

Data report

Coding of Party Manifestos and PMs speeches in Denmark

Christoffer Green-Pedersen &

Peter B. Mortensen

Department of Political Science

University of Aarhus

Bartholins Allé

8000 Aarhus C

Denmark

Phone +4589421133

Fax + 4586139839

Email [cgp@ps.au.dk/peter@ps.au.dk](mailto:cgp@ps.au.dk/peter@ps.au.dk)

Web: [www.agendasetting.dk](http://www.agendasetting.dk)

## **Introduction**

This data rapport discusses various data issues relating to coding of party manifestos and PM's speeches in Denmark. Both data types have been coded as part of the Danish "Policy Agendas Project" (see [www.agendasetting.dk](http://www.agendasetting.dk)). The aim is to measure policy agendas through content coding of different political activities, in this case party manifestos and PM speeches. All activities are coded according to the same content scheme, which is a modified version of the original American "Policy Agendas codebook (see Baumgartner, Jones & Wilkerson 2002). The modification of the codebook to Denmark is described in Green-Pedersen (2005) and the codebook is available in English at [www.agendasetting.dk](http://www.agendasetting.dk). The coding system has 19 major and 236 subtopics. The following therefore mainly relates to other questions concerning the two datasets.

The two datasets have been coded by the same three coders. The three coders went through about a week of intensive training in the coding system, where examples and problems were discussed. Several rounds of reliability tests were then conducted where the three coders coded the same documents. The training was stopped when the level of interceded reliability had reached 85% at the sub-topic level. Coders were also continuously supervised through the coding process allowing them to ask questions about coding of specific activities. Finally, a special variable allowed coders to indicate cases where they were in doubt about the coding and these cases were then checked and sometimes recoded by the supervisors.

## **Party manifestos**

Party manifestos have been used as a data source in political science for decades due to the international “Party Manifestos Project” (cf. Budge et al. 2001). Coding of party manifestos raises two specific questions, one about the party manifesto material and one about how to code them.

### **Material**

Denmark has no elaborated tradition for party manifestos and the documents used for the Danish part of the international “Party Manifesto dataset” have been criticized because many of the documents are not really manifestos (Hansen 2008). This is, however, not a problem in this project as long as the documents reflect party election issue emphasis. Further, finding new documents years back has largely been considered impossible, also because there is no guarantee that better documents exist given the lack of a manifesto tradition in Denmark. The material has kindly been given to us by Robert Klemmensen and Martin Ejnar Hansen.<sup>1</sup> Elections from 1953 to 2007 have been coded.

The most serious problem with the material is the varying length of the documents. The average manifesto is 128 sentences long, but the longest is 1065 sentence and the shortest only 6 sentences with a standard deviation of 154 sentences. Therefore, to avoid an over-weighting of long manifestoes a general strategy in terms of analyzing the data is to calculate a separate agenda for each party and then aggregate these into an agenda of the entire party system.

### **Coding**

The original manifesto project used a coding scheme that is not well-suited for policy agenda-setting analysis. There are for instance no specific category for health care, transportation, or

---

<sup>1</sup> Unfortunately manifestos for the Christian People’s Party 1971, the Conservatives 1984, and The Progress Party 1988 are missing.

refugees and immigrants. Though it can be used for agenda-setting studies (cf. Green-Pedersen 2007) to secure comparability with the other agenda datasets, the manifestos have been recoded according to the Danish policy agenda setting codebook as described above.

The original party manifesto dataset used “quasi-sentences” as coding unit (Budge et al. 2001). In the recoding based on the policy agendas project plain sentences were used instead as this allows for a more simple coding procedure, and often sentences and quasi sentences will be similar. A sentence was defined as going from dot to the next one. Textboxes and other graphic elements were not coded.

A comparison of the number of sentences resulting from the new coding with the number of quasi-sentences revealed no major difference in most cases, and no systematic tendency with regard to the number of quasi-sentences being higher or lower than the number of sentence. However, in a number of cases, there were large differences. In some cases this seems due to the fact that the original quasi-sentences codings only included part of the documents, whereas the difference in other cases were difficult to account for.<sup>2</sup>

One of Hansen’s (2008) central criticisms of the original Danish party manifesto coding is the large number of uncodeable sentences, in average about 30%. In the recoding, this number was reduced to around 15% with a standard deviation of 13%. Thus the problem of uncodeable sentences is not just due to the material, but also the way the original coding system was implemented. The uncodeable sentences in the recoding are mainly very general declarations sentences without any policy content. Thus even though the number of un-codeable sentences in the recoded data is still relatively high compared to the original coding in other countries (cf. Hansen 2008), there is no reason that this constitute a source of systematic error in the coding.

---

<sup>2</sup> The original coder, Tove-Lise Schou, has sadly passed away.

## PM speeches

Every year, the Danish PM delivers a speech at the opening of Parliament on the first Tuesday in October. These speeches have been coded using sentences as coding units from 1953 to 2008. Since 1979, the PM has also delivered a speech to parliament in connection with the last secession of parliament before recession, normally in early June. These speeches have also been coded using sentences as described above. Both datasets have a specific code (code=2999) for sentences without any policy content.

### References:

- Budge, Ian, Hans-Dieter Klingeman, Andrea Volkens, Judith Bara & Eric Tanenbaum (2001). *Mapping Policy Preferences. Estimates for Parties, Electors, and Governments 1945-1998*, Oxford: Oxford University Press.
- Hansen, Martin Ejnar (2008). "Back to the Archives? A Critique of the Danish Part of the Manifesto Dataset", *Scandinavian Political Studies*, 31, 2, pp. 201-216.
- Green-Pedersen, Christoffer (2004). *Coding of Parliamentary Activities in Denmark 1953-2003. A datarapport*, Available at [www.agendasetting.dk](http://www.agendasetting.dk)
- Green-Pedersen, Christoffer (2005). *The Political Agenda in Denmark: Measurement and trends since 1953*, Research note available at [www.agendasetting.dk](http://www.agendasetting.dk)
- Green-Pedersen, Christoffer. 2007, "The growing importance of issue competition. the changing nature of party competition in Western Europe", *Political Studies*, 55, 4, pp. 608-628.