Danish agenda-setting project

Background:

We are currently setting up the project in a new way, so we will be a group of four people working more closely on the project. Besides the fun of working together, we hope that this closer cooperation will make it easier for us to get the maximum out of the data we have assembled. We have a fairly large research grant from the Danish Social Science Research Council which provides us with funds for data collection and different activities.

The four of us are

Assistant professor Anne Binderkrantz, research professor Christoffer Green-Pedersen, assistant professor Peter B.Mortensen and assistant professor Rune Stubager.

We are currently collecting data for this project, but also have data that Christoffer assembled for a comparative agenda setting project funded by the Danish Social Science Research Council and data that Peter collected for this dissertation. We integrate these into the new project.

Data

What data do we have:

1) Data on parliamentary activities in Denmark from 1953 to 2003. We have coded all activities in the Danish parliament accept for committee activities. This means we have coded all bills/laws (the vast majority of bills in the Danish parliament are government bills that are passed), interpellations, proposal for parliamentary resolutions, governmental accounts and questions to the minister. This database has around 82.000 units (questions, laws etc). Around 63.000 of them are questions to the minister.

   We use this data to construct two measures of the political agenda in Denmark. One is the length of debate on bills, interpellations, governmental accounts and proposal for parliamentary resolutions. The other is the number of questions to the minister. There are pros and cons of both measures. The first one has the advantage that long debates only come when many parties are interested in a topic. That makes it a good measure of the (party) political agenda. Its drawback is that N especially back in time is rather small. Once you break it down to subtopics this implies 1 or 2 or even fewer debates in each category.
   A much bigger N is the advantage of the second measure, but the drawback here is that most questions come from the opposition and this may bias it as an indicator of the (party) political agenda.

2) Budget data going back to 1973 based on Statistics Denmark's central government spending data

What data are we currently collecting:

1) A media datatset: We have summaries of news features of the Danish radio news which we are currently coding to get a database on the agenda of Danish media. The database will cover the period 1984-2004 and will contain about 180.000 news features
2) A “most important problem” dataset. The MIP questions has been included in different surveys over the years and we are working on assembling these data to have a time series of MIP questions. Hopefully they will be fairly complete back to 1980
3) We also have a plan to code the opening and closing speeches of the Danish prime minister in parliament. The speeches are held in October and May every year and will give us a measure of the government agenda.

How have they been coded:

For all datasets we have coded basic descriptive information like number of bill, length of news feature etc. We also have a brief summary (a line or so) of each item and then most interestingly a content code. This code varies across the datasets.

1) Parliamentary activities. Coded into a modified version of the policy agendas code book with 235 subcategories. Generally we deleted categories that had no use in a Danish context and split a number of US categories into two or more categories. Thus with a few exemptions we can create a dataset that is completely comparable to the US one. The datarapport from the coding of parliamentary activities, which is enclosed in this document, describes in detail the modifications we made
2) Budget data. Are coded into 13 categories. Many of them match one of the 19 main categories in the policy agendas system. Recoding the budget data into the policy agendas system is largely impossible.
3) Media data: Are coded into a simple version of the scheme used for parliamentary activities. We merged categories, so we now have 60 content categories. The data are with a few exemptions completely comparable to the parliamentary activities data
4) MIP data: These are mainly open questions, which have been coded into a content scheme which is fairly comparable to the policy agenda system.

How has it been organized?

All coding is done by paid student coders who have been intensively trained for about a week. When coding we have been in close contact with them and conducted reliability test (see the attached data rapport for parliamentary activities. If we code open and closing speeches we will consider doing so through a computer program since we then have full texts, not just short sentences as we have with parliamentary activities and news features.

What have we used the data for:

The good thing about such data is that once you have them you get new ideas. Here are some examples of the kind of studies that have been done on the data by us or people who have borrowed the data

1) A study of the long-term development of the political agenda in Denmark which shows the decline of economic issues
2) To document the development of political attention to specific issues
3) A study of questioning behavior in the Danish parliament
4) A study of how increasing political attention to issues drives budget developments (Peter's dissertation
5) Comparative studies of agenda setting dynamics

What we plan to do:

1) A statistical study of causal connections between the media agenda, voter's agenda the political/parliamentary agenda
2) A study of political reactions to media stories. We want to trace what determines whether politicians respond to news stories and how their responses keep news stories alive.
3) A study of how interest groups try to influence the media agenda.

