The Benefits of Clustering Methods in a Comparison of Electoral Manifestos

Samo Kropivnik
University of Ljubljana
Faculty of Social Sciences
samo.kropivnik@fdv.uni-lj.si

Paper prepared for
RC33 8th International Conference on Social Science Methodology
University of Sydney, Sydney, Australia
July 9–13, 2012

The data collection and a part of the analyses were completed in the framework of the research project, "Pre-election campaign and democratic evolution of state and society," conducted at the Centre for Political Science Research at the Institute of Social Sciences, University of Ljubljana and financed by the Slovenian Research Agency (1/5/2009–30/4/2012)

Abstract

We address a substantive question of political science research: What are similarities and differences in the content of political party electoral programmes (manifestos)? In our analyses, we demonstrate the benefits of clustering methods in providing methodologically superior solutions of multilevel comparisons dilemmas.

Keywords: political party program, elections, policy positions, policy domains, MARPOR, comparative studies, country profiles, Euclidean distances, Galaxy, agglomerative hierarchical clustering, k-means clustering

Political Party Manifestos—the Data

Political party manifestos (electoral programmes, party programmes or electoral manifestos; we use all these terms as synonyms) play a significant role in electoral democracies because they recognize the importance of political issues, develop a party position on them, set the course of actions a party will take if elected, unite a party internally and, last but not least, advise party activists and supporters as well as inform the general electorate.[1] In party manifestos, political and policy ideas, positions and goals are recorded, publicised and documented for analysis.

In Europe, parliament party manifestos have been systematically collected since 1979. Their content has been coded, and the data are available for analysis (project MARPOR, previously MRP and CMP; http://www.mzes.uni-mannheim.de/projekte/manifestos/; 24. 3. 2012). Each parliamentary party programme is characterised according to its match with a standardised set of carefully selected, precisely defined and theoretically relevant policy positions. Subsequently, policy position codes are merged in seven mutually exclusive and theoretically exhaustive policy domains that are defined in Table 1.[2] Obviously, for each document, contextual data are also available on political party, party family, country and election year.

[1] Klingemann, Hans-Dieter, Richard I. Hoffebert in Ian Budge (1994, editors). Parties, Policies and Democracy. Boulder, Co.: Westview.

[2] For detailed description of data creation proces see Werner, Annika and Andrea Volkens (2009). Manifesto Coding Instructions (3. fully revised version). Manifesto Project (MRG/CMP/MARPOR). Berlin: Wissenschaftszentrum Berlin für Sozialforschung (WZB).

Table 1: Policy domains and policy positions (codes)

Domain 1: External Relations

101 Foreign Special Relationships: Positive
102 Foreign Special Relationships: Negative
103 Anti-Imperialism: Positive
104 Military: Positive
105 Military: Negative
106 Peace: Positive
107 Internationalism: Positive
108 European Integration: Positive
109 Internationalism: Negative
110 European Integration: Negative

Domain 2: Freedom and Democracy

201 Freedom and Human Rights: Positive
202 Democracy: Positive
203 Constitutionalism: Positive
204 Constitutionalism: Negative

Domain 3: Political System

301 Decentralisation: Positive
302 Centralisation: Positive
303 Governmental and Administrative Efficiency: Positive
304 Political Corruption: Negative
305 Political Authority: Positive

Domain 4: Economy

401 Free Enterprise: Positive
402 Incentives: Positive
403 Market Regulation: Positive
404 Economic Planning: Positive
405 Corporatism: Positive
406 Protectionism: Positive
407 Protectionism: Negative
408 Economic Goals
409 Keynesian Demand Management: Positive
410 Economic Growth
411 Technology and Infrastructure: Positive
412 Controlled Economy: Positive
413 Nationalisation: Positive
414 Economic Orthodoxy: Positive
415 Marxist Analysis: Positive
416 Anti-Growth Economy: Positive

Domain 5: Welfare and Quality of Life

501 Environmental Protection: Positive
502 Culture: Positive
503 Social Justice: Positive
504 Welfare State Expansion
505 Welfare State Limitation
506 Education Expansion
507 Education Limitation

Domain 6: Fabric of Society

601 National Way of Life: Positive
602 National Way of Life: Negative
603 Traditional Morality: Positive
604 Traditional Morality: Negative
605 Law and Order: Positive
606 Social Harmony: Positive
607 Multiculturalism: Positive
608 Multiculturalism: Negative

Domain 7: Social Groups

701 Labour Groups: Positive
702 Labour Groups: Negative
703 Agriculture: Positive
704 Middle Class and Professional Groups: Positive
705 Minority Groups: Positive
706 Non-Economic Demographic Groups: Positive

Policy domains shares are the most characteristic and most valuable feature of the dataset since when considered jointly, policy domains exhaustively cover existing national (country specific) policy issues and simultaneously offer universal comparison in time and space (e.g. among policy arenas, among countries). They have been created for this purpose and their validity has been repeatedly evaluated and confirmed in various comparative studies.[3]

[3] Studies are documented on MARPOR website (see above).

The Aim and the Structure of the Research

Drawing on the MARPOR dataset, comparisons can be carried out between national political parties competing for votes in a certain election year or even throughout longer periods with more election cycles involved.[4] On the other hand, countries (national policy arenas) can be compared on the basis of average policy domain shares, calculated for narrow or wide time intervals. In national policy arenas, political parties compete with each other, and thus they unavoidably react to each other. Therefore, their manifestos ideally reflect their own policy goals as well as their responses to policy initiatives of other major parties. Only when manifestos are combined in a national collection of policy positions do they have the potential to comprehensively describe a country's specifics in a domain structure. Comparisons among countries provide a broader perspective in electoral democracy studies, which makes them more interesting for political scientists.

What we claim and intend to demonstrate in the paper is that in order to describe competently and comprehensively the policy structure content of electoral campaigning in any country in comparison with others, one has to consider both the country level, i.e., average shares of domains in the analysed countries and the party programmes level, i.e., exact shares of domains in each party manifesto involved. The latter is required to estimate differences between programmes in a country in order to realize how the nationalization of party manifestos has progressed so far.[5]

In our study, we begin with comparisons between countries (aggregates of parties) and focus on Slovenia as a new democracy. Since Slovenia became an electoral democracy only in 1990, we limit the study to the period between

[4] Kustec Lipicer, Simona and Kropivnik, Samo (2011). Dimensions of party electoral programmes: Slovenian experience. Journal of Comparative Politics. Vol. 4, No. 1, 52-75.

[5] c.f. Lipset, Seymur Martin and Stein Rokkan (1967, editors). Party Systems and Voter Alignments. New York: Free Press.

