Poll of polls: A compositional loess model

Bergman, Jakob; Holmquist, Björn

Published in:
Scandinavian Journal of Statistics

DOI:
10.1111/sjos.12023

2014

Document Version:
Peer reviewed version (aka post-print)

Link to publication

Citation for published version (APA):
Bergman, J., & Holmquist, B. (2014). Poll of polls: A compositional loess model. *Scandinavian Journal of Statistics*, *41*(2), 301-310. https://doi.org/10.1111/sjos.12023

Total number of authors:
2

LUND UNIVERSITY

PO Box 117
221 00 Lund
+46 46-222 00 00

Poll of polls: A compositional loess model

Running title: A compositional loess model

JAKOB BERGMAN and BJÖRN HOLMQUIST
Department of Statistics, Lund University

Abstract

A large number of polls on party preferences are published today. In order to get an estimate of the changes in political opinion, the polls may be combined into a poll of polls. We discuss a method for combining polls using the fact that they are compositions and respecting the properties of the compositional sample space (the simplex). The method is easily implemented and the estimate may be computed in linear time. We provide an example with Swedish data from year 2007 to 2010. The method also allows us to present the deviations between the estimated compositions and the observed. In the data set, we note e.g. differences between different polling institutes.

Keywords

compositional weighted least sum of squares, party preferences, simplex, smoothing

1 Introduction

The results from a number of polls of political party preferences to nine Swedish groups of political parties during approximately four years are shown in Figure 1. The Swedish election system is proportional representation, and hence the total share of votes for each party is of interest. There exists, as seen in Figure 1, a large number of polls all trying to measure the same opinions. Several websites and media combine the results of multiple polls to get a (hopefully) better estimate of the current and past opinions; this is sometimes referred to as a "poll of polls". Swedish poll of polls are e.g. presented by Novus Opinion/Swedish Radio (http://www.novusgroup.se/vaeljaropinionen/ekotnovus-svensk-vaeljaropinion), Svensk Opinion (http://svenskopinion.nu) and Henrik Ekengren Oscarsson (http://www.henrikoscarsson.com). We have, however, found few suggestions in the literature how to do this. Jackman (2005) models Australian polls using a Kalman filter, and focuses on estimating the "house effects". The model is however limited to one party (proportion) and hence of less use in a situation with many parties. Thorburn & Tongur (2012) suggest using a logistic-transformed Wienerprocess.

An intuitive approach would be to fit smooth curves to the data to get estimates of the population proportions of sympathizers. One way would be to estimate each party individually and omit one party, e.g. the group "Other parties", and calculate the omitted party as one minus the sum of the rest. However, these univariate series are not independent due to the summation constraint; the sum of the proportions of all parties must always be 1. In fact, a vector of positive components summing to a constant K is a *composition*. The components of compositions are correlated (Pearson (1897) refers to this as "spurious correlation"), and hence statistical methodology disregarding this is not applicable. Dropping one group, e.g. the group "Other parties", does not remove this problem; it will still be there only less visible.

FIGURE 1 ABOUT HERE

We present a method for smoothing compositional data, which takes into account the special nature of the multivariate data being compositions. The application of the method is in no way only restricted to poll of polls, but can be utilized in any situation requiring smoothing of compositional data. This method is based on locally weighted regression (loess) introduced by

Cleveland (1979) and further developed by Cleveland & Devlin (1988). As an example we consider a data set consisting of $n = 218$ number of polls of political party preferences in Sweden. The polls extend over the time period from October 2006 to May 2010 and were performed by a number of different polling institutes including Statistics Sweden (SCB). They were all essentially telephone interviews of a number of individuals, each of which were given the question: "If it were general election today, what political party would you vote for?" The given definite answers were used to calculate proportions of the different party sympathizers, i.e. party preference compositions. The number of individuals taking part in each poll varies from poll to poll but is in general quite stable around 1000-2500 individuals. However, there also exist polls in which as many as 7000 individuals were interviewed. There are nine parts in the compositions: the four liberal/conservative parties currently in office (M, FP, C, KD), the three environmentalist/socialist parties (S, V, MP), the nationalist party (SD), and all other parties (Other).

The organization of this paper is as follows. Section 2 presents some notions in the theory of compositions, such as perturbation, power transformation and measures of distance in the simplex space. Section 3 introduces the concept of compositional loess and discuss its properties. Section 4 reports numerical results of the compositional loess technique applied to the poll of polls data of political party preferences in Sweden. The paper ends with a discussion in section 5.

2 Some theory

The sample space of a D-part composition is the simplex $\mathbb{S}^D$ defined as

$$\mathbb{S}^D = \{\boldsymbol{x} : x_i > 0, i = 1, \ldots, D, \sum_{i=1}^{D} x_i = K\}.$$

Without loss of generality, we will take $K = 1$.

