**APPENDIX**

1. **Importance of the first confidence vote for each new government**
2. **Pre-processing of the texts**
3. **Robustness check**
4. **Examples of policy-related dissimilarities**
5. **List of Italian stop-words**
6. **List of German stop-words**

**1. Importance of the first confidence vote for each new government**

Among others, Giannetti and Laver (2005), Ieraci (2006), Curini and Martelli (2009; 2010), Curini and Pinto (2013; 2017) have shown that MPs’ speeches in the moment of the first vote of confidence for a new government are one of the most reliable ways to detect party positioning in the Parliamentary arena. Moreover, this is one of the few instruments (if not the only one) allowing researchers to detect Prime Minister’s preferences without assuming that his/her position is dependent from the ones of the parties supporting his/her government (as we show later in this comment). Furthermore, reliability tests and comparative analyses with expert surveys, roll call voting studies, and with manifestos’ content analysis confirm the reliability of the analysis of such data as a source for detecting party positioning (Curini and Martelli 2009). After having recognised the appropriateness of parliamentary speeches, there is still an issue that deserves to be addressed: are the speeches connected to the first confidence vote the most suitable source of information concerning parties’ and PMs’ preferences? We do think so for a number of reasons: first of all, if we had analysed all the parliamentary debates, we would have been forced to select a certain number of issues discussed in Parliament, being not feasible to collect all the parliamentary speeches by all Italian and German MPs between 1994 and 2014 on all matters, and this would have also brought, given the quantity and the quality of such pieces of textual information, to imprecise results. Another solution would have therefore been that of selecting only debates concerning a limited number of issues. Nonetheless, if we had chosen this path, we would have introduced a selection bias in our analysis: indeed, analysing such limited number of issues would have brought us to infer MPs’ (parties’) and PMs’ positions, but this would have not allowed us to take into consideration the fact that parties (and therefore MPs) do assign different salience and do have more or less extreme (polarised) positions depending on the issue(s) taken into consideration. Moreover, which issues could have been chosen? It is evident that the answer to this question would have made us incur in a selection bias problem. On the contrary, during the debate on the first confidence vote for each new government, the Prime Minister ‘expounds in detail the programme of the government and the government’s future policy options, after which parliamentary debate is opened, various MPs of the parties represented in Parliament react to it, declaring their approval or disapproval and eventually a vote of confidence is delivered’ (Ieraci, 2006: 262). These speeches are therefore quite appropriate for our purposes (for instance, for the possibility not to assume that that PM’s policy stances are dependent from the ones of the parties that support his/her government): indeed,

parties’ programmes are obviously elaborated in the pre-electoral phase and are addressed mainly—if not exclusively—to the voters, with little consideration of the policy offers of the other parties or of the feasibility of the policies in the institutional context. In contrast, both the government programme delivered in the parliamentary arena and the party reactions to it are necessarily affected by quite a number of practical concerns. For example, the policy proposals have to be sustainable vis-à-vis the opposition, and the ‘viscosity’ of the parliamentary arena, that is the likelihood that policy proposals will go smoothly through the legislative decision- making, has to be taken into account both by parties and governments. Finally, parties and coalitions are concerned about their future options and fear the reaction of public opinion.’ (Ieraci, 2006: 262).

**2. Pre-processing of the texts**

Firstly, we have created a document-feature-matrix for each corpus, that is, a matrix presenting the documents in rows and the features’ frequencies (the unique words in the used corpus) in columns[[1]](#footnote-1). Secondly, we have removed the so-called stop-words for both languages (Italian and German), i.e. words that are non-descriptive for the topic of a document, such as *a*, *and*, *are* and *do*. Since the lists of stop-words present in the quanteda package are not exhaustive, we have decided to make them more complete: as shown in this appendix. Thirdly, we have also removed MPs’ and Government ministers’ names, since they do not provide any substantial information. Fourthly, we have stemmed words using Porter’s suffix-stripping algorithm (Porter 1980), so that words with different endings will be mapped into a single word. This algorithm is implemented in R package SnowballC. The supported languages are, at the moment, Danish, Dutch, English, Finnish, French, German, Hungarian, Italian, Norwegian, Portuguese, Romanian, Russian, Spanish, Swedish and Turkish. The underlying assumption is that different morphological variations of words with the same root/stem are thematically similar and should be treated as a single word (e.g. the words ‘legislation’, ‘legislator’, and ‘legislate’ are all reduced to the stem: legisl-). Finally, we have decided to discharge extremely rare terms, by removing features occurring less than three times in less than 5% of the documents present in each corpus. The rationale behind discarding such infrequent terms is that in many cases they do not provide a connotation of the documents and may be of little contribution in the assessment of the similarity between texts (Slapin and Proksch 2008).

