THE UNIVERSITY OF
WARWICK

Exit polling in a cold climate:
The BBC/ITV experience in Britain in 2005

John Curtice
Department of Government
University of Strathclyde

David Firth*
Department of Statistics
University of Warwick

July 16, 2007

Abstract

Conducting an exit poll to forecast the outcome of a national election in terms of both votes and seats is particularly difficult in Britain. No official information is available on how individual polling stations voted in the past; use of single member plurality means that there is no consistent relationship between votes and seats; electors can choose to vote by post; and most of those who vote in person do so late in the day. In addition, around one in every six intended exit-poll respondents refuses to participate. Methods developed to overcome these problems, and their use in the successful 2005 BBC/ITV exit poll, are described and evaluated. The methodology included: a panel design to allow the estimation of electoral change at local level; coherent multiple regression modelling of multi-party electoral change to capture systematic patterns of variation; probabilistic prediction of constituency winners in order to account for uncertainty in projected constituency-level vote shares; collection of information about the voting intentions of postal voters before polling day; and access to interviewer guesses on the voting behaviour of refusals. The coverage and accuracy of the exit poll data are critically examined, the impact of key aspects of the statistical modelling of the data is assessed, and some general lessons are drawn for the design and analysis of electoral exit polls.

Key words: general election, exit poll, forecasting, postal voting, probability calibration, Steed swing, swingometer, ternary diagram

1 Introduction

Electoral exit polls are typically conducted in order to enable a broadcasting organisation to publish, as soon as the polls close, an estimate of the share of the vote won by each party. In the case of a legislative election an attempt is also usually made to indicate how many seats each party is expected to win. These estimates provide a basis for informed on-air discussion

---

*Address for correspondence: Department of Statistics, University of Warwick, Coventry CV4 7AL, UK; email `d.firth@warwick.ac.uk`

1

during the period between the close of polling stations and the declaration of the first results; the numbers of viewers and listeners during that period are often among the highest achieved by current affairs broadcasts.

Conducting and analysing an exit poll successfully is a formidable methodological challenge at the best of times. As it has to be conducted in a sample of polling stations, the design of any such poll is inevitably heavily clustered, with highly adverse consequences for sampling error. Those who are approached to ascertain how they voted may well refuse to disclose what they have done in what after all is a secret ballot. Meanwhile the results of the interviews that are conducted have to be collated and analysed in little more than a few hours against an immovable deadline. The potential for error is considerable.

The problems facing those carrying out exit polls in the United Kingdom, however, are particularly formidable:

 (i) In contrast to most other countries, election results are not counted and declared by polling station/district. This makes it impossible to use knowledge of the election results at a previous election in order to stratify the selection of polling districts and so help ensure that they are representative of the country as a whole.

 (ii) Elections to the principal legislature, the House of Commons, are conducted using the single member plurality electoral system under which there is not necessarily a regular relationship between votes and seats.

(iii) Voters can if they wish opt to vote by post rather than attending a polling station, thereby automatically eluding the exit pollster.

(iv) As elections take place on a working day (Thursday) rather than, as is the case in many countries, a non-working day (such as a Sunday), the majority of electors vote in the early evening, thereby leaving very little time indeed to collate and analyse the bulk of the data before the polls close.

Despite these difficulties, at 10 p.m. on 5 May 2005 the two principal broadcasting organisations in the UK, the BBC and ITV, broadcast a forecast total of 356 seats for Labour — a House of Commons majority of 66 seats — based on the analysis of a jointly commissioned exit poll (see Table 1). When eventually the result of a delayed election in South Staffordshire was declared in the early hours of June 21, this proved to be exactly the arithmetic of the new parliament. Never before in the history of exit polling in Britain has the overall majority been predicted with such accuracy (see, e.g., Payne, 2003, p. 214). Such an *exact* forecast of the headline majority could not of course have been made without a fair measure of luck, since many House of Commons seats are very closely contested. Indeed, the poll-based forecast for the number of seats won by the two main opposition parties was slightly less accurate, and the national vote shares for the Liberal Democrat and Labour parties were respectively slightly underestimated and slightly overestimated. Nevertheless, overall the predictions made from the exit poll were as accurate

2

Table 1: 2005 BBC/ITV exit poll forecast and final election outcome compared

		Lab	Con	LD	Oth
Exit poll	votes (GB) (%)	37	33	22	8
	seats (UK)	356	209	53	28
Outcome	votes (GB) (%)	36.2	33.2	22.6	7.9
	seats (UK)	356	198	62	30

as could reasonably be expected given the difficulties mentioned above. This paper reports on and critically evaluates the methodology that lay behind these forecasts. In so doing it aims to identify what general lessons might be drawn for the conduct of exit polls given this apparent success in a particularly harsh electoral climate.

The paper falls into three main sections. In Section 2 we outline in more detail the problems facing exit pollsters in the United Kingdom and how the methodology of the 2005 BBC/ITV exit poll attempted to overcome those problems. In Section 3 we evaluate the accuracy of the exit poll data in estimating party performance in terms of votes won. In Section 4 we examine how the analysis and modelling of those data contributed to the accuracy of the forecast in terms of seats won. Finally in our concluding remarks we consider the lessons to be drawn from the 2005 experience.

2   Overcoming the Obstacles

2.1   Selecting Polling Locations

As we have noted, votes are never counted and published at the level of individual polling stations in Britain. We thus do not have accurate information on the political complexion of any particular polling district, information that could be used to ensure that a selection of polling districts is politically representative. At the same time the geography of polling districts does not match that of the enumeration districts for which information is available from the census. This means that the census cannot easily be used to stratify polling districts according to those social characteristics known to be correlated with voting behaviour. At best one can engage in a multi-stage process of stratification. First of all one can select a stratified sample of parliamentary constituencies for which both past general election results and census data are available. Then, within each selected constituency, one or more local authority wards can be selected after stratifying by past local election results and, again, census data. But thereafter within the selected wards polling districts have either to be selected at random, or an attempt made, for these wards, to construct a best fit of census enumeration districts with polling districts, and this then used to guide the selection of a 'typical' polling district within each selected ward. Meanwhile, neither census data nor local election results are always a reliable guide to the political complexion of an area in a general election, and local authority ward boundaries do not necessarily match those of parliamentary constituencies at any given point in time.

3

How might one overcome this apparently insoluble problem? We begin by observing that the *change* in each party's share of the vote between any adjacent pair of elections varies far less than does each party's actual share of the vote. For example, the constituency-level standard deviation of the percentage of the vote won by the three main parties in Britain in 2005 was Labour 15.1, Conservative 14.0, and Liberal Democrat, 10.4. In contrast, the equivalent statistic for the change in their share of the vote since 2001 was only Labour 4.0, Conservative 3.0 and Liberal Democrat 4.4 (Curtice *et al.*, 2005). As a result, the probability that any randomly selected set of polling stations accurately reflects the change in each party's share of the vote across Britain as a whole is much higher than is the probability that any such set accurately reflects each party's share of the vote. Even a set of polling stations that is not representative in terms of each party's share of the vote may well be representative when it comes to the *change* in their share of the vote.

But of course we can only calculate the change in each party's share of the vote in a polling district if we have information on how well the parties performed in that district at the last election — precisely the information that is not available from the official results in Britain. However, if an organisation has conducted an exit poll at the previous election, it does have some — albeit potentially imprecise — information on how well the parties performed at each polling station covered by that poll. If therefore it returns to those same stations at the subsequent election, it can use an exit poll to estimate the change in each party's share of the vote. It is this approach that was adopted by the 2005 exit poll. Its principal purpose was to estimate the local change in each party's share of the vote. Thereafter estimates of each party's share of the vote and the number of seats it would win were derived by applying the estimated changes to the actual outcome at the 2001 election in each constituency.

This was not the first time that such an approach had been adopted. ITV's exit polls had long been conducted as far as possible in the same polling districts as at the previous election. The BBC's history of exit polls is a shorter one, but in 2001 it too adopted a strategy of revisiting the same polling districts as it had done in 1997. Thus when the two organisations decided in 2004 to lay down their rivalry and commission a joint exit poll, largely on the grounds of cost, it was relatively easy for them to agree that the 2005 poll should be based on those polling districts that had been covered by one or other of them in 2001.

There had, however, been some differences between the approach used by ITV in 2001 in selecting polling districts and that used by the BBC. While it had been disproportionately conducted in marginal constituencies, and in particular those marginal between Labour and the Conservatives, the 2001 BBC poll, conducted at 90 polling stations, had covered all kinds of constituencies. In contrast the parallel ITV poll, conducted in 100 locations, had *only* visited polling stations in constituencies that had previously been marginal. Either they were constituencies that had been marginal between Labour and the Conservatives (all of which were constituencies held by Labour in 2001 with majorities over the Conservatives of 25% or less) or they were ones

4

that were closely contested by the Liberal Democrats and the Conservatives (a mixture of seats held relatively narrowly by either party). In excluding any other kind of marginal, this approach reflected the fact that these had hitherto been the two main political 'battlegrounds' at British general elections. The two organisations adopted their different approaches because their polls were designed to achieve somewhat different objectives. ITV's poll was only intended to produce a seat projection, not forecast shares of the vote, and thus it could focus exclusively on the two types of constituency where seats were most likely to change hands. In contrast, not only was the BBC's poll intended to produce a forecast vote share, but it was also used, alongside the results from those seats that declared first, in order to produce an updated forecast; see, for example, Brown *et al.* (1999). As most of the constituencies that declare early are safe (Labour) constituencies (e.g., Payne, 2003, p.202) it is only possible to make an early check on the likely accuracy of an exit poll if it contains at least some polling stations located in safe constituencies.

