Estimating Policy Positions from Political Texts

Michael Laver Trinity College Dublin
John Garry Trinity College Dublin

The analysis of policy-based party competition will not make serious progress beyond the constraints of (a) the unitary actor assumption and (b) a static approach to analyzing party competition between elections until a method is available for deriving reliable and valid time-series estimates of the policy positions of large numbers of political actors. Retrospective estimation of these positions in past party systems will require a method for estimating policy positions from political texts.

Previous hand-coding content analysis schemes deal with policy emphasis rather than policy positions. We propose a new hand-coding scheme for policy positions, together with a new English language computer-coding scheme that is compatible with this. We apply both schemes to party manifestos from Britain and Ireland in 1992 and 1997 and cross validate the resulting estimates with those derived from quite independent expert surveys and with previous manifesto analyses.

There is a high degree of cross validation between coding methods, including computer coding. This implies that it is indeed possible to use computer-coded content analysis to derive reliable and valid estimates of policy positions from political texts. This will allow vast volumes of text to be coded, including texts generated by individuals and other internal party actors, allowing the empirical elaboration of dynamic rather than static models of party competition that move beyond the unitary actor assumption.

Deriving reliable and valid estimates of the policy positions of key actors is fundamental to the analysis of political competition. Various systematic methods have been used to do this, including surveys of voters, politicians, and political scientists, and the content analysis of policy documents. Each method has advantages and disadvantages but, for both theoretical and pragmatic reasons, policy documents represent a core source of information about the policy positions of political actors.

We explore various ways to extract information about policy positions from political texts. We are particularly interested in using computer-coding techniques to derive reliable and valid estimates of the policy positions of political actors. This is not mere laziness on our part, a lack of stomach for the hard graft of expert coding. If analyses of party competition are to move beyond both static models and a view of political parties as unitary actors, this requires information on the policy positions of actors inside political parties and on the development of these over time and between elections. The laborious expert "hand-coding" of text is simply not a viable method for estimating the policy positions of huge numbers of political actors, for example, all members of a legislature. Any serious attempt to operationalize a model of internal party policy competition, or of dynamic policy-based party competition or coalition government between elections, implies using computer-coding for estimating the policy positions of key political actors.

We first review existing methods for estimating policy positions from political texts. These have for the most part concentrated on the expert coding of party manifestos. We then suggest ways to improve these, dealing with both expert- and computer-coded content analysis. We then explore the impact of our suggestions upon estimates of party policy positions derived from British and Irish manifestos issued during the 1992 and 1997 general elections in each country, positions for which a range of

Michael Laver is a Professor of Political Science and Director of the Policy Institute, Trinity College, Dublin 2, Ireland (mlaver@tcd.ie). John Garry is a Ph.D. student of Political Science, Trinity College, Dublin 2, Ireland (garryj@tcd.ie).

Versions of this paper have been presented at the ECPR workshop on "Empirical rational choice theory," Warwick, April 1998 and the ECPR workshop on "Estimating the policy positions of political actors," Mannheim, March 1999. The authors are grateful to Ken Benoit, Ian Budge, Miranda de Vries, Matt Gabel, Daniela Giannetti, John Huber, Jan Kleinnijenhuis, Michael Marsh, Michael McDonald, Leonard Ray, and many other participants at these conferences and workshops, as well as three anonymous journal referees for their helpful and constructive comments.

American Journal of Political Science, Vol. 44, No. 3, July 2000, Pp. 619–634

independent estimates are available. The results suggest that the computer coding of huge volumes of virgin text may be a viable undertaking, with obvious implications for dynamic analyses of party competition that go beyond the unitary actor assumption. We conclude with suggestions for the refinement of this approach.

Analysing Party Manifestos: The Story So Far

Party manifestos are strategic documents written by politically sophisticated party elites with many different objectives in mind. This leaves considerable scope for debate about whether party manifestos reflect the "real" positions of the parties that publish them. In our view this debate is ultimately fruitless, however, since the "real" policy position of a political actor is a fundamentally elusive, even metaphysical, notion. All we can do in practice is use evidence about policy positions in particular political contexts and make context-specific inferences from this. In this sense we keep our feet on firm ground when we study official party documents published during election campaigns. As an official document, it will be difficult for party members to resile from policies in the party manifesto, while party leaders can be charged with failure to implement published manifesto pledges when given the chance to do so. Furthermore, manifestos are typically issued by each party at each election for most of the post-war period. Manifestos thus provide historical evidence of the movement of party policy positions over time. Regardless of the merits of different methods of estimating *contemporary* party policy, manifestos offer an unparalleled way to retrieve data on party policy *in the past.*

The Manifesto Research Group Project

The Manifesto Research Group (MRG) is by far the biggest show on the road as a source of data on party manifestos. The MRG developed its own coding scheme and used this to analyse nearly all manifestos of nearly all political parties contesting nearly all elections in most post-war parliamentary democracies. This involved expert coders, fluent in the language concerned, reading each manifesto sentence by sentence and allocating each sentence to a category in the coding scheme. The project has been running for about 20 years and has acquired near-monopoly status in the field, for the very obvious reason that any attempt to redo such an analysis seems a truly Herculean task. The original motivation of the MRG, however, was to operationalize a specific model of party competition, which assumes that parties compete in terms of the *salience* of particular issues in the policy package they put to voters. Whatever subsequent use has been made of their data, MRG researchers set out to measure the *relative emphasis* placed on an issue by a party in a manifesto, not the party's *substantive position* on this issue.

Position and emphasis are quite distinct parameters of party policy. Two parties may have *quite different substantive positions on the same issue, but emphasise this issue to precisely the same extent* in their respective manifestos. Recent expressions of saliency theory do assert a strong relationship between party position on, and party emphasis of, an issue—and even that "emphases equal direction" in a particularly forthright statement of the model (Budge, 1999). This, however, is acknowledged to be an empirical proposition to be tested as part of the evaluation of saliency theory. Testing the proposition, furthermore, requires independent estimates of direction and emphasis, rather than an indicator that conflates the two.

While the analytical distinction between substantive position on a policy dimension and the emphasis given to this might seem clear-cut, the situation is more complicated in practice. The great scarcity of time-series data on party policy has created a determination to squeeze the MRG data until they yield useful information on policy positions. Baron (1991), Schofield (1993), and Warwick (1994), among a wide range of authors, explored models of coalition politics using empirical policy spaces derived from the MRG data, on the clear if implicit assumption that these could be used to estimate party policy positions. Laver and Budge (1992, chapter 2) made a more explicit attempt to extract positional information from of an extensive reanalysis of the MRG data. They used a priori reasoning supplemented by exploratory factor analyses to identify clusters of closely interrelated coding categories which they felt were defined in such a way as to convey information about substantive policy positions. The raw variables making up these apparently more positional clusters of coding categories were then aggregated and used as building blocks in the construction of a general left-right scale that Laver and Budge considered to have good face validity. A different and fundamentally inductive version of this approach has recently been proposed by Gabel and Huber (2000), who do not make any a priori assumption about which policy categories are associated with left-right ideology. They use principal factors analysis on the MRG data to extract the first factor for each country. They interpret this as being, by definition, the main left-right dimension in the country concerned and derive regression scores for each manifesto on this.

