Are interviews costing £0.08 a waste of money? Reviewing Google Surveys for Wisdom of the Crowd projects

G W Roughton and Iain MacKay

Survey data collection costs have become so low that there have been questions as to their quality. This is in addition to low response rates; professional respondents; respondent interaction and uncertainties about self-selecting ‘samples’. Some recent political polls have raised questions. The authors were curious as to whether a “Wisdom of the Crowd” approach might be an alternative. This paper looks at over 100 such surveys and reports that in five out of six cases discussed, the results from £0.08p interviews delivered results in line with known outcomes. Our results, explained in the paper, shows that such interviews are not a waste of money.

Internet surveys; response rate; political polls; wisdom of crowds

# Introduction

We are outlining our experience of carrying out a lot of surveys, 155 in fact, in which the data collection costs have been less than £0.08p per completed interviews – 80,590 interviews in total. We bought these from Google Consumer Surveys at $0.10 per complete, which is a bit less than 8p. We started a programme of surveys in October 2014 and which are continuing on various relevant subjects such as who will win the US Presidential election in November, who will succeed David Cameron, etc.

Of the 155 surveys, 128 were on subjects, such as an Election outcome, where there is a known outcome and where we can compare what our surveys tell us and then how often the survey gave a reasonable correct answer. At the same time as evaluating the survey process, we also consider an alternate “wisdom of the crowd” interviewing process. The first part of this paper describes aspects of the survey process including an analysis of the metadata. The second part evaluates some of the surveys and makes a judgement as to their effectiveness.

# Survey Metadata

The surveys we are considering are almost all of 500 completed interviews and have been conducted between the autumn of 2014 and now. Table 1 shows the number of surveys and the corresponding number of interviews:

|  |  |  |
| --- | --- | --- |
| **Row Labels** | **Surveys** | **Interviews** |
| UK | 125 | 65514 |
| US | 30 | 15076 |
| Grand Total | 155 | 80590 |

Table 1 Surveys and Interviews

Most of the surveys we have done are in the UK, but 30 are in the United States. The procedures are similar in both countries. Google also offer to do surveys in several other markets. Google review and approve the actual questions but otherwise it is almost entirely automated from start to finish. Once we tried them out, we were hooked. They publish a white paper on their methodology (McDonald, et al., n.d.) which covers many of the traditional questions one might ask any supplier of internet based surveys. We do not intend to dwell on the specifics of Google Surveys.

The subjects were mainly political but we explored some other subjects. Table 2 lists the political subjects because these are ones where there is a known outcome.

|  |  |  |
| --- | --- | --- |
| **Subject** | **Surveys** | **Outcome** |
| **UK** | 100 |   |
|  EU Referendum | 39 | Remain |
|  Labour leadership | 7 | Corbyn |
|  London Mayor | 8 | Khan |
|  UK General election 2015 | 46 | Cameron |
| **US** | 21 |   |
|  Democratic candidate | 13 | Clinton |
|  Republican candidate | 8 | Trump |
| Grand Total | 121 | 64606 |
|  |  |  |
| **% correct** | **83.3%** |

Table Surveys and Interviews

There are 6 subjects listed – four in the UK and two in the States. Most of the surveys gave results that agreed with the outcome of the Elections concerned but there were some exceptions of which our prediction of a Remain vote in the EU Referendum was the most notable. On the other hand our surveys correctly forecast, in May 2015, that David Cameron would be the next Prime Minister.

Some of the surveys were one-off, others were trackers. In the latter case, Google allows one to specify a given frequency – we have tried weekly and fortnightly so far. We have bought some on an ad-hoc basis to settle specific points.

Except for a couple of trials, these were all single question surveys. Google charges $50.00 for a 500 sample – or $0.10 per completed. Actually it is less than £0.08p per complete. Almost all the samples were 500 although in a few cases we used 1,000.

