

Issue Attention in West European Party Politics

CAP and CMP Coding Compared

Christoffer Green-Pedersen

34.1 Introduction

The CMP dataset has long been the central data source for the study of West European party politics. The discussions around the CMP dataset have been extensive and have covered many aspects of the dataset (cf. Gemenis, 2013). This includes the use of party manifestos as a data source (Helbling and Tresch, 2011), the use of quasi-sentences as a coding unit (Däubler et al., 2012), and, not least, how to estimate party positions from the data (Lowe et al., 2011; Budge and Meyer, 2013; Dinas and Gemenis, 2010). What has received much less attention are the consequences of the coding scheme on which the CMP dataset is based (though see Horn et al., 2017; Zulianello, 2014). However, the coding scheme applied is one of the most fundamental aspects of any dataset. Once the data has been coded in a certain way, this structures the use of the dataset profoundly. No matter how the data is transformed or re-scaled afterwards, one cannot make up for a category not found in a coding scheme, just as one cannot compensate for a survey question that was not asked.

The challenge with regard to the CMP coding scheme, however, is that it has been difficult to explore its consequences without having to recode the data. However, the growth of the CAP datasets offers new possibilities because, for a number of countries, the CAP datasets also include coding of party manifestos (e.g., Froio et al., 2017; Brouard et al., 2012). This opens up the possibility of comparing the coding of identical documents using different coding schemes. This chapter thus compares attention to a number of different policy issues identified in the CMP and in the CAP coding systems based on the coding of party manifestos for seven countries (Germany, France, the United Kingdom, the Netherlands, Belgium, Sweden, and Denmark) from Christoffer Green-Pedersen, *Issue Attention in West European Party Politics: CAP and CMP Coding Compared*. In: *Comparative Policy Agendas: Theory, Tools, Data*. Edited by Frank R. Baumgartner, Christian Breunig and Emiliano Grossman, Oxford University Press (2019). © Oxford University Press. DOI: 10.1093/oso/9780198835332.003.0034

1980 onwards. Put together, the manifestos in these seven countries provide a strong basis for comparing the patterns across the two coding systems.

It is important to remember that the CMP dataset was not developed with the purpose of measuring party attention to specific policy issues. The purpose of the dataset was primarily to measure political parties' ideological positions on a left–right scale. Nevertheless, the CMP dataset has increasingly been used to measure party attention to specific policy issues (see Section 34.2). Further, the creation of the left–right (rile) scale is also based on a saliency perspective (Budge and Meyer, 2013). A better understanding of the policy content of the CMP categories is therefore also important for evaluating the left–right scale.

The general finding of this chapter is that even though the CMP scheme was not originally developed to measure attention to policy issues, it is feasible to use it for that purpose; but with caution. The level of attention to various policy issues identified based on the CMP and CAP coding schemes is relatively similar, but with a number of important exceptions. A similar pattern emerges when looking at the correlations between the two measures on similar issues. The explanation for these exceptions can be found in the differences in the coding systems that again reflect their different theoretical purposes: The CAP coding scheme measures attention to policy issues whereas the CMP scheme was developed to measure ideological policy positions. This has two important implications. One is that the CMP coding focuses on the ideological goals of certain policy measures, whereas the CAP coding focuses on the policy instruments. Thus, if a party declares that it wants to improve primary schools to reduce inequality, the CMP coding would focus on the goal of reducing inequality whereas the CAP coding would focus on the attention to primary schools. The other implication is that the CMP dataset does not aim to cover the entire policy agenda. For instance, there is no specific category for healthcare; it is included in the categories relating to the welfare state. From the policy perspective of the CAP coding, this is an important drawback, but in terms of measuring ideological positions, this is not a problem as the ideological goals of different policies relating to the welfare state are often very similar. From a broader perspective, this chapter thus points to the importance of the theoretical perspectives behind different coding schemes. Classification and categorization need a theoretical foundation that must not be forgotten when datasets are later used for different purposes.

34.2 The CMP and CAP Datasets and the Discussions around them

The CMP dataset goes back to the establishment of the Manifesto Group in 1979. The debate around the dataset took off after its publication in 2001 (Budge et al., 2001). The dataset became the major source for studying party

politics in the Western world and, at the same time, a vivid debate around many aspects of the dataset has emerged (cf. Gemenis, 2013; Volkens et al., 2013). This relates to the use of party manifestos as a data source (Helbling and Tresch, 2011), the use of quasi-sentences as a coding unit (Däubler et al., 2012), the reliability of the CMP coding (Mikhaylov et al., 2012), and, not least, how to estimate party positions from the data (Lowe et al., 2011; Budge and Meyer, 2013; Dinas and Gemenis, 2010). What has received more limited attention is the CMP coding scheme (though see Horn et al., 2017; Zulianello, 2014). The scheme consists of fifty-six categories organized into seven broader domains.¹ To discuss this coding scheme, awareness of the theoretical foundation is necessary.

As stated by Volkens (2001: 96–8), the CMP coding scheme is a further development of the coding scheme that Robertson (1976: 74–5) developed to capture British party positions. The driving logic is not policy issues, but ideological ‘symbols’ like democracy, freedom, social justice, etc. The underlying idea is that parties’ mentions of such symbols are a way to measure their ideology, or more precisely, their policy preferences. Pro and con categories on some policy items were added later as a supplement to the pure saliency approach that Robertson (1976) had originally developed. Thus, the CMP dataset has—deliberately—(cf. Robertson, 1976: 72–3) not been based on the idea of policy issues, but is an attempt to capture the broader ideological preferences of political parties. The main focus of the coding is the ideological goals that parties express, not the policy means or instruments they suggest.

