UNIVERSITY OF SOUTHAMPTON

Faculty of Social Sciences
School of Economic, Social and Political Science

Charting Westminster's Bubble: An Ideological Map of Britain's Digital Elite

by

Conor Gaughan
BSocSc, MSc
ORCiD: 0000-0002-0774-0111

A thesis for the degree of Doctor of Philosophy

June 2025

University of Southampton

Abstract

Faculty of Social Sciences
School of Economic, Social and Political Science

Doctor of Philosophy

Charting Westminster's Bubble:
An Ideological Map of Britain's Digital Elite

by Conor Gaughan

The social networking site, Twitter (now rebranded as X), is a microblogging platform which has garnered a reputation for being a platform for the elites. Not only are its users disproportionately members of socially privileged groups – younger, wealthier, and more highly educated – relative to other social networking sites, it has become the favourited social media platform for many high profile media and political elites. This is no less the case than in the United Kingdom (UK) where many of the country's politicians, journalists, media outlets, and political commentators have taken to the site en masse over the last decade and a half. The recent digitisation of Britain's commentariat has presented an opportunity to trace the networks of political and media elites in the UK and exploit them to gather a better understanding of important offline phenomena. Using advanced quantitative and computational techniques, this three-paper thesis leverages original large-scale digital data from the Twitter networks of UK political and media elites to address three key concepts: (1) intra-party competition; (2) media representation; and (3) dyadic representation.

Generating a formally validated set of left/right ideological estimates of UK Members of Parliament (MP) and a wider set of elite accounts that follow them, this thesis strives to build an ideological map of the UK's digital elite. **Paper 1** of this thesis uses these left/right estimates to model candidate endorsement in the September 2022 Conservative Party leadership contest, confirming that Liz Truss drew support from the further right of the party. **Paper 2** assesses ideological representation in the guest selection of seven flagship political programmes on the UK's six major T.V broadcasters between 2022 and 2024, finding that each of the seven shows selected from the right of the average elite Twitter user. **Paper 3** makes use of contemporary developments in small area estimation in the form of multilevel regression with poststratification to assess the dyadic relationship between an MP's Twitter profile and their respective constituencies, establishing a within-party responsiveness to the left/right position of their constituents along a social dimension.

Contents

List of Figures

List of Tables

Declaration of Authorship

I declare that this thesis and the work presented in it is my own and has been generated by me as the result of my own original research.

I confirm that:

1. This work was done wholly or mainly while in candidature for a research degree at this University;
2. Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
3. Where I have consulted the published work of others, this is always clearly attributed;
4. Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
5. I have acknowledged all main sources of help;
6. Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;
7. Parts of this work have been published as:

   Gaughan, C. (2024). Estimating Ideal Points of British MPs Through Their Social Media Followership. *British Journal of Political Science*. Pages 1-13. doi:10.1017/S0007123424000450. Copyright ©The Author(s), 2024. Published by Cambridge University Press. Reprinted with permission.

Signed:.. Date:..................

Acknowledgements

There are countless people I wish to thank for supporting me on my long journey to completing my doctorate, without whom most of what I have accomplished would not have been possible.

First and foremost, I must thank my three excellent supervisors for all the help and guidance they have given me over the last three years: Will Jennings, Rebecca Hoyle and Markus Brede. Your advice and support has been invaluable during my time as a PhD student, and a special thanks goes to Will in particular for introducing me to a range of great people and opportunities in the world of politics. I must also thank the wider politics team at Southampton which is packed full of extremely talented and supportive researchers who have helped me immeasurably during my studies. In particular, I would like to acknowledge Stuart Turnbull-Dugarte, Dan Devine, Matt Ryan, Paolo Spada, Jess Smith, Lawrence McKay and Rafa Mestre who have all supported me through my journey in one way or another during my time at Southampton. Additionally, I would like to thank all the team at the South Coast Doctoral Training Partnership (SCDTP) for all of their support and the opportunities they have enabled me to undertake while training for my PhD, as well as all the members of the wider SCDTP cohort. Most importantly, I am immensely grateful to the ESRC for providing the +3 studentship to fund my PhD project, without which I would never have been able to embark on this intellectual pursuit.

Away from Southampton, a special thanks goes to Patrick English for first introducing me to the world of social data science and for teaching me how to write my first ever line of code during my MSc in Exeter. Not only that, but it was Patrick who guided me to the opportunity of a PhD project in Southampton and supported my application along with Chris Playford, without which I likely would not have done one at all. An important thank you also goes to Anatoliy Gruzd and Philip Mai at the Social Media Lab at Toronto Metropolitan University for agreeing to host me for three months as a visiting scholar. My time in Canada was an unforgettable experience and my placement at the lab has helped me develop as a social media researcher learning from your extensive expertise. My gratitude extends once again to the SCDTP for agreeing to financially support this placement. I would also like to acknowledge the many fantastic international social media scholars that made up the ECPR Joint Sessions team in Toulouse in 2023 and the follow up SPARTA workshop in Munich in 2024. I'm extremely grateful for all of your knowledge, support and feedback on my work over the last two years and it is great to be a part of such a welcoming network of researchers from across the continent.

Finally, I want to say a heartfelt thank you to all my close friends and family who have supported me throughout my PhD. From all of the great friends I have made during my time in Southampton as part of the PhD community and in Building 58, to my friends and family in Manchester, Exeter and Coventry. Most importantly, a massive thank you to my partner, Izzy: I am forever grateful for your unwavering support.

Definitions and Abbreviations

API – **A**pplication **P**rogramming **I**nterface
BBC – **B**ritish **B**roadcasting **C**orporation
BESIP – **B**ritish **E**lection **S**tudy **I**nternet **P**anel
CA – **C**orrespodence **A**nalysis
DUP – **D**emocratic **U**nionist **P**arty
DTD – **D**igital **T**race **D**ata
EDM – **E**arly **D**ays **M**otion
ERG – **E**uropean **R**esearch **G**roup
EU – **E**uropean **U**nion
HoC – **H**ouse of **C**ommons
ITV – **I**ndependent **T**elevision (Channel 3)
MCA – **M**ultiple **C**orrespondence **A**nalysis
MCMC – **M**arkov **C**hain **M**onte **C**arlo
MII – **M**ost **I**mportant **I**ssue
MP – **M**ember of **P**arliament
MRP – **M**ultilevel **R**egression with **P**oststratification
PSB – **P**ublic **S**ervice **B**roadcaster
SCG – **S**ocialist **C**ampaign **G**roup
SDLP – **S**ocial **D**emocratic and **L**abour **P**arty
SMD – **S**ingle **M**ember **D**istrict
SNP – **S**cottish **N**ational **P**arty
UK – **U**nited **K**ingdom
UN – **U**nited **N**ations
US – **U**nited **S**tates of **A**merica
UGC – **U**ser **G**enerated **C**ontent

Chapter 1

Introduction

> “A “Westminster bubble” inhabited by the political class and Westminster journalists conducts debate in a way and on terms that have little relevance to the average citizen. ... They want to see, hear and read the merits of interesting ideas by ministers or shadow ministers instead of all sorts of angles and spin and process minutiae — endlessly fascinating and exciting to the incestuous Westminster bubble but boring and dull to everyone else.”
>
> **The Lord Hain, *Time to Prick the Westminster Bubble* - (Hain, 2012)**

> ″The trouble with Twitter, the instantness of it - too many twits might make a twat.”
>
> **David Cameron talking on Absolute Radio, 2009**

1.1 Charting Westminster’s Bubble

The “Westminster bubble” is a pejorative characterisation of the political and media class in the United Kingdom (UK) as being isolated and out of touch with life outside of Parliament in Westminster, London. While it primarily relates to official Members of Parliament (MP) and Peers, it has also extended to wider members of the political commentariat such as media broadcasters, journalists, lobbyists, researchers, policymakers, pollsters, and academics. It is a term that is akin to “Inside the Beltway” in the American political lexicon and depicts a political culture that is disconnected from the general interests and priorities of the wider public. It is a popular term for UK politicians where many have previously used it to disregard issues they consider menial, or as a populist trope to slander their opposition or position themselves as outside of the bubble. When Conservative Party MP Michael Howard announced his appointment as Leader of the Party in November 2003, for instance, he sought to present himself as a leader outside of the Westminster bubble in an attempt to shift himself closer to the electorate (Hall, 2003). After receiving criticism for a controversial remark made during a party leaders' debate about immigrants arriving in the UK with HIV, Nigel Farage disregarded the outrage as unique to the Westminster bubble and not shared by the general public (Swinford, 2015). More recently, in response to the controversy surrounding the UK health secretary Matt Hancock breaking social distancing guidelines during the COVID-19 pandemic, then Prime Minister Boris Johnson dismissed Hancock's conduct as simply "stuff going on in the Westminster bubble" (Stone, 2021).

The term itself is arguably overused and not entirely accurate. Nonetheless, it is a commonly known phrase for describing the centralisation of political and socio-economic discussion into the hands of a small number of elites, many of whom derive from more socially privileged backgrounds. To emphasise the sense of disconnection between members of the Westminster bubble and the wider public, trust in politicians, government ministers and journalists to tell the truth has reached its lowest level since records began in the UK, ranking lower than advertising executives and estate agents (Clemence and King, 2023). Additionally, the majority of the British public believe that politicians do not understand the lives of people like them and that voters have far less influence over public policy decisions than more elite actors like lobbyists, the media, party donors and corporations (Patel and Quilter-Pinner, 2022, pp.21-22). In his seminal work *The Structural Transformation of the Public Sphere – An Inquiry into a Category of Bourgeois Society* in 1989, German philosopher Jürgen Habermas expressed the importance of a functioning public sphere to a healthy democracy: a free and unrestricted space where citizens engage in critical public debate free from elite intervention and a sense of public opinion can be formed (Habermas, 1989). More recent work by Habermas in the early 2000s critiqued the reality of the public sphere in modern Western society, where the asymmetrical dynamic of political communication and contemporary mass media continues to place political, media, financial, and cultural elites at the centre of political discussion and relegates the voices of ordinary people to the periphery (Habermas, 2006).

At the same time during the early 2000s, the world was experiencing the rapid emergence of what became known as Web 2.0 (Toledano, 2013). In the early stages of the World Wide Web in the 1980s, web pages hosted on the Internet were typically static and one-dimensional (Web 1.0). Web 2.0 signalled the arrival of a more dynamic and multi-dimensional digital space at the turn of the millennium which transformed Internet users from passive consumers to active creators, where websites and applications became an interactive and participatory experience. Web 2.0 placed its users at the centre and user generated content (UGC) became the dominate form of information that existed on the modern web.[1] The democratising potential of Web 2.0 and its constituent components like social networking sites, Web applications and wikis gave promise to the realisation of a genuine online public sphere (Papacharissi, 2010). No explicit barriers to entry for citizens, free from elite interference, where ordinary people could post and share their thoughts and ideas in an open digital environment. One social networking site in particular emerged during this period which many hung their hats on to fulfil the promise of a digital public sphere: *Twitter*. Twitter emerged on the social media scene in 2006 and offered a unique user experience compared to other early competitors like Facebook and Myspace with its non-reciprocal network structure and microblogging-style features. The inherently asymmetric structure of Twitter promoted a more hierarchical top-down flow of information where posts were short and could be published near-instantaneously (Gruzd et al., 2011). This made it a popular communication medium for politicians, journalists, commentators and academics to disseminate opinions, policies, news and research between themselves and to wider users. Over the next fifteen years, the platform saw significant uptake by various elite actors in both the UK and internationally and is now widely regarded as an *elite space* (Blank, 2017; Dagoula, 2019).

As of 2022, 591 out of 650 (91%) UK MPs in the House of Commons (HoC) had active accounts on Twitter accompanied by tens of thousands of wider British media and political elites (Gaughan, 2024, nd). Much of the Westminster bubble has digitally transitioned itself onto the platform over the last decade and a half and this bubble remains equivalently insular online in regard to connections and communication (Dagoula, 2019; McLoughlin, 2019; O'Malley, 2019). More importantly, studies have shown that the majority of ordinary Twitter users do not follow political opinion leaders or use the platform for political purposes (Mukerjee et al., 2022; Wojcieszak et al., 2022). Taken together, the core-periphery structure of Westminster elites and the ordinary public largely persists on the site and contradicts the expectations that Twitter would constitute an open public sphere (as is explored in more detail in **Chapter 2**). At the very least, however, the transference of much of the UK's commentariat to Twitter over the past fifteen years provides an unprecedented opportunity to digitally trace and map the Westminster bubble in a way that would have been impossible only two decades prior. The many ways in which elite British actors connect and communicate between one another on the site provides a rich source of information that can be exploited for a number of purposes.

Specifically, through the application of advanced quantitative methods, this thesis seeks to leverage the homophilic and ideologically polarised structure of elite networks on Twitter (Aiello et al., 2012; Barberá et al., 2015; Colleoni et al., 2014) to

[1] UGC is an umbrella term for any form of content posted by users on an online platform and can range from websites, blogs, wikis, and video channels to costumer reviews, podcasts, social media posts, and personal advertising.

estimate an ideological map of UK political and media elites. In doing so, it strives to improve our understanding of three key areas in British legislative and communicative politics: intra-party competition, media representation, and dyadic representation.

1.2 A Brief Summary of Twitter

Twitter is a social networking site that allows users to post and interact with messages commonly known as “tweets”. Initially developed as a platform for short microblogging-style communication, it was launched in 2006 by its four founders: Jack Dorsey, Noah Glass, Biz Stone, and Evan Williams. Over the next decade, Twitter evolved into a global phenomenon for communication, news sharing, and social networking. The defining feature of the platform was its 140-character limit for tweets, later expanded to 280 characters in 2017, reflecting its emphasis on brevity and simplicity. In its infancy, Twitter served as a platform for casual social interaction between its modest user base but as it grew, it became a popular and far more professionalised tool for journalists, politicians, celebrities, and businesses. It also developed into a hub for real-time news sharing, bolstered by its unique features for trending topics and use of “hashtags”. Unlike many other social networking sites, the Twitter platform is heavily unidirectional where users “follow” other users but they do not have to follow the user back in return (unlike platforms such as Facebook or LinkedIn where connections have to be reciprocal). Users who follow an account are commonly referred to as that account’s “followers”.

Some of its key features include:

Tweets: Publicly visible posts (by default, but users can set to private) of no more than 280 characters which can also contain other content such as URL links, images and videos. Other users can engage with public tweets in various ways including retweeting (reposting and amplifying the tweet forward) and quote tweeting (embedding the tweet inside another tweet to add additional commentary), as well as liking and/or replying to the tweet or privately bookmarking it so that it can be easily revisited later.

Timelines: There are different types of timelines in the Twitter lexicon. Generally, a Twitter timeline refers to a list of tweets displayed in chronological order from when they were posted. The Home timeline is what every user sees by default and contains lists of tweets posted by the accounts they follow, as well as recommended tweets from accounts that users may not follow but are recommended based on a variety of algorithmic signals. A user’s timeline refers to the chronological list of tweets (incl. retweets and replies to other tweets) that each user has posted themselves. This is usually publicly available but can be made private at the discretion of the user.

Hashtags: Words or phrases prefixed with a # symbol. These became a hallmark feature of Twitter in its early years and are used to group and categorise tweets together by topic . This can be useful for boosting the visibility of tweets and allowing users to track and follow trending topics or topics they are interested in.

Usernames: Each user account has a unique username (also referred to as a handle or screen name) which is prefixed by an @ symbol. These can be used in tweets to directly address another user account and this is referred to as a “mention”.

Trending Topics: A word, phrase, topic or hashtag that is mentioned in tweets at a greater rate on the platform than is typical is said to be “trending”. Trending topics become a curated list of popular discussions that users can follow to keep up-to-date with real-time global events.

Verification: Blue checkmarks given to certain user accounts that are used to identify users that are active, authentic and notable. This was initially designed to prevent impersonation accounts and usually includes profiles with some form of offline notability such as companies, brands, public figures, government organisations, political parties, and media outlets.

In October 2022, Twitter was acquired by businessman Elon Musk and rebranded to the name X in July 2023. Following its acquisition, a number of substantial changes were made to the platform including the implementation of a paid subscription service known as X Blue (now X Premium) which modified the meaning of verification, the creation of long-form text in tweets, account monetisation options, and audio-video calling.

Research for this thesis began in October 2021 and all data extracted from the platform was conducted before its rebranding to X. Thus, for the sake of clarity and simplicity, the platform will continue to be referred to as Twitter throughout the thesis. Occasionally, some figures may make reference to Twitter as Twitter/X. This is to reflect any analysis that may have been conducted after the rebranding to X was made, but this is rare.

1.3 Research Questions & Objectives

The overarching aims of this thesis are primarily two-fold:

1) to provide a methodological demonstration of the informative power of social media networks (namely, Twitter) in bettering our understanding of important political concepts that are difficult to ascertain via other mediums; and

2) to provide a substantive contribution to the fields of UK legislative and communicative politics.

To accomplish these objectives, this thesis adopts a three-paper format, where each individual paper forms a single chapter (**Chapter 4**, **5** and **6**) and each paper builds on the data and analysis from the previous one to form the empirical spine of this research.

First and foremost, this thesis seeks to generate ideological estimates (formally referred to as "ideal points") of UK political and media elites using the Twitter follower networks of UK MPs. This approach is not strictly novel and builds heavily on the earlier works of Barberá (2015) and Barberá et al. (2015). However, this has yet to be applied and explored on a large scale to the UK and, as is explored in detail in **Chapter 4,** Westminster provides an especially idiosyncratic case study owing to its restrictive parliamentary culture and strict whipping system (Hix and Noury, 2010; Spirling and McLean, 2007). It strives to illustrate the way in which large-scale follower networks extracted from Twitter can be leveraged to circumvent the restrictions of a Westminster-style legislative like the HoC to benefit our understanding of intra-party competition. In the case of **Chapter 4,** intra-party competition is assessed in the context of the September 2022 Conservative Party leadership contest but its applications can extend far beyond leadership contests to better understand a range of within-party conflict and electoral competition. In doing so, two research questions are explored:

1) Can the Twitter follower networks of UK MPs be successfully used to estimate valid measures of their left/right ideal points?

2) If so, can these estimates be used to model the ideological component to the September 2022 Conservative Party leadership contest?

Building substantively on the analysis of **Chapter 4**, the next paper follows on from this to ask an additional three exploratory questions:

1) Can the sets of UK MPs that wider elites follow on Twitter be successfully used to estimate valid measures of their left/right ideal points?

2) If so, what is the general ideological distribution of the Westminster elite on Twitter?

3) What is the ideological representation of guests who appear on the flagship political programmes of the UK's major television broadcasters?

As with the case of Westminster in **Chapter 4,** the UK's media broadcasting landscape is similarly restrictive, legislating for a strict impartiality regime which has come under increasing scrutiny in recent years for allegations of political bias on major broadcasters (Barnett and Petley, 2023). **Chapter 5** seeks to build on the work of **Chapter 4** to illustrate the way in which social media networks of political and media elites can be leveraged to better understand political representation within the media environment.

Again, further building on the works of **Chapter 3** and **Chapter 4, Chapter 6** looks to leverage both the social media networks *and* topics of discussion posted by UK MPs on the Twitter platform to assess the within-party responsiveness of MP online behaviour to the positions of their constituents. Commonly known as the concept of *dyadic representation* (Weissberg, 1978), the final paper of this thesis asks the following three questions:

1) Are the Twitter follower networks of UK MPs individually responsive to the economic and/or social left/right positions of their constituents beyond party affiliation?

2) Are the Twitter positive engagement networks of UK MPs individually responsive to the economic and/or social left/right positions of their constituents beyond party affiliation?

3) Are the topics of UK MP Twitter posts individually responsive to the most important issues of their constituents beyond party affiliation?

As one can see from the eight research questions outlined across the three separate chapters (papers), the objectives of this thesis are largely exploratory in nature. The purpose of this thesis is not to provide a causal explanation for the concepts of interest. It does not posit a theory as to why Liz Truss may have won the Sept 2022 Conservative leadership contest, nor definitively explain the causal mechanisms that drive guest selection on major television broadcasters in the UK. It does not look to establish a causal relationship between the positions of UK constituents and the Twitter behaviour of their MPs. Instead, what it does strive to do is illustrate the informative power of social media networks and the potential it unlocks for better understanding an array of important concepts in political and social research. The methodological approach to understanding the concepts of intra-party competition, media representation, and dyadic representation in this thesis are novel contributions to each space (particularly in the UK context) and, in the process of doing so, provide some illuminating insights that will be of interest to researchers in legislative study and political communication.

1.4 Outline of the Thesis

The following chapter of this thesis (**Chapter 2**) outlines the theoretical underpinnings of using Twitter data for social and political research. It serves to position the empirical research of the thesis around the basis of the Twitter platform as an elite digital space; one where, in the case of the UK particularly, its user base is disproportionately drawn from more elite members of the general population making it an especially informative data source. Specifically, the chapter explores the characteristics of British Twitter users and how this compares to the user bases of other major social media platforms such as Facebook, Instagram, YouTube and TikTok, as well as social media non-users and the general public. It then moves on to discuss the various methodological and technological biases that exist on a site like Twitter and how this violates the terms of Habermas' public sphere. Finally, it examines the reality of Twitter as a politically polarised network, the popular concepts of echo chambers and filter bubbles, and how the ideological structure of elite Twitter networks combined with the eliteness of ordinary Twitter users can be leveraged to estimate ideal points of political and media elites.

Chapter 3 then follows on to outline the methodological underpinnings of the thesis and, specifically, the data sets that form the cornerstone of the empirical analysis. It describes the process of data extraction from the Twitter platform, the search queries used to harvest the data, and how these digital data sets are stored. Two data sets are extracted from the platform which are both discussed in detail. It also discusses some of the technical and ethical issues with working with digital data and how some of these issues have been addressed in this thesis. The chapter also covers the other data sets that have been used in the three empirical chapters, namely two primary surveys that have been commissioned exclusively for this thesis and details of the main secondary survey data source used to augment the analysis (primarily in **Chapter 6**). Finally, it discusses broadly the methods that are applied in each of the three empirical chapters to estimate ideal points of political and media elites using Twitter networks **(Chapter 4)**, how these are used to assess ideological representation on major UK broadcasters **(Chapter 5)**, and how constituency position is estimated to assess dyadic representation **(Chapter 6)**.

The following three chapters then contain the three empirical papers which form the core of this thesis. **Chapter 4** (Paper 1) is a primarily methodological chapter which acts as the backbone of the succeeding two papers and is essential to the creation of the ideological map of Britain's digital elite. It describes the method of data extraction from the Twitter platform and how one can estimate ideal points of elites using such data. These ideal points are subsequently validated and then used to model the ideological component to the Sept 2022 Conservative Party leadership contest as a use case for how these estimates can be used. **Chapter 5** (Paper 2) directly builds on the analysis of **Chapter 4** by using the data and ideal point estimates generated in the first paper to assess ideological representation on major UK broadcasters between 2022 and 2024. Finally, **Chapter 6** is the most methodologically and substantively advanced of the three empirical papers. Combining a contemporary technique for small area estimation known as multilevel regression with poststratification (MRP) with the ideal point estimates for elite accounts derived from the previous two papers, MP-to-constituency responsiveness is assessed as a concept of *digital dyadic representation.* Additionally, this paper incorporates a second Twitter data set including the tweet timelines of UK MPs to assess not only the left/right ideal points of MPs based on their Twitter networks and their relationship to their constituencies, but also their general topics of discussion based on the content that they post.

Finally. **Chapter 7** contains a conclusionary section which summarises the work conducted for this thesis and discusses the key insights that it has generated. It considers the novelty of the results that each of the three empirical chapters has produced collectively and the intellectual contribution that they make to the fields of legislative politics, media and communication, and political representation. It evaluates the strengths and limitations of the research and potential new avenues that can be built upon in future work. Most importantly, the chapter serves to contextualise the findings of this thesis in light of the current state of social media and digital research in 2024 and how this strengthens the value of this research moving forward.

The next chapter will now begin the thesis by outlining the novelty of Twitter as a data source and how it can be leveraged to estimate ideological maps of political and media elites.

Chapter 2

Twitter as an Elite Digital Space

> ”Why did all the pollsters and the pundits get it so wrong? Because, fundamentally, they didn’t understand the people who make up our country. The vast majority of people aren’t obsessives, arguing at the extremes of the debate.
>
> Let me put it as simply as I can: Britain and Twitter are not the same thing”
>
> **David Cameron, Conservative Party Conference Speech, 2015**

2.1 Twitter: Is it Britain?

In his keynote speech at the Conservative Party conference in October 2015, then Prime Minister David Cameron famously took a swipe at Britain’s political commentariat. The conference had followed on the back of a surprise election victory earlier that year where many pollsters and analysts had forecasted another hung parliament. Instead, Cameron’s party went on to secure a second successive term in office and the first (albeit wafer-thin) Conservative majority in the House of Commons since 1992. It was a result that confounded the pundits and sparked a sense of contempt towards the pollsters who had spent the majority of the election campaign suggesting an entirely different outcome. An industry-wide review conducted shortly after the election by the British Polling Council found that the opinion polls leading up to the election were among the most inaccurate since surveying began and had been primarily caused by an over-sampling of Labour voters and under-sampling of Conservatives (Sturgis et al., 2016). This might then explain the Prime Minister’s disdain for Britain’s so-called political experts, but why the jibe at Twitter specifically? In his speech, Cameron went on to distance ordinary members of the public from Twitter users by referring to them as “decent, sensible and reasonable”, suggesting that the latter were anything but. The hyper-politicised nature of sections of Britain’s

"Twitterati" is well-established – particularly among its power-users (Coughlan, 2024) - and the explanation for this is relatively straightforward. When compared to the general population, British Twitter users are, on average, younger, wealthier, more highly educated, more likely to be male, more politically engaged, left-leaning, and more socially liberal (Blank and Lutz, 2016; Mellon and Prosser, 2017; Singh, 2018; Sloan, 2017). Had British Twitter users alone voted in the 2015 and 2017 General Elections, the left-wing Labour Party would have emerged victorious in both and Remain would have won the Brexit referendum by a comfortable margin of 68-32 (Singh, 2018).[1]

2.1.1 The Left-Liberal Bias of British Twitter

While many of the studies that assessed the representativeness of British Twitter users were conducted in the mid-to-late 2010s, an up-to-date examination of British Twitter users in 2024 would indicate that not much has changed. Based on Wave 28 of the British Election Study Internet Panel (BESIP) carried out in June-July 2024, the average Twitter user still illustrates a left-liberal bias when compared to non-Twitter users (See **Figure 2.1**).

[1] Singh's analysis shows that the same can be said for Facebook too, though its user base is far less unrepresentative of the population than Twitter's.

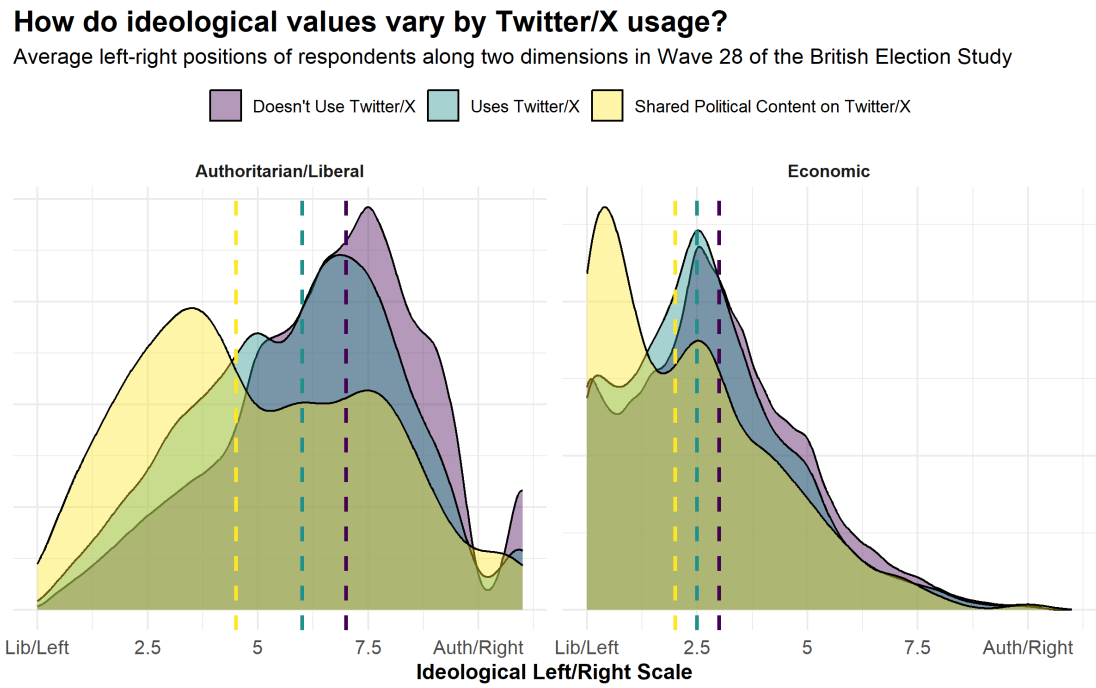


FIGURE 2.1: Density plots showing the ideological distribution of 31,582 respondents in Wave 28 of the British Election Study Internet Panel. Density curves are smoothed using kernel density estimation with a default bandwidth (nrd0). Values along the left/right scale (x-axis) represent the added and scaled score of respondents to a set of 5 value-based questions relating to each dimension respectively. Dashed lines represent the median score in each group.