Google puts out any given survey as a survey wall to people reading premium content. Here is an example of what one of our surveys looks like to a respondent – just the one question:

****

Figure Specimen question presentation

Google’s system keeps track of the amount of time a respondent spends on a questionnaire. For a single question survey, the average is about 20 seconds. Google check every questionnaire and advocate brevity and simplicity. We have not done enough surveys to form a view of factors that influence the length of interview. Table 3 outlines the interview length by category of project.

|  |  |  |
| --- | --- | --- |
| **Subject** | **Surveys** | **Length in seconds** |
| **UK** | 100 | 19 |
|  UK General election 2015 | 46 | 19 |
|  Labour leadership | 7 | 18 |
|  London Mayor | 8 | 26 |
|  EU Referendum | 39 | 18 |
| **US** | 21 | 23 |
|  Democratic candidate | 13 | 22 |
|  Republican candidate | 8 | 24 |
| **Grand Total** | **121** | **20** |

Table Length of Interview

There is a more variation within survey.

As we all know not everyone presented with such a screen will answer it. Google report the number of times a questionnaire is presented, as above, to a prospective respondent. They call these “impressions”. The proportion of interviews to impressions they call “Response rate”. Table 4 shows this for those surveys for which we have impressions data.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Subject** | **Surveys** | **Impressions** | **Interviews** | **Response rate** |
| **UK** | 100 | 313943 | 54066 | 17% |
|  UK General election 2015 | 39 | 120959 | 20340 | 17% |
|  Labour leadership | 7 | 48454 | 3504 | 7% |
|  London Mayor | 8 | 43383 | 4129 | 10% |
|  EU Referendum | 46 | 101147 | 26093 | 26% |
| **US** | 21 | 57653 | 10540 | 18% |
|  Democratic candidate | 13 | 35343 | 6516 | 18% |
|  Republican candidate | 8 | 22310 | 4024 | 18% |
| **Grand Total** | **121** | **371596** | **64606** | **17%** |

Table Impressions, Interviews & response rate

The average is 17% for those surveys for which we currently have impressions data. We have not done enough surveys to understand the response rate variations in this table. Our suspicion is that the low figure for London Mayor (7%) maybe due to the fact that this was a National Sample. The highest (26%) response rate for the EU Referendum question might arise for the greater interest in this subject. The range is of course greater at the individual survey level. We would be interested to know how the figure of 17% compares with other organisation

Response rate may vary by the number of impressions from which interviews are obtained from different sources. Table 5 shows these sources. We hope to amplify this with further analysis.

|  |  |
| --- | --- |
| **Source** | **Interviews** |
| News | 33305 |
| Other | 21866 |
| Mobile App | 11869 |
| Reference | 6220 |
| Arts & Entertainment | 6075 |
| Mobile | 1078 |
| Shopping | 177 |
| **Total** | **80590** |

Table Source of Interviews

Google apply weights to the Internet population for the majority of respondents. They base this on inferred gender, age group and country. An example of their weights is given in Table 6 below:

|  |  |  |  |
| --- | --- | --- | --- |
| **Group** | **Achieved** | **Weight** | **Internet population** |
| Male | 55.6% | 0.90 | 50.1% |
| Female | 44.4% | 1.12 | 49.9% |
| 18-24 | 23.2% | 0.70 | 16.3% |
| 25-34 | 24.2% | 0.81 | 19.7% |
| 35-44 | 19.5% | 0.94 | 18.4% |
| 45-54 | 16.1% | 1.18 | 19.0% |
| 55-64 | 9.4% | 1.49 | 14.0% |
| 65+ | 7.6% | 1.66 | 12.6% |
| England | 78.7% | 1.07 | 84.4% |
| Scotland | 13.1% | 0.64 | 8.4% |
| Wales | 5.6% | 0.84 | 4.7% |
| Northern Ireland | 2.6% | 1.00 | 2.6% |