The justification for this work is that the MRG data already exist and comprise a huge source of information about at least something to do with party policy. Some MRG coding categories do deal in a unipolar way with positional issues: "nationalisation," for example, or "law and order." In some of these cases emphasis may in practice imply position. Few who speak of "law and order," for example, advocate less law and order. Other MRG categories are bipolar and convey more explicitly positional information. Examples are "social services expansion: positive" and "social services expansion: negative." or "decentralisation: positive" and "decentralisation: negative." The MRG coding scheme does not systematically use bipolar categories, however. As we shall see, and as Gabel and Huber (2000) also show, the existence of some positional categories is why reanalysis of the raw MRG data does retrieve positional information on some aspects of party positions.

The Party Change Project

The content of party manifestos was used in a quite different way by the researchers of the Harmel-Janda "Party Change Project" (PCP), explicitly designed to extract policy positions from party manifestos. The PCP defined a set of nineteen issues of interest on a priori grounds to the researchers. Manifestos were then used as follows to identify positions on each of these. "After identifying, gathering, and carefully reading all of a manifesto's passages relevant to a given issue, coders then assigned the numerical code [on a +5 to –5 scale] which, in their best judgment, best reflected the overall content of these statements." (Harmel, Janda, and Tan, 1995, 7). In effect, the PCP data generation process was like a highly structured expert survey (see below). Instead of asking many experts to locate parties in general terms on particular policy dimensions, at most three experts were given detailed coding instructions and asked to use a party's manifesto to locate it on each policy dimension.[1]

The data generated by this process are explicitly concerned with party policy positions and enable more valid estimates *on the nineteen policy scales under investigation* than could be constructed from the MRG data on policy emphases. However, we must be content with the nineteen scales defined by the PCP, a problem with all expert survey techniques. The PCP content analyses thus cannot be encyclopaedic descriptions of party policy, nor can they chart the rise to prominence of hitherto unimportant issues. In common with expert survey techniques, furthermore, the PCP judgments are more explicitly subjective than the basic coding decision of determining whether a particular sentence is in or out of a particular coding category. It seems likely that the PCP's expert coders would have found it much more difficult to separate their subjective placement of parties on scales from their prior knowledge of policy profiles of the parties concerned.[2] Unlike expert surveys, however, which quite explicitly rely *only* on the prior knowledge of experts, the PCP data are not the result of averaging subjective judgments across number of experts. The estimates derived from the expert surveys used below, for example, are based upon between 30 and 110 expert judgments.

Expert Surveys

It may seem odd to include expert surveys in a review of methods for analysing party manifestos but, as the PCP illustrates, there is a continuum of techniques based upon expert judgments. At one end, the MRG used expert coders to analyse a manifesto on a sentence by sentence basis. In the middle is the PCP technique of using expert coders to identify substantive party positions at one of eleven points on each of nineteen issue dimensions, having read the manifesto as a whole.[3] At the other end is the expert survey technique of having experts locate parties at substantive positions on one (Castles and Mair, 1984; Huber and Inglehart, 1995) or more (Laver and Hunt, 1992) policy or ideological dimension(s), taking account of everything they think might be relevant. "Everything" presumably includes the direct and indirect impact of manifestos as well as many other things besides. However, it may also include aspects of observed behaviour (for example, coalition formation) that the data are then used to explain.

Expert surveys thus have the virtue, shared with the PCP technique, of generating unequivocally positional estimates of policy on well-defined dimensions. Since the experts are not required to study manifestos and explicitly justify every judgment they make, the expert survey technique imposes far fewer costs, allowing far more people to be consulted. The expert placement on scales "taking everything into consideration" is, however, obviously less explicit than the PCP technique. The big disadvantages of expert surveys relative to text-based coding

[1]The small number of coders and the detailed coding instructions thus make this process look much more like content analysis than an expert survey.

[2]Indeed, and almost paradoxically, it might be better for PCP-style analyses to use coders who were political scientists knowing little of the country concerned to read manifestos and allocate parties to scale positions, to ensure that prior knowledge of party positions did not color these judgments.

[3] In effect the text unit to be coded is the entire manifesto.

are, first, that text-based techniques are far more valid for the estimation of a historical party positions and, second, that a given text can typically be located at a precise time point so that a time line of cause and effect can be more confidently established.

Expert Coding of Text on Policy Positions

There are two quite distinct parts of the process of estimating policy positions from political texts; some past confusions have arisen from considering both together. First, there is a process of data reduction in which a large and complex text is reduced in a reliable way to a smaller and simpler set of coded data. This can be done by either expert coders or computers and has three stages: the design of a coding scheme; the definition of a text unit to be coded; and the coding of real text units. Second, there is a data manipulation process, during which raw data are processed into variables that give valid estimates of party policy positions. Data manipulation can equally be applied to data collected using new methods and to the reanalysis of existing data such as those generated by the MRG. We begin by considering data reduction, first in relation to expert coding and then in relation to computer coding. Then we consider the estimation of policy positions from coded texts.

A New Expert-Coding Scheme for Party Policy Positions

In the original MRG analysis, the coding scheme transformed a text into a set of sentence counts for fifty-four coding categories. No matter how long and complex the text, it was reduced in coded form to a case in a dataset with observations on fifty-four variables.[4] In the PCP analysis, data reduction used the project's coding scheme to transform the text into scores on nineteen policy scales. No matter how long and complex the text, it was reduced to a case in a dataset with observations on nineteen variables.[5] Given the complexity of the document being analysed and the skill and effort required from each expert coder, the type of raw data set produced by each approach is very coarse-grained. This is particularly the case for the PCP project, which codes data directly at the level of the scales to be estimated. There is no possibility to disaggregate these and subsequently recombine them into new policy scales. This strongly implies the need for collecting raw data using as fine-grained a coding scheme as is consistent with reliable expert coding, leaving the definition and estimation of specific policy scales explicitly to the data manipulation stage. Data generated by a fine-grained coding scheme are also far more useful for the political science community as a whole, allowing more flexible reanalysis for purposes that go beyond anything the original researchers had in mind.

A further issue is that the MRG coding scheme, as we have seen, does not consistently use bipolar coding categories. It seems to us to be axiomatic that any content analysis coding scheme designed to extract substantive information on policy *positions* should use coding categories that are at least bipolar. It is probably more useful, indeed, to ensure that all policy concerns can be coded in a *tripolar* way. This allows any mention in a manifesto to register some concern for the category involved, while all mentions can be coded into those that are pro some well-defined policy position, those that are con, and those that are neutral on it. Those whose theoretical concerns are with the emphasis attached to policy dimensions rather than positions on them can, of course, easily collapse all codings relating to a given policy category into a single variable. Those whose concerns are with policy positions cannot disaggregate data coded at the level of policy emphasis to retrieve positional information. (Indeed, to extract positional information from data that deal only with policy emphasis it is necessary to *assume* the validity of saliency theory). We therefore developed a new coding scheme for the content analysis of political texts, in which no policy category is defined without defining its antithesis, as well as a neutral position between the two. The substantive policy areas covered by the new scheme build on those of the MRG scheme but are considerably more comprehensive and fine-grained than these.[6]

To ensure coherence and systematic coverage of potential policy spaces, the new scheme is hierarchically structured, something that is also axiomatic in any text-coding scheme designed to extract information about policy. At the highest level in the hierarchy, we defined a set of nodes representing broad policy "domains." These are the economy, the political system, the social system, external relations, and a "general" domain that has to do with the cut and thrust of specific party competition, as

[4]The bipolar nature of many of the coding categories, alluded to above, meant that far fewer than fifty-four different policy *concerns* were in fact picked up by the MRG analysis.

[5]In neither project was any attention paid to the *sequence* in which references to the various coding categories appeared in the text.

