The Empirical Determinants of Manifesto Content*

Thomas Däubler
Mannheim Centre for European Social Research
thomas.daeubler@mzes.uni-mannheim.de

Kenneth Benoit
London School of Economics and Trinity College Dublin
kbenoit@lse.ac.uk

June 18, 2013

Abstract

Party manifestos form the largest source of textual data for estimating party policy positions, typically based on methods that assume that longer manifestos with more units of text provide more confident estimates. Despite using them extensively for nearly three decades, however, we know little to nothing about what explains why either the overall length of manifestos or their scope of issue coverage varies so highly across parties, elections, and contexts. Here, we critically test the notion that political context affects overall length and manifesto content. We use multi-level modeling to predict manifesto length and issue scope in a large number of coded party manifestos covering the post-war period. Our findings indicate that manifesto length and the scope of issue coverage can be largely explained by a combination of political variables related to party size, policy orientation, as well as election-specific factors related to political competition and the timing of elections.

Key Words: Party manifestos, Comparative Manifestos Project, policy positions, issue salience, saliency theory.

*Paper prepared for presentation at the Annual Meeting of the European Political Science Association, June 20–22, 2013, Barcelona. This research was supported by the European Research Council grant ERC-2011-StG 283794-QUANTESS. We thank Michael Gallagher and Alex Herzog for extensive comments and suggestions on earlier versions of this paper.

Of all of the forms of text analyzed as data in political science research, election manifestos outstrip other documents by far as sources of information about the policy emphases and positions of political parties. Manifestos lay out the party's official positions on the issues of importance prior to an election, and identify those issues the party acknowledges as salient. For these reasons, manifestos have formed the main source of textual data for both manually coded content analysis research such as the long-standing Manifesto Project (e.g. Klingemann, Volkens, Bara, Budge and McDonald, 2006) as well as numerous attempts to extract policy positions automatically using supervised (Laver, Benoit and Garry, 2003*a*) or unsupervised (Slapin and Proksch, 2008*a*; Monroe and Maeda, 2004) learning methods.

Underlying all uses of manifesto data is the assumption that manifestos containing more text are more informative: that more data provides more confident estimates than less data. Some have modelled this process explicitly (for instance Benoit, Laver and Mikhaylov, 2009), to estimate the variance of manifesto text as a function of its length. Other models, such as Laver, Benoit and Garry (2003*a*)'s "wordscores" or the Poisson scaling model of Slapin and Proksch (2008*a*) incorporate this as a feature of their estimation method. Whether words, sentences, or "quasi-sentences", models of textual data built on party manifestos share the feature of modelling observed counts of text units and using these counts to estimate features of the party's policy stances. In reality, however, very little is known about the data generating process giving rise to counts of textual units in party platforms. These platforms and the textual units covering different issues contained in them vary widely in length, across parties, countries, and time.

In this paper, we provide the first systematic attempt to explain not the proportion of content that parties devote to particular policy issues, but rather the absolute length. What makes some parties devote more space to some issues, or more text to all issues, than other parties or to other issues? While some of this variation comes from idiosyncratic features of manifesto authorship, we show that much of manifesto length can be explained as a function of systematic political factors such as party size, government status, and policy orientation, as well as election-specific factors related to political competition and the timing of elections. These specific political conditions take the form of systematic factors related to party size and governing status, a

party's policy orientation towards an issue, and the positions of other parties on the issues. Our investigation focuses on five "mainstream" issues likely to be in some form by nearly all parties—economic policy, social policy, environmental policy and deregulation policy—as well as one that is highly salient in some contexts yet absent from others: policy embrace of greater or less sovereignty for the institutions of the European Union.

By modelling the length of manifestos as well as the absolute attention devoted to specific political issues, using a count-based model, we offer insight into the nature of textual data, as well as provide a more direct understanding of the textual evidence of party position-taking in election programmes. This understanding should not only shed light on a previously unstudied political process of how policy programmes are articulated, but also improve models that rely on textual data from manifestos by providing an improved understanding of the process generating units of textual data in these documents.

We test our predictions of manifesto length on data from the Comparative Manifestos Project (Budge, Robertson and Hearl, 1987*b*; Budge et al., 2001*b*; Klingemann, Volkens, Bara, Budge and McDonald, 2006) using multilevel regression models. The results confirm that variables of substantial interest do affect manifesto size. The findings suggest that the length of election programmes depends especially on factors related to party status and resources. Parliamentary parties and larger parties publish longer documents than parties not represented in parliament and smaller parties. When parties have less preparation time because of an early election, manifestos are also considerably shorter. Length is also related to party positions on issues, in some cases with a linear and in others with a U-shaped effect.

Why study manifesto length?

The publication of manifestos by political parties before elections is common practice in many countries. The scope of the data collection by the Comparative Manifestos Project (CMP, see Budge, Robertson and Hearl (1987*a*); Budge et al. (2001*a*); Klingemann, Volkens, Bara and Budge (2006)) nicely illustrates this point. Including a recent update (Volkens et al., 2010) the database includes more than 3300 documents from 55 countries, and a further geographic

extension of coverage is on the way. The widespread use of election manifestos in polities across the globe is no surprise. The idea that political parties provide the voters with an outline of their policy plans before elections lies at the very heart of mandate models of representative democracy (e.g. APSA, 1950; Downs, 1957; Powell, 2000; see McDonald et al., 2004: 3-4 for a synthesis). In a nutshell, such models suggest that parties offer (distinct) policy packages to the electorate, voters make their choice on the basis of their preferred package, and the parties seek to implement the proposed policies after the election. In such a model of the representative process manifestos are important documents because "they form the basis for comment in the mass media and provide the cues for questions raised with party candidates at all levels, as well as staple issues for their campaigns" (Budge, 1987, 18). Because manifestos are drawn up for purposes of shaping the frames of its election campaign and setting out its policy positions, parties typically place great care in drafting these texts. For these reasons, a party manifestos is a text that "can be singled out as a uniquely representative and authoritative characterisation of party policy at a given point in time" (Budge, 1987, 18).

Political scientists therefore frequently use election manifestos to infer features of party policy. Most commonly, the interest lies in locating political parties in policy space. Researchers interested in analyzing the development of party positions over time can hardly avoid using election manifestos as data source. While there are other, non-manifesto-based approaches for estimating the policy positions of political actors (Benoit and Laver, 2006, 56-77), the regular publication and the "official" status of manifestos makes them the first choice to measure time-varying party positions. Specific methods for deriving party positions from manifestos vary. In addition to different methods using the human-coded CMP data (Laver and Budge, 1992*a*; Gabel and Huber, 2000; Franzmann and Kaiser, 2006), automated techniques based on word frequencies (Laver, Benoit and Garry, 2003*b*; Slapin and Proksch, 2008*b*) have been introduced more recently. Regardless of the technique, the length of a manifesto has important consequences for the estimation of any policy position from it. Our confidence about data-based estimates is higher when these are based on more information, and textual data is no exception. Inferences from units of text, whether simple proportions or more complicated measurement models, become more certain as the number of text units grows.

Specifically with regard to CMP data, Benoit, Laver and Mikhaylov (2009) describe the stochastic process from unobserved policy position to observed text as governed by a stochastic process where the variance is directly related to the observed text units. Correspondingly, if the observed text is longer, the researcher's inferences about the underlying message become less uncertain. Their model of stochastic text generation, however, treats the determination of textual length as fixed and unexplained; substantively, a political black box: "The total number of text units found in a manifesto appears to be, absent systematic information or prior expectation on this matter, unrelated to any political variable of interest" (Benoit, Laver and Mikhaylov, 2009, 501). Our approach, by contrast, is to model the number of textual units, first in the overall manifesto and then on policy issues, explicitly in order to critically and empirically test whether manifesto length is essentially idiosyncratic or whether it forms an endogenous part of party competition or otherwise arises from systematic political factors.

Differences in Length and Scope of Manifesto Content

The most commonly cited explanation of how parties allocate manifesto space to issues is the "saliency" theory advanced by Robertson (1976) and Budge, Robertson and Hearl (1987*b*), in which parties signal their positions on different policy dimensions. Parties are assumed to differentiate themselves by emphasising issues on which they have the best reputation with voters (Budge, 1994). Because positioning is a matter of emphasis, more emphasis can be equated with more extremity, and therefore the proportion of emphasis of a particular issue provides a direct indication of a party's relative position on that issue (Budge et al., 2001*b*, 82). The predominant focus on the relative proportion of issue-specific content, however, obscures the important question of how and why parties decide to cover any given issue in absolute terms. How many sentences, or pages, are required to outline a position on a particular issue?

