Back to the archives? A critique of the Danish part of the manifesto dataset

Martin Ejnar Hansen

Abstract:
The dataset of the Comparative Manifesto Project (CMP) is one the most known and often used datasets in comparative studies of Western Europe. However, the Danish part of the data is very problematic. On average, nearly a third of the content of the Danish manifestos is deemed impossible to code. A close examination of the Danish data reveals large variation in the number of coded sentences across time and parties. An analysis of the manifestos on which the coding is based sets all the warning lights flashing. Firstly, not all manifestos are in fact manifestos; a number consists of other political texts, i.e. speeches, drafts or even stem from different elections, for instance municipal manifestos masquerading as national manifestos. Secondly, the large variation in the length of the manifestos across time suggests that the information derived from them is not readily comparable. This leads to a distorted picture of the positions of the Danish parties. The conclusion of this paper is that in order for the Danish part of the CMP dataset to become more trustworthy it is necessary to search the archives, find the actual manifestos and recode them.

In 1979 a large European research program was born, the Manifesto Research Group, later renamed the Comparative Manifesto Project (CMP). Some 25 years later the project has developed into one of the most frequently used data sources for comparative studies of political parties and political systems. A search in the Social Science Citation Index on the book *Mapping Policy Preferences* (Budge et al. 2001) yields over 110 hits.[1]

The dataset consist of party manifestos from 25 countries covering the period since 1945, all coded into 56 different categories (Budge & Bara 2001). Each sentence of the manifestos is divided into what is called quasi-sentences. Quasi-sentences are defined as an argument in which a political idea or issue is verbally expressed, and which is covered by the 56 different categories (Volkens 2001, 34). The manifestos cover a broad array of policy positions and can thus be described as "(…) a set of key central statements of party positions" (Budge 1987, 18). This procedure has created a unique dataset, which has been used in numerous publications. The dataset has acquired almost monopoly status in the comparative study of party positions; only expert surveys can be said to be able to compete with the CMP data. In view of the scope of the data, only the national election surveys are of the same calibre as the CMP data, but they unfortunately cover not nearly as extensive a time period as the CMP data.

Many publications rely either on the complete CMP dataset or parts of thereof. Denmark is included as a case in many of these studies. The reasons for including Denmark are plentiful; Denmark has had many minority governments, has relatively many parties, and stable institutions across time. All in all the Danish case provides variation across time and, combined with the other Western European countries, across sections. Examples of the inclusion of the Danish case using the CMP dataset are, for instance Bräuninger (2005), who studies the spending preferences of parties in 19 OECD countries. Adams et al. (2004) use part of the CMP data, including the Danish case, to analyze

whether parties react to public opinion or to previous election results. In her study of the success of niche parties and the reactions of mainstream parties, Meguid (2005) also relies on parts of the CMP dataset and includes Denmark in her analysis. While plenty of studies include Denmark only a few specifically use the Danish part of the CMP data. Holmstedt & Schou (1987) use the manifesto data to analyze the Danish party positions. Schou & Hearl (1992) use the Danish government declarations, which are coded using the same principles as the CMP data. In recent years Skjæveland (2003) as well as Stubager (2006) have relied on the Danish data as independent measures of party positions.

When the Danish data is analyzed more thoroughly, however, the high inclusion rate of the Danish case proves to be somewhat problematic for the results, and it is perhaps all to the good that so few studies rely solely on the Danish case. There are significant problems with the Danish part of the CMP data on three levels: the selection of the manifestos, the coding of the manifestos, and finally, the positioning of the parties on a left-right scale. The latter result is related to the other two problems, though in all likelihood not the only cause. This paper addresses these problems. Firstly, I discuss the selection of the manifestos; secondly, I address the case-specific problems with the CMP data and I conclude with a discussion the left-right placements of the Danish political parties.

Criticism of the CMP data has been relatively rare over the years. Laver & Garry (2000) and Pelizzo (2003) were among the first to provide a comprehensive critique. Benoit & Laver (2007a) have provided arguments for a general critique, whereas Hansen (2006) has presented the initial research and arguments that called attention to the problems with the Danish part of the CMP data. In this paper it is argued and shown for the first time, by way of a thorough examination of the original manifestos for the Danish parties as collected by the CMP participants,[2] that the Danish part of the CMP dataset does

indeed need challenging, and that it may be problematic to include it in any analysis. It is further argued that a thorough regathering and recoding of Danish manifestos is required if the Danish part of the CMP dataset is to be sufficiently reliable for further research.