There are two basic operations on the simplex analogous to addition and multiplication on the real space: perturbation (Aitchison, 1982) defined as

$$\boldsymbol{x} \oplus \boldsymbol{y} = \mathcal{C}(x_1 y_1, \ldots, x_D y_D)'$$

for $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{S}^D$, and power transformation (Aitchison, 1986, p. 120) defined as

$$a \odot \boldsymbol{x} = \mathcal{C}(x_1^a, \ldots, x_D^a)'$$

for $a \in \mathbb{R}$ and $\boldsymbol{x} \in \mathbb{S}^D$. In both cases, $\mathcal{C}$ denotes the closure operation

$$\mathcal{C}(\boldsymbol{z}) = \left(\frac{z_1}{z_1 + \cdots + z_D}, \ldots, \frac{z_D}{z_1 + \cdots + z_D} \right),$$

where $\boldsymbol{z} \in \mathbb{R}_+^D$. Aitchison (2001) showed that the perturbation and power transformation define a vector space on the simplex.

For a composition $\boldsymbol{\alpha} \in \mathbb{S}^D$, the centred log-ratio transform clr introduced by Aitchison (1983) is defined as:

Definition 1. Let $\boldsymbol{\alpha} \in \mathbb{S}^D$, then the *centred log-ratio transform* clr is

$$\operatorname{clr}(\boldsymbol{\alpha}) = \left\{ \log \frac{\alpha_1}{g(\boldsymbol{\alpha})}, \ldots, \log \frac{\alpha_D}{g(\boldsymbol{\alpha})} \right\} = \boldsymbol{a}_C,$$

where $g(\boldsymbol{\alpha}) = (\alpha_1 \cdots \alpha_D)^{1/D}$, i.e. the geometric mean, and its inverse clr^{-1} is

$$\operatorname{clr}^{-1}(\boldsymbol{a}_C) = \left(\frac{\exp a_1}{\sum_{i=1}^D \exp a_i}, \ldots, \frac{\exp a_D}{\sum_{i=1}^D \exp a_i} \right) = \boldsymbol{\alpha}.$$

The distance between compositions is measured with the Aitchison (or simplicial) distance d_S defined by Aitchison (1983) and Aitchison (1986, p. 193) as

$$d_S^2(\boldsymbol{x}, \boldsymbol{y}) = d_E^2\{\operatorname{clr}(\boldsymbol{x}), \operatorname{clr}(\boldsymbol{y})\},$$

where d_E is the Euclidean distance, or equivalently defined by Aitchison (1992) and Pawlowsky-Glahn & Egozcue (2002) as

$$d_S(\boldsymbol{x}, \boldsymbol{y}) = \sqrt{\frac{1}{D} \sum_{i<j} \left(\log \frac{x_i}{x_j} - \log \frac{y_i}{y_j} \right)^2}.$$

Locally weighted regression models are based on weighted least sum of squares estimates (WLS); hence we need to define an analogue for compositional data. Given observations of a dependent variable $\boldsymbol{y}_i \in \mathbb{S}^D$ ($i = 1, \ldots, n$) and of an explanatory variable $\boldsymbol{t}_i \in \mathbb{R}^p$ ($i = 1, \ldots, n$), we want to fit an arbitrary $\mathbb{S}^D$-valued parametric function $\hat{\boldsymbol{y}}_i = f(\boldsymbol{t}_i, \boldsymbol{\theta})$, such that the squared distances between $\boldsymbol{y}_i$ and $\hat{\boldsymbol{y}}_i$ are minimized. In estimating $\boldsymbol{\theta}$

we however believe that some observations $(\boldsymbol{y}_i, \boldsymbol{t}_i)$ are more informative and hence should have a larger influence on the estimate. Each observation is therefore assigned a weight $w_i \in (0, 1)$. Using the Aitchison distance we define a compositional weighted least sum of squares estimate.

Definition 2. The *compositional weighted least sum of squares ($\mathcal{C}$-WLS)* estimate $\hat{\boldsymbol{\theta}}$ is the $\boldsymbol{\theta}$ minimizing the compositional weighted sum of squares

$$Q_{\mathcal{C}}(\boldsymbol{\theta}) = \sum_{i=1}^{n} w_i d_S^2\{\boldsymbol{y}_i, f(\boldsymbol{t}_i, \boldsymbol{\theta})\},$$

where $w_i \in (0, 1)$, $\boldsymbol{y}_i \in \mathbb{S}^D$, $\boldsymbol{t}_i \in \mathbb{R}^p$, and f is a $\mathbb{S}^D$-valued function.

