Building policy scales from manifesto data: A referential content validity approach

Christopher Prosser

University of Oxford, Oxford OX1 3UJ, United Kingdom

ABSTRACT

Researchers using scales based on MRG/CMP/MARPOR's manifesto dataset face a bewildering array of different scales. The validation of these scales has tended to focus on external, convergent validity. The actual content of these scales has received less attention and the choice of the manifesto components which make up these scales has often been conducted by either opaque or questionable methods. This article develops a critique of existing methods of component selection and proposes a new method of component selection based on the covariance of components with 'naïve' provisional scales, which are refined in an iterative process. It uses this method to construct a set of comparable one (general left–right) and two (economic and social) dimensional scales – filling a gap in the existing body of scales that will allow researchers to compare dimensionality across models without inadvertently comparing different assumptions that underlie the construction of the scales.

© 2014 Elsevier Ltd. All rights reserved.

Determining where a party sits on a policy issue or an ideological dimension is one of the most fundamental questions of measurement in political science. Whether the end research goal is one of how parties' positions influence electoral choices, intergovernmental negotiations, or budget overruns (to name but a tiny fraction of the ways party positions have been used in the literature), scholars must first determine what the relevant policy dimensions are, and then, where parties sit on them. There have been myriad answers to these basic questions which broadly draw on four methodological approaches – expert surveys, mass surveys, roll call votes, and electoral manifesto data. All of these approaches have their benefits and costs, and in an ideal world, many approaches would be used to ensure cross-validation of empirical findings. Unfortunately however, the world is far from ideal, particularly if we wish to analyse politics across a wide time span: we cannot travel back in time and conduct expert and mass surveys, and the existing surveys do not cover a long enough periods nor are

they conducted at regular enough intervals to track fine grained developments in multiple countries. Scholars who wish to tackle these questions currently only have one option, to use manifesto data, specifically that collected by the group currently known as the Manifesto Project on Political Representation (MARPOR) (Budge et al., 2001; Klingemann et al., 2006; Volkens et al., 2009).[1]

Although a certain amount of methodological hand-wringing seems to be *de rigueur* amongst scholars who use them, the general validity of scales derived from MARPOR data has now been well established. MARPOR have endeavoured to show the validity of manifesto scales – particularly their 'Rile' scale – across many different validity criteria (Budge et al., 2001; Klingemann et al., 2006). Many authors however, particularly critics of MARPOR's scaling methods,

---

[1] Depending on the date, this group has alternatively been called the Manifesto Research Group (MRG), Comparative Manifesto Project (CMP), or most recently, the Manifesto Project on Political Representation (MARPOR). MARPOR is used as a generic term to cover all of the incarnations of the project.

have tended to focus only on convergent validity, tested by cross validation with scales derived from other techniques (Bakker and Hobolt, 2013; Budge, 2001; Budge et al., 2001; Klingemann et al., 2006; Lowe et al., 2011; Ray, 2007).

Although a bewildering number of scales have been constructed from MARPOR data, their convergent validity is often tested against the same expert survey data – most commonly the Laver and Hunt/Benoit and Laver and Chapel Hill group expert surveys (Benoit and Laver, 2007; Hooghe et al., 2010; Laver and Hunt, 1992; Steenbergen and Marks, 2007). Although some authors seem to treat them as such in their validation approaches, expert surveys do not represent 'true' party positions, but rather estimates of these positions with their own strengths and weaknesses (Budge, 2000; Steenbergen and Marks, 2007; Volkens, 2007), and so their use in discriminating between competing manifesto scales is problematic. Even leaving these debates to one side, the fact that many different scales, composed of different manifesto components and combined by different methods, all apparently perform equally well in terms of convergent validity suggests that although cross-validation is an essential method of establishing the general validity of different scales, it is perhaps less useful as a technique for choosing between them.