The Comparative Manifesto Project

A central argument for using party manifestos to derive party positions on policy is that a party manifesto is the only authoritative collective statement on party policy (Budge 2001a, 50: Volkens 2001, 34). The party manifestos are usually passed by the highest organisational level of each internal party arena. The manifestos are used actively in the sense that they are written to a broader audience and are hence very straightforward (Budge 2001a, 51). Furthermore, the concept of political manifestos has been known for a long time and is still an important part of the options available to parties to make their policies known to the electorate. This makes the party manifesto suitable for research both across time and across cases. That said, one must ask whether party manifestos, speeches, or political texts more generally, can be considered reliable sources. Iain McLean (2006) gives two answers to this question, a strong answer and a weak one. The weak answer is that the data is what it is, that party manifestos allow us generate data that is better than anecdotal evidence on politics, and that is better than unsystematic sampling and far better than just ignoring it. The strong answer is that the data produces correct, reliable and valid results, which are in accordance with, for instance, data from the electorate. The party manifestos further enable to use new techniques on historical data, which is certainly not an option when using mass survey data. Furthermore, the CMP dataset is the only one that covers both policy positions and changes across time (Budge 2001a, 50). The alternatives are few, if any, if we wish to study and compare political systems with regard to policy or ideology (Mair 2001, 11), and where data alternatives do exist they will oftentimes be based on

statements by party elites or on general statements by individuals who might not readily be assumed to speak on behalf of the general party (Budge 1987, 18).

The values of the CMP data express the level of saliency accorded to various areas by the parties. This means that the party positions for or against a specific policy are not readily found in the dataset. However, this does not imply that it is not possible to measure the distance between parties. It only means that when a party does not mention a specific policy, we may only assume that this policy has a low level of salience in that party, but not that the party in question has a centre position on that policy (Volkens 2001, 40). The CMP data measures the changes in policy positions in a plausible and acceptable way (Budge 2001a, 50), since it is assumed that all parties to some extent endorse the same positions, although with varying weight (Budge 2001b, 82). While the CMP data obviously can and has been used for purposes other than positioning parties on policy dimensions, for instance to analyze political agendas (e. g. Saglie 1998; Green-Pedersen 2004), it has primarily been used as a proxy for party positions across time and countries. When the CMP data has been used it is often presented as being remarkably reliable and unproblematic for analysis since this has been done before with a great deal of success. However, as I will show in this paper that assertion is not as straightforward as it may sound if the Danish part of the data is included in an analysis.

The general critique of the CMP data
The CMP data has in later years been subject to some general criticism. Pelizzo (2003) criticizes the coding method, and the party positions derived from the data.
Another criticism is that the 56 categories were developed some 30 years ago on the basis of British general election manifestos. These categories have uncritically been employed as general categories for all other countries, and it is reasonable to assume that things have

happened that influence the political systems in Western Europe over this period. Benoit & Laver (2007a) points to the lack of an uncertainty measure connected to the coded material. In the same vein there is no measure for the reliability of the coders, and when recoding is carried out by other coders different results may well occur, as reported in Budge (2001a). Another general criticism against the CMP data is that despite being available for a period covering many years, the raw data is not readily available in the sense that it is not possible to apply other coding procedures to evaluate the CMP coding itself. It may be argued that the CMP dataset is still undergoing developments and should thus not be made available to others than members of the research group. Against that argument speaks that the CMP project has been generously supported by various public research funds, including the Danish Social Science Research Council, and many of these research funds have a policy that data collections that have received support should, after a period of time, be deposited in some publicly accessible data archive, and while the CMP dataset itself is available, the manifestos on which it is based are not.

The Danish part of the data

The general criticism of the CMP aside, it is evident that there are several problems with the Danish part of the data. First, one problem is the manifestos on which the coding is based. The selection of which manifesto to code and thus include in the dataset has not been handled in a systematic way. Indeed, there are several examples of texts included in the dataset where the term 'manifesto' would need much stretching in order to cover. Local candidates' electoral material or speeches have been included as being representative of the national party, and they are not easily describable as authoritative statements of the parties' policies.

A second problem with the Danish part of CMP data is that in the case of the Danish parties, around 30 percent of all quasi-sentences are deemed uncodeable (Volkens 2001, 41), which means that on average, one third of the contents of the Danish party manifestos is not included in the analysis. Holmstedt & Schou (1987) argue that the Danish political system is so different from those in other nations that coding the missing 30 percent would provide unusable results since they are nationally specific areas that cannot be coded. That argument is very interesting, especially when in certain areas of political science the general idea is that the Scandinavian political systems are very much alike (e. g. Lijphart 1984; 1999). As shown in table 1 there is a large difference between the highest number of uncoded sentences, which is Denmark, and the second highest number of uncoded sentences, which is found in Sweden. The standard deviation for Denmark is 18 percentage points. Not even countries with fundamentally different political cultures like Israel and Turkey where country-specific policy areas are in play that are difficult to code in the 56 categories have the same high level of uncoded sentences. Another interesting feature is that the percentage of uncoded sentences for Sweden is unknown for 1945-1985. This is also the case for Norway where we do not know the level of uncodeable quasi-sentences from 1945-1993. Thus, the percentage of uncoded sentences in these two countries might in fact be higher than the Danish case, it could be the same, or it may be lower. However, as we do not know, using data covering the periods with missing information from these two countries becomes even more problematic than it is to use data with many uncoded sentences.