We have also decided to apply the $tf∙idf$ weighting system (Salton, 1989; Manning, Raghavan and Schütze, 2008)[[2]](#footnote-2) to weight term frequencies in the corpora used. The $tf∙idf$ is a weighting system whose value increases proportionally with the number of times a word appears in a document but is balanced by the frequency of that word in the whole corpus of documents. In other words, this weighting system helps researchers adjust for words that appear more frequently in the speeches. $tf$corresponds to the normalised term frequency, measuring how frequently a term occurs in a document, and $idf$ corresponds to the inverse document frequency, measuring how important the term in a specific corpus is[[3]](#footnote-3).

**3. Robustness check**

We have run an additional regression controlling for the fact that the PM might also be the leader of a specific party in a coalition. It might be the case that if a PM is the leader of a specific party, he/she might have a different behaviour towards this party compared to the behaviour he/she might have towards other coalition parties. To control for this possibility, we have regressed PMPA on the interaction between MINISTER RATE and a dichotomous variable operationalising whether the PM is also the leader of a specific coalition party. The coefficient of the interaction is not significant, and the presence of the interaction neither changes other variables’ coefficients and their significance levels nor increases the goodness of fit of the model (see Table 1 below).

**Table 1 - Multilevel regression on PMs’ Policy Autonomy (PMPA), Germany and Italy (1994-2014)**

|  |  |  |  |
| --- | --- | --- | --- |
|  | Model I | Model II | Model III |
| MWC |  | 0.410\* | 0.404\* |
|  |  | (0.165) | (0.168) |
| POLARISATION |  | -3.885\*\*\* | -3.944\*\*\* |
|  |  | (0.466) | (0.477) |
| NUMBER OF PARTIES |  | 0.080\*\* | 0.083\*\* |
|  |  | (0.025) | (0.026) |
| MINISTER RATE |  | -0.005\*\* | -0.005\*\* |
|  |  | (0.002) | (0.002) |
| DISTANCE |  | -0.01 | -0.014 |
|  |  | (0.237) | (0.236) |
| PM SELECTION |  | 0.066 | 0.065 |
|  |  | (0.034) | (0.035) |
| PARTY LEADER |  |  | -0.385 |
|  |  |  | (0.262) |
| MINISTER RATE | PARTY LEADER |  |  | 0.007 |
|  |  |  | (0.005) |
| Constant | 4.738\*\*\* | 5.893\*\*\* | 5.915\*\*\* |
|  | (0.118) | (0.358) | (0.367) |
| AIC | 48.38 | 36.834 | 50.271 |
| Log Likelihood | -21.19 | -9.417 | -14.135 |
| Num. obs. | 71 | 71 | 71 |
| N Governments | 17 | 17 | 17 |
| Variance Governments | 0.221 | 0.015 | 0.016 |
| \*\*\*p < 0.001, \*\*p < 0.01, \*p < 0.05; Standard errors in parentheses |

Moreover, we have also tested whether the number of texts included of each party in each coalition has an impact on the PMPA. The results of the analysis can be found in Table 2 below, and show that there is not a statistically significant impact of the abovementioned variable on PMPA.

**Table 2 - Bivariate regression of the number of texts for each party on PMs’ Policy Autonomy (PMPA)**

|  |  |
| --- | --- |
|  | Model I |
| Number of texts | -0.005 |
|  | (0.013) |
| Constant | 4.85 |
|  | (0.092) |
| $$R^{2}$$ | 0.002 |
| F-statistic | 0.148 |
| \*\*\* p< 0.001; \*\* p< 0.01; \* p< 0.05; Standard Errors in Parentheses. |

**4. Examples of policy-related dissimilarities**

Berlusconi (PM) (2005) – ‘Concerning the South [of Italy], our action will focus on some precise targets: the prompt use of 22.5 billion Euros available in capital grants, the immediate start of a revolving fund to channel towards the South new and additional instruments of bank credit, the introduction of fiscal instruments in favour of firms, the acceleration towards the completion of Southern Italian infrastructures, started with the *Legge Obiettivo* [‘Target Law’]’

Northern League (2005) – ‘There was a lack in tailoring the subvention policy towards Southern regions. On the contrary, there was a preference towards the previous policy of financing via old means of intervention in Nothern regions, while the Southern regions surely need new measures. Therefore, apart from a stronger communication about the targets already accomplished, this government should have the ability and the courage (and it is not easy to say such things) to make new generations understand that surely help could come from the rest of the country and from Europe, but the biggest help should spontaneously grow in Southern regions’