Any possibility that this divergence would cause difficulty in securing agreement between the two organisations about which approach, and thus which polling districts, should comprise the locations at which interviewing was conducted (which it had been agreed should total 120) in the 2005 poll was, however, removed by a further complication. Polling district boundaries, which are determined by local returning officers, are far from being permanent. If, for example, an area's electorate grows substantially, this may make it administratively desirable that it be subdivided into two. Meanwhile the introduction of new local authority ward boundaries may split an existing polling district into two or more wards and so require an adjustment of polling district boundaries. Much of England and Wales, including all of London and all of the English metropolitan districts, had had their local authority ward boundaries redrawn since the 2001 general election, making it particularly likely that their polling district boundaries had been redrawn since then too. It was anticipated that, for this reason alone, a significant proportion of the 190 polling locations covered by the two organisations in 2001 (some of which were in any event located in the same parliamentary constituency) could not be revisited because the 2001 polling district no longer existed. In particular, there might well be fewer polling stations than desirable in London, a city where there was both a large number of marginal seats and where the change in each party's share of the vote had departed notably from the national pattern at past elections (see, for example, Curtice and Steed, 1997). Even in those locations where the polling district boundary had not changed, if a large increase or decrease in electorate had occurred since 2001 there was a danger that the partisan colour of the area might have changed as a result of population change, an eventuality that would make any exit-poll data for such an area much less informative than it might otherwise have been about the national pattern of change. In addition, if the location of a polling station had changed, even though the boundaries of the polling district were the same, the physical layout of the new location might present insuperable logistical difficulties for interviewers. In combination these considerations suggested that it was in fact quite likely that the final sample of 120 polling locations might more or less choose itself,

5

and that it might even be necessary in London in particular to visit some new polling locations.

This proved to be the case. After making extensive checks during the early months of 2005 on whether the boundaries of each of the 190 polling districts covered by ITV and the BBC in 2001 had changed, on whether the location of their polling station had changed, and after determining how much their electorate had changed between 2001 and 2005, just 116 of the 190 sampling points visited in 2001 could reasonably be revisited. Even then that figure was only reached after both adopting rather looser criteria in London in respect of the degree of change in the size of the electorate between 2001 and 2005 that would be tolerated, and after accepting too that in some instances the poll should be conducted in more than one polling location within the same parliamentary constituency. As this sample of 116 districts was still light on polling locations in London, it was augmented by the addition of four polling districts in the capital that had not been covered in 2001. These were districts where, after inspection and manipulation of the 2002 local election results and 2001 census data, it appeared likely that their political complexion was similar to that of the parliamentary constituency in which they were located, thereby making it possible to estimate the change in the share of the vote in that location by comparing the 2005 exit poll data with the result for the whole constituency in 2001.[1]

These 120 polling locations were far from being representative of the country as a whole. They inevitably over-represented marginal constituencies. For example, no less than 47% were located in constituencies where in 2001 Labour won with a lead of less than 20% over the Conservatives, whereas such constituencies comprised only 20% of all constituencies. This can be seen in Figure 1. This ternary diagram contrasts the political colour of the constituencies in which the 120 exit poll locations were situated with that of all the 628 constituencies in Great Britain. In the diagram a seat in which Labour secured all of the vote would be located in the corner marked 'Labour', and the further a constituency is located away from that corner, the lower Labour's share of the vote. Seats that fall in the sector of the triangle that is nearest to the Labour corner are those where Labour came first, though the closer they are to the lines that demarcate the Labour sector the more marginal they are. The same logic applies to the representation of the share won by the Conservatives and Labour; for further details see, for example, Upton, (1976, 1994) or Miller (1977). As can be seen, the locations covered by the exit poll were disproportionately situated in constituencies that fell within Labour's sector of the triangle but were relatively close to the Conservative one. Overall, the share of the 2001 vote won by each party in our exit-polled constituencies was Labour 40%, Conservative 36% and Liberal Democrat 20%, compared with the national (GB) outcome of Labour 42%, Conservative 33% and Liberal Democrat 19%. Evidently, the ability of the exit poll to provide the basis for an accurate estimate of the outcome rested on the premise that a sample that was unrepresentative in terms of the national share of the vote could still, after being analysed for possible differences of behaviour in different kinds of constituency, be representative in terms of change in share.
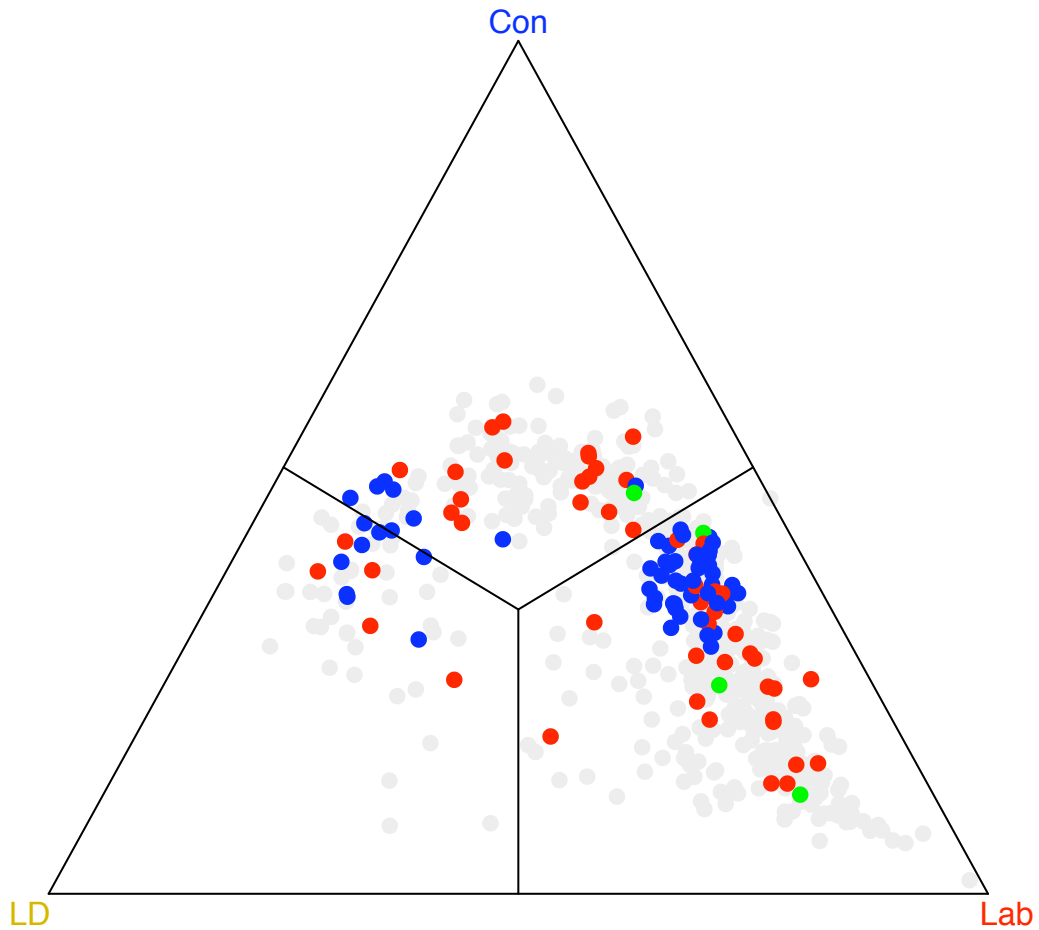
6

Figure 1: Ternary diagram of the division of the vote cast for the three main parties in each constituency in 2001. Coloured points represent constituencies covered by the 2005 exit poll; those in grey are constituencies not included in the poll. Points coloured red represent constituencies where the 2005 exit poll location was covered by the BBC 2001 exit poll, those coloured blue are constituencies in the ITV 2001 poll, and points in green are constituencies where the polling location was newly introduced in 2005.

2.2   Postal voting

The problem of selecting polling locations for an exit poll in Britain is long standing, but in 2005 a new problem presented a potentially more fundamental challenge. Exit polls can of course only acquire information from those who actually attend a polling station to vote. If a significant proportion do not vote in person but rather vote, say, by post, then it may no longer be the case that those who vote in person are representative of all those who vote at that station. Equally, and more importantly for our purposes, if there has been a significant increase since the last election in the proportion of people who vote by post there is a danger that any estimate of the changes in party shares of the vote derived by comparing the results of two exit polls may be biased. If those who newly opt to vote by post have different political preferences from those who continue to vote in person, this will have an artefactual impact on the change in the vote share recorded amongst those who vote in person. Between 2001 and 2005 there was just such an increase in the proportion voting by post. Although postal voting on demand had first been introduced just prior to the 2001 election, relatively few voters availed themselves of this new facility at that election; just 4% of the electorate were registered to vote by post (Electoral Commission, 2002).[2] That figure had, however, already begun to approach 10% by the time of the 2004 local and European elections (Electoral Commission, 2004), and it appeared from data on postal votes collected for 594 constituencies as part of the 2005 exit-poll exercise that nearly 12% were so registered at the 2005 election.[3] Even if we were happy to assume, as we were, that the change in each party's share of the vote amongst those who voted by post in both 2001 and 2005 matched the change in the share amongst those who voted in person both times, we evidently had to allow for the possibility that the estimates we derived from comparing data from the 2001 and 2005 exit polls would be biased by a change in the kind of person who voted in person.

One possible approach to this problem would have been to identify and interview, just before polling day, a sample of those living within each sampled polling district who were registered to vote by post and to ascertain both for whom they voted in 2001 and how (that is in person or by post) and for whom they voted in 2005. Such an approach was, however, clearly both costly and unlikely to yield reliable data. Instead a much simpler approach was adopted, an approach made possible by the fact that in a number of the pre-election polls it conducted in 2001 the polling company, MORI, had identified those voters who said they had registered to vote by post or intended to do so. If similar information were collected on polls conducted during the 2005 campaign it would be possible to establish whether those who were continuing to vote in person appeared to have changed their political composition relative to that found amongst postal voters. If they had, then together with information on the change in the number of postal voters at each polling station covered by the exit poll sample it would be possible to make an appropriate adjustment to the exit poll estimates. In the event the data that were collected suggested that the more widespread adoption of postal voting on demand resulted in

8

the political composition of those voting by post becoming a little less distinctively Conservative than it had been previously.[4] This meant that our exit poll ran the risk of slightly overestimating any swing to the Conservatives since 2001 (albeit only by a fraction of one percentage point[5]).

