

Supplementary Appendix

This memo contains supplementary analyses for the paper “Conceived in Harlesden: Candidate-centred campaigning in the British general elections.”

Data gathering and coding

Gathering and parsing the contents of the Electionleaflets.org database was done automatically via a script written in the Python programming language. The script automatically downloads scanned leaflet images from the Electionleaflets repository, and scrapes and stores meta-information such as constituency, date of receipt, and any keywords or content flags stored along with the leaflet images. Within this dataset, we subset along content and party lines. For example, some leaflets in the repository were issued regarding local elections. As we are only interested in political communication during the general election, these leaflets were discarded. We also discarded leaflets that were incorrectly or incompletely scanned.

The Python scraper downloaded the images of the leaflets, so we were also able to manually code additional information about each leaflet’s contents. We evaluate each leaflet on twenty-one policy dimensions: the economy, civil rights, immigration, health, agriculture, labour, education, environment, energy, transportation, law/crime/family issues, social welfare, community, banking, defence, international affairs, science, trade, government, territorial issues, and Europe. With exception of Europe and immigration, all issues dimensions are coded using the topics and subtopics used by the UK Policy Agendas Project. In addition to policy issues, we also identified whether the leaflet focuses on the traits of an individual candidate and if the leaflet includes a photograph of the candidate or leader; if the leaflet and the extent to which the publishing party directly criticises other parties or candidates.

Comparison between leaflet availability and BES leaflet contact

We compare our data to contact rates calculated using data from post-election waves of the 2014-2017 British Election Study (BES) Internet Panel. To do so, we calculate the percentage of BES respondents in the constituency who reported they received a leaflet from a given party in the previous 4 weeks. BES respondents were not asked whether the leaflets they received identified a candidate by name. To ensure that we are comparing like with like, we use the entire sample of leaflets received from a given party within a constituency, regardless of whether the candidate is mentioned by name.

When we compare these figures to the total number of leaflets we have for the party in each constituency, we find a positive and statistically significant correlation for all parties in both elections (Table S1). In other words, we have more leaflets from seats where more BES respondents reported receiving a leaflet from the given party.

Table S1. Correlation between no. of leaflets and BES leaflets contact in seat (%)

Party	2015	2017
Conservative	0.18	0.17
Labour	0.28	0.22
Liberal Democrats	0.47	0.40
SNP	0.56	--
UKIP	0.21	0.24
Green	0.46	0.44

Notes. All correlation coefficients are statistically significant at $p < 0.01$.

Evaluating the potential bias associated with self-selected data

Our dataset covers a wide range of seats in England, Scotland, and Wales, but we acknowledge that is a sample of convenience, and therefore, there is the potential for bias in the sample. To investigate the differences between the population of constituencies and the sample of constituencies used in our analysis, we conduct a series of t-tests to identify systematic differences. Table S1 compares the mean values for the constituencies that are included in our dataset versus those constituencies where we have no candidate leaflets on a number of dimensions, as well as the p-value associated with the difference between the two groups. A p-value below 0.05 suggests that constituencies for which we have data differ meaningfully from the constituencies that are absent from the dataset on that trait.

Table S2. Constituency Sample Representativeness, 2015-2017

Variables	Not included	Included	p-value	Not included	Included	p-value
<i>Socio-demographic characteristics</i>						
Population density (number per square km)	15.6	22.5	0.00	17.1	25.7	0.00
Professional occupations (% of population)	28.9	31.7	0.00	29.3	33.2	0.00
Routine/manual occupations (% of population)	35.5	31.7	0.00	34.6	30.2	0.00
Level 4 qualifications or above (% of population)	24.2	28.1	0.00	24.7	30.3	0.00
No qualifications (% of population)	25.1	22.4	0.00	24.6	21.0	0.00
<i>Region</i>						
England	80.8	86.2	0.08	82.8	88.7	0.02
Scotland	8.5	9.8	0.59	10.5	7.36	0.20
Wales	10.7	4.1	0.00	7.73	3.90	0.06
<i>Election results (current)</i>						
Margin of victory	24.8	24.0	0.48	24.5	23.2	0.35
Conservative Winner	48.4	54.2	0.17	49.1	51.9	0.50
Labour Winner	41.8	34.1	0.06	42.4	39.8	0.53
Lib Dem Winner	0.93	1.43	0.60	1.50	2.60	0.33
Other Party Winner	8.92	10.0	0.66	6.98	5.63	0.51
N	213	419		401	231	