[Figure 4 about here]

Before we develop our argument further, it is useful to introduce a stylised model of the manifesto preparation process (Figure 4).[1] To start with, parties do not make up manifestos

[1]This model builds on first hand information provided by political actors closely involved in manifesto creation in the five major Irish parties (Däubler, 2012*a*) and in the five major parties in the German state of Baden-

from scratch. At the beginning of the process, parties have to make an inventory of "concrete" party policy (governmental or parliamentary documents, formal party decisions etc.). This stock of party policy is evaluated and areas requiring updates and additions are identified. What we refer to as "latent" policy, i.e. views and ideas that have hitherto not been written down, may be added, and all the proposals need to be given a coherent framework. A first draft produced in such a way is then undergoing intra-party discussions, and revisions are made. This process may be repeated a number of times. A final draft is then proposed to the adopting body, e.g. a party committee, the party leadership or a party convention. The adopting body may amend the document and/or ask for further changes, and will finally ratify it, which results in the published manifesto.[2]

The simplified outline of the preparation process points out why we expect variation in manifesto length and issue coverage. First, not all parties approach the task of putting together a manifesto from the same starting point. The extent of concrete and latent policy parties have in stock varies. Second, the illustrated creation process requires resources, both in terms of financial means (especially to pay staff) and time. This applies both to general policy development (mainly taking place before manifesto preparation) as well as the work and coordination of the preparation of the document as such. Third, parties may vary in terms of how they use the manifesto, as manifestos can serve different functions (Kavanagh, 1981; Pappi and Seher, 2009; Däubler, 2012*a*). For example, manifestos play an important role in the coalition formation process, when parties need to agree on government policies (e.g. Laver and Budge, 1992*b*; Debus, 2007). The probability to get involved in such negotiations, however, is not the same for all parties. Differences with regard to intended use may also have an impact on how extensively parties cover certain policies.

Wuerttemberg (Däubler, 2012*b*).

[2]This outline of the manifesto creation process leaves the identity of the actors involved at the different stages deliberately open. This is because it may empirically vary considerably between parties, as was the case, for example, in Britain in the 1970s (Kavanagh, 1981) and in Ireland in 2007 (Däubler, 2012*a*). At the same time, there is no larger-scale empirical data on this matter available. See Schumacher, de Vries and Vis (2013) for a related argument about policy-making in parties more generally.

Overall Length

One thing is clear: manifesto length varies widely. In what follows, we report length as measured in the "quasi-sentence" units recorded by the Comparative Manifesto Project.[3] With time, manifesto length has grown, as shown in Figure 1. Using CMP data, it displays the distributions of manifesto size in 19 OECD countries by decade, in logged quasi-sentence counts. It can be seen that median manifesto size has strongly increased since World War II. In the period between 1945 and 1959, the median manifesto had a length of 126 quasi-sentences, whereas the median manifesto from the first decade of the 21st century was 741.5 quasi-sentences long. In other words, the size of a "typical" manifesto is these days almost six times what it used to be around the 1950s.

[Figure 1 about here]

Large differences in manifesto size can also be found between countries. Figure 2 presents box plots of manifesto length by country, for the same set of 19 OECD states, but limited to the time period from 1970 on. The countries in Figure 2 are ordered from top to bottom according to decreasing median manifesto length. The graph shows that there is considerable variation across countries. Median manifesto size ranges from a low of 89 quasi-sentences in Denmark to a high of 2213 in Greece, which means that a "typical" Greek manifesto was almost 25 times the size of its Danish counterpart. It is also clear that there are considerable differences across manifestos within countries. This holds particularly true for Portugal and Ireland. We are curious to learn more about the factors behind this variation over time, across countries, and within countries.

[Figure 2 about here]

The specific context of manifesto preparation, the resources available and devoted to the task, and the intended uses are shaped by a complex set of factors that are hard to model. However, we can consider how some key variables should impact on the overall length of manifestos. *Parties currently represented in parliament* should present longer manifestos, as they

[3] A quasi-sentence is defined as part or all of a natural sentence that states a distinct policy proposition, "an argument which is the verbal expression of one political idea or issue" Volkens (2007). Quasi-sentences are identified by human coders during the unitization stage of coding manifestos, according to explicit rules.

have a larger stock of "concrete" policies at their disposal. From their MPs' work in the plenary and the committees, these parties are familiar with the parliamentary agenda (Green-Pedersen and Mortensen, 2010). Such parties also obtain valuable resources from their parliamentary status, including funds for hiring policy experts. In addition to parliamentary status, *party size* should matter for a number of reasons. First, larger parties usually have more resources at their disposal (more donations, state funding). Second, larger parties represent a larger and thus more heterogenous group of voters. Therefore they need to offer a much broader range of policies in order to appeal to their electorate. Small parties are more likely to be built around a single issue or to cater to the special interests of a small segment of society. Some small parties may not even base a great extent of their appeal on proposing detailed policies, but rather on leader's charisma or on an "anti-establishment" notion. A third reason why larger parties can be expected to have longer manifestos stems from their higher propensity to get involved into government negotiations.

How *government status* affects overall manifesto length is ambiguous. On the one hand, government parties should have a larger stock of existing policies to work from, since actual policy-making is their everyday business. They also enjoy special access to resources and staff, also from the ministerial bureaucracy, that can assist with general policy development and more specifically with manifesto preparation. On the other hand, opposition parties may face a stronger need to present their case and explain their future policy plans.[4]

In addition to these party characteristics, there are other factors that are expected to affect manifesto length. Preparing manifestos requires time, which is scarce in case of *early elections*. This means that less work can be spent on developing the policy material to publication-ready standard. Also, in early elections there may be less need to introduce new policy ideas since some from the previous election remain up-to-date. Figure 2 suggested that there is variation in manifesto length across countries. Factors which could account for such variation include the level of economic development (*GDP per capita*), as e.g. drafting longer manifestos requires more resources. It may also take time for programmatic party competition to develop;

[4]Based on interviews with representatives of Irish parties after the 1997 elections, Garry and Mansergh (1999) also report that opposition parties had more time to work on the manifesto and were less constrained by their past policies.

therefore manifestos may be shorter in *young democracies*. As already seen from Figure 1, manifestos have increased in size over the *years*. Some authors (e.g. Topf, 1994; Kavanagh, 1996) suggested that this development reflects the trend that governments have taken over a larger number and variety of responsibilities. Another mechanism could be that technological development has lowered the marginal costs of the material production and dissemination of longer documents.

Issue-specific coverage

Manifestos are bundled into issue areas or clusters of issues, representing distinct policy dimensions (Robertson, 1976, 61), such as policy on economic redistribution, environmental policy, or foreign affairs. How parties allocate coverage to different issues has been studied in relative terms (e.g. Pogorelis et al., 2005; Rovny, 2012; Dolezal et al., 2013; Wagner, 2012), but little is known about just *how much text* a party deems necessary to convey its position on a particular issue, and what factors shape these decisions.

Indeed, it is very difficult to pry open the black box of manifesto authorship to resolve these issues, because information about the process of drafting programmatic policy statements within parties is not readily available. Some case studies exist (e.g. Däubler, 2012*a*,*b*) but their results are not necessarily generalizable. Our approach is to focus therefore not on anecdotal or first-hand accounts at a micro level, but rather to attempt to find patterns in the length of specific issue coverage at a macro level.

[Figure 3 about here]

Clearly, there is variation in coverage across dimensions, as shown in Figure 3. The graph is based on the sample that will also be analyzed below. Here, we have singled out five specific issues on which to test coverage: the two "main" dimensions of economic and social policy, two relatively single-issue dimensions of the environment and decentralisation, and a fifth, "new" issue of support or opposition to a more centralised European Union. As expected, from these five issues parties tend to write most on economic matters, followed by "social" issues and environmental policy. Less is said about European integration (although the sample

contains only European countries) and decentralisation. However, parties cover one and the same issue to different degrees. This is the case for all issues, but especially pronounced for decentralisation. The observed variation asks for an explanation.

Several possible explanations for the length of specific issue coverage exist, which we view at this stage as speculative theories to be tested. To start with, a baseline model may posit that issue coverage is simply due to idiosyncratic features which at the macro level seem entirely random, such as the writing style of the authors. Qualitative evidence, indeed, suggests that manifestos are drafted by several contributors, with different authors covering different topics (Däubler, 2012*a*). If these authors have different writing styles, this may cause sections to differ in length.

To develop a more interesting explanation, we look to previous work on relative emphasis. Before doing so, let us point out why our approach does not consider relative emphasis as the dependent variable. In our view, while relative emphasis can be seen as a result of the manifesto writing process, it does not itself provide a realistic theoretical model of manifesto writing. A "relative emphasis" model of manifesto writing assumes that parties a priori face a fixed budget for length, constrained by factors such as perceived audience expectation (and attention span), and possibly costs. A party then considers how to strategically allocate the total length budget to different issues.