Good advice:

Setting up databases is hard work and require an investment of time and research funds, but once the data is there, it is great.

A few "lessons from the trenches"

1) When coding parliamentary activities it is important to think about whether one wants to use the data as a measure of the political agenda (which could also be measured through activities not taking place in parliament) or because one wants to study for instance laws as such. The topics of laws may not necessarily be a good indicator of politics more generally as issues can receive a lot of attention without a laws being past. N tends also to be quite small when coding laws. There will in most countries be only 300-500 laws to code each year.
2) Make sure you have professional help with computer software. You can spend a long time on matters that take a professional 2 minutes.
3) If coding with more than 200 subcategories intensive work with the coders is crucial. Otherwise data quality becomes too low.

Further reading

Www.ps.au.dk/greenp contains codebooks etc. We are currently in the process of setting up a more professional web page.

Contact:

Christoffer Green-Pedersen

Research Professor, Ph.D.
Department of Political Science
University of Aarhus
Bartholins Allé
DK-8000 Århus C
Denmark
phone: +45 8942 1133/1297
fax: +45 8613 9839
email: cgp@ps.au.dk
web: www.ps.au.dk/greenp

Coding of Parliamentary Activities in Denmark 1952-2003

Data Rapport

Christoffer Green-Pedersen

Associate Professor

Department of Political Science

University of Aarhus

Bartholins Allé

8000 Aarhus C

Denmark

Phone +4589421133

Fax + 4586139839

Email cgp@ps.au.dk

Web: www.ps.au.dk/greenp

Version 1.3, December 2004

Introduction

The following is a data report for the project Party Competition, Agenda-Setting, and Public Policies in Western Europe.[1] The report serves two purposes. As background material for the publications from the project, it provides an introduction to the considerations behind the creation of the database on parliamentary activities from 1953 to 2003, which serves as the main empirical material for the project. Hopefully, the report will provide the reader with the kind of insight into the details of the database which can never be found in academic publications. Second, the report is meant as a source of inspiration for other researchers aiming to build similar databases. In particular, it should be of interest to others aiming to use the content categories scheme developed by the American policy agendas project (see www. http://www.policyagendas.org/) in other countries. The creation of the database benefited greatly from the experience of the American project (see Baumgartner, Jones and Macleod no date, http://www.policyagendas.org/Resources/lessons.html).

The aim of the database has been to provide quantitative data on the political agenda in Denmark from 1953 and onwards. The political agenda is defined as the issues that politicians or political parties pay attention to and is measured by coding the activities of the Danish parliament (Folketinget) (cf. Green-Pedersen 2004 for further discussion)

The following provides information on threeissues: what parts of parliamentary activity have been coded; the establishment of the content coding scheme, which has been a crucial aspects of the coding; and the organisation of the code work to ensure reliability of the data.

What has been coded?

The database contains all bills (11.952), parliamentary resolutions (4.418), interpellations debates (1.313), questions to a minister (63.737), questions to a minister during the weekly questioning hour established in 1997 (526)[2], and accounts by ministers (581) in the Danish parliament from October 1953 (beginning of the parliamentary session) to September 2003 (end of parliamentary session).[3]

[1] The project has received financial support from the Danish Social Science Research Council (grant nr. 24-03-0022), the Research Foundation at the University of Aarhus, and the Department of Political Science University of Aarhus.

[2] These questions, which were introduced in 1997, differ since the minister does not know the questions in advance and has to answer during the debate.

[3] The handbook of the parliament (Håndbog i folketingsarbejdet), which is available on the homepage of the parliament, www.ft.dk, contains detailed information on the activities of the parliament. It is available in Danish only, but the homepage of the parliament contains less detailed information in English.

These items have been coded because they are easily accessible through the parliamentary yearbook. For most activities, the summaries of the yearbook contain the necessary information to code the different variables, although for some specific variables it has been necessary to consult the parliamentary records, see below. The most important activities of parliament not coded are the activities of the parliamentary committees. They contain both a special part, which relates to the reading of bills and resolutions, and an ordinary part where the committee scrutinizes policy developments within its area. This is done by asking questions to the minister and sometimes by asking the minister to appear before the committee for consultation. Unfortunately, there are no systematic summaries of question, consultation etc., so coding this parliamentary activity would require going through the actual documents, which is far beyond the resources of this project. Therefore, the only information available on this activity is fairly general such as the number of questions asked by each committee.[4] However, this allows some comparison with the other activities of the parliament, especially the questions to the minister asked by individual members.

