Estimating Election Incumbency Bias in General Elections in England

Marwan Riach

Imperial College London

mlr1617@ic.ac.uk

June 11, 2021

Overview

What is Election Incumbency Bias?

Election incumbency bias is the boost (or decline) in vote share awarded to a candidate due to their seeking re-election. There are two types of incumbency:

- Personal incumbency
- Party based incumbency

Omitted Variable Bias - Hell Hath No Fury Like a Variable Scorned!

OVB is when a statistical model leaves out one or more relevant variables. The bias results in the model attributing the effect of the missing variables to those that were included. The bias appears in the estimates of the parameters in regression analysis.

To see this, we consider the following model:

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \epsilon_i, \ \epsilon_i \sim N(0, \sigma^2) \tag{1}$$

where this models the true cause-and-effect relationship.

Omitted Variable Bias - Hell Hath No Fury Like a Variable Scorned!

Now let us consider another model which ignores the confounding variate x_{2i}

$$y_i = \beta_0^* + \beta_1^* x_{1i} + \epsilon_i^* \tag{2}$$

Ascertaining the relationship between models (1) and (2) will allow us to see the OVB. Let us further consider the relationship between x_{1i} and x_{2i} as:

$$x_{2i} = \gamma_0 + \gamma_1 x_{1i} + \nu_i \tag{3}$$

Substitution of (3) into (1) yields:

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2(\gamma_0 + \gamma_1 x_{1i} + \nu_i) + \epsilon_i$$

Omitted Variable Bias - Hell Hath No Fury Like a Variable Scorned!

Grouping the terms together gives us:

$$y_i = \beta_0 + \beta_2\gamma_0 + (\beta_1 + \beta_2\gamma_1)x_{1i} + \beta_2\nu_i + \epsilon_i \tag{4}$$

Equating coefficients of x_{1i} in (2) and (4) we find that:

$$\beta_1^* = \beta_1 + \beta_2\gamma_1 \tag{5}$$

and hence $\hat{\beta}_1^*$ is unbiased ($E(\hat{\beta}_1^*) = \beta_1$) if, and only if, $\beta_2\gamma_1 = 0$. We have used the fact that β_1 is unbiased as it represents the true relationship and is therefore the ordinary least squares estimator, which we know is unbiased.

Omitted Variable Bias - Demonstration

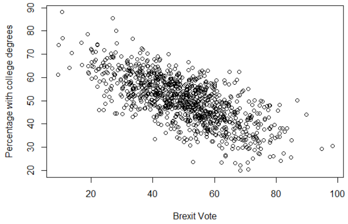
To see this in practice, we consider the following two models on simulated data:

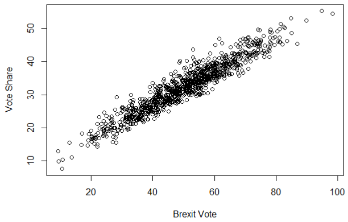
$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{3i} + \beta_4 x_{4i} + \epsilon_i \tag{6}$$
$$y_i = \beta_0^* + \beta_1^* x_{1i} + \beta_2^* x_{2i} + \beta_3^* x_{3i} + \nu_i \tag{7}$$

- y_i - vote share of the Conservative party
- x_{1i} - personal incumbency of candidate
- x_{2i} - proportion of voters with college degrees
- x_{3i} - median income
- x_{4i} - brexit vote

Model (6) is the true model and model (7) leaves out the brexit vote.

Omitted Variable Bias - Demonstration





Omitted Variable Bias - Demonstration

Predetermining the β_i values and fitting the simulated data to model (7) we see the corresponding results in the table below:

i	β_i	$E(\hat{\beta}_i^*)$	s.e.
0	20	59.7	0.000824
1	2	2.02	0.000317
2	-0.2	-0.593	$1.54 * 10^{-5}$
3	0.00007	0.0000694	$7.71 * 10^{-9}$

Table: Showing OVB through simulation

The results show us that the estimated value of $E(\hat{\beta}_2^*)$ (via Monte Carlo simulation) is negatively biased, agreeing with the theory established.