2.3 Refusals

Past experience suggests that in Britain as many as one in six of those who are approached to participate in an exit poll may refuse to complete the mock ballot paper that has been the principal mechanism by which information on how people vote has been collected in previous exit polls. If voters of one political persuasion are more likely to refuse to participate than those of another, that proportion is sufficiently large to bias the vote-share estimates derived from an exit poll appreciably — though of course if the degree of such differential refusal is approximately constant across elections it need not bias estimates of the *change* in each party's share of the vote based on comparing the data from one exit poll with those of another. The approach adopted to this problem in exit polls conducted for the BBC and ITV prior to 2005 had been somewhat different. In 2001 MORI's interviewers, working for ITV, were required to replace those who refused with the next person who left the polling station of the same age and gender as the person who had refused. NOP for the BBC in contrast made no attempt to replace refusers but instead asked interviewers to guess how they thought all those whom they had approached had voted. These data provided a means by which the voting behaviour of those who had refused could be estimated.[6] In 2005 the NOP procedure was adopted, though both the 2001 and the 2005 exit polls could be analysed on the day with the data on refusers and replacements either included or excluded.

2.4 Modelling votes and seats

Although numerous attempts have been made to suggest that there is a regular and predictable relationship under single member plurality between votes and seats across the country as a whole (e.g., Kendall and Stuart, 1950; Butler, 1951; Tufte, 1973; Laakso, 1979; Taagepera and Shugart, 1989), in truth that relationship depends on the geographical dispersion of party support (Gudgin and Taylor, 1979). Thus for example a party with a relatively small share of the vote across the country as a whole will win more seats if its vote is geographically concentrated than if it is evenly spread. In contrast a larger party will win more seats if its support is more evenly spread, such that it wins plenty of seats by small margins and relatively few by large ones. This of course is not necessarily a problem if the geography of party support is the same from one election to the next. Indeed this is the assumption that is made by the most widely used method to forecast the likely consequences of any outcome in votes, that is a model of uniform change under which each party's support is assumed to rise or fall in every constituency in line with the change in its support across the country as a whole. However, the post-war history of

9

British elections has demonstrated that non-uniform movements of party support can change the electoral geography of a party's support in a manner that significantly affects the relationship between seats and votes (Curtice and Steed, 1982; Curtice and Steed, 1986; Curtice, 1992; Curtice and Steed, 1997; Johnston *et al.*, 2001; Blau, 2004; Curtice *et al.*, 2005; Curtice, 2006).

Any attempt to derive an estimate of the outcome of a British election in terms of seats needs therefore not simply to be able to estimate accurately the level of party support across the country as a whole, but also its geographical distribution — or more precisely in our case, given our focus on the change in party support, how that distribution might have changed since the last election. We thus needed to develop from the exit poll data a parsimonious statistical model describing the systematic variation in the change in party support found in the poll, which could then be used to predict the change in each party's support in each constituency in Great Britain.[7] Such an approach would also enable us to ensure that our estimate of each party's share of the vote was not biased by the unrepresentative nature of our sample of polling locations. In order to maximise our ability to capture whatever systematic variation in party support existed, we collated a wide variety of constituency-level background data, including previous election results, demographic data from the 2001 census, and socio-economic data made available to use by the marketing agency, Experian, which could be used as predictor variables in regression models of change.

First, however, it is necessary to define what changes in vote share are to be modelled — i.e, the response variables to be used in regression models — given that in Britain there are more than two parties capable of winning significant numbers of seats. It is not possible simply to create separate regression models for the change in each party's vote share of the vote because the resulting estimates would not necessarily sum to 100%. A coherent resolution of this difficulty was achieved by using four 'nested' change variables, namely the changes in

- Oth

- $100 \times \text{Nat}/(100 - \text{Oth})$

- $100 \times \text{LD}/(100 - \text{Oth} - \text{Nat}) = 100 \times \text{LD}/(\text{LD} + \text{Lab} + \text{Con})$

- $100 \times \text{Lab}/(\text{Lab} + \text{Con})$

— the change in the last of these being the equivalent of 'Steed' or 'two-party' swing (Steed, 1965). Here 'Nat' means the percentage vote share for the Scottish National Party (SNP) in Scotland and Plaid Cymru (PC) in Wales, and 'Oth' denotes the aggregated votes for all other candidates. A hierarchy of this kind allows four separate regression equations to be constructed without fear of violating the sum-to-100 constraint.

This hierarchy is of course arbitrary, and was chosen largely because the outcome of the resulting models could easily be explained, not least to BBC and ITV journalists. For the same reason, and also for reasons of numerical stability, these four response variables were used in their 'raw' form, untransformed — rather than, for example, applying a logit, probit or other

10

such transformation — and accordingly care was taken to avoid the use of predictor functions implying vote shares outside (0, 100%). To help us meet these constraints we routinely transformed continuous-valued predictor variables; prior to their inclusion as candidate predictors in our regression models of change, such variables were typically either dichotomized, for example at their median, or else coded into three ordered categories ('low', 'medium', 'high'). Clearly such transformation degrades the available information, but this was a price we were happy to pay in order to be free to experiment with different sets of candidate predictors, under severe time pressure, without the danger of producing some unreasonable or even nonsensical constituency vote-shares by extrapolation of the predictor function beyond the range of the limited available data.

Multiple linear regression models were selected and fitted fitted separately for each of the four defined change variables. Estimation was by weighted least squares, with automated downweighting of gross outliers using Tukey's bisquare function (e.g., Mosteller and Tukey, 1977, Ch.10). The estimated four-dimensional change vector for each constituency, as predicted by our regression models, was then applied to the shares of the vote won by each of (Con, Lab, LD, Nat, Oth) in 2001.[8] This provided us with a predicted share of the vote for each major party in each 2005 constituency. Predicted national shares were then obtained by straightforward aggregation of these estimates, based on an assumption of no change in the pattern of turnout between 2001 and 2005.[9]

It might be anticipated that predicting the total number of seats to be won by each party would be equally straightforward: we could simply allocate each seat to the party with the largest predicted vote share and then sum across all constituencies. However, to do so would have been to neglect both the substantial uncertainty present in our predicted constituency vote shares, and the fact that even if our model captured adequately the systematic variation in change in party support (which it was by no means guaranteed to do), there would still be additional random variation. So instead, we used the predicted vote shares as the basis for a *probabilistic* prediction for each seat, an approach that has been central to the BBC's approach to election night forecasting since 1970 (Brown and Payne, 1975, 1984; Brown *et al.*, 1999). For example, the predicted percentage vote shares in the Ribble South constituency, on the basis of our models estimated from the exit poll, were: Con 37.8, Lab 41.8, LD 19.1, Oth 1.3. Rather than identifying Labour as the winning party in Ribble South on the basis of those predicted shares, we computed a corresponding set of probabilities as to who would win that seat: Con 0.27, Lab 0.73, LD 0.00, Oth 0.00. Thus, although our model predicted that Labour would have slightly more votes than the Conservatives in Ribble South, the uncertainty was such that we still gave the Conservatives a 27% chance of winning the seat.[10] The expected House of Commons seat totals were then obtained simply by summing each party's probabilities over all constituencies, and then rounding the result to whole numbers for presentation. It should be noted that the aim was purely one of accurate point prediction: while we could, in principle, have worked towards a rather more

11

informative prediction *interval* for the House of Commons seats totals, that was not our brief.

The probability calculation itself rests on an assessment of uncertainty about the predicted vote shares derived from our exit-poll regression models. A thorough statistical-modelling approach to this would demand a fully Bayesian analysis, with proper quantification of model uncertainty, of uncertainty about sources of bias, etc., as well as sampling variation. Partly due to lack of time, and partly for reasons of transparency — we wished to avoid such complications as multiple models and simulation-based inference — a much simpler, entirely *ad hoc* approach was adopted. The probability for party j to win in constituency i was calculated from the predicted percentage vote shares s_{ij} as $r_{ij}/\sum_j r_{ij}$, where

$$r_{ij} = \exp\left[-\left(\frac{\max_j\{s_{ij}\} - s_{ij}}{\sigma}\right)^\lambda\right], \tag{1}$$

and σ, λ are adjustable 'tuning' constants. The value of λ in this formula is chosen to reflect the anticipated shape of the distribution of errors in the estimated shares $\{s_{ij}\}$: roughly speaking, the value $\lambda = 1$ corresponds to a 'heavy-tailed' error distribution with exponential tails, while larger values of λ correspond to lighter-tailed distribution shapes in which gross errors are, relatively speaking, less frequent. The constant σ is a scale parameter, controlling the amount of dispersion in the implicit error distribution: a small value of σ indicates that the likely errors in predicted vote shares are small, and produces probabilities that are correspondingly close to 1 for the party with the highest predicted share in a seat and close to 0 for the other parties; larger values of σ represent greater uncertainty in the predicted vote shares, and yield correspondingly more moderate probabilities. The use of (1) does not — as far as we know, for any choice of λ and σ — produce probabilities that are consistent with any particular multivariate distribution for the $\{s_{ij}\}$. This means, in particular, that the constants λ and σ do not have simple interpretations in terms of familiar distributional properties; for example, the 'best' value of σ to use in (1) cannot be determined by estimating — or by eliciting belief about — the joint dispersion of likely errors in the $\{s_{ij}\}$.

On 5 May 2005 the tuning-constant values actually used when applying formula (1) were $\lambda = 1.5$ and $\sigma = 4$. Their impact when included in formula (1) is exemplified by the Ribble South probabilities quoted above. These values were chosen on the basis of some experimentation with the 2001 exit-poll data and actual election results, where it was found that the choices $\lambda = 1.5$ and $\sigma = 3$ would have produced well-calibrated probabilities based on the regression models that were actually used for the 2001 poll at the BBC. The slightly larger value of 4 for σ was used for the 2005 election since it was thought likely that there would be more unexplained variation in change than there had been in the 2001 general election.[11] For an assessment of how realistic the actual probabilities underlying the 10 p.m. forecast were, see Section 4.3.