[6]The full version of this scheme can be accessed via *http://www.politics.tcd.ie/personnel/staff/laver.html*.

TABLE 1 Abridged Section of Revised Manifesto Coding Scheme

1	ECONOMY Role of state in economy	
1 1	ECONOMY/+State+ Increase role of state	
1 1 1	ECONOMY/+State+/Budget Budget	
1 1 1 1	ECONOMY/+State+/Budget/Spending Increase public spending	
1 1 1 1 1	ECONOMY/+State+/Budget/Spending/Health	
1 1 1 1 2	ECONOMY/+State+/Budget/Spending/Educ. and training	
1 1 1 1 3	ECONOMY/+State+/Budget/Spending/Housing	
1 1 1 1 4	ECONOMY/+State+/Budget/Spending/Transport	
1 1 1 1 5	ECONOMY/+State+/Budget/Spending/Infrastructure	
1 1 1 1 6	ECONOMY/+State+/Budget/Spending/Welfare	
1 1 1 1 7	ECONOMY/+State+/Budget/Spending/Police	
1 1 1 1 8	ECONOMY/+State+/Budget/Spending/Defense	
1 1 1 1 9	ECONOMY/+State+/Budget/Spending/Culture	
1 1 1 2	ECONOMY/+State+/Budget/Taxes Increase taxes	
1 1 1 2 1	ECONOMY/+State+/Budget/Taxes/Income	
1 1 1 2 2	ECONOMY/+State+/Budget/Taxes/Payroll	
1 1 1 2 3	ECONOMY/+State+/Budget/Taxes/Company	
1 1 1 2 4	ECONOMY/+State+/Budget/Taxes/Sales	
1 1 1 2 5	ECONOMY/+State+/Budget/Taxes/Capital	
1 1 1 2 6	ECONOMY/+State+/Budget/Taxes/Capital gains	
1 1 1 3	ECONOMY/+State+/Budget/Deficit Increase budget deficit	
1 1 1 3 1	ECONOMY/+State+/Budget/Deficit/Borrow	
1 1 1 3 2	ECONOMY/+State+/Budget/Deficit/Inflation	

well as uncodable pap and waffle. Within the economic domain, the coding scheme then has four branches: to increase the role of the state in the economy; to reduce the role of the state in the economy; to be neutral on the role of the state in the economy; and to display a general concern with economic growth. Within each of the three broad policy stances on the role of the state in the economy, the coding scheme branches deal with four very general ways in which the state can intervene in the economy: the state budget, state ownership of industry and services, state regulation, and direct action by the state. Within the state budget, policy could relate to spending, taxation, or the deficit. Taxation policy can relate to income taxes, sales taxes, capital taxes, and so on. Table 1 shows an abridged section of part of the new scheme dealing with this area. Other policy domains are spanned hierarchically in the same systematic way.

There is no reason to regard this scheme as being fixed for all time. While deleting branches from its hierarchical structure might cause problems of comparison between newly coded documents and those coded before, adding new branches to suit particular local or temporal circumstances presents no problem at all. The beauty of an hierarchically structured coding scheme is that, if perfect comparability is required between a "parent" coding scheme and one that has been expanded, it is always easy to collapse the expanded scheme back to its parent.[7]

The coding scheme we propose has over 300 categories. It is thus far more fine-grained than schemes used by either the MRG or the PCP, but even so its hierarchical structure considerably simplifies expert coding. For each text unit, coding involves a sequence of straightforward decisions. Does the text unit deal with the economy, the political system, external relations, etc.? Does it deal with the budget, ownership, or regulation? Does it deal with spending, taxation, or budget deficit? Does it deal with spending on housing, education, or health? Pre-testing did not throw up particular problems for coders using this hierarchical decision-making process, each level of which is actually more straightforward than coding into the less structured fifty-four category MRG coding scheme.

Text Units to Be Coded

The MRG used manifesto "quasi sentences" as its fundamental unit of analysis. A quasi sentence is a word string that is either a complete sentence or a part sentence that could have been a complete sentence if the writer had chosen to make it so. It might seem that complete sentences should be used, since these occur unambiguously between particular punctuation marks. But this would put the analyst at the mercy of the writing style of the manifesto author(s). Comparing two manifestos, the first might appear to give more weight to some topic than the second merely because the author of the former used shorter sentences, triggering more "hits" for a text passage of the same length. The big disadvantage of this approach is that the definition of a "quasi-sentence" might itself be a source of unreliability. Accordingly, we have chosen to use words as the unit of analysis or, more precisely, word strings with an average length of ten words. There are two reasons to do this. At a practical level, it is very time-consuming to code individual words of lengthy texts without giving anything like ten times the payoff of coding word strings with an average length of ten words. At a methodological level, the coder has to read and interpret the text in context, and this is not something reliably done one word at a time.[8]

Coding Conventions

The MRG coders assigned every text unit to one and only one coding category, a convention that derived from an interest in the saliencies of policy concerns, which were estimated as the relative proportions of quasi-sentences assigned to each coding category. Since no text unit was coded into more than one category, these relative proportions always added up to 1.00. We follow the same convention in the present analysis to allow our results to be compared with those of the MRG, but we are not wedded to it as a general principle. Multiple coding of text units might well be appropriate in content analyses designed to extract policy positions on a range of dimensions; nothing intrinsic to our scheme precludes multiple coding.

A further important coding convention is that text units prima facie "neutral" on a particular policy concern are coded in context. Thus if a "neutral" text unit is embedded in a paragraph that otherwise expresses a "pro" position on some policy concern, it is coded pro. If it is embedded in a paragraph that otherwise expresses a "con" position, it is coded con.[9] Just as it makes no sense to code every occurrence of the word "the" as being neutral in the grounds that it conveys no information about a policy position, it makes no sense to take a string of ten words out of context and, because they convey no policy meaning as they stand, code them as being neutral.

Computer Coding of Text on Party Policy Positions

Comparing Expert and Computer Coding

A radical alternative to the "qualitative" expert coding of text is to use "quantitative" content analysis. Quantitative techniques use a computer to allocate text units to a coding scheme that is closely analogous to an expert-coding scheme. Expert coding uses the subjective judgment of a human coder to allocate texts units and can therefore take greater account of their substantive con-

[7]In addition, provided that appropriate computer software is used to manage the coded policy documents, it is a straightforward matter to identify passages in previously analysed documents that have been coded into nodes in the scheme to which additions have been made. It is then possible to make a decision as to whether or not to recode these passages in the light of modifications to the scheme.

[8] Computer software for assisting expert coders presents each text unit on a separate line. Anyone who has tried to read a lengthy text in which every word is presented on a separate line will know it is much harder to do this than to make sense of a text with an average of ten words per line. We prepared texts for coding using a word processor to create documents with an average line length of ten words and then input them into the NUD.IST computer package that we used for storing and retrieving text, coding scheme and codings, as well as assisting and managing the coding process.

[9] If a neutral text unit is in a paragraph that otherwise expresses a neutral position on the policy concerns at issue, then it is obviously coded neutral. If a neutral text unit is in a paragraph at a point between a pro and a con text unit dealing with the same policy concern, then it is coded neutral.

text.[10] Most quantitative approaches, in contrast, allocate text units according to mechanical criteria that typically imply taking text units out of any wider political context. This is done by defining a content analysis "dictionary" of words or phrases systematically associated with particular coding categories in relevant texts. The computer then counts the number of words or phrases associated with each coding category.