1990 and 2003, the latter being the last election year for which complete MARPOR data are available at the time of writing this paper. We employ graphically presented country profiles (line graph) and hierarchical agglomerative clustering (dendrogram) to both give a general overview and to recognise similarities and differences between Slovenia and the other fifty countries that participated in the project during the period analysed. In order to take differences inside countries into account, we return in the second step to original (basic, not aggregated) units of analysis and cluster individual party manifestos (almost 1300 units in the selected period) into groups (ideal types) using hierarchical agglomerative clustering methods and K-means clustering method. In a contingency table the obtained groups are split according to the manifestos' countries of origin to estimate variation inside countries. Furthermore, we employ Euclidean distances to present relations between Slovenia and all other countries in a Galaxy in a graphical format developed for that purpose. The central (referential) country is represented as the Sun and all other countries are represented as different planets allocated around the Sun proportionally to Euclidean distances. Countries (i.e., planets) are depicted according to the policy domains structure of their manifestos, taking into account the type and the level of homogeneity, which are both recognized drawing on the results of clustering analyses of manifestos.

Comparisons between Slovenia and Other Countries

In the table below (Table 2) average policy domains shares in the period from 1990 to 2003 are calculated from the MARPOR database for each country. The shares are modified so that their sum is always 100% in order to enable comparisons between countries. This is achieved by excluding uncoded sentences, i.e., sentences without any policy-oriented content, such as general statements, introductory remarks etc. On average, there were 8% of uncoded sentences in a program. The last row of the table shows the average shares for all countries combined.

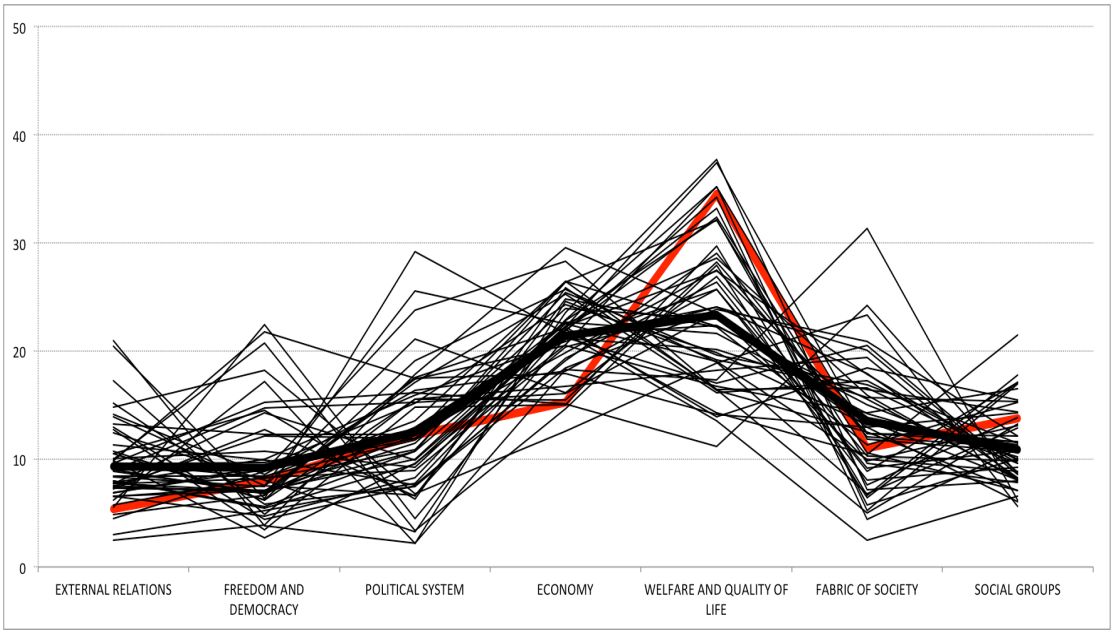
Table 2: Average policy domains shares in 51 analysed countries in the period from 1990 to 2003 (recalculated from the source)[6]

	EXTERNAL RELATIONS	FREEDOM AND DEMOCRACY	POLITICAL SYSTEM	ECONOMY	WELFARE AND QUALITY OF LIFE	FABRIC OF SOCIETY	SOCIAL GROUPS
Albania	7.70	14.24	9.90	22.00	14.30	10.38	21.48
Armenia	10.07	10.74	9.22	22.17	13.87	18.44	15.48
Australia	2.44	3.87	17.86	29.52	23.47	9.79	13.06
Austria	7.79	9.93	15.86	20.89	25.65	9.12	10.76
Azerbaijan	14.76	18.16	6.75	12.53	18.87	23.31	5.61
Belorussia	6.16	17.19	2.17	24.60	19.15	16.39	14.36
Belgium	8.34	7.43	21.09	16.14	26.87	10.28	9.84
Bosnia and Herzegovina	5.53	22.45	4.47	21.26	17.01	20.51	8.78
Bulgaria	7.70	8.19	10.84	25.40	19.57	16.84	11.45
Canada	5.71	8.32	13.55	26.46	22.28	15.54	8.15
Croatia	8.95	14.74	15.21	16.02	23.52	13.51	8.04
Cyprus	12.67	3.41	14.76	15.01	29.71	6.66	17.77
Czech Republic	9.53	9.70	12.36	21.45	22.38	14.98	9.60
Denmark	12.96	4.38	7.48	18.43	23.83	20.83	12.09
Estonia	7.34	6.91	10.69	22.48	28.61	15.92	8.06
Finland	6.54	4.67	8.29	22.84	35.17	11.80	10.69
France	11.31	9.87	7.59	19.06	29.01	12.07	11.10
Georgia	13.24	12.26	12.24	26.39	16.33	12.48	7.07
GDR	9.93	20.73	3.35	14.62	27.92	6.39	17.07
Germany	15.13	5.79	7.46	19.34	26.35	8.82	17.10
Great Britain	10.67	6.17	15.49	15.45	28.24	12.37	11.61
Greece	14.11	8.56	16.33	18.10	25.68	8.04	9.19
Hungary	7.29	7.10	11.48	23.94	22.97	12.95	14.27
Iceland	9.81	2.68	9.61	20.54	35.19	5.23	16.95
Ireland	7.94	5.56	10.67	26.37	32.09	9.59	7.79
Israel	20.94	3.81	2.20	18.25	17.31	31.29	6.20
Italy	6.51	7.55	25.56	22.51	19.02	10.00	8.85
Japan	20.40	6.32	29.16	21.59	13.54	2.47	6.53
Latvia	7.98	5.42	9.54	22.24	27.36	17.86	9.62
Lithuania	6.88	9.01	11.91	21.16	20.11	15.66	15.27
Luxembourg	6.55	7.08	9.97	22.81	32.13	7.62	13.84
Macedonia	10.50	14.49	6.28	25.80	16.43	16.57	9.93
Malta	4.85	6.61	17.51	22.09	33.16	5.74	10.05
Moldova	7.73	6.97	17.38	20.46	16.70	14.27	16.50
Montenegro	9.81	21.74	17.47	17.87	14.12	12.95	6.04
Netherlands	9.00	6.75	12.17	19.86	27.50	13.45	11.28
New Zealand	2.96	5.22	7.77	22.18	37.42	12.34	12.12
Norway	13.88	7.54	6.64	21.90	32.38	10.56	7.10
Poland	6.83	9.32	15.15	24.27	19.82	13.12	11.49
Portugal	8.85	6.95	12.92	21.82	34.18	4.38	10.91
Romania	7.24	12.74	6.56	24.76	22.18	11.28	15.24
Russia	7.48	6.87	17.11	25.26	16.13	17.26	9.89
Serbia	9.61	15.25	16.12	15.07	11.18	24.17	8.60
Slovakia	8.44	8.43	11.71	21.23	24.03	17.90	8.25
Slovenia	5.36	7.95	12.22	15.24	34.51	10.99	13.74
Spain	10.47	6.21	19.07	25.68	18.91	6.81	12.84
Sweden	12.41	8.04	3.25	23.57	37.66	7.13	7.96
Switzerland	9.68	8.28	8.96	20.38	24.02	20.08	8.59
Turkey	4.48	8.85	23.78	28.31	16.38	5.00	13.20
Ukraine	7.44	12.13	11.86	22.62	21.57	14.18	10.21
United States	17.25	4.87	15.55	16.74	18.17	19.35	8.07
AVERAGE	9.28	9.17	12.49	21.40	23.32	13.49	10.85