Table Example Google weighting

Their weighting is to the Internet population. Google do not offer other weightings as part of their standard service, but since individual observations are provided in an Excel Workbook, it is relatively easy to test your own weighting schemes. It is very easy to turn weighting on/off and immediately see the weighting effect. In general the effect of Google weighting is fairly small. Google provides an area code which maps to ONS Standard Regions in the UK for 99.5% of respondents. Table 7 shows sample distribution compared to National figures. The Google sample figures are from the aggregate of respondents in our Brexit Referendum surveys (described in more detail below).

|  |  |  |
| --- | --- | --- |
| **Standard regions** | **UK Population** | **Google sample** |
| **Number** | **%s** | **Total** | **Referendum** |
| East Midlands | 4,533,222 | 7.2% | 4.9% | 5.0% |
| East of England | 5,846,965 | 9.3% | 7.3% | 7.3% |
| London | 8,173,941 | 12.9% | 18.1% | 20.8% |
| North East | 2,596,886 | 4.1% | 4.1% | 3.9% |
| North West | 7,052,177 | 11.2% | 9.0% | 8.8% |
| Northern Ireland | 1,810,863 | 2.9% | 5.4% | 2.2% |
| Scotland | 5,295,000 | 8.4% | 13.9% | 16.9% |
| South East | 8,634,750 | 13.7% | 10.3% | 10.3% |
| South West | 5,288,935 | 8.4% | 7.0% | 7.5% |
| Wales | 3,063,456 | 4.8% | 6.3% | 3.6% |
| West Midlands | 5,601,847 | 8.9% | 7.0% | 7.2% |
| Yorkshire and the Humber | 5,283,733 | 8.4% | 6.6% | 6.5% |
| **Grand Total** | **63,181,775** | **100%** | **100%** | **100%** |
| **Base** | **64,939** | **10,761** |

Table UK Population &Google

It is fairly obvious that the areas which mainly voted to stay, particularly London and Scotland, are over-represented. Google’s built-in weighting procedure didn’t balance at a lower geographical level than “state” (UK home country) within gender, and so the imbalances within England went uncorrected.

Our final published result for the referendum is shown in Figure 2 EU referendum survey results:



Figure EU referendum survey results

When we applied weighting by standard region, remain decreased and leave increased. This would have shown a much closer result to the final outcome. When looking at answers where there are pronounced area differences, as in the case of the Referendum, one might find it worth considering the weighting strategy.

With a database of this size, there are many further tests that could be applied, we are considering further analysis of this.

We will now consider the actual survey results in greater detail.

# Why Wisdom-of-crowds?

In X-MR we began having concerns about conventional political research after the Scottish Independence referendum, where polls consistently suggested a much tighter outcome than the reality. Betting markets had made a better prediction of the Scottish referendum than conventional polling (Thomas, 2014). The idea that people are more reliable predictors of other’s behaviour than their own seemed plausible and worth exploring. “Wisdom of crowds” is an approach to prediction and estimation popularised by Surowiecki in his eponymous book (Surowiecki, 2005). Within market research in general the technique has been used successfully by the company BrainJuicer (Brainjuicer Group PLC, n.d.). In opinion research the technique was pioneered by ICM (Boon, n.d.) in 2010 and 2011, with some impressive results. Boon concludes:

*“(his paper also) shows that a Wisdom approach to regular vote intention tracking produces an interesting complement to classically conducted vote intention polls. Or, if one were to be bold, a competitor to them*”.

So let us be bold.

With the 2015 UK General Election in prospect we speculated that a "wisdom of the crowds" methodology might be an interesting tool to deploy. We began a series of surveys, initially fortnightly and later weekly, asking people who they thought would be the next Prime Minister, which became the first of a series of studies of major UK political events.