Estimating Incumbency Advantage by Individual

Notation:

- v_t^I - vote share obtained by party A at time t $(t = 1, 2)$ with an incumbent running
- v_t^O - vote share obtained by party A at time t when the seat is open
- I_t - 1 if party A incumbent runs for reelection, 0 if no incumbent runs (i.e. open seat), and -1 if party B incumbent is seeking reelection at time t
- Ψ_t - incumbency advantage granted to a candidate in election t.
- $f(v_t^O)$ - probability density function of the vote share at time t in an open race, that generates v_t^O in election t
- δ - partisan swing from time $t = 1$ to $t = 2$

Estimating Incumbency Advantage by Individual

Assumptions:

- Uniformity of swing δ in each constituency race
- Marginal density, f, is symmetric between elections 1 and 2, allowing for the swing i.e. $f(v_1, v_2 + \delta) = f(v_2, v_1 + \delta)$

The Sophomore Surge of an individual candidate running for party A is given by:

$$SS_A = v_2^I - v_1^O | I_1 = 0, I_2 = 1$$

Were this method to be unbiased, we would expect that $E(SS_A) = \Psi_2 + \delta$

Estimating Incumbency Advantage by Individual

In simulations:

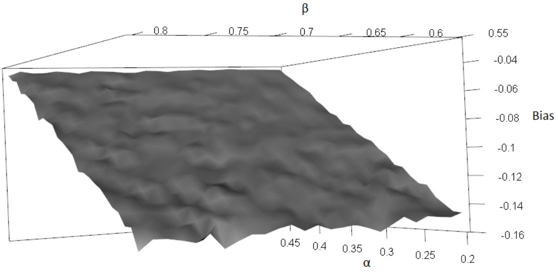


Figure: Simulation to show the bias of the sophomore surge method across parameters of α and β for a uniform distribution $U[\alpha, \beta]$.

The biases of all points on the surface are well below 0.

Estimating Incumbency Advantage by Individual

Gelman and King suggest the following unbiased regression method as an alternative to Sophomore Surge:

$$v_{i,t} = \beta_0 + \beta_1 v_{i,t-1} + \beta_2 P_{i,t} + \psi I_{i,t} + \epsilon_{i,t} \tag{8}$$

- $v_{i,t}$ - vote share at time t in constituency i
- $P_{i,t}$ - party that wins the election at time t
- $I_{i,t}$ - winning status of the incumbent candidate (0 in an open race, 1 if the Conservative is the incumbent and -1 if Labour is the incumbent)

Estimating Incumbency Advantage by Individual

- Model (8) used in USA congressional races
- We modified this for UK elections
- Only used seats where Labour or Conservative came 1st or 2nd since 1997 (selection bias)
- Only used seats with constant boundaries since 1997 (selection bias)
- Leave out seats won by the Speaker of the House

Estimating Incumbency Advantage by Individual

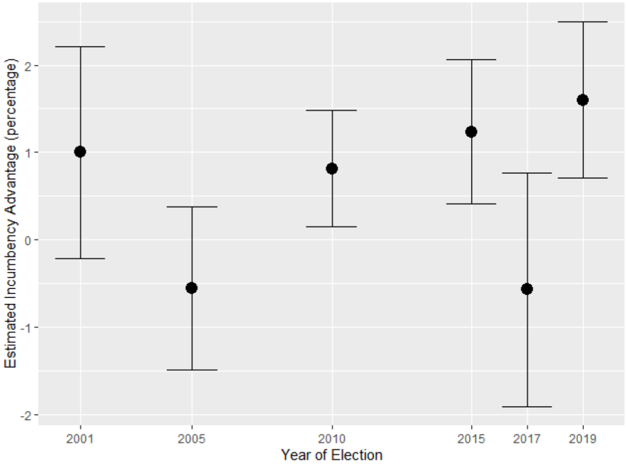


Figure: Estimates of incumbency advantage in each election from 2001-2019 with 95% confidence intervals

Estimating Incumbency Advantage by Party

What about the incumbency of the political party?

To investigate this, we implement the model in Eggers and Spirling which uses regression discontinuity design where the incumbency effect is determined by β_2

$$Y_{it} = \beta_0 + \beta_1 ConMargin_{i,t-1} + \beta_2 ConWon_{i,t-1} + \tag{9}$$