This bias is even more extreme for Twitter users who had posted or shared political content on the site at any point over the previous 4 weeks. Ideologically, Twitter users are still disproportionately to the left of the general population, even after its recent acquisition by Elon Musk in 2022. This bias translates quite clearly into electoral behaviour where there is a distinct inverse relationship between support for the Conservative and Labour parties in the 2019 General Election as Twitter usage increases (See **Figure 2.2**).

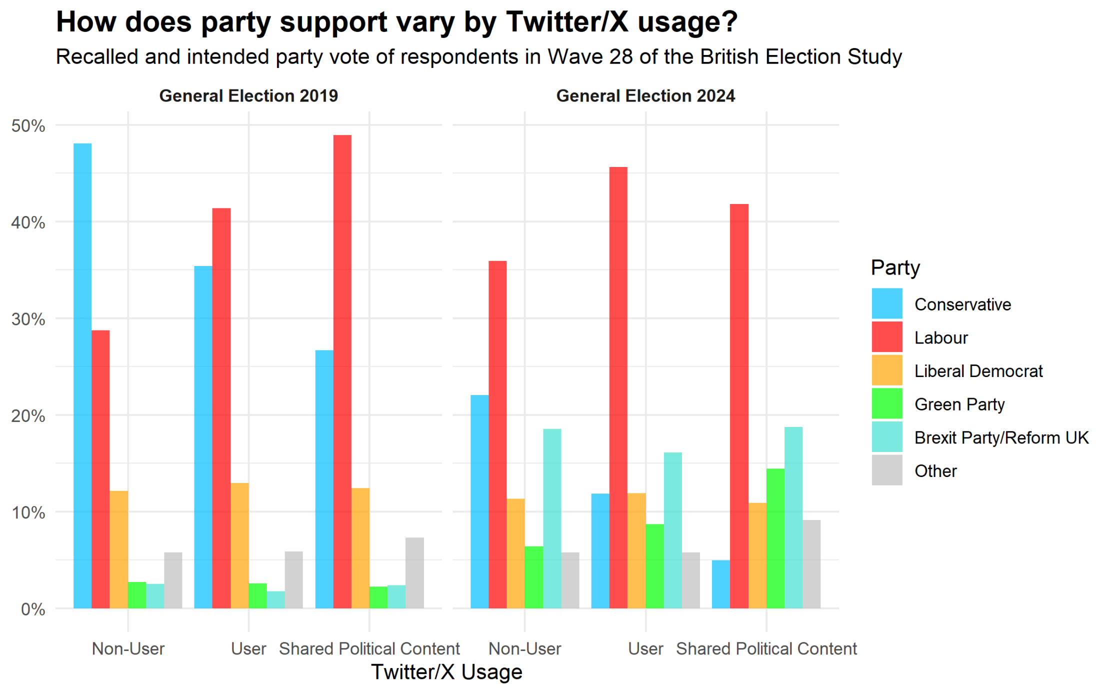


FIGURE 2.2: Clustered bar chart which illustrates the percentage of respondents within each Twitter usage group who said they either voted (GE:2019) or intended to vote (GE:2024) for each political party. This is based on 29,493 respondents in Wave 28 of the British Election Study Internet Panel and excludes those who said they would/-did not vote or didn't know.

Support for the Labour party is far higher among Twitter users than non-users and substantially higher for users who posted political content on the site compared to general users. Had only Twitter users who post political content on the platform voted in the 2019 General Election, Labour would have claimed almost 50% of the popular vote, compared to only the 32% they actually achieved. The same pattern is observed in voting intention in the 2024 General Election, where Conservative vote share drops significantly at each Twitter usage level. Among Twitter users who posted political content on the platform, less than 5% said they would vote for the Conservatives. Interestingly, however, this effect is less observed among minority parties. With the exception of the left-wing Green Party in 2024, whose vote share increases at each level of Twitter usage, support for smaller parties remain relatively stable across usage groups.

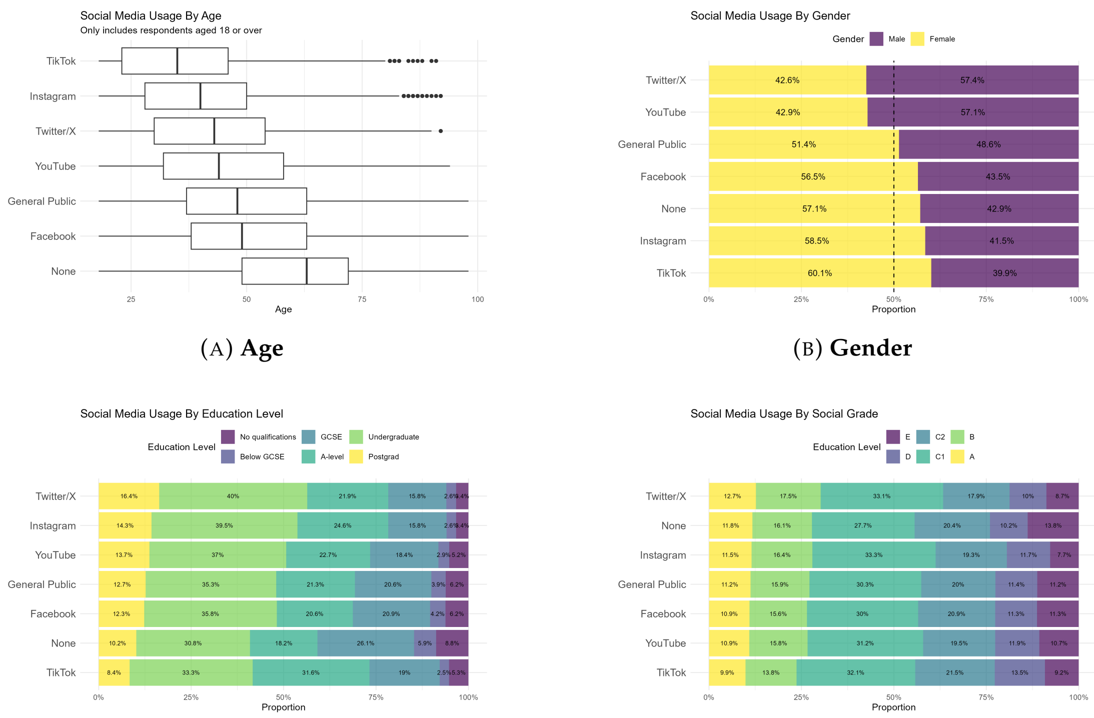
This perhaps suggests that the unrepresentativeness of British Twitter is not purely a function of ideological bias alone, particularly when considering that support for right-wing Reform UK in the 2024 election is higher among those who share political content on the site than non-users. Mellon and Prosser's work showed that, when controlling for age, gender and education, the differences between social media users and non-users in their political behaviour and values are no longer statistically significant (Mellon and Prosser, 2017). They suggest that disproportionately higher levels of support for the Labour Party among Twitter and Facebook users is primarily due to the demographic composition of social media users generally. If one combines this fact with the knowledge that Labour's vote share has risen substantially among younger and more educated populations over the last two decades (Jennings and Stoker, 2016, 2017) - a trend which has held in the previous two UK elections (McDonnell and Curtis, 2019; McDonnell, 2024) - it likely explains the left-leaning skew that still persists among British Twitter users. This dispels the notion that the platform simply orientates itself towards more left-liberal voters, instead suggesting a more organic construction of its user base.

2.1.2 Digital Divides and Self-Selection Bias

Taking all of this knowledge into account, it would confirm that Cameron's statement was entirely correct: Twitter *is not* Britain. It's demographic composition is disproportionately weighted towards a particular subset of the general public, consequently generating an unrepresentative population of users. One explanation for this may be the digital divides in the UK that perpetuate asymmetries in the access and usage of digital technologies by different social groups (Rogers, 2001). A report from the Office for National Statistics illustrates that, whilst the number of overall Internet non-users has declined over the last decade, there are still significant disparities in usage that cuts across regional, gendered, ethnic, socioeconomic, educational, disability, and age groups (Serafino, 2019). More specifically, Internet users in the UK are, on average, more likely to be younger, male, non-white, of higher socioeconomic status, highly educated, from more urban areas, and non-disabled. Many of these characteristics can be associated with socially privileged groups in the UK and this bias similarly translates onto Twitter though to an even more extreme degree. Research by Blank (2017) found that British Twitter users are younger, wealthier, and better educated than other Internet users, who are in turn younger, wealthier, and better educated than Britain's offline population. Given this fact, digital divides alone cannot account for the elite bias of Twitter's user base, especially when comparing it to other social media platforms like Facebook and Instagram that do not have such unrepresentative populations (Blank, 2017; Singh, 2018).

This discrepancy can be accounted for by a second important factor: self-selection effects. Certain social media platforms will appeal to certain groups of people more than others because of differences in technical features, user bases, and community cultures (Lohmann and Zagheni, 2023). For example, taking a look at the composition of users on the world's most popular social media platforms for 2024, different sites attract different types of users: Instagram and TikTok are the platforms of choice for under-35s while Facebook and WhatsApp are favoured by those 35 and over (Kemp, 2024).

These differences are important when studying the social media sphere and reminds researchers that social media as a concept is not monolithic. Twitter is uniquely an *elite space* (Blank, 2017; Dagoula, 2019) and a comparison of demographic compositions of users of major social networking sites in the UK in 2024 would only serve to reaffirm this.[2] Again, using data from Wave 28 of the BESIP, **Figure 2.3** illustrates the composition of users of the UK's five major social network platforms (Facebook, Instagram, Twitter, YouTube and TikTok) as well as social media non-users and the general public, split by age, gender, education level and social grade.



(A) **Age**

(B) **Gender**

(C) **Education Level**

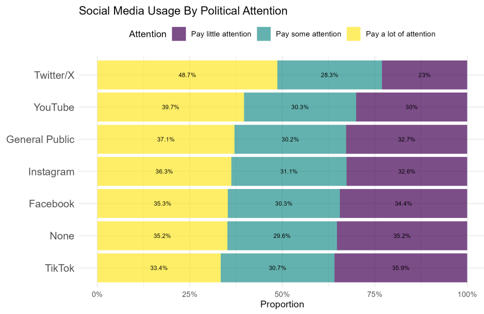
(D) **Social Grade**

FIGURE 2.3: Demographic composition of different social media user groups in the UK as well as the general public and social media non-users. Plot A illustrates boxplots of age distribution by group and plots B, C and D illustrate stacked bar charts of gender, education level and social grade. Plot A y-axis is ranked in ascending order by median age, while plots B, C and D are ranked in descending order by % Female, % Postgrad, and % Grade A respectively. Wave 28 of the BESIP, n = 31,582.

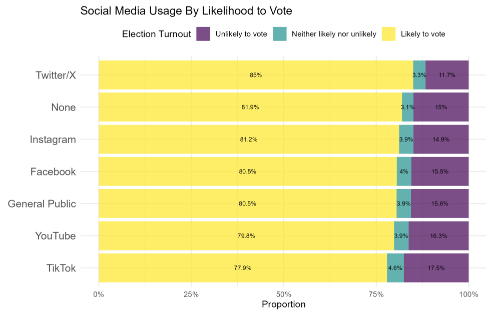
As one can see, British social media users broadly are not representative of the general population, though so is the case for those who do not use any of the five social media sites at all. Facebook appears to align the closest with the general public along the four demographic dimensions which is perhaps a result of the site having the overall largest user base. Along every dimension, Twitter users are still not reflective of the British general public in 2024: they remain younger, more likely to be male, more highly educated, and from a higher socioeconomic background. More importantly,

[2] To note, the term *elite* in reference to Twitter users is used in two different ways here. It can be in formal reference to official public figures like politicians and celebrities but also to ordinary users who demonstrate characteristics more commonly associated with socially privileged groups (Blank, 2017).

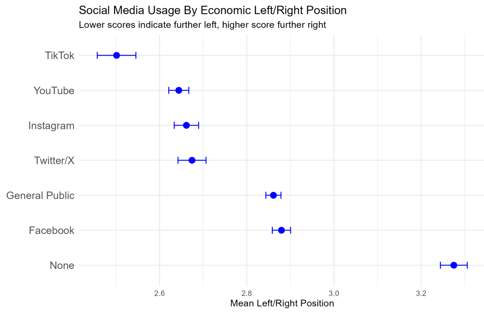
however, these facts hold even when comparing to users of other social media platforms. With the exception of age, where TikTok and Instagram users are younger on average, Twitter users have the highest proportion of males to females of any other site (57.4% male), are the most highly educated (16.4% postgrads; 40% undergrads), and have the greatest proportion of users from the highest two social grades (12.7% Grade A; 17.5% Grade B). Even for social media users, those who use Twitter in the UK are disproportionately drawn from a higher social strata. This expectedly translates into notable discrepancies along political dimensions between user groups, where Twitter users pay the greatest levels of attention to politics (48.7% pay a lot of attention) and are the most likely to turnout to vote (85%) than any other social media user group, non-users, or the general public as a whole (See **Figure 2.4**).



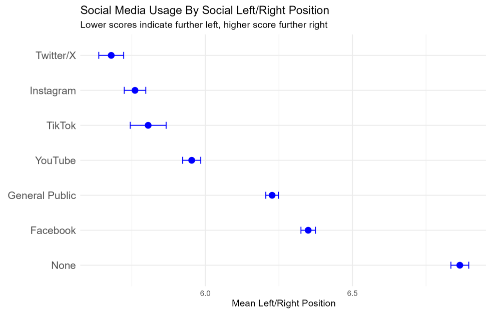
(A) **Political Attention**



(B) **Likelihood to Vote**



(C) **Economic Left/Right**



(D) **Social Left/Right**

FIGURE 2.4: Political behaviour of different social media user groups in the UK as well as the general public and social media non-users. Plots A and B illustrate stacked bar charts of political attention and likelihood t vote and plots C and D show mean points with standard error bars for economic and social (lib-auth) left/right positions. Plot A and B y-axis is ranked in descending order by % Pay a Lot of Attention and % Likely to Vote, while Plot C and D are ranked in ascending order by mean score. Wave 28 of the BESIP, n = 31,582.

Ideologically, TikTok users are the furthest to the left economically, followed by YouTube, Instagram and Twitter users who all occupy a similar ideological space. Facebook users are approximately in line with the general public economically, while social media non-users are noticeably further to the right. Along an authoritarian-liberal scale, Twitter users are more socially liberal than any other group reaffirming the site's liberal bias (at least in the case of the UK) while Facebook and non-users are actually to the right of the general public. A headline examination of global Twitter usage in 2024 highlights that these biases are not idiosyncratic to the UK. Twitter users worldwide are disproportionately members of more socially privileged groups: on average, they are wealthier and more highly educated, more likely to be male, and typically more engaged with the news and politics (Robertson, 2023).

2.2 Twitter: Town Square or Private Plaza?

Thus, it is fair to say that Twitter as a digital space is largely weighted towards society's more educated and socio-politically engaged populations. Users are in the minority in every country and the overall number of Twitter users, while substantial (circa. 590 million active users worldwide), pales in comparison to global social media giants like Facebook (over 3 billion), YouTube (2.53 billion), Instagram (2 billion) and TikTok (1.69 billion) (Kemp, 2024). Despite its relatively modest size, it has disproportionately appealed to more elite members of society, fostering an air of exclusivity about its membership. When social networking sites began to sprout on the Internet in the mid-2000s, they were initially welcomed by *"a fresh wave of technological optimism"* (Loader and Mercea, 2011) owing to their perceived potential to enhance democracy and provide open and accessible arenas for citizen-led deliberation. It was suggested that they would afford the general public new avenues to express their democratic will in a space where citizen-led or citizen-participated discussion could guide and mould the general conception of knowledge around a particular topic free from elite intervention (Christensen, 2011a; Jackson and Foucault Welles, 2016). Optimistic researchers believed that these new platforms had signalled the coming of age of non-elite actors to the forefront of public discussion, primarily in their ability to frame political narratives (Hermida, 2015; Meraz and Papacharissi, 2013). When attempting to conceptualise these emerging digital platforms, early scholars in the space leant heavily into Habermas' (1989) seminal theory of the public sphere as a logical starting point (See: Bruns and Highfield, 2015; Çela, 2015; Fuchs, 2015; Kruse et al., 2018; Papacharissi, 2008, 2010).

2.2.1 The Habermasian Public Sphere

The general guiding principle of Habermasian theory is that of a *"society engaged in critical public debate"* (Habermas, 1989, p.52). The idea, in other words, is that the public sphere should be a place in which citizens can engage freely and inclusively in public debate to a degree that a general sense of public opinion can be formed. This was considered by Habermas to be a significant departure from the more hegemonic social structures of the past where the majority of the population were excluded from public debate and governing policies were not informed by public sentiment. At the time, Habermas' theory referred to the emergence of a new public domain in the 18th Century that separated from ruling authorities and opened up new opportunities for public discussion and debate; owing in large part to advances in public education, wider accessibility of literature, and revolutionary changes in critical journalism. In the 21st Century, scholars saw an opportunity to reapply this theory to a newly emerging public sphere: the digital realm. It was alleged that the Internet and social networking sites had opened up new digital spaces for public discussion and engagement to take place and for new forms of public opinion to develop. Central to this was the theoretical deconstruction of the geographical and sociological boundaries that may have once restricted a common public space. As a result, general members of the public would be able to switch from conventionally passive consumers of new ideas and information to active participants in their creation and dissemination.

However, foundational to these claims is the upholding of equal accessibility and an ability to distribute information free from outside interference (Fuchs, 2012; Loader and Mercea, 2011). These would be necessary conditions to meet the Habermasian criteria for a legitimate public sphere: general accessibility to all and the elimination of all privileges (Habermas, 1974, p.51). When Twitter as a platform first emerged on the Internet in 2006 and explosively grew in 2009, it made a unique contribution to the social media universe as an asymmetrical microblogging service (Gruzd et al., 2011). The non-reciprocal nature of Twitter coupled with the use of short (originally 140 character-long) messages in the form of "tweets" promoted a far more top-down hierarchical network structure and flow of information than other sites. The hyper-politicised nature of Twitter relative to other platforms has meant that the promise of a genuine digital public sphere has hinged on this site more than most over the last decade. When acquiring Twitter back in October 2022, Elon Musk said this:

> *"The reason I acquired Twitter is because it is important to the future of civilization to have a common digital town square, where a wide range of beliefs can be debated in a healthy manner, without resorting to violence. There is currently great danger that social media will splinter into far right wing and far left wing echo chambers that generate more hate and divide our society"*
>
> **Embedded in a tweet by Elon Musk, October 2022 (Hern, 2022)**

If one evaluates the platform against Habermas' criteria – equal access and no elite dominance – the reality of Twitter as a genuine town square/public sphere is hard to sell. Twitter, in principle, has very little in the way of explicit restrictions in access or usage. Aside from the handful of countries where Twitter has been officially banned by the national government, anyone in theory can sign-up for a Twitter account and post content (as long as it does not violate their Terms of Service).[3]

However, as has been established, access to usage, while it may not be formally restricted, is unequally weighted in favour of socially privileged groups. Those participating in the "town square" are not representative of the population in any country and, as such, the formulation of public opinion is skewed ideologically to one side. Whether these restrictions arise because of self-selection effects or digital divides, Habermas' first principle of equal access is violated nonetheless. Additionally, equal access does not only pertain to presence on the site, but also the ability to participate. Research has shown that not only is selection bias a problem when analysing Twitter data, but so is participation (Yang et al., 2022). Even before the acquisition of Twitter by Musk – where the implementation of paid subscriptions in the form of Twitter Blue (now X Premium) can algorithmically enhance user visibility – imbalances in user participation and visibility were staggering.[4] A study conducted by Pew Research Center found that 80% of all tweets from United States (US) accounts were published by only 10% of users (Wojcik and Hughes, 2019). More recent analysis by Pew Research illustrated even more stark results, that the top 25% of US tweeters were responsible for an incredible 97% of all US tweets published on the site (McClain et al., 2021). This handful of highly active and highly engaged users drive the overwhelming majority of discussion and debate on the platform whilst the bottom 90% of users are passive consumers, publishing a median of only 2 tweets per month (Wojcik and Hughes, 2019).

Twitter is dominated by a small subset of so-called "power users" and when one couples this with the fact that the overall Twitter user base is already biased towards a small and privileged subset of national populations, it presents a starkly different reality to a free and open town square. More recent work by Habermas (2006) has been openly critical of the asymmetrical dynamic of traditional political communication in modern society. They argue that it places the political and media elites and the intelligentsia at the core of political discussion and pushes the voices of ordinary members of the general public (non-elites) to the periphery. It has been found that public opinion can suffer from an 'elite bias' where ordinary members of the public can conform readily to the policy views and sentiment of elite actors in policy fields they know little about (Bullock, 2011; Gabel and Scheve, 2007). Social media platforms like Twitter had presented cause for optimism because of their promise for decentralised arenas of communication that could flatten these asymmetries in informational power. However, research has shown that these

[3]As of writing this thesis, Twitter is formally banned in nine countries: China, Iran, Myanmar, North Korea, Pakistan, Russia, Tanzania, Turkmenistan and Venezuela. It has also been previously banned in Brazil, Egypt, Nigeria, Turkey and Uzbekistan.

[4]More information about Twitter Blue/X Premium can be found here: `https://help.x.com/en/using-x/x-premium#tbranking`

asymmetries persist on Twitter, where elite actors demonstrate continued dominance, although the decentralised and anarchic nature of Twitter has ensured that no absolute centres of communication have emerged (Tong and Zuo, 2018).

In its earlier formation, Twitter did demonstrate the potential for something approximating a town square to emerge and for a greater degree of communicative power to be handed to non-elites. One of the best examples was the early use of Twitter hashtags. Hashtags provided an explicit way for non-elite users to generate tags that could simplify ideas and symbolise movements, from #BlackLivesMatter and #MeToo to #jesuischarlie and #notinmyname (Bruns and Burgess, 2015). These hashtags could create 'sub-publics' on Twitter where trending topics pertaining to a particular idea or narrative could reverberate across other social media platforms and propagate exponentially outwards to users all around the world. Twitter is famously (and disputably) credited for helping to facilitate collective action during the Arab Spring in the early 2010s (Howard and Hussain, 2013; Wolfsfeld et al., 2013), as well as the Occupy Wall Street protest movement in 2011 (Gleason, 2013; Theocharis et al., 2015; Tremayne, 2016) and the Euromaidan uprising in Ukraine in 2013 (Bohdanova, 2014; Surzhko-Harned and Zahuranec, 2017). These so-called "Twitter Revolutions" gave non-elites the capacity to overcome traditional barriers to protest such as geographical location and physical copresence. More recently (and more negatively), Twitter has also been credited for its role in the January 6th Capitol attack in the US in 2021 and the anti-immigration riots in the UK in 2024. In both cases, it has been suggested that Twitter provided fertile ground for the proliferation and diffusion of conspiracy theories and misinformation which was attributed to fueling offline violence and coordinated action (CCDH, 2024; Griffin, 2024; Vishnuprasad et al., 2024).[5]

2.2.2 Legitimising Twitter as a News Source

The inherently asymmetrical nature of network structures on Twitter (X can follow Y, but Y does not have to follow X) coupled with its microblogging style of communication inherently favours a top-down centralisation of information and likely explains why Twitter has become the platform of choice for most offline elites. In the first few years of Twitter's adolescence it was largely informal and anarchic, but the uptake of the site by major political and media actors over the last decade has served to legitimate the site and transform it into an authoritative source of information (Molyneux and McGregor, 2022). According to Laor (2022), approximately 75% of

[5] The weight of evidence in favour of social media as a direct and substantial facilitator of offline collective action relative to other traditional mediums is uncertain. The very concept of Twitter and other social media "revolutions" is widely disputed by scholars and, while it is generally accepted that social media platforms can help to facilitate easier communication and mobilisation of collective action, the overall significance of its impact is questionable (Christensen, 2011b).

journalists in Western society maintain an active Twitter account. A recent State of Journalism report found that 77% of journalists value Twitter higher than any other social media platform and the site ranked second in top destinations for journalists to source news, beaten only by online newspapers and magazines (Mercier, 2022). Politicians have similarly migrated to the platform over the last decade, where an average of 80% of parliamentarians across 26 countries in the English-speaking world were active on the site (Van Vliet et al., 2020). Data from this thesis itself shows that 91% of UK MPs in the House of Commons were active on the platform as of 2022. Likewise, Twitter has become a popular platform for many members of the intelligentsia, where scientists, academics and researchers have adopted the site as an effective tool for the dissemination of their research to the wider public (Bisbee et al., 2022; Chugh et al., 2021). Over the course of its lifecycle, Twitter has evolved into an integral element of the 24-hour media news cycle and a significant facilitator of the continuous circulation of information online (Nielsen and Schrøder, 2014).

Simultaneously, the number of people who are turning to social media platforms (primarily Facebook, Twitter, and YouTube) for news consumption rises every year. As of 2021, the number of US adults who use social media sites for news on a regular basis stands at just under half (48%) the population (Walker and Matsa, 2021). A comprehensive report by Ofcom in 2024 found analogous findings in the UK illustrating a "generational shift" in news consumption within the population, where the overwhelming majority (96%) of respondents stated they consumed news in some form but the ways in the which they consume it are changing dramatically (Ofcom, 2024a). As of 2024, seven in ten (71%) said they had consumed online news in some capacity which now levels with traditional TV news (70%), and more than half of all UK adults said they had consumed news through social media (52%). Of the major social media platforms, Facebook was the most used as a news source (30%) in 2024 followed by YouTube (19%), Instagram (18%) and subsequently Twitter (15%). In 2022, Twitter was the second most used social media for news consumption behind Facebook but has dropped over the last two years, where Instagram and, substantially, YouTube have both risen and overtaken it. Importantly, the way in which the British public consumes news cuts asymmetrically across different intersections of the population, chiefly age where older people (55+) still prefer established T.V news providers like BBC One and ITV1 whereas "digital natives" (16-24) prefer a plurality of online news sources.

The growing digitisation of news serves to further legitimate the use of sites like Facebook, Instagram and Twitter as not just social media platforms, but authoritative hubs of information.[6] However, with the transference of the offline authority and influence of major news outlets to the social media space, this can serve to violate

[6]Though it is worth noting that despite the growth of news consumption via Internet sources such as social media, these still rank lower in levels of trust, accuracy and impartiality than traditional news sources (Ofcom, 2024a, pp.10-11).

Habermas' second key tenet of a true public sphere: the absence of elite dominance. There is evidence to suggest that the movement of formal elite actors onto these platforms has played a part in legitimising these sites as genuine informational authorities (Molyneux and McGregor, 2022; Van Dijck et al., 2018). However, unlike traditional news sources which are fairly one-dimensional, social media sites provide a far greater opportunity for elite to non-elite engagement. The use of Twitter by politicians for constituency management, public engagement, and propaganda (Jackson and Lilleker, 2011; Tromble, 2018; Woolley and Howard, 2018), and by journalists for story procurement and to monitor public sentiment (Gil de Zúñiga et al., 2018; Kim et al., 2015; Larsson and Moe, 2012) is well documented. The unmediated opportunity to directly communicate with the general public increases the ability of these elite actors to shape information diffusion by deciding how they choose to emphasise (or de-emphasise) content (Fazekas et al., 2021, p.378). Consequently, whilst the "decentralised" nature of Twitter and other social media platforms may dilute the overall unified power of traditional media and political institutions to some degree, it does afford elite actors more individual autonomy over content creation and diffusion. This matters for the way in which we choose to conceptualise Twitter as a digital space and, specifically, for the purpose of this thesis. Taking all these considerations together – the fact that Twitter is not equally *representative*, is not evenly *participatory*, and is not free from *elite intervention* - the notion that Twitter resembles anything at all like a digital town square of open deliberation and civic discourse is fallible.[7]

2.3 Twitter: A Polarised Network?

These factors, it can be said, are mostly organic. As previously mentioned, Twitter, like many social media platforms, does not have any *explicit* barriers to entry for the majority of the global population where the site is not banned. Yet, for all the reasons discussed, it has not evolved into the digital town square that had once been promised. After accounting for the fact that Twitter's user base has an elite bias, participation is driven by a handful of power-users, and network structures are mostly asymmetrical, a final consideration must be made for the nature of information flows across the platform.