A codebook has been written for each activity covering all variables except the content variables which have a separate codebook (see below). The codebooks for each activity, which are in Danish only, are available on http://www.ps.au.dk/greenp/Research/Agenda.htm. However, the coding of each activity deserves some further comments.

Bills

The bills part of the database draws on an already existing database on voting behaviour in the Danish parliament.[5] The original database contained a number of variables such as which parties voted for or against the bill, whether the bill was put forward by the government or the opposition parties, so-called "private" bills, the name of the committee scrutinizing the bills etc. Some of the variables were kept in this database whereas others, especially voting patterns were kept in the original database in order for the database of this project to be more manageable. The original database was also extended in several ways. The original database only contained bills which had been either approved or rejected at the third reading in parliament. However, around one third of the bills never reach the third reading. This is either because an election is called, and they therefore have to be put forward again in the next session, or, which is most common, because they die in the

[4] The administration of the parliament has some internal databases on the activities of the committees, but they only cover a limited part of the period and are not compatible with Microsoft Software such as Access or Excel.
[5] The voting data has been collected by several people. Palle Svensson collected 1971/1972-1979/1980 and 1982/1983 to 1986/19876. Asbjørn Skjæveland collected 1980/1981, 1981/1982 and 1987/1988 to 1997-1998. Lisbeth Hoffmann Thomsen collected 1998-1999 to 2002/2003.

parliament committee to which they are referred after the first reading. This type of bills have all been added to the database. Then a number of variables have been added. This includes the content variables which have been used with regard to all activities.

Further, the extent of the parliamentary debates has been coded by counting the number of columns which the debate covers in the parliamentary records.[6] The readings of the parliament were checked to search for changes in the layout which may affect the number of words found in each column. Only one major change was found from the parliamentary year 1993-1994 to the year 1994 to 1995. The readings went from columns to pages and stopped having a fixed left margin. The person in charge of this change in parliament, however, stated that the calculations made with regard to cost when the change was made had shown that the new system would probably allow not more than a few percent more words on each page compared to two columns. Thus, one new page would largely equal two columns. The office in parliament also confirmed that no other significant changes had been made to the layout since 1953. Altogether when pages used after 1994, are divided by two, the number of columns can be compared across time.

Finally, the original database did not contain information on the parties putting forward the "private" bills and this has been added. Most coding could be done based on the parliamentary yearbooks[7], but the extent of the debates required students to go through the parliamentary records.

Parliamentary resolutions

These resolutions are mostly put forward by the opposition and can be used by the opposition to require a minister to put forward a law etc. They thus have a semi-law status and need to be read twice in parliament and debated in one of the standing committees. However, the majority of these resolutions are put forward by the opposition and "die" in the standing committee after the first reading because it is clear that they cannot gather a majority. They are thus used by the opposition as a way to draw attention to certain issues. These resolutions have been coded with regard to basic variables such as number of the resolution, which party put forward the resolution, voting patterns etc. The resolutions have also been coded according to the content variables, and the length of debate has been coded as above. Since the majority of resolutions die in the standing committees, there is no voting on them. Further, the resolutions are also used to ratify international agreements

[6] When laws are debated together in parliament, the number of columns has been divided by the number of laws.
[7] From 1996/1997 and on, all the parliamentary material can be found on the homepage of the parliament. The coders discovered that copying the wording of questions, laws etc. from this page was easier than typing it in.

etc and this has also been coded. Almost all necessary information on the resolutions could be found in summaries in the parliamentary yearbook. Only with regard to the length of debates, did the students have to look in the parliamentary records as described above.

Accounts by ministers

These accounts are normally put forward on a minister's own initiative, but sometimes a law requires such an account, for instance, every year. Finally, the committees have the opportunity to ask the minister for an account. Except in relation to parliamentary readings of EU-directives, however, this rarely happens.[8] The EU accounts have not been coded. These accounts have been coded with basic variables such as number etc., the content variables used for all activities, and variables covering whether the accounts have been debated and the extent of the debates in the same way as regards bills and resolutions. Almost all the necessary information on the accounts could be found in summaries in the parliamentary yearbook. Only when the accounts were followed by debates, did the students have to look in the parliamentary to find the extent of the debates as described above

Interpellation debates

Interpellation debates are normally asked for by members of the opposition parties, but are also by members of the government parties. This type of debates have been coded with regard to basic variables such as number, according to the content variables, and the length of debate has been coded as above. In connection with the interpellation debates, so called "proposals to be past" may be put forward and voted on. It has been coded whether such a proposal was put forward and adopted or rejected. The voting pattern has only been coded if only one proposal was put forward and only for the period 1973 to 2003. All most all the necessary information on the interpellation debates could be found in summaries in the parliamentary yearbook. Only the extent of the debates required the students to look in the parliamentary as described above.