2.5 Getting the Data

Not only did the BBC and ITV combine their forces in commissioning an exit poll, but in responding to the invitation to tender so also did the two organisations who had conducted the fieldwork for the two broadcasters in 2001. NOP (who had conducted all of the BBC's exit polls since 1992) and MORI (who conducted ITV's 2001 poll) presented a joint bid for the 2005 poll, and divided the fieldwork between them. At each polling station, the proportion of voters who were approached to complete the mock ballot paper was in inverse proportion to the number of electors registered to vote (other than by post) at that station. This meant that the poll was designed to produce roughly equal numbers of 'interviews' per polling station, with a target of an achieved sample of 225 interviews at each station — or a total of 27,000 interviews across all 120 stations, a large proportion of which, as indicated earlier[12], would take place relatively late in the day, leaving little time for modelling and analysis.

Meeting the challenge of processing this amount of data in a short period of time has, however, been facilitated by developments in information technology. Although the original poll data were simply phoned through to NOP's telephone call centre, thereafter they could be made available to the analysis team using a secure network connection. This was used to deliver data speedily on a regular basis throughout the day, with a particular emphasis on the evening 'rush hour'. The poll data were made available at two levels — the individual respondent and, more importantly, aggregated to the level of each individual polling station. The latter data could then be linked to the wide range of contextual information that had been compiled about the parliamentary constituency in which each station was located, and it is this combined file of information that was the key resource used in the modelling of the data.

The first few data deliveries on polling day were used mainly to check data quality and identify problems to be resolved by the two polling organisations. However, from about 4 p.m., graphical exploratory methods, automated forward search and tabulation were used in order to suggest, on the basis of the data collected so far, candidate statistical models for the emerging pattern of electoral change, and to identify possible outliers in the data (which were either downweighted or removed). Thus our ability to assimilate and model a large amount of data in a short period of time depended heavily on the ability provided by relatively recent developments in information technology to write interactive programs both to visualise data and to test and retest in rapid succession a series of statistical models together with associated diagnostics.[13]

After about 7 p.m., formal statistical models were fitted to each new data drop, and a wide array of standard diagnostic tools were used to criticise these models, to identify un-modelled patterns of change, etc. By this stage we were simply working with a fairly small number — two or three — of candidate regression models for each of the two most important change variables, namely the Liberal Democrat performance and the swing from Conservative to Labour. The different candidate models used different predictors, and had different interpretations: they had in common a roughly equal degree of fit to the data. A possibility from this point on would

have been to proceed by using model averaging, in essence constructing a (possibly weighted) average of the predictions made by the various candidate models under consideration. Such an approach has strong statistical appeal in terms of likely predictive accuracy, but would have made life difficult for the journalists and presenters tasked with explaining the exit poll 'story' to the public. Ultimately the 10 p.m. forecast was made on the basis of one selected model for each of the two main change variables; the models chosen gave a predicted Labour majority in the middle of the range indicated by our various candidates, and provided a basis for briefing journalists and presenters on the patterns of systematic variation found in the exit poll.

3   How good was the exit poll?

3.1   Introduction

Having described how our methodology attempted to meet the various challenges that faced the 2005 exit poll, we now turn to an evaluation of the results. This will fall into two parts. In the present Section we assess the quality of the data provided by the exit poll itself, considering not only the accuracy of the data in the polling stations that were covered, but also the adequacy of the coverage of polling stations. Thereafter, in Section 4, we examine how well our statistical modelling approach actually worked, and assess its contribution to the accuracy of the forecast.

3.2   Estimated versus actual change

One simple test of the ability of the exit poll to estimate the change in vote share in the stations it covered is that the change as measured by the exit poll at any polling station should be a good estimate of the change in the constituency where the polling station is located. Figure 2 plots, for the 120 polling stations covered, the change in the share of the vote for each of the three main parties as estimated by the exit-poll (leaving aside the information available on refusers) against the actual change in the share of the vote in the constituency in which each station was located. The line drawn on each graph indicates where perfect estimates would lie. All three plots show fairly widely scattered estimates, though the change in the Liberal Democrat vote has the smallest error variance. For the Conservatives and Labour the correlation between the polling station estimate of change and the the actual constituency change is just 0.43 (in both cases), while for the Liberal Democrats it is 0.61. Similarly, the mean absolute deviation of estimated from actual change was 4.2% in the case of the Conservatives, 4.6% for Labour, but only 3.6% for the Liberal Democrats.

In each panel of Figure 2 it would appear that the errors are distributed more or less evenly on both sides of the line of perfect estimates, though there is some apparent evidence that the exit-poll data tended to overestimate Labour and underestimate Liberal Democrat performance. Comparison of the first and third rows of Table 2 provides a simple numerical summary: for the
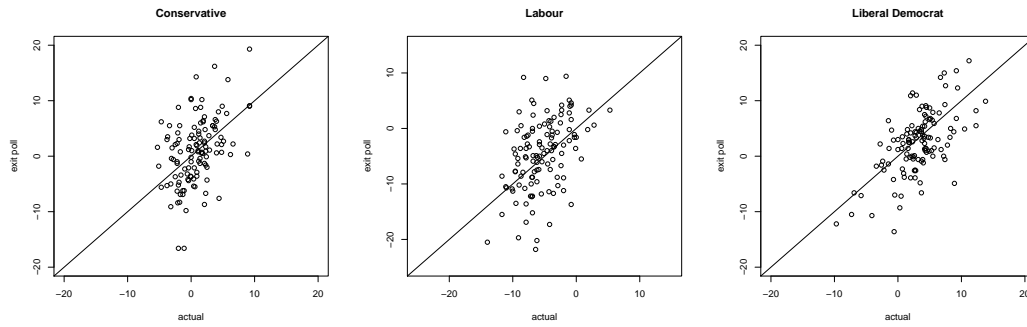
Figure 2: Exit-poll estimates of change in vote share 2001–2005 for the three main parties, versus actual change in the constituency polled.

Table 2: Accuracy of exit poll data: mean change 2001–2005 in vote share for the three main parties, in percentage points

	Con	Lab	LD
Poll ('raw' data)	+0.9	−4.6	+2.4
Poll (with guesses)	+1.3	−5.2	+2.6
Results (sampled seats)	+0.8	−5.3	+3.2
Results (all seats)	+0.2	−5.9	+4.0

constituencies of our 120 polling places, the (negative) mean change in the Labour vote share was underestimated by 0.7 points and the mean increase in the Liberal Democrat vote share was under-estimated by 0.8, while the mean change for the Conservatives in the exit poll was almost identical to that in the actual results. An indication of the strength of this apparent evidence of bias in the poll data can be obtained from a set of simple two-sided t-tests: for both the Conservatives and Labour, such a test finds the evidence to be very far from significant at conventional levels, though as a result of the smaller error variance for the change in the Liberal Democrat vote the computed p-value (0.048) in that case is just significant at the 5% level. A slightly more general, two-parameter regression test does not find any significant evidence against the hypothesis that the intercept is zero and the slope is unity, for any of the three parties. On this basis, any apparent evidence of bias in the exit poll should be regarded as inconclusive: the small biases that are suggested by Table 2 could quite easily have arisen by chance.

Also shown in Table 2, in the second row, is the mean estimated change that results from incorporating into the exit-poll data the information on interviewer guesses that was obtained in order to impute the responses of refusers. As can be seen, this brings the Labour share of the vote more or less in line with the actual outcome in the sampled constituencies and reduces the apparent Liberal Democrat error. However, it also makes the Conservative figure less accurate.[14] With the guesses included, none of the mean deviations is significantly different from zero (the Liberal Democrat mean error in particular becomes −0.6, with p-value 0.22), and it is still the case that none of the regression lines differs significantly from a unit slope through the origin. However, for two of the three parties the error variance is higher: the mean absolute deviations

15

are Con 4.7%, Lab 4.6%, and LD 4.1%. It was because of this final characteristic that, in the event, the decision was taken during polling day not to incorporate the information on the estimated behaviour of refusers in the final modelling of the poll.

3.3 Coverage

Even if the exit poll was reasonably accurate in its estimate of what happened in the constituencies that it did cover, its ability to act as an effective means of assessing the tally of votes and seats across the country as a whole could still be compromised if it did not adequately cover all the various kinds of constituencies where the change in party performance deviated appreciably from that across the country as a whole. We noted earlier that the sample of polling districts focused primarily on those seats that were marginal between Labour and the Conservatives, and between the Conservatives and the Liberal Democrats. It thus contained few seats where the Liberal Democrats were challenging Labour. There were in truth few of these in the country as a whole: in 2001 there were only 9 seats in which the Liberal Democrats were second to Labour but no more than 20 points behind, and only 52 where they were second at all. As can be seen from Figure 1 above, in the event there were only three such seats in our sample, including just one where the Liberal Democrats were less than 20% behind. But as can be seen in Figure 3, which shows how every constituency moved across a ternary diagram between 2001 and 2005, in practice there were some large movements from Labour to the Liberal Democrats in many seats where the Liberal Democrats started as second to Labour. Indeed the Liberal Democrat vote rose on average by 7.8 points in seats where they started as second to Labour, nearly twice the national average of 4.0 points, while Labour's own vote fell by 7.1 points, compared with the national average of 5.9 points, a pattern that helped the Liberal Democrats to capture no fewer than eleven seats from Labour.[15] With data on more Labour/Liberal Democrat seats, we would almost certainly have improved our model for change in the Liberal Democrat share in particular, and thus would have been more likely to predict successfully that several apparently 'safe' Labour seats would fall to the Liberal Democrats. Meanwhile, as we can see if we compare the third and fourth rows of Table 2, more data from such constituencies would also have helped bring the poll's estimate of Labour and Liberal Democrat performance across the country as a whole into line with the actual results. As it was, in the three seats covered by the poll where the Liberal Democrats were standing second to Labour, neither the Labour nor the Liberal Democrat performance proved to be particularly exceptional.[16]

