GETTING OPINION POLLS 'RIGHT'

- *What went wrong with polls in 1992?*
- *Will pollsters do better in 1997?*

The 1992 General Election result was quite different from that suggested by opinion polls and raised questions over the reliability of polling. As we approach the next General Election, interest is growing in the likely performance of polls in 1997.

This note examines how opinion polls are conducted, and the implications of changes made since the last election.

THE 1992 POLLS

General Election opinion polls were first published in 1945, and have anticipated the outcome of the 14 general elections since then fairly accurately (generally within ~2% of the result). There have been three noticeable upsets: in 1951 when the lead between the top two parties according to the polls was wrong by over 5%; in 1970 when the lead was wrong by over 6%; and most recently, in 1992 when the polls' performance was the worst in UK polling history - underestimating the Conservative lead over Labour by nearly 9%.

Looking at 1992 in more detail, 50 opinion polls were carried out during the campaign by 6 polling companies on behalf of national newspapers and broadcasters, and over four-fifths of these showed Labour to be in the lead by between 0.5% and 7% (average ~2% - **Figure 1**). Two days before the election, the Labour lead was still ~1%, but on 9 April the Conservative share of the vote was nearly 8% more than Labour's. The polls thus misjudged the gap between the top two parties by close to 9% (an error of 4.5% on each party's vote share).

After this poor performance, the Market Research Society (MRS) convened a group of experts to investigate what had happened. The report in July 1994 concluded that there were problems with the way the polling methods had been carried out but equally, patterns of public behaviour were also involved. It recommended that several sources of error should be addressed immediately, research be undertaken to develop better methods, and more attention should be paid to the **inherent limitations** of the polls by the media, politicians and the public. These questions are addressed below.

LEARNING LESSONS

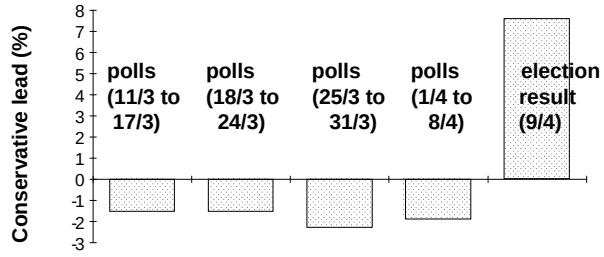
Poll Methods

Box 1 provides an outline of how polls are designed and executed, and explains the historical reliance on selecting quotas for interview which are intended to repre-

POST note **96** **March 1997**

POSTnotes are intended to give Members an overview of issues arising from science and technology. Members can obtain further details from the PARLIAMENTARY OFFICE OF SCIENCE AND TECHNOLOGY (extension 2840).

Figure 1 THE 1992 GENERAL ELECTION OPINION POLLS



Source: Market Research Society

Box 1 HOW POLLS ARE PLANNED AND CARRIED OUT

An opinion poll aims to provide a snapshot of what the population thinks about a particular issue at a particular time, and to do this pollsters select a **manageable sample** of people designed to represent the **whole population**. For practical and cost reasons, usually between 1,000 and 2,000 people are interviewed. Pollsters use data from national surveys of attitudes and behaviour to set 'quotas' of the types of people to be interviewed (main sources include the National Readership Survey (NRS), the General Household Survey, the Labour Force Survey, mid-year population estimates from the Office for National Statistics, and the latest Census). For instance, surveys show that ~52% of the population is female; ~38% is aged between 25 and 44; and ~32% is in the junior non-manual socio-economic group. The quotas are designed to ensure that the sample of people interviewed reflects these types of statistics. Polls are also structured to cover different areas (inner cities, suburbs, rural communities, etc.), and 'typical' safe and marginal seats.

The poll itself is carried out by interviews conducted face-to-face (either in the street or in people's homes) or by telephone. Questions, and the order in which they may be put, do influence the outcome to a limited extent and thus need consistent treatment in the poll design. The 'raw' results need statistical treatment before they are released. First, the actual sample interviewed may differ from the target quota (e.g. the interviewer may exceed the quota for 'working class males' and fail to meet that for 'retired females'). Results must be **weighted** to adjust for this and to 'fine-tune' the poll to take account of more subtle variables. For an **adjusted** poll, some of the 'don't knows' may be reassigned according to their past voting behaviour. A 'raw' result thus goes through several stages as illustrated below, and also needs to have a statistical error margin assigned to it (usually the **95% confidence interval**).