Questions to the minister

Questions to the minister have been coded with regard to basic variables, number, who put forward the questions, answered by the minister and whether it was answered orally or in writing. The

[8] From 1986 to 1997, this was only an option for the EU committee.

questions have also been content coded in the same way as the other activities. Unfortunately, the summaries of the parliamentary yearbook had some limitations with regard to questions to the minister. From 1986/1987 and onwards, they did not contain the exact dates the questions were asked. The library of the Danish parliament then provided a pile of computer printouts with the dates. Typing all this into the database is, however, time consuming and has therefore not yet been done. More troubling was the fact that the summaries did not contain the full text, but only summaries of the questions. Therefore, we added a variable called "doubt" which could be given the value 1, if the question could not be coded based on the summaries. This was the case for 3.8% of the questions. The only solution here is to find the full text in the parliamentary records. This, however, requires significant manpower and has until now not been done due to a lack of funds. Still, a look at the distribution of these questions across categories showed no evidence that they related to specific categories and the solution is therefore to omit these questions from the analysis and thus accept a smaller N.[9] One could regard these questions as "don't know" answers in surveys.

The content coding

All parliamentary activities have been content coded using the same content scheme and this has been a crucial part of the project. The idea of the content coding has been to specify what the bill, question etc. was about. The applied content coding scheme is a modified version of the one used by the American "policy agenda project" (Baumgartner & Jones 2000; Baumgartner, Jones & Macleod no date; Adler & Wilkerson 2002).

There were two reasons for using the American codebook as a starting point. First of all, it would facilitate comparative studies greatly and such comparative studies were one of the ideas behind this study. Secondly, the American scheme had been developed using the experience obtained through the coding of a thousand of hearings, bills etc. in the US congress. Thus, this scheme would have solved many of the teething troubles that developing such coding schemes imply.

Of course, the big problem was whether the scheme would work in a Danish context. Would we end up coding a lot of questions in the "others" category for instance, thus loosing a lot of information? The strategy was simply to try it out. Together with a very enthusiastic student, Lars Beer Nielsen, who has worked on the US dataset, Christoffer Green-Pedersen started to code

[9] If the coders had a reasonable guess on the relevant category, they would code this one as well as the "doubt" variable. If they had no clue, they would code the doubt variable and one of the "others" categories based on the minister to which the question was asked.

questions. In the beginning a huge part consisted of just getting to know the US dataset. We read the codebook, coded and discussed. We also recoded American hearings. This gave us some help in understanding the coding scheme, but we also realized that with limited knowledge of American politics and policy, we often had difficulties understanding what a hearing etc. was really about. More helpful was assistance from people working with the US dataset at the University of Washington. Bryan Jones, John Wilkerson and especially T. Jens Feeley answered all our "where would you code" questions. Heather Larsen also coded several hundred Danish questions, which had been translated, and this also proved very helpful to us.

After coding a couple of thousand questions, we started to have an impression of how this would work: we found that the coding scheme was an excellent tool which with some modifications could be used on Danish data. This was substantially, of course, also very interesting. It told us that despite differences, there were clearly so many similarities in the issues having caught political attention in the two countries that the same basic coding scheme would work.

In its newest version (Adler & Wilkerson 2002), the US coding scheme had 19 main categories and 225 subcategories. The 19 main categories come quite close to the Danish ministries and the committees in the Danish parliament: there is one for agriculture, foreign affairs, traffic, justice etc. Had we designed it without paying attention to the US scheme, there would not have been a main category for Public land and water management (21) and probably one for cultural affairs, but otherwise, it would have been similar. Therefore, we saw no reason for changing the main categories. With regard to the sub-categories this was less clear. There were categories in the US scheme that were clearly useless in the Danish context. For instance, "secondary mortgage market (1410), which we still do not really know what is. Such categories were deleted. With regard to, for instance, defence issues (16), it was clear that there were many more categories in the US scheme that we would have created if not working from the US scheme, reflecting much greater interest in defence issues in the US. On the other had, the problem was not greater than we decided to keep most categories, but deleting for instance the one on veterans issues. (1609).