[7]This analysis largely reflects the platform Twitter before Elon Musk's acquisition in October 2022. Studies have shown that since then, hate speech and misinformation have increased substantially on the site which further weakens the notion of a digital town square (Hickey et al., 2023; Miller et al., 2023).

2.3.1 Online Echo Chambers and Filter Bubbles

Rarely can social media be discussed without reference to two commonly used buzzwords: “echo chambers” and “filter bubbles”. Or, as Axel Bruns (2024, p.64) describes them, *“the dumbest metaphor[s] on the Internet”*. Both these phrases have been popularised over the last decade – along with a host of other digital media metaphors like town squares, rabbit holes, frontiers, trolls, and digital natives (Farkas and Maloney, 2024) - and, although often used interchangeably, are in fact distinct concepts. An early definition of the term “echo chamber” stretches as far back as 2008: a *“bounded and enclosed group that magnifies the internal voices and insulates them from rebuttal.”* (Jamieson and Cappella, 2008, p.76). Commonly, this term has been confused with another concept known as an “epistemic bubble” which is a *”social epistemic structure in which some relevant voices have been excluded through omission”* (Nguyen, 2020, p.142).

In both cases, echo chambers and epistemic bubbles share an important trait: the omission of alternative voices or ideas. In such informational environments, this can generate false narratives among its members and has raised concerns about how this could fuel polarisation and extremism without adequate exposure to challenging or counter-attitudinal information. However, what primarily distinguishes an echo chamber from an epistemic bubble is intent. While epistemic bubbles can organically emerge as an ordinary consequence of social selection and community formation, echo chambers are arguably more insidious, where opposing ideas and alternative voices are actively suppressed and/or discredited. Members of echo chambers are systematically isolated from all outside epistemic sources and share an united distrust of all those outside of their group (Ibid. p.142). One could argue that everyone exists to some degree within an epistemic bubble, where natural processes of social selection mould our social networks and subsequently influence our understanding and exposure to knowledge of the world around us. Conversely, echo chambers are deliberately manufactured through manipulation of information flows and are seemingly driven by politicised or ideological intent. Nguyen likens the nature of echo chambers to some accounts of *”cult indoctrination”* (Ibid.p.142) and a number of studies have used the concept to explain the evolution of conspiracy networks like COVID anti-vaxxers (Di Marco et al., 2021; Jennings et al., 2021; Jiang et al., 2021), climate change denial (Jasny et al., 2015; Treen et al., 2020; Walter et al., 2018), and QAnon (Bleakley, 2023; Priniski et al., 2021). Due to the greater freedom afforded to Internet users to search out, curate, and block information for themselves, the concern over digital echo chambers has been especially prevalent over the last two decades.

The term “filter bubble”, however, has a distinctly different meaning. Specifically, the phrase relates to the filtration effects that arise from the use of personalisation technology online. The concept of epistemic bubbles and echo chambers refer broadly to social self-selection effects and how this can create siloed information flows. Eli Pariser originally defined a filter bubble as the synthetic creation of online echo chambers through the algorithmic mechanisms employed by online websites and social network platforms to filter and streamline the information users are exposed to (Pariser, 2011). It is well-known that search engines and social networking sites employ algorithms to tailor the content that users are exposed to based on personal information collected through a multitude of sources including a user's search history, platform usage, privacy settings, and online habits. In the specific case of Twitter, the site openly made the switch from a reverse-chronological timeline for users in 2016 to algorithmic content selection and research shows that it significantly distorted the information that users were exposed to (Bartley et al., 2021; Dujeancourt and Garz, 2023). Following US President Donald Trump's permanent suspension from the platform in 2021 for alleged incitement of violence, accusations of anti-conservative bias were levied against the platform by Republicans. However, an internal report by the platform actually found the opposite effect: mainstream parties and outlets on the political right enjoyed higher levels of "algorithmic amplification" than their left-wing counterparts in six out of the seven countries that were studied (Huszár et al., 2022). Concurrently, a report by researchers at NYU poured further water on the accusations of conservative censorship by social media companies in the wake of Trump's exclusion from the platform, finding no reliable evidence at all to support the claim and likened it to a form of disinformation (Barrett and Sims, 2021).

2.3.2 Myths vs. Reality

The ongoing debate around online echo chambers and filter bubbles have consistently dominated the conversation around social media platforms and politics, but the significance of their impact is disputed by scholars. As early as 2013, Garrett definitively stated that the "*idea that we live in an era of political echo chambers, in which news consumers seek out likeminded partisans while systematically shielding themselves from other viewpoints, is both prevalent and wrong.*" (Garrett, 2013, p.248). More recently, Axel Bruns has been perhaps the most outspoken critic of the commonly used terms. Across multiple literature, they attempt to provide a more realistic account of the existence of echo chambers and filter bubbles online: they cite the scant evidence for the case of well-developed, exclusive echo chambers or filter bubbles on social media platforms (Bruns, 2017; Bruns and Enli, 2018), criticise the lack of clarity and conceptual vagueness around the terms (Bruns, 2021) and accuse purveyors of the term with sparking unfounded moral panic and scapegoating technologies for a much wider societal problem (Bruns, 2019, 2021). Accordingly, a number of studies have shown that the use of social media does not only not reduce the amount of counter-attitudinal information users are exposed to but can in fact increase incidental exposure to news and opinions from the opposite side of the political spectrum (Flaxman et al., 2016;

Fletcher and Nielsen, 2018; Weeks et al., 2017). Gentzkow and Shapiro (2011) similarly found that internet news consumers with ideologically narrow news diets are rare and that most users are exposed to a relatively diverse spread of opposing information online. Similar findings were noted in multiple other studies which tracked the informational networks and digital data of social media and internet users, highlighting a significant diversity within the news diets of participants online and a reasonably strong degree of audience overlap between news outlets (Fletcher et al., 2023; Guess, 2016; Nelson and Webster, 2017; Webster and Ksiazek, 2012).

A study conducted by Guess (2021) which examined US online media consumption found that most people on the political spectrum have relatively moderate media diets. This is with the exception of a small number of partisans who drive a disproportionate amount of traffic to ideologically slanted websites. A cross-national study of Internet users in the US and six countries in Europe found evidence to counter claims of both online filter bubbles and echo chambers: Internet users routinely expose themselves to a plurality of opinions and viewpoints through a diversity of media and a sizeable number admit to often reading news they disagree with and less than 20% had blocked/unfriended someone with differing political views (Dutton et al., 2017). Most interestingly, there is evidence to suggest that exposure to counter-attitudinal information on social media does not necessarily guarantee a softening of an individual's sentiment to the alternative point-of-view and may even have the opposite effect. A study by Bail et al. (2018) showed that after routine exposure of tweets published by elite political actors from the opposing party in the US, Republican participants had developed more conservative views post-treatment while Democrat voters grew more liberal. It is important to note that much of an Internet user's media diet is moderated by their level of political interest and their partisan preferences (Boulianne, 2011; Strömbäck et al., 2013; Stroud, 2008) and that, while the prevalence of online echo chambers appears to be overstated, those with a combination of low political interest *and* low media diversity have a higher likelihood of finding themselves trapped inside one (Dubois and Blank, 2018).[8] Those with higher levels of political interest are less likely to avoid conflicting and counter-attitudinal information given that they see a greater value in being exposed to such ideas (Knobloch-Westerwick and Kleinman, 2012).

2.3.3 Political Polarisation on Twitter

The diverse and high-choice media environment available to Internet users helps to prevent against the evolution of embedded online echo chambers, where most users

[8] Dubois and Blank's study found that of their respondents, 8% reported a media diversity score of less than 10 (out of a possible 48) as well as being disinterested in politics. They theorise that while this small segment of the population might be more vulnerable to an echo chamber, they may still be able to escape them through more politically engaged family and friends (Dubois and Blank, 2018, pp.741-742)

are routinely exposed to a plurality of information. However, as Dubois and Blank (2018, pp.740-741) also state, this specifically relates to users across *multiple* online media platforms where singular social media sites are only a small part of the overall digital media universe. They note that political information can be sourced through a variety of media channels both on and offline and a social media site like Twitter or Facebook will likely be only a fragment of a user's overall media diet. Referring to Twitter specifically, they say:

> *"It seems likely that networks on Twitter are polarized ... and networks on other social media may be equally polarized. But social media are only part of the environment, and they are the least trusted part. ... Twitter may be a place where individuals talk to people with the same political opinions. But a study of Twitter says little about the political information one is exposed to when they watch CNN or BBC news, or visit the Economist website or the Washington Post. These are places where individuals may be exposed to a wider variety of information and political views."* (p.741)

If one considers the fact that Twitter users are, on average, more highly educated and politically engaged than the general population, it is unlikely that they would consume their political information from only one (typically less trusted) source. Thus, the idea that Twitter users widely exist in echo chambers is questionable if one accounts for their media diet across the board. Data from Wave 28 of the BESIP would largely confirm this for British Twitter users at the very least, where Twitter users, on average, get their political information from a more diverse range of both on and offline sources. Between television, radio, newspapers (incl. online) and the Internet, Twitter users derive information from a mean 2.28 of them, significantly higher than any other social media user group or the general public. (See **Figure 2.5**)

With that in mind, it is difficult to suggest that Twitter users are widely trapped in informational or ideological echo chambers, certainly in the case of British users. However, as Dubois and Blank noted, this only holds when taking a multi-source approach to media diversity but does not say much for each platform individually. While it is unlikely that Twitter users who do not rely solely on the platform for their political information exist inside an echo chamber, this does not mean that they do not exist on the site itself. In its relative infancy in 2011, Conover and colleagues were one of the first to assess political polarisation on Twitter and the reality of its networked public sphere. Their study of over 250,000 (re)tweets from the six weeks leading up to the 2010 US congressional midterm elections showed that the network of political retweets exhibited a highly partisan structure, with extremely limited connectivity between left- and right-leaning users (Conover et al., 2011). A following paper by Colleoni and colleagues in 2014 looked to assess the same concept and found analogous results: within Twitter networks of US accounts identified as Republican or

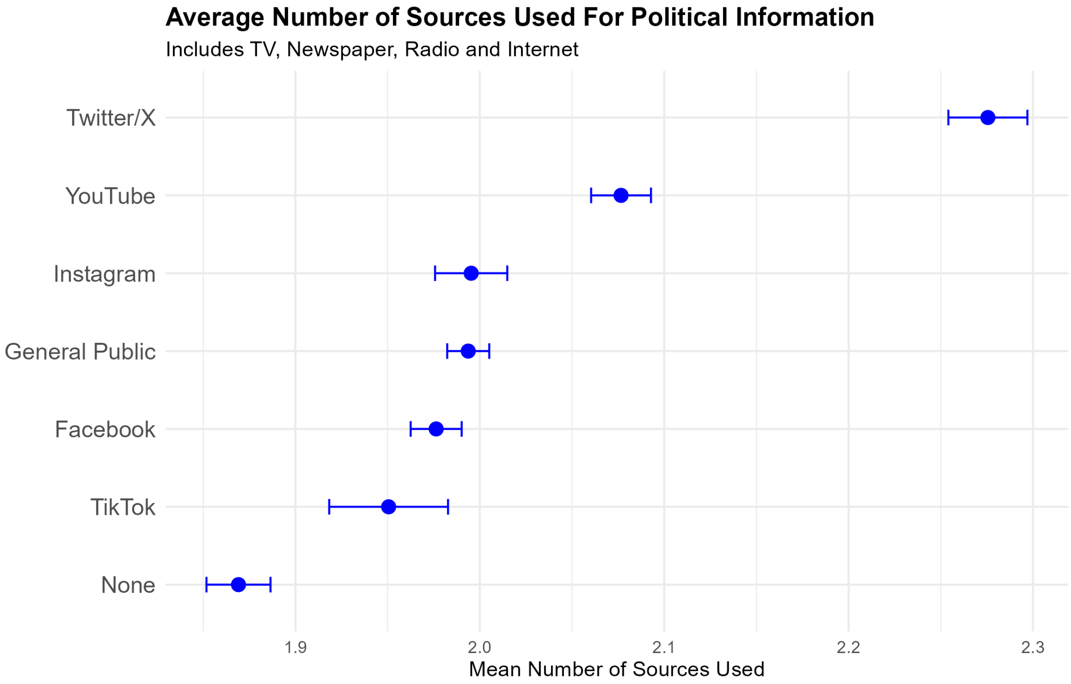


FIGURE 2.5: Mean number of sources used for political information by respondents in Wave 28 of the BESIP per social media user group. Options range from 0 to 4 including None, T.V, Newspaper (inc. online), Radio and the Internet. n = 31,582

Democrat, political homophily (the tendency to connect with others more politically similar to themselves (McPherson et al., 2001)) was extremely high (Colleoni et al., 2014).

Loosely defined ideas around echo chambers and online polarisation can likely be better understood through more traditional theories within media and communication studies. In particular, Klapper's seminal theory of selective exposure moves users away from being passive consumers of media towards active selectors of the information they wish to consume and from whom (Klapper, 1960). Conventionally, it is posited that this phenomenon is driven by a subconscious desire to reduce cognitive dissonance (Festinger, 1957; Sears and Freedman, 1967) - the psychological stress induced by holding two or more conflicting ideas or beliefs - which can generate a form of confirmation bias: where individuals selectively search for information that only serves to reinforce or confirm what they already believe to be true (Nickerson, 1998). In other words, it is typical for people to gravitate towards information that aligns with their pre-existing beliefs or values to reduce their cognitive discomfort. This is not a uniquely online phenomenon and fuels much of the criticism towards social media naysayers. Individuals will always prefer information that they agree with and this will be reflected in their media consumption choices, whether online or otherwise.

The importance of quality and credibility is also highly influential in how individuals select their news, where individuals will gravitate towards information that they deem to be more credible or of a higher quality (Fischer et al., 2008; Kahan et al., 2010; Metzger et al., 2020). However, it is typical that an individual's perception of credibility tends to be highly correlated with pro-attitudinal information (Clark and Evans, 2014; Schweiger and Cress, 2019). When applied to the congeniality of politically motivated information, this form of selective exposure shifts to what can be described as partisan exposure: individual's will choose their political media in accordance with their political predispositions, assuming a reasonable degree of political interest and knowledge (Stroud, 2008) In the US, a recent study conducted by MIT found that Twitter users were over three times more likely to follow other Twitter accounts that aligned with themselves politically, emphasising the influence of shared partisanship on social tie formation on social media (Mosleh et al., 2021).

A number of studies have reaffirmed the existence of homophilic network structures on Twitter and the tendency for partisan clusters to form around both followership and content sharing (Bail et al., 2018; Darius, 2022; Garimella and Weber, 2017; Hong and Kim, 2016; Wojcieszak et al., 2021); though some scholars have found that the overall extent of ideological segregation ("polarisation") may have been overstated and that users are more willing to engage in cross-ideological dissemination than had been previously suggested (Barberá et al., 2015; Gruzd and Roy, 2014; Shore et al., 2018). Urman (2020) speculated that the uncertainty over the degree of polarisation on Twitter was largely because of differences in local contexts and the fact that each study tended to focus on single country perspectives. Studying the political Twittersphere of 16 different countries together, they were able to establish varying degrees of polarisation from nation to nation based on the degree of audience duplication (overlapping of followers) between the major political parties of each country. Their study illustrated that Twitter polarisation can vary substantially between nations of users and this is reflective to a large degree of the level of polarisation within each country's own political context. The study also showed that the US political Twittersphere was the most intensely polarised (along with South Korea and Jamaica) and raised questions about the generalisability of Twitter polarisation studies where the majority are focused on a US context. Specific research on Twitter polarisation in the context of British users is relatively scant and almost exclusively centred on Brexit, where a handful of studies have shown ideological segregation of users between Leave and Remain identities (Bastos et al., 2018; Bouko and Garcia, 2020; Grčar et al., 2017; Mora-Cantallops et al., 2021; North et al., 2021; Wellings et al., 2024). Though, in accordance with Urman (2020) topology, British Twitter was categorised as polarised with limited audience duplication between the two major parties (Conservative and Labour).[9]

[9]Though the study did show that the Liberal Democrats (Britain's traditional "3rd party") play an important bridging role between the two as well as the Green Party, reducing the degree of segregation between users.

2.4 Harnessing Twitter's Eliteness for Political Research

The previous section sought to outline the reality of Twitter as an online platform and digital space, specifically in the case of the UK, and what this means in regard to the composition of its user base, participation by its users, the structure of its networks, and flows of information. This section will now consider how such a platform can (and has) been used for the benefit of political research. At the same time that offline political and media elites began to realise the potential of a site like Twitter for information diffusion, particularly in an increasingly digitised media environment, social researchers began to recognise its value as a data source. Rarely had there been an observational space in which both elite and non-elite actors existed in tandem, and their behaviour could be directly extracted and measured via what became commonly known as "digital trace data" (DTD) (Howison et al., 2011). These records of digital activity could be extracted in the millions and the potential to harness this social media "big data" posed a new frontier for social science research.

2.4.1 Using Social Media Data For Research

Over the years, DTD has been extracted from social media sites like Twitter, Facebook, Reddit, Instagram, YouTube and Telegram for a swathe of purposes including health research (Korda and Itani, 2013; Paul and Dredze, 2011; Reece et al., 2017), disaster response (Houston et al., 2015; Gao et al., 2011), sports prediction (Kampakis and Adamides, 2014; Wakefield, 2016), tracking criminal networks (Andrews et al., 2018; Drury et al., 2022; Resende de Mendonça et al., 2020), monitoring financial markets (Almog and Shmueli, 2019; Jin et al., 2017), and for studying online extremism (Benigni et al., 2017; Ferrara et al., 2016; Urman and Katz, 2022). However, while social media platforms like Twitter have offered researchers access to "naturalistic" observational data on an unprecedentedly large scale, failing to adequately acknowledge the implicit biases that tend to exist within them risks severely undermining the research. In the context of political analysis, for example, a study by Burnap et al. (2016) made a genuine attempt at forecasting the outcome of the 2015 UK General Election by using the summed sentiment of tweets about each party and party leader as proxies for overall vote share. The study's projections were predictably inaccurate with results akin to the polls: a hung parliament with Labour holding the majority of seats. Without adequately adjusting in some way for the selection and participation biases inherent within British Twitter's user population, studies like this will fall foul of similar errors to the pollsters in 2015 General Election: over-sampling left-wing voters and under-sampling the right.

Since political researchers first began to recognise the informative value of social media "big data", a plethora of studies have looked to harness it for election forecasting and gauging public opinion (see, for instance, Caldarelli et al., 2014; Larsson and Moe, 2012; McKelvey et al., 2016; Sang and Bos, 2012; Skoric et al., 2012; Tumasjan et al., 2010). A systematic review of the literature on predicting elections using social media data by Brito et al. (2021) showed that the earliest published attempt to do so by Tumasjan et al. (2010) remains the most commonly used method (volume and sentiment of tweets as proxies for vote share) but with consistently low success rates. An umbrella review of the nine surveys which assessed the state-of-the-art on election prediction using social media between 2013 to 2021 indicated that the majority of research on Twitter faces a host of methodological and technical limitations and made election prediction no better than chance (Brito et al., 2024). The review suggests that without a diversification of input data sources (both across different social media platforms and offline) and a movement towards using social media data as part of a statistical model instead of directly for polling, researchers will continue to repeat the straw poll errors performed before 1936 (Crossley, 1937).

2.4.2 Exploiting Twitter's Unrepresentativeness

So, at the very least in the case of political analysis, one might question why political researchers would use social media as a data source at all given its myriad of drawbacks. Firstly, one needs to account for a number of technical limitations associated with data extraction from social media (or any DTD generally) – following Bail (nd):

1) **Inaccessibility:** a lot of DTD is private or access is restricted by the platforms themselves;

2) **Drifting:** platforms shift in their popularity and usage over time and between user groups;

3) **Algorithmic Confounding:** observed human behaviour is not necessarily "natural" if it changes in response to platform algorithms;

4) **Unstructured:** DTD is typically large and unstructured ;

5) **Sensitivity:** a lot of DTD is often sensitive;

6) **Incomplete:** DTD is often incomplete due to deletion, redaction or removal;

7) **Platform "Black Boxes":** data quality assurance it difficult to establish if platform data governance is not clear.

Notwithstanding the technical issues associated with extracting and using data from social media platforms, one has already established that the study of British Twitter is fraught with representational constraints: its general user base is widely unrepresentative of the general population, content is disproportionately produced by a small subset of power-users, elite bias is considerable, and networks and information flows are ideologically segregated. An additional bias yet to be accounted for is also positivity bias: the idea that communication on social networking sites is more performative and less likely to be authentic (Reinecke and Trepte, 2014; Schreurs and Vandenbosch, 2021). Fortunately, rather than treating the unrepresentativeness of Britain's Twitter population as a limitation to be adjusted for, this thesis looks to *directly harness* it.

This idea was articulated directly in seminal work by Pablo Barberá in 2015:

> *"The use of Twitter data presents ... advantages over other sources of information about preferences. First, the large number of active users on this social networking site can be exploited [for estimation], if we consider users as "experts" ... Twitter users are not a representative sample of the voting age population. ... Citizens who discuss politics on Twitter are more likely to be educated and politically interested, and that makes them a particularly useful source of information."* (Barberá, 2015, p.77)

In this way, if the aim of the research is not to directly generalise what is studied on the platform to the "real world", then Twitter offers an extremely powerful source of information, where the behaviour of millions of highly educated and politically informed users can be crowdsourced for a multitude of purposes.

2.4.2.1 Ideal Point Estimation

The central aim of this thesis is to estimate an ideological map of British political and media elites by exploiting the ideological informativeness of Twitter networks. The platform offers two distinct advantages compared to other potential sources of information: 1) a heavy presence of official British political, media and cultural elites on the site; and 2) a suitably polarised ideological structure to network and information flows which can be leveraged for ideology estimation. To do so relies on an existing theory which has profoundly shaped the study of elections and voter behaviour: *The Spatial Theory of Voting*. Classical spatial theory is originally rooted in the seminal works of Hotelling's (1929) *Stability in Competition*, Downs' (1957) *Economic Theory of Democracy* and Black's (1958) *Median Voter Theorem*. The essence of classical spatial theory is that voters make decisions through a comparison of their own preferences on issues/policies against the perceived positions of candidates or parties on those same issues/policies. In a spatial context, voters and candidates/parties are said to both have an "ideal point" somewhere in the latent space and rational voters will seek to minimise the distance between their ideal point and that of the candidate/party. Simply put, rational voters prefer candidates/parties whose ideal point is closest to theirs (Enelow and Hinich, 1984). It is common for people to conceive of ideology in spatial terms, where the words "left" and "right" are conventionally used to describe positions along the political spectrum. An individual's ideal point is an empirical representation of their position within that latent dimension.

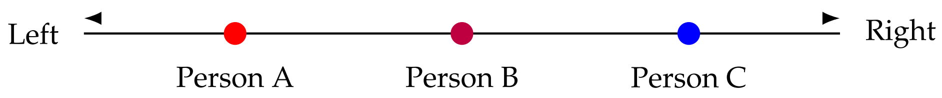


FIGURE 2.6: Simple one-dimensional ideal point spectrum

Although interrelated and often interchangeably used, the concepts of ideal points and ideology are distinct in the context of political science. Ideology is the much broader and widely known of the two; though, despite its common usage, is less well-defined and much more contested. The origin of the term ideology can be traced back to French philosopher Destutt de Tracy in the late 18$^{\text{th}}$ Century who first conceived of the notion as a *"science of ideas"*, a term to denote a new discipline dedicated to the study of ideas: *idèologie* (van Dijk, 2006, p.729). Broadly, it can be said that ideology encompasses a system of beliefs, values, and ideas that can shape an individual or group's worldview and approach to a set of political, economic, social and cultural issues (ibid.). Terms like "liberal", "socialist", "communist" and "progressive" might be commonly used to describe those who occupy spaces on the left of the spectrum, while terms like "conservative", "capitalist", "fascist", and "traditionalist" tend to be associated with those on the right. Terms like "moderate" and "centrist" are frequently used to describe those who occupy the middle-ground of the spectrum, while newer phrases such as "populist" and "neoliberal" have also become commonplace in modern political discourse. However, the major weakness with ideology as a term is the lack of a clear definition about exactly what it means, as well as the loaded nature of the word and its associated groupings.[10] Additionally, ideology tends to be culture-specific, where the meaning of the word "liberal" or "conservative" might embody entirely different sets of values and ideas depending on the context in which it is used, and by whom. In some cases, the term ideology even carries a negative connotation, where it has been used as a pejorative to discredit an individual or group.[11] Adherents to an ideology – "ideologues" – can be seen as rigid, uncompromising and dogmatic in their views and, as recent examples in the UK, was a term used to criticise the polices of major party leaders Jeremy Corbyn (Lab: 2015-2019) and Liz Truss (Con 2022).

[10]Many social scientists have tried to coalesce around a single definition of ideology but a consensus has not been reached. The debate is well explored in Eagleton (1991).

[11]Marx-Engels conceived of ideology as a form of *"false consciousness"*: that there is an objective truth, and then there are ideologies which are rigid, partisan, and misguided (van Dijk, 2006, p.728)

Furthermore, there remains a debate over the exact nature of preference formation among the general public and how well ordinary people conceive of a vague concept like ideology. Seminal work by Converse (1964) laid the foundation for a modern understanding of mass opinion, generally concluding that ordinary members of the public are not ideologically coherent. Citizens tend to lack knowledge on "what goes with what" and rarely hold stable policy views that could be connected together into a wider ideological view (Freeder et al., 2019; Yeung and Quek, 2024; Zaller, 1992). Zaller's work in particular acknowledged the importance of competing considerations, where an individual's position on an issue is a function of multiple factors which does include personal values, but is primarily influenced by exposure to elite-driven communication and mediated by factors such as political attention and issue salience (Zaller, 1992). Ideal points, on the other hand, do not subscribe so much to subjective characterisations of ideology and instead provide a much more agnostic measure of position. Spatial models of voting adopt a straightforward issue-based approach to understanding voter behaviour, assuming that the majority of electoral competition can be explained empirically by issue-preference.[12] Conventionally, these models focus on positional issues which invoke a natural degree of diversity (Enelow and Hinich, 1984): to what degree should the government intervene in the economy? Should the government allow more or less immigration into the country? Should drug laws be loosened or restricted? These issues do not necessarily have to combine into a grander ideology or worldview, but one's position on a set of issues can be measured as a single point along a particular policy dimension.

The spatial model of voting fundamentally relies on the choices of voters to act as observable signals about their latent ideal points. In this sense, ideal points can act as reasonable proxies for ideological position within a low-dimensional space, if one has a reasonable idea of what that dimension(s) pertains too. For example, if one estimates individual ideal points based on positions on issues relating to taxation, employment, inflation and public spending, one could reasonably suggest that this dimension is economic. If one estimates individual ideal points based on positions on issues relating to LGBTQ+ rights, abortion laws, social inequality and human rights, one might infer this to be a social dimension, and so on. Thus, the interpretation of one's ideal point is a reflection of the source of information from which individual preference has been derived. If one seeks to estimate the ideal points of legislators, the most obvious source of issue preference is roll-call data. Perhaps the most famous example of this is Poole & Rosenthal's DW-NOMINATE (1985; 1991; 1997), a multidimensional scaling technique which is used to estimate the spatial distances between legislators based on observed political choices. In this case, nominal vote choice where legislators are faced with a binary option (yea or nay) on a set of policies.

[12]Further work has expanded on the traditional spatial model of voting (proximity-based measures of issue/policy distance) to also evaluate individual choices based on valence (competency): it's not what parties/candidates stand for per se, but which parties/candidates are best placed to deliver (Stiers, 2022).

The model follows the same fundamental assumptions as traditional spatial modelling strategies: the alternative choices that voters face can be projected onto a low-dimensional latent space, all voters have an ideal point somewhere within this space, and when faced with a set of alternative choices (i.e: voting yea or nay on a particular policy), voters will opt for the choice which is most closest to their own ideal point within that space (Clinton et al., 2004). In terms of roll-call analysis, DW-NOMINATE was first applied to the historic voting records of members of the US Congress to generate spatial estimates of member's ideal points which can provide reasonable approximations of their ideological position in a unidimensional or multidimensional space.

