Measuring issue salience for political parties using LLMs

Kenneth Benoit
Singapore Management University and London School of Economics

Michael Laver
New York University and London School of Economics
(Corresponding author: ml127@nyu.edu)

ABSTRACT

Understanding both party positions on issues and the relative importance, or salience, they attach to these is central to the study of party competition. Recent work shows that large language models (LLMs) can estimate *issue positions* from manifestos. We ask here whether LLMs can also capture *issue salience*. Building on work by Benoit *et al*., we adapt LLM-based methods to measure relative issue salience, applying these to a multilingual corpus of manifestos. We test alternative strategies—including unconstrained scoring, ranking, saliency budgets, and pairwise comparisons—and assess their validity compared to benchmark expert surveys, as well as estimates derived from human-annotated manifestos. Results show that LLMs can produce meaningful estimates of relative issue salience, though with lower correspondence to expert judgements than for issue positions. This highlights the promise of LLMs, the need for careful thinking about the concepts of issue importance and issue salience, and about whether party manifestos, as strategic documents, are the best source of information abut the "true" importance parties attached to particular issues.

Keywords: Issue salience, party competition, party manifestos, large language models (LLMs), expert surveys, text-as-data

Funding: The authors gratefully acknowledge funding from the Data Science Institute, London School of Economics

Biographical details:

Ken Benoit is Dean and Professor of Computational Social Science at the School of Social Sciences, Singapore Management University. He was Director of the Data Science Institute at LSE from 2020 to 2024.

Mik Laver is Emeritus Professor of Politics at New York University, where he was Dean for the Social Sciences 2012-2018, and Visiting Professor jointly in the Data Science Institute and Department of Methodology at the London School of Economics.

INTRODUCTION

Measuring where political parties stand on key issues—and how much they care about each issue—is central to the study of party competition. Early analyses of the text in party manifestos by the Manifesto Project (MP) were motivated by a "saliency theory" of party competition which focussed on measuring the *relative importance*, or *salience*, that parties attach to different issues (Budge, et al. 1987). More recently, while much progress has been made in estimating parties' policy *positions* (Laver 2014), less attention has been paid to measuring issue salience. Salience matters, however. Party competition and government formation often concern not just where parties stand on important issues, but which of these they prioritize.

In what follows, we develop and evaluate a new method for estimating issue saliences for individual political parties, using large language models (LLMs) to read and analyse party manifestos. Our approach builds on recent work using LLMs to score issue positions in party manifestos (Benoit, et al. 2026), modifying this for what turns out to be the substantially more difficult task of inferring how salient different issues are for each party. One reason the task is more difficult is that issue salience is inherently relative. The salience of one issue only has meaning relative to the salience of other issues. Issue saliences cannot be measured independently of each other. Another reason is that, while issue positions tend to be set out relatively explicitly in party manifestos, the relative saliences of different issues tend to be more implicit, with the result that manifestos may be less reliable sources of information about issue saliences than they are about issue positions.

Analysing a multilingual corpus of party manifestos, we compare LLM-based estimates of issue salience with those derived from two different benchmarks. The first is the set of expert surveys used by Benoit et al (2026) as benchmarks for issue positions. The second is a labelling of the same documents we analyse by the MP's human expert coders (Budge, et al. 1987, Laver and Budge 1992, Klingemann, et al. 1994, Budge, et al. 2001, Klingemann, et al. 2006). Comparing our LLM estimates with both benchmarks allows us to infer how much information manifestos contain about issue saliences as opposed to issue positions – relative to the aggregated wisdom of expert survey respondents. This exercise demonstrates both the promise and the challenges of using AI-based tools to move beyond existing "text as data" (TaDa) methods toward deeper semantic understanding of political texts. (For overviews of the TaDa program see Grimmer and Stewart 2013, Gentzkow, et al. 2019, Benoit 2020).

To see why we often need to know both issue positions and issue saliences of political parties, consider the hypothetical setting for government formation shown in Figure 1. There are three parties A, B and C and two policy dimensions. No single party commands a legislative majority. Which majority coalition is most likely to form? The answer to this question depends entirely on the relative salience of each dimension for each party. If both dimensions are of equal salience for all parties, then absent any further information most models of government formation predict that coalition AC will form. If economic policy is much more salient than social policy for parties B and C, then coalition BC is most likely to form. Both parties get what they want on economic policy and care much less about their differences on social policy. If economic policy is much more salient for party A, but social policy is much more salient for party B, then they can "logroll" to form coalition AB. Party A can get its ideal economic policy, while Party B can get its ideal social policy (De Marchi and Laver 2023). For precisely the same configuration of party policy positions, *any* of the possible two-party coalitions might form, depending on the salience of each issue dimension for each party. More generally, in the multidimensional settings for party competition and government formation which are of course generic, measuring the "policy loss" between any agent's ideal point and actual or potential political outcomes requires weighting each issue dimension by its salience for the agent concerned.

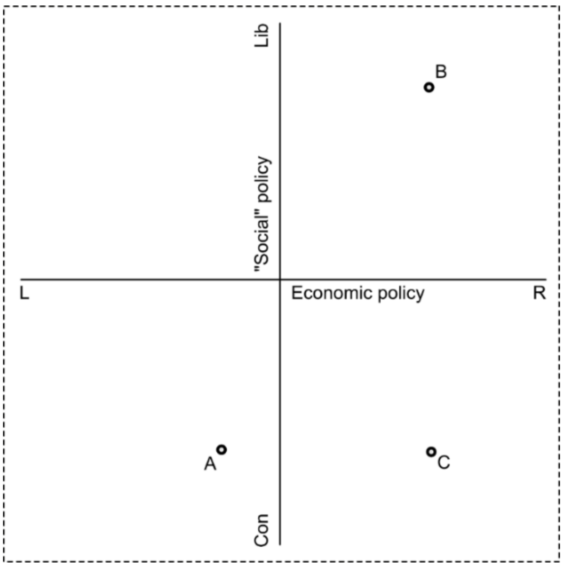


Figure 1: Government formation with two issue dimensions

In what follows, we:

- discuss the distinction between issue salience and the more informal notion of "issue importance";
- discuss the associated methodological problem of "measurement separability", which makes measuring issue salience more challenging than measuring issue positions;
- identify a text corpus and set of issue dimensions that allow us to compare LLM estimates of issue salience with estimates derived from benchmark expert surveys and MP expert labelling;
- build on Benoit *et al.* (2026) to develop methods for measuring relative issue salience;
- discuss general implications of new AI/LLM based methods for political science research.

"MEASUREMENT SEPARABILITY" OF ISSUE POSITIONS AND SALIENCES

When we measure a politicians' positions on multiple latent issue dimensions, we typically assume, albeit implicitly, that *measured positions on one issue are independent of measured positions on any other.* This is consistent with a deeper *theoretical* assumption, explicitly deployed by most spatial modelers, that fundamental political preferences are "separable": that preferred outcomes for issue *X* do not depend on preferred outcomes for issue(s). The analogous measurement assumption is that politicians' measured positions on one issue dimension do not depend on their measured positions on other issue dimensions. We can think of this as the "measurement separability" of issue positions.

The meaning of a politician's *position* on some issue is relatively uncontroversial common ground among scholars of party competition and voting behaviour. The meaning of the *importance*, or *salience*, of the same issue for the same politician is harder to pin down (Gemenis 2013, Steenbergen and Marks 2007). While we sometimes use the terms importance and salience interchangeably, there is a distinction worth noting. In survey research, the more informal notion of issue *importance* is often conceived, implicitly, in absolute terms. Survey respondents or experts may judge many issues to be simultaneously "very important". Issue salience, as opposed to issue importance, is a more precise concept. In the loss functions of well-specified spatial models, for example, issue salience is treated as *inherently relative*, even if this is implicit. The natural way to express the salience of issue X for me at time t is to compare how strongly I feel about X at time t with how strongly I feel about other issues, Y and Z, at the same time—or to compare how I feel about X at time t relative to how I felt about X at t-1. In what follows, therefore, when we refer to "issue salience", we are referring to a more precise definition of the general notion of issue importance, which expresses the *relative importance* of an issue, compared to that of a fixed set of other issues.

If issue salience—as opposed to issue importance—is relative rather than absolute, then the salience of one issue only has meaning in the context of the salience of other issues. We can express this in two different ways. The first is *ordinal*. We can say that issue X is more important to me than issue Y today—or than X was to me last year—without saying *how much* more. More generally, we can *rank* the salience of a given set of issues. A second way to measure relative salience is at the *cardinal* level, using an interval or ratio scale. *How much* more salient is issue X to me than issue Y, where zero salience means that the issue is of absolutely no importance for me? One way to think about relative salience in behavioural terms sees politicians as having a finite "saliency budget", a fixed stock of time and effort which they can credibly allocate across issue areas to optimize overall outcomes. Given a finite saliency budget and a fixed set of issues and/or time points, there is a vector of relative issue saliences which sum to unity.

