIC. Signorelli and Wit (2016) showed that overall BIC appeared to be the best selection tool for these political networks. We will abbreviate the process of penalization and BIC by PEB. The parameter estimates we get are those represented by the columns of X. We then use the identifiability conditions to calculate the effects we took out ((α_1 or γ_1), $\phi_{11}, \phi_{22}, \ldots, \phi_{pp}$).

The next step is to summarize the estimates by representing the network as a reduced graph (Anderson et al. (1992), Signorelli and Wit (2016)). Each node will represent a party, where its node size is proportional to the group productivity. In the case of the basic model, this is simply α_r. In the extended model, we take an averaged sum of the productivity γ_i for $i \in r$. The node size is adjusted to group size, such that it scales with proportional, rather than absolute productivity. The part we are most interested in is ϕ_{rs}, the preferences parties have to collaborate with other parties. Each edge (r, s) in the graph represents a positive ϕ_{rs}. These graphs are then analyzed using information about the network and the parameter estimates are inspected. Parties have been given abbreviated names and colors manually. For each data set we show the basic blockmodel and the extended blockmodel, using both the MLE and PEB procedures.

2.4 Separation

A possible problem that can arise is the one of separation (Santos Silva and Tenreyro (2010)). Separation occurs in certain data configurations in which one or more regressors x_{ia} correspond to zero when y_i is positive, while otherwise they are non-negative with at least one positive observation. The problem is that no maximum likelihood estimator exists in such a situation. In our case this occurs whenever there is a pair of uncollaborating parties. For logistic and binary regression solutions exists (Heinze and Schemper (2002), Zorn (2005)), but these solutions are not helpful in our case. So when we are dealing with some pair of parties (r, s) for which: $y_{ij} = 0$ for all $i \in r, j \in s$, we have separation. In the estimation process, μ_{ij} will be 'placed at $-\infty$', because we have a log link and the probability is estimated to be 0. For networks without separation the estimates are small ($|\beta_{ij}| < 4$) and the standard errors are a lot smaller. In networks with separation ϕ_{rs} becomes large in absolute value and gets huge standard errors. As we will see, all other standard errors are also inflated. This is one of the reasons we ignore parties consisting of merely one or two members in the data, another reason being that they are not appropriate for the stochastic blockmodel. Even without these tiny parties, separation can still occur with larger parties and represents a problem for our model. We investigate this further with some toy examples in the next section.

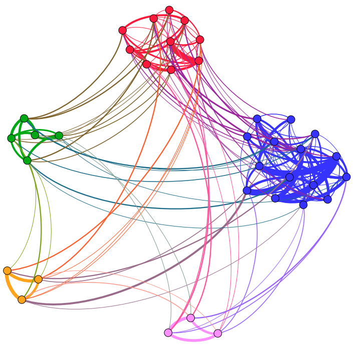
3 Toy Examples

The four toy examples are based on the amount of deputies shown in the table. The nodes stay the same, the edges are changed as we move from one example to another. We show how well the models work for structurally different networks. We first show a realistic example for which the model works nicely. The second example is to show what extreme separation does to the models. The third example shows that if just a single pair of parties do not collaborate, the models are still affected. Finally we show a full network which is drawn from a single Poisson distribution, to see if we can fool our models. Using Gephi we produce an image corresponding to the network based on the created node list and edge list. The nodes are placed and given color so the graph looks similar to the model graphs, thicker edges represent multiple cosponsorships.

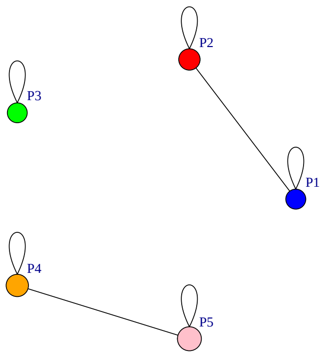
	# seats
P1	15
P2	10
P3	5
P4	3
P5	3

3.1 No Separation

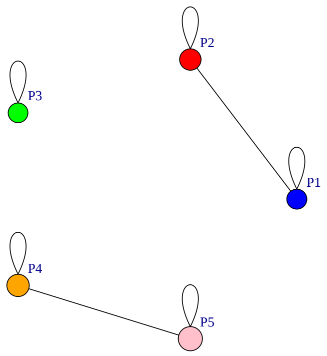
This example shows a fully connected network viewed from a group perspective. We see that parties (1,2) are strongly connected. Also, the weights of the edges between group (4,5) are large, so we can expect a link there as well. Other than that, the parties work together among themselves, so we expect a lot of self-links. These self-links are typically expected in political networks, since most parties are groups of like-minded individuals and they are in parties together for good reasons. These tables



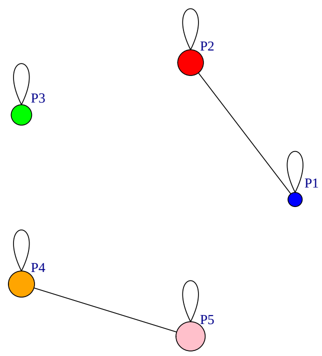
Gephi Graph



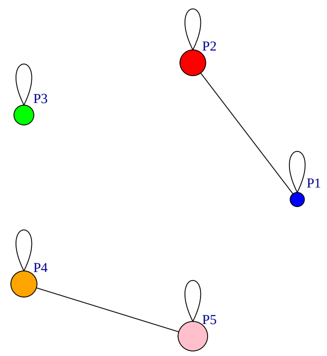
MLE for Basic Blockmodel



PEB for Basic Blockmodel



MLE for Extendedc Blockmodel



PEB for Extendedc Blockmodel

represent the parameter estimates that are given as output by the functions *glm* (the MLE estimate and the associated standard errors) and *glmnet* (the parameter estimates of the model selected by PEB). We do not consider standard errors for the penalized model. Standard errors are uninteresting for estimation procedures that produce biased parameter estimates. The missing coefficients (α_1 or γ_1), $\phi_{11}, \phi_{22}, \ldots, \phi_{pp}$ can be calculated by taking the negative sum of their counterparts. In the tables we can see that parameter estimates for networks without separation are indeed small. What we typically see is that the estimation process seems to influence the outcome more than the choice of model, in other words: the basic model with MLE estimation looks similar to the extended model with MLE estimation and the same applies for PEB estimation.

For this network the parameter estimates are very similar, regardless of the model or estimation process. The estimates are small, as are the standard errors. If we focus on the ϕ_{rs} we note that their estimates and standard errors are all very similar across all four models. So the models agree on the same graph for this network, which is a nice result. When the network gets more complicated we can expect slight variations, but for a simple network like this it would be strange to have different results. We will investigate if this remains the case when we change the network.