(Surowiecki, 2005) makes four stipulations for a crowd to be effectively wise (Table 8).

|  |  |  |
| --- | --- | --- |
| **Criterion** | **Explanation** | **How addressed by sample surveys** |
| Diversity of opinion | Each person should have private information even if it's just an eccentric interpretation of the known facts. | We draw a random sample; people are not selected based on any relationship to the issue in question. |
| Independence | People's opinions aren't determined by the opinions of those around them. | This is more questionable, especially if we are asking people what they believe is the consensus opinion of others. |
| Decentralization | People are able to specialize and draw on local knowledge. | We (attempt to) draw a nationally representative sample |
| Aggregation | Some mechanism exists for turning private judgments into a collective decision. | We arrive at the collective decision by simple aggregation of responses. |

Table Wisdom of crowd prescriptions

By using a sample survey as a “crowd” we should realise the diversity, independence and decentralisation criteria. Tabulation of survey responses provides aggregation.

In our surveys we have asked questions of the form “Who/which {*alternative*} do you think most likely to {*winning outcome*}, **whatever your personal preference**”, and our results show percentages for each alternative. The outcome is usually an election result, and the alternatives the candidates.

It is not obvious how to interpret the percentages achieved by each alternative.

The alternative considered most likely by most people is perhaps the most likely outcome; and by extension the next highest ranked alternative the next most likely and so on.

Extending this reasoning, the percentages are perhaps a measure of the probability in some sense.

If the percentages are a measure of the probability, then how does this probability relate to share-of-vote?

 In this paper we are not proposing any explanatory model, but only presenting some data that hopefully will inspire further investigation and analysis.

Our tables compare rankings and share of vote in the outcome to our own rankings and percentages, which are often very similar, but we do not propose or believe that there is any simple relationship.

(Surowiecki, 2005) asserts that a properly composed crowd can be wiser than any of its members, who need not be expert in the domain in question. We shall see how that worked out for us.

# Why Google Surveys

We chose Google Consumer Surveys (GCS) as the vehicle for this research because:

* GCS encourage, indeed mandate, short surveys of simple questions.
* By their pricing GCS encourage single question surveys.
* GCS provide the interviewing engine and a dynamically balanced sample of respondents, including post-survey weighting to key demographics (demographic variables are inferred as part of the regular Google relationship with their users)
* Last, and far from least, the surveys we contemplated were inexpensive and so we could undertake them as a pure research exercise with no external sponsorship.

# Our political tracking studies

In this paper we focus on our UK political research. We have also surveyed, and continue to survey, aspects of the US Presidential Election which are not described here.

In our results quoted below, values are weighted and based on an unweighted sample of 500 unless otherwise stated. The weighted sample size is typically rather smaller (75-80%) because not all respondents have imputed demographics as a basis for the weighting. The GCS weighting procedure uses different combinations of demographics depending on the sample sizes available. Weights are typically close to unity because GCS balance sampling dynamically as the survey progresses.

## The UK General Election 2015

In our first survey, in early October 2014, we asked people what sort of government they expected to emerge from the Election:



Figure Experimental General Election survey

The conspicuous problem here was that we had no idea what our respondents thought the most probable outcome (“Other”) might actually be. In retrospect we should perhaps have explored “Other” as an open-ended, however our decision was to rephrase the question.

From October 10th onwards we asked:

“*Who do you think most likely to be UK Prime Minister after the May 2015 election (whatever your personal preference)?*” with alternatives:

* *David Cameron*
* *Ed Miliband*
* *Nick Clegg*
* *Nigel Farage*
* *Other (specify)*

The alternatives apart from the final “Other” were randomised in order of presentation. Figure 4 shows the results for the 8 months leading up to the Election. The consistent result of these surveys was that David Cameron was always considered by far the most likely Prime Minister. This was quite contrary to the accepted wisdom of the polls during this period; readers who were in the UK during this period will recall the debates about hung Parliaments, the SNP holding the balance of power etcetera, with very little prospect of the Conservatives being able to build a government.