[6] Klingemann, Hans-Dieter (2006). Mapping policy preferences II: estimates for parties, electors, and governments in Eastern Europe, European Union, and OECD 1990-2003. Oxford, New York: Oxford University Press

Policy domains shares for Slovenia can be easily read from the table, although taken separately they have no substantial meaning. It is clear that the domains shown are not equally important (i.e., if a domain share in a document is taken as the indicator of importance). The welfare and quality of life domains are overrepresented and external relations, freedom and democracy, and fabric of society domains are more or less underrepresented (according to an ideal 14.3% share in the case of uniform distribution). Apart from that basic description, we cannot make any conclusions regarding commonness or uniqueness of the domain structure of Slovene manifestos. Clearly, we have to compare them with the overall structure (average) and further to all other countries. Figure 1, depicting national profiles and the average profile of policy domains importance, is created for that purpose. The Slovene profile is red and all other profiles are black. Among the black profile lines, the average profile is shown by the thick line.

Figure 1: National profiles, Slovene profile and the average profile of the importance of policy domains (focused graphical presentation of Table 2)



Graphical presentations convey more pieces of information at the same time and enable multiple comparisons on various levels.[7] Figure 2 readily shows that there is quite a variation in national profiles and that although the Slovene profile is not very different from the average profile, its specific characteristics cannot be neglected. Its most distinctive features are a considerably higher share of welfare and quality of life domains and substantially lower shares of external relations and economy domains. Regarding all three distinctive domains, the Slovene profile is quite extreme in comparison with other countries. Only Sweden and New Zealand pay more attention to the welfare and quality of life domain; only Australia and New Zealand pay less attention to the external relations domain; and Azerbaijan alone to the economy domain (these countries are recognized as shown in Table 2).

Apart from these Slovene characteristics, less obvious distinguishing features can be recognized only by more detailed and more formal analyses. Among them, direct comparisons domain by domain to determine particularities are straightforward and could be performed drawing on Table 2, in the same way as demonstrated above for the most obvious Slovene specific characteristics.

However, a more productive method is to employ a measure of profiles (countries) similarities or dissimilarities according to all seven domains taken jointly. Among possible measures, Euclidean distance[8] proved ideal for the purpose as it captures resemblances and differences in profiles in a way that is similar to the human eye and mind. It pays less attention to small differences and puts more stress on large differences. For example, small differences in all seven domains will result in lower Euclidean distance between two profiles than an exact match in six domains and a noticeable difference in one domain, even if a noticeable difference is much smaller than

[7] Henry, Gary T. (1995) Graphing Data. Techniques for Display and Analysis. London: Sage, and Tufte, Edward R. (1983) The Visual Display of Quantitative Information. Cheshire, CT: Graphic Press.

[8] See e.g., Johnson, A. Richard in Dean W. Wichern (1992). *Applied Multivariate Statistical Analysis*. London: Prentice Hall.

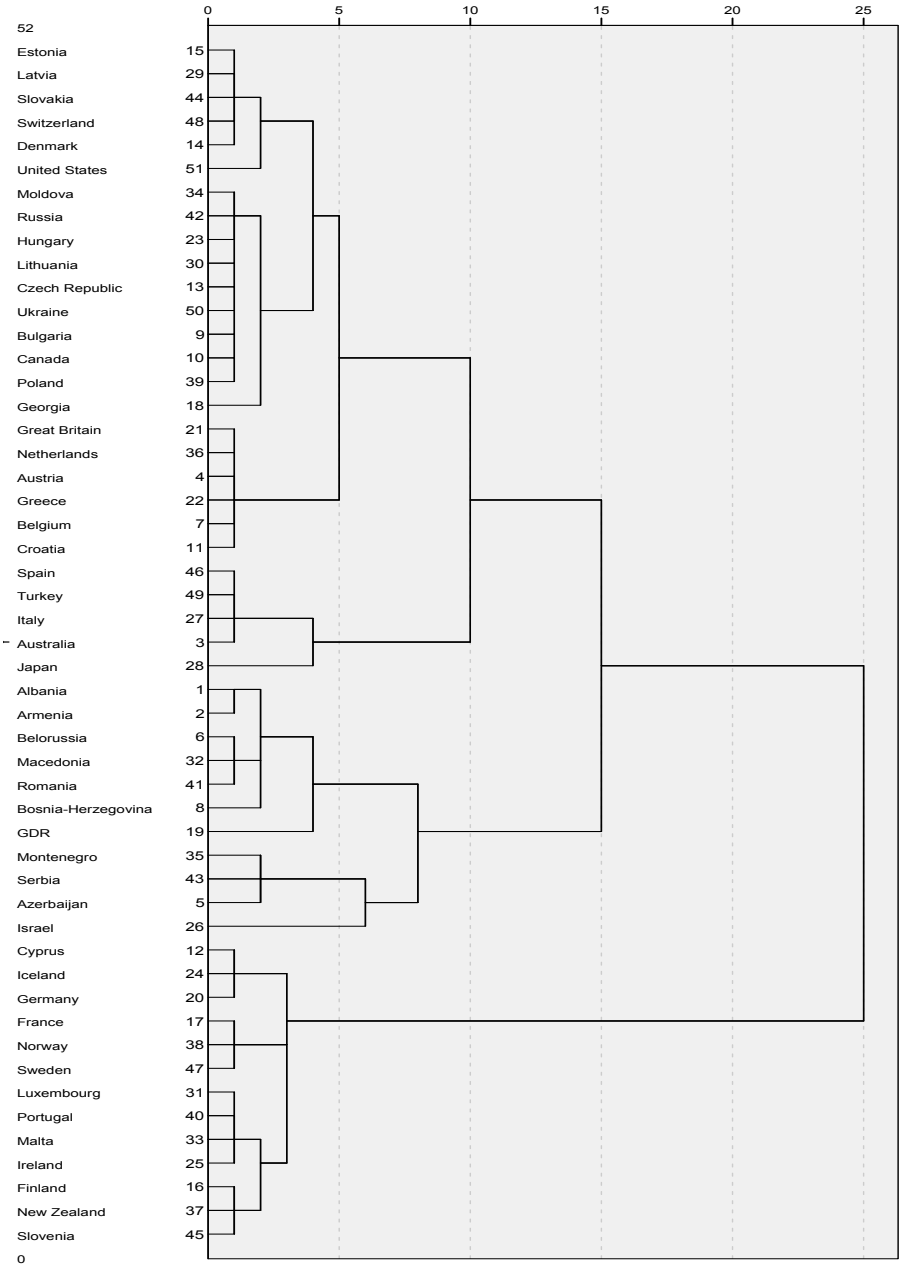
the sum of the seven small differences in the first case. Euclidean distance precisely and realistically captures differences in profiles and expresses them as numerical values that can be used for more formal comparisons among manifesto policy domains structures in different countries.