Toy Example No Separation

β	MLE	Std. Err.	PEB
θ_0	-0.452	0.085	-0.453
α_2	0.234	0.094	0.246
α_3	-0.234	0.109	-0.275
α_4	0.207	0.113	0.210
α_5	0.439	0.107	0.462
ϕ_{12}	1.171	0.125	1.152
ϕ_{13}	-1.194	0.300	-1.436
ϕ_{14}	-1.529	0.357	-1.305
ϕ_{15}	-1.202	0.281	-1.164
ϕ_{23}	-0.202	0.175	0.000
ϕ_{24}	-0.904	0.226	-0.961
ϕ_{25}	-0.731	0.194	-0.820
ϕ_{34}	-0.619	0.325	-0.758
ϕ_{35}	-0.158	0.245	0.000
ϕ_{45}	0.977	0.180	0.939

Parameter Estimates for Basic Blockmodel

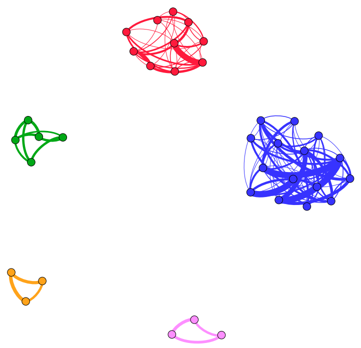
Toy Example No Separation

β	MLE	Std. Err.	PEB
θ_0	-0.924	0.103	-0.922
γ_2	-0.125	0.146	-0.121
γ_3	-0.272	0.153	-0.268
γ_4	-0.289	0.154	-0.286
γ_5	-0.768	0.181	-0.764
γ_6	-0.592	0.170	-0.589
γ_7	-0.289	0.154	-0.286
γ_8	-0.188	0.149	-0.184
γ_9	-0.272	0.153	-0.268
γ_{10}	-0.741	0.179	-0.738
γ_{11}	-0.569	0.168	-0.566
γ_{12}	-1.083	0.204	-1.080
γ_{13}	-0.946	0.193	-0.943
γ_{14}	-0.689	0.176	-0.686
γ_{15}	-0.272	0.153	-0.269
γ_{16}	0.849	0.157	0.858
γ_{17}	0.230	0.192	0.239
γ_{18}	0.665	0.166	0.673
γ_{19}	0.369	0.183	0.378
γ_{20}	0.624	0.168	0.632
γ_{21}	-0.202	0.226	-0.194
γ_{22}	-0.345	0.239	-0.337
γ_{23}	0.797	0.160	0.805
γ_{24}	0.931	0.154	0.939
γ_{25}	-0.117	0.218	-0.109
γ_{26}	0.030	0.233	-0.015
γ_{27}	-0.020	0.236	-0.064
γ_{28}	0.030	0.233	-0.013
γ_{29}	-0.072	0.240	-0.114
γ_{30}	0.030	0.233	-0.012
γ_{31}	0.547	0.202	0.548
γ_{32}	0.664	0.195	0.666
γ_{33}	0.020	0.236	0.021
γ_{34}	0.599	0.204	0.624
γ_{35}	0.285	0.229	0.312
γ_{36}	1.024	0.178	1.046
ϕ_{12}	1.170	0.125	1.151
ϕ_{13}	-1.192	0.300	-1.434
ϕ_{14}	-1.533	0.357	-1.304
ϕ_{15}	-1.209	0.281	-1.169
ϕ_{23}	-0.201	0.175	0.000
ϕ_{24}	-0.910	0.226	-0.967
ϕ_{25}	-0.740	0.194	-0.831
ϕ_{34}	-0.621	0.325	-0.764
ϕ_{35}	-0.164	0.246	0.000
ϕ_{45}	0.965	0.180	0.926

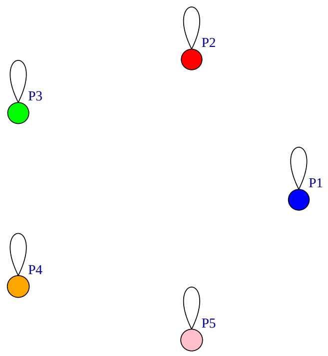
Parameter Estimates for Extended Blockmodel

3.2 Total Separation

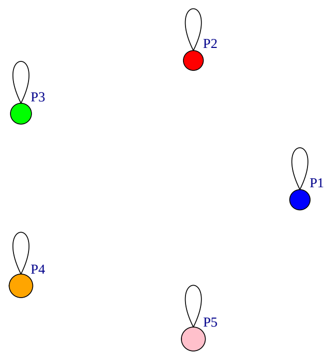
This is an extreme case of separation: there are zero observed cosponsorships between any two members from different parties. However, if we look at the graphs they actually represent the network perfectly. There are no edges between any two different groups and all groups have self-links. Therefore the models all seem to work. However, all estimates have huge standard errors, especially when smaller groups are involved. This is probably due to the larger groups having more cosponsorhips. It seems that *glm* (MLE) suffers a lot more than *glmnet* (PEB). Of course this data example is extreme and unrealistic.



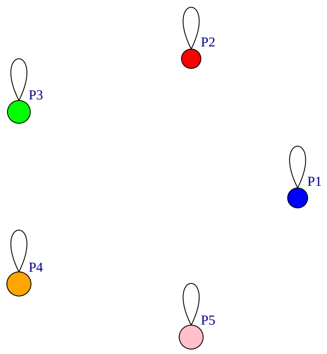
Gephi Graph



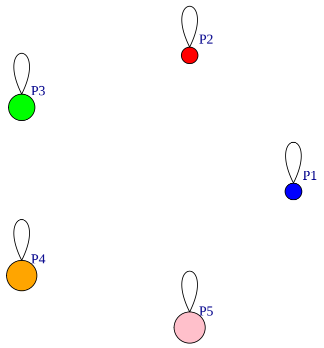
MLE for Basic Blockmodel



PEB for Basic Blockmodel



MLE for Extendedc Blockmodel



PEB for Extendedc Blockmodel

The interpretation would be a parliament in which no parties collaborate with other parties at all. This would mean that nothing gets done, unless one of the parties has majority by itself. In that case it would be uninteresting to study the parliament anyway, since interactions between groups would not be required for that party to push their bills. A more interesting case is the one in which there is just a pair of parties that is not collaborating. This is realistic, as we actually have data of multiple legislatures in which this happened. The next example will cover this type of separation, which we will call single separation.

Toy Example Total Separation

β	MLE	Std. Err.	PEB
θ_0	-15.959	781.982	-5.607
α_2	-0.142	889.861	-0.474
α_3	0.026	966.210	0.017
α_4	0.124	1030.722	0.383
α_5	0.106	1030.722	0.437
ϕ_{12}	-4.087	1274.728	-2.244
ϕ_{13}	-4.256	1573.272	-1.839
ϕ_{14}	-4.353	1883.809	-1.591
ϕ_{15}	-4.335	1883.810	-1.503
ϕ_{23}	-4.228	1799.876	-1.822
ϕ_{24}	-4.326	2187.572	-1.637
ϕ_{25}	-4.307	2187.572	-1.555
ϕ_{34}	-4.494	2910.310	-1.745
ϕ_{35}	-4.476	2910.311	-1.713
ϕ_{45}	-4.574	3649.982	-1.905