--

Merkel (PM) (2005) – ‘On the benefits side, however, we will make quick changes. We want more freedom of contract and design options from the patients about the health insurance companies to the practices and hospitals’

SPD (2005) – ‘[O]ur traditional institutions, pension insurance, health insurance and long-term care insurance, have such a long tradition - and this applies to the first two - and that it is worthwhile continuing to be among the most important guarantors of the welfare state in Germany’

--

Merkel (PM) (2005) – ‘The strengthening of international institutions is vital in the face of globalization. A policy that claims to shape globalization - that is what we have to claim, even if many people feel that politics cannot do this - must not go beyond international institutions, but must include international institutions to empower globalization’

SPD (2005) – ‘Moreover, in the storm of globalization, many citizens no longer see international institutions (among them the EU) as a social protector, but rather as a Trojan horse of neoliberalism’

--

Letta (PM) (2013) – ‘It is necessary to overcome the actual system of taxation for the first home, starting from June instalment to all the government and the parliament to devise and rapidly implement an overall reform to give oxygen to families, especially the less well-off ones’

Partito Democratico (2013) – ‘I believe, for instance, that it is not right that the instalment for the first home will be suspended for some us, given our income, in June’

**5. List of Italian stop-words**

"ad","al","allo","ai","agli","all","agl","alla","fra","quella","de",

 "alle","con","col","coi","da","dal","dallo","dai","ce",

 "alcuna","alcuno","alcuni","alcune","alcun","almeno","allora",

 "dagli","dall","dalla","dalle","di","del","dello","quegli",

 "dei","degli","dell","degl", "della","delle","in","esso","essi","esse",

 "nel","nello","nei","negli","nell","negl","nella",

 "nelle","su","sul","sullo","sui","sugli","sull",

 "sugl","sulla","sulle","per","tra","contro","io","tu",

 "lui","lei","noi","voi","loro","moi","mia", "miei","mie",

 "tuo","tua","tuoi","tue","suo","sua","sue","suoi","nostro",

 "nostra","nostri","nostre","vostro","vostra","vostri",

 "vostre","mi","ti","ci","vi","lo","la","li","le","gli","ne",

 "il","uno","una", "ma","ed","se","perché","anche","come","dov",

 "dove","che","chi","cui","non","più","quale","quanto","quante",

 "quanti","quanta","quello","quelli","quelle","questo","questi",

 "questa","queste","codesto","codeste","codesta","codesti","si",

 "tutto","tutte","tutti","tutta","a", "e","é", "c","i","l","o",

 "ho","hai", "ha","abbiamo","avete","hanno","abbia","abbiate",

 "avrete","abbiano","avrò","avrai","avrà","avremo","avranno",

 "avrei","avresti","avrebbe","avremmo","avreste","avrebbero",

 "avevo","avevi","aveva","avevamo","avevate","avevano","ebbi",

 "avesti","ebbe","avemmo","aveste","ebbero","avessi","avesse",

 "avessimo","avessero","avendo","avuto","avuta","avuti","avute",

 "sono","sei","è","siamo","siete","sia","siate","siano","sarò",

 "sarai","sarà","saremo","sarete","saranno","sarei","saresti",

 "sarebbe","saremmo","sareste","sarebbero","ero","eri","era",

 "eravamo","eravate","erano","fui","fosti","fu","fummo","foste",

 "furono","fossi","fosse","fossimo","fossero","essendo","faccio",

 "fai","facciamo","fanno","faccia","facciate","facciano","farò",

 "farai","farà","faremo","farete","faranno","farei","faresti",

 "farebbe","faremmo","fareste","farebbero","facevo","facevi",

 "faceva","facevamo","facevate","facevano","feci","facesti","laddove",

 "fece","facemmo","faceste","fecero","facessi","facesse","facessimo",

 "facessero","facendo","sto","stai","sta","stiamo","stanno","stia",

 "stiate","stiano","starò","starai","starà","staremo","starete",

 "staranno","starei","staresti","starebbe","staremmo","stareste",

 "starebbero","stavo","stavi","stava","stavamo","stavate","stavano",

 "stetti","stesti","stette","stemmo","steste","stettero","stessi",

 "stesse","stessimo","stessero","stando","all", "dall","dell","l",

 "un","nell", "sull","l","pag","i","x","v","ii","iii","iv","vi",

 "vii","viii","ix","xi","xii","xiii","xiv","xv","xvi","xvii","xviii",

 "xix","xx","xxi","xxii","xxiii","xxiv","xxv","xxvi","xxvii","xxviii",

 "xxix","xxx","xxxi","xxxii","xxxiii","xxxiv","xxxv","xxxvi","xxxvii",

 "xxxviii","xxxix","xl","l","c","uno","due","tre","quattro","cinque","sei",

 "sette","otto","nove","dieci","undici","dodici","tredici","quattordici",

 "quindici","sedici","diciassette","diciotto","diciannove","venti",

 "ventuno","ventidue","ventitre","ventiquattro","venticinque","ventisei",

 "ventisette","ventotto","ventinove","trenta","qualsiasi","qualunque",

 "gennaio","febbraio","marzo","aprile","maggio","giugno","luglio","agosto" "settembre","ottobre","novembre","dicembre","lunedì","martedì","mercoledì", "giovedì","venerdì","sabato","domenica","oggi","ieri","domani","Applausi","gruppi","gruppo","Popolo","Libertà","Partito","Democratico","Lega","Nord","Italia","Valori","Unione","Centro","anche","Signor","Signori","Signore","Signora","onorevoli","colleghi","collega","ministri","ministro","Presidente","Consiglio","consiglio","governo","membri","membro","deputati","deputato","deputata","parlamentare","parlamentari","a","b","c","d","e","f","g","h","i","l","m","n","o","p","q","r","s","t","u","v","z","w","x","y","j","k", "n.","p.", "h."