Figure Answers to the Prime Minister question

This is based on 30 surveys grouped by month. The consistent result was that David Cameron was always considered by far the most likely Prime Minister. This was quite contrary to the accepted wisdom of the polls during this period; readers who were in the UK during this period will recall the debates about hung Parliaments, the SNP holding the balance of power etcetera, with very little prospect of the Conservatives being able to build a government.

Figure 5, from (Wikipedia, 2016), summarises the conventional polls conducted in the years leading up to the election:



Figure Summary of opinion polls

However, as we all know, Cameron returned as Prime Minister on the back of an overall Conservative majority. No great surprise to X-MR, but a big surprise to political researchers in general.

The actual election results were as shown in Table 9.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Who will be next PM?** | **Party** | **X-MR final %** | **Share of vote** | **Share of seats** |
| David Cameron | Conservative | 48% | 37% | 51% |
| Ed Miliband | Labour | 35% | 30% | 36% |
| Nick Clegg | Liberal Democrat | 5% | 8% | 1% |
| Nigel Farage | UKIP | 10% | 13% | 0% |
| Other | Other Parliamentary | 3% | 9% | 12% |

Table Summary of general election results

### What might the numbers mean?

As discussed above, our studies do not purport to estimate share of vote, only to rank the outcomes in order of perceived likelihood. Further inferences from our results to e.g. winning odds or share-of-vote would require some modelling process.

For Prime Minister of course, the share that counts is the share of Parliamentary seats and not the share of votes – not just for his or her own party but for potential alliances with other parties. We were asking our respondents for their judgement of the likelihood of this complex outcome so we should not expect to make any quantitative interpretation of the results.

Nevertheless, our respondents came very close to estimating the actual share of seats for the two major parties, the main determiner of the answer to the question who will be the next Prime Minister. Figures for the minor parties do not compare so well. Whatever relationship our results have to quantified outcomes beyond rank ordering is not simple.

A few weeks before the election we published some more details of our attempts to investigate the reason for our contrarian results (MacKay, 2015).

In the closing weeks we asked respondents how they believed voters for the different parties would behave in the forthcoming election, and applied these vote-switching expectations on a constituency-by-constituency basis. These had some interesting results (X-MR, 2015) which we published on the eve of the election. Inter alia the imminent destruction of the Liberal Democrats in Parliament was obvious.

### Comparison with other survey platforms

While we were tracking the Prime Minister question we made two experiments to explore whether our results were related to the Google Consumer Survey methodology rather than the reality of public sentiment.

#### **Fly Research**

Fly Research conducted an online survey on our behalf using their panel and the same question in the period 13th-15th February 2015.

#### **Crowdflower**

Crowdflower provide a crowdsourcing platform to delegate simple online tasks to their team of workers around the world. This is similar to the long-established Amazon Turk service. Normally a small number of workers will be given a large number of tasks, e.g. transcripts for sentiment analysis, phrases to translate etcetera. Instead we gave the Crowdflower platform instructions to ask workers in the UK to answer our survey (using our own online survey engine), for which we paid them a small amount, stipulating that each worker should answer the survey only once. In order to have five hundred different workers undertake the task we had to wait three weeks.

|  |  |  |  |
| --- | --- | --- | --- |
| **Potential PM** | **Fly Research13-15th February** | **Crowdflower1-21st January** | **X-MR Google survey30-31st January** |
| Base (unweighted) | 1002 | 509 | 501 |
| David Cameron | 55% | 53% | 53% |
| Ed Miliband | 30% | 32% | 27% |
| Nigel Farage | 4% | 5% | 10% |
| Nick Clegg | 2% | 7% | 7% |
| Other | 10% | 4% | 4% |

Table Comparing GCS with others

Note from Table 10 that our Google respondents rated the prospects of the UKIP leader, Nigel Farage much higher than we found in the other sources, but the order of the candidates was the same. Our contrarian expectation that Cameron was the most likely Prime Minister was reinforced by these studies.

