

Methodology for Projection

Part One Acquiring the data

Step One polling

A country-wide polling average was aggregated by using data from the following polling firms.

You Gov (April 13-14)

More In Common (April 11-14)

Focal data (April 11-14)

Lord Ashcroft (April 10-14)

Opinium (April 9-11)

Techne (April 9-10)

Find Out Now (April 9)

This national polling average is:

LAB: 25

CON: 22

REF: 24

LD: 13

GRN: 9

Step two: acquire demographic data.

- Demographic data was acquired from NOMIS (ONS) on the following metrics: Social deprivation, Highest Level of qualification, Age By single year, ethnic group.
- This data was acquired on a “district” or “unitary” electoral ward level (importantly **NOT on the county council ward level**).
- This data is from the 2021 census.

Here are the following definitions for the demographic data used.

Social Deprivation (in multiple dimensions)

Definition: a household is classified as deprived in multiple dimensions at least two dimensions if at least two of the following 4 conditions are met.

- Education: A household is classified as deprived in the education dimension if no one has at least level 2 education and no one aged 16 to 18 years is a full-time student.
- Employment: A household is classified as deprived in the employment dimension if any member, not a full-time student, is either unemployed or economically inactive due to long-term sickness or disability.
- Health: A household is classified as deprived in the health dimension if any person in the household has general health that is bad or very bad or is identified as disabled. People who have assessed their day-to-day activities as limited by long-term physical or mental health conditions or illnesses are considered disabled. This definition of a disabled person meets the harmonised standard for measuring disability and is in line with the Equality Act (2010).
- Housing: A household is classified as deprived in the housing dimension if the household's accommodation is either overcrowded, in a shared dwelling, or has no central heating. (NOMIS)

Data set

Highest Level of Qualification

“This dataset provides Census 2021 estimates that classify usual residents aged 16 years and over in England and Wales by their highest level of qualification. The estimates are as at Census Day, 21 March 2021.”(NOMIS)

The levels of Qualification are

- Level 1 and entry level qualifications: 1 to 4 GCSEs grade A* to C , Any GCSEs at other grades, O levels or CSEs (any grades), 1 AS level, NVQ level 1, Foundation GNVQ, Basic or Essential Skills
- Level 2 qualifications: 5 or more GCSEs (A* to C or 9 to 4), O levels (passes), CSEs (grade 1), School Certification, 1 A level, 2 to 3 AS levels, VCEs, Intermediate or Higher Diploma, Welsh Baccalaureate Intermediate Diploma, NVQ level 2, Intermediate GNVQ, City and Guilds Craft, BTEC First or General Diploma, RSA Diploma
- Level 3 qualifications: 2 or more A levels or VCEs, 4 or more AS levels, Higher School Certificate, Progression or Advanced Diploma, Welsh Baccalaureate

Advance Diploma, NVQ level 3; Advanced GNVQ, City and Guilds Advanced Craft, ONC, OND, BTEC National, RSA Advanced Diploma

- Level 4 qualifications or above: degree (BA, BSc), higher degree (MA, PhD, PGCE), NVQ level 4 to 5, HNC, HND, RSA Higher Diploma, BTEC Higher level, professional qualifications (for example, teaching, nursing, accountancy)
- Other qualifications: vocational or work-related qualifications, other qualifications achieved in England or Wales, qualifications achieved outside England or Wales (equivalent not stated or unknown)

Data set

Age by single year

This dataset provides Census 2021 estimates that classify usual residents in England and Wales by single year of age. The estimates are as at Census Day, 21 March 2021. (NOMIS)

Data set

Ethnic Group

“Description: The ethnic group that the person completing the census feels they belong to. This could be based on their culture, family background, identity or physical appearance. Respondents could choose one out of 19 tick-box response categories, including write-in response options.” (NOMIS)

Data set

Step three Acquiring previous electoral data

Three main sources of electoral data were used for this project.

2021 Local election data

First, and most importantly was the electoral data from the 2021 local elections for every seat currently being contested. This data was taken from the Commons Library archive on the 2021 local election. This was used as the baseline to calculate electoral swings for each electoral district.

Data set

Westminster results

Secondly, the constituency-level results from the 2024 General election and their swings compared to 2019 were compiled and aggregated for each council. This was one tool used to infer how significant the council-wide swing would be likely to be as the change in “party coalitions” from 2019-2024 is pretty similar to 2021-2025*. Ie Think of how the 2019 conservative party electorate was much more working-class than what ultimately became the 2024 conservative coalition. In this same vein, the 2021 conservative electorate was similar to its’ 2019 coalition while (according to opinion polls) the current conservative coalition is broadly similar to its 24’ GE coalition. Thus a similar trend of greater declines with working-class former conservatives is likely to repeat itself in the 2025 local elections.

Data set

District elections.

Finally, most multi-tier county councils had either district, borough, or city-level elections from 2022-2024. The results of these elections were used to infer where potential undersampled seat gains could be made by different parties. This was especially useful for predicting potential Lib Dem & Green gains as there has been a precedent of these parties gaining at the district level and consolidating at the county-wide level (or vice versa).

Data Set, (2), (3)

Part two Simplifying, formatting, and combining the data

Step 4 Simplifying demographic data

The previous demographic data was simplified to create an “index” to be used to measure the demographic of a ward. This section will explain how it was done and what it was used for.