$$\beta_3(ConWon_{i,t-1} \times ConMargin_{i,t-1}) + \epsilon_{it}$$

Estimating Incumbency Advantage by Party

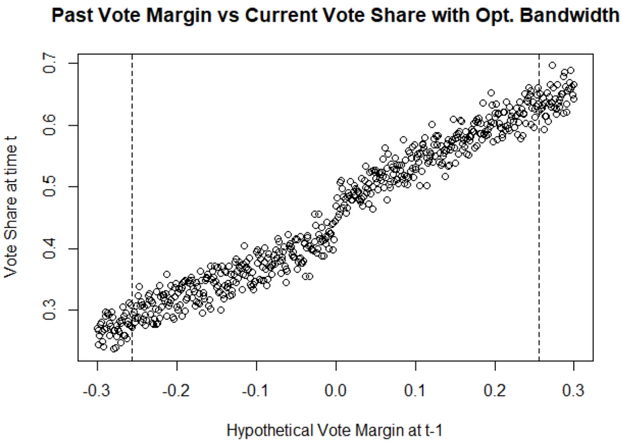


Figure: The intercept here clearly differs depending on whether or not the Conservative won at time $t-1$, with the change in intercept representing the incumbency effect.

Estimating Incumbency Advantage by Party

- Assume uniformity of swing
- β_2 strictly measures the sum of the party based incumbency for both the Conservative party and the Conservative party's opponent
- For precise party based incumbencies we could solve the following:

$$\beta_{2,C,L} = I_C + I_L$$
$$\beta_{2,L,LD} = I_L + I_{LD}$$
$$\beta_{2,C,LD} = I_C + I_{LD}$$

where I_C, I_L and I_{LD} are the specific party based incumbencies

Estimating Incumbency Advantage by Party

We propose that the effects of incumbency are more pronounced, the more marginal the seat.

To account for this, we implement a locally weighted regression with a triangular kernel:

$$\min_{\beta} \sum_{i=1}^{N} K_h(x_i)[y_i - \beta_0 - \beta_1 x_i - \beta_2 z_i - \beta_3 x_i z_i]^2$$

$$K_h(x) = 1_{|x| \leq h}\left(1 - \frac{|x|}{h}\right)$$

Estimating Incumbency Advantage by Party

We find optimal h using consistent estimates of the non-constant terms in the (10):

$$\text{AMSE}(h) = C_1 \cdot h^4 \cdot \left(m_+^{(2)}(0) - m_-^{(2)}(0) \right)^2 + \frac{C_2}{N \cdot h} \cdot \left(\frac{\sigma_+^2(0)}{f(0)} + \frac{\sigma_-^2(0)}{f(0)} \right) \tag{10}$$

- $m_{+/-}(0)$ - limit from right/left of the conditional mean of y_i at the boundary of 0
- $\sigma^2_{+/-}(0)$ - limit from right/left of the conditional variance of y_i at the boundary of 0
- $f(0)$ - marginal distribution of the forcing variable at 0

Estimating Incumbency Advantage by Party

Conservative Party Results:

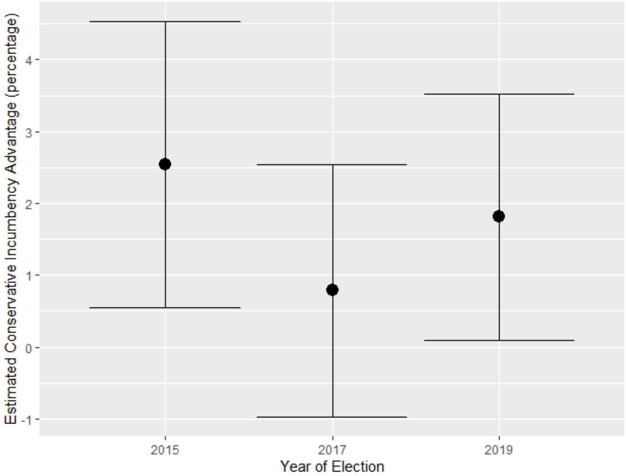


Figure: Conservative incumbency advantage estimates with 95% confidence intervals

Estimating Incumbency Advantage by Party

Labour Party Results:

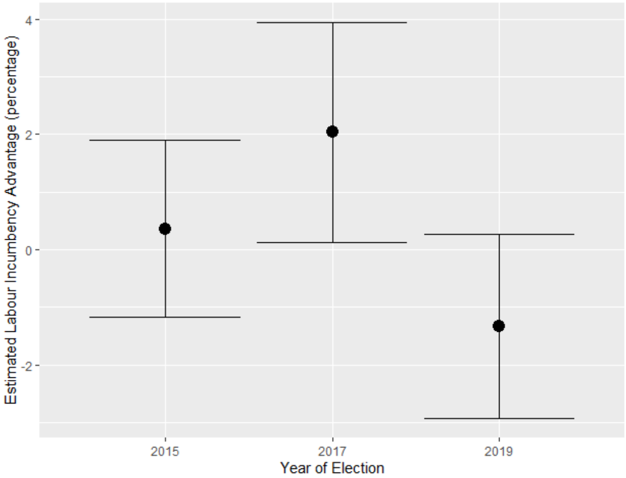


Figure: Labour incumbency advantage estimates with 95% confidence intervals

Estimating Incumbency Advantage by Party and Individual

- OVB suggests using both types of incumbency is better
- Personal incumbency $\Rightarrow$ party incumbency
- Previous results suggest vote share depends on both types of incumbency