A very important difference between ordinal and cardinal measures of relative salience is that, for cardinal measures where relative saliences sum to a constant across some set of issues and/or time points, *adding or subtracting from this set changes relative saliency weights for all remaining issues or time points* since there is less, or more, saliency weight to go around. This in turn means that we cannot make the same measurement assumption about cardinal level issue saliences that we make about issue positions. Cardinal measurements of relative saliences are *not* independent of each other. They do not have measurement separability. This has important implications for how we measure them. The saliences we estimate now depend, for reasons we have just seen, on the precise set of issue dimensions and/or time points we are investigating. Measuring cardinal issue saliences therefore requires committing to a set of issue dimensions and time points. *Adding new issues or time points potentially changes these measurements.*

Whether we need ordinal or cardinal measures of relative salience for a particular set of issues depends upon why we want the data. We may be satisfied with ordinal rankings of issue saliences, perhaps because we are modelling logrolling (de Marchi and Laver 2020, 2023). Even though salience scores for a set of issues do not sum to the same constant for each party, we can still compare their saliency *rankings* between parties. For example, if three issues are ranked in the order *x, y, z*, then adding issue *q* to the choice set at any point in the ranking does not change the ordering of *x, y* and *z*. Similarly, subtracting *y* from the choice set does not change the ordering of *x* and *z*.[1] If on the other hand we want compute salience-weighted policy distances between politicians and stimulus points, perhaps for one of the many spatial models of party competition or government formation, then we need *cardinal* saliency weights to support the necessary multiplications. In contrast to the measurement of "separable" issue positions, measuring cardinal issue saliences requires *committing to a set of issue dimensions and time points*.

BENCHMARK ISSUE DIMENSIONS AND DATA

In what follows, we commit to the set of six issues for which party positions were estimated using LLMs by Benoit *et al.* (2026). It is important to note that these issues were selected, not because they are assumed span the salient issue space for all elections in postwar Europe, but because this is a proof-of-concept exercise. The six benchmark issues were consistently deployed in expert surveys over the time period under investigation, presumably because survey designers did indeed feel they are important. In common with Benoit et al. (2026), this allows us to compare LLM estimates with an expert benchmark. Having established proof-of-concept, we plan in future work to tackle the fundamentally more challenging task of answering the open-ended question of *which particular issues are salient in a given setting of interest*.

The six issues were[2]:

- economic: policies on taxation, spending, trade, economic growth, and labour markets;
- social: social welfare, healthcare, education, equality, and cultural policy;
- immigration: immigration policy, border control, integration, asylum;
- EU: position on the European Union and related institutions;
- environment: environmental protection, climate change, energy policy; and
- decentralization: distribution of powers between national, regional, and local governments.

As a control, we added an issue area deemed important enough for governments to allocate substantial budgets to it, but one that we confidently expected not to be mentioned in any party manifesto:

- planetary defence: impact mitigation for near-earth objects (asteroids, comets, etc.).

Our benchmark expert survey series begins with Laver and Hunt, continuing with Benoit and Laver, and the Chapel Hill Expert Survey, CHES (Laver and Hunt 1992, Benoit and Laver 2006, Hooghe, et al. 2010, Jolly, et al. 2022). Experts in these surveys are asked for their judgments of parties' issue positions and importances without any instruction to refer to party manifestos – and without any reference to *relative* issue salience. By implication, they are asked about a party's "true" issue positions and saliences, which may differ from those set out in manifestos. While we cannot directly observe these true positions and saliences, we know that observed words and deeds may well be strategic misrepresentations. Expert country specialists are more likely to rate parties' issue positions and saliences in the round, however, if they read between the lines of strategic statements in party manifestos to give something closer to the true party position. As expert benchmarks, we used both the "raw" scores from the expert survey data and, to capture the *relative* salience of each issue in relation to the other five issues, we rescaled these raw scores to sum to unity across the six-issue set to which we are committed.

Our analysis is complicated by the fact that, unfortunately, the benchmark CHES often did not ask experts to rate issue saliences[3] for all six benchmark issues, sometimes only asking this for only three or four of these. Indeed, we only have expert survey ratings of the salience of all six issues for 72 manifestos in the corpus. We return below to discuss the implications of this.

For a more direct comparison between our manifesto-based results and another canonical measure derived directly from party manifestos, and following Benoit et al. (2026), we draw on the MP dataset for our second benchmark. This measures issue saliences in terms of proportions of mentions of the issue in each manifesto. We aligned the MP issue categories with the six expert survey dimensions (see Appendix B for details), measured issue saliences in the same way as the MP in terms of the proportion of manifesto references labelled as dealing with each issue, then transformed the relevant proportions into "budgeted" scores by rescaling them to sum to unity across the six issue dimensions.

Text corpus and issue dimensions

Our texts are party manifestos, collected and made available by the Manifesto Project (MP).[4] Our corpus is based on that of Benoit *et al.* (2026); we also limited the sample to party manifestos issued in election years for which an expert survey was also conducted, (see Appendix A for a full listing). Our dataset covers a total of 246 party manifestos, from 21 different countries, written in 22 different languages. The country coverage of this corpus is described in Table 1, along with the total token counts for each country. Individually, the manifestos ranged from 391 to 246,338 tokens in length, with an average of 26,266 tokens.

Table 1: Description of manifesto corpus for LLM analysis

Country	*Number of Manifestos*	*Main Languages*	*Total Tokens*[5]
Austria	9	German	167,273
Belgium	33	Dutch, French	2,567,757
Bulgaria	7	Bulgarian	37,987
Croatia	8	Croatian	69,552
Czech	11	Czech	208,689
Denmark	10	Danish	140,081
Estonia	5	Estonian	59,444
Finland	16	Finnish, Swedish	116,632
Greece	7	Greek	104,461
Hungary	15	Hungarian	400,374
Iceland	5	Icelandic	9,372
Ireland	5	English	76,483
Netherlands	32	Dutch	799,755
Norway	8	Norwegian, Nynorsk	226,078
Poland	6	Polish	101,053
Portugal	7	Portuguese	361,274
Slovakia	12	Slovak	205,979
Slovenia	7	Slovenian	86,443
Spain	5	Spanish	219,456
Sweden	23	Swedish	158,310
UK	15	English	344,949
Total	246	22 languages	6,461,402

USING LLMS TO READ AND INTERPRET PARTY MANIFESTOS ISSUE CONTENT

In what follows, we use LLMs to approximate the human qualitative analysis of party manifestos, answering queries about relative issue saliences expressed in these. Our approach is that of "holistic grading", based on a reading of the document as a whole rather than sentence-based annotation, as in the Manifesto Project (Budge, et al. 1987, Laver and Budge 1992, Budge, et al. 2001, Klingemann, et al. 2006, Volkens, et al. 2013) or the Comparative Policy Agendas Project (Baumgartner, et al. 2006, Baumgartner, et al. 2019).

"Holistic grading" is an approach used by the Global Populism Database project, whereby trained human coders read and evaluate entire speeches or documents to identify varying degrees of populist discourse. This approach, adapted from educational psychology, allows for the capture of subtle, dispersed rhetorical features—such as tone, framing, and moral juxtapositions—that would elude sentence- or keyword-based methods. It also supports consistent cross-language application and provides a nuanced, continuous measure of populist rhetoric (Hawkins, et al. 2019). By using LLMs to assess all text in a manifesto holistically, we can capture the nuances, emphases, relative intensities, and other linguistic and rhetorical devices that make up the political communication of salience. While this approach is the natural way a human reader would approach reading a text for measuring relative issue salience, it scales very badly because of the extreme demands of time, attention, and linguistic abilities demanded of human readers of manifestos that often run into the hundreds of pages. With modern LLMs, however, we show that we can approximate this process while eliminating the barriers to implementing it consistently and at scale.

Benoit *et al.* (2026) used an ensemble of LLMs (GPT-4o, Claude 3.5 Sonnet, and Gemini 1.5 Pro) to measure *issue positions* in party manifestos. This involves asking each LLM first to summarize what the manifesto has to say about each issue under investigation. For each issue summary for each manifesto, each LLM is then asked to score the issue position on a seven-point scale, where each scale point is given a substantive interpretation. The multiple issue scores generated in this way are summarized using an ensemble mean LLM-based issue position score for each issue in each manifesto.

This method yielded extremely highly correlations for five of the issues investigated—of the order of 0.90—with the benchmark expert ensemble means.[6] Benoit *et al.* showed that these correlations are at the upper bound of what can reasonably be expected, established by estimating distributions of split-sample correlations between two "independent" expert surveys. Testing the robustness of their findings, Benoit *et al.* almost perfectly replicated their analysis using three powerful open-weight LLMs (Llama-3.3, Gemma-3 and Deepseek-V1).