Parameter Estimates for Basic Blockmodel

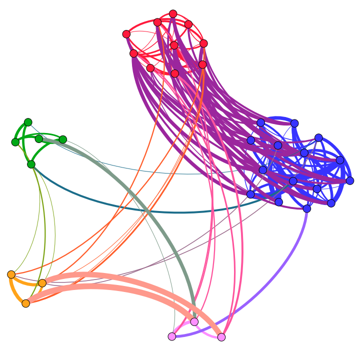
Toy Example Total Separation

β	MLE	Std. Err.	PEB
θ_0	-16.177	594.294	-6.224
γ_2	0.794	569.282	0.593
γ_3	0.017	569.282	-0.184
γ_4	0.259	569.282	0.058
γ_5	-0.533	569.282	-0.734
γ_6	0.113	569.282	-0.088
γ_7	0.017	569.282	-0.184
γ_8	0.050	569.282	-0.151
γ_9	0.339	569.282	0.137
γ_{10}	0.389	569.282	0.188
γ_{11}	-0.648	569.282	-0.850
γ_{12}	-0.648	569.282	-0.850
γ_{13}	-0.711	569.282	-0.913
γ_{14}	-0.533	569.282	-0.735
γ_{15}	-0.533	569.282	-0.735
γ_{16}	-1.176	725.501	-1.351
γ_{17}	-0.450	725.501	-0.624
γ_{18}	0.717	725.501	0.544
γ_{19}	-0.644	725.501	-0.817
γ_{20}	-0.209	725.501	-0.382
γ_{21}	-0.209	725.501	-0.382
γ_{22}	0.053	725.501	-0.120
γ_{23}	0.717	725.501	0.544
γ_{24}	0.272	725.501	0.099
γ_{25}	-0.138	725.501	-0.311
γ_{26}	0.010	1038.629	0.317
γ_{27}	0.361	1038.629	0.667
γ_{28}	-0.066	1038.629	0.241
γ_{29}	0.225	1038.629	0.531
γ_{30}	0.155	1038.629	0.461
γ_{31}	0.058	1223.632	0.559
γ_{32}	0.176	1223.632	0.677
γ_{33}	0.464	1223.632	0.964
γ_{34}	0.078	1224.374	0.652
γ_{35}	0.366	1224.374	0.939
γ_{36}	0.211	1224.374	0.785
ϕ_{12}	-4.036	1202.508	-2.347
ϕ_{13}	-4.237	1517.445	-1.876
ϕ_{14}	-4.338	1814.483	-1.640
ϕ_{15}	-4.319	1817.245	-1.516
ϕ_{23}	-4.204	1719.061	-1.689
ϕ_{24}	-4.304	2085.512	-1.634
ϕ_{25}	-4.286	2088.931	-1.554
ϕ_{34}	-4.505	2880.351	-1.650
ϕ_{35}	-4.487	2885.596	-1.651
ϕ_{45}	-4.587	3623.376	-1.855

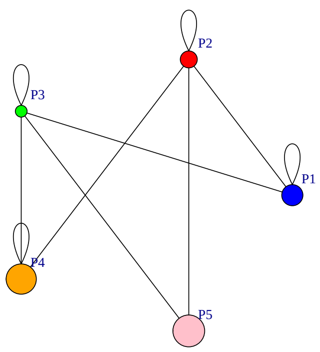
Parameter Estimates for Extended Model

3.3 Single Separation

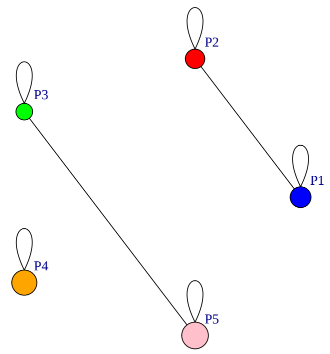
In the previous example all interactions were equally problematic, but in this case there is only one problematic pair of parties, namely (2,3). This is perhaps the most interesting and realistic case of separation, so we go into most detail on this example. This example also shows the severity of the situation: it only takes one uncollaborating pair of parties to cause estimation problems. For the first time there is variation in the graphs other than some fluctuations in node size. Also apparent is the increase in edges in the MLE graphs, which look identical. The PEB graphs also look very similar, though the PEB for Extendedc Blockmodel graph has an edge (2,5), which is not present in the PEB for Basic Blockmodel graph. This edge is from a weak coefficient $\phi_{25} = 0.176$, so there is no terribly big difference between the two. Interestingly, the nodes $P2$ and $P3$ in the MLE graphs have links with all other parties, except between themselves. This last part is logical as they are the source of separation, but apparently there is overcompensation for this in the MLE estimation process which expresses itself through more links with other groups. This is not something we see in the PEB graphs. Knowing this, we can take a look at these models while keeping in mind that not all links with the parties that are problematic should be accepted. If we then compare the four graphs they are actually very similar again. Looking at the estimates, the weakest four edges are actually (1,3), (2,4), (3,4) and (3,5) in order of weak to strong. These are exactly the edges that are not represented by the PEB graphs (the fourth one is, in the PEB for Extendedc Blockmodel graph). From the separation in (2,3) we have $y_{ij} = 0$ for



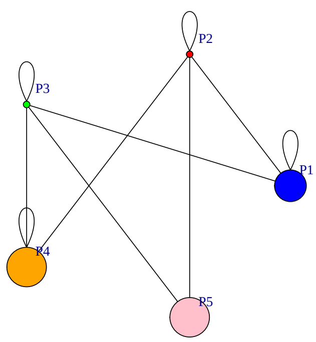
Gephi Graph



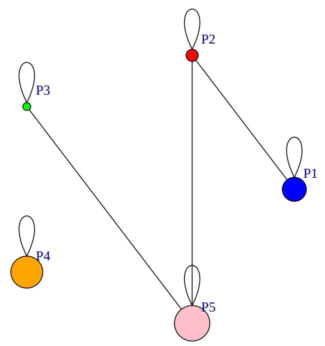
MLE for Basic Blockmodel



PEB for Basic Blockmodel



MLE for Extendedc Blockmodel



PEB for Extendedc Blockmodel

all $i \in 2, j \in 3$. The tables show inflated coefficient estimates and standard errors. This is expected for ϕ_{23}, but in fact all the parameters have large standard errors. My reasoning is that the error sizes are proportional to how often the parameters are involved with the problematic y_{ij}. The model is

$$log(\lambda_{ij}) = \theta_0 + \alpha_r + \alpha_s + \phi_{rs},$$

so the left hand side goes to $-\infty$ for $y_{ij} = 0$. The parameter that is influenced least of all is θ_0, because it is associated with all y_{ij} and many of them are non-problematic. We would expect no influence on α_4, α_5, ϕ_{14}, ϕ_{15} and ϕ_{45} because they are not directly involved with the $y_{ij} = 0$, but they might be influenced indirectly, albeit the least. Parameters that are influenced more are α_2 and α_3 and all the ϕ's corresponding with either one of the problematic groups. They are associated with both problematic and non-problematic entries of y_{ij}, so their standard errors should be inflated more. Finally the biggest victim will be ϕ_{23} since it is only involved with problematic y_{ij}. The parameter estimates of the MLE for the basic blockmodel seem to be in line with this reasoning. Interestingly enough, if we look at the MLE estimates of the extended blockmodel we see a similar pattern in standard error size, which is not unexpected, but there is also a flat increase of about 50% in size.