Furthermore, drawing on Euclidean distance and employing clustering methods,[9] we are able to uncover patterns, i.e., recognizable and relevant combinations that draw more attention to certain policy domains and less to others, which are typical of certain countries. In other words, we not only can realize how similar each country is to others but also can put together those countries that are very similar: countries in a group have to be as similar as possible and groups (types) as different (unique) as possible. Because this is a problem of optimization, both statistical and conceptual criteria affect the final solution and we understand the involvement of conceptual criteria to be an advantage. Thus, we can create a relevant typology of party manifesto structures, associate each country with the most appropriate type, and realize a country position in a type (group) as being more central, more peripheral or anywhere in the middle. In the case of agglomerative hierarchical clustering, a country membership and position can be understandably presented in the form of an agglomeration tree (dendrogram), i.e., a popular graphical presentation of a clustering procedure and result. To our knowledge, there is no other single approach or method that conveys that many relevant pieces of information in so condensed and straightforward form.

In the dendrogram below, countries are clustered according to their average policy domains structures on the basis of squared Euclidean distance (large differences become even more important) and the Ward method (balanced clustering criteria that respect group homogeneity and between-groups differences).

[9] Regarding the method description and clustering notions we refer to Johnson, A. Richard in Dean W. Wichern (1992). *Applied Multivariate Statistical Analysis*. London: Prentice Hall.

Figure 2: Agglomerative hierarchical clustering tree of countries



The dendrogram shown in Figure 2 yields the following:

First, it is reasonable to distinguish from two to five types of countries. However, three types seems to be the most balanced solution, as two are too superficial, four are too close to three or five, and five adds another small group to the previous four groups, among which one is already small.

Second, one group is always the same. The most stable group is composed of thirteen countries at the bottom of the dendrogram. Since all

countries have joined the group at a low level of dissimilarity, the group is very homogeneous. Only on a very low level of dissimilarity (which is irrelevant) can we recognize four subgroups.

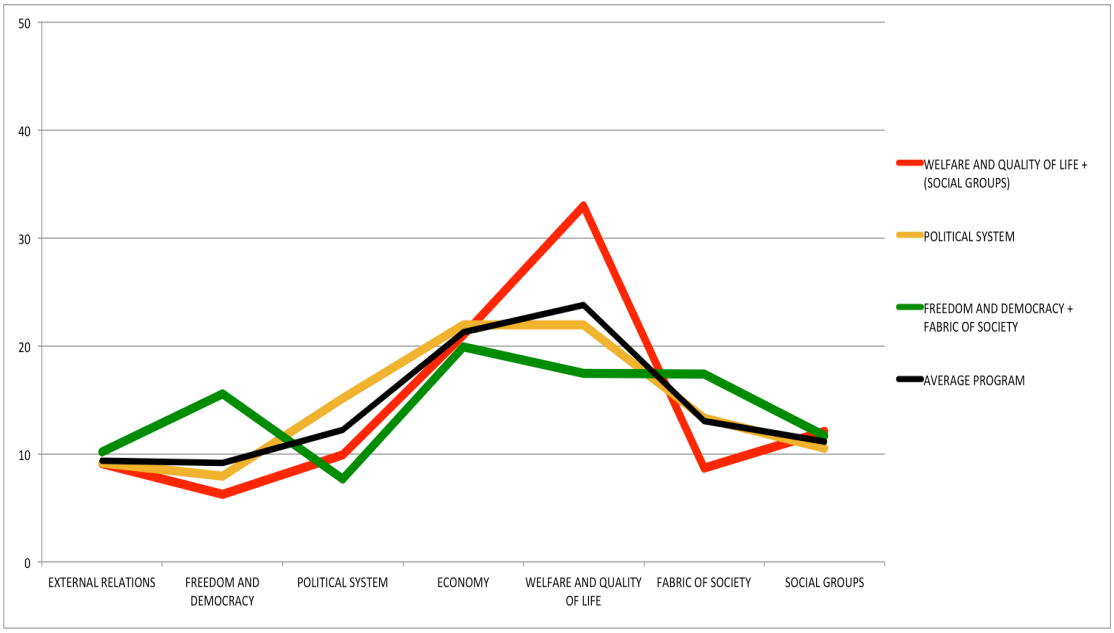
Third, Slovenia is in the most stable and homogeneous group. Inside the group, it most resembles Finland and New Zealand, followed by three subgroups of countries, the first including Luxemburg, Ireland, Portugal and Malta, the second France, Norway and Sweden and the last Germany, Iceland and Cyprus.

Fourth, next from the bottom to the top of the dendrogram shows the least stable group, which becomes a group in the case of the three-group solution and splits in two subgroups in the case of the five-group solution. Clearly it is less homogeneous. The members comprise eleven countries from Israel to Albania (from the bottom of the dendrogram up).

Fifth, the largest group includes twenty-seven countries from Estonia to Japan (listed from the top down) and splits into two unequal subgroups in the case of the four- or five-groups solutions. It is a large and a heterogeneous group.

To enable recognition of the character of each type (group), typical (ideal or average) profiles indicating the importance of policy domains in each group are presented in the graph below. The average profile of all countries is included as well.

