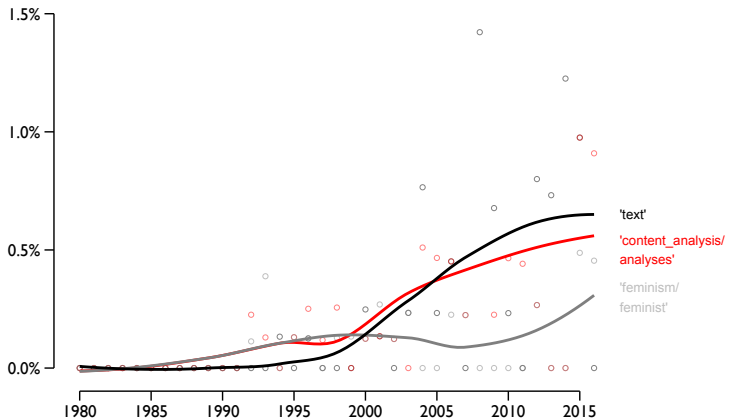


'Text as Data' in Computational Social Science

Bjørn Høyland

September 23, 2016

Article abstracts containing search terms in APSR, AJPS, JOP, PA, BJPS, EJPR, POQ



What Political Scientists do with text

- Extracting policy positions from text
- Typologizing authors on the basis of their text
- Grouping of texts by topics
- Classification by other text features (difficulty, tone, relationship between entities)

This is often done along a temporal dimension

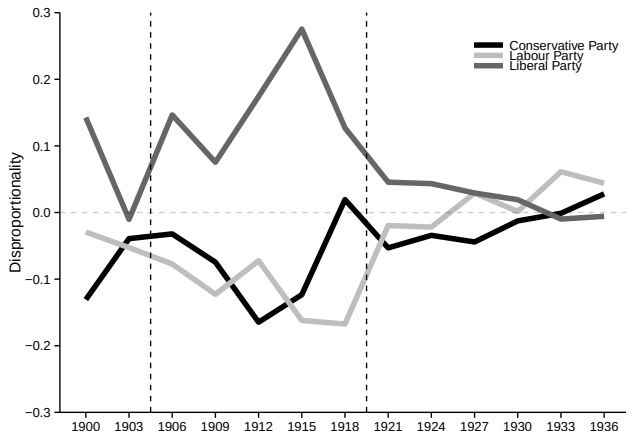
Example: Electoral Reform and Parliamentary Debates: The Norwegian Parliament before and after 1919

(with Martin G. Søyland)

- Electoral system shape legislative behavior through incentives
- **Question:** Does the electoral system matter for what gets debated in parliament?
- **Approach:** Compare topic composition before and after the 1919 Norwegian electoral reform

Electoral Reform of 1919

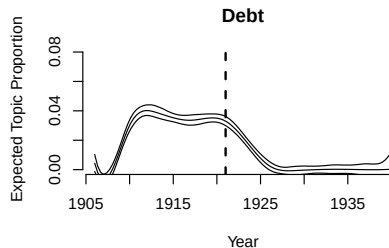
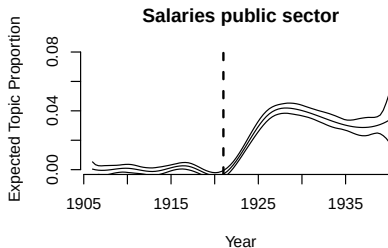
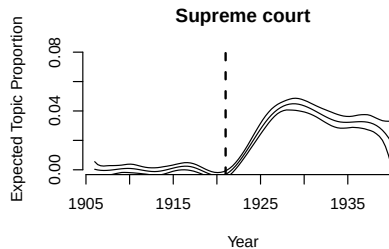
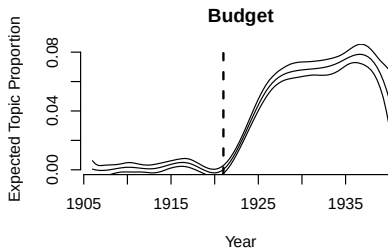
- from two-round SMD to a multimember proportional
- aim: improve seat- /vote-share ratio
- Strengthen parties



Method: Structural Topic Models

- Topic models aim at classifying debates into topics on the basis of relative word frequencies
- Unsupervised: Topics are not entered into the model!
- Mixed membership: each debate may relate to several topics
- Structural Topic Models with explicit modeling of time-trend and debating chamber

Examples of topic porportion change (preliminary)



Summary: From constituency interests to party contestation

- 1919 Electoral Reform in Norway: From two-rounds SMD to PR
- Data: All Parliamentary debates from 1906 - 1940
- Method: Structural Topic Model with time-trend, length of debate and chamber of the debate
- Finding: Many topics change substantively with the reform
 - From constituency-centric to party contestation?→ Need more refined tools and the associated computation power