--- TABLE 1 AROUND HERE ---

One explanation for the 30 percent uncodeable sentences in the Danish case may be that there is high variation across time. And if only the results from the last 10-15 years are used the problem becomes minor. Although the data from the later period contains a lower percentage of uncodeable sentences, it is still significantly higher than the data for the other countries. Figure 1 provides a view on the development in the percentage of uncodeable sentences. We see that the percentage of uncodeable sentences does not drop below 20 percent until the election of 1994. And a drop below 10 percent is not achieved before the last period 1998 election, and is lost again at the 2001 election. If we want to use a long time series, then the average number of uncodeable sentences is 30.2 percent, with a standard deviation of 18.5 percentage points. This is compared with only 6.1 percent for the CMP dataset as a whole excluding Denmark.

--- FIGURE 1 AROUND HERE ---

The relevant question to pose here is of course: do the 30 percent uncoded sentences actually have an impact? Or is it just a problem for statisticians and strict data disciplinarians? That 30 percent is missing is a problem for the data is indisputable. But if in the coding procedure some form of sampling was used we might be able to accept a high level of missing data. Nowhere in the descriptions of the CMP data is there any mention of the use of some form of sampling technique. From table 2 we also learn that if any form of sampling did occur it was entirely unsystematic. And since we do not know which 30 percent was not coded we cannot solve the problem by correcting the data statistically. This raises an even more important question: if 30 percent of all sentences were not coded correctly, how can we be sure that the remaining 70 percent were coded correctly? To answer this question we can begin by analyzing whether the problem is

evenly distributed among all parties or is related to particular parties. From table 2 it is very clear that the uncoded sentences occur in all parties, and within the parties there is a high degree of variation from manifesto to manifesto, standard deviations of more that 20 percentage points in the direction of serious errors connected with the data. Only thorough empirical studies of the coded manifestos will confirm or reject whether or not the 30 percent uncoded sentences has an impact on reliability. However, as will be discussed later on, it is not as easy a task as it may seem.

--- TABLE 2 AROUND HERE ---

The selection and collection of Danish manifestos

Another problem with the Danish part of the CMP data is the actual manifestos that were coded. Firstly, we must distinguish between manifestos that are election manifestos and party programmes. The latter are the parties' general statements of their goals, and one which may not undergo as much change as the election manifestos. For instance, the Social Democrats had the same party programme for nearly 50 years (Holmstedt & Schou 1987, 180). Two problems arise when using manifestos as basis for the coding. Manifestos do not necessarily include statements of the parties' goals and attitudes in all policy areas. A manifesto comprises party positions in those areas that the party believes to be important in the coming election. Secondly, it is not unreasonable to assume that the parties, especially in recent years, have more than one manifesto on the shelf aimed at different segments of the electorate.

Brief examples of these two problems are the Social Liberal manifesto from 1981 where youth unemployment is the main theme, whereas the Christian Democrats in the same year think that getting a new government is the most important issue. In 1990 we

find two different sources for the Social Democratic manifestos, an advertisement or insert in the party newspaper *Det Fri Aktuelt,*[3] and another, and probably more general, manifesto. In 1990, the Social Liberals believe that unemployment is the most important issue in the campaign. The Liberals believe that it is the policy toward Europe that is important, while the Conservatives have a general manifesto. The three parties in question started the campaign as members of the same coalition government. Finally, there is no information as to which manifestos were coded for the 1984 election. This missing information is increasingly problematic should one wish to replicate the coding using other tools. In addition, the lack of a systematic choice of manifestos does not make any it easier to judge which manifestos may have been coded in 1984. Incidentally, some of these examples also highlight the problems with the saliency theory on which the CMP dataset is supposedly grounded. Budge (2001b, 82) notes that all party programmes to an extent endorse the same position. If this were the case we would expect all manifestos at an election to endorse one position on one policy dimension, although with varying emphasis. And while there undoubtedly is something to this argument in some elections we cannot on the whole accept this argument since there is much variation in both the issues covered in the manifestos and the length of them.

The varying length of the manifestos is another issue that leads to further frustration. Some variance in the length of party manifestos is of course to be expected, and from table 3 we learn that this average differs from party to party, but we also find that there are very large differences between the minimum and maximum number of words among the parties.[4] Oftentimes the maximum length of a manifesto is more than ten times the length of the shortest one recorded for a particular party. This may not be problem in itself; however, it is problematic that a manifesto of 594 words is compared with one of 15273 or, for that matter, 738 with 4241. The higher the number of words, the more

information can be extracted from a manifesto. We also see that across time the average manifesto from a left-aligned party is some 500 words longer than the average one from a party aligned to the right of centre.

This difference is not problematic as such. However, in the averages are hidden the exact same extremes described previously. If we move to consider the standard deviations of each party, we see that it is not just a case of a single outlier. In most of the cases, the standard deviations are very high and this indicates that there might be a serious problem with the differences in the information provided.