The CAP system was originally developed by Baumgartner and Jones (see Chapters 1 and 2) based on policy agenda-setting theory (see Baumgartner et al., 2002). It is an explicit attempt to cover the entire policy agenda and to provide a coding system that can travel across time and countries. The primary focus is on the policy content or policy instruments that receive attention. The coding scheme has been used to code very different political activities, such as executive speeches, media news, parliamentary activities (hearings, bills, laws, parliamentary questions, etc.), and local council agendas, as well as party platforms or manifestos. The CAP coding scheme is further based on the differences between main topics (or issues) such as the economy, transportation, and education, and subtopics within each major topic, such as inflation, monetary policy, and unemployment as subtopics within the economy, and railways, air transportation, or sea transportation within transportation. The latest version of the coding system has twenty-one major topics and 213 subtopics. As each unit is coded into a subtopic, these can be aggregated in whatever way suits a particular research purpose.

The CMP and CAP dataset thus start from quite different theoretical concepts—ideological goals or symbols vs. policy instruments or policy issues. Still, sometimes ideological symbols and policy issues are difficult to distinguish.

Education, culture, and European integration feature in the CMP dataset, which in reality thus also includes a number of policy issues. At the same time, the CMP dataset is based on a saliency approach, where the salience of the different items in the coding is used to calculate party positions. Thus, despite the focus on party positions, the CMP coding system is similar to the CAP coding system in the sense that the data measures the saliency of different ideological symbols, which are often in fact policy instruments or issues.

One difference between the coding schemes lies in the fact that the CMP dataset also tries to capture positions by coding both positive and negative references to a number of issues (e.g., European integration and education). However, there are also a number of items, such as the environment, where the CMP does not distinguish between positive and negative mentions. Instead, positions on the environment are constructed by subtracting other categories from the categories measuring attention to the environment (per416 and per501). Lowe et al. (2011) and Abou-Chadi (2016) both subtract productivity (per410), whereas Meguid (2008: 89) subtracts positive mentions of free enterprise (per401), support for agriculture (per703), and negative mentions of internationalism (per109). Further, the left–right (rile) scale developed from the CMP data—probably the most widely used measure derived from the data—is based precisely on subtracting attention to different items (e.g., economic incentives (per402)—economic planning (per404)). Thus, despite the differences in the coding systems, they in fact both attempt to measure the saliency of different policy instruments or issues.

Therefore, it is not surprising that the CMP coding system has increasingly been used to study attention to different policy issues. This is the case, for instance, with regard to the welfare state (Jensen and Seeberg, 2015), education (Jakobi, 2011; Busemeyer et al., 2013), the environment (Abou-Chadi, 2016; Spoon et al., 2014), corruption (Curini and Martelli, 2015), decentralization (Toubeau and Wagner, 2016), and immigration (Abou-Chadi, 2016; Green-Pedersen and Krogstrup, 2008; Breunig and Luedtke, 2008).

In sum, both with regard to the ‘core use’ of the CMP dataset for deriving a left–right scale and the broader use of the dataset to measure attention to particular policy issues, a detailed understanding of the consequences of the CMP coding scheme is important. However, there has hitherto been little debate about the CMP coding scheme and how it deals with different policy issues. Horn et al. (2017) is a rare attempt to dig into the coding system. There are three aspects into which one can dig even more deeply.

The first one has to do with the delimitation of the policy issues, i.e., what are the definitional boundaries between issues. This is of course a central question for any coding scheme, but it becomes even more pressing for a coding system that is only partly based on policy issues. With education, for

instance, one issue of delimitation is what distinguishes education from labor market policy, i.e., training, active labor market policy, etc. However, the CMP does not have a category for labor market policy. It has categories per701 and per702 for positive and negative mentions of labor groups, but not labor market policy. Further, technical training is supposed to be coded with 411.² As a coder, one looks for the best available category for coding a given sentence,³ so the fact that no direct category for labor market policy exists may affect the number of sentences in the education category. The lack of a labor market policy category may make education the most appropriate category, which might not be the case if a category for labor market policy existed. Thus the fact that only some policy issues have a specific code in the CMP scheme, i.e., the exhaustiveness of the scheme, might also affect the coding of the policy issues that do have their own category.

The second issue relates to the interpretation of some of the categories as policy issues. Thus, the categories per601 and per602 (national way of life positive and negative) and per607 and per608 (multiculturalism positive and negative) have frequently been used to study the issue of immigration (e.g., Abou-Chadi, 2016; Green-Pedersen and Krogstrup, 2008; Alonso and Fonseca, 2012; Breunig and Luedtke, 2008). The concern about this use of the categories is that they appear to have broader coverage than just immigration.⁴ Thus, the categories may sometimes include attention to issues other than immigration. This raises questions about their validity as measures of attention to immigration. Further, additional categories like 705 (underprivileged minority groups) (Green-Pedersen and Krogstrup, 2008) or even law and order (Alonso and Fonseca, 2012) are sometimes also added.

The third issue relates to the use of the CMP dataset to provide a general overview of the development of attention to policy issues. Given that only some policy issues have distinct categories, such overviews will necessarily have to focus on very broad categories such as material vs. non-material or left/right-oriented attention (e.g., Ward et al., 2015; Tavits and Potter, 2015; Albright, 2010; Green-Pedersen, 2007). The coding system also has two broad categories for positive and negative mentions of the welfare state (per504 and per505), which include diverse policy issues such as healthcare, pensions, and social housing. The CMP coding scheme thus makes difficult a general investigation of which issues rise and decline over time. Important issue developments may be hidden in the broad categories.