In contrast, we argue that a more realistic model of manifesto preparation departs from the idea that manifesto sections are produced independent from each other in the first place. A section covering a certain topic is meant to address issues in that policy area, diagnose policy problems and present solutions. The length of text in a certain policy area is therefore initially determined by topic-specific factors and considerations. After putting together the different manifesto sections, a party may compare their relative length and possibly shorten some of them – in case there is a budget constraint. Qualitative evidence from interviews (Däubler, 2012*a*,*b*) suggests that only some parties pay attention to the total length of their manifestos. Our approach is therefore to formulate expectations about absolute length, while allowing for some interrelation across issues.

At this point, we suggest three factors that may explain issue coverage. First, parties might

devote more text to topics that are perceived to be the most important issues of the day. In practice, policy change is often associated with deficiencies in the status quo (Robertson, 1976; Vis and Kersbergen, 2007). Exogenous parameters such as macroeconomic conditions, demographics or basic societal values may have changed and require an adjustment in policies. Decreasing birth rates, for example, pose a threat to social insurance schemes, and parties may think about ways how to alter policies in order to make the number of births rise again. The amount of such "problem pressure" can differ across policy areas and this may explain differences in coverage.

Second, parties have different social constituencies and may represent different societal cleavages. Parties' core constituencies are a key explanatory factor of issue ownership (Petrocik, 1996; Pogorelis et al., 2005). Parties may therefore cover more extensively those topics that are of key interest to their electorate. This may reflect that parties have special expertise in these areas, that they address issues their core voters especially care about or that they strategically emphasize topics on which they are perceived as competent (Green and Hobolt, 2008; Geys, 2012).

A third type of explanation is based on the spatial nature of party competition. A first argument from this group suggests that parties emphasize policy positions that are advantageous to them in electoral terms, and that this is the case for positions at the extremes of a dimension (Wagner, 2012; Rovny, 2012). This allows parties to engage in a strategy of product differentiation (Kitschelt, 1994). A second argument predicts differences across issues on the basis of the seminal argument by Stokes (1963, 373–374). Some issues may be explicitly "positional", in a sense that parties take different positions along *one* ordered dimension. Other issues are better described as valence issues, where all parties agree on the ends, so competition is over the priority given to *different* issues. "Saliency theory" (Robertson, 1976; Budge and Farlie, 1983) takes this view to the extreme, suggesting that "all party programmes endorse the same position, with only minor exceptions" (Budge et al., 2001*b*, p.82). We do not share the view that all issues are valence issues, but rather that some issues are characterized by valence competition while others are not. Such a view is in line with the empirical findings in a recent study of Austrian manifestos by Dolezal et al. (2013).

From the issues we study, we expect that environmental policy tends to be a valence issue. Few parties should explicitly advocate anti-environmental positions, so a pro-environmental position should be reflected in longer coverage of that issue and a less pro-environmental position in shorter coverage. In addition to the party competition argument, we also expect some differences across issues with regard to the other mechanisms. Economic policy is an area where we would expect little variation in importance across countries. The opposite is the case for decentralization, which should be high on the agenda in some countries while of little relevance in others. In terms of the effects of party-level importance, we expect to find them especially for environmental policy, since Green parties attach special importance to this policy field.

Results: Overall Length

In order to test these expectations about the determinants of overall manifesto length, we run a number of multilevel log-normal regression models.[5] The models include random intercepts at the country-level and in addition at either the election-level or party-level.[6] Results are shown in Table 1. The table displays estimates of the coefficients and variance parameters. Exponentiating the coefficients gives values representing the multiplicative change in manifesto length if the independent variable changes by one unit, which is easily amenable to an interpretation of relative change in percent. Table 1 presents the results of four different models. Models 1 and 2 are null models insofar they merely contain the random intercepts, Models 3 and 4 include all independent variables. Models 1 and 3 use an election-level random intercept at level two, Models 2 and 4 a party-level random intercept.

[Table 1 about here]

By means of calculating the intra-class correlation (ICC) (Gelman and Hill, 2007, :448), the null models allow to examine how much variation in manifesto length can be found at the different levels. Model 1 suggests that approximately 39% of variation in manifesto length occurs

[5]The data represent counts, but the mean of the dependent variable in the sample (635 quasi-sentences) is large enough to justify neglecting the discreteness aspect of the data.

[6]The models were estimated by maximum-likelihood using STATA's `xtmixed` command.

at the country-level ($ICC = .535/(.535 + .316 + .536) \approx .39$), whereas the respective figure at the election level is merely 23%. Put differently, election manifestos from a single country tend to be more similar in length to each other than are documents from one and the same election. Model 2 shows that there is less clustering at the party-level. In this specification, 37% of variance can be found at the country-level, but only 15% at the party-level.

The results from Models 3 and 4 are in line with the expectations referring to party status. In Model 3 the coefficient for parliamentary status is .262. Since $exp(.262) \approx 1.30$, this means that parliamentary parties produce manifestos which are 30% longer than those by parties not represented in parliament. The respective coefficient in Model 4 is almost of the same size. There is also clear support for the argument that larger parties publish longer manifestos. For instance, an increase in seat share of ten percentage points (i.e. by 0.1 on the scale of the variable) results in a 9.9% (Model 3) respectively 13.1% (Model 4) increase in manifesto length. Above, it was argued that it is theoretically ambiguous if outgoing government parties publish longer or shorter manifestos than opposition parties. The respective coefficient is positive (in both models suggesting a 7% increase in length that goes along with government status), but remains statistically insignificant.

Very clearly, manifestos issued for early elections are shorter than their counterparts from "regular" elections. Manifestos from early elections are shorter by about 23.7%, which suggests that the time at hand for preparing the document is an important factor for explaining its length. Also, in younger democracies manifestos are considerably shorter. The estimate from Model 3 suggests that during the first ten years after transition, manifestos have a relative length of only 68% of their counterparts in later years. Election manifestos tend to be longer in wealthier countries, although the estimate is only statistically significant in Model 4. According to that coefficient, a difference equal to the sample range (45.55 between the minimum in Finland, 1951 and the US, 2008) is associated with an increase in manifesto length of 89%. Considerable temporal variation remains despite including economic affluence and the other variables. An additional year later in terms of the election date is associated with a 1.6% (since $exp(.016) \approx 1.016$) increase in length (based on Model 3).

Models 3 and 4 include one additional variable that controls for the relative verbosity of the language the manifesto is written in (see Appendix 1). The findings for the verbosity variable cast doubt on the idea that the method of parsing manifestos into quasi-sentences removes all language effects. Manifestos written in "wordier" languages tend to be divided up into more quasi-sentences. A one-percent increase in verbosity (relative to English) is associated with a 1.7% increase in the number of quasi-sentences. This means that on average manifestos written in the wordiest language in the sample (Spanish, 20% wordier than English) yield about 2.1 times the number of quasi-sentences than those written in the least wordy language in the sample (Finnish, 23% less verbose than English).

The two null models (1 and 2) are useful also for assessing how well the included covariates can explain variation in manifesto length at the various levels. It can be seen that the inclusion of the covariates reduces the error variance at the country level by approximately 22%/21% (from .535 to .415/from .511 to .404). Respective values at the election-level are 57% (Models 1 and 3, from .316 to .136) and at the party-level 27% (Models 2 and 4, from .203 to .149). The covariates as a whole therefore indeed contribute to an explanation of variation in manifesto length. This finding clearly demonstrates that document length is not entirely idiosyncratic.

Results: Explaining issue coverage

Our interest lies in analyzing how the absolute coverage of a specific policy area in a manifesto is influenced by the authoring party's position on the associated dimension and the importance the party attaches to this dimension. To that end, we treat position and importance as given and fixed at the time of manifesto writing and take respective data from the expert surveys by Laver and Hunt (1992) and Benoit and Laver (2006). Expert survey information is matched to the manifesto from the election closest to the date of the survey (i.e. the election may have taken place before or after the survey). We do not consider any manifestos from elections that fall outside of a time window ranging from three years before to three years after the survey. In addition, the current version of the dimension-specific analysis includes only cases based on manifestos that were also part of the sample considered for the analysis of overall length in

the first part of the paper. Four variables are based on the expert survey data. *Party position* is simply the mean position of the party on the respective dimension, which is measured on a 1-20 scale. To test for nonlinear effects of this variable, we also include it in its squared variant. *Importance in country* reflects the mean of the importance each party in that country attaches to the dimension, weighted by the vote-share the party received in the previous election. The underlying scale also has a range from 1-20, with higher scores reflecting higher levels of importance. *Party importance* of a dimension is defined relative to the importance a party attaches to other dimensions, as the deviation of the party importance score from the median importance score across all other dimensions included in the expert survey.

