

British Election Study: Technical Document

The Sample

YouGov uses its own system, PanMan, for sampling. This system is a propriety system designed by YouGov to ensure respondents are always able to take a survey. YouGov uses targeted quota sampling as opposed to random probability sampling. Using advanced analytical techniques and taking into account several key factors, YouGov's samples are most frequently assigned to achieve representative samples at the end of fieldwork. The software looks at all surveys that currently need panel members, and calculates how many people to send invites to every 30 minutes. Whilst a panel member is invited because of a certain surveys requirements, it doesn't mean they will definitely be sent to that survey. For example, if that particular job had closed, they will be taken instead to another survey that requires someone of those demographics. This helps to ensure our panel members are kept highly engaged in our research, and means they have no need to try and screen themselves into any one survey. Sampling frames are drawn according to the population being researched, and will generally contain the same target quotas as we wish to wait to. These can be applied either individually or interlaced as a job requires.

This system also removes the problem of fast responder bias as respondents are sent a link to the system, and not to individual surveys, and therefore an invite can be in a respondents email inbox for several days before they enter the system. Due to the way jobs are sampled we do not have a per survey response rate, however, the average response time for a clicked email is 19 hours from the point of sending.

The YouGov panel is recruited from a wide range of sources in order to ensure we accurately reflect the views of the nation. The website is free for anyone to sign up to at any time, and this is typically the way people who have seen our polls published come to the site, however we also run targeted advertising campaigns to help attract people who wouldn't have otherwise sought us out. Once someone has joined the site they are also given the opportunity to refer friends and family members to the site.

The combination of all these methods means we have a large, sustainable panel, from which we can both draw large representative samples and target smaller niche groups as the research needs require.

YouGov also invited all respondents still registered on the YouGov panel who had taken previous British Election Study surveys to this project.

Data Collection

Data is collected via the latest version of our propriety scripting system Gryphon, and this is written directly back to a database written in C++ (MongoDB). Survey data can then be accessed via a number of tools including SPSS, Dimensions, and Excel by both researchers and the data processing team. Respondent data is stored against a unique identifier assigned to each panel member, which allows us to match back to demographic information each time, without storing the data sets together in order to ensure we observe data protection policies as required.

Response and Weighting

YouGov use a rim weighting system for all work. i.e. we use an iterative process in order to ensure the data we publish is in the correct proportions for each of the major demographics (or rims). Typically these rims are Age/Gender, Region, Social Grade, with targets drawn from sources such as the ONS and NRS. We can expand these to weight on any other variables that may affect the data, assuming we can find reliable information to base this upon. Target percentages are derived from three sources:

1. Census data
2. National Readership Survey (a random probability survey comprising 34,000 random face-to-face interviews conducted annually)
3. YouGov internal analysis conducted on more than 60,000 responses to YouGov surveys at, and shortly after, the May 2010 general election, when respondents were asked both i) whether they generally thought of themselves as Labour, Conservative, Liberal Democrat etc; ii) which party they would support, or had supported, in the 2010 general election; and iii) election results

at the time. From this analysis, along with 2010 election results, YouGov derives its party identification target percentages.

For the British Election Study, rather than sample and weight the data to the profile of all British adults, four separate regional surveys were fielded; Scotland, Wales, London and England excluding London. Each of these four surveys were separately sampled and weighted to the profiles included as an appendix to this document.

Each of the four samples were weighted three times, with a process of data cleaning undertaken between each of these. The first weighting was with the entire sample and produces the highest weight factors. Following this, a random selection of the overrepresented groups in the sample were removed and the data were re-weighted. A second process of removing a random selection of the over represented groups was again undertaken before a third and final weight was applied. The below table highlights the total sample sizes for each weighting scheme, the highest weight factor, the weighing efficiency and the effective sample size for each of the four geographic regions as well as the total sample.