Figure 3: Profiles of three types of countries and the average profile

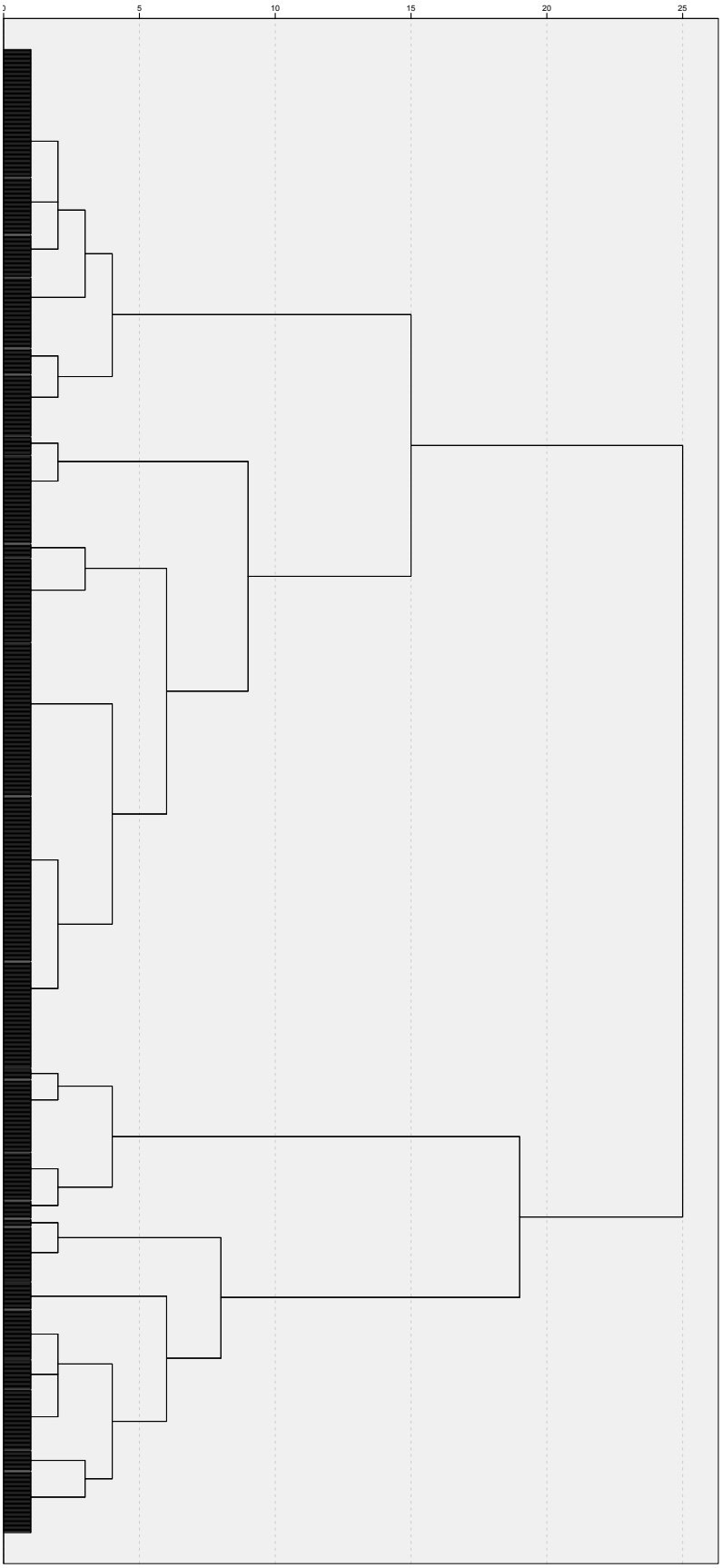


Differences between the groups and the average as well as differences among the groups are modest., however characteristics of the types can be recognized. The most stable group (the one including Slovenia), presented by the red profile, is the most specific because of the highest and most exposed average share of the welfare and quality of life domain in party manifestos, accompanied by lower shares of the fabric of society, freedom and democracy and political system domains. Clearly, for the countries in the “red” group, welfare and quality of life policy domain topics are the most important issue. The least stable group, presented by the green profile (including eleven countries in the middle of the dendrogram), is characterized by the highest share of two domains, namely the freedom and democracy domain and the fabric of society domain. In contrast to the first group, welfare and quality of life are the least important for the “green” countries. The largest group of countries, presented by the orange profile, is close to the average (obviously due to the size) and is specific only because of the highest (but not very eye-catching) interest paid to the policy domain topics.

The classification of countries on the basis of average shares of policy domains provides fundamental information for comparisons among countries but ignores differences inside countries, i.e., the level of nationalization of

party manifesto structures. It is important to determine whether all party manifestos are similarly structured and therefore almost identical to the average structure that credibly represents a country, or manifestos demonstrate significantly different structures and consequently the average is nothing more than an artificial construct that does not truly represent a country. In order to take the differences inside countries into account, in the second step of the study we return to original (basic, not aggregated) units of analysis and cluster individual party manifestos (almost 1300 units in the selected period) into groups (ideal types) using hierarchical agglomerative clustering method in the same way as we did before using countries as units of analysis.

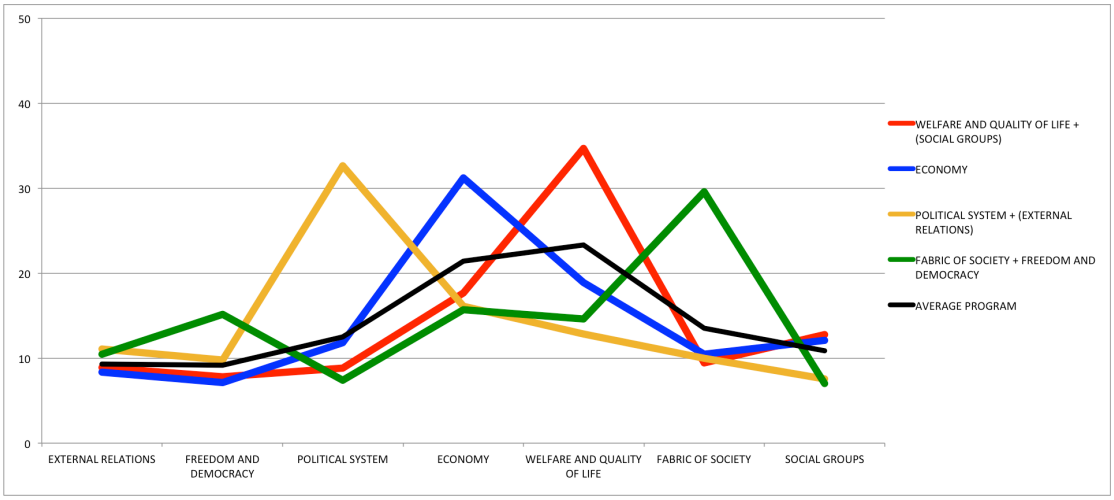
Figure 4: Agglomerative hierarchical clustering tree of party electoral programmes



The dendrogram above shows that the four-group solution is the most balanced and reasonable. Two- and three-groups solutions are too superficial since it is obvious that quite different groups are still joined together, five is too close to six, and the six-groups solution appears to be too particular since only a few manifestos split from two of the previous four groups, and the levels of dissimilarity are quite low. Additionally, the four-group solution has been confirmed by K-means method, as the largest drop in Ward criterion function arises when four groups replace three groups. Further, the solution (i.e. the classification of manifestos into four groups) has been optimized by K-means method.