Estimating Incumbency Advantage by Party and Individual

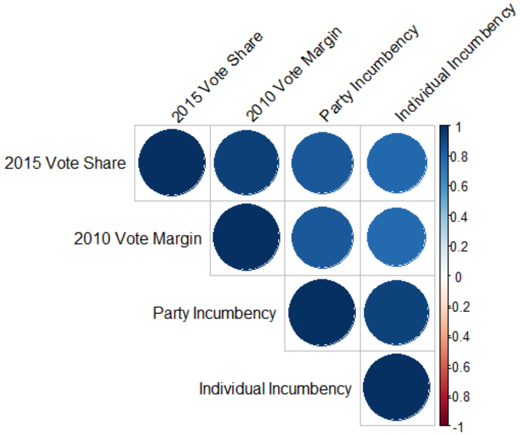


Figure: Correlations between covariates to show cause of OVB. The data used is from the 2015 election and the party examined is the Conservative party.

Estimating Incumbency Advantage by Party and Individual

We propose the following new and hitherto untested model which accounts for both types of incumbency:

$$Y_{it} = \beta_0 + \beta_1 ConMargin_{i,t-1} + \beta_2 ConWon_{i,t-1} \tag{11}$$

$$+\beta_3(ConWon_{i,t-1} \times ConMargin_{i,t-1}) + \beta_4 I_{i,t} + \epsilon_{it}$$

where I_{it} represents the status of the incumbent candidate (which is set to zero in an open race, 1 if the Conservative is the incumbent and -1 if the incumbent is not a Conservative). The same bandwidths as in the previous section are used.

Estimating Incumbency Advantage by Party and Individual

Conservative Party Results:

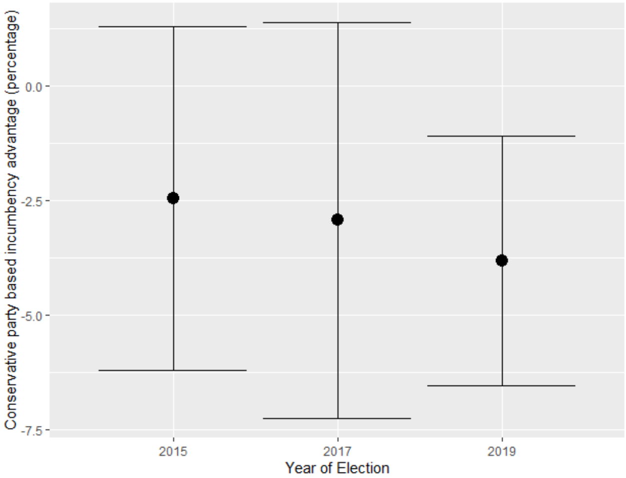


Figure: 95% confidence intervals for constituency incumbency of the Conservative party