USING LLMS TO READ AND INTERPRET MANIFESTOS FOR PARTIES' ISSUE SALIENCES

There are various ways in which relative issue salience may be expressed in political text. One way is purely quantitative—how frequently is the issue mentioned? The Manifesto Project (MP) measures relative issue salience in precisely this way, drawing on a "saliency theory" of party competition which assumes parties compete, not by changing their issue positions but by attempting to increase the salience of issues they feel favour them by increasing the proportions of their party manifestos devoted to these issues (Budge, et al. 1987, Budge, et al. 2001). The MP therefore measures relative issue salience as the proportion of sentences in a text mentioning the issue in question, *regardless of the substantive content of each mention*. Issue salience in a text may, however, also be expressed in the intensity of a statement, apart from its frequency, with one strongly worded sentence potentially eclipsing a dozen milder mentions. Issue salience may also manifest in terms of a text's structure—where in the text the issue appears, whether a separate heading is devoted to it, and so on.

A further complication is that, while issue position is typically explicit, *issue salience is often implicit*. Politicians tend to state issue positions with at least some degree of clarity, endorsing or criticizing specific policy positions. In contrast, politicians seldom explicitly state which issues are most important to them and which, by implication, are relatively unimportant. One reason for this could be that politicians seek to avoid alienating potential supporters who have different issue priorities. On the other hand, political rhetoric also carefully aims to signal priorities to one's political base, by communicating some statement recognising their importance. Navigating these tendencies for different parties across different issues, in a multidimensional context, is a necessarily complex and more qualitative task. It suggests that salience must typically be inferred—from structure, repetition, placement, and rhetorical emphasis within the text—rather than observed directly.

An additional concern is the *lack of measurement separability* of issue saliences. For the set of six issues identified as our benchmark, we are asking LLMs not simply to rate a manifesto position on each issue in isolation, but to rate each issue's salience in the manifesto relative to five other issues; six issues must be evaluated simultaneously.

All of this makes salience harder to measure using traditional methods, which often rely on surface features such as word counts or sentence-level topic labels. The LLM method we describe below is however particularly suited for this task, interrogating a text *holistically* in terms of the entire document's rhetorical structure and thematic focus. This allows us to ask not just how often some issue is mentioned but how salient it seems in context—a closer approximation to the kind of judgments that human readers make when assessing which issues a party most cares about—even when that emphasis is not explicitly declared.

Taking all of this into account, LLM methods we develop here are grounded in two different measurement assumptions, both resting on the notion that issue salience can only meaningfully be expressed in relative terms. The first, capturing the notion of a "saliency budget" is that the estimated vector of *cardinal issue saliences for a given set of issues must sum to unity*. The second is based on the insight that, both behaviourally and methodologically, *relative salience is most unambiguously expressed in terms of pairwise comparisons*. Forced to choose between getting your ideal policy on either issue X or issue Y, which issue would you choose?

Our baseline method for using LLMs to measure relative issue saliences was a version of the Benoit *et al.* approach. In light of our discussions of the meaning of relative issue salience, we modified this method in three different ways. The first modification involved an LLM prompt asking LLMs to *rank order* the salience of each issue for each party. The second modification imposed a "saliency budget" on the LLM, prompting for cardinal issue saliences, but requiring that these sum to unity across the set of issues.

The third modification presented the LLM with each possible pair of issues in the issue set, asking which is the more important. (All prompts can be found in Appendix C). The resulting data matrix recording "wins" for each issue pair contains the information we need to compute the relative salience of each issue in the set. We could, for example, compute a simple "win rate" for each issue—the relative frequency with which each issue is preferred in the set of pairwise comparisons. A widely-used and more refined approach, which exploits more of the information in the matrix of pairwise issue wins, is the Plackett-Luce model (Plackett 1975, Luce 1959, Turner, et al. 2020). In essence, this assumes each issue has an unobserved weight, with the probability of one issue "beating" the other being determined by their relative weights. The algorithm computes the set of salience weights that maximizes the probability of observing the relevant win matrix.

Since Benoit et al. found little difference in the performance of the three main proprietary LLMs—Claude, Gemini and GPT—we used the LLM which, at the time of our research, had the largest context window. This was Gemini 2.5 Pro, accessed via its API. Released by Google in June 2025, Gemini 2.5 Pro is a multimodal transformer model featuring a native two million-token context window. This vast context capacity makes it exceptionally well-suited for analysing long documents in their entirety without segmentation[7], enabling exactly the holistic approach to measuring salience argued for above. This allowed even the largest manifestos to be analysed without summarizing or "chunking", particularly important since estimating relative issue saliences requires simultaneous consideration of all six issues, as these are discussed throughout the manifesto.

Checking for "contamination" of manifesto-based estimates from the LLM's training data

Versions of LLMs deployed via a chat interface, as opposed to those deployed via APIs, are increasingly "agentic", optimising their answer to any prompt not only using their training data but also using web searches and other information. We use API versions of the LLM to analyse large collections of documents based on a single prompt. These are not agentic, but for this application we also want the LLM to *confine itself strictly to manifesto texts* when answering our questions, rather than relying on its extensive "general knowledge". Our prompts therefore specifically instructed the LLM to confine itself to manifesto content. We tested that it was in fact doing this using an experiment which added a seventh "control" dimension to the issue set: planetary defence, as described above. Planetary defence from Near Earth Objects (NEOs) is generally deemed of huge importance, despite the low likelihood of a strike, due to the catastrophic potential consequences to humanity from asteroids or comet impacts—a fact clearly acknowledged by Gemini when we queried it on this issue through the chat interface.[8] It is, however, not a "salient" issue likely to have been mentioned in any party manifesto. If the LLM is strictly following instructions, therefore, then it should ignore its general knowledge of this issue's importance and report Planetary Defence as having zero salience in the text. Table 2 below shows that we found that this was invariably the case and concluded that potential contamination of this type is not a concern.

RESULTS

Table 2 shows mean salience scores across all manifestos for each of the six issue dimensions under investigation, computed using output from Gemini for the various approaches we discuss above. The first shows issue saliences estimated by prompting for a score between 0 and 10 for each dimension, without constraining these to sum to a fixed "saliency budget". This is the prompt directly analogous both to the Benoit *et al.* method for estimating issue positions, and to the way in which expert survey respondents were prompted. It is in this sense the most directly comparable method. We see economic and social policy rated as the clearly the most important, followed, in rank order, by environmental protection, decentralization, the European Union, and immigration. Our control dimension, "planetary defence", rated at the bottom on every scale, indicating that Gemini was following prompt instructions to base its importance scores strictly on the text provided.

The primary importance of the economic and social policy dimensions is replicated by each of the methods we used: cardinal issue importance constrained to a budget of unity; ranked issue importance; the proportion of manifestos ranking the issue first; Plackett-Luce scaling of the relative importance of each pair of issues. Environmental protection is always rated the third most important issue, with some minor permutations in the rankings of the three least salient issues.

Table 2: Mean importance scores, by issue dimension

Dimension	*Unbudgeted cardinal 1-10*	*Budgeted cardinal 0-1*	*Rank*	*Proportion ranking first*	*Plackett-Luce 0-1*
Economic	9.53	0.32	1.69	0.51	0.39
Social	9.43	0.33	1.89	0.30	0.32
Environment	7.06	0.13	3.79	0.11	0.12
Decentralization	5.80	0.08	4.40	0.03	0.06
EU	6.47	0.08	4.35	0.02	0.06
Immigration	4.75	0.06	4.88	0.04	0.05
Planetary defence	0.00	0.00	7.00	0.00	0.00

Table 3 reports correlations between our different LLM-based measures of relative salience for each issue dimension and our benchmark expert survey estimates of the same thing. The first three columns show very similar correlations with expert scores for: unbudgeted cardinal salience scores; budgeted cardinal scores; ordinal salience rankings. These results have two striking features. First, correlations of LLM *salience* scores with the expert benchmark are substantially worse than the equivalent correlations Benoit *et al.* found for issue *positions.* This is consistent with the much greater difficulty of both defining and measuring issue salience, discussed above.

Second, correlations are noticeably higher for the issues of immigration and the environment. One plausible conjecture is that this happens because there are parties (far-right parties for immigration and green parties for the environment) whose manifestos place unambiguous emphasis on these issues, making issue importance much easier to estimate in these cases. Appendix A1 shows that there are very few observations for any one individual party, so we cannot break out statistically meaningful estimates by party. We can however compute estimates for *groups* of parties. We did this and found that correlations between

LLM and expert scores for the salience of immigration policy were significantly higher for far-right than for non-far-right parties – a pattern not repeated for any other issue. The analogous effect for Green parties and environmental policy was somewhat different. In contrast to the situation for immigration policy and the far right, where there was some variation between parties on the salience of immigration, there was little variation in respective estimates of the (high) salience of environmental policy for Green parties. Despite high *correspondence* between LLM and expert estimates of the salience of environmental policy for Green parties, lack of variation in these meant that *correlations* between these quantities were low. So Green parties' manifestos are not driving the results in Table 3.