Average importance profiles of the policy domains of each group are presented in the graph below to enable recognition of each type's character in the same way as in the case of countries (Figure 3). The average profile of all manifestos (nearly identical to the average of countries) is also included.

Figure 5: Profiles of four types of party manifestos and the average profile



Differences between groups and the average as well as differences among groups are now far more substantial, and type characteristics can be recognized clearly.

Regarding their interpretation, three profiles are the same as those shown in Figure 3 (profiles of groups of countries). As before, the red profile represents the group of manifestos with the highest average share of welfare and quality of life policy domain. As the Figure 3 and 5 show, both red profiles appear to be almost the same, which indicates that the group of manifestos highly resembles the group of countries. Notably, the “red” group is the largest one, containing 488 documents. The green profile is also very much the same as shown in Figure 3, i.e., it is characterized by the highest share of two domains, namely the freedom and democracy domain and the fabric of society domain. However, noticeably, in the case of clustering manifestos, the type has a more distinctive character since differences from the average profile and from the profiles of other groups are now far more obvious. The group contains 235 party programmes, which makes it a middle-sized group. Although the third, orange group of manifestos does not stand out in Figure 3, it also becomes far more distinctive in Figure 5 as the group’s interest paid to the political system domain topics is now more distinguishing. However, with 162 documents, it becomes the smallest group.

The fourth, blue profile is new and typical only of manifesto clustering. Notably, the blue group is the second largest group, containing 408 manifestos, so it cannot be overlooked in any case. Its main characteristic is the highest share in the economy domain, but in general, the blue profile is slightly less distinctive and closer to the average.

In the contingency table below (Table 3), the obtained four groups of manifestos (columns) are split according to the origin countries of the manifestos (rows) both to estimate differences in manifesto structures inside countries and to categorize countries according to in-country characteristics of manifesto structures. The cells contain row percentages and the total is shown in the last row.

Countries (rows) are arranged according to the three types of countries previously established (Figures 2 and 3). Country type is marked in front of country name in a consistent colour, as shown in Figure 3: 1/ in red for

welfare and quality of life country type; 2/ in green for freedom and democracy plus the fabric of society country type; and 3/ in orange for political system country type.

The country's name is written in a colour consistent with its in-country characteristic type of manifesto structure (same colours as in Figure 5). Also the column headings in Table 3 are shown in the same colours: red for welfare and quality of life manifesto type; green for freedom and democracy plus fabric of society manifesto type; orange for political system manifesto type; and blue for economy manifesto type. If a country cannot be categorized in a single category because it has two characteristic types of manifestos, both colours are used (half of the name is in one colour, and the other half is in the other colour). However, the first part of the name is in the dominant characteristic type colour. If a country cannot be categorized because of the lack of distinctive characteristics, its name is printed in black.

In-country characteristics of manifesto type are recognized by row percentages to describe the relative impact of a manifesto type in a country. Cells containing high row percentages are filled with colours. Red, orange and yellow are used to indicate an absolute dominance (a majority) of a certain type of manifesto on different intervals: red is used for extremely high shares, making all other types irrelevant (70% or more); dark orange is used for very high shares on a slightly lower interval, making all other types hard to affect the country type (60% to 70%); and yellow is used for high shares on an even lower interval that allows other types to be relatively influential (50% to 60%). All other shares representing the relative prevalence of a certain manifesto type in a country are coloured grey if they are at least 10 percentage points higher than the total.

The final recognition of in-country characteristics of manifesto type (categorization of countries) is based on a subjective estimation of the importance of a manifesto type in a country, which draws on both the absolute supremacy of a type on different levels and on the relative over-presence of manifesto types (one or more).

Table 3: Countries by types of manifestos

MANIFESTO TYPE / COUNTRY	WELFARE AND QUALITY OF LIFE	ECONOMY	POLITICAL SYSTEM	FABRIC OF SOCIETY + FREEDOM AND DEMOCRACY	NUMBER OF MANIFESTOS 1990 - 2003
1/ Slovenia	**73.3 %**	10.0 %	6.7 %	10.0 %	30
1/ New Zealand	**80.0 %**	20.0 %			25
1/ (Malta)	**100.0 %**				4
1/ Portugal	**77.3 %**	22.7 %			22
1/ Luxembourg	**70.0 %**	30.0 %			10
1/ Sweden	**79.3 %**	20.7 %			29
1/ Norway	**81.0 %**	19.0 %			21
1/ Iceland	**75.0 %**	20.0 %	5.0 %		20
1/ Cyprus	**70.0 %**		**30.0 %**		10
1/ Finland	**64.7 %**	20.6 %	2.9 %	11.8 %	34
1/ France	**63.2 %**	21.1 %		15.8 %	19
1/ Germany	**63.2 %**	36.8 %			19
1/ Ireland	**50.0 %**	**50.0 %**			18
2/ GDR	**86.7 %**	6.7 %		6.7 %	15
2/ Bosnia-Herzegovina	12.9 %	19.4 %		**67.7 %**	31
2/ (Azerbaijan)	22.2 %	11.1 %		**66.7 %**	9
2/ Israel	27.3 %	12.1 %		**60.6 %**	33
2/ (Belorussia)	28.6 %	28.6 %		**42.9 %**	7
2/ Macedonia	8.0 %	**40.0 %**	8.0 %	**44.0 %**	25
2/ Montenegro	8.3 %	25.0 %	**25.0 %**	**41.7 %**	24
2/ Serbia	17.1 %	17.1 %	**25.7 %**	**40.0 %**	35
2/ Armenia	17.6 %	**52.9 %**		**29.4 %**	17
2/ Albania	5.4 %	**62.2 %**	13.5 %	18.9 %	37
2/ Romania	35.3 %	**44.1 %**		20.6 %	34
3/ Japan	6.3 %	31.3 %	**62.5 %**		32
3/ Croatia	27.3 %	21.2 %	**30.3 %**	21.2 %	33
3/ Italy	12.5 %	**45.8 %**	**37.5 %**	4.2 %	48
3/ Russia	8.3 %	**52.8 %**	**22.2 %**	16.7 %	36
3/ Turkey	9.5 %	**57.1 %**	**33.3 %**		21
3/ (Moldova)	25.0 %	**50.0 %**	**25.0 %**		4
3/ Belgium	**54.1 %**	8.1 %	**32.4 %**	5.4 %	37
3/ Greece	**52.9 %**	17.6 %	**29.4 %**		17
3/ (United States)	12.5 %		**37.5 %**	**50.0 %**	8
3/ Great Britain	**70.8 %**	12.5 %	12.5 %	4.2 %	24
3/ Netherlands	**62.1 %**	27.6 %	6.9 %	3.4 %	29
3/ Estonia	**55.2 %**	20.7 %	3.4 %	20.7 %	29
3/ Hungary	**53.8 %**	23.1 %	7.7 %	15.4 %	26
3/ Latvia	**53.3 %**	26.7 %	3.3 %	16.7 %	30
3/ Austria	**47.8 %**	30.4 %	17.4 %	4.3 %	23
3/ Spain	3.3 %	**66.7 %**	**30.0 %**		30
3/ Australia	22.2 %	**55.6 %**	16.7 %	5.6 %	18
3/ Bulgaria	17.9 %	**53.6 %**	3.6 %	25.0 %	28
3/ Czech Republic	18.8 %	**50.0 %**	9.4 %	21.9 %	32
3/ Canada	26.7 %	**40.0 %**	13.3 %	20.0 %	15
3/ Georgia	17.2 %	**44.8 %**	17.2 %	20.7 %	29
3/ Lithuania	20.8 %	**45.8 %**	4.2 %	**29.2 %**	24
3/ Denmark	41.2 %	29.4 %		**29.4 %**	34
3/ Switzerland	45.5 %	22.7 %	4.5 %	**27.3 %**	44
3/ Ukraine	29.4 %	32.4 %	5.9 %	**32.4 %**	34
3/ Poland	29.3 %	39.0 %	12.2 %	19.5 %	41
3/ Slovakia	33.3 %	35.9 %	7.7 %	23.1 %	39
TOTAL	37.7 %	31.6 %	12.5 %	18.2 %	1293