### Weighting

The tenets of Wisdom-of-Crowds mandate diversity, rather than balance. We balance samples in normal research on the basis of characteristics that affect people’s decision in respect of the matter in hand. In Wisdom-of-Crowds it is the quality of people’s judgement of others that matter. We have found that conventional demographics don’t strongly affect people’s judgement of outcomes. For instance, we have compared age and gender across all the Prime Minister surveys in Figure 6 and Figure 7.

Figure Age breakdown of expected PM

Figure Gender breakdown of expected PM

We find that age is the demographic that most influences these political judgements, but in our surveys the demographic differences are not usually significant, and where significant are not substantial.

This consistency in judgement means that there is very little difference between our weighted and unweighted results.

## The Labour Leadership election

We ran weekly polls asking “Who do you think is most likely to be elected Leader of the Labour Party in the September 2015 election (whatever your personal preference)?” in each week from the announcement of the nominated candidates until the week of the election on September 10th.

This was a particularly interesting study because the universe (Labour party members, affiliated members and registered supporters) was very hard to access by conventional means. There were just under 600 thousand people eligible to vote - only 1% of UK adults, and very expensive to find by sampling the adult population. This universe was also a moving target as thousands of registered supporters were joining each day as the campaign progressed. The benefit of wisdom-of-crowds here is that we didn’t need to try to reach these people – just our usual national sample.

The results we obtained from our eight weeks of polling are shown in Figure 8, and the final results in Table 11:

Figure Answers to Labour Leadership question

|  |  |  |  |
| --- | --- | --- | --- |
| **Candidate** | **X-MR final poll** | **YouGov poll** | **Result** |
| **04-Sep** | **[[1]](#endnote-1)6-10 Aug** | **12-Sep** |
| Jeremy Corbyn | 52% | 53% | 59% |
| Andy Burnham | 20% | 21% | 19% |
| Yvette Cooper | 16% | 18% | 17% |
| Liz Kendall | 12% | 8% | 4.50% |

Table Labour Leadership Election

So far as we can determine, only the YouGov poll mentioned sampled the actual electorate.

Numerous other polls were conducted in the election period asking the general public or Labour voters questions like “If you had a vote in the Labour Leadership election, who would you vote for”, and unsurprisingly obtained figures very different from the actual outcome.

Our respondents were a sample of the UK Internet Population, with nothing asked or known about their affiliation to the Labour Party or their interest in the Labour Party Leadership campaign. We offer these results as support for the Wisdom of Crowds “non-expert” tenet.

Notes:

* The YouGov poll in early August was a very good predictor of the outcome a month later
* Public perception of the outcome as measured by our polls lagged the sentiment of the electorate as measured by YouGov by about a week, and remained very stable thereafter.
* Only around 5 of the 500 respondents in each of our surveys might be expected to have a vote in the actual election.
* The total cost of this series of surveys was $400.

## The London Mayoralty election

We used this as a further opportunity to test the tenet that a crowd does not need to have specialist knowledge. From the time nominations closed until the election we asked a sample of all UK adults “Who do you think most likely to be elected Mayor of London in May 2016 (whatever your personal preference)?”. **Our sample universe was adults in England** – only 19% of whom were (imputed) London residents and therefore eligible to vote.

The results we obtained are shown in Figure 9 and Figure 10 and the outcome in Table 12:

Figure – Adults in England (weighted, sample size average 392 per survey)

Figure Adults in London (weighted) – small sample size, averaging 73 per survey

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **First round result** | **Final X-MR poll(Londoners)** | **Final X-MR poll(all England)** | [**Final conventionalPoll 26/4 -1/5[i]**](#_edn1) | **Final poll, adjusted** |
| Caroline Pidgeon (Lib Dem) | 5% | 9% | 6% | 3% | 4% |
| Peter Whittle (UKIP) | 4% | 11% | 11% | 4% | 5% |
| **Sadiq Khan (Labour)** | **44%** | **38%** | **39%** | **35%** | **48%** |
| Siân Berry (Green) | 6% | 15% | 10% | 4% | 5% |
| Zac Goldsmith (Conservative) | 35% | 23% | 32% | 26% | 35% |
| Other | 8% | Candidate | 4% | 2% | 3% |