3 Compositional loess ($\mathcal{C}$-loess)

Returning to our data, we denote by $\boldsymbol{y}_i$ the different proportions of party sympathizers at time t_i, i.e. $\boldsymbol{y}_i \in \mathbb{S}^9$. In our case, the compositions are assumed to be dependent of time, but the procedure could of course use any explanatory variable or variables. Suppose that the data is generated by

$$\boldsymbol{y}_i = f(t_i) \oplus \boldsymbol{\varepsilon}_i,$$

where f is assumed to be a smooth $\mathbb{S}^9$-valued function of the real-valued variable t and $\boldsymbol{\varepsilon}_i \in \mathbb{S}^9$ is the error term. The idea is to estimate f at an arbitrary time point t_k by locally fitting a first degree polynomial $\boldsymbol{\beta}_0 \oplus (t \odot \boldsymbol{\beta}_1)$ in the simplex space. This is done by letting the closest points influence the estimates the most. Polynomials of higher order could also be fitted, if that is deemed necessary.

At each time point t_k we find compositions $\boldsymbol{\beta}_0 \in \mathbb{S}^9$ and $\boldsymbol{\beta}_1 \in \mathbb{S}^9$ that minimize

$$Q_{\mathcal{C}}(\boldsymbol{\beta}_0, \boldsymbol{\beta}_1) = \sum_{i=1}^{n} w_i(t_k) d_S^2\{\boldsymbol{y}_i, \boldsymbol{\beta}_0 \oplus (t_i \odot \boldsymbol{\beta}_1)\},$$

where $w_i(t_k)$ are the weights defined by a weight function W

$$w_i(t_k) = W\left\{\frac{d_E(t_k, t_i)}{d(t_k)}\right\},$$

where in turn $d(t_k) = d_E(t_k, \text{the } q \text{ closest } t_i)$ for some given integer q, $1 \leq q \leq n$. Here closeness is measured in a metric of the space of t, and thus

$d_E(t', t)$ is the Euclidean distance between t' and t. Following Cleveland (1979) we have chosen to use the "tricube" weight function

$$W(u) = \begin{cases} (1 - u^3)^3, & \text{if } 0 \leq u < 1 \\ 0, & \text{otherwise.} \end{cases}$$

Any other weight function satisfying the properties (2.1) in Cleveland (1979) can of course be used. The squared distance measure used here is

$$d_S^2\{\boldsymbol{y}_i, \boldsymbol{\beta}_0 \oplus (t_i \odot \boldsymbol{\beta}_1)\} = \frac{1}{D} \sum_{j<l} \left(\log \frac{y_j}{y_l} - \log \frac{\beta_{0,j}}{\beta_{0,l}} - t_i \log \frac{\beta_{1,j}}{\beta_{1,l}} \right)^2$$

where $\beta_{0,i}$ and $\beta_{1,i}$ are the components of $\boldsymbol{\beta}_0$ and $\boldsymbol{\beta}_1$, respectively.

Here $\boldsymbol{\beta}_0$ and $\boldsymbol{\beta}_1$ are compositions in $\mathbb{S}^D$ and hence the minimization of $Q_{\mathcal{C}}$ is subject to the simplicial constraints.

The local compositional weighted least sum of squares ($\mathcal{C}$-WLS) estimates corresponding to t_k are

$$(\hat{\boldsymbol{\beta}}_{0k}, \hat{\boldsymbol{\beta}}_{1k}) = \underset{\boldsymbol{\beta}_0, \boldsymbol{\beta}_1}{\arg\min}\, Q_{\mathcal{C}}(\boldsymbol{\beta}_0, \boldsymbol{\beta}_1)\,.$$

It is possible to find explicit expressions for the minimizing $\boldsymbol{\beta}_0$ and $\boldsymbol{\beta}_1$ when using the distance measure d_S^2 (see Appendix).

The method proposed here is not restricted to the use of d_S and it is possible to use other distance measures defined on $\mathbb{S}^D$. Martín-Fernández *et al.* (1999) for instance has suggested a distance measure based on divergence. Aitchison (1992) discusses the properties of various distance measures. If other distance measures are used, numerical minimization algorithms allowing for the simplicial constraints probably have to be used.

Repeating the procedure for each t_k in some set we obtain a set of estimated compositions. The locally fitted first degree polynomial composition at t_k is denoted $\hat{\boldsymbol{y}}_k = \hat{\boldsymbol{y}}(t_k)$ and is equal to $\hat{\boldsymbol{\beta}}_{0k} \oplus (t_k \odot \hat{\boldsymbol{\beta}}_{1k})$, where $\hat{\boldsymbol{\beta}}_{0k}$ and $\hat{\boldsymbol{\beta}}_{1k}$ are the locally fitted estimates corresponding to time t_k. This is then a compositional locally weighted regression, which we name $\mathcal{C}$-loess.