A more pressing problem was that with regard to a number of issues, we could not find a good category in the US system and ended up coding these questions under "general (00) or other (99). This included, for instance, issues regarding the Danish national church, private homeownership, planning issues etc. Generally, we also found that the US categories dealing with welfare state issues (health, unemployment insurance etc.) did not work that well. In most cases, this, of course, reflects substantial differences in policy. We, therefore, created new categories, for

instance 210 for issues with regard to the Danish national church, etc. As long as what we did was to create a specific category for items that had otherwise been coded together with items dealing with what we considered other issues, this is unproblematic for comparability. One can mechanically recode all the items found in the new categories into the old (American) ones.

The most problematic was when we were looking for a category, for instance, fishing issues that for long time had had its own ministry in Denmark, but in the US scheme would be spread over several categories. If a new category was created covering several US categories, comparability could not be secured through mechanic recoding. Consequently, we were reluctant to do so, but on the other hand more important was that the coding book made sense in a Danish context and we did create three such categories, one for fishing issues (408), one for private homeownership (1411) and for planning issues (712), but these are the only new categories not created by splitting others.[10] For all main categories we also kept a general (00) and an "other" category (99). Finally, we moved some of the subcategories from one main category to another. For instance, the one on immigration and refugees issues was moved from main category 5 (Labour market issues) to 2, (civil rights etc). This was done to make the placement of the subcategories more logical from a Danish perspective and thus facilitate the coding process.

Subsequently, we wrote a draft of the Danish codebook, coded more questions, revised it, until we decided that more testing was unnecessary. At that point, we had coded about 2500 questions and a few hundred other items. We had come up with a codebook (in Danish) with 19 main categories and 236 subcategories. Lars Beer Nielsen then wrote a final codebook with examples from our test-coding.

We also followed the American coding rules according to which all items should be coded in one category only. If it was impossible to decide which of two categories, the lowest should be chosen, if the categories came from different main categories, and the general (00) if they were from the same main category. We also typed in the wording of the questions, bill etc. allowing for keyword search in the database. For all items, except the interpellation debates, where summaries have been written, the exact wording of a question, bill etc. has been typed.

With regard to taxation issues, we generally followed the general rule of the American coding system that taxation of special groups and items with their own category should be coded there, for instance taxation of small business should go in 1521, whereas 107 was meant for general taxation issues and for taxation of groups and items without their own category. However, with

[10] Looking at questions to the minister only, the three categories cover around 2.4% of the questions so the comparability problem with the US data is tiny.

regard to taxation of tobacco, alcohol, cars, gas and private pension schemes, an exemption was made as these were coded 107. The reason is that taxation of such items play a huge fiscal role in Denmark and is seen more as a fiscal issue than, for instance, a health issue with regard to tobacco and alcohol.

The EU created a special problem. In the US coding scheme there was simply a sub category (1910) for Western Europe and Common Market issues. With the importance of the EU for Danish politics and policy, this was clearly not good enough. The EU is today both a substantial and an institutional aspect of Danish politics. The institutional structure of the EU, enlargement, Danish referendums are themselves political issues. However, substantial areas such as the environment, agriculture etc. are also affected by the EU. Therefore, a double system was created. If an item had EU as a substantial part, for instance was about the role of the Commission, it should be coded 1910. If it was about a substantial area such as the environment, it should be coded after this, but then we added an EU dummy variable which should be coded 1 if EU was either substantially or institutionally involved.

In appendix 1 there is a complete overview of the changes made to the US coding system. After coding thousands of items, it is possible to evaluate the functioning of the coding system. The feedback from the students doing the actual coding provided an impression that the scheme worked but afterwards, a more precise evaluation is possible by looking at the use of the general and especially the others categories. The table below shows the percentage of questions[11] in each main category coded in the 00 and the 99 categories.