Toy Example Single Separation

β	MLE	Std. Err.	PEB
θ_0	-1.704	23.797	-0.874
α_2	-1.644	35.695	-0.605
α_3	-2.112	35.695	-0.977
α_4	1.458	23.797	0.634
α_5	1.691	23.797	0.854
ϕ_{12}	3.049	35.695	1.674
ϕ_{13}	0.684	35.696	0.000
ϕ_{14}	-2.781	23.800	-1.839
ϕ_{15}	-2.454	23.798	-1.518
ϕ_{23}	-10.843	202.273	-4.466
ϕ_{24}	0.973	35.696	0.000
ϕ_{25}	1.146	35.696	0.000
ϕ_{34}	1.259	35.697	0.000
ϕ_{35}	1.720	35.696	0.460
ϕ_{45}	-0.275	23.798	0.000

Parameter Estimates for Basic Blockmodel

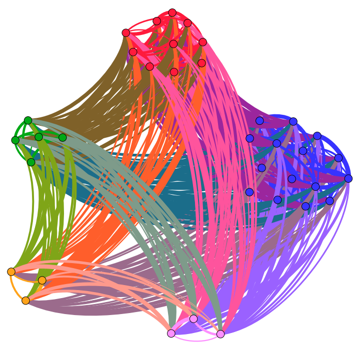
Toy Example Single Separation

β	MLE	Std. Err.	PEB
θ_0	-2.356	37.646	-1.395
γ_2	1.247	36.198	0.632
γ_3	1.100	36.198	0.485
γ_4	1.082	36.198	0.467
γ_5	0.603	36.198	-0.012
γ_6	0.779	36.198	0.163
γ_7	1.082	36.198	0.466
γ_8	1.184	36.198	0.567
γ_9	1.100	36.198	0.483
γ_{10}	0.630	36.198	0.013
γ_{11}	0.802	36.198	0.184
γ_{12}	0.288	36.198	-0.331
γ_{13}	0.426	36.198	-0.194
γ_{14}	0.682	36.198	0.063
γ_{15}	1.100	36.198	0.481
γ_{16}	-1.180	50.677	-0.234
γ_{17}	-1.743	50.677	-0.799
γ_{18}	-1.180	50.677	-0.233
γ_{19}	-1.588	50.677	-0.642
γ_{20}	-1.221	50.677	-0.274
γ_{21}	-2.144	50.678	-1.199
γ_{22}	-2.360	50.678	-1.415
γ_{23}	-1.102	50.677	-0.153
γ_{24}	-0.911	50.677	0.039
γ_{25}	-1.965	50.677	-1.016
γ_{26}	-1.861	50.678	-0.854
γ_{27}	-1.796	50.678	-0.788
γ_{28}	-1.998	50.678	-0.984
γ_{29}	-2.227	50.678	-1.207
γ_{30}	-1.733	50.678	-0.719
γ_{31}	1.919	36.198	1.029
γ_{32}	2.036	36.198	1.158
γ_{33}	1.392	36.199	0.478
γ_{34}	1.971	36.198	1.100
γ_{35}	1.657	36.199	0.775
γ_{36}	2.396	36.198	1.555
ϕ_{12}	3.139	52.125	1.833
ϕ_{13}	0.776	52.126	0.000
ϕ_{14}	-2.845	34.752	-1.805
ϕ_{15}	-2.521	34.751	-1.740
ϕ_{23}	-11.358	295.374	-5.101
ϕ_{24}	1.058	52.125	0.000
ϕ_{25}	1.228	52.125	0.176
ϕ_{34}	1.347	52.126	0.000
ϕ_{35}	1.804	52.125	0.699
ϕ_{45}	-0.347	34.750	0.000

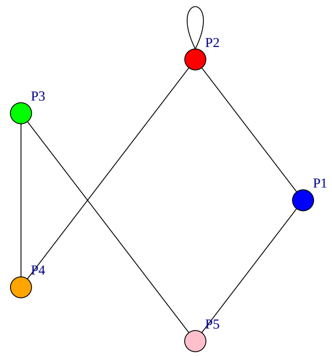
Parameter Estimates for Extended Blockmodel

3.4 Perfect Spread

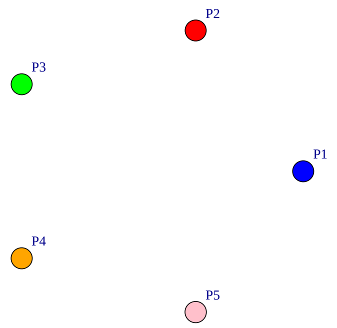
In this network all deputies have a random number of cosponsorships based on a Poisson(7) distribution. This means that we expect a large θ_0 and all other effects are expected to be zero. For the MLE the parameter estimates would change slightly upon repetition, so with really small effects the MLE does not always converge to the same graph. This example demonstrates that penalized estimation is nice, because it can set parameter estimates equal to zero. Since all interactions come from the same distribution, it is expected that all preference parameters ϕ_{rs} are zero, which is shown in the tables. The α's and γ's are not zero, because they are not penalized. There is a way to allow the MLE to also 'set estimates to zero' by adding a significance test. A reason to include a test is it can prevent false positive parameter estimates. There is also a reason to not include a test, because when we start looking at real data, the observations we have as data are not a sample, but they are the whole population. So if an effect is shown, even though it's small, we can not simply omit it. On the other hand, we view the cosponsorhips as random draws from a Poisson distribution, so if we have a small positive parameter estimate, it could be simply coincidence that it is positive. To be consistent with the networks in which separation is present, we do not omit links based on a significance test, since in those cases we would show empty graphs. In the tables we show which edges should be omitted after a significance test with an asterisk.



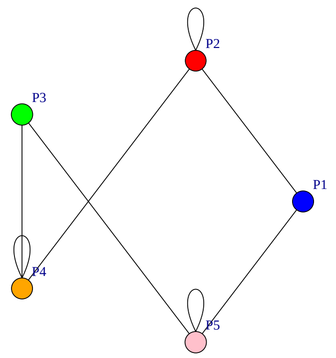
Gephi Graph



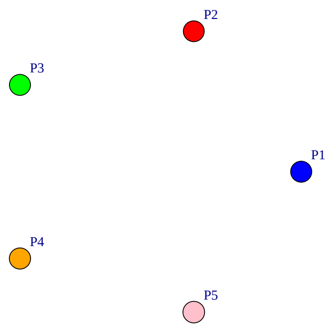
MLE for Basic Blockmodel



PEB for Basic Blockmodel



MLE for Extendedc Blockmodel



PEB for Extendedc Blockmodel

Toy Example Perfect Spread

β	MLE	Std. Err.	PEB
θ_0	1.809	0.025	1.814
α_2	-0.025	0.029	-0.019
α_3	0.018	0.036	-0.010
α_4	-0.008	0.048	0.003
α_5	0.019	0.047	0.034
ϕ_{12}	0.024*	0.038	0.000
ϕ_{13}	-0.057	0.049	0.000
ϕ_{14}	-0.002	0.063	0.000
ϕ_{15}	0.045*	0.062	0.000
ϕ_{23}	-0.013	0.055	0.000
ϕ_{24}	0.038*	0.070	0.000
ϕ_{25}	-0.056	0.071	0.000
ϕ_{34}	0.109*	0.085	0.000
ϕ_{35}	0.156*	0.083	0.000
ϕ_{45}	-0.145	0.114	0.000