The $\mathcal{C}$-loess estimate $\hat{\boldsymbol{y}}_k$ is not a linear combination of the $\boldsymbol{y}_i$'s as it would be for unrestricted spaces. However, it can be shown that it can be written as

$$\bigoplus_{i=1}^{n} \{\ell_i(t_k) \odot \boldsymbol{y}_i\},$$

i.e. a compositional linear combination, where the $\ell_i(t_k)$ depend on $t_1, \ldots, t_n, W$, and q (and t_k) but not on the $\boldsymbol{y}_i$'s (see Appendix).

If we fit $\mathcal{C}$-loess estimates for the same time points as we have observed, i.e. $t_k = t_i$, we may define the squared residual deviations $d_i^2 = d_S^2(\boldsymbol{y}_i, \hat{\boldsymbol{y}}_i)$. As a measure of lack-of-fit we may use the average

$$s_{\mathrm{LOF}}^2 = \frac{1}{n} \sum_{i=1}^{n} d_S^2(\boldsymbol{y}_i, \hat{\boldsymbol{y}}_i)$$

based on the residual deviations between the fitted and observed compositions at $t_1, \ldots, t_n$. It may also sometimes be useful to plot the individual deviations d_i^2 or d_i against t_i to have a summary of the deviations over the whole range of observations. This plot may reveal time points where the smoothed result deviate very much from what is observed. Also a smoothed version of such a scatter plot may give information on the local deviations from the estimated level.

The larger q, the smoother result will be obtained from the $\mathcal{C}$-loess procedure and the larger value of s_{LOF}^2. An "optimum" smoothing parameter q can be obtained by minimizing a function of s_{LOF}^2 suitably penalized by sample size n and q.

4 Fitting the data

FIGURE 2 ABOUT HERE

We modelled the data using a number of different values of the smoothing parameter q, ranging from 10 to 150. We choose to use $q = 40$ as it seems to give a good balance between capturing the changes in trend while not being too sensitive to individual polls.

Figure 2 shows the smoothed series using $q = 40$. For the largest party in office (M) we see a clear increase around the start of the economic crisis in the second half of 2008, and at the same time a large decrease for the largest opposition party (S). We see a steady decline for C and an almost doubling in size for MP. Two parties are constantly close to the election threshold: KD is just above four per cent and SD is approaching four per cent from below. The increase for Other parties around early summer 2009 is likely due to the Pirate party's success in the European parliament elections in

June that year (see e.g. Barber, 2009).

FIGURE 3 ABOUT HERE

The sequence of residual deviations d_i for the smoothed series in Figures 2 is shown in Figure 3. There does not seem to be any major changes in deviations due to time. This is confirmed by the smoothed line which is roughly constant over time; the smaller deviation values at the beginning and the end of the series are due to edge effects.

FIGURE 4 ABOUT HERE

FIGURE 5 ABOUT HERE

In Figure 4 we see the deviations plotted for the different polling institutes. Using analysis of variance we find that there exist significant differences in mean deviations between the various institutes ($p < 0.001$), and this even when controlling for different sample sizes, which also has a significant effect ($p < 0.001$). Differences due to different institutes are sometimes referred to as “house effects”. The small deviations for Statistics Sweden (SCB) are, at least in part, explained by the fact SCB uses much larger sample sizes: approximately 6000-7000 compared to the others institutes’ sample sizes of 1000-2000. The large deviations of United Minds are probably explained by the fact that they use a web panel and not a telephone interviewing scheme. We believe that the other differences are probably due to different degrees of weighting of the proportions using previous election results and various demographic statistics.

5 Discussion

An important aspect of all data analysis is to respect the data generation process. If the question “What party would you vote for?” is asked, it seems strange to perform the analysis as if the questions

“Would you vote for party A?”
“Would you vote for party B?”
“Would you vote for party C?”
etc.

had been asked for all parties. The two set-ups might appear identical, but result in two different sample spaces. The analysis should of course always be consistent with the data generation process. In our case this means that the data is generated by counting processes where every poll results in a nine part composition.

We have introduced a method for smoothing such compositional data. The method has been developed for proportions of party preferences depending of time but can be applied to any compositional series with an explanatory variable. We have illustrated it for a one-dimensional explanatory variable (time), but the method can easily be generalized to accommodate not just one-dimensional but also multi-dimensional explanatory variables, e.g. using Cartesian coordinates to smooth spatial compositional data.

The method essentially provides a point estimate of the population proportions. (Looking at the many poll of polls seen today, this seems to be the statistic given the most attention.) As with any loess model estimating the uncertainty is very hard without distributional assumptions; in the poll of polls example it would be even harder due to different sampling designs and institute weighting procedures. This would of course be an interesting but demanding field for further research.

We have demonstrated that the smoothed series can be easily computed. The method respects the inherent properties and constraints of the sample space (the simplex). In general it will result in different values than when the traditional loess model is applied componentwise and possibly scaled to sum to 1. We therefore suggest that $\mathcal{C}$-loess is used for compositional time series and other situations which require smoothing of compositional data.