2.4.3 The Westminster Case

This method of ideal point estimation using roll-call data works well in the US because of its atypically loose party cohesion and individualistic electoral culture (Bawn and Thies, 2003). This afford legislators a much higher level of freedom to vote independently from their party, and has also shown promise in other legislative bodies including European Parliament (Hix, 2001; Lo, 2018; Martin, 2021), the UN General Assembly (Binder and Payton, 2022), the Supreme Court (Cameron and Park, 2009), and in the national assemblies of other countries (Clerici, 2021; Rosenthal and Voeten, 2004).). Evidently, these models work best in environments where party influence is less dominant, optimising the informativeness of the roll-call data. However, in a Westminster-style legislative body like the UK House of Commons, parliamentary culture is characteristically rigid with a strict whipping system, tight party cohesion, and a "government vs. opposition" mentality (Spirling and McLean, 2007). Members tend to compete electorally on national policy platforms and voters (typically) select parties for government instead of local representatives (Bawn and Thies, 2003; Cox, 1987). In a system where parties have strong top-down agenda-setting powers and MPs toe the party line and are reluctant to rebel, obtaining unobtrusive measures of legislator position beyond their party position using roll-call votes is extremely difficult (Franklin and Tappin, 1977; Hix and Noury, 2010).[13]

2.4.3.1 Papers 1 and 3: Intra-Party Competition and Dyadic Representation

Paper 1 and **Paper 3** of this thesis both require estimates of MP left/right position for the purpose of testing intra-party competition and dyadic representation respectively. MPs do not represent their party or their constituents in isolation; their decision-making is often complex and can be influenced by multiple factors. In some

[13] In cases of extreme salience, MPs have shown themselves to be willing to break from the party whip (E.g. Iraq War, Brexit Deals, Same-Sex Marriage) but rebellions in the HoC are still extremely rare - less than 1% of all votes (Slapin et al., 2018, p.20).

cases these are relatively straightforward to measure - partisanship, faction membership, social demographics (Barrett and Cook, 1991; Miller and Stokes, 1963; **?**; Saalfeld, 2014) – and in other cases these can be more intricate and difficult to quantify: political ambition, access to knowledge, moral code, social connections, ideology, and so on (Benedetto and Hix, 2007; Herrick and Moore, 1993; Kam, 2001; Ringe et al., 2013). Trying to understand why legislators behave and vote the way that they do has been a subject of focus for many scholars in the field of political science for decades, and great strides have been made in developing increasingly more advanced models of legislator behaviour. However, some of what drives both intra- (*within*) and inter- (*between*) party competition is practically impossible to observe without access to the unique social dynamics that exist behind the scenes of legislative systems. How do political alliances form? What social processes help forge political adversaries? How much is legislator behaviour swayed by constituency opinion? In two-party majoritarian-style Westminster systems like the UK HoC, strong party loyalty tends to overwhelm these other variables. But, as studies have shown, within-party responsiveness can still be observed in these systems if one can find the right signals (Blumenau, 2021; Kellermann, 2012; Hanretty et al., 2017).

Twitter as an elite digital space with its highly polarised network structures and heavy presence of political figures can provide such data. **Paper 1** describes in more detail the way in which ideal points can be estimated using data extracted from this platform and how this can be used to compliment models that estimate intra-party competition. **Paper 3** outlines existing theory and literature around dyadic representation and expands on **Paper 1** to assess MP-to-constituency responsiveness, both of which are assessed in the context of a restrictive Westminster system.

2.4.3.2 Paper 2: Media Representation

Paper 2 makes a slightly different contribution to the thesis, departing from the study of legislative behaviour to examine political representation on public service broadcasters (PSB) in the UK. In the same way that the UK's characteristically strict legislative system makes it an interesting case study for assessing intra-party competition and dyadic representation, Britain also has a tightly regulated broadcasting regime. The UK has a long and proud of tradition of public service broadcasting and has historically legislated for an impartiality regime in television and radio broadcasting enforced by an independent regulator: *Ofcom* (Barnett and Petley, 2023, p.18). The emergence of two new ostensibly partisan news channels – GB News and TalkTV – over the last three years has placed increasing pressure on Ofcom's willingness and authority to legislate on alleged violations of impartiality. Despite a host of literature assessing media impartiality in the UK, they almost exclusively rely on thematic content analysis to assess media representation (Berry, 2016; Cushion et al., 2018; Cushion and Lewis, 2017; Hughes et al., 2023; Lewis and Cushion, 2019; Morani et al., 2022). Twitter has the unique benefit of a heavy presence of not only political elites but also media and cultural elites who regularly appear as discussants on the political programmes of UK T.V broadcasters. In the same vein as **Papers 1** and **3**, **Paper 2** outlines existing literature and methods for assessing media representation before describing the way in which Twitter can be leveraged as a data source to make a contribution to this space.

Within this comes the conception of the "Westminster bubble". Expanding out from just official Members of Parliament, this thesis estimates an ideal point map of the broader political commentariat in the UK – politicians, journalists, policy advisers, commentators, academics – many of whom have made Twitter their digital home over the last decade. The UK offers a unique case study for this research: 1) its legislative body is characteristically strict; 2) its media broadcasting system is tightly regulated; 3) the majority of its major political, media and cultural figures use Twitter; and 4) ordinary British Twitter users are disproportionately more educated and political engaged. The following section will now outline the data and methods used to do so.

Chapter 3

Data & Methods

The three empirical papers presented in the main body of this thesis strive to advance our understanding of three important elements of the UK's political and media system: intra-party competition, media representation, and dyadic representation. The strict nature of the UK's legislative and broadcasting environments makes for an interesting case study and Twitter provides an ample data source for circumventing these restrictions. This chapter will now outline broadly the data and methods used in the three papers to construct an ideological map of British elites. The first section details the key data sets used in the analysis as well as details about the sources from which they are derived and the methods for obtaining the data. The second section will discuss the advanced quantitative methods used to analyse the data, derive ideal point estimates, and to assess the three elements of interest. This chapter is comparatively brief as each empirical paper contains a comprehensive data and methods section specific to that paper and this section aims not to repeat this. Additionally, it is difficult to provide a too broad overview of the data and methods of this thesis given that there is notable variation from paper to paper. However, the three papers remain linked in their overlapping use of the same data sets and similarities in methods used. As far as possible, this chapter will provide additional information about these that were not provided in detail within the three separate papers.

3.1 Data Sources

This thesis relies on a number of different data sets to map the UK's ideological landscape. Specifically, it combines the use of two primary digital trace data sets extracted from Twitter with two primary survey data sets commissioned exclusively for this thesis, along with secondary survey data from the British Election Study Internet Panel, and a number of wider secondary data sources specific to each individual paper. The bulk of the analysis in this thesis depends heavily on the two digital data sets extracted from Twitter and the data and extraction methods will be outlined in the following sub-section.

3.1.1 Digital Data: Twitter

Two data sets were extracted from the Twitter platform for this thesis at two separate time points. Both these data sets centre on the active Twitter profiles of both sitting (2019-2024) and former (2015-2017; 2017-2019) UK MPs in the House of Commons. **Data Set 1** contains the complete follower network data of the sitting MPs *only* (n = 591) as of the 22/08/2022 and **Data Set 2** augments this data with the tweet timelines of both sitting *and* former MPs (n = 797) as of the 21/04/2023 stretching back to one year prior to the 2015 General Election (07/05/2014).[1] As of the time of data collection for both data sets, Twitter provided a well-supported Developer Platform which allowed researchers and developers access to Twitter's open, global, real-time and historical platform of data. This gave researchers and developers with approved accounts comprehensive access to a suite of tools and data to build applications, stream data, and conduct research on a large scale. For the purposes of this thesis, engagement with the Twitter Developer Platform was purely extractive, harvesting both current and historical data from the platform to conduct the necessary analysis. Interaction with, and extraction from, the Twitter Developer Platform was conducted through the platform's Application Programming Interface (API) which, in essence, is an intermediary gateway for developers to programmatically communicate with Twitter.

Twitter research for this thesis was conducted through the Twitter Academic Research product track. This product track no longer exists as of mid-2023 but, at the time, qualified researchers access to all v2 endpoints, a maximum extraction cap of 10 million tweets per month, and complete access to the entire historical archive of Twitter data which extends back to the first ever tweet post in March 2006.[2]. Interaction with the Twitter API was programmatically conducted via the *Python* programming language and simplified using the *tweepy* package. The Twitter API allowed for access to virtually all publicly available data on the site, including public tweets and their accompanying metadata, public user profiles and their accompanying metadata, as well as data on public lists, spaces, media, and polls.

[1]Although this is the overall scope of the data set, this thesis only makes use of data from MPs over the last parliamentary period (2019-2024). However, future research may wish to make use of the additional data covering the 2015 and 2017 parliaments also.

[2]Although the Academic Research product track no longer exists, a technical overview of what it offered can still be found here: `https://devcommunity.x.com/t/introducing-the-new-academic-research-product-track/148632`. A current overview of the Twitter API v.2 functionality, features and endpoints can be found here: `https://developer.x.com/en/docs/x-api/migrate/whats-new`

Private data could not be accessed, however, including tweets and user profiles that had been set to private, as well as any data that had been removed or deleted. One point of note here is the implicit bias that is introduced into the data set because of an inability to access private data and the omission of any deleted or removed data that is unknown to the researcher. This is a common concern when studying human behaviour using digital trace data and should be addressed here. In the background discussion of the potential errors that can arise when using DTD by Sen et al. (2021, pp.402-406), they identify a number of common issues:

1) DTD is *non-reactive* meaning that they exist independently from any research design. This can raise concerns over construct validity which is important for this research given the aim of deriving ideal points from Twitter DTD. This issue is directly addressed using external validation from both experts in **Paper 1** and the general public in **Paper 2.**

2) Measurement errors can arise from a failure to account for *platform affordances* (how platform-specific norms and technical features can influence human behaviour) and *platform coverage* (the mismatch between a platform's user population and the target population). These issues have been largely addressed in the previous chapter where the uniqueness of Twitter's affordances and user population actually strengthens its value in the specific nature of this research.

3) User and trace *selection errors* raise concerns in the data collection phase where researcher queries fail to capture all relevant data (or capture irrelevant data). The sampling frame used to query the data from the Twitter API for this thesis is explicit and restrictive, only gathering followers and tweet timelines of pre-defined accounts (UK MPs). This helps to significantly reduce the over-capture of irrelevant data where the only major concern is the inclusion of non-human followers (*bots*). This is explicitly addressed in **Paper 1.** The only other major concern in this regard is the omission of data due to deletion, privatisation, or removal. Followers of MPs may have been under-sampled because of the deletion or banning of accounts.

For MPs, aside from the obvious issue with not all MPs being present on the platform (as well as some wider elites that are of interest for **Paper 2**), there is a high probability that most of their tweet timelines are incomplete due to deletion. As is discussed in the literature review in **Paper 3,** MPs will use their digital profiles as a form of impression management (Goffman, 1959) and will thus likely curate their timelines professionally. Political communication is often tightly engineered and fine-tuned and examining what is left out or, more specifically in this case, what is retroactively deleted is impossible to access if it is not captured before deletion. In many cases, deleted tweets may be composed of typos and corrections, but in other cases they may be strategic removals which would be of significant research value (Meeks, 2018).

A tool developed by the Open Foundation called *Politwoops* addressed this issue by recording, storing and archiving tweets posted and erased by politicians across more than 30 countries since 2012. For a substantial period of time, this tool made a valuable contribution to holding politicians to account and for fact-checking misinformation but no longer works since changes to the Twitter API in 2023.[3] Attrition rates of tweet data is an issue for the representational aspects of a data set derived from the platform and a number of studies have discussed this in detail (Almuhimedi et al., 2013; Elmas, 2023; Küpfer, 2024; Meeks, 2018; Zubiaga, 2018). These studies generally concur that tweet persistence is *non-random* and that highly sensitive and controversial data sets are the most vulnerable to attrition due to platform moderation practices and are subsequently the hardest to replicate. In this sense, the tweet timelines of MPs should not be too susceptible to attrition due to the (mostly) professionalised nature of MP tweet content.

Research has also shown that there is a significant temporal bias to the persistence and reliability of Twitter data sets, where representativity increasingly fades over time and this raises particular concerns when using the Academic API historical archive where time periods further in the past are under-represented (Pfeffer et al., 2023; Zubiaga, 2018). It is uncertain exactly how much this affects the UK MP tweet timeline data extracted for this thesis. It is unlikely that the tweet data will have been substantially affected by platform moderation and, as for deletion through personal curation, research by Noonan (2022, p.36) showed that of the 730,530 tweets posted by UK MPs in 2018, 590,646 still remained in 2021 indicating a missingness of 19%. Among all parliamentarians across six Western European countries the rate was relatively similar, around 22% of tweets had been deleted. Importantly, the study also confirmed that MPs do not typically engage in mass deletion and the ones that do have usually left parliament. An attrition rate of approximately 1-in-5 is reasonably high, though this is expectedly not spread evenly across MPs: there is significant inequality in tweet volume and this is a major predictor of tweet deletion, where MPs who tweet more, delete more. Additionally, populist rhetoric was also a significant predictor of deletion which potentially aligns with the fact that more controversial tweets are more likely to be deleted (Noonan, 2022).

Overall, there is no direct way to know the exact degree of representativity of either data set; they are only a complete representation of the Twitter platform at the exact time point in which they were extracted. Moving forward with this caveat in mind, the following two sub-sections will now outline the data sets in more detail.

[3] Link to the Politwoops website is here: `https://openstate.eu/en/projects-tools-data/elections/politwoops-2/`. An explanation for its decommissing can be read here: `https://www.propublica.org/article/politwoops-deleted-tweets-twitter-politicians-musk`. It is unclear if a historical archive of deleted tweets by UK MPs can be made available.

3.1.1.1 Data Set 1: UK MP Follower Networks

Data Set 1 was extracted from the Twitter platform via the developer API over a three-and-a-half-week period commencing 22/08/2022 and completing on 15/09/2022. At the time of data collection, 591 of the 650 UK MPs sitting in the House of Commons had an active Twitter account (91%). Before harvesting each of the 591 MP's complete sets of followers, metadata was extracted for each profile which contains high-level metrics about their account such as profile name, profile description, verification status, location, number of followers, number of accounts following, and number of tweets posted. For this data collection, the *tweepy* function *get_user* was used with two parameters: *username* and *user_fields.* For the *username* field, each MP's unique Twitter profile handle was individually provided and the following user fields were requested:

["created_at","description","id","location","name","pinned_tweet_id", "profile_image_url","protected","public_metrics","url","username","verified","withheld"]

where *public_metrics* were further unnested to extract the following:

["followers_count","following_count","tweet_count","listed_count"]

Each individual set of profile metadata for each MP was iteratively added to a Python dictionary and subsequently concatenated into a single data frame with 591 rows and 15 columns. **Paper 1** contains summary statistics of profile metadata by political party. **Figure 3.1** illustrates the top twenty accounts with the greatest number of followers:

Unsurprisingly, former major party leaders Boris Johnson and Jeremy Corbyn lead the way in terms of followers, trailed by other major political figures including Theresa May, Keir Starmer, David Lammy, Ed Miliband and Rishi Sunak, all of whom have held the position of major party leader in the UK at one point or another (with the exception of David Lammy who currently holds the position of UK Foreign Secretary). **Figure 3.2**, however, also illustrates that the most followed accounts are not necessarily the biggest tweeters. Interestingly, while no SNP MPs feature in the top twenty most followed accounts, five appear in the top twenty biggest tweeters.

As for extracting follower networks, MP profile usernames were once again used as the lookup search string. Lists of followers were requesting using the *tweepy* function *get_users_followers* with four parameters: *username*, *max_results*, *user_fields*, and *pagination_token.* The *username* and *user_fields* parameters remain the same as for the profile metadata lookup, where user fields in this case relate to the profile metadata metrics requested for each *follower* account. The maximum number of follower accounts that can be extracted in one single API call using *get_users_followers* is 1,000 and the maximum is used for the *max_results* parameter. Followers are extracted for each MP in descending chronological order starting from most recent follower. After a

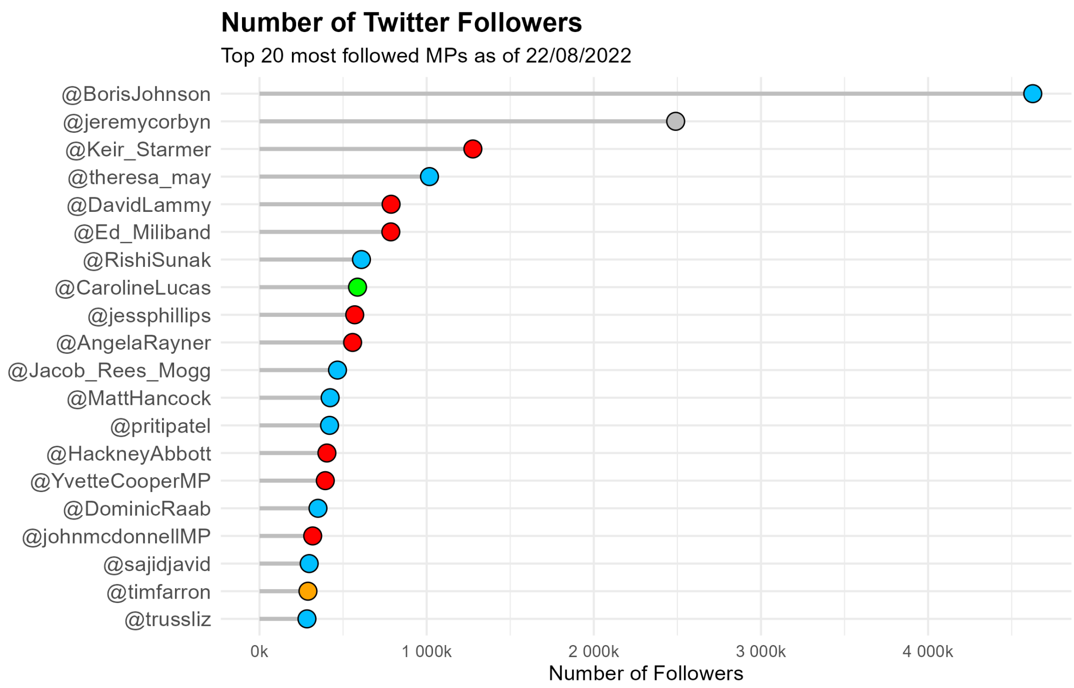


FIGURE 3.1: Lollipop chart illustrating the top 20 most followed MPs on Twitter as of the 22/08/2022.

single call is made, returning 1,000 accounts, if an MP has more remaining followers to be extracted, a pagination token is also returned which can be used to paginate through the next set of followers in the list until the entire list is exhausted. For follower network extraction, MP followers are iteratively extracted in batches of 1,000 followers at a time, updating the *pagination_token* field with the latest pagination token until all followers for each MP has been harvested. For every follower account, accompanying profile metadata is also returned providing high-level details about each profile. Follower data is stored in a simple two-column adjacency list, where the *source_node* column contains the Twitter username of a follower account and the *target_node* column contains the username of the followed MP.

Source Node		**Target Node**
$User_A$	$\rightarrow$	MP_1
$User_B$	$\rightarrow$	MP_2
$User_C$	$\rightarrow$	MP_3
$User_D$	$\rightarrow$	MP_4

TABLE 3.1: Example of a two-column network adjacency list.

These network adjacency lists can be converted to adjacency matrices when required for analysis, where each unique source node (the account from which the connection is created) form the rows and each unique target node (the account which is connected to) form the columns. These adjacency matrices can then be populated with integers to represent the existence and weight of connections between the two nodes. In unweighted networks such as Twitter follower connections - where connections can

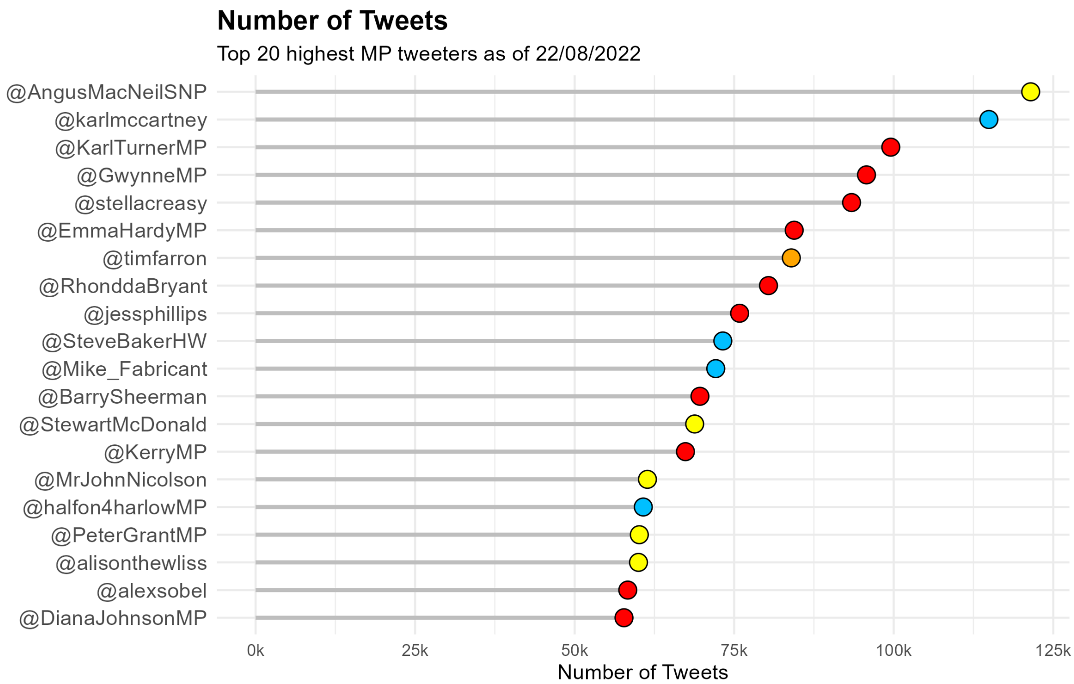


FIGURE 3.2: Lollipop chart illustrating the top 20 MPs with the highest number of tweets posted on Twitter as of 22/08/2022.

simply either exist or not exist - integers take a binary form of **0** or **1** depending on whether or not source user i follows target MP j.

	MP_1	MP_2	MP_3	MP_4
$User_A$	1	0	1	0
$User_B$	0	1	0	1
$User_C$	1	1	0	0
$User_D$	0	0	1	1

TABLE 3.2: Example of a two-dimensional network adjacency matrix.

Some networks can theoretically contain weighted connections where integers can take other forms than just **0** or **1**, giving some connections more weight than others. For example, a retweet network adjacency list can contain connections between users where user i may have retweeted user j which can also be weighted by how many times user i retweeted them.

Overall, this data contains a total combined number of followers m = 34,653,181 between the n = 591 MPs. Of these 34,653,181 followers, 11,071,104 were unique accounts. Accompanying profile metadata is also provided alongside these accounts for additional information about these users. Further details about this data set and how it is used for analysis is covered in the data and methods sections of **Paper 1** and **Paper 2.**

3.1.1.2 Data Set 2: UK MP Timelines

Data Set 2 was extracted from the Twitter platform via the developer API on 21/04/2023. Unlike when harvesting follower network data which had a strict API limit of only 15 calls every 15 minutes accounting for the three-and-a-half week long extraction period, the API rate limit for harvesting tweet data was far more liberal, allowing for 900 requests every 15 minutes. Additionally, the volume of tweets to be extracted was substantially smaller than followers, enabling all tweets to be harvested on the same day. An observational timeframe of just under nine years was decided, starting from 07/05/2014 and ending 21/04/2023. The start date of the 07/05/2014 was chosen because it is exactly one year prior to the date of the 2015 UK General Election (07/05/2015) and the 21/04/2023 is simply the date in which data collection took place. The *search_all_tweets* function was used to extract the MP tweet timelines with five parameter: *query, start_time, max_results, next_token, tweet_fields.* This function is strictly exclusive to the Academic API track and has the distinct benefit of allowing for timeline extraction as far back as required, whereas the conventional *get_timeline* function is capped at the 3,200 most recent tweets for a single profile.

The search query used to gather the tweets was straightforward and as follows:

"from:@mp_username"

This query returns every tweet posted by the specified username's account. The start time was set at:

"2014-05-07T00:00:00Z"

where no end time was specified, meaning all tweets were returned from the moment of extraction going back as far as the designated start time. The max number of tweets that can be returned in a single call is 500, which was the value used for the *max_results* parameter, and the *next_token* parameter is the pagination token used to iteratively return the next set of tweets in the chronological timeline in the same way as Data Set 1. Tweet timelines include retweets, quote tweets, and replies to other tweets. Finally, the *tweet_fields* requested were as follows:

["id","created_at","author_id","text","public_metrics","entities","attachments"]

where *public_metrics* were unnested to extract the following:

["retweet_count","reply_count","like_count","quote_count","impression_count"]

From the election of the new parliament following the 2015 General Election up to the date of data collection, 917 different MPs had sat in the House of Commons at one point or another. 797 (87%) of these had an active Twitter account at the time of extraction and their tweet timelines were extracted one by one and subsequently merged into a single data frame. Accompanying profile metadata was also extracted for each account prior to timeline extraction on the 09/04/2023. Overall, a combined total of 9,216,735 tweets were posted between the 797 MPs across the observational timeframe, where just over half are retweets (51%).

3.1.1.3 Ethical Considerations

Before moving on to discussing the survey data, it is important to provide a brief note on ethical considerations with using DTD. Along with the host of methodological challenges researchers face when using DTD as previously discussed, using DTD also brings a number of ethical issues. Ethical frameworks for working with social media data as researchers are scarce (Taylor and Pagliari, 2018), and a large reason for this is because working with DTD – social media data specifically – falls between primary and secondary analysis:

"It is primary data in the sense that there is no formal curational activity standing between the participant providing the data and the researcher who wishes to use it, yet secondary data because the data is gathered not from the participants but from an aggregating intermediary (Twitter)" – (Gold, 2022, p.5)

This can make the ethical guidelines around usage unclear. Consulting guidance by Gold (2022, p.3), they stress that the key primary ethical principle that is engaged with in relation to Twitter data specifically is *autonomy*: the right of participants to be treated as independent moral agents and to determine their own best interests. This coincides with important concepts around informed consent and participant expectations. For consent, a strict position might require a researcher to ascertain clear consent from each and every person who's Twitter data will be used. This is however extremely impractical and so in this case, informed consent to use such data is **implicit**. This leans into the debate around whether or not social media data is in fact public or private. On the one hand, Twitter makes it clear to users when they sign up to the platform of what will happen to their data in their Terms of Service, including its potential use for research purposes Gold (2022, p.6). On the other hand, there is refutation to the idea that this would constitute *informed* consent given that most users will not have read the Terms of Service (Townsend and Wallace, 2016, p.6), and that just because data is publicly accessible does not necessarily make it public data (Boyd and Crawford, 2012). Gold (2022, p.6) rectifies this conflict by suggesting that ethical consent is reasonable as long as two particular conditions are met: the data being used has *current* public disposition and the use of that data respects the agreement under which is was made available to the researcher.

A key tenet of informed consent rests on the ability of participants to withdraw their consent at any time. This can come into conflict with the first condition wherein if a user wishes to withdraw their consent at any point (i.e: delete their data) this can be easily done on the platform itself but the data would still exist in the data set that has been extracted. Additionally, given that condition two emphasises the importance of respecting the agreement under which data was made available to the researcher, the Developer API agreement states:

"You should be careful about using X data to derive or infer potentially sensitive characteristics about Twitter users. Never derive or infer, or store derived or inferred, information about a X user's: Health (including pregnancy), Negative financial status or condition, Political affiliation or beliefs, Racial or ethnic origin, or beliefs, Sex life or sexual orientation, Trade union membership, Alleged or actual commission of a crime" - Twitter Developer Terms of Service.[4]

Given that this thesis looks to estimate ideal points of user accounts, this could be classified as inferring sensitive information about each a user's political affiliation. However, the Developer agreement also goes on to say:

"Aggregate analysis of X content that does not store any personal data (for example, user IDs, usernames, and other identifiers) is permitted, provided that the analysis also complies with applicable laws and all parts of the Developer Agreement and Policy."