The problem was of course not one that could readily have been avoided, since the polling locations were inherited from the separate 2001 exit polls. However, even if the 2005 locations were to be used again in a similar exercise at the next election, then — boundary changes apart[17] — the problem would be rather less severe since the void has now been occupied to some extent by the movement of several constituencies in 2005 from 'safe' Labour territory to the Liberal Democrat target zone.
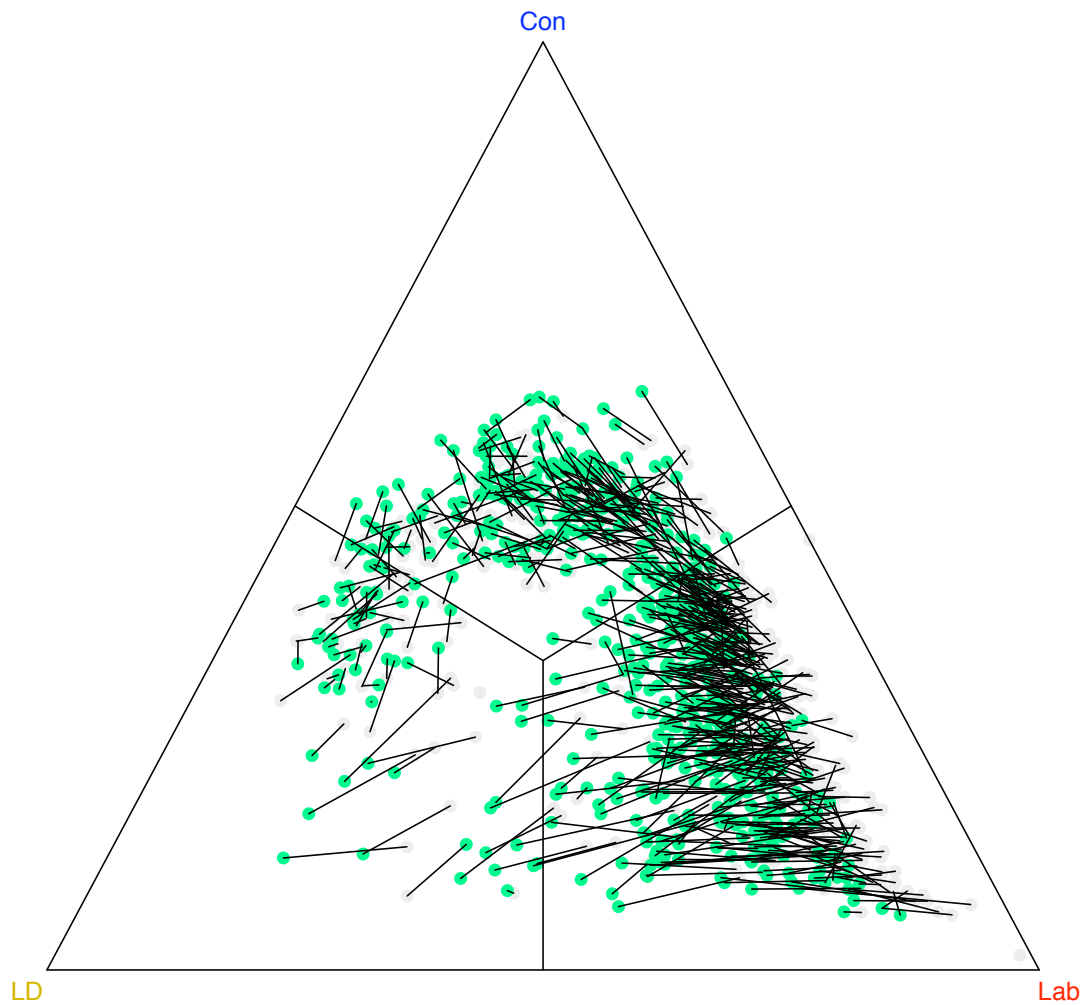
16

Figure 3: Ternary diagram showing the change in actual vote shares for the three main parties. Green points are 2005 positions of constituencies; the black 'tails' show the movement from their 2001 positions. A small number of 'special' seats — see note 7 — are excluded from the graph.

Table 3: Estimated and actual vote shares for the constituencies of the four new polling places in London. The exit-poll estimates exclude interviewer guesses for refusers.

	Exit-poll estimates (%)			Constituency result (%)		
Constituency	Con	Lab	LD	Con	Lab	LD
Camberwell and Peckham	5.2	62.8	23.0	9.8	65.3	18.8
Dulwich and West Norwood	13.8	47.9	25.1	21.9	45.4	24.4
Enfield North	29.4	47.7	7.5	39.6	44.3	11.4
Old Bexley and Sidcup	45.4	30.9	12.1	49.8	27.5	14.7

3.4 The four new London locations

We noted earlier that one particular difficulty facing the poll was a relative paucity of sampling points in London, and that as a result in the capital four new polling districts thought to be representative of the constituency as a whole were included in the sample. Table 3 assesses how effectively this was achieved by showing the shares of the total vote, as estimated by the exit poll in the selected polling district, and actual outcome in the constituency in which they were located.

Table 3 suggests that there was some success in achieving that aim, except that the Conservative vote was appreciably under-estimated in all four cases. It might be thought that this discrepancy could be due to refusal bias, and that estimates derived from interviewers' guesses of refusers' votes would help remove it. However, the estimated Conservative shares with this information included are, at 4.9%, 13.8%, 29.2% and 44.7%, scarcely different from the data seen in the first column of Table 3. In the event, the estimates from these four locations had little impact on the forecast derived from the exit poll, since they were given reduced weight in our regression analyses to reflect the lack of 2001 'baseline' exit-poll data. Their presence in the panel of polling places should, however, be beneficial if a similar exit poll takes place at the next general election.

4 From exit poll to seats forecast

4.1 Models for exit-poll change

We now turn to an assessment of how the exit poll data were modelled. We described earlier that our initial step was to develop a model of change in four measures of party support. However, the exit poll itself contained almost no information about the variation in Other, SNP or PC voting, so in practice most of the modelling effort was focused on the other two change variables, namely the change in the Liberal Democrat share of the three-party (LD, Con., Lab.) vote, and the change in the Labour share of the two-party (Con., Lab.) vote. Figures 4 and 5 display graphically these two estimates of change produced by the exit poll in each selected polling district. In Figure 4 we simply show the estimated change in the Liberal Democrats' share of the vote for the three main parties against the change in Labour's share of the two-party vote. The figure has three key

18

features:

(i) in general, a fall in Labour's vote share of the two-party vote (median change -3.0)

(ii) in general, an increase in Liberal Democrat share of the three-party vote (median $+3.5$), and

(iii) a very wide scatter of points (especially in the case of the Liberal Democrat change variable), rather wider than would be expected if there were simply sampling variation around a pattern of uniform change.

The third feature certainly suggests that there was a need to try to identify and model the existence of systematic variation in the exit poll data in order to improve the accuracy of the forecast. Meanwhile, Figure 5, which presents the exit poll estimates of change in a ternary diagram, suggests that the pattern of change since 2001 did differ according to the strength of the parties in a constituency in 2001. For example, the poll suggested that the Liberal Democrats were gaining votes in 'safe' Labour seats (where Labour were performing particularly badly) but were doing less well in seats they were contesting with the Conservatives.

The estimated coefficients for the two models actually used for the 10 p.m. forecast were as shown in Table 4.[18] The model for Con.-Lab. swing was in fact relatively simple, with just two predictor variables. First it suggested that Labour was losing ground rather less sharply relative to the Conservatives in seats where turnout had been relatively low in 2001. Seats with low turnouts are typically Labour strongholds in large cities, while those with higher turnouts tend to be Conservative-held seats in rural areas (Curtice and Steed, 2001). In short, it would appear that the poll anticipated that the two-party swing away from Labour was least in some of its strongholds. At the same time, however, the model also suggested that after taking this relationship into account Labour did in fact perform rather better in seats that were particularly rural.[19].

In contrast, the final model for change in Liberal Democrat support, which as we have seen was more variable, was more complex and contained some fairly large estimated effects. In particular the party was expected to do less well in seats the party was defending and in those where it was challenging the Conservatives — though this latter pattern was expected to be less strong in more marginal seats (defined as those where the Liberal Democrats were less than 10 percentage points behind the Conservatives). At the same time there was a substantial predicted bonus for the Liberal Democrats in Scotland.