One advantage of using $\mathcal{C}$-loess for creating poll of polls is that we can get a single-valued measurement of the deviation of the poll compared to the estimation: a sort of residual estimate. The analysis of these deviations showed significant differences between the different polling institutes, i.e. house effects. A further investigation of the causes of the differences would of course be interesting and perhaps revealing. Another interesting idea would be to use the inverse mean deviation for each polling institute in weighting the polls. Hence polling institutes with greater precision would have a larger influence than institute with less precision. This, however, remains as future research.

References

Aitchison, J. (1982). The statistical analysis of compositional data. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **44**, 139–177.

Aitchison, J. (1983). Principal component analysis of compositional data. *Biometrika* **70**, 57–65.

Aitchison, J. (1986). *The statistical analysis of compositional data*. Monographs on statistics and applied probability. Chapman and Hall, London. (Reprinted in 2003 with additional material by The Blackburn Press).

Aitchison, J. (1992). On criteria for measures of compositional difference. *Math. Geosci.* **24**, 365–379.

Aitchison, J. (2001). Simplicial inference. In *Algebraic methods in statistics and probability (Notre Dame, IN, 2000)*, vol. 287 of *Contemp. Math.*, 1–22. Amer. Math. Soc., Providence, RI.

Barber, T. (2009). Centre-right parties victorious. Financial Times.

Cleveland, W. S. (1979). Robust locally weighted regression and smoothing scatterplots. *J. Amer. Statist. Assoc.* **74**, 829–837.

Cleveland, W. S. & Devlin, S. J. (1988). Locally weighted regression: An approach to regression analysis by local fitting. *J. Amer. Statist. Assoc.* **83**, 596–610.

Jackman, S. (2005). Pooling the polls over an election campaign. *Australian Journal of Political Science* **40**, 499–517.

Martín-Fernández, J., Bren, M., Barceló-Vidal, C. & Pawlowsky-Glahn, V. (1999). A measure of difference for compositional data based on measures of divergence. In *Proceedings of the 5th annual conference of the International association for mathematical geology* (eds. S. Lippard, A. Næss & R. Sinding-Larsen), vol. 1, 211–215. Trondheim, Norway.

Pawlowsky-Glahn, V. & Egozcue, J. J. (2002). BLU estimators and compositional data. *Math. Geosci.* **34**, 259–274.

Pearson, K. (1897). Mathematical contributions to the theory of evolution.—On a form of spurious correlation which may arise when indices are used in the measurement of organs. *Proc. R. Soc. Lond.* **LX**, 489–498.

Thorburn, D. & Tongur, C. (2012). Combination of sample surveys or projections of political opinions. In *Workshop of Baltic-Nordic-Ukrainian Network on Survey Statistics*, 191–197. University of Latvia, Central Statistical Bureau of Latvia, Riga.

Jakob Bergman, Department of Statistics, Box 743, SE-220 07 Lund, Sweden.
E-mail: jakob.bergman@stat.lu.se

Appendix

For each t_k we minimize

$$\begin{aligned} Q_{\mathcal{C}} &= \sum_{i=1}^{n} w_i(t_k) d_S^2\{\boldsymbol{y}_i; \boldsymbol{\alpha} \oplus (t_i \odot \boldsymbol{\beta})\} \\ &= \frac{1}{2D} \sum_{i=1}^{n} w_i \sum_{k=1}^{D} \sum_{j=1}^{D} \left(\log \frac{y_{ik}}{y_{ij}} - \log \frac{\alpha_k}{\alpha_j} - t_i \log \frac{\beta_k}{\beta_j} \right)^2 \end{aligned}$$

where y_{ij}, α_j and β_j are the components of $\boldsymbol{y}_i$, $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$, respectively, and where we in the notation have suppressed the dependence of w_i from t_k. Let $\boldsymbol{\eta}_{Ci} = \mathrm{clr}(\boldsymbol{y}_i)$, $\boldsymbol{a}_C = \mathrm{clr}(\boldsymbol{\alpha})$ and $\boldsymbol{b}_C = \mathrm{clr}(\boldsymbol{b})$, then

$$\begin{aligned} Q_{\mathcal{C}} = \sum_{i=1}^{n} w_i \boldsymbol{\eta}_{Ci}' \boldsymbol{\eta}_{Ci} - 2 \sum_{i=1}^{n} w_i \boldsymbol{\eta}_{Ci}' \boldsymbol{a}_C - 2 \sum_{i=1}^{n} w_i t_i \boldsymbol{\eta}_{Ci}' \boldsymbol{b}_C + \\ + \sum_{i=1}^{n} w_i \boldsymbol{a}_C' \boldsymbol{a}_C + 2 \sum_{i=1}^{n} w_i t_i \boldsymbol{a}_C' \boldsymbol{b}_C + \sum_{i=1}^{n} w_i t_i^2 \boldsymbol{b}_C' \boldsymbol{b}_C. \end{aligned}$$