As outlined above, we analyze coverage in five dimensions, as the absolute number of quasi-sentences in the following CMP categories: Economy (all categories in the 400s except the one subsumed under Environment), Social (Traditional Morality Positive [603], Traditional Morality Negative [604], Law and Order Positive [605]), Environment (Environmental Protection Positive [501], Anti-Growth Economy Positive [416]), Decentralisation (Decentralisation Positive [301] and Centralisation Positive [302]) and Europe (European Integration Positive [108] and European Integration Negative [110]). From the expert survey, we use the following dimensions: Taxes vs. Spending for Economy; Social Liberalism for Social; the EU dimension that best matches with European Integration (i.e. EU Strengthening for Ireland, EU Larger/Stronger for France and EU Authority for the other countries - these are only available in Benoit and Laver (2006) but not in Laver and Hunt (1992)); Environment; Decentralization. In addition to the dimension-specific variables based on the expert survey, the models will include the variables used to explain overall length in the first part of this paper (with the exception of the time trend, since there is little variation in that variable in the reduced sample).

When analyzing dimension-specific coverage, we have – unlike in the analysis of overall length – multiple observations per manifesto and count outcomes with smaller means. To address these issues, we employ a multilevel overdispersed Poisson model that includes random intercepts for country/election-dimension combinations (country and election are equivalent here since there is only one election per country) and party-dimension combinations. Our Bayesian model allows for correlation between the random intercepts across the dimensions

at each level. Inference is via Markov chain Monte Carlo simulations implemented in R (R Core Team, 2012) using the package MCMCglmm (Hadfield, 2010). More specifically, if y_{jk} is the number of quasi-sentences in manifesto $j = 1,...,J$ (from country-election c) covering dimension $k = 1,...,K$ our model is

$$Pr(y|\mu_{jk}) = \frac{e^{-\mu_{jk}}\mu_{jk}^{y}}{y!}$$

$$log(\mu_{jk}) = \mathbf{X}\beta + \gamma_{c[j]k} + \varepsilon_{jk}$$

$$\gamma_{cj} \sim \mathbf{N}_K(\mathbf{0}, G)$$

$$\varepsilon_{jk} \sim \mathbf{N}_K(\mathbf{0}, R)$$

As priors, we choose a multivariate-normal distribution with large variances on the diagonal, for the variance-covariance matrices we specify inverse-Wishart distributions:

$$\beta \sim \mathbf{N}(\mathbf{0}, I * 10^8)$$

$$G \sim Wishart^{-1}(I, K+1)$$

$$R \sim Wishart^{-1}(I, K+1)$$

[Table 2 about here]

The results are presented separately for the Laver-Hunt and the Benoit-Laver data (spaced approximately 10 years apart). Table 2 shows the results of the four issues (excluding the EU dimension) matched with the Laver-Hunt expert surveys. As with the aggregate results, younger democracies were more likely to have shorter manifestos, across all issues, and those written in more verbose languages were likely to have slightly longer texts. There are no clear-cut effects of the importance variables. The means of the respective posterior distributions are mostly positive, but we cannot be very certain about the direction of the effects. On the other hand, for the positional variables we find interesting relationships. As expected, we observe differences across issues. Both for social issues and environmental policy, we find a U-shaped relationship between position and absolute coverage. That is, parties at the extremes cover these topics more extensively. The minima of the U-shaped functions are at 7.4 for social

issues and 12.4 for environmental policy on the original 1-20 scales (when calculating this from the mean-deviated data). Decentralization, on the other hand, shows a different pattern. Here, parties less in favour of decentralization of political control devote fewer sentences to that issue. Interestingly, the independent variables included in the model contribute little to explaining variation in coverage of economic affairs, although the baseline constant of 4.97 indicates that this issue received on average far more coverage than did the other issues we examined.

[Table 3 about here]

In Table 3, we present similar results from the models in the 2000s using the Benoit-Laver data. The general pattern of findings is fairly similar to the Laver-Hunt data. The independent variables seem not to affect how extensively parties cover economic policy, and we find no clear effects of any of the importance variables. For social issues and environmental policy we do find positional effects. Unlike in the first sample, there is evidence for a simple linear effect, with more conservative parties devoting more statements to social issues and more pro-environment parties talking as expected more about the environment. With regard to European integration, positions at the extremes are covered more extensively (the minimum of the U-shaped function liest at 11.25 on the original 1-20 scale). For decentralization, no clear-cut effects appear in this sample.

Finally, in both samples we observe considerable unobserved heterogeneity both at the country-level and party-level. What is most interesting here from a theoretical perspective is the fact that all pairs of random intercepts are correlated positively. This evidence does not support the idea that there is a budget constraint in terms of total length. If this was the case, we would expect negative correlations suggesting that different issues compete for the same space on the manifesto paper. This seems not to be the case.

Conclusions

Election manifestos play a special role in the representative process and are one of the most important data sources for inferring party positions. While the methodological literature has

pointed out the relevance of manifesto length for the uncertainty of inferences about party policy, it has treated this variable as an idiosyncratic feature. This paper, however, shows that there is systematic variation both in terms of manifesto length at a macro level, and of the length that parties devote to specific issues.

In other words, the amount of text we observe in the form of party programmatic statements is not only idiosyncratic, but also explained by the political contexts in which parties operate. More specifically, manifesto length appears to be driven especially by parliamentary status, party size and available preparation time. The determinants of the absolute coverage of specific issues differ according to issue: the mainstream issue of economic policy appears to have only a weak relationship with a party's position or perceived relative importance of the issue, while more specific policy areas such as social issues, environmental protection, decentralization and policy on the European Union appear to tie text length to a party's position on those issues.

The findings of this study have a number of important implications. The general finding that manifesto length is partly a consequence of the political context corroborates the claim to pay more attention to the uncertainty of party position estimates (Benoit, Laver and Mikhaylov, 2009). Length is not a completely idiosyncratic text feature. We have more information about party policy in some cases than in others, because of systematic variation between the contexts in which this information was provided. Differences in uncertainty associated with the inferences about party policy should therefore not be neglected. In addition, our findings have a number of important implications for the way we tie issue emphasis to positions, and these will differ according to the nature of the issue. We are still thinking this through, and welcome suggestions.

References

APSA. 1950. *Toward a more responsible two-party model*. Washington: American Political Science Association.

Benoit, Kenneth and Michael Laver. 2006. *Party Policy in Modern Democracies*. London: Routledge.

Benoit, Kenneth, Michael Laver and Slava Mikhaylov. 2009. "Treating Words as Data with Error: Uncertainty in Text Statements of Policy Positions." *American Journal of Political Science* 53(2):495–513.

Bergman, Torbjørn, Wolfgang C. Müller, Kaare Strøm and Markus Blomgren. 2006. Democratic delegation and accountability: cross-national patterns. In *Delegation and accountability in parliamentary democracies*, ed. Kaare Strøm, Wolfgang C. Müller and Torbjørn Bergman. Oxford: Oxford University Press pp. 109–220.

Budge, Ian. 1987. The internal analysis of election programmes. In *Ideology, strategy and party change. Spatial analysis of post-war election programmes in 19 democracies*, ed. Ian Budge, David Robertson and Derek Hearl. Cambridge: Cambridge University Press pp. 15–38.

Budge, Ian. 1994. "A new spatial theory of party competition: Uncertainty, ideology and policy equilibria viewed comparatively and temporally." *British Journal of Political Science* 24(4):443–467.

Budge, Ian, David Robertson and Derek Hearl. 1987*a*. *Ideology, strategy, and party change in 19 democracies*. Cambridge: Cambridge University Press.

Budge, Ian, David Robertson and Derek Hearl. 1987*b*. *Ideology, Strategy and Party Change: Spatial Analyses of Post-War Election Programmes in 19 Democracies*. Cambridge: Cambridge University Press.

Budge, Ian and Dennis Farlie. 1983. *Explaining and predicting elections: issue effects and party strategies in twenty-three democracies*. London: George Allan & Unwin.

Budge, Ian, Hans-Dieter Klingemann, Andrea Volkens, Judith Bara and Eric Tanenbaum. 2001*a*. *Mapping Policy Preferences - Estimates for Parties, Electors and Governments 1945-1998*. Oxford University Press.

Budge, Ian, Hans-Dieter Klingemann, Andrea Volkens, Judith Bara and Eric Tanenbaum. 2001*b*. *Mapping Policy Preferences: Estimates for Parties, Electors, and Governments 1945–1998*. Oxford: Oxford University Press.

Cusack, Thomas R. and Lutz Engelhardt. 2002. *The PGL file collection: File structures and procedures*. Berlin: Wissenschaftszentrum Berlin für Sozialforschung.

Däubler, Thomas. 2012*a*. "The preparation and use of election manifestos: Learning from the Irish case." *Irish Political Studies* 27(1):51–70.

Däubler, Thomas. 2012*b*. "Wie entstehen Wahlprogramme? Eine Untersuchung zur Landtagswahl in Baden-Wï¿½rttemberg 2006." *Zeitschrift für Politikwissenschaft* 22(3).

Debus, Marc. 2007. *Pre-Electoral Alliances, Coalition Rejections, and Multiparty Governments*. Baden-Baden: Nomos.