But how well did these models capture the variation that actually happened across the country as a whole? (For a comprehensive analysis of the variation in the change in party support between 2001 and 2005, see Curtice *et al.*, 2005). Given that we have already noted that Labour often performed relatively badly in some of its strongest seats, the fact that our model suggested that the swing away from Labour was lower in such seats might appear erroneous. However reference back to Figure 3 indicates that in many of these seats there was also a movement away from the Conservatives. Moreover, this fall constituted a relatively large proportion of the existing small Conservative vote in such seats, and thus was sufficient to push the swing away from Labour in
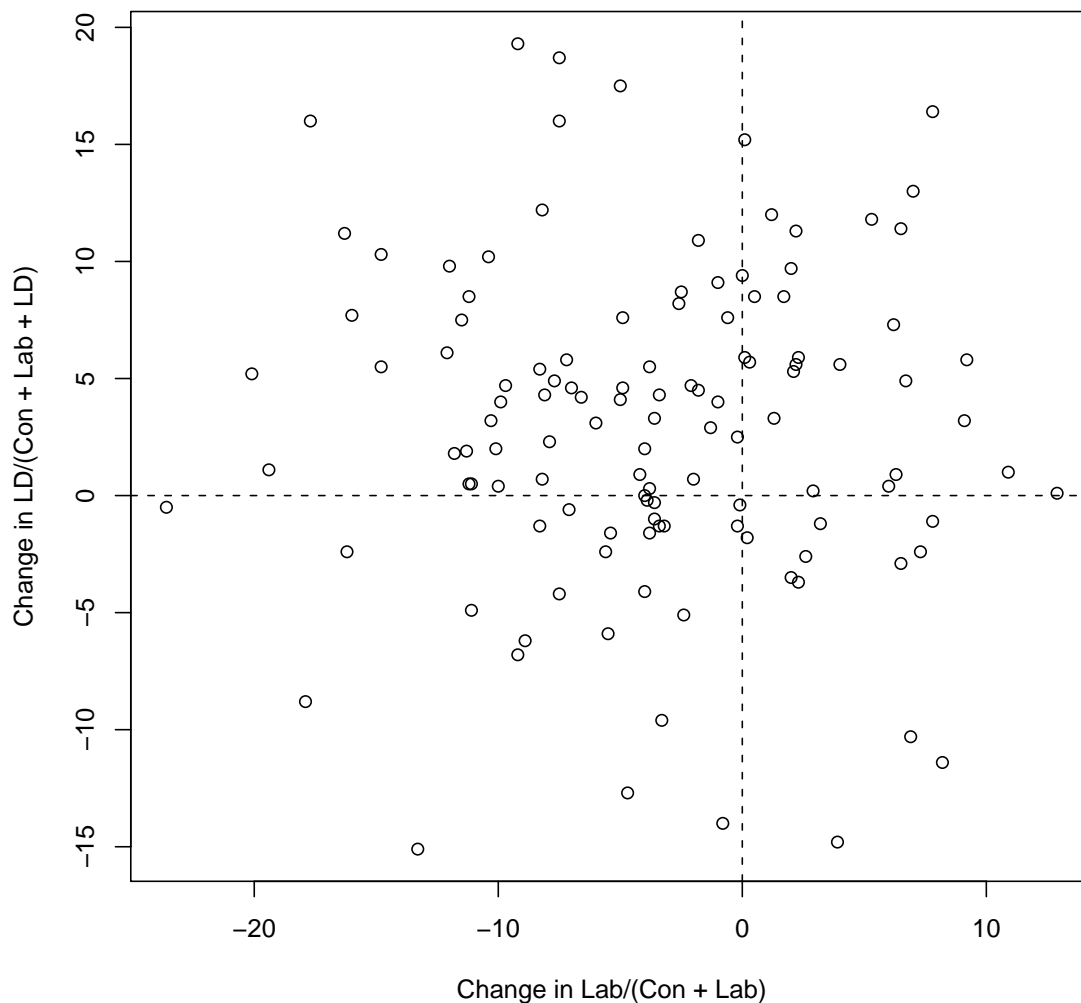
19

Figure 4: Estimated change 2001–2005 in shares of the three-party vote, expressed in percentage points, at the exit-poll locations. Data as at 9.25 p.m. on 5 May 2005.

Table 4: Estimated coefficients in the two regression models that were selected as the basis for prediction.

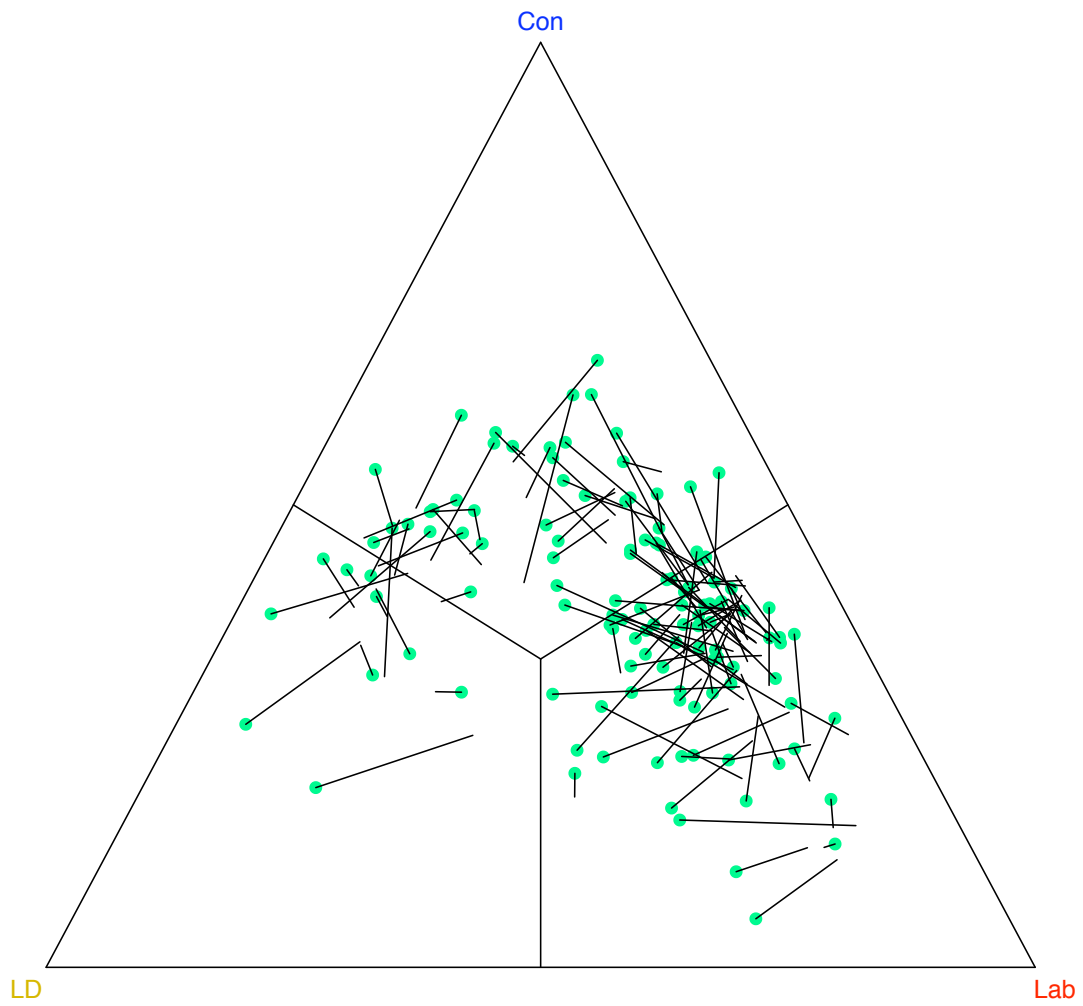| Change in Lab/(Con + Lab)		Change in LD/(LD + Con + Lab)	
predictor	estimate	predictor	estimate
Intercept	−4.3	Intercept	3.9
'Turnout 2001'	−1.8	Scotland	13.7
'Country dwellers'	2.5	'Con first; LD second'	−9.5
		'LD first'	−6.6
		'Con/LD marginal'	5.7

20

Figure 5: Ternary diagram showing the impact of estimated change on three-party vote shares, for the constituencies represented in the exit poll. Green points are estimated 2005 positions of the exit-polled constituencies; the black 'tails' show their 2001 positions. Data as at 9.25 p.m. on 5 May 2005.

21

such seats to below the national average. Overall the two-party swing in seats that Labour was defending was just −2.4 compared with −3.0 across the country as a whole. On the other hand, Labour's vote did hold up relatively well in those seats where the Liberal Democrats were fending off a Conservative challenger (two-party swing −0.7); such seats typically are particularly rural. Thus the two variables in the model appear to have helped identify some key variation in the pattern of two-party swing.

However, the model did fail to incorporate one important pattern that also occurred. This was that the Conservatives performed relatively well in much of the south-eastern corner of England. The mean two-party swing in London together with the South East government region was −4.8%, almost two points higher than the mean for the country as a whole. Moreover, thanks to the relatively high proportion of marginal seats in this part of the country, this above average performance brought the Conservatives an important dividend in terms of seats (Curtice, 2005b). Yet in the exit poll the average swing in polling districts in London and the South East was indistinguishable from that found in the poll as a whole.

The Liberal Democrat model also identified some important features of that party's performance. As we already have had reason to note, the party performed better in seats where it was in competition with Labour than it did in those where it was attempting to challenge the Conservatives. Indeed on average the party made little or no progress in seats where it shared first or second place with the Conservatives, and it actually suffered a net loss of two seats to them. The Liberal Democrats also advanced more strongly in Scotland. However, in each case the size of our estimated coefficient was too strong. For example, rather than the Liberal Democrat party performing around seven percentage points worse than average in seats that were Conservative/Liberal Democrat marginals, it did so by only just over three points. Equally it advanced more strongly in Scotland than it did in England, but only by some three points, not fourteen![20]

Meanwhile here too there are some apparent gaps in the model. Two of the most striking features of the Liberal Democrat performance were that the party's share rose particularly strongly in seats with relatively large numbers of Muslims and also in those with a relatively high proportion of students. In the case of the former this may well have been because the exit poll only covered two polling districts in constituencies where, according to the 2001 census, more than 10% of the population was Muslim. The Liberal Democrat share of the three party vote was indeed estimated to have risen by well above the national average in these locations (by 18 points in one instance and 12 in another), but the evidence of just two sampling points was insufficient to justify the inclusion of a Muslim variable in the model. Given that there are only 39 constituencies in total where more than 10% of the population is Muslim, having sufficient sampling points in relevant constituencies to detect such a pattern is always going to be difficult. In the case of constituencies with large student populations the poll was somewhat more representative: it contained 15 districts in constituencies where more than 10% of the population were students according to the 2001 Census. Unfortunately, however, in the sampled constituency with the

22

Table 5: How the forecast Labour majority responds to two key aspects of the modelling/prediction approach

	Non-Probabilistic $(\sigma = 0)$	Probabilistic $(\sigma = 4)$
Uniform change	100	88
Regression models	84	66

largest proportion of students, Oxford West and Abingdon (a seat already held by the Liberal Democrats), the Liberal Democrat share of the vote actually fell back by one point, an outcome (exaggeratedly) anticipated by the poll. This helped ensure that our model did not anticipate what proved to be a 3 point above average increase in the Liberal Democrat vote in seats with more than 10% students.

4.2 The route to '66'

Our attempts to identify the systematic sources of variation in party performance were evidently only partly successful. We might then wonder how effective our modeling was in enabling us to improve the accuracy of our prediction of the Labour majority. At the same time we might ask what the impact was of using a probabilistic approach to forecasting the outcome. These questions are addressed in Table 5, which compares our actual forecast majority with what we would have forecast if we had assumed that the change in party performance was uniform rather than variable as anticipated by our regression models, and what would have happened if we had simply assigned each seat to the party with the largest predicted share of the vote rather than using the probabilistic approach.