Table Results of the 2016 London Mayoral Election

Notes:

* Second column adjusts for DK and “will not vote” to obtain expected share of vote
* Our respondents substantially exaggerated the significance of the smaller parties and UKIP in particular.
* Our all-England sample out-performed the London sample – however, the London sample was very small. This was a weakness of our budget rather than the Google system – targeting samples below UK home country level is considerably more expensive.
* With the exception of UKIP, the candidates were listed in the correct order by both our national and London respondents.
* At the bottom of the order, the Green Party candidate was correctly identified as ahead of the Liberal Democrat candidate by both our local and national respondents.
* It seems that Zac Goldsmith was overtaken in favour by Sadiq Khan only in the closing weeks of the campaign.
* The total cost of this series of surveys was $400

## The Brexit referendum

We began surveying this topic on May 8th 2015, the day after the General Election with an incoming government committed to holding a referendum on EU membership, repeating fortnightly to track changes in UK sentiment on this issue.

We used the question:

“*In a "UK in/out of the EU" referendum, what do you think most voters in the UK will choose (whatever your own preference)?”*

With alternative answers:

* *UK remains in the EU*
* *UK leaves the EU*
* *There won't be a referendum* (fixed position)

The percentage believing there would be not be a referendum was high (16%), so we felt obliged to retain the answer until scepticism had dropped below 10%.

This did not occur for almost a year, until April 2016 in fact, several months even after the passage of the referendum Act and only as campaigning began. This is a perhaps a measure of public cynicism over political promises. Results from this first wording are shown in Figure 11.

Since clearly the UK would remain in the EU in the absence of a referendum, we see belief in Brexit remaining consistently below 40% apart from one tiny excursion in September, though showing a gentle upward trend.

Figure EU referendum sentiment (phase 1). Average weighted sample size 425

 At the end of February, with the referendum date and wording confirmed, we began a parallel study using the wording of the referendum itself. We settled on using the ballot paper order after we ran two studies on the same days and found insignificant differences between the results of fixed ballot paper order and randomising the two alternatives.

Our new question was:

## “*In the referendum on Thursday 23rd June, which answer do you think most voters in the UK will choose (whatever your own preference)?”*

With alternatives:

* “*Remain a member of the European Union*”
* “*Leave the European Union”*

Figure EU Referendum phase II. Average weighted sample size 394.

As can be seen in Figure 12 our respondents continued to be consistent in their expectation of a “Remain” victory, very steadily so until campaigning began in mid-April. In the campaign period there were fluctuations until the week of the referendum itself when our results returned to where they were before the campaign started – at no time getting even close to a Brexit majority. On referendum eve we (and the betting markets) therefore expected a Bremain outcome with the share of votes being quite clear-cut rather than “too close to call” according to all the conventional polls (Wikipedia, 2016).

The outcome of course, was a “Leave” victory, quantified in Table 13. We compared Google administration of our question with a conventional online survey. (Usurv, 2016) asked our question to sample of 500 people in the last two weeks of the campaign and reported very similar (weighted) results:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **Outcome** | **X-MR results on 20-June** | **USurv on 21-June** | **Final poll-of-polls** |
| Leave | 52% | 39% | 41% | 48% |
| Remain | 48% | 61% | 59% | 52% |

Table EU referendum result

Had we further weighted by Standard Region, the leave proportion would be more than 5% higher.

# Conclusion

We believe our experience demonstrates that Wisdom-of-the-Crowd can provide inexpensive but qualitatively accurate predictions of the outcomes of major political events.