Table 3 reveals the following information:

First, starting from previously established groups of countries, the first group marked as 1/ and characterized as the welfare and quality of life policy domain group proved to be very well founded and perfectly stable. For the second time, all thirteen countries are categorized as a group, and the dominance of the welfare and quality of life policy domain is in all cases absolute and on a high scale at mostly over 70% (red cells). The only less convincing case is Ireland, where we observe a 50/50 split between the group specific domain and the economy domain. The second group, marked as 2/ and labelled the freedom and democracy and the fabric of society domains group becomes less definite and less stable since only three out of eleven countries are categorized in the same group on the basis of a convincing absolute dominance of the characteristic two domains (orange cells). One is included on the basis of their relative prevalence (grey cell); four are split between the type and other types, and the final three are categorized in another group (one, i.e., GDR clearly in the welfare and quality domain group with 86% of characteristic domain manifestos). In the case of the third group, marked as 3/ and named the political system domain group, the evidence is even weaker. Only two countries can be categorized in the same group, one on the basis of the absolute prevalence of political system domain manifestos and the other on the basis of their relative over-presence. Among the others, seven are split between this type and another type, sixteen are categorized in one of the other three types, and two remain uncategorized (there is not even a relatively higher segment of any type).

Second, the welfare and quality of life group as well as the freedom and democracy and fabric of society domains group show new members. The first of the two includes more new countries, and as a rule they enter with an absolute dominance of the characteristic domain type (red, orange or yellow cells), and only two are split between manifesto types. The second group has fewer new countries and typically they join with merely relative over-presence of the type (grey cells). The policy system group has no new members, but it loses a great number of countries due to their relocation into the other three groups. The fourth group, the economy domain group, materializes from the previous freedom and democracy and fabric of society domains group of

countries and from the weakening policy domain group. The economy domain type of manifesto prevalence is absolute (orange and yellow cells) as well as relative (grey cells).

Third, twelve countries cannot be categorized in a single type since party manifestos belong to two significant types. Additionally, two countries cannot be classified since their manifestos express no pattern. These results do not occur when countries are classified on the basis of an average manifesto structure (so they are overlooked).

Fourth and finally, based on previous points, it is safe to conclude that a categorization of countries drawing on the classification of party manifestos yields better results than a classification of countries on the basis of their average manifesto profiles. The results of both procedures are reasonably similar only in the case of the welfare and quality of life group, although this group gains new country members with clear crucial domain dominance, which were misclassified when the countries were clustered. Additionally, when countries are categorized on the basis of the influence of in-country types of party manifestos, countries of the same type can be distinguished and presented according to the strength of their link to a type, i.e., according to the level of prevalence of a manifesto type, which is expressed in Table 3 as different cell colours. Moreover, countries can be categorized as split between types, which is a unique but very realistic feature. Similar is true for uncategorized countries.

Final Results—the Galaxy

To employ features described in the last paragraph and provide a clear, conclusive picture of the manifesto structures of Slovenian parliamentary parties in comparison with those of other countries, we depict a galaxy i.e. a graph developed for that purpose. The central (referential) country, in our case Slovenia, is shown as the Sun and all other countries are shown as different planets allocated around the Sun proportionate to Euclidean distances. Countries (planets) are depicted according to their manifesto policy domain structures, taking into account the type and the level of homogeneity,

both of which are recognized drawing on clustering results for the manifestos. The countries (planets) are coloured according to their prevailing types of manifesto structures as follows:

