

# Compositional time series with varying number of parts

**J. Bergman**

Lund University, Lund, Sweden

## Abstract

Political opinion polling plays an important rôle in many countries, especially before an up-coming election or referendum. In a referendum or a bipartisan system it is usually clear what shares the pollsters report, but in a multiparty system, especially with a multitude of small parties, the choice of which parties to report is at the discretion of the pollster (or the party commissioning the poll). This can lead to different pollsters reporting different parties. As an illustration of the fact that even stable democracies can have quite volatile party structures, we note that in the last eight elections in Denmark (1990–2015) only two have been contested by the same parties.

In order to e.g. predict the outcome of elections, it has become fairly common practice to combine some or all polls to attain better information of the political opinion. However, if different pollsters report different parties, or even the same pollster reports different parties at different time points, we end up with a compositional time series with varying number of parts. A simple remedy is to look at the subcomposition of parties which are reported in all polls. However, if the goal of the analysis is a prediction of the election outcome, this is not really an option as we are trying to predict the full composition. Another option would be to amalgamate all partially reported parties into, or with, Other parties. This is a less tractable option as we lose information. Furthermore, if the election threshold is low, small parties may be important to predict, since they can affect the distribution of seats in the parliament (assuming a proportional election system). Instead, we discuss other options e.g. imputing the unreported parties. This means imputing parts in time series that may be assumed to be changing over time. The ideas are exemplified using data from Denmark and Sweden.