--- TABLE 3 AROUND HERE ---

As the length of the manifestos does provide us with information it is necessary to take a more in-depth look at the manifestos that have actually been coded before we can pass judgement on the data as a whole. In this section, I address issues with the manifestos party by party from the 1953 election until the 2001 election, focusing not only on the parties that gained representation but on all manifestos found in the dataset. Some issues are dealt with at length while other are just mentioned. Some of the examples are that the 1987 manifesto for the Centre Democrats is an issue of their party magazine which, besides the policy goals, also in the foreign policy section includes the verbatim text of a telegram received by party leader Erhard Jacobsen from President Ronald Reagan. Comparing the manifestos of the Communist party from the September 1953 election and the 1960 election, we find that in the former, it is the central committee of the party that issues the manifesto, while the latter consists of speeches or statements by the party leader, which may be party policy, but is no manifesto in the sense of the other Communist manifestos.

The manifesto of the Justice Party for the September 1953 election has in its first lines a disclaimer that clearly says where the full manifesto by the parliamentary group and the board of the party organisation can be found: in the party newspaper. However, the manifesto included in the data is a selection of policies from that manifesto made by a Social Democratic think tank,[5] and some of their comments are also included. The coding instructions for the CMP coders state that if a manifesto could not be obtained, a selection of it from newspapers or other sources may be appropriate. However, in this particular instance this raises two questions: first, and actually valid in general as well, how can we be sure that the selection truly represents the whole of the party manifestos, and second, how can we ensure that political opponents' representation of a manifesto is equal to that of the party's own? The answers to either question are troublesome for the CMP data. The answer to the second question should be somewhat straightforward. If we accept that political parties compete for votes, then political parties cannot be trusted to represent the cases of other parties. The answer to the first question is more nuanced, even though any selection of contents in manifestos must be made at random if we are to accept such a selection. A selection made by a newspaper is not random, nor does it necessarily represent the salience that the party puts on issues; it may just as well be a selection of issues that corresponds to the issues selected by other parties, and if that is so, this approach violates one of the key components of the whole CMP project, namely that parties find different issues salient.

For the 1994 election, which was held in September, the Unity List manifesto stems from the local election in Copenhagen in November 1993. Very few national issues are mentioned and setting manifestos from local elections equal to those from national elections does not enhance the reliability of the data. The manifesto coded for the Liberals in the 1964 election is not a manifesto but rather a speech by the later minister Anders

Andersen circulated in the Hornslet constituency. Anders Andersen was not party leader in that election. The programme coded for the same party in 1966 is a letter to the electorate of the Give constituency which, although important, is not where general party policy is made.

In 1987 the coded manifesto from the Socialist Peoples' Party is a draft, not a final manifesto, which is also true for the Social Liberal manifesto of the 1964 election. Along the same lines we find the manifesto from the Independent party for the September 1953 election, which is not a party manifesto but a declaration of intent by an exploratory committee to form a new party, while a policy mentioned in this declaration lacks certain features of the manifesto, for instance, no mention of the name of the party for which it is supposed to be a manifesto. The 1966 and 1968 manifestos for the same party are advertisements from a local paper, which does contain some policy goals but seems to be somewhat different from what would normally constitute a manifesto. The 1973 manifesto for the Left Socialists is not an election manifesto but the party programme. At other elections the Left Socialists present a specific election manifesto instead of their party programme.

The positioning of the Danish parties

After each manifesto has been coded we are left with the values of 56 categories. In order to produce a left-right scale of the parties some sort of transformation of these values must take place. This can be done in numerous ways. Gabel and Huber (2000) present four different methods, including the method originally suggested by Laver and Budge (1992), which is probably used the most. It is a simple summation of 13 categories characterized as right wing subtracted from the summation of 13 other categories characterized as left wing (Budge and Klingemann 2001, 21-24). The categories have been selected by the CMP

participants on both theoretical grounds and on the basis of factor analysis of the data. Even though the method is simple, it has a high degree of explanatory power (Gabel and Huber 2000).

There are of course other methods. Recently Frantzmann and Kaiser (2006) have suggested a new method for deriving party positions which takes into account the valence issues not included in the left-right positions. This method would, like Elff (2002), make it possible to derive party positions on other dimensions as the left-right. Such a feature would be useful if the political system is indeed multi-dimensional, as claimed by some electoral researchers. Multi-dimensionality does not appear to have an influence on the left-right dimension recovered from the CMP data. In the research utilizing the CMP data, the left-right scores are by far the most used. Thus, in this article we limit ourselves to the left-right placement of the Danish parties as reflected in the CMP data. The left-right placement of the Danish parties is shown in figure 2.