Parameter Estimates for Basic Blockmodel

Toy Example Perfect Spread

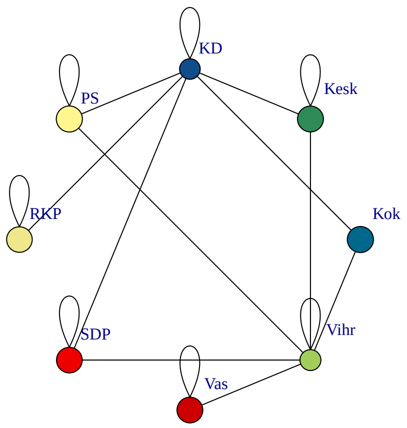
β	MLE	Std. Err.	PEB
θ	1.797	0.018	1.798
γ_2	-0.065	0.072	-0.068
γ_3	0.089	0.067	0.087
γ_4	0.024	0.069	0.022
γ_5	-0.045	0.071	-0.047
γ_6	0.029	0.069	0.027
γ_7	0.000	0.070	-0.002
γ_8	-0.065	0.072	-0.067
γ_9	0.038	0.069	0.037
γ_{10}	-0.076	0.072	-0.077
γ_{11}	0.029	0.069	0.027
γ_{12}	0.062	0.068	0.060
γ_{13}	0.075	0.067	0.073
γ_{14}	-0.092	0.073	-0.094
γ_{15}	-0.030	0.071	-0.032
γ_{16}	-0.081	0.072	-0.074
γ_{17}	-0.102	0.073	-0.095
γ_{18}	-0.025	0.071	-0.018
γ_{19}	-0.006	0.070	0.002
γ_{20}	-0.055	0.072	-0.048
γ_{21}	0.069	0.068	0.077
γ_{22}	-0.016	0.070	-0.008
γ_{23}	-0.011	0.070	-0.003
γ_{24}	0.013	0.069	0.021
γ_{25}	0.013	0.069	0.021
γ_{26}	0.008	0.073	-0.018
γ_{27}	-0.002	0.073	-0.028
γ_{28}	0.070	0.071	0.044
γ_{29}	0.018	0.072	-0.009
γ_{30}	0.023	0.072	-0.004
γ_{31}	0.058	0.077	0.072
γ_{32}	-0.032	0.080	-0.018
γ_{33}	-0.037	0.080	-0.023
γ_{34}	-0.059	0.080	-0.043
γ_{35}	0.056	0.077	0.072
γ_{36}	0.074	0.076	0.090
ϕ_{12}	0.024*	0.038	0.000
ϕ_{13}	-0.056	0.049	0.000
ϕ_{14}	-0.002	0.063	0.000
ϕ_{15}	0.045*	0.062	0.000
ϕ_{23}	-0.013	0.055	0.000
ϕ_{24}	0.038*	0.070	0.000
ϕ_{25}	-0.056	0.071	0.000
ϕ_{34}	0.109*	0.085	0.000
ϕ_{35}	0.156*	0.083	0.000
ϕ_{45}	-0.145	0.114	0.000

Parameter Estimates for Extended Blockmodel

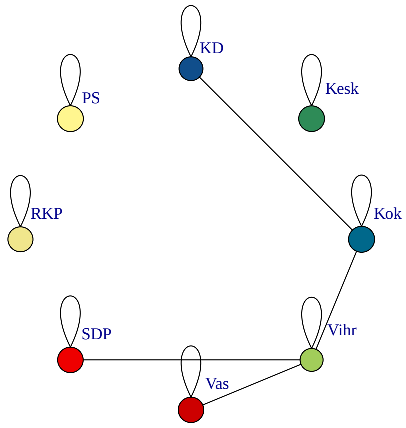
4 Real Data

4.1 Finland 2011-2014

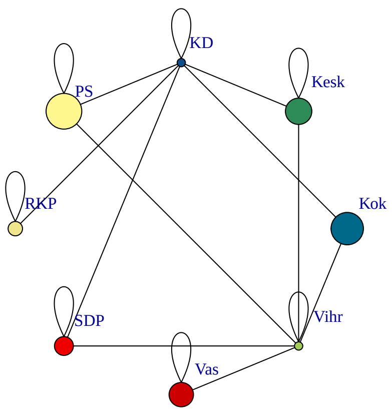
The PEB for Extended Blockmodel graph is not accurate for the PEB process. There is a strange convergence problem with *glmnet* when not penalizing the γ's. I have tried all kinds of different settings, including *lambda.min.ratio*, *nlambda*, *lambda* and even *thresh* but no sequence would lead to convergence. I investigated the condition number of X to inspect if there was a case of multicollinearity, but the condition number of Italy was far larger than the one for Finland and Italy has no issues with *glmnet* whatsoever, so I do not think there is any multicollinearity involved. However, when I let the γ's be penalized by a very tiny p_γ, there would be convergence, leading to an empty graph. So instead of not penalizing the γ's, we penalize them with the weights given by the MLE estimates. The produced graph is shown as the PEB for Extended Blockmodel.



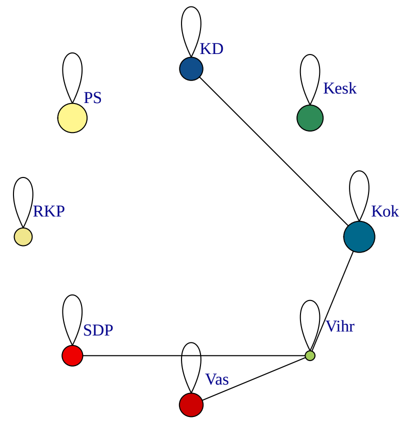
MLE for Basic Blockmodel



PEB for Basic Blockmodel



MLE for Extended Blockmodel



PEB for Extended Blockmodel

4.1.1 Interpretation

In the data there are actually ten parties, but after deleting one- and two-member parties, eight remain. The party names are the abbreviated names in the Finnish language. Also note that in this legislature we have a case of single separation (KD,Vihr). This implies we should consider the MLE models with

Information about political parties in Finland 2011-2014 according to Wikipedia (Feb 2017)
`https://en.wikipedia.org/wiki/Finnish_parliamentary_election,_2011`

Party	Ideology	Political Position	#
Kok	Liberalism, Liberal conservatism, Europeanism	Center-right	44
Kesk	Centrism, Liberalism, Agrarianism (Nordic)	Center	35
KD	Christian democracy, Social conservatism	Center, Center-right	6
PS	Finnish and Economic nationalism, Euroskepticism, Social conservatism, Right-wing populism	Social: Right-wing, Economic: Center-left	39
RKP	Swedish speaking minority interests, (Social) Liberalism	Center	9
SDP	Social democracy	Center-left	42
Vas	Democratic socialism, Eco-socialism	Left-wing	14
Vihr	Green politics, Social liberalism, Europeanism	Center-left	10

caution when it comes to edges with Vihr and KD. The Katainen cabinet was a coalition between Kok, KD, RKP, SDP, Vas and Vihr until march 2014 when Vas left.

In the MLE models, which are identical, we see some links that belong to this coalition. We see Kok-KD, KD-RKP, KD-SDP, Kok-Vihr, Vas-Vihr and SDP-Vihr. Oddly enough there is no self-link for Kok, but it does have a link with both problematic parties so that could be related. The weakest links in the network in order from weak to strong are: Kesk-Vihr, KD-PS, KD-SDP, Kesk-KD, PS-Vihr, which are the four links outside the coalition and KD-SDP. The largest of them is almost 1, but it is hard to judge its significance due to the large associated standard errors. Note that all of these are related to either KD or Vihr, the parties responsible for separation in the network, so they are likely false positives due to estimation problems.

In PEB for Basic Blockmodel we only see edges within the coalition. Kok-KD, Kok-Vihr, Vas-Vihr, SDP-Vihr all seem very reasonable given their ideological values and political positions. Even though there is separation the parameter estimates are not huge, the largest being ϕ_{38} which corresponds to the separation point (KD,Vihr).

As mentioned before the graph from the PEB for Extended Blockmodel is not reliable in the sense that it used penalized γ's which it should not have. But what it portrays makes sense, the graph is equal to the one in PEB for Basic Blockmodel.

The tables containing the parameter estimates are shown on the next page. We omit the γ's as there are roughly two hundred of them.