Table 3: Correlations of LLM estimates with expert benchmark

Dimension	*Unbudgeted cardinal 1-10*	*Budgeted cardinal 0-1*	*Rank*	*Plackett-Luce*	*n*	*Budgeted expert-budgeted LLM*[9]
Economic	0.40	0.37	0.39	0.15	143	0.68
Social	0.17	0.08	-0.05	0.10	143	0.24
Environment	0.59	0.71	0.64	0.47	169	0.81
Immigration	0.64	0.70	0.67	0.55	172	0.77
EU	0.27	0.38	0.45	0.15	205	0.50
Decentralization	0.40	0.47	0.45	0.15	143	0.60
Fisher z-average	0.43	0.48	0.45	0.27		0.63

Taking all issues together, the method most closely aligning with expert judgements used a prompt which asked for a cardinal relative salience score for each issue, constraining the set of six scores to sum to unity—the "budgeted" cardinal score. Adding this additional constraint to the prompt, over and above simply asking for a salience score for each dimension, improved LLM performance. We can see this by comparing the Fisher z-average scores, a way of summarising correlation values across dimensions, for each measure type (a method preferable to simply averaging raw correlations, since it accounts for the nonlinearity of the correlation scale).[10] In Table 3, the budgeted LLM score correlates most highly with the benchmark expert surveys, with a z-score of 0.48.

The fourth column of Table 3 reports correlations between the expert benchmark and Placket-Luce estimates of issue "worth" derived from pairwise rankings of issue importance generated by LLMs. These correlations are clearly weaker than for the other three measures, although they remain noticeably higher for immigration and environmental policy. We do not know precisely why the pairwise method did not perform as well as expected. One likely reason is that many LLM salience scores were very close to each other, producing artificial wins and losses. For example, an issue with salience 0.20 would 'beat' one scored 0.19 just as decisively as if their saliences were 0.40 and 0.01, even though the former pair are essentially tied. Consistent with this, more than half of the interval-scale assignments gave the same integer values to three or more issues (versus just under 21 percent in the budgeted scores), while ranked and pairwise judgments did not allow ties. In this sense, the underlying salience values—if they can be estimated directly—contain more information than the implied win matrix. Because LLMs can generate such direct estimates, relying on pairwise comparisons turned out to be an inferior strategy.

The experts who generated our benchmark salience scores were not asked to subject their judgements about issue importance to a budget constraint—they could in theory have ranked every issue as having the utmost importance. To improve correspondence between our "budgeted" LLM estimates and the expert benchmark, we therefore rescaled raw mean expert judgments to sum to unity over the six issues under investigation, generating a "budget-transformed" expert benchmark score. Given the lack of measurement separability for relative issue salience, discussed above, we can only make a valid comparison of budgeted scores for the 72 cases for which expert judgments of issue salience were collected for all six issues. The results are shown in the final column of Table 3. Comparing like with like, the budgeted LLM estimates correlate substantially more highly with the budgeted expert benchmark, with a Fisher z-average of 0.63 across all six dimensions. These correlations come much closer to matching the high correlations reported by Benoit *et al*. for LLM estimates of issue positions. On the basis of these results, this is our preferred method, indicating that issue saliences should be measured relative to one another, subject to an overall saliency budget.

Are experts seeing aspects of issue salience that are not in the manifestos?

Benoit *et al.* found evidence, particularly regarding decentralization, that experts were seeing aspects of "true" party policy positions that were not evident in party manifestos. We find this to be even more evident for issue saliences. For a more direct comparison between our results and measures derived directly from party manifestos, we use the MP dataset as our second benchmark, as described above. Figure 1 compares the distributions of these "budget-transformed" expert- and MP-based scores, to the budgeted LLM-estimated issue saliences.

The strong pattern is that the LLM analyses of manifesto content "over-estimate" the salience of economic and social policy *against a benchmark of expert judgements*. Given the finite salience budget, these over-estimates necessarily produce "under-estimates" of the saliences of other issues.[11] The LLM analyses do not, however, overestimate the salience of economic and social policy *against a benchmark of human manifesto labelling*. This is consistent with a general tendency for party manifestos to include discursive *tours d'horizon* of economic and social policy—a level of manifesto content out of proportion to what experts evidently see to be the "true" relative importance of these matters—while setting out party positions on other important issues in a more succinct manner. Such an outcome would arise if the LLMs were attaching considerable weight when estimating relative salience to the proportion of the manifesto devoted to each issue—effectively using a method to estimate issue salience similar to that used by the MP.

Figure 1: Distributions of "budgeted" saliences

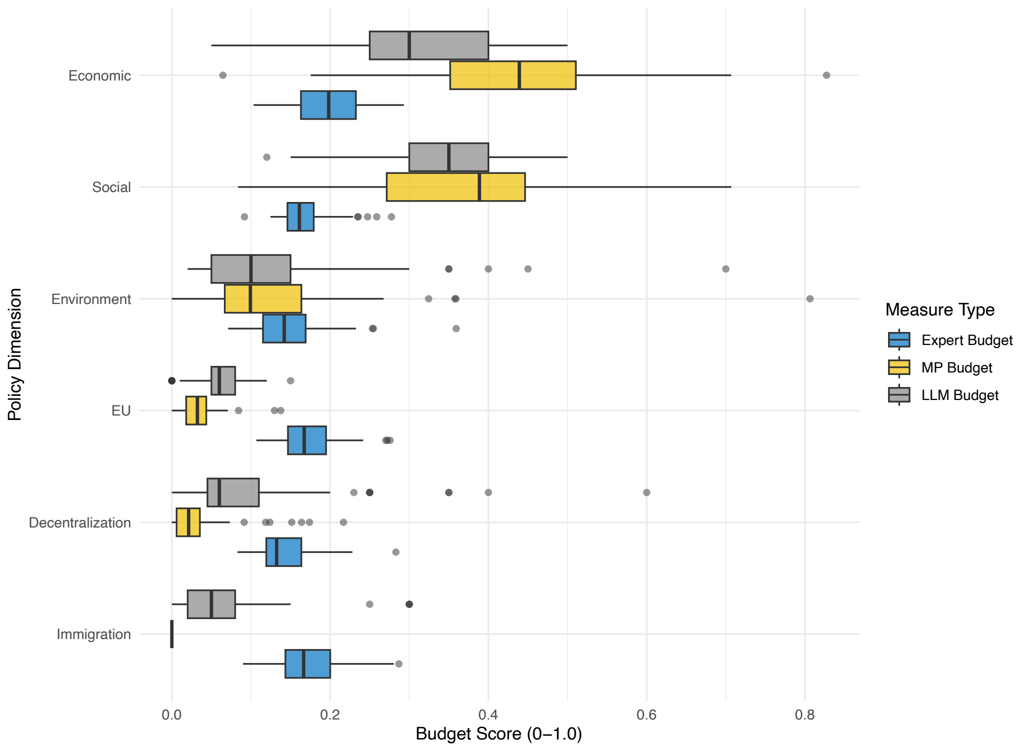


Figure 1 is consistent with this conjecture, showing that LLM and MP estimates of issue salience have similar distributions. Both methods "over-weight" economic and social policy relative to expert judgments and consequently "under-weight" the relative salience of the three least important issue dimensions. Table 4 shows that the similar distributions of LLM and MP salience weights are also associated with similar correlations between both LLM (see Table 3) and MP estimates and the expert benchmark. Figure 2 and Table 4 are consistent with a view that the LLMs are doing something similar to the MP's text human annotators—albeit doing so many orders of magnitude faster and more cheaply.

We draw three broad conclusions from these results. First, manifesto-based LLM analyses of issue *salience* did not have the same remarkable success as Benoit *et al.*'s use of essentially the same methods to estimate parties' issue *positions*—although results were substantially improved by imposing a salience budget on the cardinal scores estimated by the LLM. There are likely two main reasons for this. One is that issue saliences do seem to be much more implicit in party manifestos than issue positions, once we move beyond the simple "percentage of mentions" approach pioneered by the MP. Another reason is that the lack pf measurement separability for issue saliences means that these must be computed simultaneously for the six benchmark issues. The Benoit *et al.,* method for using LLMs to compute issue position, in contrast, computes issue positions one at a time.

Table 4: Correlations between "budgeted" LLM, MP and expert salience scores

Dimension	*Budgeted MP-budgeted expert (n=72)*	*Budgeted MP v budgeted LLM (n=63)*
Economic	0.68	0.49
Social	0.24	-0.07
Environment	0.81	0.87
Immigration	0.77	NA
EU	0.50	0.63
Decentralization	0.60	0.55
Fisher z-average	**0.63**	**0.56**

A second striking feature of the results is that the approach based on pairwise issue comparisons did not work particularly well. This may be because the ideal use case for this approach is when the underlying saliences cannot be measured directly and completely, but must be inferred from observations of the outcomes of an incomplete set of pairwise match-ups. In our case the LLM (and the MP) can estimate the full set of issue saliences directly.