--- FIGURE 2 AROUND HERE ---

Regrettably, the party placements in figure 2 are not easily recognizable for anyone with just a smattering of knowledge about Danish politics since 1945. However, evaluating party positions is no easy task. The CMP dataset is often used as a "gold standard" (Benoit and Laver 2007b), and it is indeed the only dataset covering the entire period. That leaves two options when evaluating the performance of the data in correctly placing the parties from left to right. First, we can turn to the various expert surveys that exist. This provides independent estimates of party placement, but unfortunately, such surveys have not been systematically collected for the entire period. It is only for the last twenty-something years that we have valid and useful expert estimates on party positions. The second possibility is

13

to analyse party placement based on the face validity of the placements. This is the method applied in this article, and as will be shown, it is a powerful and very simple tool to point up the problematic nature of the left-right placement of the Danish parties.

That the Socialdemocrats in 1971 under the leadership of Jens-Otto Krag should be the most left wing party is quite surprising. It is equally surprising that the entry of the Progress Party in the 1973 landslide election should have catapulted the party into the very centre of Danish politics. Most students of Danish politics would place the Progress Party on the right, probably as the most right wing party of all. It is therefore also a novelty that the Liberal Party should be the most right wing party in 1987, 1988 and 1994. Indeed, if we compare the positions of the Progress Party and the Liberal Party over the entire period, we find that the Progress Party is more often placed to the left of the Liberal Party than to the right. It should be the other way round. It is not only to the right of the political spectrum that we find interesting party positions. In 1987 and 1990, the Social Liberal Party is supposedly more left wing than the Socialist People's Party; especially the 1990 position is very unusual – it is not standard procedure in Danish politics that a party, which at the time of the publication of its manifesto still participated in the bourgeois government, is more to the left than a committed socialist party. Another party on the left is the Unity list consisting of former members of the Danish Communist Party, the Socialist Workers Party and the Left Socialists. In the 1998 election, they are in the CMP data placed to the right of the Socialist People's Party and three years later in the 2001 election to the right of both the Socialist People's Party and the Social Democrats. Likewise the Socialist People's Party is placed to the left of the Left Socialists at every election from 1975 to 1987. If these positions, and thus the value of each category, are right it may very well shake the foundations of the left wing of Danish politics. Overall, very few of the party positions can be accepted without a liberal interpretation of Danish

party system. Yet, surprisingly Budge and Klingemann (2001, 39) in their discussion of the face validity of the left-right positions, finds that the Danish party positions are "…quite plausible." In fact, the attempt by Budge and Klingemann to assert the validity of the Danish data is less than successful, and marred by the exclusion of any party formed after 1960 from their discussion.

The problems with the positions of the Danish parties can either be grounded in problems with the data from which the positions are generated, or we can accept the positions as realistic and valid. The latter is not an option if we are to give any weight to the arguments presented so far. If problems with the CMP data were found only in the Danish part of data, we might let it pass. However, Pelizzo (2003) has found similar problems with the positions of the Italian parties. Whether this problem constitutes as a large a problem in other countries as it does for the Danish case is beyond the scope of this article. Nevertheless, the argument must be that the problems with the positions of the Danish parties are probably based on the selection and coding of the manifestos.

Is there an alternative to the CMP data?

Whether as measures for saliency or positions, the CMP data is much used by scholars comparing the political systems of Western Europe. In 2003, the CMP dataset won the award for best comparative dataset from the Comparative section of the American Political Science Association. A tremendous amount of work has been invested in creating the dataset. However, as argued in this paper there are problems with the data, both at a general level and at the country specific level. The CMP data is nevertheless used widely, which may be because the problems found are case specific rather than general or it may be that there are relatively few other alternatives. Let us briefly consider those alternatives.

15

The alternatives must be of a kind that can be assumed to equal the manifestos of the parties, and with that as the limit there are three alternatives: 1) speeches by party leaders, 2) expert surveys and 3) parliamentary voting behaviour. Alternative 2 on expert surveys was addressed previously, and to recapitulate, there are many benefits associated with using expert surveys. However, they are normally only available for a limited time period and there are problems with the regularity with which they are performed. The first alternative was used by, for instance, Laver & Benoit (2002) and Giannetti & Laver (2005), who utilized speeches by parliamentarians to derive the ideal point positions of the parties or individual legislators. Nevertheless, there are also disadvantages to this approach. First of all, the speeches must be analyzed by some method. Both Laver & Benoit (2002) and Giannetti & Laver (2005) use the wordscores approach formulated by Laver, Benoit & Garry (2003) which may be the best available option at this point. Secondly, a choice has to be made concerning which speeches to include on what topics and by whom, and though it is now fairly common to find speeches online, that is not the case if we wish to expand our analysis across time.