**6. List of German stop-words**

"zwölf","zwischen","zweites","zweiter","zweiten","zweite","zwei","zwar","zwar","zwanzig","zusammen","zurück","zur","zunächst","zum","zum","zugleich","zuerst","zu","zeit","zehntes","zehnter","zehnten","zehnte","zehn","z","z.b","y","x","würden","wurden","würde","wurde","worden","wollten","wollte","wollt","wollen","wohl","wo","wirst","wirklich","wird","wir","willst","will","wieder","wie","wie","wessen","werdet","werden","werde","wer","wenn","wenn","wenigstens","weniges","weniger","wenige","wenig","wenig","wen","wem","welches","welcher","welchen","welchem","welche","weiteres","weiteren","weitere","weiter","weit","weil","wegen","was","warum","wart","waren","wäre","war","wann","währenddessen","währenddem","während","wahr","w","vor","von","vom","viertes","vierter","vierten","vierte","vier","vielleicht","vielen","vielem","viele","viel","vergangenen","v","unter","unserer","unsere","unser","uns","und","und","um","uhr","übrigens","überhaupt","über","u","tun","trotzdem","tritt","tel","teil","tat","tagen","tage","tag","t","statt","später","sowie","sonst","sondern","sollten","sollte","sollen","soll","solches","solcher","solchen","solchem","solche","solang","so","sind","siebentes","siebenter","siebenten","siebente","sieben","sie","sich","selbst",“seitdem","seit","seines","seiner","seinen","seinem","seine","sein","seien","seid","sei","sehr","sechstes","sechster","sechsten","sechste","sechs","schon","Schluss","schlecht","satt","sah","sagte","sagt","sache","sa","s","rund","richtig","rechtes","rechter","rechten","rechte","recht","r","q","p","Ordnung","ohne","oft","oft","offen","oder","oben","ob",,"o","nur","nun","noch","niemanden","niemandem","niemand","nie","nichts","nicht","neuntes","neunter","neunten","neunte","neun","neuen","neue","nein","neben","natürlich","nahm","nachdem","nach","na","n","mussten","musste","müsst","musst","müssen","muß","muss","morgen","mögt","möglich","mögen","mochten","möchte","mochte","mittel","mit","mir","mich","menschen","mensch","meines","meiner","meinen","meinem","meine","mein","mehr","mann","manches","mancher","manchen","manchem","manche","man","mahn","magst","mag","machte","macht","machen","m","los","lieber","leide","leicht","lange","lang","l","kurz","konnten","könnte","konnte","könnt","können","kommt","kommen","kleines","kleiner","kleinen","kleine","keiner","keinen","keinem","keine","kein","kaum","kannst","kann","kam","k","jetzt","jenes","jener","jenen","jenem","jene","jemanden","jemandem","jemand","jedoch","jedermanns","jedermann","jeder","jeden","jedem","jede","je","jahren","jahre","jahr","ja","j","ist","irgend","ins","infolgedessen","indem","in","immer","im","ihres","ihrer","ihren","ihrem","ihre","ihr","ihnen","ihn","ihm","ich","i","hoch","hinter","hin","hier","heute","her","heisst","hätten","hatten","hätte","hatte","hat","hast","habt","haben","habe","h","gutes","guter","gute","gut","großes","grosses","großer","grosser","großen","grossen","große","grosse","groß","gross","gott","gleich","ging","gibt","geworden","gewollt","gewesen","geschweige","gesagt","gern","gerade","genug","gemusst","gemocht","gemacht","gekonnt","gekannt","geht","gehen","gehabt","gegenüber","gegen","gedurft","gar","ganzes","ganzer","ganzen","ganze","ganz","gab","g","für","fünftes","fünfter","fünften","fünfte","fünf","früher","f","euch","etwas","etwa","es","erstes","erster","ersten","erste","erst","Ernst","er","entweder","endlich","ende","en","elf","eins","einmal",“einiges","einiger","einigen","einige","eines","einer","einen","einem","eine","einander","ein","eigenes","eigener","eigenen","eigene","eigen","ei",,"ehrlich","ebenso","eben","e","durften","durfte","dürft","dürfen","durchaus","durch","du","drittes","dritter","dritten","dritte","drin","drei","dort","doch","dir","dieses","dieser","diesen","diesem","dieselben","dieselbe","diese","dies","diejenigen","diejenige","die","dich","deswegen","dessen","desselben","deshalb","des","derselben","derselbe","dermaßen","dermassen","derjenigen","derjenige","deren","der","denselben","denn","denen","den","demzufolge","demselben","demgemäß","demgemäss","demgegenüber","dementsprechend","dem","deiner","deinem","deine","dein","dazwischen","dazu","davor","davon","dasselbe","daß","dass","daselbst","dasein","das","darunter","darum","darüber","darin","darfst","darf","daraus","darauf","daran","dann","dank","daneben","danach","damit","damals","dahinter","dahin","daher","dagegen","dafür","dadurch","dabei","da","d","d.h","c","bist","bisher","bis","bin","besten","besser","besonders","bereits","bekannt","beispiel","beim","beiden","beide","bei","bald","b","außerdem","ausserdem","außer","ausser","aus","auf","auch","auch","au","anders","andern","anderen","andere","an","am","also","als","allgemeinen","alles","allerdings","aller","allen","allem","allein","alle","ag","achtes","achter","achten","achte","acht","ach","aber","aber","ab","a".