Finland 2011-2014

β	MLE	Std. Err.	PEB
θ_0	-2.459	3.420	-2.194
α_2	0.533	3.420	0.327
α_3	-0.967	10.259	-0.337
α_4	0.799	3.420	0.534
α_5	0.012	3.422	-0.208
α_6	0.239	3.420	-0.028
α_7	0.128	3.421	-0.057
α_8	-1.792	10.260	-1.011
ϕ_{12}	-0.483	3.421	-0.276
ϕ_{13}	1.260	10.260	0.592
ϕ_{14}	-0.488	3.420	-0.217
ϕ_{15}	-0.270	3.423	0.000
ϕ_{16}	-0.510	3.421	-0.229
ϕ_{17}	-0.524	3.422	-0.294
ϕ_{18}	1.176	10.260	0.322
ϕ_{23}	0.617	10.261	-0.146
ϕ_{24}	-0.363	3.421	-0.143
ϕ_{25}	-0.263	3.424	0.000
ϕ_{26}	-0.671	3.421	-0.478
ϕ_{27}	-1.287	3.425	-1.444
ϕ_{28}	0.107	10.263	0.000
ϕ_{34}	0.471	10.261	0.000
ϕ_{35}	1.448	10.264	0.000
ϕ_{36}	0.548	10.261	0.000
ϕ_{37}	-0.183	10.274	0.000
ϕ_{38}	-9.084	85.492	-3.878
ϕ_{45}	-0.486	3.424	-0.256
ϕ_{46}	-0.666	3.421	-0.393
ϕ_{47}	-1.007	3.423	-0.847
ϕ_{48}	0.946	10.261	0.000
ϕ_{56}	-0.654	3.426	-0.436
ϕ_{57}	-0.603	3.435	-0.336
ϕ_{58}	-0.366	10.289	0.000
ϕ_{67}	-0.150	3.423	0.000
ϕ_{68}	1.230	10.261	0.395
ϕ_{78}	1.453	10.263	0.509

Parameter Estimates for Basic Blockmodel

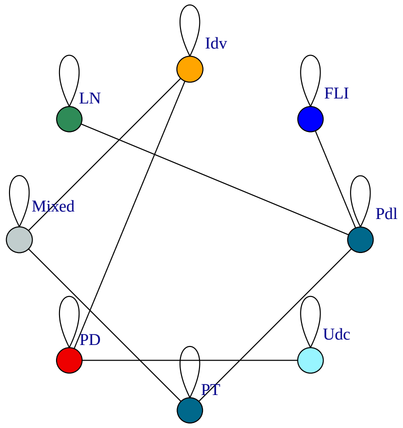
Finland 2011-2014

β	MLE	Std. Err.	PEB
θ_0	-2.455	2.126	-2.425
ϕ_{12}	-0.525	5.355	-0.298
ϕ_{13}	1.349	16.063	0.441
ϕ_{14}	-0.510	5.355	0.000
ϕ_{15}	-0.332	5.356	0.000
ϕ_{16}	-0.535	5.355	-0.293
ϕ_{17}	-0.547	5.356	-0.163
ϕ_{18}	1.262	16.063	0.158
ϕ_{23}	0.688	16.064	-0.177
ϕ_{24}	-0.404	5.355	0.000
ϕ_{25}	-0.343	5.357	0.000
ϕ_{26}	-0.715	5.355	-0.571
ϕ_{27}	-1.329	5.358	-1.506
ϕ_{28}	0.175	16.065	0.000
ϕ_{34}	0.562	16.063	-0.274
ϕ_{35}	1.499	16.066	-0.184
ϕ_{36}	0.636	16.064	0.000
ϕ_{37}	-0.093	16.072	0.000
ϕ_{38}	-9.781	133.852	-2.936
ϕ_{45}	-0.546	5.357	-0.362
ϕ_{46}	-0.689	5.355	-0.275
ϕ_{47}	-1.029	5.356	-1.038
ϕ_{48}	1.034	16.063	-0.418
ϕ_{56}	-0.717	5.358	-0.246
ϕ_{57}	-0.664	5.364	-0.058
ϕ_{58}	-0.318	16.082	0.000
ϕ_{67}	-0.175	5.356	0.000
ϕ_{68}	1.316	16.063	0.441
ϕ_{78}	1.540	16.065	0.254

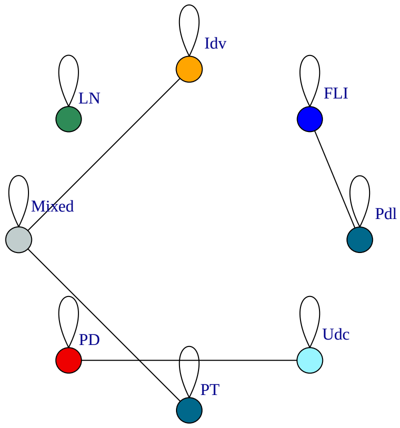
Parameter Estimates for Extended Blockmodel

4.2 Italy 2008-2013

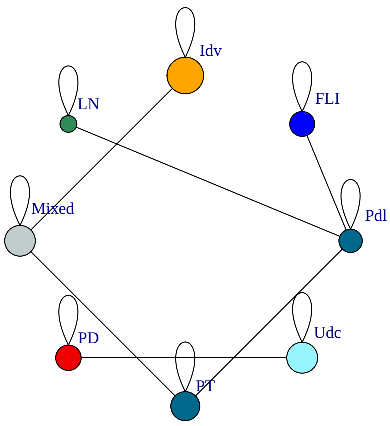
Three noteworthy things must be said for this parliament. Some small parties are grouped together as "Mixed and minor groups", this is something we get from the data. Secondly, we have corrected the group allocation of two members of this Mixed group to Idv manually. Thirdly, there was a split in 2010 in Pdl, which changed the structure of the parliament. From the split in Pdl in 2010 multiple parties were formed, such as FLI and PT. Due to this the number of seats changes over time, for varies parties. The table containing information about the parties shows the number of deputies in each party according to the data of the website of Briatte (2016).



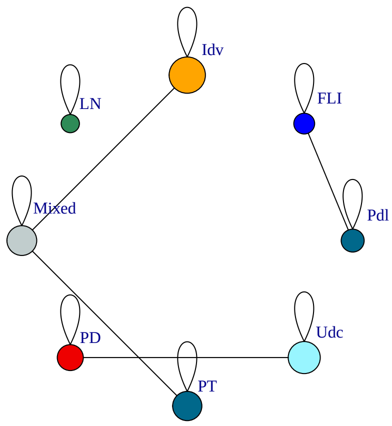
MLE for Basic Blockmodel



PEB for Basic Blockmodel



MLE for Extended Blockmodel



PEB for Extended Blockmodel

4.2.1 Interpretation

There was a center-right coalition between LN and Pdl. This link is only shown by the MLE models and is associated to a small value $\phi_{14} = 0.122$ (basic) and $\phi_{14} = 0.123$ (extended). If its standard error was slightly larger it would have been marked as omitted after a significance test, so this is a very weak link. Another link that is only present in the MLE graphs is PT-Pdl. The split party PT was supportive of the center-right coalition, but this is a weak effect $\phi_{17} = 0.185$ which would also be deemed insignificant by a test. Then there is only one more link that is not shown in the PEB graphs as well, Idv-PD which is a very weak link ($\phi_{36} = 0.001$), even though the two parties did form a center-left coalition.