Dolezal, Martin, Laurenz Ennser-Jedenastik, Wolfgang C Müller and Anna Katharina Winkler. 2013. "How parties compete for votes: A test of saliency theory." *European Journal of Political Research* .

Downs, Anthony. 1957. *An Economic Theory of Democracy*. New York: Harper and Row.

Franzmann, Simon and Andrï¿½ Kaiser. 2006. "Locating political parties in policy space. A reanalysis of party manifesto data." *Party Politics* 12(2):163–188.

Gabel, Matthew J. and John D. Huber. 2000. "Putting Parties in Their Place: Inferring Party Left-Right Ideological Positions from Party Manifestos Data." *American Journal of Political Science* 44(1):94–103.

Garry, John and Lucy E. Mansergh. 1999. Irish party manifestos. In *How Ireland voted 1997*, ed. Michael Marsh and Paul Mitchell. Boulder: Westview pp. 82–106.

Gelman, Andrew and Jennifer Hill. 2007. *Data analysis using regression and multilevel/hierarchical models*. Cambridge: Cambridge University Press.

Gemenis, Kostas. 2011. "Garbage in, garbage out? Document selection bias in the Comparative Manifestos Project." Paper presented at the 2011 ECPR Joint Sessions, St.Gallen, Switzerland.

Geys, Benny. 2012. "Success and failure in electoral competition: Selective issue emphasis under incomplete issue ownership." *Electoral Studies* 31(2):406–412.

Green, Jane and Sara B. Hobolt. 2008. "Owning the issue agenda: Party strategies and vote choices in British elections." *Electoral Studies* 27(3):460–476.

Green-Pedersen, Christoffer and Peter B. Mortensen. 2010. "Who sets the agenda and who responds to it in the Danish parliament? A new model of issue competition and agenda-setting." *European Journal of Political Research* 49(2):257–281.

Hadfield, Jarrod D. 2010. "MCMC methods for multi-response generalized linear mixed models: the MCMCglmm R package." *Journal of Statistical Software* 33(2):1–22.

Hansen, Martin. 2008. "Back to the archives? A critique of the Danish part of the manifesto dataset." *Scandinavian Political Studies* 31(2):201–216.

Heston, Alan, Robert Summers and Bettina Aten. 2011. "Penn World Table Version 7.0, May 2011.".

Huber, John D. and Charles R. Shipan. 2002. *Deliberate discretion?: the institutional foundations of bureaucratic autonomy*. Cambridge: Cambridge University Press.

Kavanagh, Dennis. 1981. "The politics of manifestos." *Parliamentary Affairs* 34(1):7–27.

Kavanagh, Dennis. 1996. *British politics: continuities and change*. Oxford: Oxford University Press.

Kitschelt, Herbert. 1994. *The transformation of European social democracy*. Cambridge: Cambridge University Press.

Klingemann, Hans-Dieter, Andrea Volkens, Judith Bara and Ian Budge. 2006. *Mapping Policy Preferences II. Estimates for Parties, Electors and Governments in Central and Eastern Europe, European Union and OECD 1990-2003*. Oxford: Oxford University Press.

Klingemann, Hans-Dieter, Andrea Volkens, Judith Bara, Ian Budge and Michael McDonald. 2006. *Mapping Policy Preferences II: Estimates for Parties, Electors, and Governments in Eastern Europe, European Union and OECD 1990-2003*. Oxford: Oxford University Press.

Laver, Michael and Ian Budge. 1992*a*. Measuring Policy Distances and Modelling Coalition Formation. In *Party policy and government coalitions*, ed. Michael Laver and Ian Budge. Houndmills/Basingstoke: Macmillan pp. 15–40.

Laver, Michael J. and Ian Budge. 1992*b*. *Party, Policy and Government Coalitions*. St. Martin's Press.

Laver, Michael J. and W. Ben Hunt. 1992. *Policy and Party Competition*. New York: Routledge.

Laver, Michael, Kenneth Benoit and John Garry. 2003*a*. "Estimating the policy positions of political actors using words as data." *American Political Science Review* 97(2):311–331.

Laver, Michael, Kenneth Benoit and John Garry. 2003*b*. "Extracting Policy Positions from Political Texts Using Words as Data." *American Political Science Review* 97(2):311–331.

McDonald, Michael D., Silvia M. Mendes and Ian Budge. 2004. "What are elections for? Conferring the median mandate." *British Journal of Political Science* 34(1):1–26.

Monroe, Burt and Ko Maeda. 2004. "Talk's cheap: Text-based estimation of rhetorical ideal-points." Working paper: Michigan State University.

Nohlen, Dieter and Philip Stï¿½ver, eds. 2010. *Elections in Europe. A data handbook*. Baden-Baden: Nomos.

Pappi, Franz Urban and Nicole Michaela Seher. 2009. "Party Election Programmes, Signalling Policies and Salience of Specific Policy Domains: The German Parties from 1990 to 2005." *German Politics* 18(3):403–425.

Petrocik, John R. 1996. "Issue ownership in presidential electins, with a 1980 case study." *American Journal of Political Science* 40(3):825–850.

Pogorelis, Robertas, Bart Maddens, Wilfried Swenden and Elodie Fabre. 2005. "Issue salience in regional and national party manifestos in the UK." *West European Politics* 28(5):992–1014.

Powell, G. Bingham Jr. 2000. *Elections as instruments of democracy*. New Haven: Yale University Press.

R Core Team. 2012. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. ISBN 3-900051-07-0.
URL: *http://www.R-project.org/*

Robertson, David. 1976. *A Theory of Party Competition*. London and New York: Wiley.

Rovny, Jan. 2012. "Who emphasizes and who blurs? Party strategy in multidimensional competition." *European Union Politics* 13(2):269–292.

Schumacher, Gijs, Catherine de Vries and Barbara Vis. 2013. "Why Do Parties Change Position? Party Organization and Environmental Incentives." *Journal of Politics* .

Slapin, J. B. and S.-O. Proksch. 2008*a*. "A scaling model for estimating time-series party positions from texts." *American Journal of Political Science* 52(3):705–722.

Slapin, Jonathan B. and Sven-Oliver Proksch. 2008*b*. "A Scaling Model for Estimating Time-Series Party Positions from Texts." *American Journal of Political Science* 52(3):705–722.

Stokes, Donald E. 1963. "Spatial models of party competition." *The American Political Science Review* 57(2):368–377.

Strøm, Kaare and Stephen M. Swindle. 2002. "Strategic parliamentary dissolution." *American Political Science Review* 96(3):575–591.

Strøm, Kaare, Wolfgang C. Müller and Torbjørn Bergman, eds. 2006. *Delegation and accountability in parliamentary democracies*. Oxford: Oxford University Press.

Topf, Richard. 1994. Party manifestos. In *Labour's last chance? The 1992 election and beyond*, ed. Anthony Heath, Roger Jowell, John Curtice and Bridget with Taylor. Aldershot: Dartmouth pp. 149–171.

Vis, Barbara and Kees Van Kersbergen. 2007. "Why and how do political actors pursue risky reforms?" *Journal of Theoretical Politics* 19(2):153–172.

Volkens, Andrea. 2007. "Strengths and weaknesses of approaches to measuring policy positions of parties." *Electoral Studies* 26(1):108–120.

Volkens, Andrea, Onawa Lacewell, Sven Regel, Henrike Schultze and Annika Werner. 2010. "The Manifesto Data Collection. Manifesto Project (MRG/CMP/MARPOR)." Wissenschaftszentrum Berlin für Sozialforschung (WZB).

Wagner, Markus. 2012. "When do parties emphasise extreme positions? How strategic incentives for policy differentiation influence issue importance." *European Journal of Political Research* 51(1):64–88.

Woldendorp, Jaap, Hans Keman and Ian Budge. 2000. *Party government in 48 democracies (1945-1998) : composition, duration, personnel*. Dordrecht: Kluwer Academic.

Table 1: Multilevel log-normal models of manifesto length

	Null, el. RE	Null, p. RE	Main, el. RE	Main, p. RE
	(1)	(2)	(3)	(4)
Parl. party (0/1)			.262**	.273**
			(.084)	(.090)
Prev. seat share			.945**	1.230**
			(.168)	(.244)
Gov. party (0/1)			.065	.072
			(.054)	(.057)
Early election (0/1)			-.277**	-.274**
			(.081)	(.055)
Young democracy			-.389*	-.354**
			(.153)	(.105)
GDP			.010	.014*
			(.009)	(.007)
Year			.016*	.013**
			(.007)	(.005)
Verbosity			.017**	.017*
			(.006)	(.007)
Constant	5.872**	5.862**	5.675**	5.632**
	(.167)	(.164)	(.166)	(.166)
σ^2_c	.535	.511*	.415*	.404**
	(.180)	(.172)	(.142)	(.140)
σ^2_{el}	.316**		.136**	
	(.042)		(.024)	
σ^2_p		.203**		.149**
		(.042)		(.033)
σ^2_e	.536**	.673**	.499**	.517**
	(.024)	(.030)	(.023)	(.023)
Log likelihood	-1528.6	-1596.3	-1428.3	-1434.8
N countries	21	21	21	21
N elections	249		249	
N parties		218		218
N total	1211	1211	1211	1211

Note: Cell entries are coefficients, standard errors in parentheses. $*p < .05, **p < .01$. Mixed effects models estimated using maximum likelihood; continuous variables were mean-deviated.