If one instead looks at the CAP coding system, its strengths are its inbuilt policy logic and its detailed coverage of the entire policy agenda. However, it is important to stress that the CAP scheme should not be considered a 'gold standard' to which the CMP dataset should be compared. As stated above, the main purpose of the CMP coding scheme has never been to measure attention to policy issues. This was a way of using the data that emerged after the design

of the coding scheme. The CAP coding scheme also comes with limitations. Not every relevant policy topic has its own subtopic code. For instance, the CAP coding scheme does not offer a specific subtopic for climate change, but the subtopic 705 covers air pollution, global warming, and noise pollution (Carter et al., 2017). Further, from a party perspective, the major limitation of the CAP coding system is clearly that it does not provide any measure of direction that makes it possible to measure party positions directly. Still, the theoretical starting purpose of the CAP coding scheme has never been to study particular actors such as political parties, but to track attention to policy issues.⁵

To summarize, the CMP and CAP coding systems have quite different origins. The CMP coding system was set up to capture party ideology by coding policy goals. The CAP coding system was designed to measure attention to policy issues by focusing on policy instruments. However, in practice both systems are used to study party attention to different policy issues, though the CMP dataset is by far the most used scheme because the data has been available for a large number of countries for a longer period of time.

34.3 CMP and CAP Compared

Since the CMP dataset has clearly been the most widely used scheme for studying party attention, the following comparison of issue attention based on the two coding schemes is structured around the three questions raised concerning this use of the CMP dataset. Thus, the first part compares issue attention to education, crime and justice, European integration, and the environment, where categories in the CMP coding scheme exist that, at least judging by their names, appear comparable to the policy issues in the CAP scheme. The next part then looks at attention to immigration, an issue where the CMP dataset has no direct category, but where other categories are used. Finally, the analysis looks at attention to healthcare to discuss the implications of the CMP coding scheme not having a category for this policy issue. Taken together, looking at these different questions allows us to evaluate the delimitation and exhaustiveness of the CMP coding scheme.

Before comparing attention to the different policy issues across the two coding systems, a few other questions concerning the comparability of the data are necessary. The comparison below is based on seven countries from 1980 onwards where CAP coding of party manifestos exist. These are Denmark, Sweden, the United Kingdom, France, Germany, the Netherlands, and Belgium. Together, they offer an extensive empirical basis for comparing the issue attention of political parties. The CMP dataset is based on quasi-sentences. This is also the case for most CAP manifesto datasets

(the United Kingdom, Sweden, Belgium, and France), but Denmark and Germany use natural sentences as coding unit (dot to dot). As shown by Däubler et al. (2012), this makes only a limited practical difference. The Dutch data is coded based on paragraphs, which must be taken into consideration when comparing with the CMP data.⁶

In terms of the parties covered, there are some differences across the two datasets. The most important differences are that the British CAP data only includes the three major parties (the Labour party, the Conservatives and the Liberal Democrats), unlike the CMP dataset, which also includes other parties represented in parliament. On the other hand, the coverage of minor Dutch parties is more extensive in the CAP dataset. Finally, concerning the French right, common programs of the parties are used for the CAP coding when they exist, and the same is the case for joint programs of the Swedish right (Alliancen) and left (Left-Alliance) in recent elections. In the analysis, only parties included in both datasets are compared.⁷

The purpose of the analysis in the following is to provide an overview of similarities and differences in both the level of attention to different policy issues identified by the two different coding schemes, and the correlation between the two measures. Further, potential explanations for differences will be discussed by looking at the coding schemes and including some examples of specific party manifestos where large differences were identified.⁸

Starting with education, the results presented in Table 34.1 indicate a significant difference across the two coding systems.⁹ In all countries, the CAP coding scheme identified higher average levels of attention to education than did the CMP dataset. The difference is both statistically and substantially significant for five out of seven countries. The following discussion examines why this is the case.

Table 34.1. Comparison of CAP and CMP party attention to education (paired t-test and correlations)

Country	CAP mean	CMP mean	Difference	SE	Correlation (Pearson r)	N (no. of manifestos/countries)
Belgium	5.0	4.7	0.3	0.32	0.54	85
Denmark	7.7	5.4	2.3***	0.46	0.77	94
France	5.5	4.8	0.7	0.66	0.41	34
Germany	5.5	3.4	2.1***	0.26	0.88	46
Netherlands	7.1	5.0	2.1***	0.29	0.52	84
UK	9.2	5.6	3.6***	0.55	0.68	21
Sweden	6.2	4.5	1.7***	0.32	0.81	48
Average (across country)	6.6	4.8			0.66	7

Note: * p < 0.1, ** p < 0.05, *** p < 0.01.

Sources: CAP data, see Green-Pedersen (2018); CMP data from <https://manifesto-project.wzb.eu/May 2016>.

The fact that the pattern is quite consistent across all countries would indicate that the explanation should be found in the codebook itself and not particular interpretations of it in a given country.¹⁰ In the CAP coding scheme, the major topic of education covers all policy questions related to primary education, secondary education, universities (including students), tertiary education, and vocational education. In the CMP system, education is covered by the categories per506 and per507, referring to education expansion and limitations. However, the CMP coding scheme also contains a number of other categories that may be potentially relevant with regard to education. For instance, per706 covers non-economic demographic groups like university students. The CMP coding scheme also contains categories like per503, which relates to equality and the removal of barriers for underprivileged groups. Statements about education are likely to be framed in exactly such a way and may therefore be coded here (see also Horn et al., 2017).