Information about political parties in Italy 2008-2013 according to Wikipedia (Feb 2017)
https://en.wikipedia.org/wiki/Italian_general_election,_2008

Party	Ideology	Political Position	#
Pdl	Liberal conservatism, Christian democracy, Liberalism	Center-right	256
FLI	Liberal and National conservatism, Liberalism	Center-right	24
Idv	Big tent, Centrism, Populism, Anti-corruption politics	Center	23
LN	Regionalism, Federalism, Populism, Anti-immigration, Euroskepticism, Anti-globalization,	Right-wing	69
Mix	Mixed	Mixed	35
PD	Social democracy, Christian left, Social liberalism	Center-left	208
PT	Christian democracy, regionalism, liberalism, National and Liberal conservatism	Center-right	6
Udc	Christian democracy, Social conservatism	Center(-right)	42

The rest of the edges are also shown in the PEB graphs, so the models almost agree on the same structure. The other split party, FLI, does show a link with Pdl. There is an edge between PD and Udc, which is credible since christian democrats play an important part in the PD (they come from parties that merged into PD). The other two links are Idv-Mixed and PT-Mixed, which are hard to judge, because we do not know much about the ideology of the deputies in the Mixed group. All groups have self-links, which is to be expected. If we take a look on the website of Briatte (2016) we can click on several nodes to find out that LN mostly has cosponsorships within the party itself and Idv shows some interest in cosponsoring with the PD, but not overwhelmingly so. This explains that Pdl-LN and Idv-PD are not present in the PEB graphs.

Because Italy has already been studied in Signorelli and Wit (2016) we can compare results. The model they used was like our basic blockmodel with added covariates. For paramter estimation they used the adaptive lasso and BIC, juse like we do here, so we can compare their graph with our Basic Blockmodel PEB graph. The links their graph showed are PD-Udc, Pdl-FLI, IdV-Mixed, IdV-PT and all self-links. So their graph is very similar to ours, the difference being that they show a link between IdV and PT whereas we show a link between Mixed and PT. Apparently the slight differences in links can be explained through the addition of covariates, since that is the difference between the models used in Signorelli and Wit (2016) and the basic blockmodel we use here.

The tables containing the parameter estimates are shown on the next page. We omit the γ's as there are more than six hundred of them.

Italy 2008-2013

β	MLE	Std. Err.	PEB
θ_0	-2.519	0.033	-2.529
α_2	-0.192	0.077	-0.408
α_3	0.420	0.054	0.415
α_4	-0.640	0.050	-0.539
α_5	0.541	0.040	0.510
α_6	-0.199	0.036	-0.155
α_7	-0.055	0.094	-0.019
α_8	0.035	0.046	0.113
ϕ_{12}	1.348	0.080	1.584
ϕ_{13}	-1.189	0.075	-1.142
ϕ_{14}	0.122	0.058	0.000
ϕ_{15}	-0.618	0.050	-0.561
ϕ_{16}	-0.907	0.043	-0.935
ϕ_{17}	0.185*	0.113	0.000
ϕ_{18}	-0.376	0.056	-0.403
ϕ_{23}	-1.132	0.202	-1.545
ϕ_{24}	-0.018	0.134	0.000
ϕ_{25}	-0.692	0.137	-0.599
ϕ_{26}	-0.407	0.098	-0.189
ϕ_{27}	-1.511	0.557	-2.427
ϕ_{28}	-0.411	0.141	-0.074
ϕ_{34}	-0.659	0.125	-0.575
ϕ_{35}	0.906	0.070	0.906
ϕ_{36}	0.001*	0.067	0.000
ϕ_{37}	-0.827	0.309	-0.452
ϕ_{38}	-0.681	0.117	-0.789
ϕ_{45}	-0.981	0.110	-1.030
ϕ_{46}	-1.265	0.084	-1.411
ϕ_{47}	-0.509	0.264	0.000
ϕ_{48}	-1.292	0.143	-1.397
ϕ_{56}	-0.030	0.051	0.000
ϕ_{57}	0.764	0.139	0.539
ϕ_{58}	-0.056	0.075	0.000
ϕ_{67}	-0.167	0.134	0.000
ϕ_{68}	0.344	0.056	0.182
ϕ_{78}	-0.157	0.212	0.000

Parameter Estimates for Basic Blockmodel

Italy 2008-2013

β	MLE	Std. Err.	PEB
θ_0	-3.666	0.027	-3.636
ϕ_{12}	1.348	0.080	1.594
ϕ_{13}	-1.194	0.075	-1.167
ϕ_{14}	0.123	0.058	0.000
ϕ_{15}	-0.623	0.050	-0.561
ϕ_{16}	-0.904	0.043	-0.938
ϕ_{17}	0.178*	0.113	0.000
ϕ_{18}	-0.374	0.056	-0.396
ϕ_{23}	-1.139	0.202	-1.672
ϕ_{24}	-0.019	0.134	0.000
ϕ_{25}	-0.699	0.137	-0.546
ϕ_{26}	-0.407	0.098	-0.133
ϕ_{27}	-1.521	0.557	-2.540
ϕ_{28}	-0.412	0.141	0.000
ϕ_{34}	-0.665	0.125	-0.546
ϕ_{35}	0.895	0.070	0.880
ϕ_{36}	-0.002	0.067	0.000
ϕ_{37}	-0.841	0.309	-0.333
ϕ_{38}	-0.686	0.117	-0.797
ϕ_{45}	-0.987	0.110	-1.025
ϕ_{46}	-1.263	0.084	-1.422
ϕ_{47}	-0.517	0.264	0.000
ϕ_{48}	-1.291	0.143	-1.424
ϕ_{56}	-0.034	0.051	0.000
ϕ_{57}	0.749	0.139	0.448
ϕ_{58}	-0.062	0.075	0.000
ϕ_{67}	-0.173	0.134	0.000
ϕ_{68}	0.346	0.056	0.145
ϕ_{78}	-0.165	0.212	0.000

Parameter Estimates for Extended Blockmodel

5 Discussion

To recap we have used the basic model from (1):

$$log(\mu_{ij}) = \theta_0 + \alpha_r + \alpha_s + \phi_{rs}$$

and extended it by replacing the group productivity α_r with individual effects γ_i, leading to a model much like the model of Wang and Wong (1987). We have seen the performance of the models for various toy examples. The models actually show agreement on most of them, but in the case of single separation there is an issue with the parties associated to the separation point.

In these cases the MLE graphs of the basic and extended model agree on the same graph, as do the PEB graphs, but there are differences between the MLE and PEB graphs. The MLE graphs in this case show more links associated to the problematic parties than there probably should be. We have shown that in the case of single separation, the MLE methods are less stable, leading to large standard errors for all parameters in the model. These errors are increasing for parameters that are not-, indirectly-, and directly involved with the source of separation. In this case, the PEB graphs are definitely the better option because penalized likelihood is less vulnerable to the phenomenon of separation. The real data of Finland clearly showed exactly these issues with separation, the problematic parties had too many links with other parties.

An issue with the MLE process is that it can not select zeros for parameter estimates. Parameter estimates may fluctuate slightly, which can mean that tiny effects can change sign. To counter this we can add a significance test on the parameter estimates, but this does not work in the case of separation. The reason for this being that all standard errors are inflated, making every edge insignificant if the test would be performed. The PEB model do not need a significance test, as parameters are selected through the process of penalized likelihood and BIC.

However there was another issue that remains unresolved. The problem is that sometimes *glmnet* did not converge when penalizing all the γ's, which is strange. I have tried solving the problem by using different lambda sequences, but to no avail. My guess is that the problem involves the $penalty.factor$ of *glmnet*, but I could not pinpoint the exact issue. The only thing I know is that it only occurred in cases where there was separation involved. Either way, it prevents us from comparing the basic and extended PEB models, as the estimation process was altered for the latter.