Table 2: Multilevel Poisson models of dimension coverage: Late 1980s/early 1990s

	All/Economy			**Social**			**Environment**			**Decentralisation**		
	2.5%	Mean	97.5%	2.5%	Mean	97.5%	2.5%	Mean	97.5%	2.5%	Mean	97.5%
Parl. party (0/1)	-1.95	-0.06	1.99									
Prev. seat share	-0.09	1.58	3.17									
Early election (0/1)	-1.64	-0.64	0.36									
Young democracy (0/1)	-4.82	**-2.96**	-0.92									
GDP	-0.00	-0.00	0.00									
Verbosity	0.03	**0.07**	0.11									
Gov. party (0/1)	-0.83	-0.25	0.30	-0.95	-0.15	0.67	-0.66	-0.11	0.56	-0.74	0.13	0.95
Position	-0.04	0.01	0.07	0.01	**0.08**	0.17	-0.17	-0.06	0.04	-0.27	**-0.14**	-0.03
Position squared	-0.01	-0.00	0.01	0.00	**0.02**	0.03	0.00	**0.02**	0.03	-0.02	-0.00	0.02
Imp. in country	-0.45	-0.14	0.16	-0.38	-0.02	0.30	-0.23	0.02	0.28	-0.22	0.03	0.28
Imp. to party	-0.08	0.03	0.14	-0.14	0.04	0.20	-0.07	0.06	0.22	-0.08	0.08	0.23
Constant	2.85	**4.97**	6.99	0.37	**2.51**	4.68	1.49	**3.47**	5.61	0.21	**2.26**	4.43
σ^2_c	0.29	0.83	1.61	0.34	1.09	2.18	0.30	0.92	1.76	0.33	1.21	2.39
σ^2_p	0.51	0.82	1.18	0.93	1.63	2.51	0.50	0.88	1.30	0.98	1.83	2.98
Correlation of REs (country level/party level)												
Economy		1										
Social		.33/.54			1							
Environment		.39/.62			.37/.43			1				
Decentralization		.30/.49			.32/.49			.21/.51			1	
N countries/elections		16										
N parties/manifestos		69										
N total/parties*dimensions		276										

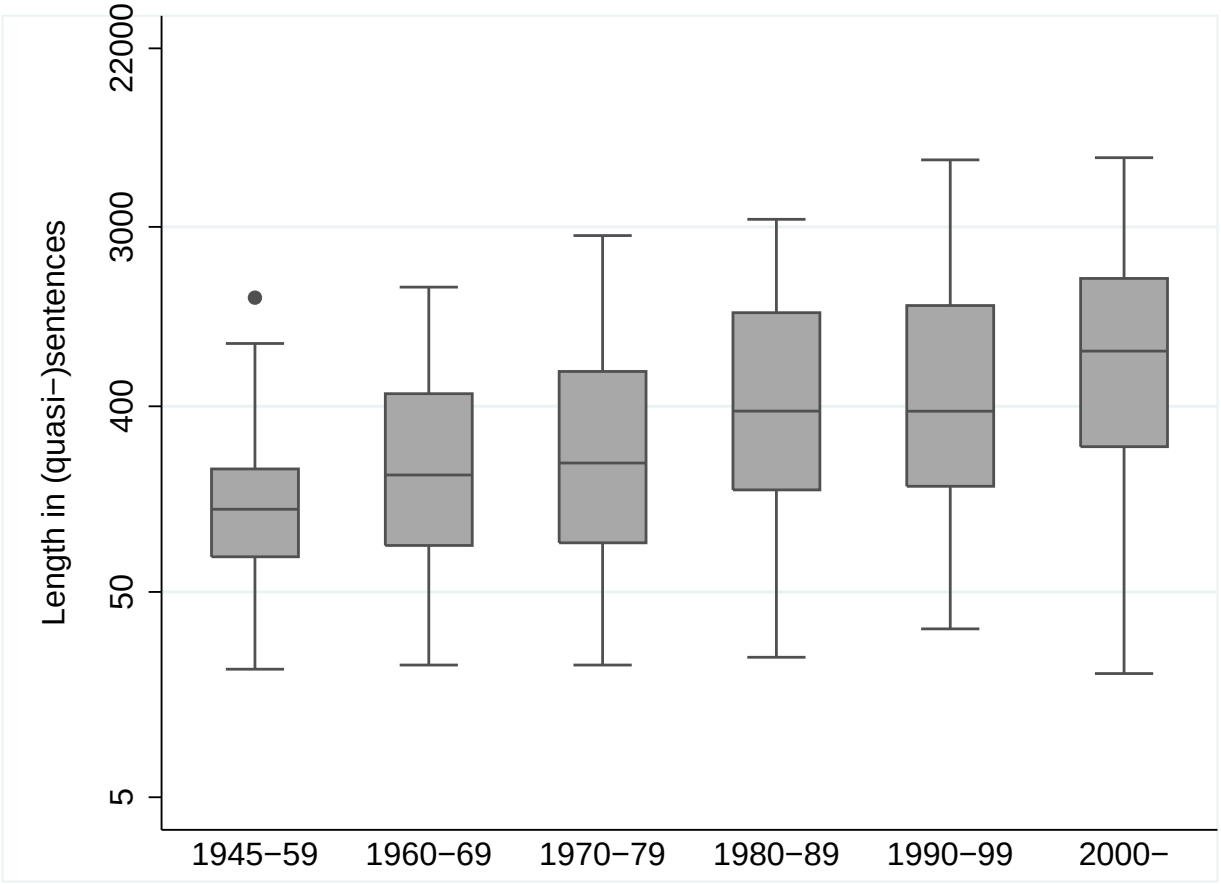
Note: Cell entries are coefficients. Bayesian model as described in text. Continuous variables were mean-deviated.

Table 3: Multilevel Poisson models of dimension coverage: Early 2000s

	All/Economy			**Social**			**Environment**			**Decentralisation**			**Europe**		
	2.5%	Mean	97.5%	2.5%	Mean	97.5%	2.5%	Mean	97.5%	2.5%	Mean	97.5%	2.5%	Mean	97.5%
Parl. party (0/1)		(all obs. one)													
Prev. seat share	-0.02	2.00	3.85												
Early election (0/1)	-0.61	0.57	1.70												
Young democracy (0/1)		(all obs. zero)													
GDP	-0.00	0.00	0.00												
Verbosity	0.00	**0.04**	0.09												
Gov. party (0/1)	-0.11	0.48	1.09	-0.17	0.45	1.11	-0.30	0.41	1.19	-0.59	0.33	1.16	-0.13	0.48	1.14
Position	-0.01	0.03	0.07	0.02	**0.06**	0.10	-0.22	**-0.13**	-0.04	-0.23	-0.11	0.02	-0.07	-0.02	0.03
Position squared	-0.01	0.00	0.02	-0.01	0.00	0.01	-0.01	0.00	0.02	-0.02	0.02	0.06	0.01	**0.02**	0.04
Imp. in country	-0.76	0.05	0.82	-0.50	0.06	0.62	-0.50	0.02	0.59	-0.19	0.36	0.87	-0.54	-0.07	0.44
Imp. to party	-0.10	0.01	0.12	-0.02	0.07	0.16	-0.13	0.04	0.21	-0.13	0.11	0.35	-0.23	-0.05	0.12
Constant	3.49	**4.38**	5.23	2.27	**3.24**	4.13	2.07	**3.04**	3.98	0.07	**1.32**	2.47	1.02	**1.86**	2.76
σ^2_c	0.52	1.72	3.47	0.32	1.08	2.24	0.57	1.90	3.76	0.74	2.74	5.61	0.44	1.44	2.94
σ^2_p	0.57	0.91	1.28	0.66	1.04	1.51	0.92	1.47	2.12	1.11	2.00	3.10	0.53	0.93	1.37
Correlation of REs (country level/party level)															
Economy		1													
Social		.37/.66			1										
Environment		.58/.62			.31/.56			1							
Decentralization		.60/.57			.35/.45			.61/.51			1				
Europe		.50/.54			.32/.40			.45/.53			.42/.46			1	
N countries/elections		12													
N parties/manifestos		73													
N total/parties*dimensions		365													