**REFERENCES**

Curini, L. and Pinto, L. (2013), ‘Government formation under the shadow of a core party. The case of the First Italian Republic’, *Party Politics* 19(3): 502-522.

Curini, L. and Pinto, L. (2017), *L’arte di fare (e disfare) i governi: da De Gasperi a Renzi, 70 anni di politica italiana*, Milano: Egea.

Curini, L. and Martelli, P. (2009), *I Partiti della Prima Repubblica. Maggioranze e Governi dalla Costituente a Tangentopoli*, Roma: Carocci.

Curini, L. and Martelli, P. (2010), ‘Ideological proximity and valence competition. Negative campaigning through allegation of corruption in the Italian legislative arena from 1946 to 1994’, *Electoral Studies* 29(4): 636-647.

Giannetti, D. and Laver, M. (2005), ‘Policy position and jobs in the government’, *European Journal of Political Research* 44(1): 91-120.

Ieraci, G. (2006), ‘Governments, Policy Space and Party Positions in the Italian Parliament (1996-2001): An Inductive Approach to Parliamentary Debate and Votes of Investiture’, *South European Society and Politics* 11(2): 261-285.

Porter, M. F. (1980), ‘An algorithm for suffix stripping’, *Program* 14(3): 130-137.

Slapin J. B. and Proksch S.-O. (2008), ‘A Scaling Model for Estimating Time-Series Party Positions from Texts’, *American Journal of Political Science* 52(3): 705–722.

1. The package quanteda in R has been used. [↑](#footnote-ref-1)
2. In mathematical terms, the is given by , where represents the weight,is a term in a document , is the frequency of a specific term and represents the total number of documents (speeches) in a given corpus. [↑](#footnote-ref-2)
3. More specifically, let D be a set (a corpus) of documents { }, and T be the set of distinct terms { } occurring in D. Let each document be represented as a vector of terms ; moreover, let (d, t) be the weighted frequency of the terms t in the document . The resulting vector representation of each text contained in each corpus is then:

|  |  |
| --- | --- |
| = (*tfidf* (d, t1); *tfidf* (d, t2)… *tfidf* (d, tn)) |  |

 [↑](#footnote-ref-3)