It is quite possible, using quantitative techniques, to take greater account of the textual context of any unit being coded. For example, rather than including individual words, the dictionary can include phrases or more complex text strings. There is, however, a significant trade-off to be faced if this is done. First, particular phrases and word strings are likely to be repeated far less frequently than individual words in any well-written text, greatly reducing the amount of data generated by the coding process. Second, the use of given phrases and word strings is more stylistically idiosyncratic to a particular author. This raises reliability issues when relating text from different authors to the same underlying policy position. Thus, despite a longstanding gut instinct within the profession that more complex text units should be incorporated into quantitative content analysis, the reality is that much valuable information can be extracted from texts by using individual words as the fundamental unit of analysis. Any alternative faces very serious problems of its own.[11] Accordingly, we use individual words as our units of analysis and do not include longer word strings in our dictionary. Since, as we shall see, the level of cross-validation between our estimates and those derived from completely independent sources was high, we see no reason at this stage to move to a more complex unit of analysis.

People who come to computer-coded text analysis for the first time are often understandably sceptical. They immediately think of words that have several quite different meanings in different contexts—race, state, or class, for example. They then think of words that are often qualified by their context to have contradictory ideological meanings—taxes, spending, or services, for example. In an abstract sense this scepticism seems well justified, but closer familiarity with the technique, combined with actual patterns of word use in real texts, tends to allay at least some of these worries. There is absolutely no need to code all words in the text under investigation. Indeed ambiguous words are typically not coded at all, and a good quantitative content analysis dictionary will consist of words with as little ambiguity as possible.[12] Furthermore, once we turn to real texts, there are far fewer ambiguous words than might on the face of things be expected. Of many theoretically possible meanings of a word, in practice one meaning tends to dominate in the texts analyzed. Most uses of the word "taxes" in party manifestos, for example, are in practice associated with arguments in favour of cutting taxes. Far fewer actual uses of the word taxes are found in discussions of the need to raise taxes, despite the fact that, in the abstract, both types of occurrence might seem equally likely. Thus if we assign the word "taxes" to a coding category dealing with cutting taxes, we will not always be right, but the number of times we are right will far outweigh the number of times we are wrong. Assigning the word "taxes" in this way and analysing its occurrence in party manifestos gives us valuable information about the texts under investigation.[13] In practice, however, most of the words used in our dictionary have a relatively unambiguous meaning.

With a well designed coding scheme and its associated dictionary, therefore, quantitative content analysis can give us a lot of information about the substantive content of texts. It does, furthermore, score over expert coding in two important respects. The first and most obvious is reliability. Computer coding is 100 percent reliable, while levels of intercoder reliability among experts, and even the intracoder reliability of the same expert coding the same text at different times, can leave a lot to be desired. Indeed, such is the cost, in terms of time and effort, of expert coding that most studies engage in very little systematic evaluation of either intercoder or intracoder reliability. Thus, while mechanically analysing words out of context may on the face of things seem to have an obvious cost in terms of the *validity* of data generated, this is offset by a very significant gain in their *reliability*.

[10]As we will shortly argue, however, it is very difficult to confine the context taken into account by a human coder to the text being coded, as opposed to other knowledge the coder has about that text and its author.

[11]For an excellent review of this and many other issues in computer assisted content analysis, see Alexa (1997).

[12]If the analyst feels that certain crucial words are ambiguous, then these can be "disambiguated" by an expert coder—for example, into "race#1" (as in running) and "race#2" (as in ethnicity). This requires intervention from an expert coder and thus reduces the comparative advantage of computer over expert coding. What results is, in a sense, a hybrid technique.

[13]It is possible to generate a probabilistic dictionary that assigns a probability that a particular word comes from a particular manifesto in a given set of calibration texts. (For an early attempt to do this using Dutch texts, see Kleinnijenhuis and Pennings [1999]). Procedures for doing this are far more complex than those we propose below, which in effect assign probabilities of 1.0 that given words are associated with particular coding categories, and there is no indication as yet that they produce better results. The future development of our proposed technique certainly does not preclude probabilistic dictionaries, however.

Second, even when we consider validity, there are important ways in which expert coding may be less valid than computer coding. An expert coder, by definition, comes to a text with prior knowledge of its context. Knowing a particular text to come from a left-wing party, for example, an expert coder might be more inclined to allocate certain text units to a left-wing coding category. The same coder might have allocated the same words in different way if he or she had known that they came from a right-wing party. Inevitably, expert coding impounds the prior opinions of the coder about the text being coded, as well as its actual content. A computer that codes words out of context comes to a text with no prior opinions—what is coded is the text and nothing else. This suggests, for example, that radical but previously unannounced shifts in policy positions might be more effectively identified in computer coding than by an expert coming to a text with certain expectations about the policy positions under investigation.

Designing a Quantitative Content Analysis Coding Scheme and Dictionary

Clearly the design of a good dictionary is vital to good quantitative content analysis. Unfortunately, existing dictionaries were not well suited to the task of extracting information on substantive policy positions from political texts. We thus designed a procedure for designing an English language dictionary to do this job for us. What we present here is our first attempt at such an enterprise, certainly not the last word on the subject and one we will continue to refine. Most important, given changing political meanings of words over time and space, is the *procedure* for deriving a dictionary rather than the substantive content of any given dictionary. Moreover, as we shall argue in the concluding section, there may be a good case, since we view the enterprise as one of estimating a priori policy scales, for reconstructing the dictionary as frequently as we have independent data sources against which to cross-validate it. What remains constant over time is thus the dictionary generation procedure, not the actual word lists in the dictionary. Indeed changes in the word lists associated with policy positions are matters that should be of considerable substantive interest in their own right.[14]

We based the categories used in our dictionary on a collapsed version of the new expert-coding scheme discussed above to allow us to compare the results of computer coding with those generated by expert coders. This gave us categories for economic policy (pro- and con- state intervention in the economy) and social values (liberal and conservative). We included categories on political reform (radical and defensive), environmental policy, and law and order, as well as a number of other matters that we do not deal with here. We defined our dictionary by allocating words to these categories using a combination of a priori and empirical criteria. Empirically, we used the British Conservative and Labour manifestos of 1992 as a pool of key words. These were fat manifestos with lots of words, issued by parties that we expected, a priori, to have different policies on many issues. Before selecting any specific word for the dictionary, we looked at all words used in the two manifestos, comparing the relative frequencies with which each used all possible words and focusing on words that one party used more than twice as often as the other. We then set about allocating this subset of words to the dictionary, using a combination of a priori reasoning and our empirical knowledge of whether the observed frequencies identified the word in question as a "Labour" or a "Conservative" word. Thus the word "taxes" was used twenty-two times in the 1992 Conservative manifesto and only once in the 1992 Labour manifesto. We thus allocated it to the category: "reduce state involvement in the economy." To take another example of a word that in the abstract can be used in many contexts, the word "choice" appeared thirty-eight times in the 1992 Conservative manifesto and only three times in the Labour manifesto. Thus we also allocated this to the category: "reduce state involvement in the economy." We used similar criteria to allocate words to the full coding dictionary.[15] No word was allocated to a coding category unless it had a clear substantive meaning in terms of that category. And no word was allocated to a category unless it was used by one party twice as often as by the other. Finally, no word was allocated to more than one coding category although, as we mentioned above, we have no principled objection to multiple coding.

It is important to note that our use of the 1992 British Labour and Conservative manifestos in the definition of our dictionary means that any substantive codings of these manifestos reported below contain no new information. The coded manifestos appear very different from each other precisely because they were used in the process of defining the dictionary; indeed the dic-

[14] This method need not be confined to English-language dictionaries; associated researchers have recently used the same technique to derive dictionaries for coding political texts written in Dutch (de Vries, 1999), German (Garry, 1999), Italian (Giannetti, 1999), and Norwegian (Garry, 1999).