The method is not infallible, as we discovered in the EU referendum. The crowd seems to be not always wise, though our experience shows it usually is, and we don’t have a proposal presently to identify such ‘unwise’ situations.

The particular benefits are:

* The results are not very sensitive to the composition of the sample, so it is not necessary to use large or rigorously selected and balanced samples.
* Wisdom of Crowds can be used to investigate an outcome where the decision makers are not easily accessible at all, let alone for probability sampling – for instance those eligible to vote in the Labour Leadership election
* Wisdom of Crowds can be used to investigate complex outcomes – such as the Prime Minister emerging from a Parliamentary election, which can only be derived from opinion polls by using elaborate modelling. Such outcomes are the indirect consequence of the individual behaviour being investigated in a poll. The crowd however can be asked the question directly.

These benefits are well realised by a tool like Google Consumer Surveys that delivers short simple surveys to a mass audience at a very low cost.

In conclusion, we believe our experience confirms that

**£0.08p INTERVIEWS NEED NOT BE A WASTE OF MONEY**

Boon, M., n.d. Predicting Elections, a Wisdom of Crowds approach. *International Journal of Market Research,* 54(4).

Brainjuicer Group PLC, n.d. [Online]
Available at: http://www.brainjuicer.com
[Accessed 14 July 2016].

MacKay, I. M., 2015. *UK Election 2015 – wisdom of crowds?.* [Online]
Available at: http://www.x-mr.com/assets/UK\_Election\_2015\_2.pdf
[Accessed 15 July 2016].

McDonald, P., Mohebbi, M. & Slatkin, B., n.d. *Comparing Google Consumer Surveys to Existing Probability and Non-Probability Based Internet Surveys.* [Online]
Available at: https://www.google.com/insights/consumersurveys/static/consumer\_surveys\_whitepape
[Accessed 14 July 2016].

Surowiecki, J., 2005. *The Wisdom of Crowds: Why the Many Are Smarter Than the Few and How Collective Wisdom Shapes Business, Economies, Societies and Nation.* s.l.:Anchor.

Thomas, D., 2014. *Scottish Independence Betting: Betfair Exchange proved correct again!.* [Online]
Available at: https://betting.betfair.com/politics/scottish-independence-referendum-betting/scottish-independence-betting-betfair-exchange-proved-correct-as-scotland-votes-no-190914-51.html
[Accessed 14 July 2016].

Usurv, 2016. *Survey results.* [Online]
Available at: https://www.usurv.com/C07ScAABX
[Accessed 15 July 2016].

Wikipedia, 2016. *Opinion polling for the 2015 General Election.* [Online]
Available at: https://en.wikipedia.org/wiki/Opinion\_polling\_for\_the\_2015\_United\_Kingdom\_general\_election
[Accessed 15 July 2016].

Wikipedia, 2016. *Opinion polling for the United Kingdom European Union membership referendum.* [Online]
Available at: https://en.wikipedia.org/wiki/Opinion\_polling\_for\_the\_United\_Kingdom\_European\_Union\_membership\_referendum
[Accessed 15 July 2016].

X-MR, 2015. *Pre-election forecasts.* [Online]
Available at: http://www.x-mr.com/blog/20150506\_ukelection.html
[Accessed 14 July 2016].

1. s

**G W Roughton, MA(Cantab),** Fellow of the Market Research Society and member since 1955, Fellow of the Royal Statistical Society, Trustees of AMSR and founder of MAS Survey Research (now absorbed into TNS) who were the first MR company in Europe to have its own in-house computer in 1966 (an IBM 1130). Subsequently a Director and Chairman Pulse Train Ltd until its acquisition by Confirmit. CEO of X-MR Ltd.

**I M Mackay. lain MacKay. MA (Cantab.) Computer Sciences**. Co-founder of Pulse Train Limited in 1976 and served as Director until 1998. Lead developer of survey systems STAR and Bellview. Technical Director X- MR Ltd

1. [↑](#endnote-ref-1)