In accordance with the Developer API agreement and the ethics approval conditions agreed with the University of Southampton's ERGO Committee, the conditions for consent to use Twitter user data is met by ensuring that any and all analysis derived from this thesis is conducted at the **aggregate** level *unless* it pertains to an account that can be considered a public figure. The threshold for classification of an account as a public figure is if the account holds an offline position in society that can be reasonably considered a public role (i.e: politician, journalist, celebrity, etc.) or holds an online position of authority which can be reasonably considered as significant. This second threshold is established by two criteria: 1) if an account was verified, which under the previous Twitter verification criteria when the data was extracted meant the user was "active, notable, and authentic"; or 2) had a minimum of 30,000 followers based on the UK's Advertising Standards Authority 2019 ruling of celebrity influencers on social media.[5]

Any data that is published as part of this research is fully anonymised so that no accounts can be identified from the data set unless they are identified as a public figure account. Anonymisation is ensured by removing any identifiable information such as user IDs, usernames, profile names, location data, or profile descriptions. Where required, user IDs are pseudonymised to ensure that network and profile data

[4] Link here: `https://developer.x.com/en/developer-terms/more-on-restricted-use-cases`

[5] Link to Twitter details about verfication can be found here: `https://help.twitter.com/en/managing-your-account/about-x-verified-accounts`. The link to the ASA's ruling on celebrity influencers on social media can be found here: `https://www.theverge.com/2019/7/4/20682087/instagram-twitter-celebrity-30000-followersadvertising-standards-authority-uk`.

can still be merged. The only profile metadata of non-public figures that is made public is non-identifiable metrics such as follower and following counts, tweet counts, verification status, and the date of profile creation. No tweet data for ordinary accounts is used in this thesis, where only MP tweets are extracted and these can be considered public accounts.

3.1.2 Primary Survey Data

As part of this thesis, two original surveys were commissioned and conducted with the purpose of providing validation data for Paper's 1 and 2. Directly addressing the concern with construct validity and DTD as outlined in the opening section of this chapter, validation data sets were ascertained from two sources: experts and the general British public. The exact use of these data sets are made clear in the two papers respectively. The next two sub-sections will outline the details of how both surveys were conducted.

3.1.2.1 Expert Survey

A short survey was circulated to 133 experts in British politics and political science on the 08/11/2022. The survey was generated through Qualtrics and made accessible to participants via a link attached to an email that was distributed to each of them individually. All responses were anonymous. The experts chosen to participate in the survey were all academics with substantive research expertise in electoral, legislative, or parliamentary politics and/or public opinion with a specific focus on the UK. In order to source these academics and their contact information, the politics faculty websites of all UK universities who are either members of the Russell Group and/or ranked in the top 30 UK universities for politics in any of the Times Higher Education World University Rankings, the Guardian University Guide, or the Complete University Guide were used. Although this sampling method is slightly restrictive, it helped to narrow down the process as efficiently as possible and provide a straightforward way of sourcing an appropriate number of individuals who could be justifiably considered credible sources of expert opinion on British politics. Of the 133 experts contacted and invited to participate, 70 participants completed the survey in full indicating a 53% response rate. The survey asked participants to place a sub-sample of 30 UK MPs, 12 UK political parties, and 13 UK media organisations on an 11-point scale between 0-10 based on where they believed each MP/party/organisation sat on the left/right ideological spectrum. Following the same format and questioning as the BESIP, 0 represented the furthest left and 10 represented the furthest right on the spectrum. All responses were optional and a Don't Know option was also provided. A starting question also asked participants to rank their level of knowledge of British parliamentary politics on a 6-point likert scale. This question was purely exploratory and was not used to filter the responses in any way.

Of the 30 MPs, 13 were members of the Conservative Party, 13 were members of the Labour Party, 2 were members of the Liberal Democrats, and 1 was a member of the Green Party. The final MP chosen was an Independent but was formerly a member of the Labour Party. There was a deliberate attempt to balance the sample of MPs as much as possible across the left/right spectrum (based on personal knowledge) as well as between more well-known MPs such as party leaders and cabinet ministers and lesser-known backbenchers. The complete survey format distributed to participants as follows:

Q1: What would best describe your level of knowledge of British parliamentary politics?

[None, Know a limited amount, Know a moderate amount, Know a fair amount, Know a great deal, Expert]

Q2: In politics people sometimes talk of left and right. Where would you place each of these MPs on the following scale?

[0 - Left, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 - Right, Don't Know]

Q3: Where would you place each of these political parties on the following scale?

[0 - Left, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 - Right, Don't Know]

Q4: Where would you place each of these media organisations on the following scale?

[0 - Left, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 - Right, Don't Know]

This survey data is used in **Paper 1** of the thesis.

3.1.2.2 *YouGov* Survey

Following the same general structure as the expert survey, a survey of the general British public was conducted in collaboration with British market research company *YouGov*. A representative sample of 2,068 members of the adult (18+) British population were surveyed on 21/02/2024, asking them to place 6 UK political parties, 24 UK media organisations, and 30 individual media personalities on a [0-10] scale, where 0 represents the furthest left and 10 represents the furthest right. These questions were asked in the same way as the expert survey, with the additional response "Never heard of them" provided alongside "Don't know". Individual media personalities were primarily selected by largest number of Twitter followers, making sure to only include individuals who were British, regularly appear in the British media, and could reasonably be identified ideologically. Also, an attempt was made to ensure that a reasonable spread of individuals from across the political spectrum were selected. For UK political parties and media organisations, only major parties/organisations were selected that could reasonably be identified ideologically by members of the general public.

This survey data is used in **Paper 2** of the thesis.

3.1.3 Secondary Survey Data

In addition to the primary digital and survey data acquired as part of thesis, the analysis conducted in **Paper 3** was augmented with secondary data from the BESIP.

3.1.3.1 British Election Study Internet Panel (BESIP)

The British Election Study Internet Panel is a large online panel study which administers online questionnaires to a sample of adults designed to be as representative of the adult British population as possible. These online panels are conducted at regular time intervals ("waves") where approximately 30,000 adults are sampled in each wave. The first wave of the BESIP was conducted by the British Election Study team between 20th February and 9th March 2014, followed by another 28 waves over the next ten years at roughly equal time intervals leading up to the most recent wave 29 conducted in July 2024. The BESIP has offered unparalleled insight into the attitudes and electoral behaviour of the British period over a turbulent last decade and strives to resample as many of the same respondents from the previous wave as possible to allow for the study of within-person change over time. At each wave, every initial BESIP respondent have been invited to take part, and, in order to maintain cross-sectional representativeness, replacement respondents are sampled who resemble those who have dropped out. Survey sampling and panel attrition are non-probabilistic and so panel weights are used to adjust the raw survey data to be more closely represent the target population (Fieldhouse et al., 2023).

For this thesis, the BESIP is used for two purposes. Firstly, data from Wave 28 has been used in the previous section to provide an up-to-date examination of social media usage in the UK and how Twitter users compare to other social media user groups and the general public. Within the three empirical papers, BESIP data is not used in Paper 1 or Paper 2 but is used extensively in Paper 3. Paper 3 requires estimation of UK constituency left/right ideal point positions as well as probability estimates of their most important issues, spanning the majority of the last parliament (2019-2024). To generate these estimates, data on both relevant individual and constituency-level variables are required. The BESIP provides rich data on individual attitudes to both economic and social values which are useful for estimating left/right ideal points, while the BESIP asks respondents an open-ended question in every wave of what they believe to be the most important issue facing the country at the present time.

Additionally, the BESIP also contains a range of relevant demographic variables for each respondent and the repeated waves throughout the BESIP allows for coverage of changing attitudinal variables over the previous parliamentary period (2019-2024; W20–W25). Importantly, the BESIP also records the constituency code that each respondent belongs to. This can be merged with the BESIP's constituency results file which compiles data from a number of sources on previous election results in each constituency from 2005 to 2019, other electoral details such as voter turnout, European Union (EU) referendum vote share and majority (Fieldhouse et al., 2019). It also contains contextual information derived from the 2011 Census which includes data on population size, demographic composition, and socio-economic variables. Along with individual-level data from the BESIP, the constituency-level variables from the BES constituency file are necessary for modelling constituency positions, as is outlined in more detail in the methods section of **Paper 3**.

3.2 Paper 1: Methods for Estimating Ideal Points Using MP Twitter Follower Networks

The aim of **Paper 1** is to generate left/right ideal point estimates of UK MPs sitting in the HoC over the last parliamentary period (2019-2024). This paper outlines general spatial theory which underpins ideal point estimation, the history of this method in the estimation of ideal points for legislators, and the limitations of such methods in a restrictive legislative system like the UK HoC. It subsequently moves to exploring how researchers have begun using social media networks as a medium for circumventing these restrictions and then outlines what is known as the *Bayesian Spatial Following* model (Barberá, 2015). **Paper 1** is the methodological backbone of the thesis and provides ideal points estimates which are central to the analysis in **Paper 2** and **3**. In Barbera's original paper, ideal points of US legislators are estimated using Bayesian Markov Chain Monte Carlo (MCMC) methods but, as is discussed in the paper, a far more computationally efficient method is adopted instead which has shown to produce a close approximation of traditional spatial models (Barberá et al., 2015; Bonica, 2014).

Correspondence analysis (CA) is a multidimensional scaling technique which can be applied to a two-way contingency table (such as a network adjacency matrix) to transform the data into relational data which can visualise the relationship between variables (Greenacre, 2010). It is conceptually similar to principal components analysis but can be applied to categorical data instead of continuous. This method of dimensionality reduction is applied to the follower network adjacency matrix of the 591 UK MPs constructed from the follower adjacency list extracted from Twitter (Data Set 1) to generate left/right ideal point estimates of MPs. This exploits the fact that connections on Twitter between politically engaged users is driven heavily by an ideological component, where high-profile political elites closer to one another ideologically have a higher probability of sharing similar sets of followers (Barberá, 2015). Thus, in capturing the relationship between MPs' follower networks, CA is able to capture this underlying ideological component as a single spatial coordinate (ideal point) in a latent left/right ideological dimension. This is explained in more detail in Paper 1 and validated using both face validity assessment and through regression analysis against the expert survey data. These left/right ideal points are then used as the primary independent variable in a logistic regression model predicting candidate endorsement in the September 2022 Conservative Party leadership contest.

3.3 Paper 2: Methods for Assessing Media Representation Using Twitter Ideal Points Estimates

Importantly, when generating left/right ideal point estimates of UK MPs using CA on their follower network adjacency matrix, this simultaneously generates ideal points for the follower accounts in tandem. Where MP ideal points are estimated based on the similarity of their sets of followers, ordinary user ideal points are estimated based on the similarity of the sets of MPs they follow. This provides a set of left/right estimates for 424,297 ordinary accounts that followed at least 10 MPs which contains a large number of wider political, media and cultural elites. Subsequently, **Paper 2** draws directly from the user ideal point estimates generated in **Paper 1** to assess descriptive representation of guests on the flagship political programmes of seven major T.V broadcasters in the UK between 2022 and 2024. Seven political programmes across six T.V broadcasters are selected for study: BBC One: Question Time, BBC One: Sunday with Laura Kuenssberg, BBC Two: Politics Live, ITV One: Peston, Channel 4: The Andrew Neil Show, Sky News: Sophy Ridge on Sunday, and GB News: The Camilla Tominey Show. **Paper 2** begins by outlining the history of public service broadcasting in the UK and how its history of strict impartiality has been challenged by accusations of growing partisanship on mainstream broadcasters like the BBC and Channel 4, as well as the emergence of new channels such as GB News and TalkTV. It then outlines previous methods for assessing representation in the media and how social media networks can be leveraged to provide measures of left/right position beyond the need for manual coding.

As with **Paper 1**, the construct validity of these user ideal points as adequate reflections of ideology are assessed using face validity checks, as well as a formal regression analysis against the *YouGov* general public survey data. To formally test the descriptive ideological representation of guest selection on the seven selected shows, complete guest lists for each programme are manually gathered from various sources including the programme websites, Spotify podcast descriptions, and Twitter feeds. Guests are then matched to their corresponding ideal point by their Twitter username (if they have one) and each programme's overall guest list ideal point distribution is compared against one another, as well as the distribution of ordinary Twitter users, and the distribution of Twitter elites. Two key statistical tests are used to confirm differences between different ideal point distributions: a Mann-Whitney *U* test (McKnight and Najab, 2010) is used to test whether the difference between two distributions is statistically significant, and a Hartigan's Dip Test of Unimodality (Hartigan and Hartigan, 1985) is used to test the presence of bimodality of each distribution. Guest lists across the two year period are also split into quarterly intervals to assess any significant difference in guest representation over time.

3.4 Paper 3: Methods for Assessing Dyadic Representation Using Twitter Data

Paper 3 subsequently builds on the work of Paper 1 and Paper 2 to assess the relationship between an MP's Twitter profile and the ideal point position of their constituency. This is conducted across three sub-studies within the paper: where Study 1 makes direct use of the ideal point estimates generated and validated in **Paper 1** for MP left/right position based on their follower networks; Study 2 makes use of the user ideal points generated in **Paper 1** and validated in **Paper 2** for MP left/right position based on their retweet networks; and Study 3 makes the first use of Data Set 2 to generate an MP "general issue" ideal point based on the topics of their tweets. While the DTD for Study 1 and Study 2 is derived directly from the data generated in **Paper 1**, Study 3 categorises the tweet timelines of MPs over the last parliamentary period (2019-2024) into ten topics, matching the most important issue code schema used in the BESIP. This includes topics such as health, economy, immigration and Europe and tweets are coded into topics based on a keyword dictionary containing 436 terms relating to each topic. Each MP's timeline is split across these ten topics based on the % of each set of tweets that belong to each category.

Assessing MP position along these three dimensions, these are regressed against the economic and social left/right ideal points of their respective constituencies, as well as their most important issue probabilities to test for within-party responsiveness of MP Twitter profiles to their constituents. The key methodological contribution of Paper 3 that departs from Papers 1 and 2 is the technique used to estimate constituency-level positions. Drawing directly on respondent data from Wave 20 to 25 of the BESIP separately, MRP is used to model the position of each UK constituency at five time points across the previous parliament. Initial respondent left/right ideal points and their most important issue are estimated by fitting Bayesian multilevel regression models with MCMC to respondent economic and social value preferences and their most important issue, incorporating individual and constituency-level predictors into the model. These fitted models are subsequently post-stratified using a poststratification frame to estimate constituency-level positions based on the joint distribution of respondent demographic characteristics by local area (Hanretty, 2020). The principal aim of this paper is not to establish any particular causality between MP digital position and the position of their constituency. Instead, it is an exploratory study that looks to demonstrate the capacity of social media analysis to circumvent the restrictions of a restrictive legislative system such as the HoC to assess an important political concept like dyadic representation. As with **Paper 1** and **Paper 2**, more specific details of this method can be found in the main body of **Paper 3**.

Having now outlined the data sets and general methods used in the three empirical papers that form the spine of this thesis, and how they connect together, these three papers are now presented in order and in detail in the following three chapters.

Chapter 4

Estimating Ideal Points of British MPs Through Their Social Media Followership - Paper 1

Ideal points of MPs in the UK House of Commons are characteristically difficult to ascertain due to tight party discipline and strategic voting by opposition members. This paper generates left/right ideal point estimates for 591 British MPs sitting in the HoC as of 22/08/2022, ascertained through their social media followership. Specifically, estimates are derived by conducting correspondence analysis (CA) on MP Twitter follower networks which are subsequently validated against an expert survey, confirming that these estimates have a high degree of between-party ($R^2 = 0.93$) and within-party (Con: $r = 0.84$; Lab: $r = 0.81$) accuracy. The informative value of these estimates is then demonstrated by predicting candidate endorsement in the Sept 2022 Conservative leadership contest, confirming that an MP's ideal point was a statistically significant predictor of candidate endorsement, with Liz Truss drawing support primarily from the further right of the party.

A version of this paper was published as a letter in the *British Journal of Political Science*:

Gaughan, C. (2024). Estimating Ideal Points of British MPs Through Their Social Media Followership. *British Journal of Political Science*. Pages 1-13. doi:10.1017/S0007123424000450.

4.1 Introduction

When people talk of ideology, they typically conceive of it in spatial terms between the “left” and the “right”. In political science, an *ideal point* refers to an individual’s position within this latent ideological dimension, where spatial voting models will attempt to estimate ideal points of legislators using the votes they cast on legislation (Enelow and Hinich, 1984). Spatial voting models operate under the guiding principle that political preferences can be represented as points in a uni- or multi-dimensional space, where legislators will consistently favour policy outcomes that most closely align with their own ideal point (Clinton et al., 2004). Classical spatial theory derives from the seminal works of Hotelling (1929), Downs (1957) and Black (1958) and has since been foundational to the study of legislative and electoral politics. Evaluation of legislator behaviour in the US Congress using Poole and Rosenthal’s (1985; 1991; 1997) DW-NOMINATE is fundamentally built on the principles of spatial voting, estimating ideal points of members of Congress through their roll-call data. Beyond US Congress, this method has also been applied to other legislative bodies including European Parliament (Hix, 2001; Lo, 2018; Martin, 2021), the UN General Assembly (Binder and Payton, 2022), the Supreme Court (Cameron and Park, 2009), and in the national assemblies of other countries (Clerici, 2021; Rosenthal and Voeten, 2004).

In the UK, attempting to estimate the political preferences of members of the House of Commons (HoC) using this method has proven to be extremely difficult. Spatial voting models fundamentally rely on the assumption of *sincere voting*: that actors will cast a vote for the outcome that they prefer above all others (Clinton et al., 2004). In doing so, their votes act as unobtrusive measures of their genuine policy preferences. As such, these models work most effectively in environments were legislators can vote independently from their party, optimising the informativeness of the roll-call data. In the British case, parties are characterised as tightly disciplined, with a strong whipping system and “government vs. opposition” parliamentary culture fostering a greater degree of strategic voting (Spirling and McLean, 2007).The tendency of MPs to toe the party line renders their position largely indistinguishable from the general position of their party in most circumstances (Hix and Noury, 2010). In certain cases of extreme salience such as the Iraq War, Brexit, or Same-Sex Marriage, MPs have shown themselves to be willing to break from the party whip and rebel, but these instances are extremely rare (Slapin et al., 2018, p.20).

Franklin and Tappin (1977) were one of the first to identify this as a significant limitation for measuring legislator preferences in the UK and suggested the use of Early Days Motions (EDMs) as an alternative to roll-call data. EDMs are formal, non-binding expressions of opinion by members of the HoC which other MPs can subsequently sign to indicate their support for the motion (Kellermann, 2012, p.760). Topics of EDMs can vary greatly and can be used for a variety of purposes such as drawing attention to specific events or to criticise other MPs or Lords.[1] Having the distinct benefit of being open and unwhipped, they are generally considered to be largely unobtrusive measures of backbench opinion in the HoC and thus an adequate source of legislator preference. EDMs have been used in multiple studies in the UK to measure legislator policy preference (Berrington, 1973; Finer et al., 1961; Hanretty et al., 2017; Kellermann, 2012; Leece and Berrington, 1977) but, as Kellermann (2012, p.758) points out, the unique features of EDMs can inhibit the ability of established methods to map these directly to the conventional left-right spectrum.

Subsequently, this paper seeks to make a contribution to this space by generating ideal point estimates of UK MPs through their Twitter follower networks. Adopting an approach established by Barberá (2015), complete follower bases for all 591 members of the HoC with an active Twitter account as of the 22nd August 2022 are harvested through Twitter's API. MPs are then mapped across a unidimensional left-right spectrum, with ideal points scaled using correspondence analysis (CA) on their follower bases. The accuracy of these estimates are then validated against a set of ideological scores for a sub-sample of 30 MPs, ascertained through a survey of experts in the field of British politics.The myriad potential applications for this scale in the analysis of British politics is then demonstrated by modelling the ideological component to candidate endorsements in the September 2022 Conservative party leadership contest.

The results of this research make a two-fold contribution to the field:

1) Validates the efficacy of employing multidimensional scaling methods on UK MP Twitter networks to scale ideological position, and may encourage the use of this method on other legislative bodies with similar restrictive parliamentary cultures.

2) Confirms the significance of ideology in party leadership endorsement, specifically in the case of the September 2022 Conservative Party leadership contest.

4.2 Literature Review

4.2.1 Existing Approaches to Ideology Estimation of Political Elites

The purpose of this paper is to demonstrate an alternative to measuring ideal points of legislators in the UK beyond traditional methods such as roll-call analysis. Research on ideal point estimation is well-documented, as are the problems associated with attempting to estimate them for legislators in Westminster systems. Ideal point estimation matters for seeking to better understand legislative politics and lacking

[1] A more detailed description of EDMs can be found on the UK Parliament website: https://www.parliament.uk/about/how/business/edms/.

such estimates poses a limitation for political scientists and academics trying to study these bodies. It is important to note that a plethora of contemporary measures for ideology estimation do exist and these will be briefly discussed along with their limitations, before addressing the specific difficulties with ideal point estimation in the case of Britain. For example, one of the more traditional methods for inferring the ideological position of legislators remains the use of expert surveys. First popularised by Mair and Castles (1984), they have since been an established mode of gauging party and legislator ideology (Benoit and Laver, 2006; Huber and Inglehart, 1984; Marks and Steenbergen, 1984; Ray, 1999; Whitefield et al., 2007).[2] A select group of politically informed experts within a particular country are surveyed, asking them to locate political parties or representatives on a pre-defined ideological dimension(s) and then these scores are averaged to build the scale. The benefits of using such a method are evident: they are relatively quick and simple to conduct (compared to other scaling methods), the use of experts reduces ambiguity around estimation strategies, and they have shown to be largely reliable measures of ideological position that remain comparatively stable across countries and over time (Bakker et al., 2014). However, concerns have been raised over the validity of these measures given the lack of coherence around how each expert interprets the left/right ideological scale - or the overall salience of this scale - and the impact of the ideological position of the experts themselves in biasing their placements (Budge, 2001; Curini, 2010).

One possible alternative is through the analysis of rhetoric. An increasingly popular technique for ideological estimation is text-based scaling where ideology can be inferred from the language and themes used within bodies of text. One of the first and most prominent text-based ideological scaling procedures was developed by Slapin and Proksch (2008) called WORDFISH, using an unsupervised modelling strategy to scale ideology based on word frequencies. Text-based scaling for ideological inference is also central to the work conducted by the Manifesto Project, tracking the ideological position of political parties over time through the annotation of quasi-sentences from manifesto excerpts by political country experts.[3] One of the primary benefits of using text-based scaling methods is their versatility in application to a variety of different corpora, making it suitable for ideological estimation in many situations where other scaling methods would not be appropriate.

In previous research, text-based scaling has been used on newspaper articles to estimate media bias (Gentzkow and Shapiro, 2010; Groseclose and Milyo, 2005; Kaneko et al., 2021), within political speeches (Iyyer et al., 2014; Monroe et al., 2008; Sim et al., 2013), through congressional debates (Bayram et al., 2019; Thomas et al.,

[2]Also, see the Chapel Hill Expert Survey. CHES estimate party positioning on ideology and policy issues, and international relations for national parties in countries across the world. The trend file stretches across 6 waves between 1999-2019 and is one of the most robust datasets for ideological scaling based on expert surveys anywhere in the world. Access: https://www.chesdata.eu/

[3]The Manifesto Project is a renowned research project which collects, analyses, and distributes the manifestos of parties from countries all around the world. Access: https://manifesto-project.wzb.eu/

2006), legislative speeches (Beauchamp, 2011; Lauderdale and Herzog, 2016) and, more recently, on social media content (Bond and Messing, 2015; Djemili et al., 2014; Fagni and Cresci, 2022; Preoţiuc-Pietro et al., 2017; Temporão et al., 2018; Zhang and Counts, 2015). However, one of the key concern with text-based scaling is that textual data can be high-dimensional, meaning that not all language in a body of text maps onto the same latent concept (Goet, 2019, p.520). Language can be *noisy*, especially if drawn from text that is not explicitly political such as social media posts, media articles, and so on.

Roll-call estimation strategies largely circumvent the issues associated with survey or discourse-based methods by directly estimating the ideological position of legislators from their voting records (Berry et al., 1998; Carson and Oppenheimer, 1984; Krehbiel, 1986; Poole and Rosenthal, 1985, 1991, 1997). The spatial model of voting positions voters in a single or multi-dimensional space, where each voter is modelled as having an ideal point somewhere in this space. Rational voters will subsequently 'vote' for candidates that are closest to them within the space (Enelow and Hinich, 1984). In the case of roll-call analysis, legislator voting records can act as observable signals about their latent ideological position, where legislators will consistently opt for policies that are closest to them in the ideological dimension (Clinton et al., 2004). As noted, the applicability of these models to legislative bodies which foster a high-degree of party discipline, strategic voting, and government vs. opposition alignments is extremely limited. Thus, opportunities for legislators in strict majoritarian-style parliamentary systems to signal their latent ideological position distinctively from their party are scarce.

4.2.2 Ideal Point Estimation in a Westminster System

This has proven to be the case in the UK's Westminster system, where MP positions become largely indistinguishable from the general position of their party (Hix and Noury, 2010; Spirling and McLean, 2007). Westminster systems generally appear to be far more challenging environments for measuring within-party policy preferences, with highly disciplined parties competing electorally on national policy platforms (Bawn and Thies, 2003; Hanretty et al., 2017). The top-down party dominance of Westminster systems such as the HoC poses an interesting question about how successfully legislator policy preferences can be disentangled from their parties. Justification for the importance of gauging within-party policy responsiveness in the UK can be rooted in the last decade alone, where complex intra-party dynamics of Britain's major parties have contributed to much of Westminster's turbulence since 2010. Sizeable opposition to Jeremy Corbyn's leadership of the Labour Party between 2015 and 2020 from backbench MPs and from within his own shadow cabinet led to mass resignations, multiple rebellions, and a vote of no confidence during his tenure as Leader of the Opposition (Slapin and Proksch, 2008). Concurrently, the long-standing divisions within the Conservative Party over Britain's relationship to Europe and the EU has played a substantial role in support for, or opposition against, negotiated withdrawal agreements from the EU after Brexit, as well as for the four party leaders elected since 2016 (Jeffery et al., 2018, 2022, 2023; Roe-Crines et al., 2021; Booth et al., 2023).

It could be argued that now more than ever, the need for valid estimates of within-party ideal points of MPs is crucial to improving our understanding of legislative politics in the UK. The limitations of traditional methods of roll-call estimation in restrictive Westminster systems such as the HoC has already been noted. The difficulties of exploiting EDMs as an alternative data source has been flagged by Kellermann (2012) for their degree of idiosyncrasy. In their paper, they do well to build on traditional attempts to derive ideal point estimates for British parliamentarians using EDMs by incorporating a Bayesian framework into the estimation strategy, producing a far more accurate set of ideal point estimates than previous work. This strategy was extended by Hanretty et al. (2017) to examine the degree of legislator responsiveness to constituency opinion in the HoC, outside of the dominant party agenda. However, criticism towards EDMs has been growing in recent years, having been described as "parliamentary graffiti" and "politically impotent", bringing about a silent boycott among a sizeable number of MPs in the HoC who consider them an ineffective and wasteful lobbying tool (Goodhind, 2023). The recent decline in MPs tabling or signing EDMs poses a significant problem for their use in ideal point estimation moving forward. Other workarounds for measuring individual legislator policy positions in the UK have included direct surveys of parliamentary candidates in previous elections asking them to rank their policy preferences along multiple dimensions (Norris and Lovenduski, 1992, 1997, 2001). However, these types of surveys traditionally suffer from low response rates (Kam et al., 2010), and responses that are given can still suffer from the same bias of insincerity generated by strict party discipline (Hanretty et al., 2017, p.237).

Spirling and Quinn (2009) make a technical attempt at evading the drawbacks of a tightly-disciplined legislative by adopting a strategy based on a Dirichlet process mixture model, allowing them to estimate a probable number of intra-party voting blocs within the Labour Party.[4] Schwarz et al. (2017) make the case that as long as legislators toe the party line in regards to votes on legislation, they are relatively free to express themselves through legislative speeches. This fact is exploited by Goet (2019) to measure political preferences of MPs in the HoC through parliamentary speech records, using textual analysis of over 6.2 million speeches between 1811-2015 to measure polarisation in the UK. Another possibility would be to scale ideology

[4] A Dirichlet process is a stochastic process used in Bayesian non-parametric models of data. It is a distribution over distributions, making it a particularly useful modelling strategy for cluster analysis when the prior distribution of clusters is unknown.

based on cases where MPs have broken party cohesion and rebelled against the whip. This could act as a signal of an MP's latent ideology, treating rebellion as a direct rejection of the policy being voted on. Unfortunately, MP rebellions in the HoC are extremely rare – less than 1% of all recorded votes are rebellious (Slapin and Proksch, 2008, p.20). There is also some debate around how much MP rebellion is a representation of their ideological preference versus other relevant factors such as strength of party affiliation, length of time in office, or perceived future party prospects (Benedetto and Hix, 2007; Cowley, 2002; Kam, 2001).