We can see that if we had both assumed that the change in party support was uniform and adopted a non-probabilistic approach to forecasting the outcome in seats, we would have forecast a majority of no less than 100 for Labour. This is close to what would have been produced by familiar 'swingometer'-type calculations[21]; for example a 3% swing from Labour to Conservative would have implied a Labour majority of 104 seats in the 2005 parliament, according to the BBC Online interactive calculator, which at the time of writing can still be found at `http://news.bbc.co.uk/1/shared/vote2005/swingometer/html/labcon.stm`. Using our regression models but without using a probabilistic approach helps to reduce the error, but still leaves Labour with a forecast majority of 84. Using our probabilistic approach (with $\sigma = 4$) without the regression models would also have reduced the error, but would still have left Labour with a forecast majority of 88. Clearly both aspects of the modelling approach, that is the modelling of both systematic and 'random' variation in change, made appreciable contributions to the accuracy of the forecast majority.

It is worth examining in particular why using a probabilistic approach proved to be important at this election. The distribution of seats around the average two-party swing of 3% was highly asymmetric: whereas there were just 18 seats that would switch from Labour to Conservative
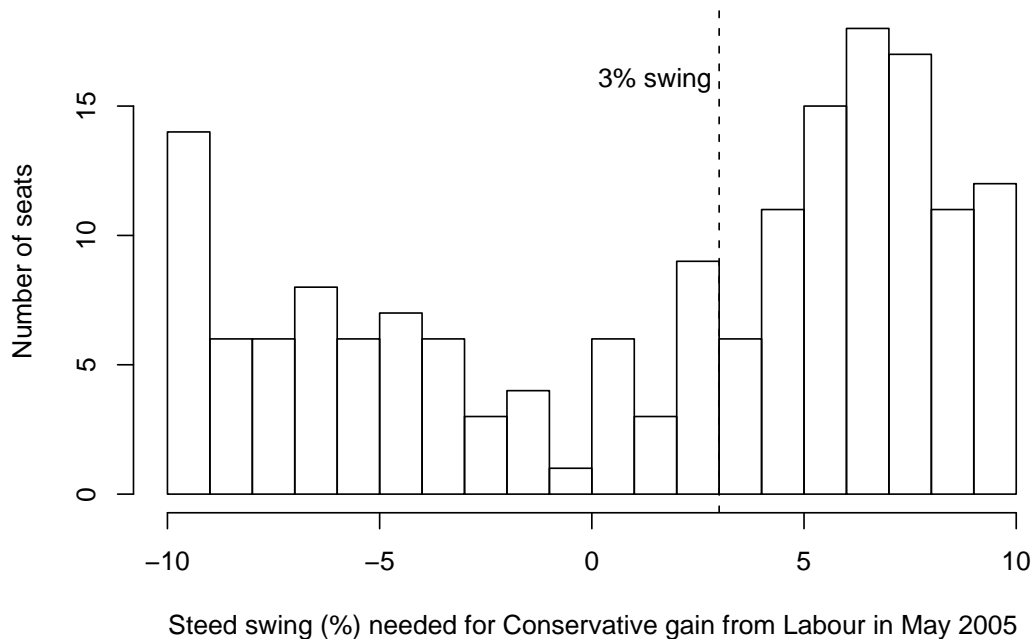
Figure 6: Parliamentary seats closely contested between Labour and Conservative parties. Swing needed is calculated as half the difference in share (%) of the two-party vote at the 2001 election.

on a swing of 0–3%, there were no fewer than 32 that would switch on a swing of between 3–6% (see Figure 6). Thus any variation around the national average was almost certain to result in the Conservatives winning more seats than would be anticipated if the swing were entirely uniform. As it happened a relatively large proportion (38%) of the seats vulnerable to a swing of between 3% and 6% were in London and the South East, and thus in practice our probabilistic approach helped us to overcome the fact that our regression model for the poll did not anticipate the somewhat better Conservative performance in that part of the country.

We can also assess the contribution made to our forecast by the adjustment that was ultimately made for the potential effect of increased postal voting. As that adjustment was only a small one, it inevitably had only a small impact on the forecast. Without the adjustment, the 10 p.m. forecast majority would have been 62 seats rather than 66.

4.3   How well calibrated were the probabilities?

We have seen in Section 4.2 that the use of probabilities in predicting party seat totals was important. It might though still be asked, in retrospect, 'How good were the probabilities?' A good probability statement should, minimally, be 'well calibrated' (e.g., Dawid, 1986); for example, if a statement is made that some event has a probability of 0.8, then the event should occur in approximately 80% of relevant cases. To illustrate the calibration properties of the probabilities underlying our 10 p.m. forecast, Table 6 summarises all of the Labour probabilities that were

24

Table 6: Calibration performance of Labour probabilities. Seats where the Labour probability was less than 0.001 or greater than 0.999 are excluded.

Lab probability (%)	Lab wins/possible	(%)
(0, 10]	0/16	0
(10, 20]	0/7	0
(20, 30]	2/8	25
(30, 40]	2/7	29
(40, 50]	8/19	42
(50, 60]	9/14	64
(60, 70]	8/10	80
(70, 80]	15/19	79
(80, 90]	21/25	84
(90, 100)	31/32	97

calculated to be between 0.001 and 0.999. For example, there were 8 seats where Labour was given a forecast probability between 20% and 30% of winning, and they eventually won 2 of those 8 (i.e., 25% of them). The degree of calibration exhibited in Table 6 seems remarkably good, especially in view of the rather *ad hoc* nature of the probability calculation itself. The proportion of seats actually won by Labour in any probability interval was inside the relevant range in 6 out of the 10 cases, and in no case was it above or below the end points of the range by more than 10 points. Certainly Table 6 does not indicate gross mis-calibration; and the same applies to the corresponding probabilities, not shown here, for the Conservative and Liberal Democrat parties.

5   Concluding remarks

At the beginning of this paper, we identified four particular challenges facing the conduct of an exit poll in Britain at the time of the 2005 election: the lack of data on actual vote counts in individual polling districts; the uncertainty about the relationship between seats and votes created by the single member plurality electoral system; a marked increase in the proportion of the electorate opting to vote by post; and a tendency for votes to be cast late in the day. These complications were of course in addition to those that face exit polls the world over, such as the danger that supporters of one party may be more inclined than supporters of another to refuse to say how they voted, or perhaps to respond untruthfully. Our principal solutions to these problems were as follows:

(i) to focus on estimating the change in party support rather than the level, using data from a previous exit poll as a baseline,

(ii) both to attempt to estimate systematic variation in the change in party support and to adopt a probabilistic approach to forecasting the outcome in seats,

(iii) to collect additional data in advance of polling day to estimate any change in the political composition of postal voters, and

(iv) to exploit techniques of graphical visualisation and model diagnosis and re-estimation

25

made possible by relatively recent developments in information technology.

At the same time, an attempt was made to estimate the likely behaviour of refusers by getting interviewers to guess the voting preference of both respondents and non-respondents.

Many of these solutions proved to be important in helping us produce a good forecast of the majority in 2005. Thus, although the large increase in the number of postal voters resulted in only a small change in their political composition, if we had not estimated it we would have slightly overestimated the swing against Labour. Without the ability to diagnose model deficiencies and re-formulate models quickly it would have been impossible to analyse adequately the rush of exit poll data that arrived late in the day. The use of such modelling proved to be essential in arriving at an accurate estimate of the government's majority, as was the use of a probabilistic approach to generating that estimate. In this latter respect our experience echoes in particular that of Brown and Payne (1975, 1984) and Brown *et al.* (1999) in their work on election night forecasting for the BBC at previous British elections.

But perhaps the most important single element of the approach was to use the poll, by comparing it with its predecessors in 2001, to estimate the change in party performance rather than its level. This certainly freed us from the virtually impossible requirement in the British context of selecting a nationally representative sample of polling locations. But it had other advantages too. Any biases in the conduct of the poll, such as differential refusal, would not matter so long as much the same biases were present also in the previous poll.[22] By focusing on a variable with less variance we had in any case more prospect of accurately estimating what had happened across the country as a whole in terms of votes.

Of course, we fully acknowledge that getting the headline forecast *exactly* right owed a good deal to luck, as well as to methodology. Even so, it was no mere flash in the pan. At the previous election in 2001 the BBC undertook a similarly designed if somewhat smaller exit poll. It was analysed using much the same methods as those described here, and when broadcast at 10p.m. the resulting forecast of the size of the Labour majority was accurate to within 6 seats (Payne, 2003, p.214). Those methods are designed to reduce substantially the impact of certain potential sources of bias and variation, and as such they are generic and might be expected to perform well under similar circumstances elsewhere and in the future. While the use of modelling and a probabilistic approach may only be of particular importance where plurality or majoritarian electoral systems are used, others of our solutions could usefully be adopted irrespective of the electoral system. In particular, we would suggest that using information from a previous exit poll rather than just the current one is an approach that should be considered even where official information is available about how individual polling districts voted at the previous election, and can be used to construct a nationally representative sample of polling districts. For if exit polls suffer from methodological bias, perhaps for example as a result of differential refusal, they may still be inaccurate even if they are conducted in a perfectly representative sample of polling places. However, if that methodological bias is approximately constant between elections, then

it should be capable of being corrected by taking into account the results of the previous exit poll. Exit polling is not only about forecasting the future, but also learning the lessons of the past.

Acknowledgments

Notes

[1] This approach had been used in previous ITV exit polls. The choice of which constituencies should be covered by this exercise was determined by making a random selection from a list of London constituencies rank-ordered by Labour's share of the vote in 2001 — though this had to be done twice before four constituencies could be found where it was possible to identify a polling district that was apparently representative of the constituency in which it was located. We would acknowledge in particular the work of Roger Mortimore of MORI in this exercise.