The second alternative, expert surveys, has probably been the most successful competitor to the CMP data. The expert survey by Castles & Mair (1984) was the first of a series of large scale expert surveys of party positions, and it was followed by Laver & Hunt (1992); Huber and Inglehart (1995), and lastly, Benoit & Laver (2006). Expert surveys benefit from the multiple observations of each case limited only by the number of experts available, thus enabling not only a point estimate of a position but also an error term. Another argument in favour of using expert surveys is that they can be performed independently of elections. However, the downside is that they cannot be performed back in time, so the time series available only run from around 1980 until today, and nor are the same questions posed in all surveys, which might make comparisons

problematic. While expert surveys are much used in the comparative study of European political systems, which is not the case for the third alternative. This is either due to the argument made by Michael Laver in a review article (Laver 2006), that parliamentary votes are heavily influenced by institutions such as party discipline and government participation, so they cannot really be used, or it maybe because it is very costly to acquire parliamentary voting data and that the analytical tools needed to use such data are beyond the reach of many researchers. The conclusion of the review of the alternatives is thus that few exist, and that they are either costly and/or lack the time-series perspective represented by the CMP data. Regrettably, the trustworthiness of the CMP data can not be as high as its frequent use should prerequisite. The problems put forward in this paper, albeit case specific, does raise questions about the usefulness of the dataset.

Concluding remarks

When certain theories or data acquire a dominant position within an area it is normally because they perform well and cover large areas, or it may simply be lack of competition. However, the hegemony of one theory or dataset does not imply that it is beyond criticism. And as has been shown in this paper, there is good evidence in support of handling the CMP dataset with some caution.

The problems with the CMP data falls in two parts: general and case-specific. The general problems may be assumed to be less important to people who want to utilize the long time-series in the CMP data. However, the case-specific problems for Denmark as presented in this paper ought to serve as a warning before including the Danish part of the data. The large number of uncoded sentences is the first sign that something is not as it should be. The large differences in the lengths of the coded manifestos also indicate that something is amiss. Calculating left-right positions of Danish parties generates results

which in many years produce a completely distorted picture of the actual positions of Danish political parties, and when it is further found that it is in some cases not actual manifestos but drafts, speeches, local platform or opposition research notes that were coded, then the bright red warning lights should really begin to light up.

We must go back to the archives and find the real and full manifestos of the Danish parties and completely recode the correct manifestos to eliminate the obscure number of uncoded sentences before we can rely on Danish part of the CMP data. In fact, it might be an idea to comb through the entire CMP dataset and weed out any faults and flaws in other countries.

References

Adams, J., Clark, M., Ezrow, L. & Glasgow, G. (2004): 'Understanding Change and Stability in Party Ideologies: Do Parties Respond to Public Opinion or to Past Election Results?', *British Journal of Political Science* 34, 589-610

Benoit, K. & Laver, M. (2007a): 'Estimating party policy positions: Comparing expert surveys and hand-coded content analysis', *Electoral Studies* 26, 90-107

Benoit, K. & Laver, M. (2007b): 'Benchmarks for text analysis: A response to Budge and Pennings', *Electoral Studies* 26, 130-135

Benoit, K. & Laver, M. (2006): *Party Policy in Modern Democracies*, London: Routledge

Bräuninger, T. (2005): 'A partisan model of government expenditure', *Public Choice* 125, 409-429

Budge, I. (1987): 'The internal analysis of election programmes' in Budge, I., Robertson, D. & Hearl, D. (eds.): *Ideology, Strategy and Party Change: Spatial Analyses of Post-War Election Programmes in 19 Democracies*, Cambridge: Cambridge University Press

Budge, I. (2001a): 'Validating the Manifesto Research Group approach: theoretical assumptions and empirical confirmations' in Laver, M. (ed.): *Estimating the Policy Position of Political Actors*, London: Routledge

Budge, I. (2001b): 'Theory and measurement of party policy positions', in Budge, I., Klingemann, H-D., Volkens, A., Bara, J. & Tanenbaum, E.: *Mapping Policy Preferences*, Oxford: Oxford University Press

Budge, I. & Bara, J. (2001): 'Introduction: Content Analysis and Political Texts', in Budge, I., Klingemann, H-D., Volkens, A., Bara, J. & Tanenbaum, E.: *Mapping Policy Preferences*, Oxford: Oxford University Press

Budge, I. & Klingemann, H-D. (2001): 'Finally! Comparative over-time mapping of party policy movement' in Budge, I., Klingemann, H-D., Volkens, A., Bara, J. & Tanenbaum, E.: *Mapping Policy Preferences*, Oxford: Oxford University Press

Budge, I., Klingemann, H-D., Volkens, A., Bara, J. & Tanenbaum, E.: *Mapping Policy Preferences*, Oxford: Oxford University Press

Castles, F. G. & Mair, P. (1984): 'Left-Right Political Scales: Some "Expert" Judgements', *European Journal of Political Research* 12, 73-88

Elff, M. (2002): 'Parteinsystem, Sozialstruktur und Wahlabsicht', in Brettschneider, F., van Deth, J. W. & Roller, E. (eds.), *Das Ende der politisierten Sozialstruktur?*, Opladen: Leske und Budrich

Franzmann, S. & Kaiser, A. (2006): 'Locating Political Parties in Policy Space: A Reanalysis of Party Manifesto Data', *Party Politics* 12, 163-188