Weighting 1:

Sample	Sample size	Highest weight	Weighting efficiency	Effective sample size
Wales	3,477	14.96	48%	1,667
Scotland	5,864	9.16	61%	3,556
London	3,683	20.12	42%	1,546
Rest of England	17,216	11.00	64%	10,941
TOTAL SAMPLE	30,239	20.12	57%	17,322

Weighting 2:

Sample	Sample size	Highest weight	Weighting efficiency	Effective sample size
Wales	2,255	6.65	77%	1,707
Scotland	4,137	4.12	84%	3,488
London	1,915	4.47	83%	1,590
Rest of England	12,574	5.16	85%	10,739
TOTAL SAMPLE	20,881	6.65	84%	17,500

6.65

Weighting 3:

Sample	Sample size	Highest weight	Weighting efficiency	Effective sample size
Wales	1,413	1.59	97%	1,373
Scotland	3,430	2.44	94%	3,236
London	1,237	1.10	98%	1,209
Rest of England	9,790	2.03	98%	9,578
TOTAL SAMPLE	15,870	2.44	97%	15,390

As part of the project, YouGov oversampled in London, Scotland and Wales to ensure robust samples for analysis in these regions. Therefore, an additional weight has been computed to correct for this at the overall level.

APPENDIX – WEIGHTS

Age & Gender	Scotland	Wales	London	England
Male 18-24	6	6.2	6	6.0
Male 25-39	11.9	11.2	18.3	11.8
Male 40-59	17.3	16.6	16.1	16.9
Male 60+	12.4	14.5	8.9	13.7
Female 18-24	5.8	6	6	5.9
Female 25-39	12.3	11.2	17.5	12.1
Female 40-59	18.4	17.2	16	17.4
Female 60+	15.9	17.1	11.2	16.1

Social Grade	Scotland	Wales	London	England
AB	19.5	18	26	28.4
C1	27.5	29	33	29.5
C2	20	20	11	22.7
DE	33	33	30	19.4

Newspaper	Scotland	Wales	London	England
Express / Mail	12	14	17	14.9
Sun / Star	16	16	20	22
Mirror/Record	20	21	14	15.4
Guardian / Indy	6	2	5	2.8
FT / Times / Teleg.	6	5	10	8.5
Other	20	21	9	11.2
No paper	20	21	25	25.2

Region	Scotland	Wales	London	England
North	N/A	N/A	N/A	33.5
Midlands	N/A	N/A	N/A	22.5
East	N/A	N/A	N/A	13
South	N/A	N/A	N/A	31

Welsh Region	Scotland	Wales	London	England
Mid and West Wales	N/A	19.1	N/A	N/A
North Wales	N/A	20.7	N/A	N/A
Cardiff	N/A	11.6	N/A	N/A
South Wales Central	N/A	10.7	N/A	N/A
South Wales East	N/A	20.7	N/A	N/A
South Wales West	N/A	17.2	N/A	N/A

Party id 2010	Scotland	Wales	London	England
Labour	N/A	N/A	N/A	30.7
Con	N/A	N/A	N/A	31
Lib Dem	N/A	N/A	N/A	12.5
Oth	N/A	N/A	N/A	1.8
None / DK	N/A	N/A	N/A	24

Holyrood Past Vote	Scot	Wales	London	England
Con	11.1	N/A	N/A	N/A
Lab	25.4	N/A	N/A	N/A
LD	6.3	N/A	N/A	N/A
SNP	26.2	N/A	N/A	N/A
SNP (Holyrood) & Lab (wmster)	10.1	N/A	N/A	N/A
Oth	0.9	N/A	N/A	N/A
DNV	20	N/A	N/A	N/A

Party id 2010	Scotland	Wales	London	England
Lab	N/A	37	N/A	N/A
Con	N/A	20	N/A	N/A
Lib Dem	N/A	8	N/A	N/A
SNP/PC	N/A	7.5	N/A	N/A
Oth	N/A	2.5	N/A	N/A
None / DK	N/A	25	N/A	N/A

Party id 2010 London	Scotland	Wales	London	England
Labour loyal	N/A	N/A	26	N/A
Labour disloyal	N/A	N/A	12	N/A
Con	N/A	N/A	28	N/A
Lib Dem	N/A	N/A	11	N/A
Oth	N/A	N/A	3	N/A
None / DK	N/A	N/A	20	N/A

Ethnicity	Scotland	Wales	London	England
White	N/A	N/A	70	N/A
Indian	N/A	N/A	7	N/A
Pakistani/Bangladeshi	N/A	N/A	5	N/A
Black	N/A	N/A	11	N/A
Other and Mixed	N/A	N/A	5.5	N/A
Refused	N/A	N/A	1.5	N/A