[15] The complete version of this dictionary can be downloaded from *http://www.politics.tcd.ie/personnel/staff/laver.html.*

tionary was defined to make them different. The codings have substantive meaning only for the "virgin" texts we analyse below.[16]

Estimating Policy Positions from "Raw" Codings

How should information extracted from policy documents be used to estimate the positions of political actors? One way of doing this is largely inductive and approaches the text with no prior assumption about the substantive content of any underlying ideological dimension. This is the general style of analysis adopted by the original MRG researchers, who analysed their data by first using factor analysis to summarise patterns in the empirical frequencies of text units in different coding categories. They then used both factor loadings and the factor scores of manifestos on these dimensions to interpret them in substantive terms. A careful and persuasive development of this approach has recently been proposed by Gabel and Huber (2000), who review and compare a number of inductive methods of estimating positions on a general left-right scale, using MRG data.

The essential purpose of the Gabel-Huber method is to estimate a general left-right scale on which political actors can be placed, quite explicitly making no assumption about the policy concerns that might form part of such a scale, and quite explicitly not interpreting the resulting scale in terms of substantive policy concerns. They make no attempt to estimate party positions on any other dimension. Gabel and Huber use the entire fifty-four-category MRG dataset to find the underlying dimension that best accounts for the covariation in party positions on the policy categories. They define this as the main left-right dimension without regard to, or indeed any reporting of, its policy content. For this reason they refer to their method as the "vanilla" method. Since the vanilla method is explicitly mute on the substantive policy content of the left-right dimension it may have everything, or nothing, to do with economic policy, social values, or anything else. The key objective for Gabel and Huber is to use the MRG data to generate the best unidimensional account of covariation in party policy positions. Within this frame of reference, they operationalize and test the vanilla method against alternative methods of extracting a left-right scale from MRG data, measuring success in terms of ability to predict independent estimates of party positions on a general left-right scale.[17] If what is required is a single, general left-right scale, then their results are very encouraging indeed, though it should never be forgotten that any given implementation of this inductive technique will be entirely dependent upon the choice of cases to analyse. Applying the inductive Gabel-Huber technique to different parties and different time periods will produce left-right scales with different substantive meanings. This means that interpretations of party movements on the scale over time must be made with very great care.

In contrast to this explicitly inductive and unidimensional approach, we prefer a more a priori, substantive, and multidimensional method for estimating party positions. We see ourselves as engaged in the estimation of party positions on the type of exogenously defined policy dimensions that have become part of the currency of "multidimensional" spatial models of party competition. These dimensions give structure and meaning to the strategic moves that parties make and typically deal with policy concerns such as economic policy, social values, foreign policy, and so on.

We thus address the problem of analysing political texts to estimate the position, P_C, of some party, P, on some policy concern C. We do this by attempting to model the process by which actors might extract policy information from political texts. Once a policy document has been read by a political actor (or coded by an analyst), a certain number of text units, possibly zero, will have been read (or coded) as dealing with policy concern C—say P_{Ctot}. Of these, P_{Cpro} units will have been read (or coded) as being pro some substantive position on that issue, P_{Cneut} will have been read (or coded) as being neutral, and P_{Ccon} as being contra. Thus:

$$P_{Ctot} = P_{Cpro} + P_{Cneut} + P_{Ccon}$$

The original MRG approach was to use P_{Ctot}, expressed as a proportion of the total number of text units in the manifesto, as an indicator of the *emphasis* being given by the manifesto to the policy concern under consideration. This is the number that subsequent analysts have (mis)used as the basis of an estimate of the policy position of the manifesto in question. It is clear, however, that the *position* on some policy issue taken by a reader from the manifesto is much more likely to be some function of the balance of pro and contra mentions of the substantive policy concern under investigation.

[16]Computer coding was conducted using the "tagcoder" routines in the Textpack content analysis package.

[17]It is worth noting that some of the scales with which they compare their own do have substantive policy content and might be better evaluated in terms of ability to predict independent estimates of party policy positions on more substantive scales.

Consider the whole corpus of text in the document dealing with the policy concern under investigation. If every text unit had been coded as pro, it seems reasonable to assume that a reader—whether a member of the public, a journalist, or a party politician—would be in no doubt that the position of the party on the policy concern is pro. If every text unit had been con, it seems reasonable to assume that a reader would be in no doubt that the position is con. But what of policy concerns for which every relevant text unit in the manifesto is not coded in the same way, the situation that typically obtains in practice?

Imagine the reader comes to the text with no prior estimate of P_C, the position of the party under investigation on the policy concern at stake. (We will return in future work to the interesting possibility that the reader does in fact have some prior estimate of P_C.) The reader thus estimates the party to be neither pro nor con before reading the manifesto, but updates this estimate after the manifesto has been read. Assume that every time a reader encounters a "pro" text unit this increases the probability that she feels the party to have a pro position, while every time she encounters a "con" text unit, this increases the probability that she feels the party to have a con position. Assume that a "con" text unit negates the updating effect of a "pro" text unit and that a "neutral" text unit has no updating effect on her estimate of whether the party has a pro or a con position. In other words, all pro and con text units have information about the party's substantive position, and the relative balance of pro and con text units is the basis of how the reader updates her estimate of this position.

In order to use this type of information to estimate a party position on some policy scale, we arbitrarily fix the endpoints of the scale. If the position of party P on concern C is unambiguously pro, let $P_C = +1$. If the position of party P on concern C is unambiguously con, let $P_C = -1$. If the position of party P on concern C is neutral, defined as being a perfect balance between pro and con, let $P_C = 0$. We estimate the updating effect, $U(P_C)$, of the manifesto on estimates of the position of party P on concern C as the relative balance of pro and con text units, taken as a proportion of all text units conveying information on this matter. Thus:

$$U(P_C) = (P_{Cpro} - P_{Ccon}) / (P_{Cpro} + P_{Ccon})$$

On the assumption of no prior knowledge about the party policy position, the updating effect of the manifesto is the *only* information the reader has about the position of party P on concern C. Thus:

$$\begin{aligned} P_C &= U(P_C) \\ &= (P_{Cpro} - P_{Ccon}) / (P_{Cpro} + P_{Ccon}) \end{aligned}$$

Fortunately, this is a number that can easily be calculated from codings of text units in any policy document. For a given policy concern, it is the number of pro text units, minus the number of con text units, as a proportion of all text units conveying information on this matter, pro or con.

Substantive Policy Scales

We use this approach to generate two substantive policy scales from party manifestos coded using both the revised expert-coding scheme and the new computer dictionary. The first scale relates to economic policy, the second to social values. An economic left-right scale, $Econ_{LR}$, is defined as:

$$Econ_{LR} = (Econ_R - Econ_L) / (Econ_R + Econ_L)$$

This scale was calculated from both expert and computer codings by using the following data from each case to estimate its component parts:

$Econ_L$ = total text units in category: "increase role of state in the economy"

$Econ_R$ = total text units in category: "reduce role of state in the economy"

In order to compare our results with those that would have been generated by the original MRG coding scheme, we calculated directly analogous scales from the MRG data.[18] To this end, we concentrated only upon the positional clusters of coding categories identified by Laver and Budge (1992), using these clusters as the building blocks of positional policy scales. We used the Laver-Budge cluster of coding categories "state intervention" as our indicator of $Econ_L$ and their "capitalist economics" cluster as our indicator of $Econ_R$.