Thirdly, and in our view most importantly, our results suggest we need to give serious consideration to the best benchmark against which to assess any new method for measuring issue saliences. An alternative way to think about this is that these results raise the question of what, precisely, we are looking for when we analyse party manifestos. Figures 1 strongly suggests that experts are seeing aspects of issue saliences which are not revealed in the party manifestos. Distributions of the two manifesto-based estimates of issue salience, from the LLM and the MP, are similar to each other. But these both differ sharply from the distributions of expert judgments. In effect, according to the experts, there is more to issue salience than can be inferred from party manifestos. This is certainly consistent with a view that manifestos are strategic documents which do not fully convey parties' "true" issue priorities—at least as these priorities are judged by ensembles of country specialists. Manifestos clearly contain signals about these issue priorities, but such signals are imperfect. In short, any lack of correspondence between manifesto-based LLM estimates and the benchmark expert scores may well arise, not because of any shortcoming in the way LLMs read and analyse manifestos, but because these manifestos are imperfect sources of information about parties' "true" issue priorities – a finding which would have considerable implications for how these priorities are in fact measured.

CONCLUSIONS

Agents' issue positions and relative issue saliences are fundamental quantities to be estimated before applying a well-specified spatial model of party competition and government formation to any given setting. Measuring these quantities is therefore a core research programme for political science. Results we report here suggest that measuring agents' relative issue saliences using party manifestos as a source of data is a harder problem than measuring their issue positions. Issue saliences are likely treated more implicitly in party manifestos than issue positions, possibly because of the more ambiguous meaning of "salience" or "importance", but possibly also for strategic reasons. Adding to the problem, the lack of measurement separability for relative issue saliences makes the task of estimating them fundamentally more difficult. It is first necessary to commit to a set of issue areas for which *relative* saliences can be expressed. We commit to the set of six issues identified by Benoit *et al.* as having been consistently deployed in the benchmark expert surveys. Having committed to a set of issues, the LLM must consider all issues simultaneously when it reads a manifesto—rather than focusing on one issue at a time as it does when estimating separable issue positions.

Using LLMs to estimate the relative salience of different issues for different parties is, for obvious reasons, massively cheaper and faster than conventional methods relying on human experts or research assistants. Analysing a corpus of 246 manifestos, written in 22 languages, using any human-based approach would require recruiting, training, coordinating, and quality-controlling multilingual coders over a prolonged period.[12] LLM-based holistic reading and analysis of party manifestos dramatically reduces labour, time, and coordination costs while producing outputs that are fully replicable and scalable.

It might be argued in light of this that we should simply ask LLMs to score parties' issue priorities directly, based on their training data. If asked to do this, they typically do indeed give plausible answers. We do not recommend this approach, however, for epistemological reasons. The key issue for us is the *scientific traceability of our estimates to an explicit data source*. When an LLM is prompted without a document, its output cannot be meaningfully tied to observable evidence, nor can it be adjudicated, replicated, or falsified with respect to a defined corpus. In this sense, we treat LLMs as analogous to trained human research assistants: as measurement instruments applied to a clearly specified body of text, operating under explicit instructions. Asking an LLM to infer party priorities directly from its training data would instead amount to soliciting a model's latent "general knowledge", which may be plausible but is methodologically opaque and cannot be independently validated against the source material. The manifestos we analyse are freely downloadable. The added costs, in terms of both time and money, of LLM analyses based on party manifestos as opposed to their general knowledge are tiny, given the huge resulting gains in scientific traceability and replicability.

Benoit *et al.*'s very encouraging results on using LLMs to measure parties issue positions from party manifestos, doing so much more effectively than the MP's human text annotators for example, may raise unrealistic expectations. What we see here is that, addressing the thornier problem of measuring issue saliences, the LLM is performing at about the same level, and apparently in somewhat the same way, as the MP's human text annotators. This is of course no mean feat given that the LLM is producing, almost instantly and for very little money, results equivalent to those which took the MP huge amounts of time, money and human resources to generate.

Another bottom line is suggested by the overlap between the distributions of the two manifesto-based estimates of issue salience, and the disjunction between these and the "benchmark" expert surveys. For the most part, political scientists have treated manifesto data and expert surveys as two different ways of estimating essentially the same thing. We ourselves in effect assumed this when specifying expert surveys as our benchmark for assessing the performance of LLMs in analysing manifestos. Results we report above strongly suggest that party manifestos and expert surveys contain *different* information about parties' issue priorities. Experts are not referred to party manifestos when making their judgments and are in effect asked for their best estimates of parties' "true" issue positions and saliences. Party manifestos are strategic campaign documents. Comparing our findings on salience with those of Benoit *et al.* on positions, generated using very similar methods, our conjecture is that strategic incentives to be ambiguous about party positions are less potent than incentives to be ambiguous about issue saliences. This may be because, attempting to appeal to different sections of the electorate who have different issue priorities, parties may anticipate doing themselves more harm than good by being explicit about issue priorities in their manifestos. If some party supporters prioritize guns over butter, while others do the reverse, party strategists may want to avoid an explicit electoral appeal that in effect says, "we value both guns and butter but if forced to choose we'll spend the money on guns".

All of this goes to show that, while LLMs will surely become valuable research tools for social scientists, they must be deployed with the rigour and circumspection that is essential when deploying any new measurement instrument. Above all, we need to validate their output in any new application on which they are deployed. Benoit *et al.* found strong validation for using LLMs to estimate parties' issue positions from party manifestos. This suggests we can treat with some confidence estimates of issue positions generated by using the same LLM method on out-of-sample texts. We show here that we can't expect the same level of validation when deploying LLMs on the different problem of measuring issue saliences from the content of party manifestos. Possibly because the notion of issue importance is fundamentally more ambiguous, our results suggest we should have less confidence in estimates of issue saliences generated by using the same LLM method on out-of-sample texts.

We close by setting out what we see as next steps in the research programme of using LLMs to measure the issue positions and saliences expressed in political texts. Our work, like that of Benoit *et al.*, has been a proof of concept: validating LLM estimates of party positions and saliences against expert benchmarks on six key issue areas. These six were not meant to cover the full landscape of post-war party competition, but were selected because expert surveys used them and thus provided a basis for validation. To model party competition in a particular local setting, however, we need the estimates of party positions and saliences *on the issues that actually matter in each setting*. A fixed six-dimensional template cannot capture the issue structure in every modern democracy.

The core of our future work, therefore, will be to use LLMs first to *identify the key issue dimensions* in a given setting—typically a single election and post-electoral government formation in a single country—and then estimate party positions and saliences for that issue set. We see this as a core research module, producing the best possible data to model party competition and government formation in a specific time and place. Broader applications, such as building time series or cross-national comparisons, depend on this foundation. These are second-order tasks, however. The first priority is to generate the best possible estimates for each political issue context in which parties actually compete.

REFERENCES

Baumgartner, Frank R, Christian Breunig, and Emiliano Grossman. 2019. "The Comparative Agendas Project: Intellectual Roots and Current Developments."

Baumgartner, Frank R, Christoffer Green-Pedersen, and Bryan D Jones. 2006. "Comparative Studies of Policy Agendas." *Journal of European public policy* 13: 959-74.

Benoit, Ken. 2020. "Text as Data: An Overview." *The SAGE handbook of research methods in political science and international relations*: 461-97.

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

Benoit, Kenneth, Scott De Marchi, Conor Laver, Michael Laver, and Jinshuai Ma. 2026. "Using Large Language Models to Analyze Political Texts through Natural Language Understanding." *American Journal of Political Science* In press.

Benoit, Kenneth, Kohei Watanabe, Haiyan Wang, Paul Nulty, Adam Obeng, Stefan Müller, and Akitaka Matsuo. 2018. "Quanteda: An R Package for the Quantitative Analysis of Textual Data." *Journal of Open Source Software* 3: 774-74.

Budge, Ian, Hans-Dieter Klingemann, Andrea Volkens, Judith Bara, Eric Tannenbaum, Richard Fording, Derek Hearl, Hee Min Kim, Michael McDonald, and Silvia Mendes. 2001. *Mapping Policy Preferences: Parties, Electors and Governments: 1945-1998: Estimates for Parties, Electors and Governments 1945-1998.* . Oxford: Oxford University Press.

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

De Marchi, Scott, and Michael Laver. 2023. *The Governance Cycle in Parliamentary Democracies*: Cambridge University Press.

———. 2020. "Government Formation as Logrolling in High-Dimensional Issue Spaces." *The Journal of Politics* 82: 543-58.

Gemenis, Kostas. 2013. "Estimating Parties' Policy Positions through Voting Advice Applications: Some Methodological Considerations." *Acta politica* 48: 268-95.

Gentzkow, Matthew, Bryan Kelly, and Matt Taddy. 2019. "Text as Data." *Journal of Economic Literature* 57: 535-74.

Grimmer, Justin, and Brandon M Stewart. 2013. "Text as Data: The Promise and Pitfalls of Automatic Content Analysis Methods for Political Texts." *Political Analysis*.

Hawkins, Kirk A, Rosario Aguilar, Bruno Castanho Silva, Erin Kristin Jenne, Bojana Kocijan, and Cristóbal Rovira Kaltwasser. 2019. "Measuring Populist Discourse: The Global Populism Database." Paper presented at the EPSA annual conference.