That different scales have been found to have essentially equal convergent validity is not completely surprising given the large overlap in the components used to construct them, as demonstrated by the three examples of 'left–right' scales shown in Table 1, and the two 'social' scales shown in Table 2. That each of these scales has been cross-validated suggests that scholars must turn to other measures of validity in order to choose between them.

Researchers interested in questions relating to the dimensionality of politics also face another concern – the various sets of scales on offer tend to treat the underlying dimensionality of politics as *either* unidimensional or multidimensional. MARPOR's Rile scale implies a single 'super dimension' ranging from liberal/state control of the economy/left to conservative/free market/right. Many other manifesto scales approaches have also adopted a one-dimensional approach (Franzmann and Kaiser, 2006; Gabel and Huber, 2000). One dimensional scales are prevalent in many streams of the political science literature, and not only when using scales from manifesto data. However, whenever scholars set out to look for multiple ideological dimensions, they tend to find them (Bakker et al., 2012; Henjak, 2010; Schofield, 1993; Stoll, 2010a; Warwick, 2002).

Although the descriptions given to these dimensions varies from scholar to scholar a consistent finding has been the existence of an economic dimension, consisting of issues surrounding markets, labour, economic planning and regulation, and a social dimension, consisting of issues surrounding personal freedoms, human rights, morality, and traditions.[2]

Although the existing evidence supports a multidimensional interpretation of political space, the approach here remains agnostic to this question and recognises that the use of one or two dimensions in research may be driven by many reasons – from the appropriateness of multiple dimensions in different contexts to simple matters of pragmatisms surrounding the availability of other data with which researchers wish to compare their data (left–right placements in survey research tend to be unidimensional for example).

On some questions, rather than taking an a priori stance on the question of dimensionality, scholars may wish to compare the results of alternately specifying uni- or multi-dimensional models investigating party competition – for example whether party positions on a particular policy are best explained by a one or two dimensional conception of political space. A problem facing such research questions is that scales using the MARPOR data have tended to be developed using either a uni- or multi-dimensional framework and there is no one set of scales that contains both uni- and multi-dimensional scales that shares an underlying method of component choice. MARPOR's venerable 'Rile' scale may well be appropriate for a unidimensional approach, or the Benoit and Laver (2007) state intervention in the economy and social liberal-conservative scales for a multidimensional approach, but comparing the Rile scale to the Benoit and laver scales engenders not just a comparison of dimensionality, but also of different underlying assumptions about which components should be included and excluded in a scale. The results of any analysis based on such a comparison may occur not as a result of any true underlying relationship but because of different decisions about which components should be included or excluded from a scale. The aim of the present article is to produce a set of one and two dimensional scales that are uniquely comparable across conceptions of dimensionality because they are based on the same assumptions and selection method, though each scale is of course suitable for standalone use as well.

Any scholar wishing to develop MARPOR scales faces three questions: What components should be used to construct the scale? How should they be combined? Are the scales valid measures of what they are supposed to measure? Although these questions suggest a natural order of enquiry, the existing literature has most satisfactorily answered them in reverse. As has already been discussed the general validity of MARPOR scales is now well established. Similarly, several different methods of combining components have been developed and assessed (Kim and Fording, 1998; Lowe et al., 2011; Ray, 2007). However existing approaches to component selection – or in other words, the content validity of scales – are less well developed and are often conducted on an ad hoc basis. This article argues that scholars should think carefully about the ways in which different methods of counting manifesto components affect the end result of the scale, and how individual components should relate to the overall scale. It then develops a novel method for selecting components based on their exogenous correlation with provisional overall scales and uses it to construct a set of new scales.

---

[2] Common names for the economic dimension include 'left-right', and 'materialist'. Common names for the social dimension are 'liberal-conservative', 'libertarian-authoritarian', 'post-materialist', 'Green/Alternative/Libertarian-Traditional/Authoritarian/Nationalist'. Here they will be called the 'Economic Left-Right', and 'Social Liberal-Conservative' dimensions.