4.2.3 Estimating Ideal Points Through Social Media Networks

For the purpose of this research, this paper adopts a contemporary method of its own which has grown in popularity over the last few years. The advent of major social media networking platforms such as Facebook and Twitter has led to a surge in methods scaling ideology through the observation of individual users on these sites (Back et al., 2010; Barberá et al., 2015; Colleoni et al., 2014; Conover et al., 2011). Research has shown that network clustering on social media is driven heavily by an ideological component (Bright, 2016; Colleoni et al., 2014). The widescale uptake of social media – primarily, Twitter - by elite political actors over the last decade has fostered a environment in which both elite and non-elite actors now exist in the same digital space. A plethora of research has utilised this to expand our understanding of both intra-elite and elite-to-non-elite interactions (Boutet et al., 2013; Cherepnalkoski and Mozetič, 2016; Esteve Del Valle et al., 2022; Keller, 2020; Weaver et al., 2018). Seminal work by Barberá (2015) leveraged this to develop a Bayesian *Spatial Following* model, utilising MCMC methods for Bayesian inference to generate ideal point estimates for elite political actors on Twitter based on their respective follower bases. Further iterations of this model have yielded equally accurate estimates through a far simpler and more computationally efficient method, correspondence analysis (CA) (Barberá et al., 2015).[5] This modelling strategy has been further replicated in multiple studies with a high degree of success (Briatte and Gallic, 2015; Hassell et al., 2020; Liang, 2018; Pennycook et al., 2021; Souza et al., 2017).

The method proposed here is not strictly novel, building on the original work of Barberá (2015) who first proposed the use of Twitter follower networks for ideal point estimation. In his paper, ideal points are estimated for legislators in the US, alongside five European countries including the UK, which are also compared against expert estimates. However, this was an initial demonstration of the method's validity and was not comprehensively applied to the UK case, only including MPs from the three major parties (Con/Lab/Lib Dem). Furthermore, the gradual maturation of Twitter as

[5]This method has also been worked into a popular R library called *tweetscores*. The original replication code for this method can be found at: *https : //github.com/pablobarbera/twitter¡deology*

a social media platform since 2015 makes an updated application of this method in 2022 an insightful contribution, as well as including a far larger number of MPs from parties across the HoC. A primary benefit of scaling ideal points through social media networks is the ability to to dynamically track and update these estimates over time. Moreover, these estimates are not restricted to a specific legislative session in the same way as scaling ideal points through roll-call data or EDMs may be. For instance, attempting to compare ideal points of legislators between different legislative bodies or between legislators in the same legislative bodies over time might not be feasible, whereas many legislators past and present from legislative bodies all around the world exist in the same digital space on Twitter. Similarly, the existence of non-elite actors in the same space allows for the estimation of ideal points for ordinary Twitter users to be scaled in tandem.

4.3 Method

The guiding assumption that underpins more general spatial voting models is that voters will most likely opt for individuals, parties, or policies that they believe most closely align with their own ideological position (Enelow and Hinich, 1984) A network-based approach to ideological scaling follows a similar assumption: actors within a network will *connect* to other actors they believe most closely align with their own ideological position. This baseline assumption is heavily informed by the concept of *homophily*: a sociological phenomenon describing the tendency of individuals to associate and form bonds with those most similar to themselves (McPherson et al., 2001). The homophilic nature of network structures on social media platforms is well-documented (Aiello et al., 2012; Cinelli et al., 2021; Colleoni et al., 2014; Grevet et al., 2014). Barberá's (2015) *Spatial Following* model leverages this fact to estimate ideological position of actors based on their Twitter follower base. This model relies upon the decision of ordinary users to follow (or not to follow) a political actor to act as a signal about the latent ideological position of that user, as well as their perception of the ideological position of the political actor. The principal assumption is that the closer an ordinary user i perceives a political actor j to be to them in the latent ideological space, the higher the probability that they will follow them on Twitter.

Thus, in this way, the model treats ordinary users as something like quasi-'experts', who are consciously (or sub-consciously) 'rating' political elites on an ideological scale through their follower choices. An argument against this method may be the unrepresentativeness of Twitter's user population. Twitter is commonly considered to be an elite platform, not only in regards to the absolute volume and visibility of elite actors present on the platform (Ausserhofer and Maireder, 2013; Dagoula, 2019), but also in the 'eliteness' of the general Twitter population. Ordinary Twitter users are, on average, wealthier, more highly educated, and more politically engaged than the general population (Blank, 2017; Mellon and Prosser, 2017). However, given that this modelling strategy fundamentally relies upon the accuracy of ordinary users to scale the ideological space, a more highly educated and politically engaged sub-population is in fact desirable.

4.3.1 The *Spatial Following* Model

With these assumptions in mind, the *Spatial Following* model is as follows:

Letting $y_{ij} = 1$ if an individual Twitter user i independently chooses to follow a political actor j and $y_{ij} = 0$ otherwise, the probability that $y_{ij} = 1$ can be formulated as a logit model:

$$P(y_{ij} = 1 | \alpha_j, \beta_i, \gamma, \Theta_i, \Phi_j) = logit^{-1}(\alpha_j + \beta_i - \gamma \parallel \Theta_i - \Phi_j \parallel^2). \tag{4.1}$$

where α_j controls for the overall popularity of the political actor j (reflecting their likelihood of attracting more followers), β_i controls for user i's level of political interest (reflecting their likelihood to follow more political actors), γ is a normalising constant, and $\parallel \Theta_i - \Phi_j \parallel^2$ is the squared Euclidean distance in the unidimensional space between Twitter user i and political actor j. In this case, Θ_i reflects the ideal point of Twitter user i and Φ_j reflects the ideal point of political actor j. It is from these two latter parameters (Θ_i, Φ_j) that we can derive ideological estimates for both elite political actors and for the ordinary users that follow them.

Traditionally, latent space models are estimated through Bayesian methods using MCMC, and this is the case in Barberá's (2015) original *Spatial Following* model. However, as is documented in Barberá et al. (2015, p.1533), this becomes computationally intractable for extremely large network datasets such as those found on social media. Thus, CA is used instead, which has been found to produce a close approximation of a statistical ideal point model but at a much-reduced computational cost (Bonica, 2014, 369).

4.3.2 Data Collection

As of the week commencing the 22/08/2022, the number of UK MPs with an active Twitter account was n = 591. The total combined number of followers between these MPs was m = 34,653,181. Summary statistics for each political party's representation on Twitter can be seen in **Table 4.1**. Using Twitter's API, accessed through Python's *tweepy* package, the entire population of users who followed at least one UK MP was harvested over a three-and-a-half week period commencing 22/08/2022 and completing on 15/09/2022.[6] The overall number of unique accounts this returned was m = 11,071,104. A significant proportion of these accounts will not be useful due to inactivity and the presence of bots on the network, and so the sample was refined before being used for analysis. Following two steps from Barberá's 5-step approach to filtering Twitter user samples to remove potentially fake or non-active users (Barberá, 2015, p.81), profiles were discarded if they satisfied any of two criteria: 1) have sent fewer than one hundred tweets and/or 2) have less than twenty-five followers. After filtering the dataset using these criteria, this left the sample of unique accounts at m = 4,460,657.

	N	N (%)	Median Followers	Median Following	Median Tweets
All	591/650	91%	18,385	1,515	7,710
Conservative	312/359	87%	14,651	1,049	4,850
Labour	194/201	97%	27,495	2,196	12,865
SNP	45/45	100%	16,368	2,129	16,016
Lib Dem	14/14	100%	18,038	1,642	9,264
Sinn Féin	7/7	100%	20,448	1,649	7,170
DUP	6/8	75%	14,009	914	4,752
Independent	3/6	50%	72,657	6,737	16,882
Plaid Cymru	3/3	100%	10,644	1,967	25,355
Alba	2/2	100%	14,775	1,543	16,093
SDLP	2/2	100%	46,277	3,694	19,270
Alliance	1/1	100%	18,299	2,106	8,522
Green	1/1	100%	586,416	6,738	30,674
Speaker	1/1	100%	41,745	3,126	5,442

TABLE 4.1: All UK MP Twitter accounts. N (%) relative to N in the House of Commons.

Among these accounts, the median number of MPs followed is extremely low (1). The overall distribution of users by the number of MPs they follow is heavily right skewed, with over half of all users (52%) in the filtered dataset following only a single MP (See **Figure 4.1**).

Recent research has shown that the majority of Twitter users do not follow political elites, instead favouring non-political opinion leaders to political ones (Mukerjee et al., 2022; Wojcieszak et al., 2022). The heavily skewed distribution of users by how many MPs they follow suggests a concurrent pattern. However, their research did also indicate that among those users who *do* choose to follow political actors on Twitter, they demonstrate a far greater preference for ideological congruity. Given this fact, and the reliance of the model on the political knowledge and ideological informativeness of ordinary users, there is motivation to subset this sample of users to select only the most politically engaged. As part of Barberá's (2015, p.81) 5-step filtering approach, ordinary users were subsetted to only include those users who follow at least 10 political elites. For the benefit of optimising the model performance, as well as significantly reducing computational time and for general consistency with Barberá's original methodology, the sample of ordinary users is filtered to only

[6]The *Tweepy* GitHub page can be accessed here: https://github.com/tweepy/tweepy/

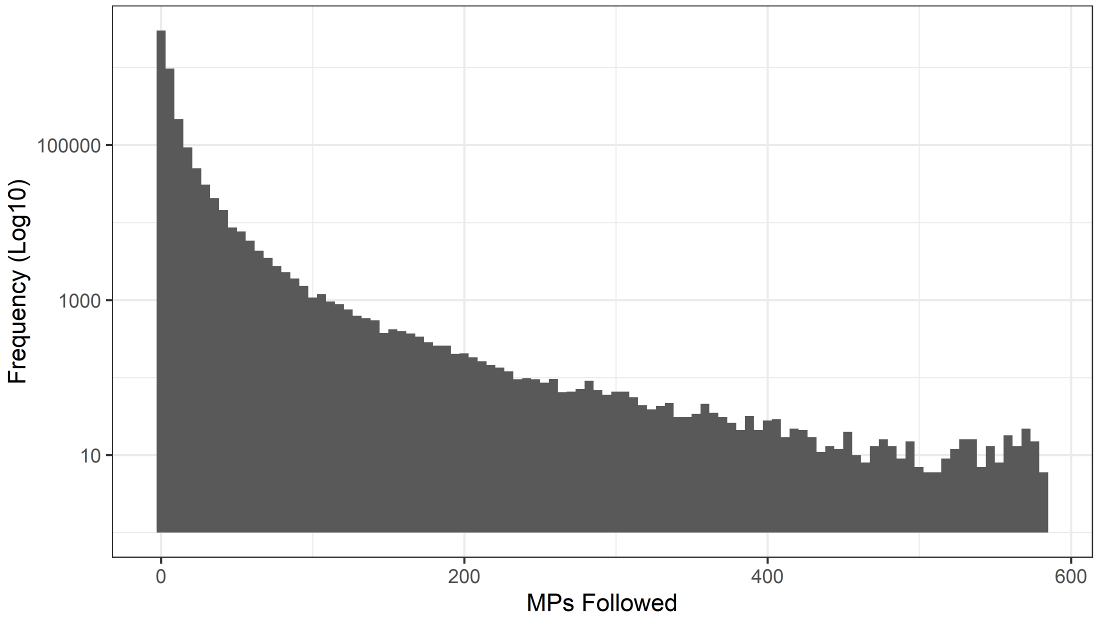


FIGURE 4.1: Histogram of users by the number of MPs they follow. Y-axis is on a base-10 log-scale and number of bins = 100.

include those users who follow at least *ten times* the median number of MPs (10). This especially informative subset contains m = 424,297 users and e = 11,443,165 unique follow connections with the n = 591 MPs.[7] Summary statistics for these accounts are shown in **Table 4.2**.

Users		Median Count				
N	Verified (%)	MPs Followed	Followers	Following	Tweets	Listed
424,297	2%	17	347	1,163	2,490	2

TABLE 4.2: Summary statistics of profile metadata for the subset of especially informative users.

Taking this final set of MPs n and their especially informative set of followers m, a network adjacency matrix **Y** is constructed where individual elements can take a binary form of integers **0** or **1**. Where an ordinary user i follows MP j $y_{ij} = 1$, otherwise $y_{ij} = 0$. Ordinary users m form the rows and MPs n form the columns, thus generating a large matrix of dimensions **424,297** x **591**. The nature of this network is fundamentally non-reciprocal with e connections only flowing in one direction from users m to MPs n.

n	***m***	***e***	**Potential *e***	**Realised *e***
591	424,297	11,443,165	250,759,527	5%

TABLE 4.3: Network adjacency matrix summary statistics. Realised e calculated as the actual e over potential e.

[7] Robustness checks are conducted to assess model performance and accuracy of the model estimates using various user sample thresholds. To see the results of these checks, see *Appendix A.1*.

With only 5% of all possible follower connections realised between users to MPs, this generates a sparse network adjacency matrix. However, even with a low network density, this can still provide an adequate informational source for constructing the spatial model.

4.3.3 Fitting a Correspondence Model

Correspondence analysis is a multidimensional scaling technique that is conceptually similar to principal components analysis but can be applied to categorical data instead of continuous (Greenacre, 2010). The general mathematical process of CA is relatively straightforward - following Barberá et al. (2015):

Consider an i x j contingency table - in this case, our **Y** follower adjacency matrix - with elements y_{ij} and total number of observations $e = \sum_{ij} y_{ij}$:

Step 1: Convert **Y** into a correspondence matrix **P** by dividing **Y** by its total sum: $\mathbf{P} = \mathbf{Y} / \sum_{ij} y_{ij}$. This converts all elements of **Y** to proportions.

Step 2: Calculate row and column *masses* of **P** by summing the elements of each individual row and column: $r_i = \sum_j p_{ij}$ and $c_j = \sum_i p_{ij}$. These are then used to construct diagonal matrices of $\mathbf{D}_r = \text{diag}(\mathbf{r})$ and $\mathbf{D}_c = \text{diag}(\mathbf{c})$.

Step 3: Compute a matrix of standardised residuals **S** where $\mathbf{S} = \mathbf{D}_r^{1/2}(\mathbf{P} - \mathbf{r}\mathbf{c}^T)\mathbf{D}_c^{1/2}$. These residuals reflect the difference between expected and observed values in each element of **P** based on their corresponding r_i and c_j. By standardising them in this way, CA is able to control for α_i (political interest of user i) and β_j (popularity of actor j) effects by re-weighting rows and columns based on how populated they are. This similarly helps to adjust for case where sample sizes are smaller (i.e: low follower counts).

Step 4: Subsequently, each element of matrix **S** contains a residual S_{ij} which is treated by the CA model as a reflection of the ideal point distance between user i (Θ_i) and actor j (Φ_j) in the latent ideological space. The singular value decomposition (SVD) of **S** is then calculated, such that $\mathbf{S} = \mathbf{U}\mathbf{D}_\alpha\mathbf{V}^\mathbf{T}$ where $\mathbf{U}^\mathbf{T}\mathbf{U} = \mathbf{V}^\mathbf{T}\mathbf{V} = \mathbf{I}$. SVD is the primary algorithmic procedure for identifying the dimensional space, onto which the column and row coordinates can be projected: $\Psi = \mathbf{D}_r^{1/2}\mathbf{U}$ for rows (ordinary users) and $\Gamma = \mathbf{D}_c^{1/2}\mathbf{V}$ for columns (MPs). In this case, these coordinates for both the rows and columns reflect their ideal points in the latent ideological space.

4.4 Scaling the UK House of Commons

When scaling the HoC, the nationalist parties pose a unique challenge. The modelling strategy here assumes that, when controlling for α_i and β_j, the primary component that explains the distance between actors j and users i is ideological. Thus, when scaling the HoC, one would expect MPs to cluster around some form of shared ideological or partisan identity. However, in the case of nationalist parties, their shared regional component appears to overwhelm their ideological one. This effect is especially dominant with the Scottish National Party (SNP), the third largest party block in the HoC, and significantly warps the overall ideological scale as well as the estimates for MPs from other parties.[8] Fortunately, another benefit of using CA for ideal point estimation is the ability to scale spatial maps using an initial subset of data and then subsequently project supplementary data onto this pre-constructed space. Therefore, the ideological space is initially scaled excluding nationalist parties to ensure that the dimension being mapped reasonably approximates left/right ideology. The nationalist party MPs are then retroactively projected onto this scale.[9] Multiple dimensions can be scaled using CA. However, for the purpose of ideal point estimation along the traditional left/right spectrum of British politics, only the first dimension is required to plot the spatial map.[10]

Figure 4.2 illustrates the distributions of MP ideal points by political party in the HoC. Plotting the spatial map demonstrates that the CA model has high face validity. There is demonstrable between-party clustering indicating that the model can successfully discriminate between MPs from different political parties. More importantly, MPs from different parties that we know officially occupy similar spaces on the ideological axis such as the Greens and Labour on the left or the DUP and Conservatives on the right are also clustered closer together. This indicates that the dimension is not simply capturing partisanship but does in fact reflect the left/right spectrum of British politics. Ordinary users are not necessarily just following MPs from the political parties they identify with, but MPs from across parties that they perceive to sit in reasonable proximity to themselves along the ideological axis.

[8] It is clear that the clustering dimension is not simply partisanship given that the same effect is also present in the Alba Party, Scotland's other nationalist party. Likewise, Scottish MPs within other non-nationalist parties also skew significantly towards the SNP, indicating a regional effect. See *Appendix A.2.2* for more details.

[9] Including the Speaker and Independents, there were 13 parties in the HoC at the time of data collection. 7 of these were nationalist parties with a combined total of 66 MPs who are treated as supplementary columns. There are 2,153 ordinary users who exclusively followed nationalist MPs requiring them to be treated as supplementary rows.

[10] For details of additional dimensions scaled by the CA model, see *Appendix A.2*.

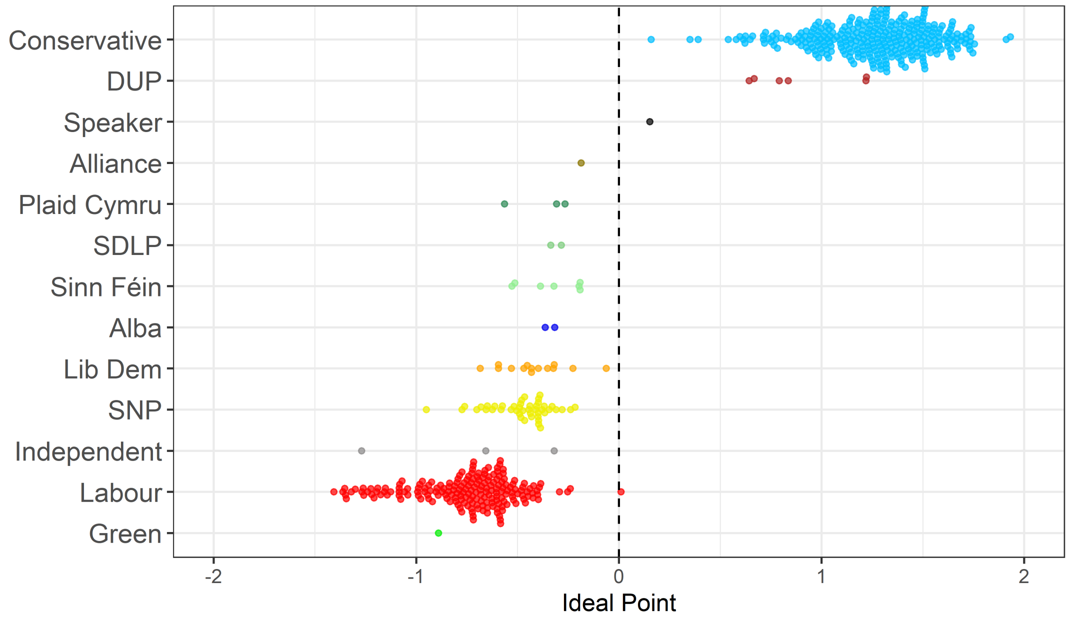


FIGURE 4.2: Beeswarm plot of the 591 MPs in the House of Commons with estimated ideal points, grouped and coloured by party affiliation. Parties are ordered along the y-axis by median party ideal point, starting at the bottom from furthest to the left up to the furthest right.

4.4.1 Model Validation

To formally validate the ideal point estimates, an expert survey was conducted. This survey was distributed to a select group of expert academics in the field of British politics. 133 experts were contacted with a 53% response rate (70). The survey asked them to place a sample of 30 MPs on an ideological scale between 0 (Left) and 10 (Right). 13 MPs were sampled from both the Conservative and Labour parties, along with two from the Liberal Democrats, one from the Green Party, and one Independent. There was an attempt to balance the sample with some more established MPs along with some that are reasonably lesser known, as well as between MPs with larger Twitter followings and those with less. This was to ensure that model performance for both left-wing and right-wing MPs could be assessed, and also to validate its accuracy when estimating placement for MPs with smaller sample sizes (low follower counts).[11] Validation of the CA model estimates for the 30 MPs against the mean estimates provided by the experts is illustrated in **Figure 4.3**.

For validation, weighted least squares regression was conducted predicting the expert ideology estimates using the ideal point estimates generated by the CA model. Model weights were applied using the standard errors of the expert estimates to account for the degree of uncertainty in the validation set itself. Model coefficients would indicate that the CA model ideal points have a high degree of between-party accuracy (R^2 =

[11] For a detailed description of the survey design, expert sample selection, and summary statistics of the expert estimates, see *Appendix A.3*.

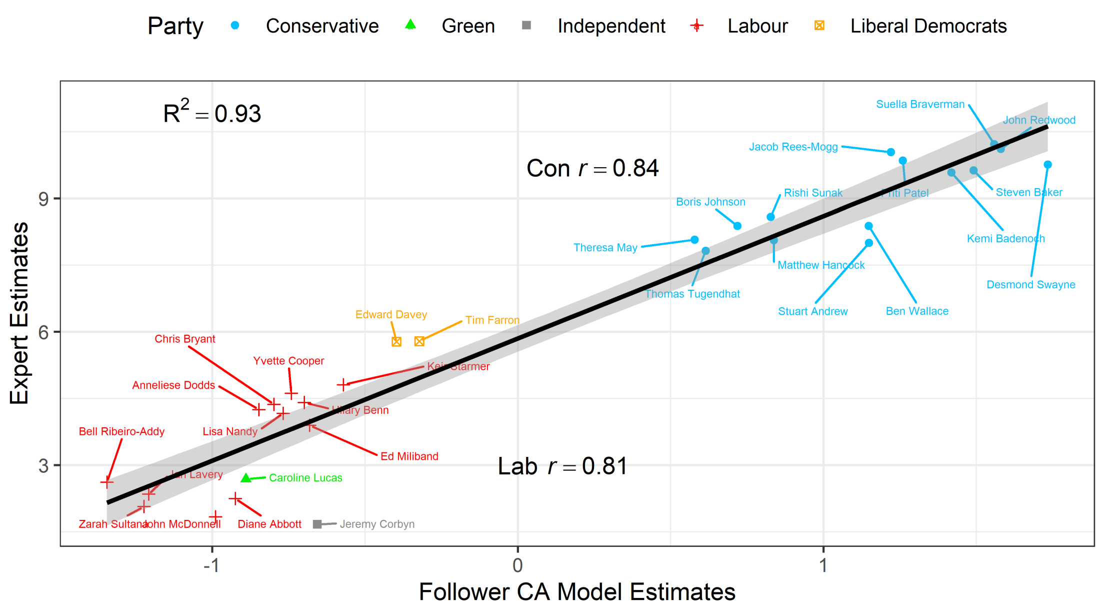


FIGURE 4.3: The ideal point estimates for the 30 MPs is plotted along the x-axis and the mean ideology estimates provided by the experts is plotted along the y-axis.

0.93). Pearson's correlation tests were used to assess the within-party accuracy of the CA model between Conservative and Labour MPs respectively, demonstrating a high degree of correlation in both cases (Con: $r = 0.84$; Lab: $r = 0.81$). Notably, the two MPs with the largest standard errors in the model validation were Jeremy Corbyn and John McDonnell, along with Diane Abbott who was fourth.[12] In all three cases, the CA model estimates placed them closer to the centre than the experts, who placed them much further to the left. This is noteworthy as all three were traditionally from the further left of the Labour Party before taking up major shadow cabinet positions under Corbyn's leadership between 2015 and 2020. In these cases, the model has possibly managed to capture an interesting effect, illustrating their movement away from their natural positions on the further left and towards the centre during their time as central opposition figures.

4.4.2 Intra-Party Factions

Traditional methods for controlling for ideological effects when assessing legislator behaviour in a Westminster system such as the HoC may be to use proxy variables such as party faction membership or nominal vote choice. Although these can be suitable measures in some cases, they are relatively simplistic and lose the granularity of discrimination between individual MPs, particularly when estimating within the same party. Commonly used variables to gauge a general left/right position along the

[12]Conservative MP Stuart Andrew had the third highest standard error, but also had by far the lowest number of responses in the expert survey and the highest variance, suggesting he was not particularly well-known or easy to place.

British political axis might be an MP's faction membership within their party, their Brexit stance, or their position on a social policy such as abortion or same-sex marriage to capture a liberal/conservative dimension. To demonstrate the within-party discriminatory power of the CA model, MP ideal points are plotted along these three dimensions, split between the two major political parties. Descriptions of these variables are as follows:

Party Faction Membership:

Both parties contain multiple internal factions and pressure groups which can help to distinguish between different ideological spaces within them. The most prominent faction within the Conservative party is the European Research Group (ERG), a hard-right caucus of MPs set up in 1993 to act as the voice of Euroscepticism within the party. Whilst there is no official membership list, support for this faction can be inferred through subscriptions paid to the ERG by MPs since 2016. 64 Conservative MPs can be identified as either current or former members of the ERG.[13] The most prominent faction within the Labour party is the Socialist Campaign Group (SCG), a hard-left socialist contingent of MPs founded in 1982. There are currently 33 official members.[14]

Brexit Stance:

The stance of Conservative MPs on the UK's membership of the EU can be sourced from various statements made by MPs on their websites, in the media, and on Twitter. Data for MPs elected in the 2015-2017 parliament are derived from a dataset collated by Cygan et al. (2021). This provides a Brexit stance [Remain/Leave] for 200 Conservative MPs in the current HoC. Unfortunately, cases where Labour MPs declared support for leave in this dataset are extremely rare and results in only 1 instance of support for leave among Labour MPs in the current HoC. Instead, Brexit stance for Labour MPs is inferred through their support for or rebellion against Boris Johnson's European Union (Future Relationships) Bill in December 2020 which allowed for the ratification of Brexit withdrawal.[15] This is an imperfect method as the support for or rebellion against this bill does not necessarily directly reflect remain/leave sentiment. However, given that supporting this bill was enforced with a three-line whip by Labour, rebellion was a high-cost decision and so can still be informative.

[13] List can be found here: https://bylinetimes.com/2021/08/23/a-party-within-a-party-calls-for-investigation-into-european-research-group-paid-quarter-of-a-million-in-taxpayer-cash-since-brexit/

[14] List can be found through the SCG's official Twitter account: https://Twitter.com/i/lists/1220096981848162305

[15] Vote data can be found through the Public Whip archive: https://www.publicwhip.org.uk/division.php?date=2020-12-30&house=commons&number=190

Abortion Stance:

To act as a binary proxy for socially liberal/conservative, MPs are also divided up by their stance on abortion. MPs are categorised as either pro-choice or pro-life depending on whether they voted in support of or against Clause 11 of the Public Order Bill 2022 making it an offence to interfere with access to abortion services.[16] There were no recorded votes against this bill from the Labour MPs but there was a 19% abstention rate. Again, using division votes as ideological proxies are imperfect (providing the motivation for this research), but can still be informative.