Working up the results

Party	No.	Weighted No.	Unadjusted share	Adjusted share	Error
Con	375	430	34%	41%	+/-3%
Lab	525	500	40%	48%	+/-3%
L.Dem.	100	70	6%	7%	+/-3%
Other	50	50	4%	5%	+/-3%
Don't know	200	200	16%	-	-

sent the wider public. One problem in 1992 was that the polled samples **did not adequately represent the national population**, because of unappreciated weaknesses in the data sources used by pollsters (Box 1) to set their quotas. For instance, some national surveys suggested that there were more in the lower social grades than was actually the case, as revealed when the 1991 Census became available after the election. As shown in **Table 1**, this caused poll samples to over-represent Labour supporters. Another problem arose because the characteristics used for setting the quotas and weighting the results (age, sex, or social grade) were not in fact, closely related to voting behaviour. Factors, such as housing tenure and car ownership, might have been better.

Now, pollsters use more up-to-date and representative data for quotas and weightings (e.g. the 1991 Census is supplemented by the quarterly Government Labour Force Survey, an improved version of the National Readership Survey, and the General Household Survey[1]). The main quota controls remain age, sex, social grade and working status, but car ownership and housing tenure are now taken into account because they are more closely related to voting behaviour, and so increase the effectiveness of the controls.

Even with these improvements, recent evidence raises questions over the effectiveness of the **quota approach** altogether, partly because of unintended biases introduced by the interviewer. The latter has free reign to decide whom to interview (provided they add up to the quota totals required), and will tend to look for the most 'efficient' way of meeting this quota. In rural/urban constituencies, for instance, effort may be concentrated in urban areas, underestimating the views of the rural voter. Interviewers may also use natural 'focal points' such as railway stations or town squares, or may be reluctant to walk up long driveways or reach the top of tower blocks. Each of these **selection biases** may skew interviews more or less towards different social groups. To counter this, and to more tightly control the types of respondents selected, MORI is conducting in-home interviews in smaller sampling areas, using Census enumeration districts or parts of local authority wards rather than whole constituencies. MORI expects this to improve accuracy and currently this approach increases apparent Conservative support by around 2%.

An alternative is to replace quota sampling by **random (or 'probability') samples** - where every adult on the electoral register has an equal chance of being selected. Individuals are selected by simple 'names out of a hat' methods or by random telephone polls based on interviewing, say, every 100th person in a telephone directory, or picking telephone numbers using random digit dialling (RDD) (when quotas are needed to select individuals from the households canvassed).

Table 1 DIFFERENCES BETWEEN DATA SOURCES

Factor		1991 NRS	1991 Census
Sex	male	48.1%	47.7%
	female	51.9%	52.3%
Age	18-24	13.4%	13.2%
	25-44	37.9%	38.1%
	45-64	28.0%	28.7%
	65+	20.6%	20.1%
Socio-Economic Group			
employer/manager		12.4%	14.7%
professional		3.0%	4.5%
junior non-manual		32.1%	32.6%
skilled manual		25.0%	21.2%
semi-skilled manual		16.1%	15.8%
unskilled manual		6.0%	5.5%
unclassified		4.7%	3.4%

Source: Market Research Society

Random samples have the advantage that they do not depend on the underlying data on social factors needed for quotas, and so they may be more accurate. Their disadvantage is that face-to-face interviews take longer because it takes time to track down the pre-selected individuals for interview. Most clients prefer speed (especially near a General Election where 'snapshots' are required more frequently), and virtually all election opinion polls for over 20 years have thus been based on quota samples. However, the spread of the telephone has led to random sampling becoming more practicable, and newer statistical methods also allow 'rules of substitution' to be applied if some of those initially selected prove difficult to contact. As a result, two polling companies (ICM and Gallup) have now abandoned quota sampling in favour of random polling[2]. They will use this method during the 1997 election campaign on the basis that the polls should be more accurate and, by using telephone RDD polling, can still be completed in an acceptable time.

The effect of a shift from quota to random polling is illustrated well by the results of two recent Gallup polls. The first, published in December 1996, was conducted by quota sampling and showed a Labour lead of 37%. The next poll, published in January 1997, was conducted using RDD telephone polling, and showed a Labour lead of 18%. This was not a halving of Labour's support, as polls by other companies had consistently shown a Labour lead of ~20%. Rather, the change indicated that the previous quota method had significant (if unintended) bias built in. The two other main polling companies (MORI and NOP) will continue to use quota sampling on the basis that the improvements they have made in other aspects of the methodology will overcome the difficulties experienced in 1992.

Voter Behaviour

In addition to the above shortcomings, the MRS report

1. There is some doubt over the future of the General Household Survey, as the Office for National Statistics will not conduct the survey in 1997.

2. These polls are not strictly random in the true statistical sense.

suggested there was a marked tendency in 1992 for Conservative supporters to be more reluctant to reveal their support than supporters of other parties. This **differential refusal** could be through refusing to be interviewed, refusing to reveal voting intentions, or answering 'don't know' as a way of avoiding the main question. As a result, polls reported significantly less support for the Conservatives than really existed.