By differentiating $Q_{\mathcal{C}}$ with respect to $\boldsymbol{a}_C$ and $\boldsymbol{b}_C$ and setting these to zero we obtain, with $d_{11} = \sum_1^n w_i, d_{12} = \sum_1^n w_i t_i$, and $d_{22} = \sum_1^n w_i t_i^2$,

$$d_{11} \boldsymbol{a}_C + d_{12} \boldsymbol{b}_C = \sum_{i=1}^{n} w_i \boldsymbol{\eta}_{Ci}$$

and

$$d_{12}\boldsymbol{a}_C + d_{22}\boldsymbol{b}_C = \sum_{i=1}^{n} w_i t_i \boldsymbol{\eta}_{Ci}.$$

These have the solutions

$$\boldsymbol{a}_C = \sum_{i=1}^{n} u_i \boldsymbol{\eta}_{Ci}$$

and

$$\boldsymbol{b}_C = \sum_{i=1}^{n} v_i \boldsymbol{\eta}_{Ci}$$

where $u_i = (d_{22}w_i - d_{12}w_i t_i)/(d_{11}d_{22} - d_{12}^2)$ and $v_i = (d_{11}w_i t_i - d_{12}w_i)/(d_{11}d_{22} - d_{12}^2)$ which both depend on t_k. Finally $\boldsymbol{\alpha} = \mathrm{clr}^{-1}(\boldsymbol{a}_C)$ and $\boldsymbol{\beta} = \mathrm{clr}^{-1}(\boldsymbol{b}_C)$.

In a similar way, expressions for compositional polynomials of higher order, e.g. $\boldsymbol{\alpha} \oplus (t_i \odot \boldsymbol{\beta}_1) \oplus (t_i^2 \odot \boldsymbol{\beta}_2)$, may also be found.

For $\boldsymbol{\alpha} \oplus (t_k \odot \boldsymbol{\beta})$ we have

$$\mathrm{clr}\{\boldsymbol{\alpha} \oplus (t_k \odot \boldsymbol{\beta})\} = \mathrm{clr}(\boldsymbol{\alpha}) + t_k\,\mathrm{clr}(\boldsymbol{\beta}) = \boldsymbol{a}_C + t_k\boldsymbol{b}_C = \sum_{i=1}^{n}(u_i + t_k v_i)\boldsymbol{\eta}_{Ci}$$

which can be written

$$\mathrm{clr}\left[\bigoplus_{i=1}^{n}\{\ell_i(t_k) \odot \boldsymbol{y}_i\}\right] = \sum_{i=1}^{n} \ell_i(t_k)\,\mathrm{clr}(\boldsymbol{y}_i) = \sum_{i=1}^{n} \ell_i(t_k)\boldsymbol{\eta}_{Ci}$$

where $\ell_i(t_k) = u_i + t_k v_i$ where u_i, v_i of course also depend on t_k as well as on W, q, and $t_1, \ldots, t_n$ but not $\boldsymbol{y}_i$.

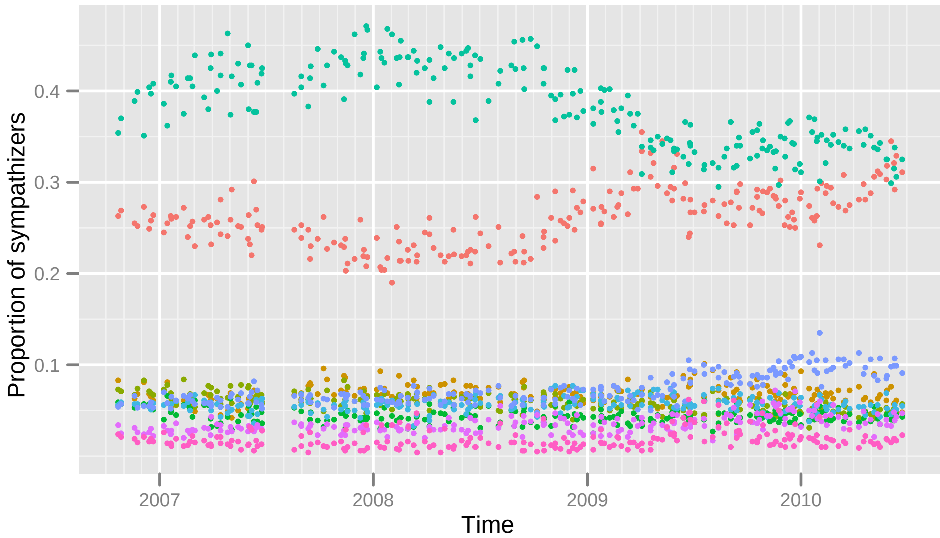

Figure 1: Polls of political party preferences to nine groups of political parties (different colours) during approximately four years.