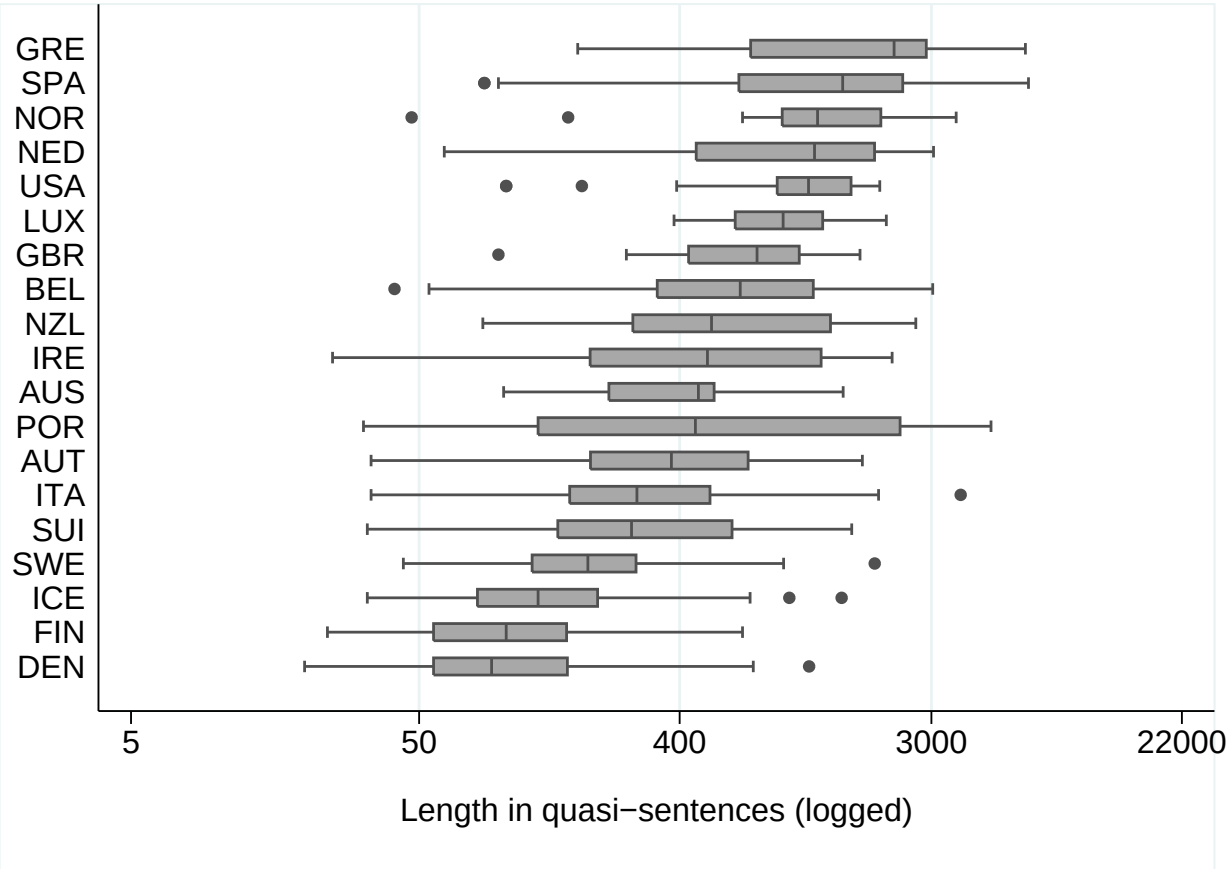
Note: Cell entries are coefficients. Bayesian model as described in text. Continuous variables were mean-deviated.

Figure 1: Development of manifesto length over time



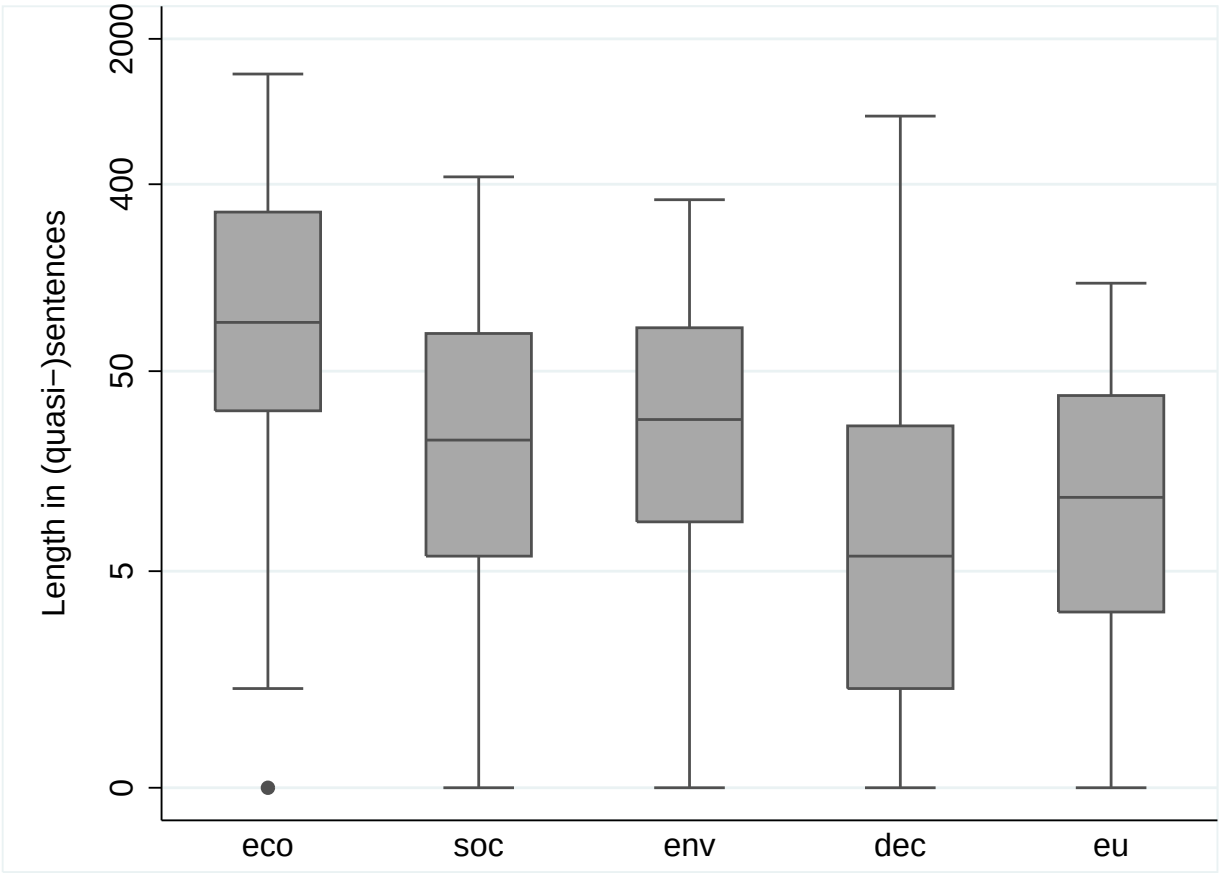
Note: Data are logged, labels on y-axis refer to natural scale. The figure refers to those 19 OECD countries which are part of the sample below *and* where manifestos were consistently parsed into quasi-sentences by the CMP. All cases which do not represent proper manifestos are excluded (see Data section for details).