Traditional market surveying methods offer a number of ways of compensating for differential refusal:

- **'Squeezing'**, where interviewers ask the question again in a slightly different way. E.g. instead of "*which party would you vote for?*" - "*which party would you be most inclined to support?*".
- Asking people to take part in a **secret ballot**. One test in September 1992 found that a ballot reduced refusals from 7% to 1%, with most of the extra accounted for by Conservative supporters. Other experiments showed no such differences, and that secret ballots offer little advantage. They will not be used during the 1997 election campaign.
- Adjusting polls by asking the 'don't knows' (DKs) how they voted at the last election, and (for those that can recall) reassigning their support to the party they voted for previously. These **adjusted polls** are regularly used in France and Belgium, and are becoming more common in the UK. Pollsters use slightly different methods of adjustment, but this can reduce the apparent Labour lead by around 2-6% (see **Figure 2**, where the average reductions in Labour's apparent lead is around 6%).

The 1992 polls also failed to detect the true size of the **late shift in support** towards the Conservatives; there was also a slightly higher turn-out among Conservative voters than for the other parties. Polls already run very close to the election[3], so there is little more that can be done directly. Rather, pollsters have to look for clues to the *possibility* of a late swing - e.g. how likely poll respondents are to vote, and how committed they are to their stated voting intentions. In 1997, polling companies will be more aware of these factors.

ISSUES

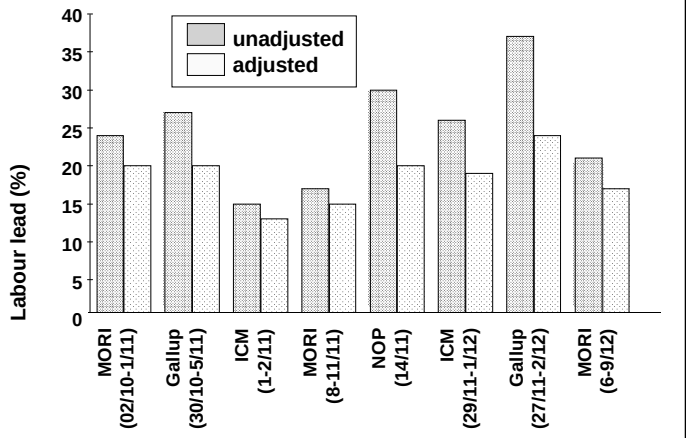
Poll Reliability

As discussed above, the 'inquest' into the 1992 polls' failure to anticipate the outcome of the election has led to the discovery that different approaches (e.g. adjusted polls, random polls) can reduce an apparent 'lead' substantially. The various changes made by the pollsters since 1992 include:

- improving quotas and weightings which will increase accuracy of the party lead by 3-5%;

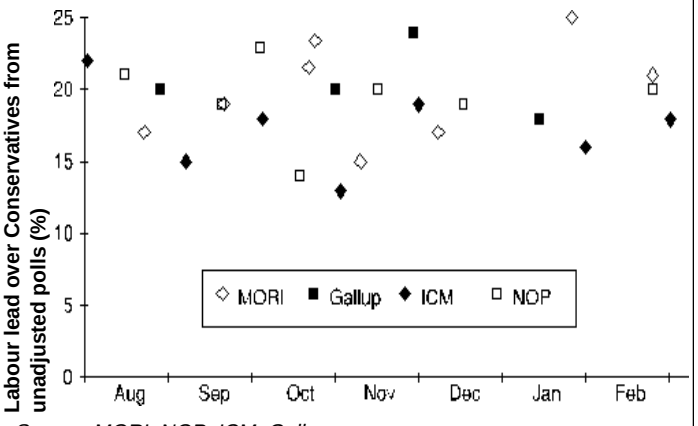
3. In contrast to France, Spain, Portugal, Belgium and Luxembourg where the publication of polls are banned in the last few days before an election (although the law is not always enforced (e.g. in Belgium)).

Figure 2 ADJUSTING THE POLLS BY PAST VOTING BEHAVIOUR



Source: MORI, Gallup, NOP, ICM

Figure 3 LABOUR'S LEAD (August 1996-February 1997)



Source: MORI, NOP, ICM, Gallup

- adjusting polls by past voting which has reduced Labour's apparent lead in recent polls by 2-6%;
- some companies shifting from quota to random polling. Here, the effects are not yet clear; the Gallup example above contrasts with ICM's use of random polling for the last 2 years, whose results are generally consistent with those based on quotas.

With the various changes now implemented, the next election provides the first major 'field-test' of newer polling methods. As shown in **Figure 3**, there has been reasonable consistency in the results of different poll methods used by different companies in recent months. **Polling companies are thus confident that the 1997 general election polls will perform much better than in 1992.** Most companies expect their polls to be within 3% of each party's share of the vote (and hopefully 1-2%) -very much less than the mismatch of 4.5% in 1992, and more in line with the general post-War record.