Table 1. Percentage of questions in each main category coded 00 or 99

	1	2	3	4	5	6	7	8	10	12	13	14	15	16	17	18	19	20	21
00	10.6	1.1	4.5	2.8	5.4	5.7	9.2	12.5	7.7	1.3	13	17.6	7.7	5.4	2.6	0.5	3.5	3.4	0
99	2.5	0.9	3.1	0.7	1.7	8.8	1.4	0.8	0.3	8.9	2.2	11.7	8.9	3	0.5	0.4	1.4	3.3	0

The main category where the 00 and 99 subcategories have been used the most is 14 (community development and housing issues). Looking at the content of the questions that have been coded here reveals some of the reasons for this. The questions found in the 1499 are mainly about the regulation of the private (non-social), market for rental housing. Further, a lot of the questions in the 1400 category relates to regulation of the rental market in general, both social and private. The 14

[11] Since the questions to the minister in numerical terms dominate the database, the question of the use of the others and general categories have been invested using the coding of these questions only.

category is thus the only main category where, with the benefit of hindsight, we would have created more categories.

The 600 main category, education, also has a lot of questions coded as 699. Around one third of these relates to a Danish school organisation, TVIND, which has long been suspected of criminal activities and therefore has attracted considerable political attention. Most of the rest of the questions in the 699 category concerns Danish folk high schools and evening classes. The 1599 category has also been used fairly extensively. The questions in this category concern a wide range of questions relating to business and industry.

With regard to the 00 categories, several of them have been used fairly extensively. When looking through the content of the questions in these categories, it may be noted that for several of them, this just reflects the fact that questions are asked about general issues within a field. Thus, there are a large number of questions regarding the economy in general (100), energy policy in general (800), environmental policy in general (700), traffic policy in general (1000), regulation of business and industry in general (1500), educational policy in general (600) and social affairs in general (1300). However, with regard to 600 and 1000, there were more specific reasons as well. The Danish business colleges offer vocational training, higher education and education at high school levels and general questions with regard to these therefore go in 600. The Danish national railway company (DSB) has until recently also been engaged in ferry transportation and public bus services and general issues regarding this company have, therefore, been coded 1000.

Coding procedures and organisation of the codebook

Given the extent of coding-work, approximately 82.000 items have been coded, it had to involve a number of student coders. On the other hand, securing reliability of the data is crucial. Most variables such as number of a bill, whether a question has been answered orally or in written are easy to code, but the content codes are much more demanding partly because of the many categories and partly because of judgement which inevitably is involved in deciding which category is more appropriate for a question. Therefore, several procedures were set up to secure a common understanding of the coding scheme.

Most coding was done in Microsoft Access. A student, Bettina Tradsborg, created forms which allowed several coders to code at the same time and the program could be given

default values and a range of values to accept. This reduces the number of typing errors.[12] Five students, Joan Osvang Pedersen, Ane Yde Skaksen, Pi Tauber, Maren Kristina Sørensen and Lise Lotte Fogh Larsen were hired to do the coding of the questions to the minister which is the lion's share of the items, 63.737. They also did the coding of parliamentary resolutions for the entire period and all laws from 1953 to 1973. Together with Lars Beer Nielsen they also did the content coding of the laws from 1973 to 2003. Christoffer Green-Pedersen coded all accounts, interpellation debates and questions during the question hour, approximately 2000 items. Other students, Rasmus Sørensen, Sune Astrup Christensen, Ulla Parbo Hefsgaard and Thomas Laursen, were hired to go through the parliamentary records to count the number of columns which a debate covered. Lisbeth Hoffman Thomsen had worked with the already existing bills database and then wrote a codebook for it as well updated it except for the content variables. Together with Joan Osvang Pedersen, she also coded the bills that had not reached the third reading, and therefore had not been coded in the already existing databases, except for the content categories.

The five students hired to code the questions to the minister went through two weeks of intensive training in the content coding scheme consisting of the discussions of the coding scheme and test coding. For the period 1973 to 2003, we split the coding according to ministries. This was possible because the questions are listed according to the minister in the parliamentary yearbook. The idea of doing this was that most question to, for instance, the minister of labour would relate to the labour market categories and that this specialization would secure a more consistent coding. The danger of course was that the coders would loose overview of the coding scheme and if questions came up that should be coded in categories not normally used by the coders, they would be miscoded. However, for most ministries this happened rarely, so the advantages of this specialization definitely outweigh the disadvantages. Some ministers, especially the prime minister, has been asked about most issues, and the student coding this ministry then had no advantage of specialization. For the period, 1953 to 1973, each coder coded entire parliamentary years. This was easier to organize and because the number of issues covered during these years was much smaller than the years after (cf. Green-Pedersen 2004), the need for specialization was smaller.