Social deprivation:

Deprivation Index

The social deprivation index is composed directly of the data from the ONS/NOMIS indicating if a ward is deprived in two or more dimensions. This score was then compared to the England-wide average of “14.34” and swings were implemented accordingly. This was done to account for wards with higher levels of social deprivation

being more likely to vote for Reform while wards with lower levels of deprivation are more likely to vote for the other parties.

Highest Level of qualification

Qualification Index

A qualification Index was created by first removing the share of the population that described their qualification as “other” and then assigning the following “scores” to each level of qualification.

No Qualifications = 1

Level 1 and entry-level qualifications = 2

Level 2 qualifications = 3

Apprenticeship (or equivalent) = 4

Level 3 qualifications = 5

Level 4 qualifications = 6

Following this a “qualification index” was created for each ward. The qualification index’s theoretical range was 0-6 with 95+% of wards having a score between 3.0-4.6. This score was then compared to the England-wide average of “3.97” and swings were implemented accordingly

Age Index

An “average voter age index” was calculated by calculating the average age of the voting age population in each ward. This was done to account for the increasing importance of age as a political divide with younger voters more likely to vote for “left-wing” parties, and older voters more likely to vote for “right-wing” parties (when controlling for other variables).

Ethnic Group

Ethnicity Index

A “per cent white” Index was created by taking the data indicating what per cent of the ward's population identified as “white”. This was done to account for the trend of diverse wards voting for “Left-wing” parties at higher frequencies in contrast to “more white” wards being likely to vote for “more right-wing” (when controlling for other variables).

Religion

Muslim Index

A “per cent Muslim” Index was also created and used in a few councils (Lancashire, Leicestershire, Hertfordshire) where a significant share of the population identifies as Muslim. This was done so that the significant shift in voting behaviour seen

by British Muslims since 2021 (a vote shift away from the Labour Party and towards independents) could be accounted for

Step 6 Combining district-level ward data, converting boundary changes, and source of error

One of the biggest issues in this project was acquiring accurate exact demographic and political data for each ward as most of the council-level data I obtained was not entirely accurate for two reasons. First, the “ward” data I got from Nomis/ONS was for lower-tier wards for Multi-tier authorities and thus not accurate to the higher-tier county council level electoral districts. Secondly, many councils are changing electoral boundaries and thus the ward data (which is for 2022) is out of date. This is not ideal but it was the best data I had available. I do however recognize that this is a source of error in the methodology and I will now outline how I accounted for this data issue.

Combining District ward data

For county councils, I combined lower-level district wards into the corresponding county council-level wards following the boundaries of maps provided by the Local Government Boundary Commission for England (LGBCE). There are two sources of data error from doing this. First, The lower tier wards did not always line up correctly with the higher tier wards meaning there are a significant number of wards that have slightly incorrect corresponding demographic and political data. Secondly, the unknown population dispersion of each ward meant that lower-level wards that crossed the boundaries of two higher-tier districts may have been (in some limited instances) inaccurately assigned to a specific higher-tier ward.

Data_Source

Boundary changes

Many councils entirely changed their electoral boundaries. I attempted to atone for this data error by converting the old boundary map into the new one using maps provided by the LGBCE in a similar manner to how I attempted to correct for the issues stated around combining district ward data. It therefore, had the same issues with data accuracy and therefore it is an important qualifier to note that the electoral and demographic data on councils that changed their boundaries have a higher margin of error on those that didn't.

While this is a source of error in the data, the demographics and political data of each higher-tier ward and electoral district that changed boundaries were broadly accurate and thus useful in projecting wards.

Part three implementing vote swings

Once all of the data was collected and aggregated into the appropriate boundaries the next step could commence which is the implementation of electoral swings.

Step 7: electoral swing

Electoral swing was used for Labour, the Conservatives, and occasionally the Lib-Dems or Greens as they stood in most (or many) seats in 2021. Region-wide vote shares for the parties were first calculated from the council's previous electoral results, current (modified*) Nationwide polling, recent regional MRP polling (More in common, X) and council-wide demographics. Following this, the demographic swing method was implemented on a ward-by-ward basis using the factors outlined in Step 4.

Step 8: Reform

Reform could not be accurately judged via “an electoral swing” as they barely stood in any wards in 2021. Instead, I calculated their scores via a demographic analysis whereby I would give them higher vote shares in wards that were demographically favourable when compared to the national average. (ie if the wards' qualification index, ethnicity index, deprivation index, and age index were ~10% more favourable to reform when compared to the national average I would give them a vote share that was ~10% higher than their current projected national vote share.

Step 9: Green and the Lib Dems

I often calculated the Greens and Lib Dem's vote share via the electoral swing method outlined in Step 7 however in many instances the baseline of their vote share was calculated based on district level results in 2022-2024 as outlined in Step 3. Finally, in the case of the Greens (who did not stand in many wards in 2021), their vote share was calculated via a similar process of demographic analysis as outlined in Step 8 inwards in which they did not previously stand.

Step 10: Independents.

There is no way to accurately predict local independents as there is essentially no polling. So I projected independent scores based on the following factors

1) previous results (if the independent was a popular incumbent, longstanding councillor or a part of a longstanding independent group I marked them upwards.

2) significant online campaign? I researched each independent candidate (took a billion years) if they had a significant online campaign I marked them upwards if not, they got marked downwards.

3) fringe parties, if the independent was not an independent but instead part of a fringe political party (such as TUSC, or Heritage) I marked them down because these groups often do worse than candidates not standing as part of these groups.

Henry Snowdon