Thus, whereas the CAP coding scheme basically captures any policy question related to education, the CMP coding system approaches it from an ideological perspective where not all statements related to educational policy may relate to its expansion or limitation, but relate to questions about equality or certain social groups like students. Thus, the CMP categories for education would seem to only partially cover education as seen from a policy perspective. Table 34.1 also reports the correlation between the two codings. This is mostly relatively high, but with some cross-national variation. Thus, the two datasets identify relatively similar dynamics in party attention, though less so in France, the Netherlands, and Belgium.

In the same way, Table 34.2 compares the CAP and CMP mean level of attention with regard to crime and justice or law and order. The table shows that attention to crime and justice is higher according to CAP, and statistically significantly so, in five of the seven countries. The differences are somewhat

Table 34.2. Comparison of CAP and CMP party attention to crime and justice (paired t-test and correlations)

Country	CAP mean	CMP mean	Difference	SE	Correlation (Pearson r)	N (no. of manifestos/countries)
Belgium	6.2	4.4	1.8***	0.48	0.53	85
Denmark	3.6	3.7	-0.1	0.35	0.77	94
France	4.7	3.8	0.9*	0.55	0.70	34
Germany	4.3	2.5	1.8***	0.30	0.68	46
Netherlands	6.7	6.2	0.5**	0.25	0.77	84
UK	6.8	5.7	1.1***	0.27	0.88	21
Sweden	3.0	3.0	0.0	0.25	0.85	48
Average (across countries)	5.0	4.2			0.74	7

Note: * p < 0.1, ** p < 0.05, *** p < 0.01.

Sources: CAP data, see Green-Pedersen (2018); CMP data from <https://manifesto-project.wzb.eu/May 2016>.

smaller than what was found with regard to education. Based on the coding scheme, one should also not expect large differences.¹¹

The CMP coding scheme has one category, per605, that covers questions about law and order such as enforcement of laws, support for the police, and importance of internal security. The CAP coding scheme also has a major topic called law and crime with subtopics for white-collar crime and organized crime (1202), criminal and civil code (1210), and police and other general domestic security responses to terrorism (1227). However, the CAP coding scheme also includes subtopics covering the broader judicial systems, i.e., court administration (1204), prisons (1205), and agencies dealing with law and crime (1201). Thus the most likely explanation for the higher levels of attention in the CAP coding is that such broader aspects of the judicial system are picked up by the CAP system. Further, the observations based on the two coding schemes correlate quite well, and with relatively limited cross-national variation.

Table 34.3 looks at European integration. The CMP has two categories, per108 and per110, that cover positive and negative mentions of European integration. The latter includes “opposition to specific European Union policies.” The CAP scheme covers European integration in two ways. First, it is captured partly through subtopic 1910, which covers “institutional” or polity-related questions regarding the European Union—enlargement, the role of Commissions, national referendums, etc.¹² Second, statements related to the policies of the European Union are covered under the relevant policy areas. Thus, environmental policy or banking policies will be coded under the relevant policy topics. For these reasons, one would expect the level of attention found in the CMP dataset to be higher.

This is indeed what is found in Table 34.3, where the higher values for the CMP measure is statistically significant for six out of the seven countries.¹³

Table 34.3. Comparison of CAP and CMP party attention to European integration (paired t-test and correlations)

Country	CAP mean	CMP mean	Difference	SE	Correlation (Pearson r)	N (no. of manifestos/countries)
Belgium	2.1	3.0	-0.9***	0.32	0.44	85
Denmark	3.3	3.2	0.1	0.46	0.84	94
France	3.0	4.2	-1.2**	0.61	0.38	34
Germany	2.1	2.9	-0.8**	0.32	0.31	46
Netherlands	2.5	3.0	-0.5***	0.20	0.65	84
UK	2.6	3.4	-0.8***	0.29	0.76	21
Sweden	2.6	3.5	-0.9***	0.27	0.84	48
Average (across countries)	2.6	3.3			0.60	7

Note: * p < 0.1, ** p < 0.05, *** p < 0.01.

Source: CAP data, see Green-Pedersen (2018); CMP data from <https://manifesto-project.wzb.eu/May 2016>.

However, in substantial terms, the differences are rather small, especially when compared to education, but also to crime and justice. Thus, the differences with regard to the coding system do not seem to make a large substantial difference. This is most likely because policies emerging from the European Union are not often discussed in manifestos. Rather, statements here are focused on the polity aspects of European integration. Policies emerging from the European Union are more likely taken up in parliamentary activities such as questions to the minister (Senninger, 2017). In terms of correlation, the coefficients exhibit more cross-national variation than for the two other issues and relatively low figures for Germany, France, and Belgium. The average correlation is also lower than for education and especially crime and justice.