We also defined a scale estimating party positions on a dimension dealing with liberal vs. conservative social values, Soc_{LC}:

$$Soc_{LC} = (Soc_C - Soc_L) / (Soc_C + Soc_L)$$

[18]This latter measure is similar in form to one independently arrived at, on the basis of a very different argument, by Kim and Fording (1998). For an evaluation of the relative reliability of ratio-based scales such as this and subtractive scales such as that proposed by Laver and Budge (1992), when applied to the MRG dataset, see McDonald and Mendes (1999).

The following data were used to estimate the component parts of this scale:

Soc_L = total text units in category: "liberal, permissive, or nontraditional social values"

Soc_C = total text units in category: "conservative, restrictive, or traditional social values"

In each case it is important to note that we did not include manifesto codings relating to crime or law and order in our social values scale, since we felt on a priori grounds that these tend to cut across other social values and are thus better treated separately. We used a similar approach in calculating a liberal-conservative social values scale from the original MRG data. We used the Laver-Budge "social conservatism" cluster of coding categories as our indicator of Soc_C and the "anti-establishment" cluster as our indicator of Soc_L. The approach we used for deriving policy scales from political texts was thus applied in an identical manner to both the original MRG data and the recoded manifestos.

Results

Estimates of policy positions based on coding British and Irish party manifestos from the 1992 and 1997 elections in the different ways described above were derived as follows. First, we used expert coders[19] and the computer coding software[20] to calculate the raw counts for expert and computer codings of the manifestos in the aggregate coding categories used to build the "economic left-right" and "liberal-conservative values" scales. Second, we used these data to construct "raw" policy scales for 1992 and 1997, as described in the previous section. We also constructed equivalent raw scales based on the original MRG data for 1992 in Britain and Ireland. All of the scales derived from content analyses are constructed on the same basis from the same text information and can be compared directly with each other. However, our completely independent source of cross-validation for the text-based scales came from expert surveys for Britain in 1989 and 1997 (Laver and Hunt, 1992; Laver, 1998a) and Ireland in 1992 and 1997 (Laver, 1994; Laver 1998b).[21] The raw scales derived from expert surveys are constructed on a different basis, using quite different information. To allow all scales to be compared directly, each was standardised across the full set of sixteen observations for each scale.[22]

These results are presented in Table 2. A systematic summary of the interrelationship between the various ways of estimating economic policy positions can be found in Table 3, which reports Pearson correlations between the economic policy scores produced by each technique for each party in each election.

Tables 2 and 3 show very encouraging cross-validation between the various economic left-right scales under investigation. The scales based upon expert-coded content analysis, under either the MRG or the revised coding scheme, are very close to each other and to scales derived completely independently from expert surveys. Correlations between pairs of these scales, for both the 1992 and 1997 elections, range from 0.94 to 0.99. The left-right scales based upon computer coding generate very similar positions to those of the other scales.[23] Correlations between computer-coded and other scales range from 0.72 to 0.94.[24]

For the British parties, the correspondence is very good, particularly given the need to treat with great caution the computer estimates for Labour and the Conservatives for 1992, since these manifestos were used to generate the computer dictionary, and will thus "artificially" separate these manifestos.

In the case of Ireland for 1997, the expert survey, revised expert coding, and computer coding all put Democratic Left (DL) firmly on the left, Labour on the centre left and the Progressive Democrats (PDs) on the right. All techniques place Fianna Fáil (FF) and Fine Gael (FG) between Labour and the PDs. The expert survey finds almost no difference between FF and FG. The revised expert coding puts FF to the left of FG. The computer coding places FG towards the centre-left, alongside its coalition partners DL and Labour. In 1992, all techniques have DL and the PDs respectively anchoring left and

[19] Two coders independently coded the documents. The coders met after completing their independent coding to discuss each text unit on which they disagreed and arrive at an agreed coding.

[20] The tagcoder routines in Textpack.

[21] No expert judgements are available for Britain in 1992.

[22] That is, eight parties in two elections.

[23] It is important to remember that the 1992 computer-generated positions for the British Labour and Conservative parties should be treated with great care, since these manifestos were used to generate the computer-coding dictionary in the first place. Differences between these manifestos, therefore, were assumed a priori rather than inferred from the data.

[24] Gabel and Huber (2000, footnote 12) regard correlations of the order of 0.88 and 0.94 as indications that policy scales are "measuring the same thing."

TABLE 2 Standardized Economic "Left-Right" and Social Values "Liberal-Conservative" Scores for 1992 and 1997 British and Irish Party Manifestos and Standardized Scores on Comparable Expert Survey

	Economic Policy				Social Policy			
	Computer	Revised Expert	MRG	Expert Survey	Computer	Revised Expert	MRG	Expert Survey
UK Lab 1992	–1.52	–0.84	–0.99	–1.18	–1.75	–0.02	0.19	–0.69
UK LD 1992	–0.15	–0.68	–0.22	–0.57	–1.19	–1.34	–1.20	–0.61
UK Con1992	2.28	1.34	1.06	1.35	0.96	1.03	1.11	1.44
UK Lab 1997	0.38	0.10		–0.12	0.21	0.98		–0.26
UK LD 1997	–0.38	–0.41		–1.09	–0.88	–0.28		–0.63
UK Con1997	0.81	1.45		0.89	0.96	1.76		0.94
Irl DL 1992	–0.95	–1.34	–1.30	–1.36	–0.90	–1.29	–1.63	–1.42
Irl Lab 1992	0.07	–0.77	–0.93	–0.85	–0.80	–0.74	0.56	–0.82
Irl FF 1992	–0.79	–0.13	0.37	0.48	1.11	0.88	1.11	1.96
Irl FG 1992	0.10	0.34	0.72	0.88	1.58	0.21	0.08	1.05
Irl PD 1992	0.82	0.64	1.28	1.44	0.82	–0.69	–0.23	0.01
Irl DL 1997	–1.38	–1.38		–1.15	–0.79	–1.48		–1.07
Irl Lab 1997	–0.68	–0.88		–0.66	–0.74	–0.75		–0.68
Irl FF 1997	0.22	0.04		0.26	0.50	–0.01		1.01
Irl FG 1997	–0.19	0.56		0.31	0.15	1.19		0.35
Irl PD 1997	1.35	1.96		1.37	0.75	0.54		–0.59

TABLE 3 Pearson Correlations between Alternative Estimates of Economic Left-Right Scale Positions, Britain and Ireland 1992–97

	Computer Codings	Revised Expert Codings	Original MRG Codings	Expert Surveys
1992				
Computer codings	1.00			
Revised expert codings	0.85	1.00		
Original MRG codings	0.72	0.94	1.00	
Expert surveys	0.75	0.95	0.99	1.00
1997				
Computer codings	1.00			
Revised expert codings	0.94	1.00		
Expert surveys	0.91	0.95	n.a	1.00