[2] This, however, still represented a doubling of the proportion who were issued with a postal vote under the old regulations in 1997. Those regulations required a voter to be ill or away from home on polling day before they could be issued with a postal vote.

[3] In the event this estimate proved to be in line with that of the official analysis published after the election by the Electoral Commission (2005). This found that 12.4% of the eligible electorate in Great Britain was issued with a postal vote in 2005.

[4] This is consistent with the fact that the increase in postal voting since 2001 was greatest in the North of England where Labour are relatively strong (Curtice, 2005a)

[5] According to the data collected by MORI and NOP in 2005 the Conservative share of vote intentions for the Conservatives and Labour alone was eight points lower amongst those voting by post in 2005 than it was amongst those who did so in 2001 — but just 3.6 points lower amongst those voting in person, a difference of 4.4 points. Given that in 2001 when 4% of the electorate was registered to vote by post, 5% of all valid votes cast were cast by post (Electoral Commission, 2002), it was estimated that with around 12% registered to vote by post in 2005, 15% of all valid votes would be cast by post, an increase of ten points (an estimate that was subsequently confirmed in Electoral Commission, 2005). This suggested that the relative decline in support for the Conservatives amongst postal voters would cost the party approximately 0.1×4.4 per cent of the votes cast for the Conservatives and Labour alone.

[6] This estimate was based on interviewers' guesses for non-respondents, adjusted by a matrix of interviewer guess by reported vote amongst respondents.

[7] Constituencies in Northern Ireland, which are not contested by the main British parties, were not covered by or predicted from the exit poll. In addition four 'special' seats in Great Britain were not predicted from the exit poll: these were the Speaker's seat, one seat held since 2001 by an Independent MP, and two constituencies where there the seat had changed hands as a result of a by-election. Predictions for these seats were made separately, using subjective probabilities.

[8] In Scotland the 2001 results could not be used because of boundary changes following a reduction in the number of constituencies from 72 to 59. Instead we relied on estimates of what the outcome of the 2001 election would have been if the 2005 constituencies had been in place

(Denver *et al.*, 2004.)

⁹ This assumption was judged to be much safer than the use of potential alternatives such as estimating a model for 2005 turnout from the number of people approached for an interview at each polling station. In the event there was only a weak (positive) relationship between the Labour share of the vote in 2001 and the increase in turnout since 2001 (Curtice, 2005a)

¹⁰ Note that the two '0.00' probabilities here are not zero, rather they are rounded figures which represent probabilities smaller than 0.005: although our model predicted that the Liberal Democrats would achieve only about 20% of the vote in Ribble South, we could not *rule out* the possibility of a Liberal Democrat gain there.

¹¹ In practice while the change in the Liberal Democrat vote did vary rather more than it had in 2001, two-party swing varied rather less (Curtice *et al.*, 2005).

¹² In the event 21,250 interviews were conducted, of which 61% took place after 4p.m. Of this total, 20,008 were available at the time of the data transfer at 9.25p.m., which was the last to be made before the published forecast at 10p.m. The final delivery of data was made at 10.45, after the close of polls; although the opportunity existed at that time to update the published forecast on the basis of the new information, in the event the forecast majority after incorporating the 10.45 data remained at 66 seats, and no public revision was made.

¹³ The data manipulation, graphics, modelling and communication with external systems were all programmed in the open-source *R* statistical computing environment (R Development Core Team, 2007).

¹⁴ It should be noted that, in calculating this estimate, for those polling districts covered by MORI in 2001 we have had to incorporate into the baseline 2001 figure the data for those who were interviewed as replacements for refusers, rather than information derived from the procedure adopted in 2005. Thus our analysis is not as clear a test of the effectiveness of the procedure of using interviewer guesses to derive an estimate of the behaviour of refusers as it might have been.

¹⁵ Including two from third place.

¹⁶ On average the Labour vote fell by 2.0 points in these three seats, while the Liberal Democrats' vote increased by just 3.4 points.

¹⁷ Constituency boundary changes will be implemented in England and Wales at the next general election.

¹⁸ This model incorporates a small adjustment to take account of the effect of postal voting as described in section2.2. This involved reducing the intercept estimate by 0.4 percentage points, from −4.3 to −3.9.

¹⁹ The actual measure used was a transformation of the estimated proportion of people living

in postcode areas classified as Mosaic Group, 'country dwellers' as measured by Experian.

[20] The exaggerated coefficients in the Liberal Democrat model are in fact symptomatic of a methodological deficiency in our modelling. It would have been better to use a coefficient-shrinkage method rather than (weighted) least squares to estimate the regression parameters. This would have helped to mitigate the degree of coefficient inflation caused by predictor-variable selection and would have improved the mean-squared-error properties of predictions (e.g., Copas, 1983). Specifically, we might have used ridge regression (as used in Brown and Payne (1975), Brown and Payne (1984) and Brown *et al.* (1999), for the estimation of regression models after some constituency results become known), or, perhaps better, a hybrid shrinkage/selection method such as the lasso (Tibshirani, 1996). It had been our intention to implement and test such methods in time for the election-day analysis of the 2005 exit poll, but ultimately the pressure of other tasks prevented it; we note it here as a strong recommendation for future exercises of this kind.

[21] See, e.g., the Wikipedia entry 'Swingometer', at `http://en.wikipedia.org/wiki/Swingometer`, for the history and further details.

[22] Note that the same point applies, *mutatis mutandi*, to the pre-election polls that were used in order to estimate the behaviour of postal voters.

References

Blau, A. (2004). A quadruple whammy for first-past-the-post. *Electoral Studies* **23**, 431–53.

Brown, P. J., Firth, D. and Payne, C. D. (1999). Forecasting on British election night 1997. *J. Roy. Statist. Soc.* A **162**, 211–226.

Brown, P. J. and Payne, C. D. (1975). Election night forecasting (with discussion). *J. Roy. Statist. Soc.* A **138**, 463–498.

Brown, P. J. and Payne, C. D. (1984). Forecasting the 1983 British General Election. *The Statistician* **33**, 217–228.

Butler, D. E. (1951). An Examination of the Results. In: Herbert Nicholas (Ed.), *The British General Election of 1950*. London: Macmillan.

Copas, J. B. (1983). Regression, prediction and shrinkage (with discussion). *J. Roy. Statist. Soc.* B **45**, 311–354

Curtice, J. (1992). The hidden surprise: The British electoral system in 1992. *Parliamentary Affairs* **45**, 466–74.

Curtice, J. (2005a). Turnout: Electors stay at home again. *Parliamentary Affairs* **58**, 776–85.

Curtice, J. (2005b). Translating poll results into seats. Paper prepared for the British Polling Council. Available at `www.britishpollingcouncil.org/papers.html`.

Curtice, J. (2006). Neither accountable nor representative: First-past-the-post in Britain. Paper presented at a conference on 'Plurality and Multiround Elections', University of California at Irvine, 17–19 February 2006.

Curtice, J. and Steed, M. (1982). Electoral choice and the production of government: The changing operation of the electoral system in the United Kingdom since 1955. *British Journal of Political Science* **12**, 249–98.

Curtice, J and Steed, M.. (1986). Proportionality and exaggeration in the British electoral system. *Electoral Studies* **5**, 209–28.

Curtice, J. and Steed, M. (1997). Appendix 2: The Results Analysed. In: Butler, D. and Kavanagh, D. (Eds.), *The British General Election of 1997*. London: Macmillan.

Curtice, J., Fisher, S. D. and Steed, M. (2005). Appendix 2: The Results Analysed. In: Kavanagh, D. and Butler, D. (Eds.), *The British General Election of 2005*. London: Palgrave Macmillan.

Dawid, A. P. (1986). Probability forecasting. In Kotz, S., Johnson, N. L. and Read, C. B. (Eds.), *Encyclopedia of Statistical Sciences* **7**, 210–218.

Denver, D., Rallings, C. and Thrasher, M. (2004). *Media Guide to the New Scottish Westminster Parliamentary Constituencies*. Plymouth: Local Government Chronicle Elections Centre.

Electoral Commission (2002). *Postal Votes, Proxy Votes and Spoilt Ballot Papers at the 2001 General Election*. London: Electoral Commission.

Electoral Commission (2004). *The 2004 European Parliamentary Elections in the United Kingdom*. London: Electoral Commission.

Electoral Commission (2005). *Election 2005: Turnout: How Many, Who and Why?* London: Electoral Commission.

Gudgin, G. and Taylor, P. (1979). *Seats, Votes and the Spatial Organisation of Elections*. London: Pion.

Johnston, R., Pattie, C., Dorling, D. and Rossiter, D. (2001). *From Votes to Seats: The Operation of the UK Electoral System since 1945*. Manchester: Manchester University Press.

Kendall, M. G. and Stuart, A. (1950). The law of cubic proportions in election results. *British Journal of Sociology* **1**, 183–196.

Laakso, M. (1979). Should a two-and-a-half law replace the cube law in British elections?. *British Journal of Political Science* **9**, 355–62.

Miller, W. L. (1977). *Electoral Dynamics in Britain Since 1921*. London: Macmillan.

Mosteller, F. and Tukey, J. W. (1977) *Data Analysis and Regression*. Cambridge, MA: Addison-Wesley.

Payne, C. D. (2003). Election forecasting in the UK: The BBC's experience. *Euramerica* **33**, 193–234.

R Development Core Team (2007). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing, URL `http://www.R-project.org`.

Steed, M. (1965). An Analysis of the Results. In: Butler, D. and King, A (Eds.), *The British General Election of 1964*. London: Macmillan.

Taagepera, R. and Shugart, M. (1989). *Seats and Votes: The Effects and Determinants of Electoral Systems*. New Haven: Yale University Press.

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *J. Royal. Statist. Soc.* B **58** 267–288.

Tufte, E. R. (1973). The relationship between seats and votes in two-party systems. *American Political Science Review* **67**, 540–54.

Upton, G. J. G. (1976). The diagrammatic representation of three-party contests. *Political Studies* **24**, 448–54.

Upton, G. J. G. (1994). Picturing the 1992 British General Election. *J. Roy. Statist. Soc.* A **157**, 231-252.