Attention to the environment is captured in the CMP coding scheme by the categories per416 (anti-growth economy) and per501 (environmental protection). The CAP coding system covers the environment through the major topic of environmental policy with sub-categories for a large number of aspects of environmental protection, e.g., drinking water and water pollution (701), waste disposal (702), and air pollution, global warming, and noise pollution (705). The CAP coding scheme also has a number of subtopics related to the environment such as international resources exploitation and resources agreements (1902) and natural resources, public lands, and forest management (2103), which are also included in the attention measure reported in Table 34.4.¹⁴

It is furthermore important to note that the two coding schemes differ in terms of the categories offered for related issues. The CAP coding scheme has a major topic for energy policy, with subtopics for nuclear power (801), coal (805), and alternative and renewable energy (806). The CMP coding scheme does not have a specific category for energy policy-related quasi-sentences. Both coding schemes have topic codes related to agriculture. The CAP coding scheme also has a major topic for agriculture-related policy, which, among others, includes a subtopic for animal and crop disease, animal welfare, and pest control (405).¹⁵ The CMP coding scheme has a category (per703) for support for farmers and agriculture, which is thus much narrower than the CAP major issue code for agriculture.

Table 34.4 shows party attention to the environment. The general tendency is for the CMP coding system to generate higher values. The difference between the CMP and the CAP measure is thus statistically significant for five of the seven countries.¹⁶ Substantially, this is most clearly the case for Sweden and the Netherlands, and to a more limited extent for Germany and France. The most likely explanation for this difference is the difference in related categories. This can be seen by looking at the Dutch Party for Animal Rights (PvdD). According to the CMP data, this party mainly pays attention to

Table 34.4. Comparison of CAP and CMP party attention to environment (paired t-test and correlations)

Country	CAP mean	CMP mean	Difference	SE	Correlation (Pearson r)	N (no. of manifestos/countries)
Belgium	5.7	6.2	-0.5	0.60	0.40	85
Denmark	6.3	6.7	-0.4	0.51	0.67	94
France	4.3	5.7	-1.4**	0.62	0.83	34
Germany	6.4	8.6	-2.2***	0.55	0.72	46
Netherlands	7.2	11.2	-4.0***	1.00	0.78	84
UK	4.8	5.2	-0.4**	0.26	0.92	21
Sweden	7.9	13.0	-5.1***	0.94	0.89	48
Average (across countries)	6.1	8.1			0.74	7

Note: * p < 0.1, ** p < 0.05, *** p < 0.01.

Source: CAP data, see Green-Pedersen (2018); CMP data from <https://manifesto-project.wzb.eu/May 2016>.

the environment: 73.8 percent of its manifesto was devoted to this issue in 2006, 74.7 percent in 2010, and 61.6 percent in 2012. According to the CAP data, the comparable figures are 24.1, 19.8, and 19.7. However, whereas the CMP coding scheme records less than 1 percent of the PvdD manifestos under attention to agriculture, the CAP coding scheme reports 38.8 percent, 38.8 percent, and 28.8 percent. Thus, a likely explanation is that various demands for protection of animal rights and regulation of agriculture made by the party are coded as related to agriculture by the CAP coding and coded as related to the environment by the CMP coding scheme, simply because no broad category for agriculture exists in the CMP coding scheme.

Another example would be the manifestos of the Swedish Green Party in 1994: Where the CMP coding scheme records 34.8 percent attention to the environment, the CAP scheme only records 15.2 percent. The CAP coding scheme then reports 5.1 percent for energy policy and 4 percent for agriculture, whereas the CMP scheme reports nothing for agriculture. Thus, it seems that the environmental categories in the CMP scheme capture statements that in the CAP scheme would be coded under energy—for instance, reduction of the use of coal to reduce CO2 emissions. It is also worth noting that countries such as Sweden and Germany that have seen intensive political debates around nuclear power (Müller and Thurner, 2017) exhibit substantially higher values for the CMP coding, which includes nuclear power. Thus, the findings on the environment clearly underline that the use of one category within either of the two coding schemes cannot be viewed in isolation from the rest of the coding scheme. In terms of correlation, these are, with the exception of Belgium, high in comparison with the other issues. Only crime and justice has the same high cross-national average correlation. This indicates that quite similar party competition dynamics are captured by the two coding schemes.

As discussed above, the CMP coding scheme has also been used relatively widely to study attention to immigration, even though no category referring directly to this in the CMP coding scheme exists. Existing studies thus use the categories referring to positive and negative mentions of national way of life (per601 and per602) and the categories referring to positive and negative mentions of multiculturalism (per607 and per608). These categories are often combined with other categories, typically per705, favorable mentions of underprivileged minority groups like immigrants, homosexuals, and the disabled (Meguid, 2008: 90; Alonso and Fonseca, 2012; Green-Pedersen and Krogstrup, 2008). The CAP coding scheme has one category related to general immigration and refugee issues. In terms of neighboring issues, both coding schemes have a category for crime and justice, as discussed earlier.

Table 34.5 shows the results from a comparison of two different measures of immigration from the CMP data. CMP1 contains per601, per602, per607 and per608, which seem mostly closely related to immigration. CMP2 then added per705. In all cases except France, CMP1 identifies less average attention to immigration than the CAP measure,¹⁷ but the differences are smaller than those found on issues such as education and the environment and not significant for France, the Netherlands and the United Kingdom. This finding would speak in favor of including more categories to capture attention when using the CMP scheme, as is typically done. However, the question then is whether this also measures attention that is not related to immigration.