Hooghe, Liesbet, Ryan Bakker, Anna Brigevich, Catherine De Vries, Erica Edwards, Gary Marks, Jan Rovny, Marco Steenbergen, and Milada Vachudova. 2010. "Reliability and Validity of the 2002 and 2006 Chapel Hill Expert Surveys on Party Positioning." *European Journal of Political Research* 49: 687-703.

Jolly, Seth, Ryan Bakker, Liesbet Hooghe, Gary Marks, Jonathan Polk, Jan Rovny, Marco Steenbergen, and Milada Anna Vachudova. 2022. "Chapel Hill Expert Survey Trend File, 1999–2019." *Electoral studies* 75: 102420.

Klingemann, Hans-Dieter, Richard I. Hofferbert, and Ian Budge. 1994. *Parties, Policies, and Democracy*, Theoretical Lenses on Public Policy. Boulder: Westview 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. 2014. "Measuring Policy Positions in Political Space." *Annual Review of Political Science* 17: 207-23.

Laver, Michael, and Ian Budge. 1992. *Party Policy and Government Coalitions*. New York, N.Y.: St. Martin's Press.

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

Luce, R Duncan. 1959. *Individual Choice Behavior*. Vol. 4: Wiley New York.

Plackett, Robin L. 1975. "The Analysis of Permutations." *Journal of the Royal Statistical Society Series C: Applied Statistics* 24: 193-202.

Steenbergen, Marco R, and Gary Marks. 2007. "Evaluating Expert Judgments." *European Journal of Political Research* 46: 347-66.

Turner, Heather L, Jacob van Etten, David Firth, and Ioannis Kosmidis. 2020. "Modelling Rankings in R: The Plackettluce Package." *Computational Statistics* 35: 1027-57.

Volkens, Andrea, Judith Bara, Ian Budge, Michael D McDonald, and Hans-Dieter Klingemann. 2013. *Mapping Policy Preferences from Texts: Statistical Solutions for Manifesto Analysts*. Vol. 3: OUP Oxford.

ONLINE APPENDICES

Appendix A: Data

Table A.1: Manifestos included in the sample, and their associated expert survey source.

Country	Party	Language	Expert Survey	Expert Survey Year	Tokens
Austria	Freiheitliche Partei Österreichs (FPÖ)	German	CHES	2006	4,278
Austria	Die Grünen	German	CHES	2006	9,074
Austria	ÖVP	German	CHES	2006	20,784
Austria	SPÖ	German	CHES	2006	7,994
Austria	Freiheitliche Partei Österreichs (FPÖ)	German	CHES	2019	3,343
Austria	Grüne	German	CHES	2019	46,720
Austria	NEOS	German	CHES	2019	22,692
Austria	Österreichische Volkspartei (ÖVP)	German	CHES	2019	15,273
Austria	SPÖ	German	CHES	2019	37,115
Belgium	Vlaams Blok	Dutch	BL	2003	27,129
Belgium	CD&V	Dutch	BL	2003	44,173
Belgium	VLD	Dutch	BL	2003	4,534
Belgium	Agalev	Dutch	BL	2003	20,143
Belgium	Ecolo	French	BL	2003	4,789
Belgium	Parti Socialiste (PS)	French	BL	2003	95,399
Belgium	Groen!	Dutch	CHES	2010	37,108
Belgium	Ecolo	French	CHES	2010	99,523
Belgium	Open Vld	Dutch	CHES	2010	27,043
Belgium	Vlaams Belang	Dutch	CHES	2010	559
Belgium	Nieuw-Vlaamse Alliantie (N-VA)	Dutch	CHES	2010	25,530
Belgium	Mouvement Réformateur	French	CHES	2010	135,178
Belgium	sp.a	Dutch	CHES	2010	15,150
Belgium	Parti Socialiste (PS)	French	CHES	2010	53,972
Belgium	CD&V	Dutch	CHES	2010	36,006
Belgium	cdH	French	CHES	2010	161,228
Belgium	Open Vld	Dutch	CHES	2014	21,639
Belgium	CD&V	Dutch	CHES	2014	91,864
Belgium	Fédération des Francophones de la Fédération Wallonie-Bruxelles (FDF)	French	CHES	2014	59,096
Belgium	Groen	Dutch	CHES	2014	226,748
Belgium	cdH	French	CHES	2014	220,238
Belgium	Nieuw-Vlaamse Alliantie (N-VA)	Dutch	CHES	2014	51,242
Belgium	Parti Populaire	French	CHES	2014	25,660
Belgium	Mouvement Réformateur	French	CHES	2014	246,338
Belgium	sp.a	Dutch	CHES	2014	126,541
Belgium	Parti Socialiste (PS)	French	CHES	2014	204,113
Belgium	PVDA+	Dutch	CHES	2014	80,850
Belgium	Open Vld	Dutch	CHES	2019	34,150
Belgium	CD&V	Dutch	CHES	2019	95,152

Belgium	Groen	Dutch	CHES	2019	47,438
Belgium	Nieuw-Vlaamse Alliantie (N-VA)	Dutch	CHES	2019	52,540
Belgium	sp.a	Dutch	CHES	2019	75,065
Belgium	PVDA	Dutch	CHES	2019	121,619
Bulgaria	АБВ (Alternative for Bulgarian Revival)	Bulgarian	CHES	2014	3,366
Bulgaria	Атака	Bulgarian	CHES	2014	1,431
Bulgaria	Coalition for Bulgaria (KB)	Bulgarian	CHES	2014	4,346
Bulgaria	ЛИДЕР	Bulgarian	CHES	2014	7,453
Bulgaria	GERB	Bulgarian	CHES	2014	12,420
Bulgaria	Patriotic Front (PF)	Bulgarian	CHES	2014	5,662
Bulgaria	Reformist Bloc	Bulgarian	CHES	2014	3,309
Croatia	Croatian Social Liberal Party and Democratic Centre coalition	Croatian	BL	2003	16,866
Croatia	Croatian People's Party (HNS)	Croatian	BL	2003	8,648
Croatia	HDZ (Hrvatska Demokratska Zajednica)	Croatian	BL	2003	19,847
Croatia	Liberalna stranka	Croatian	BL	2003	4,967
Croatia	Croatian Party of Rights (HSP)	Croatian	BL	2003	1,550
Croatia	HSS	Croatian	BL	2003	10,602
Croatia	SDP	Croatian	BL	2003	3,149
Czech	KDU-ČSL	Czech	CHES	2006	22,144
Czech	ODS	Czech	CHES	2006	25,303
Czech	Communist Party of Bohemia and Moravia (KSČM)	Czech	CHES	2006	7,536
Czech	Strana zelených	Czech	CHES	2006	30,599
Czech	ČSSD	Czech	CHES	2006	28,119
Czech	ODS	Czech	CHES	2010	17,336
Czech	Communist Party of Bohemia and Moravia (KSČM)	Czech	CHES	2010	4,261
Czech	Strana zelených	Czech	CHES	2010	34,656
Czech	Věci veřejné	Czech	CHES	2010	14,884
Czech	ČSSD	Czech	CHES	2010	12,316
Czech	TOP 09	Czech	CHES	2010	11,535
Denmark	Alternativet	Danish	CHES	2019	31,748
Denmark	Konservative Folkeparti	Danish	CHES	2019	16,527
Denmark	Dansk Folkeparti	Danish	CHES	2019	1,978
Denmark	Liberal Alliance	Danish	CHES	2019	3,924
Denmark	Venstre	Danish	CHES	2019	2,259
Denmark	Nye Borgerlige	Danish	CHES	2019	11,509
Denmark	Red-Green Alliance (Enhedslisten)	Danish	CHES	2019	5,368
Denmark	Socialdemokratiet	Danish	CHES	2019	43,065
Denmark	Radikale Venstre	Danish	CHES	2019	11,417
Denmark	Socialistisk Folkeparti (SF)	Danish	CHES	2019	12,286
Estonia	Eesti Keskerakond	Estonian	CHES	2019	14,231
Estonia	EKRE	Estonian	CHES	2019	4,162
Estonia	Isamaa	Estonian	CHES	2019	12,575
Estonia	Eesti Reformierakond	Estonian	CHES	2019	13,989
Estonia	Sotsiaaldemokraatlik Erakond	Estonian	CHES	2019	14,487
Finland	Suomen Keskusta	Finnish	BL	2003	1,571