Gabel, M. J. & Huber, J.D. (2000): 'Putting Parties in Their Place: Inferring Party Left-Right Ideological Positions from Party Manifesto Data', *American Journal of Political Science* 44, 94-103

Giannetti, D. & Laver, M. (2005): 'Policy Positions and Jobs in the Government', *European Journal of Political Research* 44, 91-120

Green-Pedersen, C. (2004): 'Center Parties, Party Competition, and the Implosion of Party Systems. A study of centripetal tendencies in multiparty systems', *Political Studies* 52, 324-341

Hansen, M. E. (2006): 'Er dårlige data bedre end ingen data? En analyse af brugbarheden af de danske valgprogramdata', *Politica* 38, 210-223

Holmstedt, M. & Schou, T-L. (1987): 'Sweden and Denmark 1945-1982: election programmes in the Scandinavian setting' in Budge, I., Robertson, D. & Hearl, D. (eds.): *Ideology, Strategy and Party Change: Spatial Analyses of Post-War Election Programmes in 19 Democracies*, Cambridge: Cambridge University Press

Huber, J. D. and Inglehart, R. (1995): 'Expert Interpretations of Party Space and Party Locations', *Party Politics* 1, 73-112

Laver, M. (2006): 'Legislatures and parliaments in a comparative context', in Weingast, B. & Wittman, D. (eds.): *Oxford Handbook of Political Economy*, Oxford: Oxford University Press

Laver, M. & Budge, I. (1992): 'Measuring Policy Distances and Modelling Coalition Formation' in Laver, M. & Budge, I.: *Party Policy and Government Coalitions*, London: Macmillan

Laver, M. & Hunt, W.B. (1992): *Policy and Party Competition*, New York: Routledge

Laver, M. & Garry, J. (2000): 'Estimating policy positions from political texts', *American Journal of Political Science* 44, 619-634

Laver, M. & Benoit, K. (2002): 'Locating TDs in Policy Spaces: Wordscoring Dáil Speeches', *Irish Political Studies* 17, 59-73

Laver, M., Benoit, K. & Garry, J. (2003): 'Extracting Policy Positions from Political Texts Using Words as Data', *American Political Science Review* 97, 311-331

Lijphart, A. (1984): 'Measures of Cabinet Durability – A Conceptual and Empirical Evaluation', *Comparative Political Studies* 17, 265-279

Lijphart, A. (1999): *Patterns of Democracy*, New Haven: Yale University Press

Mair, P. (2001): 'Searching for the positions of political actors: a review of approaches and a critical evaluation of expert surveys' in Laver, M. (ed.): *Estimating the Policy Position of Political Actors*, London: Routledge

McLean, I. (2006): 'The Dimensionality of Party Ideologies', in Bara, J. & Weale, A. (eds.): *Democratic Politics and Party Competition*, London: Routledge

Meguid, B. M. (2005): 'Competition Between Unequals: The Role of Mainstream Party Strategy in Niche Party Success', *American Political Science Review* 99, 347-359

Pelizzo, R. (2003): 'Party Positions or Party Direction? An Analysis of Party Manifesto Data', *West European Politics* 26, 67-89

Saglie, J. (1998): 'A Struggle for the Agenda', *Party Politics* 4, 347-366

Schou, T-L & Hearl, D. (1992): 'Party and Coalition Policy in Denmark', in Laver, M. & Budge, I. (eds.) *Party Policy and Government Coalitions*, Basingstoke: St. Martin's Press

Skjæveland, A. (2003): *Government Formation in Denmark 1953-1998*, Århus: Politica

Stubager, R. (2006): *The Education Cleavage: New Politics in Denmark*, Århus: Politica

Volkens, A. (2001): 'Manifesto research since 1979: from reliability to validity' in Laver, M. (ed.): *Estimating the Policy Position of Political Actors*, London: Routledge

Volkens, A. (2002): Dataset CMP2002, Comparative Manifestos Project, Science Center Berlin, Research Unit Institutions and Social Change (Director H.-D. Klingemann) in cooperation with the Manifesto Research Group (Chairman I. Budge)
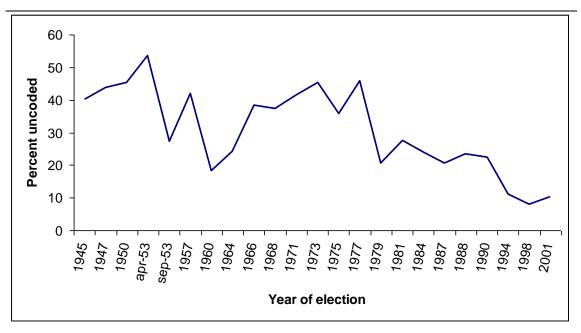
Table 1 Percent of quasi-sentences deemed uncodeable

Country	Pct. uncodeable	Standard deviation (pct. points)
Denmark	30.2	18.5
Norway[a]	0.5	36.0
Sweden[a]	12.5	42.9
Finland	2.3	3.5
Iceland	4.2	7.1
Turkey	0.5	1.1
Israel	7.7	10.5
Japan	6.0	7.5
Ireland	3.0	4.8

Source: Derived from Volkens (2002).