Ultimately, however, polls can only be reliable if respondents know their own minds, and here it is evident that the **British electorate has become more volatile** in recent years, with fewer saying that they will vote and with less commitment to stated voting intentions. Thus, a month before the 1992 election, MORI found that 69% of people interviewed said they were certain to vote, but by November 1996, this had fallen to 57%. Over the same period, the proportion "not likely or unsure"

Table 2 REALLOCATION IN A POLL

Party	Poll in reality	Poll as reported	Adjusted poll
Labour	47%	56%	51%
Conservative	22%	26%	31%
Lib Dem	10%	12.5%	12%
Other	4%	5.4%	6%
Don't know	17%	-	-
Total	100%	100%	100%
Labour lead	25%	30%	20%

Source: November Gallup 9000 poll (Daily Telegraph 6/12/96)

whether to vote rose from 23% to 34%, and those "certain not to vote" from 8% to 9% - giving up to 43% not likely to or unsure about voting in the coming election. There is also reduced commitment to stated voting intention. Thus, MORI found that while 12.5% of people changed their minds during the 1979 election campaign, this increased to 21% in 1992. This increases the inherent uncertainties of interpreting poll results.

Are polls presented properly?

Polls are generally commissioned for a customer who distils detailed poll results into a newspaper article or TV slot. This process has scope for **significantly distorting the results**. Headlines can claim findings that are not there (e.g. "*Poll Predicts Conservative/Labour Victory*"), when in reality **a poll never predicts** anything, but merely presents a snapshot of public opinion at the time the poll was carried out. The article may not include the questions actually asked in the poll, making it difficult to judge what the poll is actually telling the reader. Technical details may be omitted; e.g. the date of the poll, how many interviewed, and the error margin. Including the latter is especially important when parties are fairly even or when trying to see trends in support, and all polls should be presented **with associated uncertainties (e.g. 33 +/-2%), or as ranges (e.g. 31-35%).** It is also misleading to present polls with a **spurious accuracy**. To say that one party has 35.2% of the vote is only justified if error margins are 0.1% or less; most poll sample sizes are too small to provide such precision, and 1-2% is a more common error margin.

A key feature of published polls is how the 'don't knows' (DKs) are presented. **Table 2** shows an example of a recent poll. Column 1 shows the actual results from the 7,997 people interviewed; 17% were DKs, and the picture given can vary significantly depending on how they are treated. If left in, Labour's lead over Conservative is 25%. However, when published, the newspaper left out the DKs, and only showed the support among the 83% who **expressed a preference**; this gave a Labour 'lead' of 30% (column 2). The third approach is to take account of the problem of differential refusal discussed earlier, and adjust the poll by reassigning some of the DKs according to their past voting behaviour. If this is done, Labour's apparent lead drops to 20% (column 3).

Which of these three different ways is best is open to dispute. Some argue that the majority of DKs are unlikely to vote and thus it is reasonable to express the results in terms of those expressing a preference (i.e. col 2). Previous experience however leads many to see the adjusted polls (col 3) as more reliable since they have taken account of those DKs who voted at earlier elections and who may thus well vote again. What is important is that **the basis of the figures is clearly stated** when the results are presented. Adjusted polls clearly cannot be compared with unadjusted polls, and there may also be different methods used for adjusting polls by different companies.

Problems such as these lead pollsters to argue that **the publishers of opinion polls should do more to ensure that polls are not distorted in media reports**, and have *inter alia* produced a guide for journalists on the 'do's and don'ts' of poll reporting. To date, the response has been mixed, as illustrated by the not uncommon examples such as that in Table 2. Even where the primary clients follow pollster's ground rules (often allowing the pollsters final approval of text before publication), secondary reporting by other newspapers or broadcasters seldom retains the original safeguards. **Readers thus need to remain alert to the sources of distortion that can arise.**

Other sources of potential distortion come into play when people use the results of opinion polls to predict the outcome of an election, mostly by assuming that the swing of support from one party to another is uniform across the country. In reality, the swing varies widely across the country, and is often significantly different in marginal constituencies. Thus in the 1992 election, a uniform swing would have given an outcome of Conservative 356; Labour 253; Liberal Democrat 18, whereas in reality, seats were Conservative 336, 271 and 20 respectively (others 24).

One way to improve seat predictions from the polls might be to **concentrate more polling in marginal constituencies**, and some polls do take this into account and assume that the swing can vary. Such predictions do, however, have large margins of error and should only give a forecast of the likely **range** of seats for each party rather than precise numbers.

Overall, despite the measures introduced to make the 1997 polls more reliable, the optimism of pollsters about improved accuracy must be tempered with the uncertainties arising from the electorate being more volatile. This remains an unpredictable and uncontrollable source of uncertainty which pollsters cannot readily take into account. This ultimately may limit the ability of pundits to predict the outcome accurately.