Electoral Studies 30 (2011) 269–277

Contents lists available at ScienceDirect

Electoral Studies

journal homepage: www.elsevier.com/locate/electstud

Forecasting the 2010 general election using aggregate local election data☆

Colin Rallings*, Michael Thrasher, Galina Borisyuk, Elena Long

Department of Political Science, LGC Elections Centre, University of Plymouth, PL4 8AA, UK

ARTICLE INFO

Article history:
Received 17 September 2010
Accepted 22 September 2010

Keywords:
British election
Election forecasting
Political party

ABSTRACT

The paper presents a revised method for estimating national vote shares using aggregate data from local government by-elections. The model was originally developed to forecast the annual outcome of local elections but was adapted in time to provide an accurate forecast of Labour's landslide victory at the 1997 general election. However, over the past decade the changing pattern of party competition which has seen parties becoming more selective about which elections to contest has led to more elections being excluded from the modelling because they failed to meet the exacting criteria that all three major parties, Conservative, Labour and Liberal Democrats, had contested both the by-election and the previous main election, normally held in May. Relaxing these criteria, although increasing the number of available cases would adversely affect the forecast, over- or under-estimating party votes. Instead, the revised method overcomes the problem of differential competition by estimating vote shares for parties that contest one but not both elections. A further innovation is the calculation of a weighted moving quarterly average which takes account of the number of days elapsed between the by-election date and the date of forecast. Using the new method we provide estimates for likely party shares for the 2010 general election.

1. Introduction

Unlike most other election forecasting models, the example described here is primarily designed not to forecast a national parliamentary election but instead to forecast national equivalent votes at annual local elections. Judged by this criterion it has proved successful. The model uses aggregate level data obtained from local council by-election results from the early 1980s onwards that take place in virtually every week of the year. It operates by calculating change in vote share across two elections, the main election and the subsequent by-election, using cases that feature candidates from all three of the main parties, viz., Conservative, Labour and Liberal Democrats at both types of contest.. The theoretical basis behind the approach is that, unlike their parliamentary equivalents that generate a media circus and become a vehicle for voters' protests, local council by-elections are relatively straightforward electoral events where people behave in a normal manner and where any idiosyncratic outcomes have a way of cancelling themselves out over the longer term.

From 1995, when we began using the model for forecasting the May local elections, it immediately demonstrated its value, so much so that we used it to forecast the 1997 general election and were pleased to see that it outperformed the national polling companies (Rallings and Thrasher, 1999). A retrospective look at the 1992 election, however, found that in common with the main polling companies the model forecast was a narrow Labour victory instead of an eight-point Conservative lead (Rallings and Thrasher, 1999).

☆ International Symposium on *Methods and Models for Election Forecasting*: *The UK General Election of 2010*, organised by the Democracy, Citizens and Elections Research Network (DCERN) in association with Institute for Social Change and Politics and the Elections, Public Opinion and Parties (EPOP) Specialist Group of the UK Political Studies Association. University of Manchester, March 19–20th 2010.

* Corresponding author.
E-mail address: c.rallings@plymouth.ac.uk (C. Rallings).

0261-3794/$ – see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.electstud.2010.09.011

Prior to the 2001 general election we encountered problems caused by an outbreak of foot and mouth disease which led some local authorities to restrict local by-election activity. This meant that from the end of February to the synchronous local/general election in June there were just 51 cases, many fewer than normal and none of which were held in the month prior to the general election itself. Nevertheless, our forecast for The Sunday Times published on June 2 used our analysis of split-ticket voting at the synchronous 1997 elections to generate figures from the model data (Rallings and Thrasher, 1998, 2001, 2003). The forecast read Labour 41% (42.0% actual), Conservative 32% (32.7%) and Liberal Democrats 20% (18.8%). Interestingly, three of the five companies conducting national polls over-estimated Labour by 3–5 points in their eve of poll surveys, continuing the pattern from the early 1990s.