Finland	Vihreän liitto	Finnish	BL	2003	3,909
Finland	Vasemmistoliitto	Finnish	BL	2003	391
Finland	KOKOOMUS	Finnish	BL	2003	3,207
Finland	Social Democratic Party of Finland	Finnish	BL	2003	2,839
Finland	Swedish People's Party of Finland (RKP)	Finnish	BL	2003	3,182
Finland	Perussuomalaiset (PS)	Finnish	BL	2003	2,253
Finland	Kristillisdemokraatit	Finnish	BL	2003	558
Finland	Keskusta	Finnish	CHES	2019	4,683
Finland	Vihreät	Finnish	CHES	2019	3,719
Finland	Left Alliance	Finnish	CHES	2019	5,975
Finland	Kokoomus	Finnish	CHES	2019	2,751
Finland	Social Democratic Party of Finland	Finnish	CHES	2019	35,190
Finland	Swedish People's Party of Finland (SFP)	Swedish	CHES	2019	5,707
Finland	Perussuomalaiset (PS)	Finnish	CHES	2019	35,101
Finland	Kristillisdemokraatit	Finnish	CHES	2019	5,596
Greece	KKE	Greek	CHES	2019	4,444
Greece	ΜέΡΑ25	Greek	CHES	2019	15,671
Greece	Χρυσή Αυγή	Greek	CHES	2019	9,194
Greece	Ελληνική Λύση	Greek	CHES	2019	66,208
Greece	Κίνημα Αλλαγής (Movement for Change)	Greek	CHES	2019	2,625
Greece	Νέα Δημοκρατία	Greek	CHES	2019	3,840
Greece	ΣΥΡΙΖΑ - Συνασπισμός Ριζοσπαστικής Αριστεράς	Greek	CHES	2019	2,479
Hungary	Magyar Demokrata Fórum (MDF)	Hungarian	CHES	2006	15,631
Hungary	Fidesz	Hungarian	CHES	2006	10,391
Hungary	Magyar Szocialista Párt (MSZP)	Hungarian	CHES	2006	9,303
Hungary	Szabad Demokraták Szövetsége (SZDSZ)	Hungarian	CHES	2006	12,620
Hungary	Fidesz	Hungarian	CHES	2010	20,704
Hungary	Jobbik	Hungarian	CHES	2010	52,916
Hungary	LMP - Magyarország Zöld Pártja	Hungarian	CHES	2010	78,933
Hungary	Hungarian Socialist Party (MSZP)	Hungarian	CHES	2010	25,027
Hungary	Fidesz	Hungarian	CHES	2014	3,861
Hungary	Jobbik	Hungarian	CHES	2014	63,792
Hungary	Democratic Coalition	Hungarian	CHES	2014	28,649
Hungary	LMP - Lehet Más a Politika (Politics Can Be Different)	Hungarian	CHES	2014	30,493
Hungary	MSZP	Hungarian	CHES	2014	18,637
Hungary	Együtt 2014-Párbeszéd Magyarországért	Hungarian	CHES	2014	20,601
Iceland	Sjálfstæðisflokkurinn (The Independence Party)	Icelandic	BL	2003	2,317
Iceland	Vinstrihreyfingin grænt framboð	Icelandic	BL	2003	2,684
Iceland	Frjálslyndi flokkurinn	Icelandic	BL	2003	479
Iceland	Framsóknarflokkurinn	Icelandic	BL	2003	1,524
Iceland	Samfylkingin	Icelandic	BL	2003	2,368
Ireland	Fianna Fail	English	LH	1989	28,037
Ireland	Fine Gael	English	LH	1989	21,932
Ireland	Irish Labour Party	English	LH	1989	4,383
Ireland	Progressive Democrats	English	LH	1989	5,550
Ireland	Workers' Party	English	LH	1989	16,581

Netherlands	CDA	Dutch	LH	1989	42,935
Netherlands	D66	Dutch	LH	1989	34,094
Netherlands	PvdA	Dutch	LH	1989	21,945
Netherlands	RPF	Dutch	LH	1989	41,104
Netherlands	GPV	Dutch	LH	1989	22,769
Netherlands	SGP	Dutch	LH	1989	34,273
Netherlands	VVD	Dutch	LH	1989	25,474
Netherlands	CDA	Dutch	BL	2003	23,800
Netherlands	Democraten 66 (D66)	Dutch	BL	2003	4,282
Netherlands	GroenLinks	Dutch	BL	2003	7,974
Netherlands	Lijst Pim Fortuyn (LPF)	Dutch	BL	2003	15,410
Netherlands	PvdA	Dutch	BL	2003	4,773
Netherlands	Socialistische Partij	Dutch	BL	2003	38,185
Netherlands	VVD	Dutch	BL	2003	4,025
Netherlands	CDA	Dutch	CHES	2006	38,385
Netherlands	Democraten 66 (D66)	Dutch	CHES	2006	26,098
Netherlands	Partij voor de Vrijheid (PVV)	Dutch	CHES	2006	1,352
Netherlands	GroenLinks	Dutch	CHES	2006	16,417
Netherlands	PvdA	Dutch	CHES	2006	35,616
Netherlands	Socialistische Partij (SP)	Dutch	CHES	2006	29,887
Netherlands	VVD	Dutch	CHES	2006	7,301
Netherlands	ChristenUnie	Dutch	CHES	2006	35,590
Netherlands	Partij voor de Dieren	Dutch	CHES	2010	25,310
Netherlands	Christen-Democratisch Appèl (CDA)	Dutch	CHES	2010	35,761
Netherlands	Democraten 66 (D66)	Dutch	CHES	2010	49,579
Netherlands	PVV	Dutch	CHES	2010	7,937
Netherlands	GroenLinks	Dutch	CHES	2010	24,259
Netherlands	PvdA	Dutch	CHES	2010	25,444
Netherlands	SGP	Dutch	CHES	2010	32,537
Netherlands	Socialistische Partij	Dutch	CHES	2010	13,491
Netherlands	VVD	Dutch	CHES	2010	25,076
Netherlands	ChristenUnie	Dutch	CHES	2010	48,672
Norway	Senterpartiet	Nynorsk	LH	1989	26,555
Norway	Høyre	Norwegian	LH	1989	48,305
Norway	Arbeiderpartiet (Labour Party)	Norwegian	LH	1989	34,211
Norway	Venstre	Norwegian	LH	1989	22,964
Norway	Fremskrittspartiet	Norwegian	LH	1989	19,606
Norway	Sosialistisk Venstreparti	Norwegian	LH	1989	28,997
Norway	KRF	Norwegian	LH	1989	16,956
Poland	Koalicja Obywatelska (Civic Coalition)	Polish	CHES	2019	17,335
Poland	Konfederacja Wolność i Niepodległość	Polish	CHES	2019	2,182
Poland	Prawo i Sprawiedliwość (PiS)	Polish	CHES	2019	68,028
Poland	Lewica	Polish	CHES	2019	2,393
Poland	RZECZPOSPOLITA OBYWATELSKA	Polish	CHES	2019	7,433
Poland	Wiosna	Polish	CHES	2019	3,682
Portugal	PAN	Portuguese	CHES	2019	56,452

Portugal	PCP	Portuguese	CHES	2019	39,459
Portugal	Partido Ecologista "Os Verdes" (PEV)	Portuguese	CHES	2019	4,087
Portugal	Left Bloc	Portuguese	CHES	2019	57,301
Portugal	CDS-PP	Portuguese	CHES	2019	72,925
Portugal	Partido Social Democrata (PSD)	Portuguese	CHES	2019	54,005
Portugal	Partido Socialista (PS)	Portuguese	CHES	2019	77,045
Slovakia	Ľudová strana – HZDS	Slovak	CHES	2006	8,719
Slovakia	SDKÚ-DS	Slovak	CHES	2006	20,439
Slovakia	SMER – sociálna demokracia	Slovak	CHES	2006	11,776
Slovakia	Strana maďarskej koalície (SMK)	Slovak	CHES	2006	24,179
Slovakia	Slovenská národná strana (SNS)	Slovak	CHES	2006	27,497
Slovakia	KDH	Slovak	CHES	2006	9,198
Slovakia	MOST-HÍD	Slovak	CHES	2010	15,420
Slovakia	SDKÚ-DS	Slovak	CHES	2010	14,914
Slovakia	SMER – sociálna demokracia	Slovak	CHES	2010	16,404
Slovakia	Sloboda a Solidarita (SaS)	Slovak	CHES	2010	24,309
Slovakia	Slovak National Party (SNS)	Slovak	CHES	2010	14,713
Slovakia	Kresťanskodemokratické hnutie (KDH)	Slovak	CHES	2010	18,411
Slovenia	Stranka mladih - Zeleni Evrope	Slovenian	CHES	2014	3,940
Slovenia	Zavezništvo Alenke Bratušek	Slovenian	CHES	2014	8,365
Slovenia	DeSUS	Slovenian	CHES	2014	6,105
Slovenia	Slovenska demokratska stranka (SDS)	Slovenian	CHES	2014	6,751
Slovenia	Social Democrats (SD)	Slovenian	CHES	2014	2,762
Slovenia	Združena levica	Slovenian	CHES	2014	7,697
Slovenia	Nova Slovenija - Krščanski demokrati	Slovenian	CHES	2014	50,823
Spain	CDS	Spanish	LH	1989	56,560
Spain	Convergencia i Unió (CiU)	Spanish	LH	1989	48,913
Spain	Partido Popular	Spanish	LH	1989	44,465
Spain	Partido Socialista Obrero Español (PSOE)	Spanish	LH	1989	26,396
Spain	Izquierda Unida (IU)	Spanish	LH	1989	43,122
Sweden	Centerpartiet	Swedish	CHES	2006	4,503
Sweden	Miljöpartiet de Gröna	Swedish	CHES	2006	1,847
Sweden	Vansterpartiet (Left Party)	Swedish	CHES	2006	3,638
Sweden	Folkpartiet liberalerna	Swedish	CHES	2006	8,288
Sweden	Allians för Sverige	Swedish	CHES	2006	11,739
Sweden	Socialdemokraterna	Swedish	CHES	2006	4,336
Sweden	Kristdemokraterna	Swedish	CHES	2006	1,283
Sweden	Centerpartiet	Swedish	CHES	2010	7,525
Sweden	Miljöpartiet de Gröna	Swedish	CHES	2010	2,395
Sweden	Vansterpartiet	Swedish	CHES	2010	2,768
Sweden	Folkpartiet	Swedish	CHES	2010	10,500
Sweden	Alliansen	Swedish	CHES	2010	17,794
Sweden	Socialdemokraterna	Swedish	CHES	2010	2,788
Sweden	Sweden Democrats	Swedish	CHES	2010	1,860
Sweden	Kristdemokraterna	Swedish	CHES	2010	6,281
Sweden	Centerpartiet	Swedish	CHES	2014	6,152