Table 34.5 would indicate that this could indeed be the case for some countries. For all countries other than Sweden and Denmark, the CMP2 measure produces values that are higher, and statistically significantly so, than the CAP measure.¹⁸ The results would thus indicate that by including per705, one also captures attention to underprivileged minority groups other

Table 34.5. Comparison of CAP and CMP party attention to immigration (paired t-test and correlations)

Country	CAP mean	CMP1 mean	Difference	SE	Correlation (Pearson r) (CAP/CMP1)	CMP2 mean	Correlation (Pearson r) (CAP/CMP2)	N (no. of manifestos/countries)
Belgium	3.0	2.2	0.8***	0.34	0.30	4.1***	0.32	85
Denmark	6.0	5.1	0.9*	0.62	0.74	6.2	0.79	94
France	2.7	3.3	-0.6	0.62	0.54	5.5***	0.44	34
Germany	3.1	1.7	1.4***	0.27	0.47	3.7***	0.71	46
Netherlands	3.6	3.3	0.3	0.48	0.37	6.0***	0.36	84
UK	1.9	1.3	0.6	0.58	-0.1	2.7*	0.12	21
Sweden	2.0	1.0	1.0**	0.41	0.38	1.8	0.38	48
Average (per country)	3.2	2.6			0.39	4.3	0.45	7

Note: * p < 0.1, ** p < 0.05, *** p < 0.01.

Sources: CAP data, see Green-Pedersen (2018); CMP data from <https://manifesto-project.wzb.eu/May2016>.

than immigrants, i.e., the disabled or homosexuals. Compared to the other issues, the correlations between the CAP and CMP measures are generally much weaker than for the other issues, especially for the United Kingdom. The correlation is mostly better when including the per705 category, though only for Germany and the United Kingdom is it significantly improved and for France it is weaker.

To summarize, it seems that attention to immigration is captured by a number of different categories in the CMP dataset. Including the four categories related to national way of life and multiculturalism provide estimations relatively close to the CAP measure in terms of levels of attention, though typically lower. This makes the inclusion of further categories appealing, but Table 34.5 would indicate that including per705 implies capturing attention that is not immigration related. No matter which solution is applied, the correlation is clearly weaker than for the other issues. Thus, the immigration measures that can be developed from the CMP dataset seem the most different from the CAP ones. This is not surprising, given that they were also not designed with only immigration-related content in mind.

The third and final question raised above with regard to the CMP coding scheme was the implication of certain policy issues being completely absent from the scheme, in the sense that they have no specific category. Energy policy was mentioned above, and another example is healthcare. In the CMP coding scheme, attention to healthcare is handled by per504 and per505, which capture statements about welfare-state expansion and welfare-state

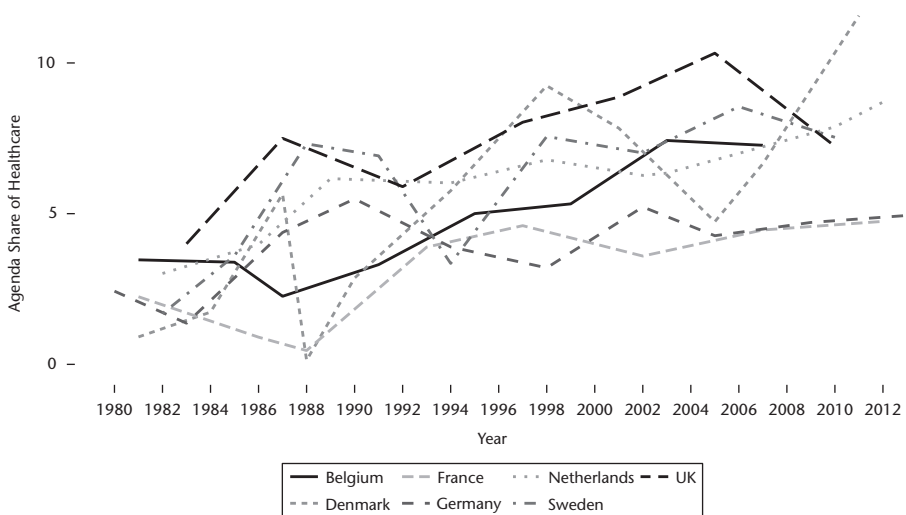


Figure 34.1. Attention to healthcare in seven countries, 1980–2013

Source: Comparative Agendas Project

limitation. The CAP coding scheme has a major topic for healthcare and a series of subtopics capturing different aspects of healthcare such as facilities construction, regulation, and payments (322) and disease prevention, treatment, and health promotion (331). Figure 34.1 shows the development of attention to healthcare in the seven countries based on the CAP data. In all countries, attention to healthcare has risen, and in several countries quite substantially. This significant development cannot be observed through the CMP dataset.

34.4 Conclusions

The comparison of the level of attention to different policy issues in the CMP and the CAP coding schemes show relatively similar levels of attention to many issues. Furthermore, the differences were typically in the same direction, indicating that they were due to one coding system systematically leading to higher or lower levels of attention in all countries. Relatively high levels of correlation between the CMP and the CAP measures were also generally found, but with significant exceptions. The biggest differences in terms of levels of attention were found with regard to the environment, whereas the weakest correlations were with regard to immigration, no matter which measure of immigration from the CMP dataset was used. Generally, crime and justice was the issue where the two measures were closest to each other, both in terms of levels of attention and correlation, and immigration was the issue where they were furthest from each other. This is not surprising when looking at the codebooks. The categories for crime and justice are quite close to each other in the two coding schemes, whereas they differ substantially with regard to immigration. Here, the issue is only measured indirectly in the CMP coding scheme.