When the coding was initiated, we set up two procedures to secure the reliability of the content coding. First, the students could put a "?" in the comments field for each items where

[12] The bills database is built on top of an earlier database which were copied into the access base. The bills that had not been included in the earlier database were coded through a form in Access. The variables that were added were coded in excel and then also copied into the Access database.

there was doubt about the coding. When each session had been coded, Christoffer Green-Pedersen went through all the items where the coders had been in doubt and sometimes recoded them. Via e-mail or orally, we also discussed the items that had caused the problem.[13] Second, Christoffer Green-Pedersen made checks by recoding randomly selected items already coded by the coders. This was done six times: five times with regard to questions and once with regard to bills, each time recoding 500 items and then checking the match with the coding by the students. In no case did the "match percentage" for the 100 questions that were recoded for each coders drop below 85% and the average for all five coders were always above 90%. If one only looked at the main categories, match percentage was always above 95%. A match in content coding required agreement with regard to both the 236 subcategories and whether or not the EU variable was coded 0 or 1. With regard to the other variables, the yearbook provided us with statistics on the number of questions asked orally, the number of questions asked by each party in parliament in each session, etc. These figures were then matched with the results from the database. The deviants here were generally minor, less than 1%. In a few cases, we found more serious disagreement and then the cause of the disagreement was found and the items recoded. By way of example, the reasons for such disagreements could be that a coder had turned around the values relating to the political party having asked a question.

Finally, a number of crosschecks have been made in the database after finishing the coding. For instance, checks have been made whether all items coded 1910 were also coded as EU related, a search was made for questions containing EU in the wording but not coded as EU related. The items found in the 99 categories used frequently have also been checked to look for systematic coding mistakes as the explanation for the extensive use of the categories. Such systematic mistakes have not been found, but in all cases a handful of items have been recoded. At the very end, Rasmus Sørensen also made a final check of the database securing that we had the same number of laws, questions etc. as the parliamentary yearbooks and checking whether typing mistakes had been made with regard to the number of columns. This also led to a few minor corrections of the database.

[13] Even though the codebook had been developed by coding approximately 2500 items, we of course constantly ran into issues that required some general decision about how they should be coded. Still, we found no revisions to the codebook necessary.

Appendix 1

Changes made to the US categories

Table 1. New categories added

New subcategory number	Content	US number
210	Questions relating to the Danish national church	207
211	Questions relating to private ownership, e.g. expropriation	200
230	Category 530 in the US scheme	530
326	Government regulation of private health care providers	321
327	Waiting lists	321
337	Diagnose and treatment of specific diseases such as cancer, hart diseases etc.	336
406	Welfare and treatment of farm animals	709
407	Environmental problems of agricultural production	700
408	Fishing issues	709, 1007, 1902
507	Public, labour market related welfare benefits, e.g. unemployment benefits, early-retirement benefits	503
508	Questions of unemployment and employment within specific sectors	502 (US 508 category changed to 1308)
712	Planning issues	200, 1403, 1405
804	Questions relating to district heating and heating supply	800
1008	Questions relating to the shipbuilding industry	1007
1308	Category 508 in the US scheme	
1411	Private homeownership	200, 1400
1910	Covers only EU issues	
1913	Non EU related issues regarding other West-European countries	1910
1916	US related questions	
2016	Regulation of municipalities and counties. 2001 covers central/local government economic relations	2001
2105	Covers issues relating to Greenland and the Faroe Islands.	

Table 2. Deleted US subcategories

Content	Deleted subcategory
Migrant and seasonal workers	529
Veterans housing	1407
Secondary housing market	1410
VA issues	1609
Census	2013
District of Columbia Affairs	2014
Native American affairs	2102

References

Adler, Scott & John Wilkerson (2002) Legislative Bills Project List of Topic and Subtopic Codes CAPP, University of Washington.

Baumgartner, R. Frank & Bryan D. Jones (2000) List of Topic and Subtopic Codes, CAPP, University of Washington.

Baumgartner, R. Frank, Bryan D. Jones & Michael C. Macleod (no date) Lessons from the Trenches. Ensuring Quality, Reliability and Usability in the Creation of a New Data Source, CAPP, University of Washington.

Green-Pedersen, Christoffer (2004) The Political Agenda in Denmark: trends since 1953, working paper Department of Political Science, University of Aarhus, Denmark.

.

Web addresses

http://www.policyagendas.org/

http://www.policyagendas.org/Resources/lessons.html

http://www.ps.au.dk/greenp/Research/Agenda.htm

www.ft.dk