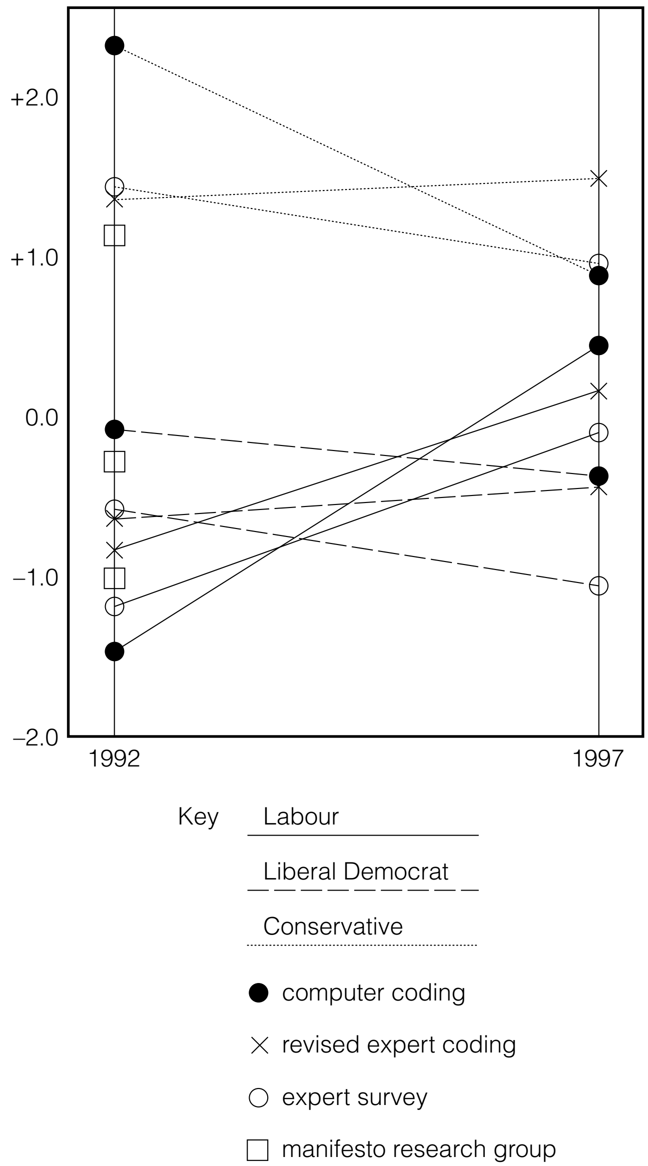
right, but computer coding of FF does appear deviant, placing it to the left of Labour while the other techniques place it to the right.

The most striking feature of these results, however, and a clear-cut test of the face validity of the computer coding technique, concerns the widespread informal perception that the British Labour Party shifted sharply towards the centre of the economic policy spectrum in 1997. This is most clearly shown in Figure 1. The expert survey of British party policy positions in 1997 showed Labour making a major move to the centre, with the Liberal Democrats shifting somewhat towards the left. The net result was that the Labour Party was, in 1997, placed by the experts between the Liberal Democrats and the Conservatives on economic policy, rather than to the left of the other two parties as previously.

Table 2 and Figure 1 show that the computer-generated scales did indeed pick up this important shift in British party policy positions. The 1989 and 1997 expert estimates of Labour Party policy are very closely mirrored by the independent computer-generated estimates. The heavy lines in Figure 1 show that the rightwards shift in Labour policy is picked up very clearly by all techniques. The expert surveys and computer coded content analysis also imply a leftwards shift between 1992 and 1997 in the economic policies of both the Liberal Democrats and the Conservatives. The scales based upon expert-coded content analysis of manifestos are more equivocal about this, implying that the Conservatives and Liberal Democrats had more or less remained in the same place.

Whichever technique is used, however, the reversal of the positions of Labour and the Liberal Democrats shows up in a very striking way. Considering that one technique involves averaging the subjective judgments of political scientists, another involves the analysis of party manifestos by expert coders, while another involves the computer counting of key words, the techniques correspond to a remarkable degree.

FIGURE 1 Standardized Expert Survey, Computer Coded and Expert Coded Estimates of Party Policy Positions in Britain 1992–97



Turning to party positions on social values, all scales under investigation are in broad agreement for the British parties. The main anomaly is that both the MRG and revised expert manifesto codings place the Liberal Democrats firmly on the liberal side of Labour in 1992, in contrast to the centre-liberal position generated by both the expert judgments and computer-generated scales. A similar pattern is found in Ireland—there is broad agreement between most scales on the positions of all parties except the PDs. The PDs are estimated as a liberal party on social values by the expert text codings, as a centre party on this dimension by the expert judgments, and as a conservative party by the computer-generated scales. The MRG codings, furthermore, identify Labour as having conservative social values, a result that does not on the face of things seem plausible.

A systematic summary of the interrelationship between the various ways of estimating social policy positions can be found in Table 4, which reports Pearson correlations between the various social policy scores. The correlations between the expert survey results and the various manifesto-based estimates are somewhat lower than those for economic policy, but are very respectable nonetheless, ranging from 0.67 to 0.88. For the election

TABLE 4 Pearson Correlations between Alternative Estimates of Liberal-Conservative Social Values Scale Positions, Britain and Ireland 1992-97

	Computer Codings	Revised Expert Codings	Original MRG Codings	Expert Surveys
1992				
Computer codings	1.00			
Revised expert codings	0.62	1.00		
Original MRG codings	0.49	0.87	1.00	
Expert surveys	0.84	0.88	0.74	1.00
1997				
Computer codings	1.00			
Revised expert codings	0.80	1.00		
Expert surveys	0.71	0.67	n.a	1.00

for which we have MRG data (1992) it is the MRG codings that are the least well correlated with the expert survey estimates.

Perhaps the most striking overall finding from Tables 2–4 is that there is a high degree of cross-validation between the three quite independent techniques for generating data on party policy. Expert-coded content analysis, computer-coded content analysis and expert surveys all produce consistent results once an explicitly positional approach is used to defining the scales with which to estimate policy positions. No technique stands out as producing deviant results, and all techniques seem very sensitive to the striking shift in the policy position of the British Labour Party that we should expect any valid technique to pick up.

This conclusion also applies the MRG data on economic policy positions, once these data are reanalysed using the new substantive policy scales, and thus offers hope that careful reanalysis of the MRG data on policy *emphases* might yield useful estimates of party *positions* on an economic left-right policy scale. This almost certainly arises from the more explicitly positional nature of at least some of the MRG coding categories relating to economic policy and offers the prospect of putting an existing huge data set to better use than hitherto. The social policy scale generated from the 1992 MRG data was, however, the most deviant. Here we see the impact of the MRG's saliency approach most clearly. While the MRG scheme did cover economic policy from many different angles, it did not code *positions*, as opposed to *emphases*, on a wide range of other matters, including social values. The revised schemes proposed in this paper do set out to do this, and the payoffs become apparent for the main noneconomic policy dimension we consider.

The Way Forward

Three headline conclusions can be drawn from these results. The first, and for us the most exciting, is that even a very simple form of computer-coded content analysis, one that can be used right now, can generate estimates of policy positions that can be cross-validated against quite independent sources. The second is that parts of the MRG saliency data can be reanalysed using a priori positional scales to derive estimates of policy positions that can also be cross-validated. The third is that the new explicitly positional expert-coding scheme works well and may offer clear advantages in noneconomic policy areas for which the original MRG scheme offers fewer quasi-positional categories.

The latter conclusions are important because they imply that the MRG data, in which the profession has already made a huge investment over many years, do represent a valuable source of information on certain substantive policy positions, provided that care is taken to derive estimates in appropriate ways. The a priori and substantive approach to building scales used here is based on assumptions about how real people might extract information from real manifestos. Reanalysing the MRG data in this way, we have come close to quite independent estimates of substantive party policy positions derived from expert judgments. Our findings thus echo, for more substantive policy scales, those of Gabel and Huber (2000) concerning a single general left-right dimension.