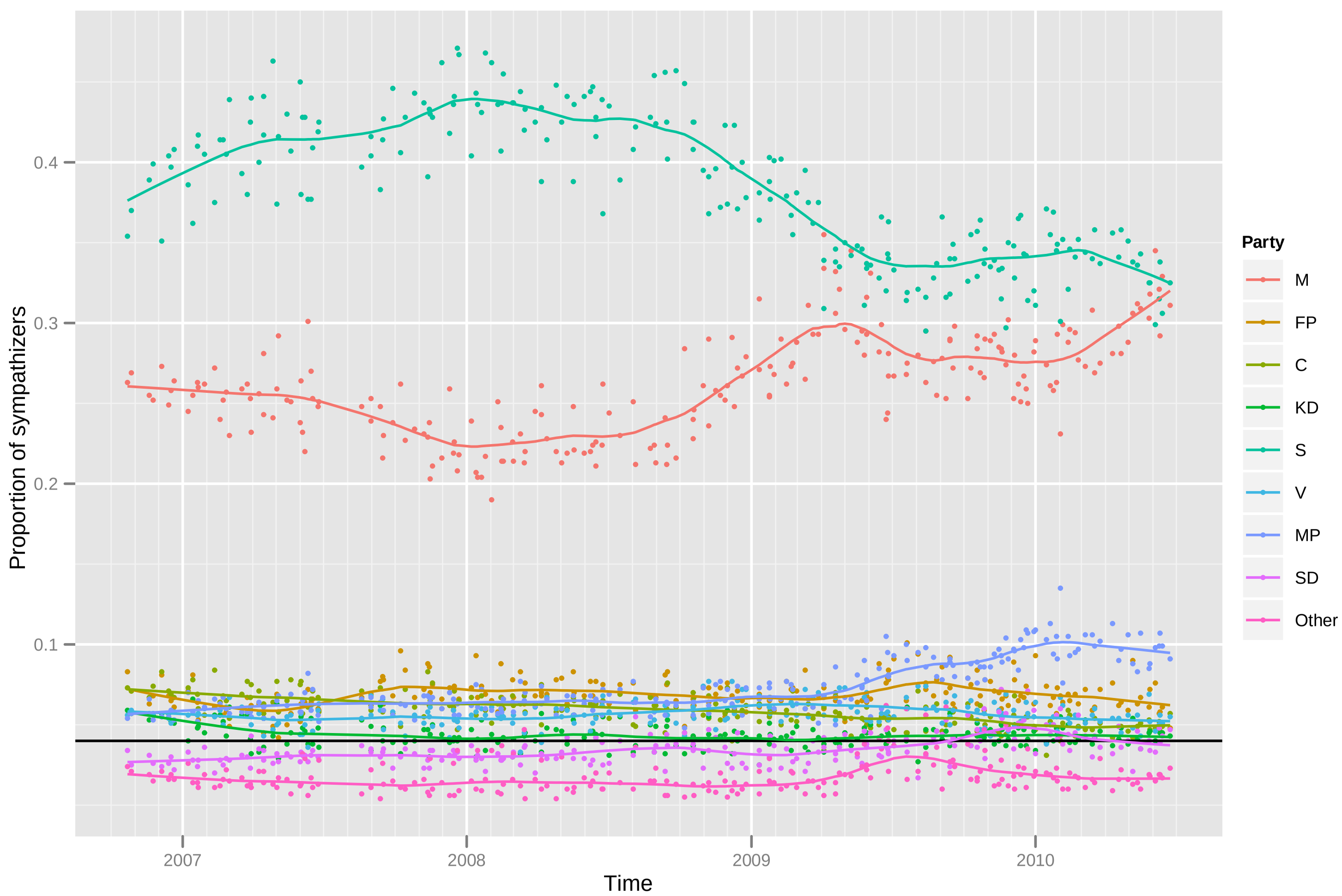

Figure 2: Smoothed Swedish opinion poll series with $q = 40$ with original observations. The solid black line represents the four per cent election threshold to win seats in the Swedish parliament.

