Obtaining Party Positions on Immigration: Comparing Different Methods

Didier Ruedin didier.ruedin@wolfson.oxon.org

29 April 2014
Lausanne, FORS/UNIL Methods and Research Meetings
Political Text (Manifesto) $\Rightarrow$ Party Position
Different Methods

- expert positions, pooled (for comparison)
- manually coding manifests
- automatic coding
- sections on immigration
  - specific issue
  - short texts
  - emphasis of negative positions only?
- not yet done: rescaling (empirical)
Expert Positions
Manual Coding

- sentence by sentence
  - mean
  - interpolated median
- checklist
  - manifesto as unit
  - 19 questions
  - mean
  - adjustment for ‘issue space’
Automatic Coding

- dictionary of keywords (Yoshikoder)
- Wordscores
- Wordfish
- salience (being adventurous)
Data

- 8 countries: AT, BE, CH, ES, FR, IE, NL, UK
  - a priori variance in the salience of immigration
- elections between 1993 and 2013: 20 years
  - relevant parties
  - 283 manifestos, 43 elections
  - 7303 sentences coded manually
- language: only a minor problem (Switzerland)

Results
Everything Pooled

- high correlations between experts and manual (0.85), checklist (0.84)
- factor analysis
  - one factor is enough (VSS, scree)
  - same construct
  - differences in placement
  - salience (relative word count) also associated
Everything Pooled
Country-Level

- generally same patterns as overall
  - manual and checklist stable over time
- automatic methods work in some contexts
  - especially Wordscores (BE, CH, FR, NL, UK)
  - usually not stable over time
  - Wordscores consistently high in UK
- checklist > manual when very short texts (ES, IE)
Meta-Analysis

- ‘true’ correlation coefficient
- $\bar{r} = \frac{\sum r}{n}$
- Fisher z-transformation: $\bar{Z}_r = \frac{\sum Z_r}{n}$
- weighted: number of manifestos

<table>
<thead>
<tr>
<th>Experts</th>
<th>$\bar{r}$</th>
<th>$\bar{Z}_r$</th>
<th>Weighted</th>
<th>Min</th>
<th>Max</th>
<th>Median</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manual</td>
<td>0.78</td>
<td>0.83</td>
<td>0.79</td>
<td>0.42</td>
<td>0.95</td>
<td>0.86</td>
</tr>
<tr>
<td>Checklist</td>
<td>0.82</td>
<td>0.84</td>
<td>0.83</td>
<td>0.57</td>
<td>0.93</td>
<td>0.85</td>
</tr>
<tr>
<td>Wordscores</td>
<td>0.50</td>
<td>0.55</td>
<td>0.46</td>
<td>0.12</td>
<td>0.90</td>
<td>0.52</td>
</tr>
<tr>
<td>Wordfish</td>
<td>0.28</td>
<td>0.34</td>
<td>0.29</td>
<td>−0.33</td>
<td>0.81</td>
<td>0.20</td>
</tr>
<tr>
<td>Dictionary</td>
<td>0.08</td>
<td>0.08</td>
<td>0.12</td>
<td>−0.28</td>
<td>0.44</td>
<td>0.08</td>
</tr>
<tr>
<td>Salience</td>
<td>0.34</td>
<td>0.37</td>
<td>0.34</td>
<td>−0.23</td>
<td>0.78</td>
<td>0.43</td>
</tr>
</tbody>
</table>
Meta-Analysis: Countries

Correlation with Expert Positions

- Manual
- Checklist
- Wordfish
- Dictionary

Salience
Meta-Analysis: Elections

Correlation with Expert Positions

- Manual
- Checklist
- Wordscores
- Wordfish
- Dictionary
- Salience

-1.0 -0.5 0.0 0.5 1.0
Conclusion
Conclusion

- manual coding (sentence as unit of analysis)
- checklist coding (manifesto as unit of analysis)
  - resource friendly
  - ‘quite good’ for short texts
- automatic approaches with limitations
  - research question
- know your method!
- can we trust experts when salience is low?
  - using left-right positions as heuristics
- is there a ‘true’ position?