These differences partly relate to the different theoretical perspectives behind the coding schemes. The ambition of CMP coding with regard to education, for instance, is not to capture all attention to education policy, but rather to measure the ideological preferences of the parties. Therefore, statements related to educational policy are also likely to be coded under, for instance, equality. The differences identified here partly relate to the different categories of the two coding schemes, where not all policy issues are covered by the CMP dataset. Thus, the level of attention paid to the environment was typically higher in the CMP coding because statements about energy or agriculture are likely to be coded here, whereas they will be coded under energy policy and agriculture in the CAP scheme. This also stresses the importance of evaluating the coding schemes in total. What is coded under one category depends on the alternative categories offered.

It is also worth noting that the differences in coding units do not seem to make a substantial difference. The CMP data use quasi-sentences as coding units, which is also the coding unit used in the CAP coding of party manifestos in the United Kingdom, Sweden, France, and Belgium. However, for Germany and Denmark, natural sentences were used as coding units for the CAP coding, and in the Dutch case, paragraphs were used. The tables shown above do not indicate that the relationships between the CAP and CMP were any weaker for the three countries with a different coding unit for the CAP data—not even for the Netherlands, where paragraphs were used. This follows the findings of Däubler et al. (2012), who also conclude that the use of natural sentences instead of quasi-sentences provides very similar results in substantial terms.

To conclude, in terms of the use of the CMP dataset to study policy issues, the findings here suggest that this should be done, but with caution. The CMP coding scheme is not based on a policy logic, and in some (but not all) cases, this makes a substantial difference.

Acknowledgements

Thanks to Christian Breunig and Nicolas Merz for very constructive criticism of an earlier version of this chapter.

Notes

1. The codebook can be found here: https://manifestoproject.wzb.eu/download/data/2016b/codebooks/codebook_MPDataset_MPDS2016b.pdf. A newer version of the codebook exists with more subcategories, but these codes have only been used for the most recent documents.
2. The detailed coding instruction of the CMP dataset can be found at https://manifestoproject.wzb.eu/download/papers/handbook_2014_version_5.pdf
3. The CMP project uses quasi-sentences as the coding unit.
4. This worry is supported by the fact that the recent Version 5.0 of the codebook has divided the four categories into two each, one general and one immigration related.
5. Potentially, one could also develop positional measures from the CAP scheme by subtracting the saliency of from the different policy issues.
6. Both coding schemes only allow one code for each coding unit.
7. The CAP coding scheme has a Master Codebook that can be found at <http://sbevan.com/cap-master-codebook.html> together with the crosswalk, which ties national versions of the codebook to the Master Codebook. The description of the CAP manifesto dataset (Green-Pedersen 2018) contains an overview of how the different subtopics from the CAP coding scheme have been combined into the policy issues.

8. What has not been feasible is a systematic sentence-by-sentence comparison of two different codings of the same manifesto. Even for the countries where quasi-sentences were also used as the coding unit for the CAP coding, the coding into quasi-sentences was done separately for the two datasets, implying differences in the exact quasi-sentences identified. Therefore, a direct comparison of the coding at the sentence level would first require an inspection of the data to ensure that only the coding of identical quasi-sentences is compared.
9. In the CMP dataset, downloaded from <https://manifesto-project.wzb.eu/May2016>, un-codeable sentences are reported as a category alongside the substantial categories. In the following, the percentages in the CMP dataset have been recoded as percentages of total coded sentences, i.e., excluding the un-codeable sentences. This is the typical approach of the CAP scheme, where a total agenda summing to 100 is calculated based only on coded units.
10. For individual manifestos, 69 percent of the cases have higher values based on the CAP scheme.
11. For individual manifestos, 62 percent of the cases have higher values based on the CAP scheme.
12. In the CAP Master Codebook, 1910 refers to Western Europe and European integration, but for most countries, this sub-code has been split into 1910 for European integration and 1913 for Western Europe.
13. For individual manifestos, 71 percent of the cases have higher values based on the CMP scheme.
14. The CAP categories added to the environment are 1902, 2100, 2101, 2013, and 2199 and, for some countries, 407 (regulation of agriculture).
15. In the CAP coding scheme, regulation of agriculture, including for environmental purposes, will be coded under agriculture. However, some of the countries have generated a specific subtopic, 407, for environmental regulation of agriculture, and attention in this subtopic has been included in environmental attention in this chapter.
16. For individual manifestos, 65 percent of the cases have higher values based on the CMP scheme.
17. For individual manifestos, 59 percent of the cases have higher values based on the CAP scheme.
18. For individual manifestos, 69 percent of the cases have higher values based on the CMP scheme.

References

- Abou-Chadi, T. (2016). Niche Party Success and Mainstream Party Policy Shifts. *British Journal of Political Science*, 46(2): 417–36.
- Albright, J. (2010). The Multidimensional Nature of Party Competition. *Party Politics*, 16(6): 699–719.