Note a: Sweden: only manifestos from 1985 and onwards are included. Norway: only manifestos from 1993 and onwards are included.

Figure 1 The number of uncodeable sentences in Danish manifestos 1945-2001

Table 2 Average percentage of uncoded quasi-sentences for Danish parties 1945-2001

Party	Pct.	Standard deviation (pct. point)
Socialdemocratic Party	29.1	19.7
Social Liberals	27.2	15.9
Conservative	21.4	13.2
Centre Democrats	34.6	13.0
Justice Party	32.7	21.2
Socialist Peoples Party	24.7	15.9
Communist Party	44.0	11.6
Liberal Centre	16.8	-
Danish Peoples Party	19.6	21.8
Common Course	5.8	-
Christian Democrats	16.8	6.8
Independent Party	25.9	12.3
Liberals	24.8	14.3
Left Socialists	44.2	24.1
Progress Party	30.5	16.6
Unity List	6.8	3.8

Source: Derived on the basis of Volkens (2002).

Note: For Liberal Centre and Common Course there is only one observation.

Table 3: Average number of words in the coded Danish party manifestos 1945-2001

Party	Average	Min. length	Max. length	Std. dev.
Socialdemocrats (A)	2524.13	594	15273	3124.88
Social Liberals (B)	1102.42	349	4190	813.92
Conservatives (C)	1532.79	110	4101	1243.13
Centre Democrats (D)	3404.23	561	17513	4540.99
Justice Party (E)	1417.63	445	3674	930.42
Socialist Peoples Party (F)	3656	958	10413	2368.25
Communists (K)	987.65	268	2689	683.13
Liberal Centre (L)	5017	5017	5017	-
Danish Peoples Party (O)	574.50	551	598	33.23
Common Course (P)	903	903	903	-
Christian Democrats (Q)	3145.93	452	16371	4296.04
Danish Unity (S)	1353	1077	1629	392.90
Independent party (U)	1868.33	438	3506	1186.95
Liberal party (V)	1841.17	586	9521	1939.97
Left Socialists (Y)	2425.88	738	4241	1353.93
Progress Party (Z)	2775.91	368	11089	3034.08
Unity List (Ø)	1914.50	1142	2768	830.23
Average Right	1963.72	-	-	2473.58
Average Left	2377.01	-	-	2397.57

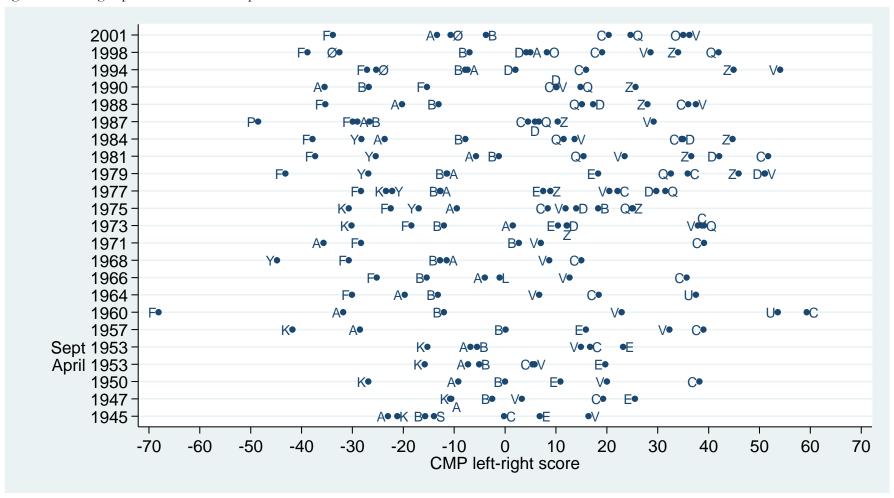Note: Number of words is calculated using the Wordscores application.

Figure 2: Left-right positions of Danish parties 1945-2001

Note: Derived from Volkens (2002). Only parties who won representation in the parliament in the election are included. Party abbreviations as in table 4

Notes

1 The exact number of hits was 114 on 2007-03-24.

2 If the manifestos were not found in the archives of the Danish participant in the manifesto project, they were obtained either from the Danish National Archive or from the parties. The titles of the manifestos analyzed in this paper correspond to the titles found in the codebooks of the CMP project.

3 *Det Fri Aktuelt* was a newspaper, now defunct, that supported the Social Democrats.

4 I have relied on the number of words instead of the number of quasi-sentences. However, as quasi-sentences to some extent are functions of the number of words and various grammatical stops, it should not alter the results significantly.

5 The think-tank was AIC (*Arbejdernes Informationscentral*) which was created to monitor and produce political propaganda against political adversaries, especially the communists. AIC also worked together with the national intelligence services.