Sweden	Miljöpartiet de Gröna	Swedish	CHES	2014	7,048
Sweden	Vansterpartiet	Swedish	CHES	2014	5,529
Sweden	Folkpartiet liberalerna	Swedish	CHES	2014	5,604
Sweden	Alliansen	Swedish	CHES	2014	26,767
Sweden	Socialdemokraterna	Swedish	CHES	2014	7,222
Sweden	Sweden Democrats	Swedish	CHES	2014	5,881
Sweden	Kristdemokraterna	Swedish	CHES	2014	6,562
UK	Conservative Party	English	CHES	2019	21,446
UK	Green Party	English	CHES	2019	25,742
UK	Labour	English	CHES	2019	27,006
UK	Liberal Democrats	English	CHES	2019	29,199
UK	Plaid Cymru	English	CHES	2019	19,454
UK	Scottish National Party	English	CHES	2019	23,320
UK	British National Party	English	KB	2010	33,661
UK	Conservative Party	English	KB	2010	31,429
UK	Greens	English	KB	2010	22,719
UK	Labour Party	English	KB	2010	33,448
UK	Liberal Democrats	English	KB	2010	32,459
UK	Plaid Cymru	English	KB	2010	7,704
UK	Scottish National Party	English	KB	2010	9,472
UK	UK Independence Party	English	KB	2010	10,517

Appendix B: MP Scale Definitions

These scales are drawn from Lowe et al. (2011), who defined equivalent scales from MP categories that we used in computing the dimensional saliences shown in Figure 1 and Table 4. These are:

Table B.1: Dimension Scales from the MP Dataset

Dimension	Lowe et al. (2011) Label	Manifesto Categories
Economic	Planned v. Market Economy	403, 404, 406, 412, 413, 504, 506, 701, 401, 402, 407, 414, 505
Social	Social Liberal-Conservative	103, 105, 106, 107, 202, 604 104, 201, 203, 305, 601, 603, 605, 606
Environment	Environmental Protection	501, 416, 410
Immigration	(not defined)	601.2, 602.2
EU	EU	108, 110
Decentralization	Decentralisation	301, 302

Appendix C: System prompts

The system prompt for every case began with: "You are a political scientist trained in manifesto analysis. Judge relative importance strictly from the provided manifesto text. Do NOT use outside knowledge. Use the full scale where justified. The policy issues are defined as: [insert policy labels and definitions]."

We then followed this with scale prompt. The direct scale prompt was: Rate the importance of each issue on a 0-10 scale. 0 = completely absent or irrelevant; 2-3 = mentioned only briefly or tangentially; 5 = moderate importance, discussed but not central; 7-8 = substantial importance, given sustained attention; 10 = top priority, central theme of the manifesto."

The prompt for rank ordering issue saliences was as follows. "First, assign a complete ranking of all issues, from 1 (most important) to 7 (least important). Each issue must have a unique rank. Then, provide a 1–2 sentence justification per issue, citing evidence from the text. Weak or absent evidence should result in a lower rank, with the justification noting the lack of discussion."

The prompt for imposing a saliency budget was as follows. "Distribute importance across the issues as a set of weights between 0.0 and 1.0. Rules: Each issue must receive a numeric score in the range [0.0, 1.0]; a score of 0.0 means the issue is completely absent or irrelevant in the manifesto; higher scores mean more relative importance; the more central an issue is to the manifesto, the larger its weight; all issue weights must sum to exactly 1.0 across the six issues. Also provide a brief justification for each issue's score."

The prompt asking for pairwise comparison was as follows. "When presented with pairs of the policy issues, select which is more important in the manifesto. No ties are allowed. Evidence each decision with a brief justification, with reference to the text. If evidence is weak, choose the best-supported option and explain the uncertainty."

Appendix D: Correlation Analysis by Manifesto Length

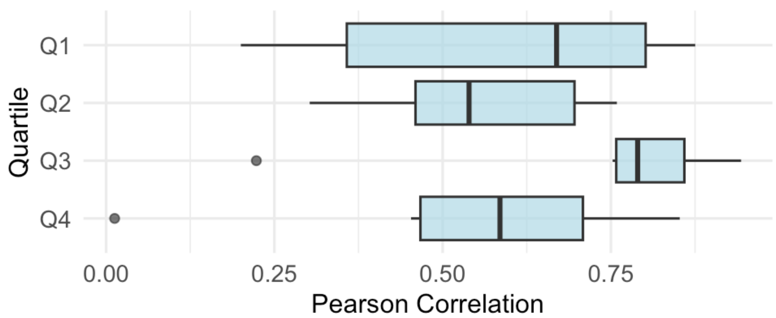
To test whether manifesto length was related to the ability of the LLMs to track expert survey values, we stratified the manifestos into length quartiles and computed the correlations separately. If longer manifestos imposed less of a constraint on signalling relative importance, then correlations might decrease with longer manifestos.

In Table D.1 and Figure D.1, we compare the expert-budget scores with the LLM budget values, for those for which at least three dimensions were scored by the experts (just as in the manuscript). Each quartile contained 18 manifestos matching this criterion. There is no evidence of a relationship between manifesto length and the level of the LLM scores of importance and the expert surveys.

Table D.1. Correlations by length quartile. Each quartile contains 18 manifestos.

	Pearson Correlations			
Dimension	**Q1**	**Q2**	**Q3**	**Q4**
Economic	0.80	0.59	0.88	0.45
Social	0.30	0.30	0.22	0.01
Immigration	0.88	0.76	0.77	0.72
EU	0.20	0.49	0.75	0.66
Environment	0.54	0.73	0.94	0.85
Decentralization	0.80	0.45	0.81	0.51
Mean	0.59	0.55	0.73	0.54
Median	0.67	0.54	0.79	0.59
Minimum	0.20	0.30	0.22	0.01
Maximum	0.88	0.76	0.94	0.85

Figure D.1. Overall correlations by quartile of manifesto length.



ENDNOTES

[1] This depends on a psychological assumption, widely implicit in models of party competition in political science, of *transitive* preferences. If I prefer A to B, then adding C to the choice set does not now make me prefer B to A.
[2] Using verbatim text from the LLM prompts
[3] As opposed to issue positions.
[4] https://manifesto-project.wzb.eu/information/documents/corpus
[5] Tokens include all numbers, punctuation, and special characters, as counted by the quanteda package (Benoit, et al. 2018). The token count from LLMs is higher, as they count sub-word elements as tokens, to the extent that the count from the Gemini LLM, for instance, is approximately 1.3 to1.4 times greater than those listed in Table 1.
[6] The exception was decentralization policy, a matter to which we rerun when discussing LLM estimates of issue salience.
[7] See Google (June2025). Gemini 2.5 Pro: Technical Report. Google AI
[8] Gemini was able to summon these facts about planetary defence: Europe contributes substantively through its budget for the European Space Agency, which pairs its Space Safety (SSA) infrastructure and the NEO Coordination Centre in Italy, coordinating threat detection and risk analysis capacities. While no European country publishes a standalone line item for planetary defence, ESA's total budget for 2024–25 is approximately €7.7–7.8 billion per year. Within that, the Space Safety (formerly Space Situational Awareness) programme—including Planetary Defence—comprises around 3.2–3.5% of the total, meaning ESA spends in the region of €250–270 million annually on this constellation of safety initiatives.
[9] $n = 72$
[10] The Fisher z-average summarizes the correlations across the six issue dimensions by first transforming each correlation to Fisher's z scale, averaging these values, and then back-transforming to provide a single overall measure of association.
[11] Less so for environmental policy. It bears repeating that we put "over-estimate" in quotes here because we do not know the "true" values of the fundamentally unobservable latent quantities of interest.
[12] Benoit *et al.* (2026) provide a careful cost comparison of using LLMs and humans for the same task in Appendix G of their paper. They find, unsurprisingly, that LLMs are several orders of magnitude cheaper.