- Alonso, S., and da Fonseca, S. C. (2012). Immigration Left and Right. *Party Politics*, 18(6): 865–86.
- Baumgartner, F. R., Jones, B. D., and Wilkerson, J. D. (2002). Studying Policy Dynamics. In *Policy Dynamics*, ed. F. R. Baumgartner and B. D. Jones. Chicago: University of Chicago Press, 29–46.
- Breunig, C., and Luedtke, A. (2008). What Motivates the Gatekeepers? Explaining Governing Party Preferences on Immigration. *Governance*, 21(1): 123–46.
- Brouard, S., Grossman, E., and Guinaudeau, I. (2012). “La compétition partisane française au prisme des priorités électorales.” *Revue française de science politique*, 62(2): 255–76.
- Budge, I., Klingeman, Hans-D., Volkens, A. et al. (2001). *Mapping Policy Preferences: Estimates for Parties, Electors, and Governments 1945–1998*. Oxford: Oxford University Press.
- Budge, I., and Meyer, T. (2013). Understanding and Validating the Left–Right Scale (RILE). In *Mapping Policy Preferences from Texts III*, ed. A. Volkens, J. Bara, I. Budge et al. Oxford: Oxford University Press, 85–106.
- Busemeyer, M. R., Simon, T. F., and Garritzmann, J. L. (2013). Who Owns Education? Cleavage Structures in the Partisan Competition over Educational Expansion. *West European Politics*, 36(3): 521–46.
- Carter, N., Ladrech, R., Little, C., and Tsagkroni, V. (2017). Political Parties and Climate Policy: A New Approach to Measuring Parties’ Climate Policy Preferences. *Party Politics*. <https://doi.org/10.1177/1354068817697630>
- Curini, L., and Martelli, P. (2015). A Case of Valence Competition in Elections: Parties’ Emphasis on Corruption in Electoral Manifestos. *Party Politics*, 21(5): 686–98.
- Däubler, T., Benoit, K., Mikhaylov, S., and Laver, M. (2012). Natural Sentences as Valid Units for Coded Political Texts. *British Journal of Political Science*, 42(4): 937–51.
- Dinas, E., and Gemenis, K. (2010). Measuring Parties Ideological Positions with Manifesto Data. *Party Politics*, 16(4): 427–50.
- Froio, C., Bevan, S., and Jennings, W. (2017). Party Mandates and the Politics of Attention: Party Platforms, Public Priorities and the Policy Agenda in Britain. *Party Politics*, 23(6): 692–703.
- Gemenis, K. (2013). What to Do and Not to Do with the Comparative Manifesto Dataset. *Political Studies*, 61(1): 3–23.
- Green-Pedersen, C. (2007). The Growing Importance of Issue Competition: The Changing Nature of Party Competition in Western Europe. *Political Studies*, 55(4): 608–28.
- Green-Pedersen, C. (2018). The CAP Party Manifesto Dataset. Department of Political Science, Aarhus University available at agendasetting.dk.
- Green-Pedersen, C., and Krogstrup, J. (2008). Immigration as a Political Issue in Denmark and Sweden: How Party Competition Shapes Political Agendas. *European Journal of Political Research*, 47 (5): 610–34.
- Helbling, M., and Tresch, A. (2011). Measuring Party Positions and Issue Salience from Media Coverage: Discussing and Cross-validating New Indicators. *Electoral Studies*, 30(1): 174–83.
- Horn, A., Kevins, A., Jensen, C., and van Kersbergen, K. (2017). Peeping at the Corpus: What Is Really Going on behind the Equality and Welfare Items of the Manifesto Project? *Journal of European Social Policy*, 27(5): 1–15.

- Jakobi, A. P. (2011). Political Parties and the Institutionalization of Education. *Comparative Education Review*, 55(2): 189–209.
- Jensen, C., and Seeberg, H. (2015). The Power of Talk and the Welfare State. *Socio-Economic Review*, 13 (2): 215–33.
- Lowe, W., Benoit, K., Mikhaylov, S., and Laver, M. (2011). Scaling Policy Preferences from Coded Political Texts. *Legislative Studies Quarterly*, 36(1): 123–54.
- Meguid, B. (2008). *Party Competition between Unequals: Strategies and Electoral Fortunes in Western Europe*. Cambridge: Cambridge University Press.
- Mikhaylov, S., Laver, M., and Benoit, K. R. (2012). Coder Reliability and Misclassification in the Human Coding of Party Manifestos. *Political Analysis*, 20(1): 78–91.
- Müller, W. C., and Thurner, P. W. (eds.). (2017). *The Politics of Nuclear Energy in Western Europe*. Oxford: Oxford University Press.
- Robertson, David. 1976. *A Theory of Party Competition*. London: Wiley.
- Senninger, R. (2017). Issue Expansion and Selective Scrutiny—How Opposition Parties Used Parliamentary Questions about the European Union in the National Arena from 1973 to 2013. *European Union Politics*, 18(2): 283–306.
- Spoon, J. J., Hobolt, S. B., and de Vries, C. (2014). Going Green: Explaining Issue Competition on the Environment. *European Journal of Political Research*, 53(2): 363–80.
- Tavits, M., and Potter, J. D. (2015). The Effect of Inequality and Identity on Party Strategies. *American Journal of Political Science*, 59(3): 744–58.
- Toubeau, S., and Wagner, M. (2016). Party Competition over Decentralisation: The Influence of Ideology and Electoral Incentives on Issue Emphasis. *European Journal of Political Research*, 55(2): 340–57.
- Volkens, A. (2001). Quantifying the Election Programmes: Coding Procedures and Controls. In *Mapping Policy Preferences: Estimates for Parties, Electors, and Governments 1945–1998*, ed. I. Budge, Hans-D. Klingeman, A. Volkens et al. Oxford: Oxford University Press, 93–109.
- Volkens, A., Bara, J., Budge, I. et al. (eds) (2013). *Mapping Policy Preferences from Texts III*. Oxford: Oxford University Press.
- Ward, D., Kim, J. H., Tavits, M., and Graham, M. (2015). How Economic Integration Affects Party Competition. *Comparative Political Studies*, 48(10): 1227–59.
- Zulianello, M. (2014). Analyzing Party Competition through the Comparative Manifesto Data: Some Theoretical and Methodological Considerations. *Quality & Quantity*, 48(3): 1723–37.