Figure 2: Manifesto length by country, 1970-



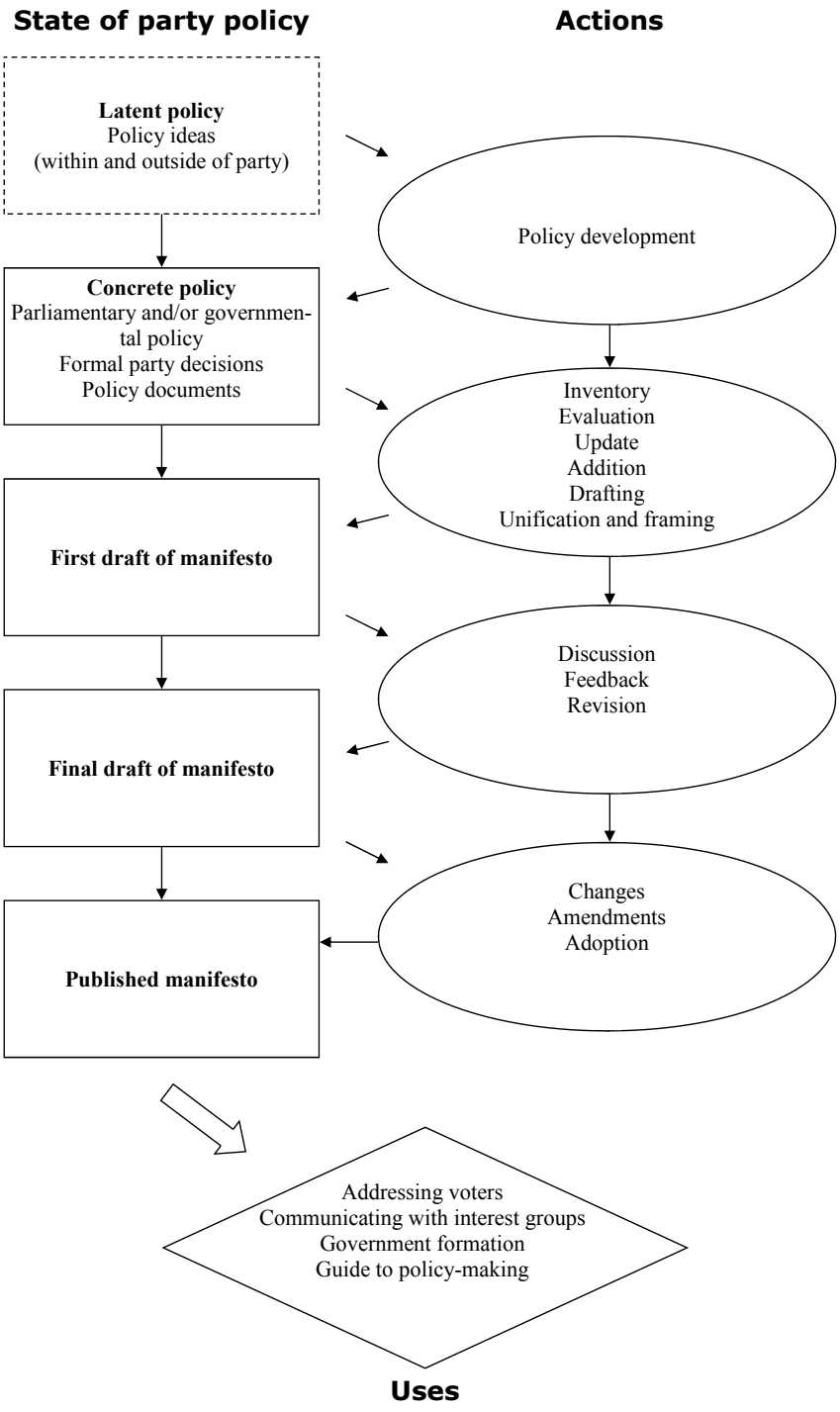
Note: Data are logged, labels on x-axis refer to natural scale. See also note to Figure 1.

Figure 3: Issue coverage in five dimensions



Note: Data are logged (after adding .5), labels on y-axis refer to natural scale. Sample includes the cases analyzed in the issue-specific regressions below (both expert surveys pooled). See also note to Figure 1.

Figure 4: Stylized representation of manifesto creation and uses



Appendix 1: Data for analysis of total length

To analyze overall manifesto length we rely on data from the Comparative Manifestos Project (CMP, see Budge, Robertson and Hearl (1987*a*); Budge et al. (2001*a*); Klingemann, Volkens, Bara and Budge (2006); Volkens et al. (2010)). Not only are these data very frequently used to measure parties' policy positions and issue emphasis. They also appear to be especially well-suited for examining the length of election manifestos, since the CMP uses the "quasi-sentence" as unit of coding, which is "defined as an argument - i.e. the verbal expression of one political idea or issue" (Budge et al., 2001*a*, :217-218). Compared to text length in words, this measure has the advantage that it should not depend on variation in the verbosity of languages (see Huber and Shipan, 2002, :178); if this actually holds true will be tested below. The unitization into quasi-sentences also should alleviate bias arising from the fact that some manifesto writers may use more words than others to express one and the same thing.

The party manifestos whose total length is analyzed are a subset of the CMP dataset and include 1211 post-1945 election manifestos issued by 218 different parties in 249 different elections in 21 OECD countries. On the time dimension, coverage varies across countries but in the majority of cases starts in the 1950s and extends to the 2000s.[7] This sample is determined by four factors. First, we removed observations from the CMP that represent estimates or are based on substitute documents rather than "proper" manifestos. The decision was made on the basis of our own coding of the information provided in Appendix 5 of Klingemann, Volkens, Bara and Budge (2006) and the respective updated document by the CMP group.[8] Also removed are cases which Hansen (2008) and Gemenis (2011) identify as not being based on proper manifestos. Second, not all documents covered by the CMP dataset were coded into quasi-sentences. The codings of many French documents and of a high number of Canadian manifestos (from 1984 on) are based on entire paragraphs, and the unitization scheme used for all German documents differs from the standard quasi-sentence approach for reasons related to the history of the project (Klingemann et al. 2006: Appendix 5, personal communication with Andrea Volkens from the CMP group). All cases that were not parsed into quasi-sentences are therefore excluded from the analysis. Third, the availability of information with regard to the independent variables described in the following paragraphs puts limits on the extent of the analysis sample.

The independent variables are measured as follows. Party status is measured with an indicator variable for representation in the previous parliament, with the proportion of parliamenty seats held prior to the election and with a dummy variable which equals one if a party was part of the previous government. This information is taken from Cusack and Engelhardt (2002) and was updated using Nohlen and Stï¿½ver (2010) and online sources.[9]

[7] More specifically, the country-elections covered are: AUS 1990, 1996-1998; AUT 1953, 1959-1979, 1986-2002; BEL 1954-1965, 1971-2003; CAN 1962-1963, 1968; DEN 1953(Apr)-1981; 1987-2007; ESP 1977-2008, FIN 1951-2003, FRA 1986, 1993-2002; GRE 1989-2000; ICE 1953-1999; IRE 1965, 1977-1981, 1982(Nov)-2007; ITA 1958, 1976-2006; LUX 1964, 1974-1999; NED 1952-2003; NOR 1953-1985, 1993-2001; NZL 1957, 1963-2002; POR 1979-2009; SUI 1951-1959, 1967-2003; SWE 1952-2010; UK 1951-2010; USA 1952-2008.

[8] This measure of document type is likely to contain some error, since the information on the documents given is not in an explicitly explained or consistent format and appears to mix up information on document type and document source as e.g. in "Homepage of party".

[9] Coding of these variables is complicated by the fact that parties split and merge. In case of electoral alliances and mergers of whole parties, parliamentary representation and government participation are coded one if at least one of the original parties was in parliament or government, and the seatshares are added. The three variables were set to zero in case of all party splits or mergers of split parties, since in most cases no information that would allow for a more appropriate coding is available. Since there is no full continuity in party organization in these cases, this approach should not be problematic.

Table 4: Descriptive statistics for sample

Variable	Min	Mean	Median	Max	N
Manifesto length	20	634.77	284	6471	1211
Manifesto length (logged)	3.00	5.74	5.65	8.78	1211
Parliamentary party	0	0.91	1	1	1211
Previous seat share	0	0.18	0.12	0.78	1211
Government party	0	0.35	0	1	1211
Early election	0	0.32	0	1	1211
Young democracy	0	0.07	0	1	1211
GDP	0.91	14.19	12.33	46.46	1211
Year	1951	1983.36	1985	2010	1211
Verbosity	73.42	100.68	100	120.38	1211

The dummy variable for early elections uses information on term length from Strøm, Müller and Bergman (2006) and Woldendorp, Keman and Budge (2000). The measure tries to grasp whether or not an election was early in a sense that it was unexpected (rather than in terms of the underlying reasons). Therefore, a primarily temporal definition is used. Political systems, however, differ with regard to the discretion they grant governments to set an election time (Strøm and Swindle, 2002). For countries where governments have considerable lee-way in calling elections, an election taking place towards the end of the legislative term may not be that unexpected (as illustrated by the British convention not to consider elections in the last year of the government term to be early (Bergman et al., 2006, :167)). We therefore use two different temporal thresholds. If the reason for the termination of the previous government as coded by Woldendorp, Keman and Budge (2000) is not an election (but e.g. lack of parliamentary support), then the election is considered early/unexpected if it takes place in the first nine tenth of the legislative term (counted from the previous election date). If the reason for the end of the previous goverment is given as "election", this election is considered early/unexpected only if it occurs in the first four fifths of the parliamentary term.

The indicator variable for young democracies equals one for elections taking place in the first ten years after the country was first/again given a Polity Score greater than six. GDP per capita information (PPP converted, in 1000 international $ at current prices) is taken from Heston, Summers and Aten (2011). The variable Election year is self-explaining. Finally, in order to test whether or not the parsing into quasi-sentences successfully removes language effects, a measure of the relative verbosity of languages is added. It is matched by language the document is written in (from CMP documentation). The variable reflects how many words a language requires in order to express the equivalent of a 100 English words.[10] Descriptive statistics for all these variables are presented in Table 4.

[10]We derived this measure using word counts across translations of several political documents (the UN declaration of human rights, one OECD policy document, and various election manifestos from European party groups from the 2009 European elections), and averaging across all available sources for the respective language.