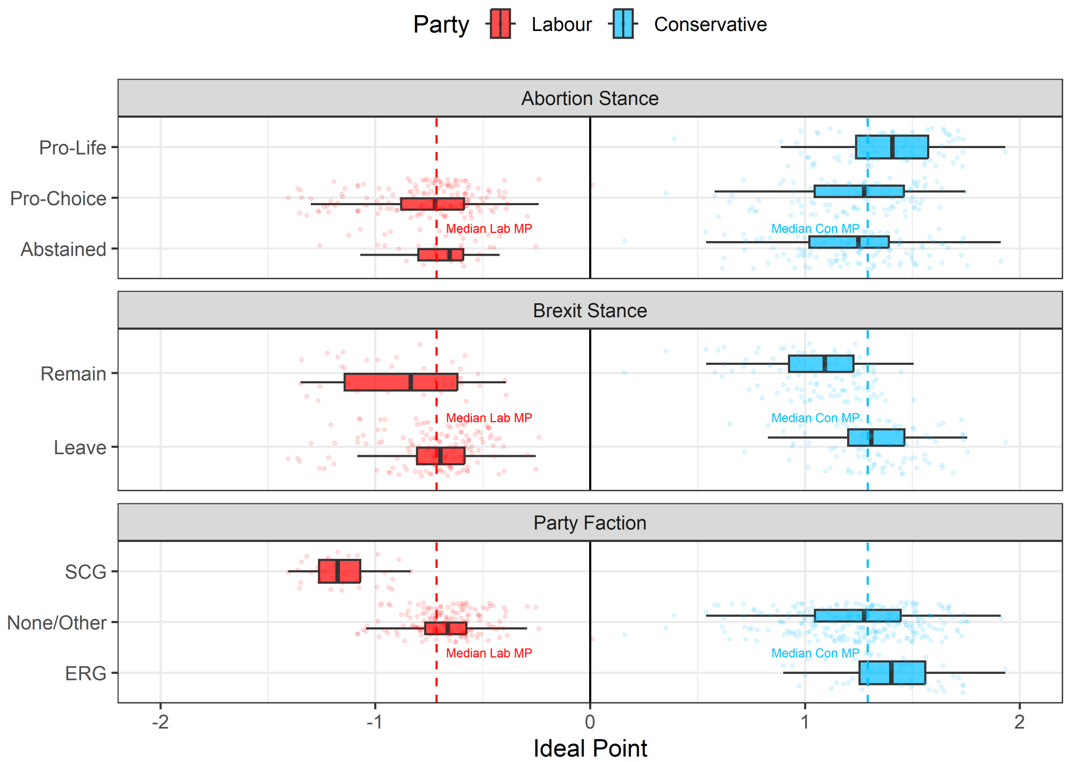


FIGURE 4.4: Jittered boxplots of MPs from the Conservative (right-side) and Labour (left-side) parties grouped by their ideological factions, inferred through voting patterns, declarations of support, and party subgroup membership. All NAs were removed.

The distributions in **Figure 4.4** add credence to the fact that these ideal point estimates reasonably approximate the left/right ideological spectrum of British politics. In both cases, party faction membership, on average, places these MPs notably further to the right and to the left of the median position of their parties respectively. As for abortion stance, the median ideal point of pro-choice MPs is to the left of the average pro-life MP in both parties (assuming abstaining MPs in the Labour party are pro-life). The model also places Conservative MPs who declared their support for remain towards the left of the party, and this effect is the same to a lesser degree in the case of Labour

[16]Vote data can be found through the Public Whip archive: https://www.publicwhip.org.uk/division.php?date=2022-10-18&house=commons&number=62

MPs who rebelled against Johnson's Brexit Bill. However, it is also evident that there is significant range across these distributions, emphasising the importance of having a unique ideal point for each individual MP.

4.5 Modelling Intra-Party Competition - Leadership Endorsements

With confirmation that the model ideal point estimates have a high degree of accuracy, the final section of this paper will seek to model the ideological component to candidate endorsement in the September 2022 Conservative party leadership contest. Following a string of controversies and party scandals during Boris Johnson's premiership, along with multiple by-election defeats and poor showing in the 2022 local elections, the Prime Minister announced his resignation on July 7th 2022. This sparked a party leadership contest between Conservative MPs that ran from July to September, with Liz Truss eventually elected leader of the party and Prime Minister of Great Britain on September 5th. However, after the announcement of a controversial and widely criticised 'mini-budget' proposing widescale tax cuts that led to a sharp fall in the value of the pound, Liz Truss resigned as Prime Minister on October 20th, after just 45 days in office. This subsequently sparked a second party leadership contest, albeit far shorter, with Rishi Sunak elected Prime Minister unopposed on October 24th. Given the fact that the October contest ran for less than a week and eventually went unopposed, this paper will not focus on this contest.

The academic literature surrounding the complex dynamics of party leadership contests in the UK is extensive (Crines et al., 2018; Denham and Dorey, 2018; Heppell, 2008, 2010, 2022; Jeffery et al., 2018, 2022, 2023; Stark, 1996) To adequately understand any one of them would require a systemic examination; one that would be far beyond the scope of this paper. Rather, this paper treats the September Conservative leadership election as an ideal context in which to demonstrate the potential applications of these ideal points to legislative study in Westminster systems such as the HoC. Therefore, the modelling strategy here is relatively simple, focusing on the relationship between an MP's ideal point and their candidate endorsement as opposed to offering a comprehensive explanation of endorsement choice as a whole.

4.5.1 July - September 2022 Party Leadership Contest

The first Conservative leadership election was triggered by Johnson on July 7th 2022. Voting opened on July 13th and was initially contested by 8 MPs, travelling through five separate MP ballots, eventually narrowing down to two final candidates: Liz

Truss and Rishi Sunak (See: **Table 4.4**). The resulting members' ballot elected Truss as leader on September 5th by a margin of 57%/43%.[17]

	Round 1		**Round 2**		**Round 3**		**Round 4**		**Round 5**	
	N	**%**	**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**
Rishi Sunak	88	25%	101	28%	115	32%	118	33%	137	38%
Liz Truss	50	14%	64	18%	71	20%	86	24%	113	32%
Penny Mordaunt	67	19%	83	23%	82	23%	92	26%	105	29%
Kemi Badenoch	40	11%	49	14%	58	16%	59	17%	-	-
Tom Tugendhat	37	10%	32	9%	31	9%	-	-	-	-
Suella Braverman	32	9%	27	8%	-	-	-	-	-	-
Nadhim Zahawi	25	7%	-	-	-	-	-	-	-	-
Jeremy Hunt	18	5%	-	-	-	-	-	-	-	-
Votes Cast	**357**	**100%**	**356**	**99%**	**357**	**100%**	**355**	**99%**	**355**	**99%**

TABLE 4.4: Table of Conservative Parliamentary Eliminative Ballots.

Leadership nominations are conducted through a closed ballot. Therefore, official leadership nominations cannot be obtained. Instead, endorsements are inferred through open declarations of support from MPs, taken from the Conservative Home official website.[18]. Unfortunately, the reliance on open declarations of support means that the nomination data used in the analysis is incomplete given that many MPs did not declare who they would be voting for. Nonetheless, of the 357 Conservative MPs eligible to vote at the time of the contest, leadership endorsements were ascertained for 319 (89%) in the first round of voting, leaving 38 undeclared. Of those 319 MPs, 278 had Twitter accounts from which ideal points could be estimated.

4.5.1.1 Initial Endorsements - First Round

Ideal point distributions of candidate endorsers from the first round of the leadership contest would suggest that an ideological component to nomination behaviour was present (See: **Figure 4.5**). Notable further right MPs Kemi Badenoch and Suella Braverman appear to have drawn their support from MPs similarly to the further right of the Conservative party, as does eventual winner Liz Truss. Conversely, Truss' two primary competitors, Sunak and Mordaunt, look to have drawn their support from closer to the centre of the party in the early stages of the contest.

[17] Official leadership nomination data obtained from https://www.politico.eu/conservative-leadership-election-2022/.

[18] Accessed here: https://conservativehome.com/2022/08/25/next-tory-leader-whos-backing-whom-our-working-list/

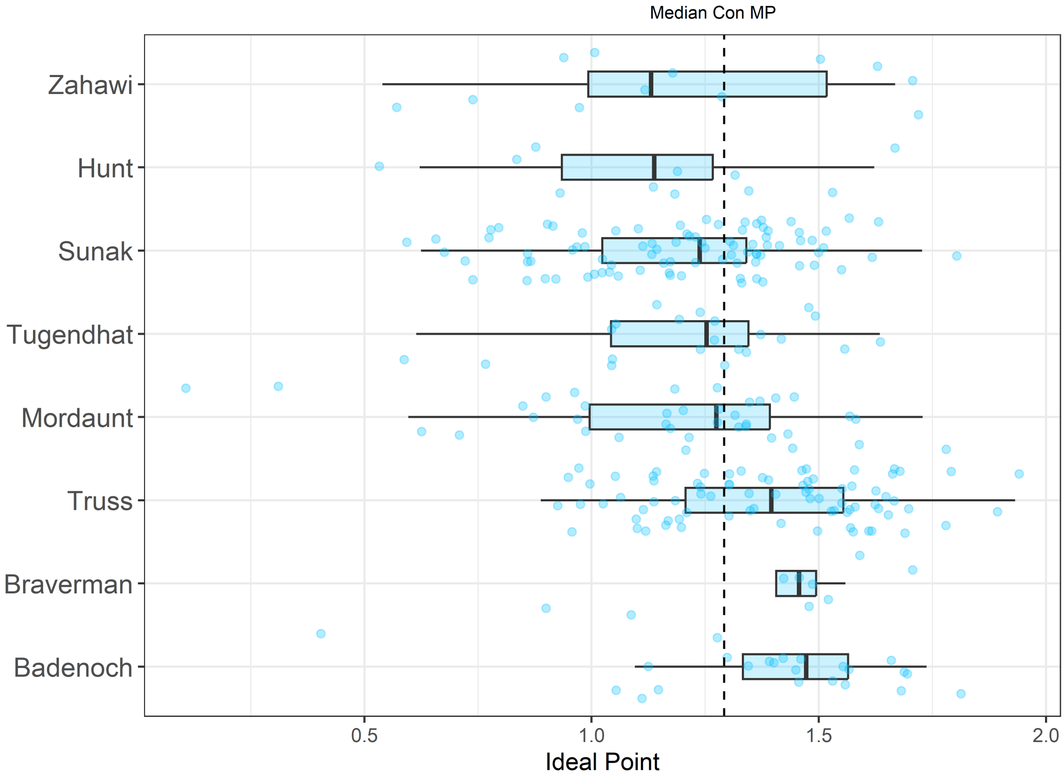


FIGURE 4.5: Jitter boxplot distributions of the ideal points of MPs who publicly endorsed each of the 8 initial Conservative Party leadership candidates.

4.5.1.2 Final Endorsements - Membership Round

Over five rounds of MP ballots, 6 candidates were eliminated, leaving Liz Truss and Rishi Sunak to contest the final members' ballot. As candidates were eliminated, many MPs switched their initial endorsements to other remaining candidates. This provides a reasonably balanced sample of endorsements split between the two final candidates. Endorsements could be obtained for 245 MPs with Twitter accounts in the final round, leaving 112 undeclared or with no ideal point.

As is illustrated in **Figure 4.6**, the median ideal point of supporters for both Rishi Sunak and Liz Truss do not change substantially from their scores in the first round of endorsements. Both Sunak's (**1.24** - **1.23**) and Truss' (**1.4** - **1.39**) lower slightly, suggesting a mild moderating effect as they gained more endorsements from MPs across the party. What remains clear in both instances is that Liz Truss appears to have drawn more support from MPs further to the right of Conservative Party.

4.5.2 Modelling Strategy

To formally model the ideological component to leadership endorsements, a simple binary logistic regression will be used where the model predicts support for eventual

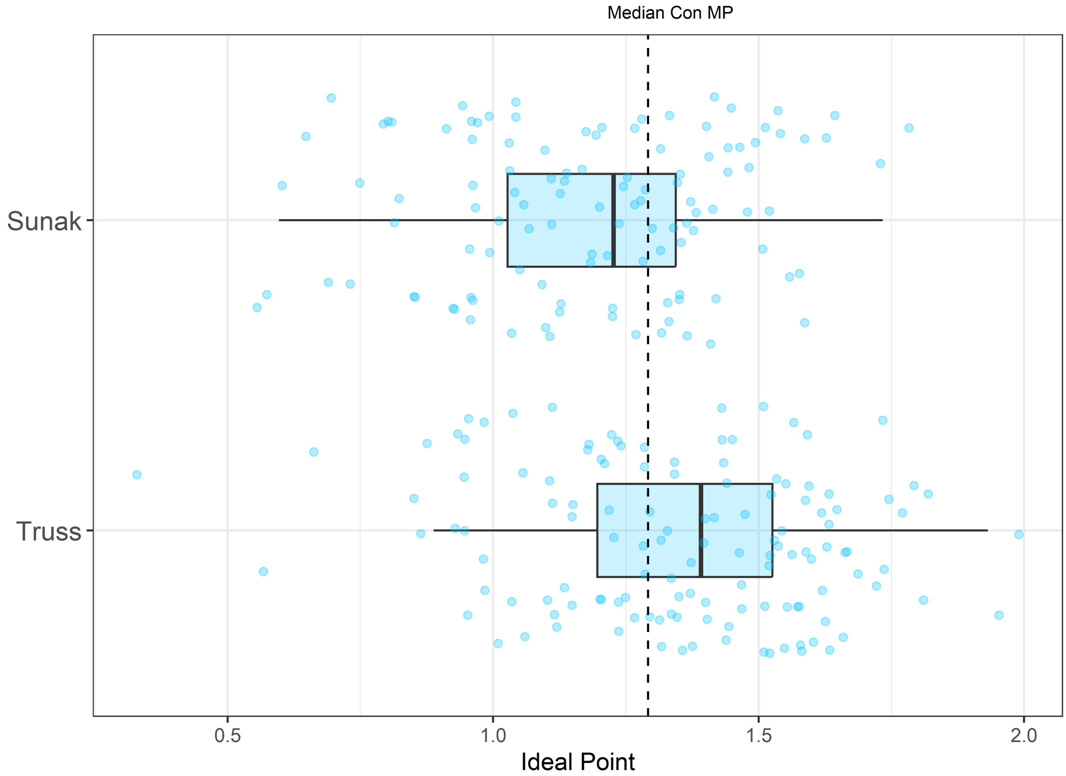


FIGURE 4.6: Jitter boxplot distributions of the ideal points of MPs who publicly endorsed each of the two final Conservative Party leadership candidates.

winner, Truss, relative to the unsuccessful candidate, Sunak. A recent paper by Jeffery et al. (2023) examined the candidate preferences of Conservative MPs in the September 2022 contest, exploring a wide range of personal, political and ideological factors that may have influenced the vote. Their study found that, contrary to common media narrative, loyalty to previous leader Boris Johnson was *not* a significant factor in determining support for Truss. However, they do note the importance of ideological factors in this contest. In place of the ideological proxies they use in their model such as ERG membership and support for the levelling-up agenda, this model will strictly use an MP's ideal point. Beyond that, this model will also control for a simple set of demographic and political variables, in-keeping with relevant factors also present in previous studies by Jeffery et al. (2018, 2022) on the voting motivations of Conservative MPs in the 2016 and 2019 party leadership contests.

4.5.2.1 Model Variables

The four demographic variables chosen to control for are an MP's age, gender, ethnicity, and educational background. Social variables are primarily sourced from the *Representative Audit of Britain* (RAB) - a dataset collating together research on parliamentary candidates in Great Britain - drawing on demographic data for MPs elected as part of the 2019 parliamentary cohort.[19] This includes their gender, ethnicity, year of birth, their secondary school type [State, Grammar, Private], and their university type [None, Non-Russell Group, Russell Group, Oxbridge]. For a handful of MPs elected after 2019 via by-elections, these data are manually sourced from publicly available online sources. Three variables relating to an MP's political background are selected for inclusion in the model: election cohort, ministerial background, and size of constituency majority. Election cohort data can also be derived from the RAB dataset and is categorised into five distinct groups [Pre-1997, 1997-2010, 2010-2015, 2015-2019, 2019]. Ministerial background is sourced manually from publicly available online sources and categorised into three groups: [Never (been a minister), Former, Current (at time of the contest)]. Constituency majority size is also sourced manually from publicly available online sources and is treated as a continuous variable (%). Summary statistics of these variables for MPs who endorsed each candidate can be found in *Appendix A.4*.

4.5.2.2 Model Results

Model coefficients are the log odds of an MP endorsing the eventual winner of the September election contest, Liz Truss, over the runner-up, Rishi Sunak. Three separate models are estimated: the first modelling only an MP's ideal point estimate as a predictor, the second also controlling for social variables, and the third also controlling for political variables. Model results can be found in **Table 4.5**.

Model results would indicate that an MP's ideal point was a statistically significant predictor of endorsement choice, with each unit increase further to the right associated with a higher likelihood of endorsement for Truss over Sunak. This would remain consistent with the descriptive statistics which illustrated that Truss had ostensibly drawn her support from the further right of the Conservative Party, whilst Sunak had drawn his primarily from closer to the centre. This relationship remains statistically significant when controlling for both social and political variables. These results would confirm that ideological position was a significant component of candidate endorsement in the September Conservative leadership election and that the further to the right of the party an MP leaned, the increasingly more likely they were to favour Truss over Sunak. The model also demonstrates that a gendered component to candidate endorsement was present, with female MPs significantly more likely to endorse Truss over Sunak.

[19] The official citation for this dataset is the *Representative Audit of Britain*, ESRC, Grant Number ES/L016508/1, led by Prof. Rosie Campbell, Prof. Jennifer Hudson and Dr. Wolfgang Rudig.

	Support for Truss		
	Model 1	Model 2	Model 3
Ideal Point	**12.10*** **	**11.82*** **	**18.27*** **
Social Variables			
Gender (Female)		**3.75** **	**3.81** **
Ethnicity (Minority)		1.53	1.42
Year of Birth		1	1.01
School Type (relative to private)			
Grammar		0.81	0.84
State		0.93	0.98
University Type (relative to Oxbridge)			
Russell Group		1.79	1.93
Non-Russell Group		1.43	1.50
None		4.89	5.67
Political Variables			
Cohort (relative to pre-1997)			
1997 - 2010			0.48
2010 - 2015			0.84
2015 - 2019			0.71
2019			0.42
Minister (relative to current ministers)			
Former Minister			0.73
Never			0.57
Majority (%)			0.99
AIC	**318.73**	**249.33**	**257.6**
Psuedo R2	**0.07**	**0.14**	**0.16**
N	**245**	**192**	**192**
Signif. Codes: ≤ 0.05 *, ≤ 0.01 **, ≤ 0.001 ***			

TABLE 4.5: Logistic Regression Model Coefficients - Predicting Support for Truss Relative to Sunak

4.6 Conclusions

The overall purpose of this research is to demonstrate the suitability of Twitter networks for ideal point estimation in restrictive parliamentary cultures such as those found in Westminster systems like the UK. Scholars in political science have long grappled with the difficulty of estimating ideal points for legislators in these types of systems, failing to adequately replicate the success of roll-call estimation strategies dominant in countries such as the US. Consequently, researchers have turned to a variety of alternative methods to fill this gap including the use of surveys and text-based scaling or, in the British case, the use of EDMs as a replacement for roll-call data. Work by Kellermann (2012) generated the most valid ideal points estimates using EDMs to date and, for a time, suggested these could be an adequate replacement for roll-call data in the UK. However, with the decline of participation in EDMs by MPs in the last few years, the need for an appropriate method of ideal point estimation still remains. Building on original work by Barberá (2015), this paper set out to illustrate the way in which Twitter follower networks can be successfully exploited to generate valid ideal point estimates of MPs in the UK and provide motivation to scholars studying other Westminster-style legislatives.

The results of this research have demonstrated two things: 1) Twitter follower networks can be successfully used to estimate valid ideal points of MPs in the UK, and 2) ideology was a significant predictor of support for Liz Truss in the September 2022 Conservative Party leadership contest. In both cases, existing research has already explored these areas in detail. The value of this particular research lies in the application of one to the other, where a matured social network of British parliamentarians has been leveraged to predict endorsement patterns in an especially salient leadership contest. Analysing a combined follower base of over 34 million accounts, ideal points for 591 MPs were estimated through correspondence analysis and validated against ideological estimates of 30 MPs provided by 70 experts in the field of British politics. These expert estimates confirmed that these ideal points have a high degree of both between and within-party accuracy. Application of these ideal points to assess the ideological component of the September 2022 Conservative Party leadership contest exhibits one of the many ways in which these estimates could be used to improve our understanding of legislative politics in the UK.

Another unique feature of the UK parliamentary system is the presence of nationalist parties in Scotland, Wales and Northern Ireland. Some of these parties such as the SNP present themselves as 'big tent' parties, having members that cover a broad spectrum of beliefs and/or having a nationalist component that transcends the traditional left/right of British politics. Positioning these parties and their MPs along the same left/right axis as non-nationalist parties and MPs may be difficult to do if their nationalist preferences overwhelm their ideological ones. Estimating ideal points through Twitter networks has the distinct benefit of circumventing this issue. The ability to scale multiple dimensions beyond ideology is another unique strength of this particular methodology, potentially allowing scholars to explore questions of dimensionality beyond the conventional left/right axis which may well be idiosyncratic to each country or legislative.

The applications of this method both within and beyond the British case are extensive, owing to the flexibility of the data source that the estimates are derived from. Generating estimates from network data found on social media platforms means a data source where legislators both past and present exist in the same digital space, and where networks are ever-changing. This provides the potential for ideal points to be dynamically tracked and updated over time, and between legislators who no longer sit in the HoC or who sit in other legislative bodies around the world. Moreover, such a method does not necessarily constrain itself to legislators but could also have applications to wider political elites such as media organisations, journalists, academics and political commentators. The nature of correspondence analysis also allows for the estimation of ideal points for row variables as well as the columns within an adjacency matrix, thus generating estimates for followers as well as elite accounts. This opens up additional avenues of research on the ideal point distributions of ordinary accounts on social media platforms.[20]

At the time of writing this paper, Twitter - now officially known as X - announced significant changes to access of the API which was previously used to freely extract the data used in this research. These changes have financially restricted access to the platform's data thus reducing the ability for many researchers to extract data again in the future to update these estimates. The network data used to generate the estimates for this paper were extracted at the high point of the Twitter research life cycle before these changes were enacted, meaning these social networks reflect peak maturity before restrictions took place. Therefore, this study may be one of the last to analyse such matured social networks of UK political elites for some time and makes available a set of validated ideal point estimates for 591 sitting MPs in the House of Commons (as of 22/08/2022), as well as for almost half a million ordinary users and wider political elites. Whilst the future of research on Twitter remains uncertain, alternative avenues for social media study will no doubt present themselves again in the future and new platforms may emerge as networking spaces for political elites. Whatever may come, this research has illustrated the power of harnessing social media data for improving our understanding of legislator behaviour in Westminster systems, and how the growing presence of political elites online can be exploited to the benefit of political study.

[20] *Appendix A.5* for more details on ordinary user ideal points.

Chapter 5

From Networks to Newsrooms: Assessing Media Representation Using Elite Twitter Networks - Paper 2

Ideological representation on TV broadcasters can be important for public trust in the media. However, objective measures of ideology can be difficult to operationalise. This article proposes a novel technique for assessing the political representation of broadcasters by estimating the ideological position of the guests they select to appear on their political programmes using their social media networks. In a case study of the United Kingdom (UK), a nation with a long and proud tradition of public broadcasting, left-right ideological estimates are derived from the Twitter follow networks of guests who appeared on the flagship political programmes of its six major broadcasters. These estimates are formally validated through a survey of the general British public (R^2 = 0.61). Assessing guest selection across seven political programmes between 2022 and 2024, results would indicate that each show selected guests significantly to the right of the average elite Twitter users more generally. The purpose of this research is to highlight the value of using social media networks to better understand elite representation in politics and the media.

5.1 Introduction

Media representation can be a difficult concept to operationalise. This is especially the case when combined with a latent variable like ideology. Over the years, several studies have sought to investigate the ideological representation of new coverage, primarily to test for the presence of media bias or impartiality. However, as is covered extensively by Hopmann and colleagues, such scholarly attention to the subject has not led to a common understanding of what political balance actually means or what unbiased news coverage would look like (Hopmann et al., 2012, p.243). The extensive variation in definitions of political balance and the units of analysis to focus on within the media's coverage of news has resulted in a myriad of different operationalisations.

This article seeks to depart from conventional studies of media representation by using a novel method of ideology estimation via elite social media networks to assess political representation on TV broadcasters. It adopts an actor-focused visibility approach, assessing the descriptive representation of the conventional left-right political axis through the guests selected to appear on the political news programmes of major TV broadcasters. While this approach thus neglects the substantive content of the programmes themselves, the strong link between descriptive and substantive representation is well-established in the political literature (e.g., Hahn, 2024; Mansbridge, 1999; Reher, 2022; Reynolds, 2013; Sobolewska et al., 2018). Public perceptions of media representation and trust in the media to report the news objectively can have a significant impact on the attitudes of the general public (Eberl et al., 2017; Ho et al., 2011; Hoffman and Wallach, 2007; Morris, 2007).

Owing to its long and proud tradition of public service broadcasting, and the fact that the BBC in particular is widely viewed as the model PSB around the world (Guardian Correspondents, 2020; Gunter, 2024), the United Kingdom (UK) is selected as a "gold standard" case study. Despite the growing controversy around the regulation of media broadcasting in the UK, and of the lasting relevancy of PSBs in their current format, studies continue to show that the BBC, ITV, and Channel 4 remain three of the most trusted news sources in the country (Nielsen et al., 2023; Smith, 2023). Even in an international hyper-partisan news environment like the United States (US), BBC News still ranks as one of the most trusted sources of news (Sanders, 2022). This highlights the continued value that ordinary citizens place in PSBs as impartial news sources, as well as the global influence of the BBC.

Drawing on existing research and openly available data from Gaughan (2024), left-right ideological estimates of British political and media elites are derived from the Twitter/X follower networks of UK parliamentarians. An original survey is commissioned through polling company *YouGov* for this study, ascertaining left-right ideology estimates of a subset of political and media actors and organisations from the general British public to formally validate the ideology estimates elicited via social media.

Guest lists of the flagship political programmes for the UK's four primary PSB TV channels (BBC One, BBC Two, ITV One, Channel 4) and two major non-PSB news channels (Sky News, GB News) are collated from a number of sources between the 01/01/2022 and 01/01/2024. Guests are matched to their ideology estimates via their Twitter/X usernames and the ideological representativeness of seven political programmes is descriptively examined. The purpose of this study is to demonstrate a novel application of social media network data to assess media representation in lieu of manual coding strategies that can be time consuming, hard to replicate, and subject to coder reliability. Additionally, the results of this research can also lend evidence to the growing debate in the UK around the political representation of its major TV broadcasters.

5.1.1 Existing Approaches to Assessing Media Representation

The majority of studies that looked to assess media representation have done so to find evidence of media bias or political imbalance to one side or the other. However, an extensive review of the body of studies relating to media representation in Western democracies by Hopmann et al. (2012) highlighted the difficulty in standardising concepts such as "political balance" or "impartiality". Across the literature they find that definitions of politically balanced news in Western democracies tend to vary based on either the structure of the political system (Semetko and Canel, 1997; Wilke and Reinemann, 2006; Zeldes et al., 2008), country-specific regulations (Hanretty, 2007; Semetko, 2003), or media culture and routines (McQuail, 1992; Schoenbach et al., 2001).

Norris (2009, p.336) acknowledges the challenging state of comparative political communication, noting that studies on media representation using content analysis are so different in their methodological approach that replicability is difficult, and that normative interpretations of representation are so varied from context to context that comparisons are hard to make. Nonetheless, the work of Hoppmann and colleagues was able to identify three key areas of focus that media representation studies typically tend to focus on (Hopmann et al., 2012, pp.8-10): visibility (of actors), favourability/evaluation (towards/of actors), and issue coverage. Most commonly, manual thematic coding is the method of choice for media and communication scholars, coding political affiliation of partisan actors who appear in news stories (Buyens and Van Aelst, 2022; Lewis and Cushion, 2019; Sheafer and Weimann, 2018), how much time they afforded or the favourability in which they are presented (Hopmann et al., 2011; Hughes et al., 2023; Zeldes et al., 2008), or by what (or how) issues are covered (Covert and Wasburn, 2007; Maurer et al., 2022; Wahl-Jorgensen et al., 2017).