Real-world research resources are always scarce. These findings imply that, while the new expert-coding scheme we propose offers a methodologically more appropriate way to estimate substantive policy positions, the huge costs of hand-recoding the entire corpus of the

MRG dataset may not justify the benefits. Very serious questions do have to be asked, however, in respect to future work. It does not seem to us to be appropriate to continue the laborious and costly process of hand-coding manifestos using *only* a substantially nonpositional coding scheme designed to estimate the relative salience of different policy concerns.

By far the most radical conclusion to be drawn from the present analysis, however, concerns the way in which computer codings of the manifestos can be used to generate estimates of party positions that are very similar to those estimated using much more resource-hungry expert techniques. This suggests that it may be possible to refine computer-coding techniques to streamline the coding of the vast amount of policy-relevant text now available in machine-readable form. This in turn opens up the possibility of using computer coding to generate systematic estimates of the policy positions of the factions, even the individual members, of political parties, and to chart the development of these positions over time. Such estimates simply do not exist at present, and there is no realistic way they will be generated unless computer-coding techniques, or some reliable and valid working alternative, can be developed. The effective computer coding of political texts will thus represent a significant breakthrough for those concerned with analysing party competition. It will allow the operationalization of models that go beyond the unitary actor assumption, as well as of those that deal with movements in substantive policy positions between elections.

We are acutely aware that we are at an early stage in this research programme. We feel the best way to proceed, since the ideological meaning of key words in the political lexicon clearly changes over space and time, is to formalise the general procedure that we have reported here for generating a computer dictionary. First, party election manifestos can be expert coded using an appropriate positional coding scheme. Scales generated from these codings can also be cross-validated against expert surveys conducted during the same period, although this is not central to the technique. Second, a computer dictionary of key words can be derived from the same manifestos, with words allocated to coding categories in a way that enables computer coding to replicate expert codings as closely as possible. This means, of course, that no new information about *manifestos* will be derived from the computer coding of them. Rather, the computer coding of manifestos for which independent positional estimates are available would be used to "calibrate" the dictionary of key words, adjusting this for changes in the political lexicon between elections.

Having calibrated the computer dictionary in this way for a particular election, this can then be used to computer code vast volumes of "virgin" text in quantities that would be quite beyond the resources of any expert-coding technique. For example, policy statements issued by all members of parliament could be computer-coded to enable the researcher to draw a detailed ideological map of the internal policy spaces of political parties and look at movements in these over time. Alternatively, new party policy statements could be coded in an attempt to track shifts of policy between elections. This process represents the type of interaction between more qualitative expert-coding techniques and more quantitative computer-coding procedures that is now regarded as best practice in the analysis of political texts (Alexa, 1997). The expert coding of manifestos keeps the computer dictionary in touch with realities that may change across space and time, the computer-coding procedure then allows otherwise unmanageable volumes of text to be processed.

An obviously important issue concerns the cross-national extension of this technique, especially to non-English language texts. Early indications in this regard are most encouraging. Researchers associated with this programme have generated computer dictionaries in Dutch (de Vries 1999), German (Garry, 1999), Italian (Giannetti 1999), and Norwegian (Garry, 1999), as well as analysing English language government declarations (Mansergh, 1999). Preliminary findings suggest that computer-generated estimates of party positions on economic policy can be cross-validated against independent estimates at levels of correlation very similar to those reported in Tables 3 and 4 above. These results imply that the technique we propose is relatively robust to changes in political context and does have the potential to be used in cross-national research.

The successful development of this technique could thus have significant implications for the systematic empirical analysis of intraparty politics in particular and party competition in general. The possibilities are both enormous and exciting, once we have at our disposal a technique for the fast, valid, and reliable coding of huge volumes of text containing information on the policy positions of political actors. It will allow empirical accounts of party politics to go beyond the unitary actor assumption. It will also allow them to go beyond static models of party competition that do not recognise any interesting political activity between elections to dynamic models of continuous movement over time, movement driven by an intraparty political game. If the computer coding of political texts can be refined, applied, and accepted within the profession, the prize is simply enormous.

Manuscript submitted December 18, 1998.
Final manuscript received July 19, 1999.

References

Alexa, Melina. 1997. "Computer Assisted Text Analysis Methodology in the Social Sciences." *ZUMA Arbeitsbericht.* 97/07. Mannheim: ZUMA

Baron, David. 1991. "A Spatial Bargaining Theory of Government Formation in Parliamentary Systems." *American Political Science Review* 85:137–164.

Budge, Ian. 1999. "Estimating Party Policy Positions: From Ad Hoc Measures to Theoretically Validated Standards." Presented to the workshop "Estimating the Policy Positions of Political Actors." ECPR Joint Sessions, Mannheim.

Castles, Francis, and Peter Mair. 1984. "Left-Right Political Scales: Some Expert Judgments." *European Journal of Political Research.* 12:73–88.

DeVries, Miranda. 1999. "Computer-Based Content Analysis of the Party Manifestos of the 1998 Dutch Elections." Presented at the workshop on "Estimating the Policy Positions of Political Actors." ECPR Joint Sessions, Mannheim.

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

Garry, John. 1999. "Computer Coded Content Analysis Of German and Norwegian Manifestos." Presented at the workshop on "Estimating the Policy Positions of Political Actors." ECPR Joint Sessions, Mannheim.

Giannetti, Daniela. 1999. "The Computer Coding of Italian Political Texts." Presented at the workshop on "Estimating the Policy Positions of Political Actors." ECPR Joint Sessions, Mannheim.

Harmel, Robert, Kenneth Janda, and Alexander Tan. 1995. "Substance vs. Packaging: An Empirical Analysis of Parties' Issue Profiles." Presented at the Annual Meeting of the American Political Science Association.

Huber, John, and Ronald Inglehart. 1995. "Expert Interpretations of Party Space and Party Locations in 42 Societies." *Party Politics* 1:73–111.

Kim, Heemin, and Richard Fording. 1998. "Voter Ideology in Western Democracies, 1946–1989." *European Journal of Political Research* 33:1, 73–97

Kleinnijenhuis, Jan, and Paul Pennings. 1999. "Measurement of Policy Positions On The Basis Of Party Programmes, Media Coverage and Voter Perceptions." Presented at the workshop on "Estimating the Policy Positions of Political Actors." ECPR Joint Sessions, Mannheim.

Laver, Michael. 1994. "Party Policy and Cabinet Portfolios in Ireland 1992: Results from an Expert Survey." *Irish Political Studies* 9:157–164.

Laver, Michael. 1998a. "Party Policy in Britain, 1997: Results from an Expert Survey." *Political Studies* 46:336–347.

Laver, Michael. 1998b. "Party Policy in Ireland, 1997: Results from an Expert Survey." *Irish Political Studies* 13:159–171.

Laver, Michael, and Ian Budge (ed.). 1992. *Party Policy and Government Coalitions.* London: St. Martin's Press.

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

Mansergh, Lucy. 1999. "Using Computer and Hand Coding to Compare the Policy Positions of Government Declarations and Parties in Government." Presented at the workshop on "Estimating the Policy Positions of Political Actors." ECPR Joint Sessions, Mannheim.

McDonald, Michael, and Silvia Mendes. 1999. "The Policy Space of Party Manifestos." Presented at the workshop on "Estimating the Policy Positions of Political Actors." ECPR Joint Sessions, Mannheim.

Schofield, Norman. 1993. "Political Competition and Multiparty Coalition Governments." *European Journal of Political Research*, 23:1–35.

Warwick, Paul. 1994. *Government Survival in Parliamentary Democracies.* Cambridge: Cambridge University Press.