Following the 2001 election, where recorded turnout fell below 60%, voter apathy appeared to spread to the local parties. It was not that there were many fewer by-elections than before but that the pattern of party competition started to change. Where three-party contests had once been commonplace they now became less so as one or other of the main parties failed to present candidates. Furthermore, in some areas more candidates from minor parties began to participate and secure significant electoral support, thereby making such cases unsuitable for national forecasting purposes. In short, an increasing number of cases were being excluded from the by-election modelling because the pattern of party competition at both the by-election and its May predecessor were incompatible with the task of estimating national vote shares. On May 1, 2005 our Sunday Times forecast was again adjusted on the basis of split-ticket voting (now 2001 as well as 1997 aggregate level data were available). The forecast was Labour to win a majority of 96 seats having polled 37% (36.1%), the Conservatives 34% (33.2%), and Liberal Democrats 21% (22.6%). The eve of poll findings from the polling companies proved as accurate.

During the recent parliament, however, the need to address the problem of declining case selection has become imperative. In the following section we outline the initial by-election model before reporting on our efforts to address the problems affecting the admission/exclusion of cases. Next, we introduce the revised model and then assess its utility by applying forecasts retrospectively both to the May local electoral cycle from the 1990s onwards. Finally, we use recent evidence to forecast the likely outcome of the 2010 general election. Although we are committed to the value of these data for forecasting since these are, "real votes in real ballot boxes" and constitute in Austin Ranney's terms the 'hardest' data political scientists can get (Ranney, 1962) the model remains a work in progress. We are still testing whether the number and location of by-elections, to an extent affected by the wider electoral cycle, is a factor that influences forecasts and, if so, in what direction. We are also working with data that records the cause of the by-election vacancy since the circumstances (enforced resignation, retirement or death of incumbent for example) may affect the distribution of subsequent party support.

2. The original by-election model

We have been collecting local council by-election results from across Britain since the mid 1980s. Each year sees an average of 290 vacancies although this fluctuates with the broader electoral cycle. The forecast model requires information about both the by-election result and the outcome at the preceding May-election for each ward (the local electoral district). Clearly, assuming that the pattern of party competition is identical across elections it is straightforward to calculate change in vote share and swing for a given ward but forecasts, generalising from the particular, require a set of benchmark figures that are common across a range of wards. This comes in the guise of the 'national equivalent vote' (NEV) an estimate of how the country as a whole might have voted extrapolated from actual local election voting in any given year (Curtice and Payne, 1991; Rallings and Thrasher, 1993).

For any given ward election in May, therefore, we know both the distribution of party support in the ward and how that compares with the country as a whole. The original method used only by-elections which featured candidates from Conservative, Labour and Liberal Democrat parties at both the May election and the by-election. An additional caveat was that cases would be excluded from consideration where votes for other parties and Independents at either the May or by-election were greater than 10% of the total vote. The exception to this was cases where the intervention and support for other parties/independents was consistent across the two elections. Because of a non-uniform local electoral cycle it was important to note the particular year when a by-election ward had held its May election since this became an important part of the calculation. A worked example is provided in Table 1 while the method is described formally in the Appendix.

3. The revised by-election model

From 2001 the rather dramatic change to the pattern of party competition had implications for modelling. In 2000 some 68% of by-elections featured candidates from all three main parties; this dropped to 62% in 2001 with a further fall to 55% in 2002. Although the proportion of three-party contests recovered from this low point it did not match the

Table 1
Calculating the current national equivalent vote from the November, 1996 by-election result in Ixworth ward, St Edmundsbury Council.

Steps in the method	Conservative	Labour	Lib Dem
a) By-election share in ward	43.4	26.2	30.4
b) 1995 May vote in the ward	36.1	29.6	34.3
c) Change in vote share (a−b)	+7.3	−3.4	−3.9
d) 1995 National Equivalent Vote	25	47	23
e) Estimate of current NEV (d + c)	32.3	43.6	19.1

Of course, in a given ward it is possible that the change in a party's vote share may be greater than its NEV for a particular year, leading to a current NEV estimate that is nonsensical. However, by averaging the estimates across all by-elections over a month/three month period, any extreme results are smoothed out. In essence, therefore, between the May and the subsequent by-election the model is calculating change in each party's vote share, adding/subtracting that change to the NEV value for the relevant year and averaging across cases to estimate a current NEV for a given point in time.