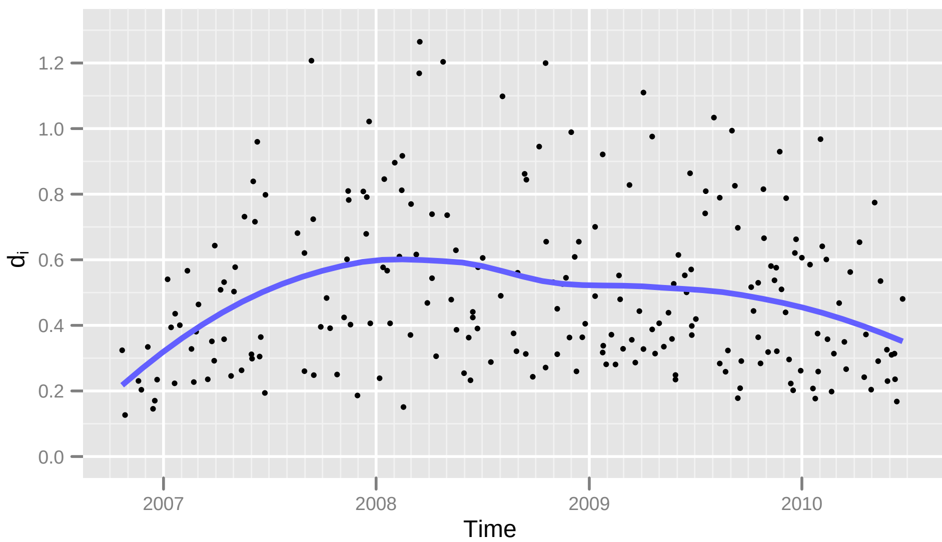

Figure 3: The residual deviations d_i from the smoothed series for $q = 40$ plotted versus time.

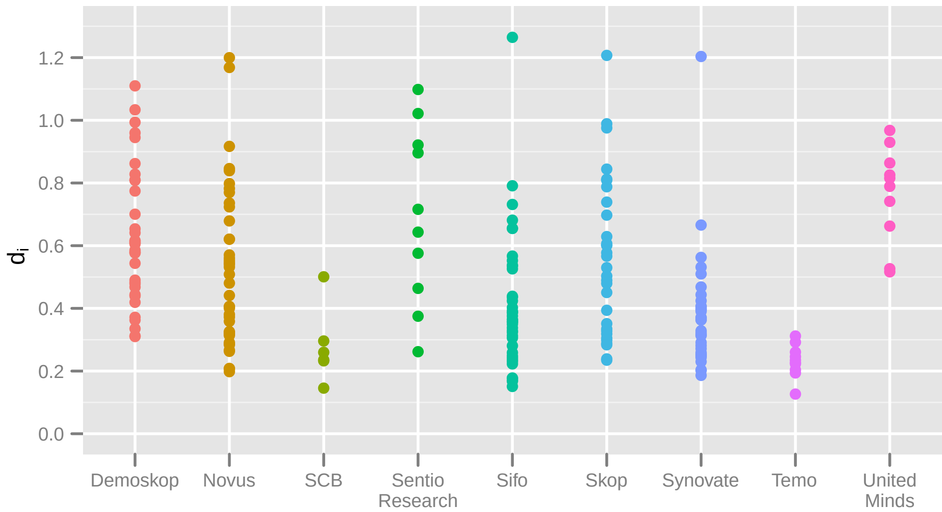

Figure 4: The residual deviations d_i from the smoothed series for $q = 40$ plotted for the different institutes.

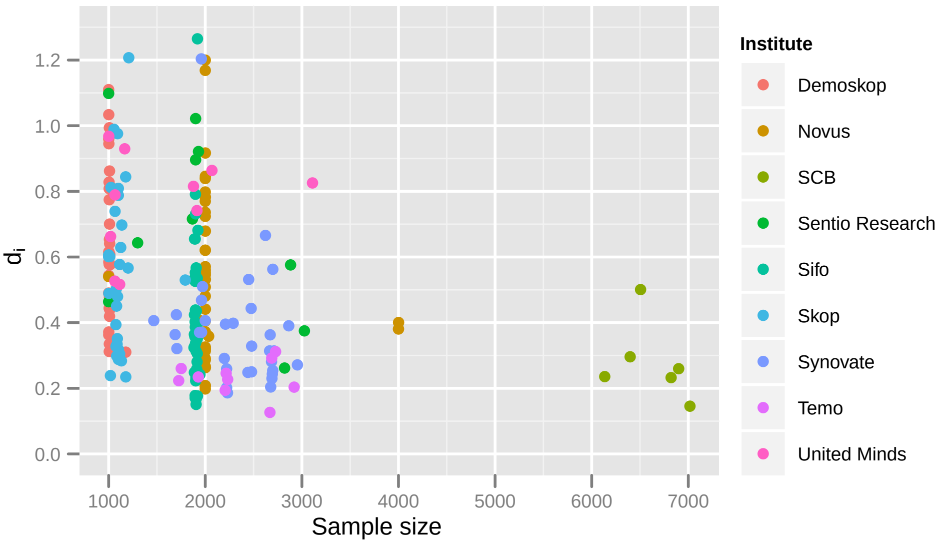

Figure 5: The residual deviations d_i from the smoothed series for $q = 40$ plotted versus sample size for the different institutes.