Labour Party Results

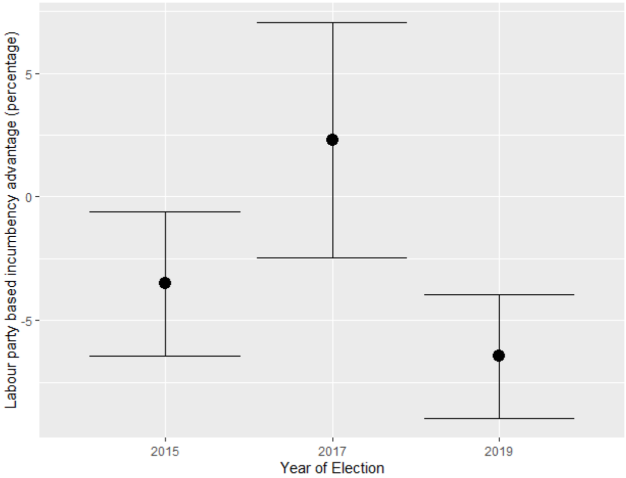


Figure: 95% confidence intervals for constituency incumbency of the Labour party

Personal Results

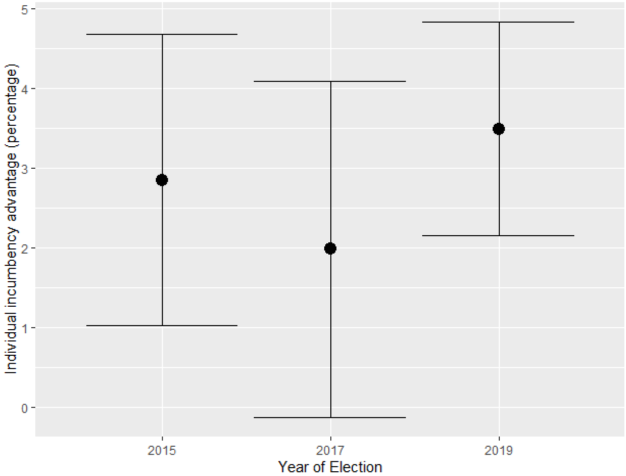


Figure: 95% confidence intervals for constituency incumbency of the individual candidate

Final Remarks

- Personal incumbency boost is around 3% (95% C.I. [1.01%,4.53%])
- Party based incumbency is much more mixed
- Precision of estimates could be improved by working with more granular data (e.g. local elections: 9,000 wards vs 531 constituencies) to accommodate swing uniformity assumption
- Choice of kernel could be investigated to better reflect the variance of vote shares at different margins of victories
- Brexit vote could be included in all models (it was omitted in the analysis in this report as it is only a recent phenomenon)

Final Remarks

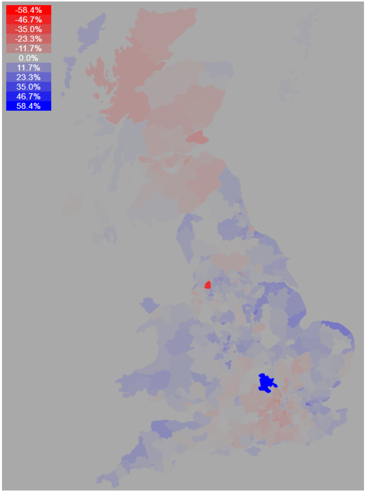


Figure: Vote share change in each constituency from 2017 to 2019 for the Conservative party

Final Remarks

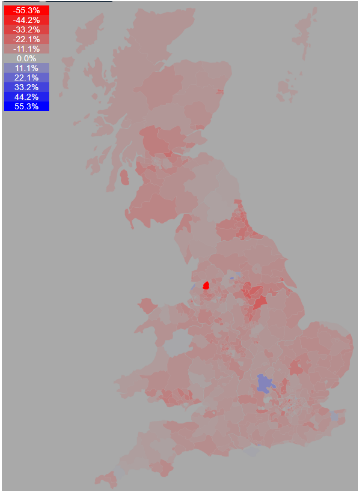


Figure: Vote share change in each constituency from 2017 to 2019 for the Labour party

Any Questions?

Appendix

Theory behind the bias in the Sophomore Surge method:

Considering $E(SS_A)$:

$$E(SS_A) = E(v_2^O + \Psi_2 - v_1^O | v_1^O > 0.5)$$

$$= \Psi_2 + \frac{1}{P(v_1 > 0.5)} \int_{0.5}^{\infty} dv_1 \int_{-\infty}^{\infty} (v_2 - v_1) f(v_1, v_2) dv_2$$

$$(v_2 = u_2 + \delta)$$

$$= \Psi_2 + \delta + \frac{1}{P(v_1 > 0.5)} \int_{0.5}^{\infty} dv_1 \int_{-\infty}^{\infty} (u_2 - v_1) f(v_1, u_2 + \delta) du_2$$

$$\leq \Psi_2 + \delta$$

(see report for full proof)

Appendix

Theory behind the simulated example:

$$
\begin{aligned}
E(SS_A) - \Psi_2 &= \frac{1}{P(v_1 > 0.5)} \int_{0.5}^{\infty} dv_1 \int_{-\infty}^{\infty} (v_2 - v_1) f(v_1, v_2) dv_2 \\
&= \frac{\beta - \alpha}{\beta - \frac{1}{2}} \int_{0.5}^{\beta} dv_1 \int_{\alpha}^{\beta} (v_2 - v_1) \cdot \frac{1}{(\beta - \alpha)^2} dv_2 \\
&= \frac{\alpha}{2} - \frac{1}{4}
\end{aligned}
$$