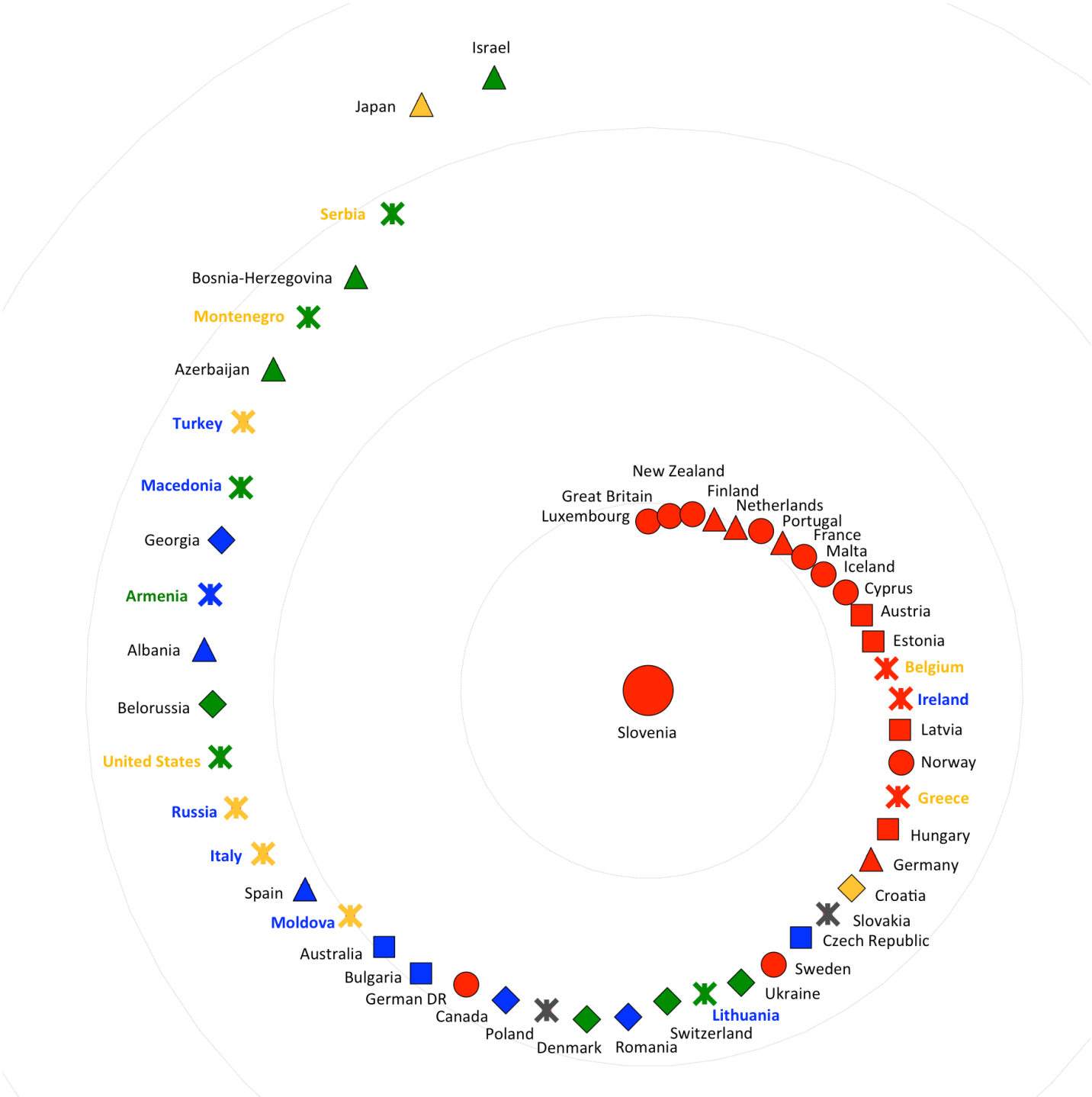
- Red for the welfare and quality of life domain group
- Green for the freedom and democracy and fabric of society domains group
- Orange for the political system group,
- Blue for the economy domain group
- Grey for the two uncategorisable countries

If there is more than one characteristic policy domain, a country name is written in the colour of the other, as a rule the less characteristic one.

Different shapes of planets represent different levels of manifesto structure homogeneity in a country, which can also be taken as an indicator of the strength of the country's categorization:

- A circle is used for the highest homogeneity (absolute dominance of a single manifesto structure type with an over 70% share)
- A triangle stands for high homogeneity (absolute dominance of a single manifesto type with a 60% to 70% share)
- A square represents modest homogeneity (absolute dominance of a single manifesto type with a 50% to 60% share and the absence of any other over-presented type)
- A rhombus is used to indicate modest variety (relative prevalence of a single manifesto type with a share at least 10 percentage points higher than the total and the absence of any other over-presented type
- A star represents a clear split in manifesto types (two types of manifestos are characteristic for a country).

Figure 6: The Galaxy with Slovenia in a central position



Drawing on the Galaxy shown in Figure 6 and comparing Slovenia with other countries, one can recognize Slovenia as a country with a clear dominance of parliamentary party manifestos that favour the welfare and quality of life policy domain. Most Slovenian party manifestos, more than 70%, are classified as

emphasizing the welfare and quality of life policy domain (red circle). Therefore, Slovenia can be categorized as a country with a homogenous party manifesto structure that is strongly associated with the welfare and quality of life domain (a strong and unmistakable case of group membership). Countries that are most similar to Slovenia are Luxemburg, Great Britain and New Zealand. All three are strong cases of belonging to the welfare and quality of life domain type (red circles). The next two similar countries are Finland and Netherlands, both belonging to the same welfare and quality of life manifesto type, but the characteristic domain prevalence is lower (red triangles). The further from Slovenia we move, the less welfare and quality of life type countries and the more heterogeneous countries we find. Readers can interpret other countries categorisation and their level of similarity with Slovenia in the same manner as the graph speaks for itself.

Conclusion

The clustering approach has enabled us to compare party manifestos as both individual documents and country representatives. It has also provided a method to estimate the level of similarity between units of interest on both levels of comparison and in general.

The strong points of the applied approach can be summarized as follows:

First, political party manifestos and countries are analysed according to all seven policy domains simultaneously in a multivariate manner (the opposite would be to consider each domain separately).

Second, Euclidean distance represents a very realistic view of the level of similarity or difference between units (countries or manifestos), which is close to what we understand as distance in everyday life. In addition, clustering algorithms use Euclidean distance in an easy to understand way and produce vivid graphical outputs, which make the research results both comprehensible and convincing (the opposite would be to use latent concepts and rather abstract notions of covariation in multidimensional space, which are difficult to comprehend for less empirically oriented scholars).

Third, in our quest for the best classification we can chose among different suggested solutions from more general (less types equals less in-group homogeneity and more differences among groups) to more precise (more types equals less differences among groups and more homogeneous groups). In defining the best solution, i.e., one that is subjectively considered the most balanced, we can (and we usually have to) apply additional, contextual and theoretical criteria (e.g., the minimum size of average differences in percentage points that we understand as a difference and do not neglect in interpretations). The use of contextual and theoretical principles together with statistical indicators leads to more convincing results.

Fourth, in our case the clustering results on the level of manifestos are used to categorize countries, drawing on our understanding of the size and the meaning of structures of in-country manifesto types.

Fifth, in line with the prevailing Euclidean space based analyses, Figure 6 (the Galaxy) vividly summarizes the results and enables a focused comparison of a selected country with all other countries without any falsification (the opposite would be, e.g., any kind of projection of multiple space into two dimensions).

In more substantial manner, regarding past, present and future political science manifesto studies of comparisons among countries, the most important conclusion is that in some countries, manifestos belong to one dominant type, and therefore these countries make sense as aggregates of manifestos when analyses on country levels are performed. These countries are authentically represented by an average policy domains structure. On the contrary, in other countries, manifestos (parties) are clearly split between types. Therefore, using these countries as aggregates does not make sense because the structures of their manifestos are too different. Moreover, an average structure inadequately represents actual manifestos. Unfortunately for studies using countries as aggregates of manifestos, the latter (the heterogeneous) countries are in majority.