The real data of Italy showed how well these models perform on networks where there is no separation. The MLE graphs show more links as expected and they might not all be significant. But the only differences came in the form of added weak links in the MLE graphs. On the website of Briatte (2016) we clicked on several nodes to see that these links are indeed weak and can possibly be omitted. This however, does mean that the PEB graphs are slightly more strict, and are less likely to show an unlikely edge than the MLE graphs. However, the MLE might accept weaker links, which could give a more complete picture when compared to the PEB graphs. Weak links could still be interesting, we can not simply ignore them. The same goes for would-be-insignificant links, they might be interesting depending on their strength.

Generally, in our examples (albeit toy examples or real data) the basic and extended MLE graphs agree and the PEB graphs do too. And even the MLE and PEB graphs are really similar. The only differences lie in weak links. This means that the extended blockmodel does not provide more insight into the block structure of the parliaments (the ϕ_{rs}) than the basic blockmodel. We might be allowing for some heterogeneity with this model, but apparently that does not influence the graphs at all. Additionally it requires much more computational power, especially for large parliaments. This is due to the number of parties always being somewhat the same, regardless of parliament size, whereas the number of individuals can become quite large. This means that our basic structure matrix X, which has dimensions $[n(n-1)/2] \times [(p-1) + p(p-1)/2]$ is really small compared to the extended structure matrix, which has dimensions $[n(n-1)/2] \times [(n-1) + p(p-1)/2]$. The change from p to n can be (for example in the case of Italy) a change from 8 to 663. The required additional memory actually prohibited me from doing the extended MLE computations at home for large countries on an 8GB ram computer. For this reason I learned to use the Peregrine High Performance Cluster of the RUG to do the calculations for me. Thanks to *glmnet*, for the PEB computations we can make use of sparse matrices, demanding far less memory and time for computations.

From the graphs there are no apparent differences between the basic blockmodel and the extended blockmodel, however this does not necessarily mean that the extended model was a failure. It does allow for heterogeneity even though it does not show different graphs. This just means that most likely the group effects are a really important factor in political networks. If we would predict the interactions y_{ij} for a new legislature based on the parameter estimates given by the different models they would be completely different. To show how vastly different the data would be we computed the squared correlation coefficients and sums of squared differences between the vector y_{ij} from the data and the predicted vector $\hat{y}_{ij}$ which we take to be

$$\hat{y}_{ij} = \begin{cases} exp(\hat{\theta}_0 + \hat{\alpha}_r + \hat{\alpha}_s + \hat{\phi}_{rs}) & \text{for the Basic Blockmodel,} \\ exp(\hat{\theta}_0 + \hat{\gamma}_i + \hat{\gamma}_j + \hat{\phi}_{rs}) & \text{for the Extended Blockmodel,} \end{cases}$$

with $i \in r, j \in s$. The squared correlation coefficients R^2 and the sums of squared differences SSD between y and $\hat{y}$ are shown for each model in the table below for both Finland and Italy. We can see

Finland	Basic model	Extended model	Italy	Basic model	Extended model
R^2 MLE	0.08454548	0.55717309	R^2 MLE	0.1203164	0.5215911
R^2 PEB	0.08397603	0.52435231	R^2 PEB	0.1199954	0.5201511
SSD MLE	67820.8	21235.2	SSD MLE	221577.3	142850.1
SSD PEB	67822.5	21840.8	SSD PEB	221581.1	142930.9

that there is a huge difference between the basic blockmodel and the extended blockmodel, whereas the estimation procedure only influences the fit slightly. So even though the differences are not apparent from the graphs, the fit we get from the extended model is much better.

5.1 Conclusion

In the end, we can not tell the differences between the extended model and the basic model from the graphs, but the fit of the extended model is far superior to that of the basic model. However it is computationally more demanding and it also affects the ability for *glmnet* to properly converge in some cases that are connected to separation. The MLE process is vulnerable to separation, but the deviations of the graph are somewhat predictable: they occur around the separation points. This could potentially become messy when there is more than one separation point. The PEB process seems to be more stable in this regard. For research on political models where many parliaments are compared the basic model given by (1) with penalized likelihood estimation and BIC selection might be more useful as it is computationally less demanding. However, the extended model gives parameter estimates that correlate a lot better to the data, so especially when prediction is involved the extended model should provide far superior results. Though on a side note it might also a lot more susceptible to changes in deputies and group sizes over time. For usage of the PEB process the MLE estimates are needed for the adaptive weights, so the MLE results can easily be taken into account as well. Those results can be compared, favoring the PEB graph to be more accurate, especially in cases where there is separation. Perhaps other extensions can enhance the heterogeneity we allowed for with our extended model. Research can be done along the lines of covariates or by introducing completely new parameters to the model. The problem of separation is a persistent one, which can not be circumvented without manipulating or omitting data. It could be interesting to investigate certain types of data manipulation, such that the separation disappears and the influence on the data is minimal, but even so it is not desired. Another path of investigation leads to enrichment analysis, which is not influenced by separation so it might provide an alternative solution. A drawback is that enrichment analysis is designed for binary graphs and we have weighted graphs, so there will be some complications along that route as well.

References

Anderson, C. J., Wasserman, S., and Faust, K. (1992). Building stochastic blockmodels. *Social Networks*, 14:137–161.

Bastian, M., Heymann, S., and Jacomy, M. (2009). *Gephi: an open source software for exploring and manipulating networks.* International AAAI Conference on Weblogs and Social Media.

Briatte, F. (2016). Network patterns of legislative collaboration in twenty parliaments. *Network Science*, 4:266–271.

Fienberg, S. E. and Wasserman, S. (1981). Categorical data analysis of single sociometric relations. *Sociological methodology*, 12:156–192.

Frank, O. and Strauss, D. (1986). Markov graphs. *Journal of the American Statistical Association*, 81:832–842.

Heinze, G. and Schemper, M. (2002). A solution to the problem of separation in logistic regression. *Statistics in medicine*, 21:2409–2419.

Holland, P. W., Laskey, K. B., and Leinhardt, S. (1983). Stochastic blockmodels: First steps. *Social networks*, 5:109–137.

Holland, P. W. and Leinhardt, S. (1981). An exponential family of probability distributions for directed graphs. *Journal of the American Statistical Association*, 76:33–50.

Krivitsky, P. N., Handcock, M. S., Raftery, A. E., and Hoff, P. D. (2009). Representing degree distributions, clustering, and homophily in social networks with latent cluster random effects models. *Social networks*, 31:204–213.

Santos Silva, J. and Tenreyro, S. (2010). On the existence of the maximum likelihood estimates in poisson regression. *Economics Letters*, 107:310–312.

Signorelli, M. and Wit, E. C. (2016). A penalized inference approach to stochastic blockmodelling of community structure in the italian parliament. *arXiv*, ArXiv preprint: arXiv:1607.08743.

Team, R. C. (2016). *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing, Vienna, Austria.

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society*, 58:267–288.

Wang, Y. J. and Wong, G. Y. (1987). Stochastic blockmodels for directed graphs. *Journal of the American Statistical Association*, 82:8–19.

Wikipedia Contributors (2017). *Finnish parliamentary election, 2011. Italian general election, 2008.* Wikipedia, The Free Encyclopedia.

Zorn, C. (2005). A solution to separation in binary response models. *Political Analysis*, 13:157–170.

Zou, H. (2006). The adaptive lasso and its oracle properties. *Journal of the American Statistical Association*, 101:476:1418–1429.