Some studies have departed from conventional methods in recent years. In particular, contemporary technological developments have made the study of media representation more popular in computer science, where automated identification of ideological leaning have grew in popularity (See Hamborg et al. (2019) and Rodrigo-Ginés et al. (2024) for comprehensive overviews). This can help to significantly increase both the speed and scale of assessing media representation compared with manual coding and can help to test multiple areas of media leaning on a wide scale. This includes event selection (Bourgeois et al., 2018; Saez-Trumper et al., 2013), labelling and word choice (Agirre et al., 2016; Papacharissi and de Fatima Oliveira, 2008), ideological spin/slant (Budak et al., 2016; Park et al., 2011), and for detecting fake or misleading information (Shu et al., 2017; Lin et al., 2019). This has also made the availability of online media bias and fact checking websites more common (e.g: Snopes, Politifact, Media Bias/Fact Check). Automated methods of media representation checking do provide the additional benefit of simplifying the process of standardisation. However, as is acknowledged by Hamborg et al. (2019), the study of media representation in the field of computer science is still in its infancy. Many of these methods are limited in their ability to adequately capture the complexity of media representation as a concept.

5.1.2 Estimating Left-Right Position Through Social Media Networks

This research looks to contribute to this space with a unique method of its own, estimating left-right positions of elite actors through their social media networks. Measuring ideological position is an intrinsically difficult task, both because it is a latent variable and because the very concept itself is contentious. The exact meaning of the term *"ideology"* has been long debated in the field of political science, spawning a wide variety of definitions, terminologies, and conceptions of dimensionality (Eagleton, 1991; Gerring, 1997; McLellan, 1986). The list of methods for operationalising it is equally as extensive, ranging from more traditional forms such as survey-based (Huber and Inglehart, 1984; Jolly et al., 2022; Mair and Castles, 1984) and roll-call analysis (Clinton et al., 2004; Poole and Rosenthal, 1985; Rosenthal and Voeten, 2004) to more contemporary techniques which leverage advanced computational tools such as machine learning for cluster analysis (Spirling and Quinn, 2010), natural language processing (Grimmer and Stewart, 2013; Németh, 2023; Slapin and Proksch, 2008), and large-scale network analysis (Barberá et al., 2015; Bonica, 2014).Along with advances in potential methodologies for measuring ideology has come a growth in the variety of data sources for estimates to be derived, moving from a reliance on conventional sources such as legislative transcripts and roll-call votes to media articles, financial campaign contributions, and social media data.

An array of studies have confirmed the existence of homophilic network structures on Twitter/X and the tendency for ideological clusters to form around both followership and content engagement/sharing (Bail et al., 2018; Barberá et al., 2015; Colleoni et al., 2014; Conover et al., 2011; Gaughan, 2024). These ideologically structured networks can be directly leveraged for ideology estimation of elites. One of the first papers to do so was Barberá (2015) applying the general principles of traditional spatial voting models to the follower networks of US legislators on Twitter/X to generate *ideal points*, scaling the left-right ideological dimension through the relative 'distance' between their individual networks. In the same way that spatial models of voting assume ideal points can be represented by the (dis)similarity between each legislator's voting record (Enelow and Hinich, 1984), Barbera's *Spatial Following* model assumes the same for each legislator's follower base.

This principle assumption is guided by the expectation that legislators and other political elites closer to one another in the latent ideological space will have a higher probability of sharing similar follower networks, where politically engaged users are more likely to follow other political accounts they perceive to be closer to their own ideal point.[1] This model was applied more recently to the British case by Gaughan (2024) generating left-right estimates of UK MPs to model candidate endorsement in the Sept 2022 Conservative Party leadership contest. Given the proven effectiveness of such a method for estimating ideological position of British elites, this article proposes a similar strategy for measuring media representation.

5.1.3 Case Study: United Kingdom

"Public service broadcasting (PSB) has a long and proud tradition in the UK, delivering impartial and trusted news, UK-originated programmes and distinctive content." (Ofcom, nd)

The tradition of public service broadcasting in the UK has roots stretching back as far as the early 1920s, establishing its first and largest PSB by a Royal Charter in 1922: the British Broadcasting Company (BBC). Self-proclaimed as, *"the world's leading public service broadcaster"* (BBC, ndb), the BBC is principally funded through a legally enforceable licence fee paid by UK households and broadcasts to weekly global audiences of over 300 million people (BBC, nda). Considered a *"giant"* in the global broadcasting market and frequently cited by foreign politicians as a model to be imitated (Guardian Correspondents, 2020), it played a significant role in the

[1] A more detailed description of the assumptions and estimation strategy for the *Spatial Following* model can be found in the original paper by (Barberá, 2015, pp.77-79).

advancement of public and semi-public broadcasting corporations around the world including ABC in Australia, CBC in Canada, and NZBC in New Zealand (Potter, 2012). It was subsequently followed by the creation of the UK's first commercial broadcaster, ITV (STV in Scotland), in 1955, then Channel 4 (and its Welsh counterpart S4C) in 1982, and finally its newest and smallest public broadcaster, Channel 5, in 1997. Unlike the BBC, ITV, Channel 4 and Channel 5 are independent broadcasters that do not receive any public funding, relying on self-generated commercial revenue. Crucially, however, in return for the licence to broadcast in the UK, these TV networks are still required to operate under a public service remit, formally qualifying them as PSBs.[2] Although this remit has, at times, been poorly defined and misunderstood, under the Communications Act 2003 the main terrestrial TV channel for each PSB – BBC One, BBC Two, ITV One, Channel 4, S4C, Channel 5 – ,*"must deliver programmes and services which cover a wide range of subjects and which meet the needs and interests of many different audiences"* (Ofcom, 2004, p.3). There is a general expectation that these broadcasters deliver high-quality programming that educates and entertains, and that adequately represents the lives of different regions and communities across the country.

Along with the UKs four major PSBs, the UK also launched its first purely commercial non-PSB in the early 1990s, *Sky TV*. Historically, the UK has legislated for an impartiality regime in television and radio broadcasting, enforced by an independent regulatory and competition authority (Barnett and Petley, 2023, p.18). Since the introduction of the Communications Act 2003, the government-approved regulator has been the Office of Communications, commonly known as *Ofcom*. Embedded in Ofcom's Broadcasting Code: Section Five is the rule of "due impartiality", the fundamental principle of which is to, *"ensure that the news, in whatever form, is reported with due accuracy and presented with due impartiality"* (Ofcom, 2021). This principle is strictly applied to *all* the news broadcasters in the UK, PSB or otherwise, and is designed to prevent against the type of hyper-partisan and polarised media landscape that is characteristic of the United States (Jurkowitz et al., 2020). Given that most studies on ideological representation in news coverage originate from the USA – where there is a far weaker public broadcasting culture - the UK's tightly regulated broadcasting landscape and strong tradition of PSB makes for an interesting case study.

Moreover, the arrival of two new ostensibly right-wing TV news channels to the UK – GB News and TalkTV – since 2021 has shined a growing spotlight on Ofcom's willingness and ability to adequately enforce its own standards (Barnett and Petley, 2023; Waterston, 2023a). Concerns about a potential "Foxification" of news coverage in the UK has been on the rise now for the last decade (Cushion and Lewis, 2009; Petley, 2020). Alongside the emergence of contemporary partisan news channels such

[2]PSB, whilst not especially well-defined, typically refers to forms of broadcasting designed for non-profit public benefit and not to serve commercial interests (Booth, 2020, p.325)

as GB News and TalkTV, traditional PSBs such as the BBC, ITV and Channel 4 are increasingly facing accusations of political bias. PSBs are inherently easy targets for accusations of left-wing bias due to state-mandated regulation over broadcasting content, while others argue that, in their effort to maintain impartiality, PSBs inevitably morph into establishment mouthpieces that simply uphold the right-wing status quo (Jones, 2014; Mills, 2020).

Since 2010, a number of studies have assessed the representativeness of UK broadcasters (primarily the BBC) in their coverage of contemporary economic and political issues. Research by Cushion et al. (2018) found that during the 2015 UK general election campaign, broadcasters followed a policy agenda more closely aligned to that of right-wing newspapers, while Hughes et al. (2023, pp.1722-1723) found that during the 2019 and 2020 UK and US election campaigns, Johnson and Trump received significantly more coverage than that of their challengers, Corbyn and Biden. In a similar vein, a study into media representation during the 2016 EU referendum campaign illustrated that, while UK broadcasters were generally even-handed in their coverage of the Leave and Remain campaigns, they significantly favoured right-wing sources, marginalising the left-wing case for both choices (Cushion and Lewis, 2017).

More recently, research by Morani et al. (2022) found that during the height of the COVID-19 pandemic UK broadcasters still relied heavily on political sources, with health and scientific experts receiving limited coverage. Among these political sources, the majority were Conservative MPs while voices of the main opposition party, Labour, received limited airtime. Ongoing analysis of descriptive representation specifically on the BBC's flagship topical debate programme, Question Time, has found a significant over-representation of right wing voices over the last nine years (Walsh, 2024). Conversely, recent research by Brandenburg et al. (2024) found no clear evidence of systematic bias in panel selection on the programme, though they did find a disproportionate over-representation of voices from privileged educational backgrounds.

5.1.4 Hypotheses

Whilst the results of these studies are not necessarily conclusive, a general pattern of political imbalance is evident. Although there is a strong case to be made that, as the incumbent political party from 2010 to 2024, Conservative MPs were expectedly more relevant and more newsworthy (Hopmann et al., 2011), this does not necessarily account for the over-representation of non-political right-wing sources. Moreover, a number of studies have shown that, when comparing with political balance of BBC news sources before 2010 (when the Labour Party were incumbents), the UK's main broadcaster has demonstrated a disproportionate bias towards the political right post-2010 (Berry, 2016; Lewis and Cushion, 2019; Wahl-Jorgensen et al., 2017).

When striving to empirically test ideological representation on UK broadcasters, this paper deliberately steers clear of normative assessments of political bias or imbalance. The methodology employed in this study is not comprehensive enough to make such a judgment. Instead, it looks to test whether the selection of guests on each programme differs significantly from the general pool of elites that were active on Twitter, and to assess the direction of this difference. The comparison with general Twitter elites is informative as this is general pool of users that one would expect guests to have been drawn from (86% of guests were elite accounts). It also looks to assess whether siginfcant differences exist *between* each programme.

The first exploratory question this paper looks to answer is as follows:

Research Question 1: Does the ideological distribution of guests who appear on each programme differ significantly from the ideological distribution of general Twitter elites?

If the ideological selection of guests on each programme *is* significantly different from general Twitter elites, this study then expects the following:

H1: *Each political programme will demonstrate an average ideological distribution of guests further to the right of general Twitter elites.*

This is informed by the fact that the right-wing Conservative Party were the incumbents during the observation period 01/01/2022 to 01/01/2024. Thus, this increases the probability of selecting guests further to the right.

The second question this paper asks is as follows:

Research Question 2: Does the ideological distribution of guests who appear on each programme differ significantly from one another?

When considering differences between each programme, there are a number of factors that may influence why certain guests are selected to appear on TV more often than others such as their media viability (King, 2002) or their relevancy (which would account for the incumbency bonus) but these factors should, in theory, influence all media outlets evenly (Eberl et al., 2017; Van Dalen, 2012). Thus, this study also expects the following:

H2a: *There will be no statistically significant difference in the ideological distribution of guests on each of the seven programmes*

However, if the distributions of guest selection do differ significantly from one another:

H2b: *The ideological distribution of guests on non-PSB programmes will be further to the right than PSB programmes..*

In the UK, all broadcast news is expected to remain politically impartial. However, the implicit public service remit that PSBs carry informs the expectation that if differences in guest selection *were* to exist between programmes, the incumbency effect that motivates Hypothesis 1 will be stronger on non-PSBs.

5.2 Data & Methods

5.2.1 Ideal Points of UK Elites

To test these hypotheses, left-right ideal points of UK elites are ascertained from a pre-existing dataset produced as part of Gaughan's research (Gaughan, 2024). Following an identical estimation strategy to Barberá et al. (2015), ideal points were generated for 591 UK MPs elicited via their social media followership. As of the 22/08/2022, the 591 MPs with active Twitter/X accounts shared a combined follower network of over 34 million connections and over 11 million unique user accounts (p.6). These accounts were filtered to only include *politically informative* accounts with at least 25 followers, having posted at least 100 tweets, and who followed a minimum of 10 UK MPs, leaving a final follower dataset of 424,297 user accounts (p.6-8). Converting this dataset to a follower adjacency matrix 424,297 x 591, where MPs n form the columns and users m form the rows, correspondence analysis (CA) was applied to estimate ideal points of MPs via the relative distances between their follower networks. These are represented as spatial coordinates in the latent left-right ideological space.

Given the nature of ideal point estimation via CA, positions are also estimated for the follower accounts in tandem, based on the relative distances between the sets of MPs that each account chooses to follow. This makes available a set of ideal point estimates for 424,297 user accounts who were active and followed at least 10 MPs as of 22/08/2022, which will be used for the purposes of this research.[3] These accounts can be reasonably separated into two categories: ordinary users and elites. For this study, elite users are identified as those accounts which meet one of two criteria: 1) verified and/or 2) has a minimum of 30,000 followers.[4] The overwhelming majority of

[3]A handful of accounts (2,153) are in the user dataset but did not have an estimated ideal point for reasons unknown.

[4]This data was extracted from Twitter/X under the previous verification program where accounts were verified under the criteria of being "active, notable, and authentic" (https://help.Twitter/X.com/en/managing-your-account/about-x-verified-accounts). The threshold for 30,000 followers is strict, but motivated by the 2019 *ASA* ruling of celebrity influencers on social media (https://www.theverge.com/2019/7/4/20682087/instagram-Twitter/X-celebrity-30000-followers-advertising-standards-authority-uk).

accounts are ordinary users (97%) while 11,525 (3%) were classified as elite users. As these ideal points are relative measures, their absolute values are meaningless; only their value in relation to every other user in the dimensional space is informative of their ideological position.[5]

The general ideal point distribution of both types of accounts is illustrated in **Figure 5.1**, while summary statistics of their Twitter/X profile metadata can be found in Appendix A.1.

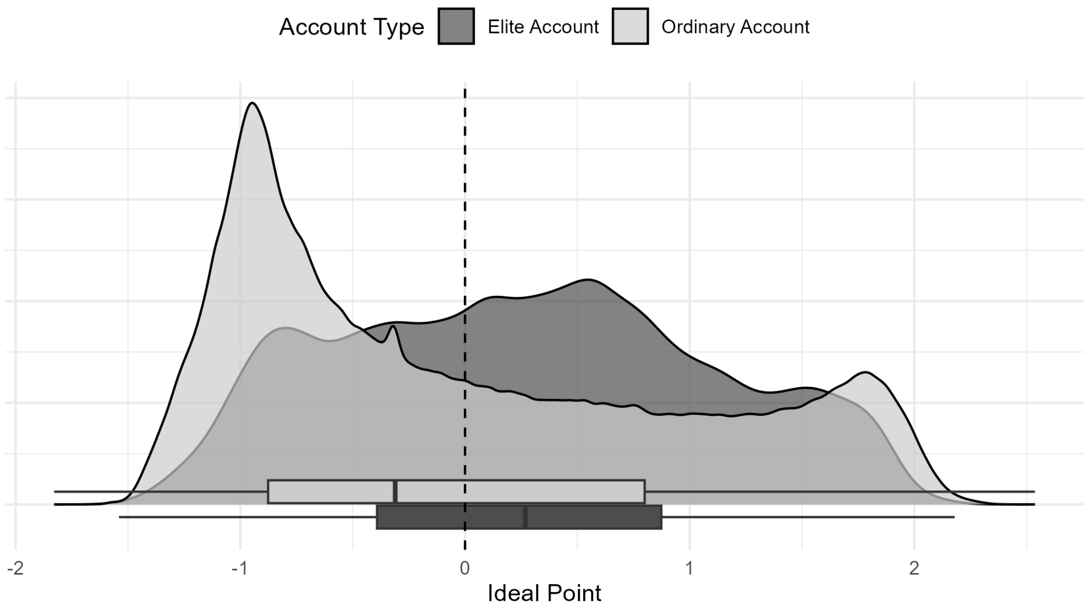


FIGURE 5.1: The ideal point distribution of ordinary users is bimodal with a median point leaning to the left, whereas elite accounts are normally distributed across the axis but leaning to the right. Density plots were computed using kernel density estimation, following the Sheather and Jones (1991) bandwidth selection method.

[5]For better clarity, ideal points are rescaled using z-scale normalisation. This makes interpretation of ideal points easier in that 0 becomes the mean ideal point position, and all other points reflect how many standard deviations each user is from the mean position.

5.2.2 Validating the Ideal Point Scale

To verify that these user ideal points align with the conventional left-right axis (in the British context), both face and formal validation tests are conducted. First, textual analysis of the top twenty recurring words, hashtags, @mentions, and emojis in the Twitter/X bios of accounts from the left and right of the distribution is assessed. Standard text preprocessing steps were taken prior to analysis, where all text was converted to lower case, punctuation was stripped (except @ and #) along with any HTML and URL links, and stop words were removed.[6] Left-wing accounts are identified as those with ideal points in the lowest third quantile and right-wing accounts in the highest third. The results are illustrated in **Figure 5.2**:

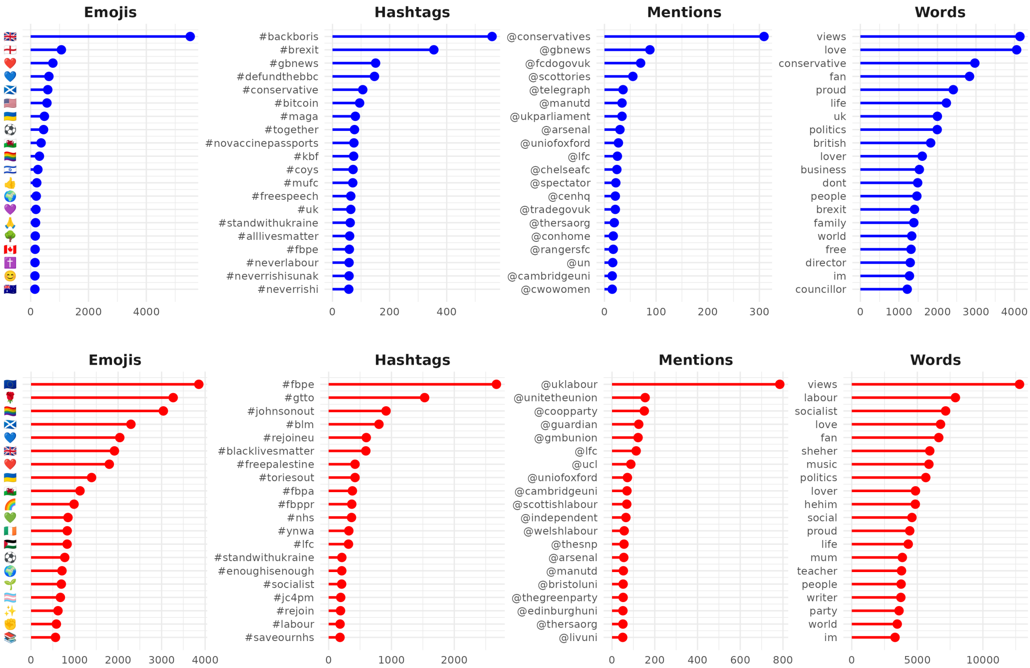


FIGURE 5.2: The top lollipop chart in red shows the 20 most recurring emojis, hashtags, @mentions, and words by all Twitter accounts with ideal points in the lowest third. The bottom lollipop chart in blue illustrates the same for all Twitter accounts with ideal points in the highest third.

Analysis of Twitter/X bio text would indicate that the ideal point distribution aligns well with the British conventional left-right ideological scale. Left-wing accounts most frequently reference words such as ”socialist” and ”labour”, hashtags such as ”#blacklivesmatter”, ”#freepalestine” and ”#toriesout”, and mention left-wing accounts such as ”@uklabour”, ”@unitetheunion”, ”@coopparty” and ”@guardian”. Conversely, right-wing accounts frequently reference words such as ”conservative” and ”brexit”, hashtags such as ”#backboris”, ”#gbnews” and ”#maga”, and mention

[6] Stop words are commonly occurring words that are typically insignificant and filtered out during text preprocessing. These are words such as “and”, “they”, “but”, and “from”. Stop words for this analysis were filtered out according to the *stop_words* English lexicon in the *tidytext* library in R.

right-wing accounts such as ”@conservatives”, ”@gbnews”, ”@scottories” and “@telegraph”. Emoji analysis also yields similarly intuitive results, where the most commonly recurring emojis for left-wing accounts are the EU flag, the red rose (the Labour Party emblem), and the rainbow pride flag, whilst for right-wing accounts it is overwhelmingly the Union Jack, followed by the flag of England.

To formally validate the ideal point scale, a survey of the general public was conducted in collaboration with polling company *YouGov*. A sample of 2,068 members of the adult British population were surveyed, asking them to place 6 UK political parties, 24 UK media organisations, and 30 individual media personalities on a 0-10 scale, where 0 represents the furthest left and 10 represents the furthest right. Individual media personalities were primarily selected by largest number of Twitter/X followers in the user dataset downwards, making sure to only include individuals who were British, regularly appear in the British media, and could reasonably be identified ideologically. Also, an attempt was made to ensure that a reasonable spread of individuals from across the political spectrum were selected. For UK political parties and media organisations, only major parties/organisations were selected that could be placed ideologically by members of the general public. All 30 individual media personalities have a corresponding ideal point, as do a number of parties and media organisations where their official Twitter/X accounts have met the inclusion criteria for an ideal point to be generated. However, 1 political party and 14 media organisations do not have an associated ideal point. In these cases, the mean ideal point of affiliated Twitter/X accounts is used as a proxy, where affiliated accounts are identified as elite accounts which @mention the party/media organisation’s username in their Twitter/X bio. The association between user ideal points and the mean ideological estimate of the British public is modelled using OLS linear regression. Results are illustrated in **Figure 5.3**:

Model results would indicate a strong degree of positive association between the Twitter/X user ideal points and the mean ideological estimates of the British public (R^2 = 0.61). It is worth acknowledging the ongoing debate in the literature around how well general members of the public actually understand the concept of ideology, as well as how accurately they could place themselves and others on a conventional ideological axis (Converse, 1964; Freeder et al., 2019; Yeung and Quek, 2024; Zaller, 1992). This is evidenced by the range of uncertainty demonstrated in the survey response data (see Appendix B), and by the fact that, on average, 43% of responses to each party/organisation/individual was “Don’t Know”. As such, the general public estimates may not necessarily reflect the “ground truth” ideological position of the target. However, a strong correlation between the Twitter/X ideal point estimates and the general public estimates does confirm that the ideal points align relatively well with public perception. Thus, when assessing ideological representation of guests on

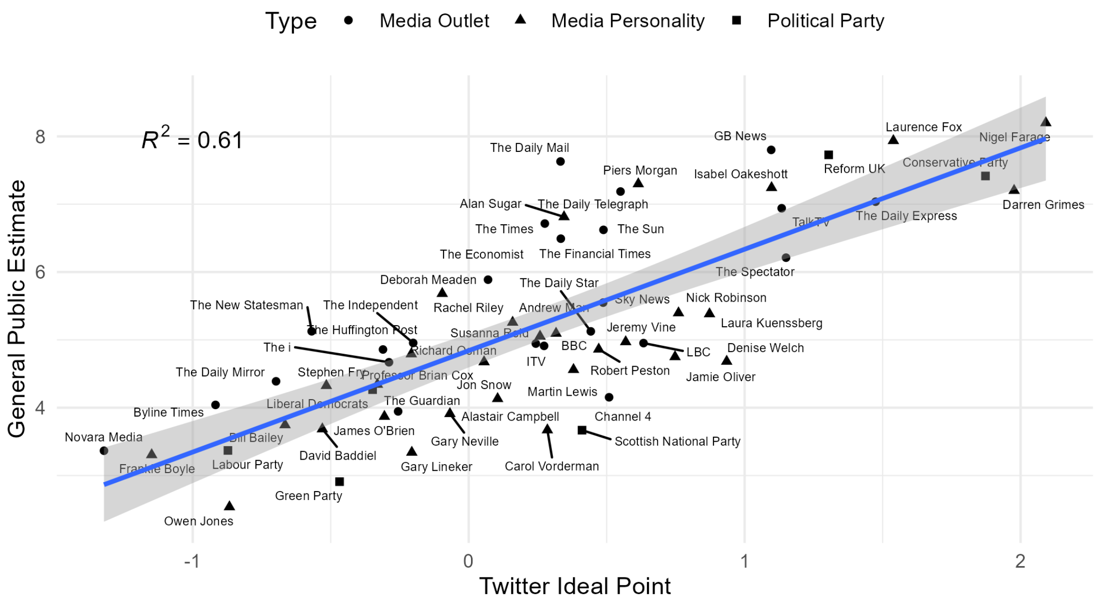


FIGURE 5.3: Weighted least squares regression predicting the mean ideological estimate of the general public (y-axis) for each target by their associated ideal point (x-axis) yields a high degree of correlation: $R^2 = 0.61$. Plot points are shaped by their category type.

each programme, this will at the very least align with how the general public conceive of it.

5.2.3 Guest List Data

Seven flagship political programmes from four PSBs and two non-PSBs are selected for analysis. Two programmes are selected for BBC One, while each other broadcaster has one. Guest lists for each show are collated across a two year period between the 01/01/2022 and the 01/01/2024. The motivation for this timeframe is primarily two-fold: 1) the ideal points were estimated based on Twitter/X networks as of 22/08/2022. Thus, analysing guest lists further back than 01/01/2022 faces the issue of increasingly inaccurate ideal points, if one assumes that ideological position does not remain fixed over time; and 2) some guests lists had to be collated manually, which is an extremely time-consuming task, subsequently restricting how far back guest lists could be gathered. All guests that appeared on each show at any time during this two-year period appear in each list, along with the accompanying date in which they appeared. Many of the guests on each show appeared more than once on different dates, and many guests appeared on multiple different shows. Further details of each programme and the data source for guest appearances on each episode can be found in Appendix C. For BBC Two: Politics Live, guest list data was derived from Twitter/X. The search query used returned tweets containing guest appearances as far back as 22/05/2023 before the tweets became too irregular to be useful.

However, Politics Live is the only show in the dataset which airs daily instead of weekly, meaning that there is still an adequate sample of guest data collated over the six month timeframe. A breakdown of the guest list data for each of the seven shows can be found in **Table 5.1**::

Broadcaster	Programme	PSB?	Frequency	Date Range	*n*	Ideal Point (%)
BBC One	Question Time	Y	Weekly	13/01/2022 - 14/12/2023	362	82%
BBC One	Sunday with Laura Kuenssberg	Y	Weekly	04/09/2022 - 10/12/2023	387	62%
BBC Two	Politics Live	Y	Daily	22/05/2023 - 19/12/2023	357	91%
ITV One	Peston	Y	Weekly	12/01/2022 - 22/11/2023	377	76%
Channel 4	The Andrew Neil Show	Y	Weekly	08/05/2022 - 02/04/2023	144	90%
Sky News	Sophy Ridge on Sunday	N	Weekly	20/03/2022 - 16/07/2023	153	85%
GB News	The Camilla Tominey Show	N	Weekly	08/01/2023 - 17/12/2023	254	64%

TABLE 5.1: Summary information for each selected programme for analysis. Columns 1-4 convey details about the programme itself, whilst columns 5-7 convey information about the guest data including observation timeframe, number of observations, and percentage of available ideal point data.

Overall, there are 825 unique guests who appeared at least once on any one of the seven shows. Basic details about these guests were manually collated from various online websites including their Wikipedia page and their social media pages including Twitter/X and LinkedIn. These details include their Twitter/X username (for ideal point matching), the primary organisation they represent, and their nationality. Guests were also categorised by their general occupation type. These categories were kept relatively broad and some guests could feasibly fit into multiple categories, but their primary occupational focus was used to separate them out. **Figure 5.4** illustrates the breakdown of guests by primary category type. For a more detailed description of their sub-categories and the distribution of these categories by ideal point, see Appendix D.

As for organisational representation, 235 unique organisations were represented by guests across the seven shows ranging from political parties, media outlets, and trade unions to think tanks, NGOs, campaign groups, and corporations. **Figure 5.5** illustrates the top 30 organisations represented by guests as a percentage of each programme. Expectedly, political parties and media outlets dominate broadcaster guest selection, with the Conservative and Labour parties the most represented group. The UK's minority parties, Lib Dems, SNP, Reform UK, and Greens, were also well represented across most of the seven shows along with a number of the UK's leading media outlets such as The Telegraph, The Guardian, GB News, The Spectator and the Financial Times. The two major political parties in the US were also relatively well represented along with the Holos Party in Ukraine. A handful of non-political or media based organisations also appeared frequently including major trade unions and think tanks. Notably, there is evidence of variation in organisational representation across each of the seven shows. Many studies that assess media representation will code the position of the individual actor based on the ideological slant of the organisation they represent. However, one of the primary strengths of using

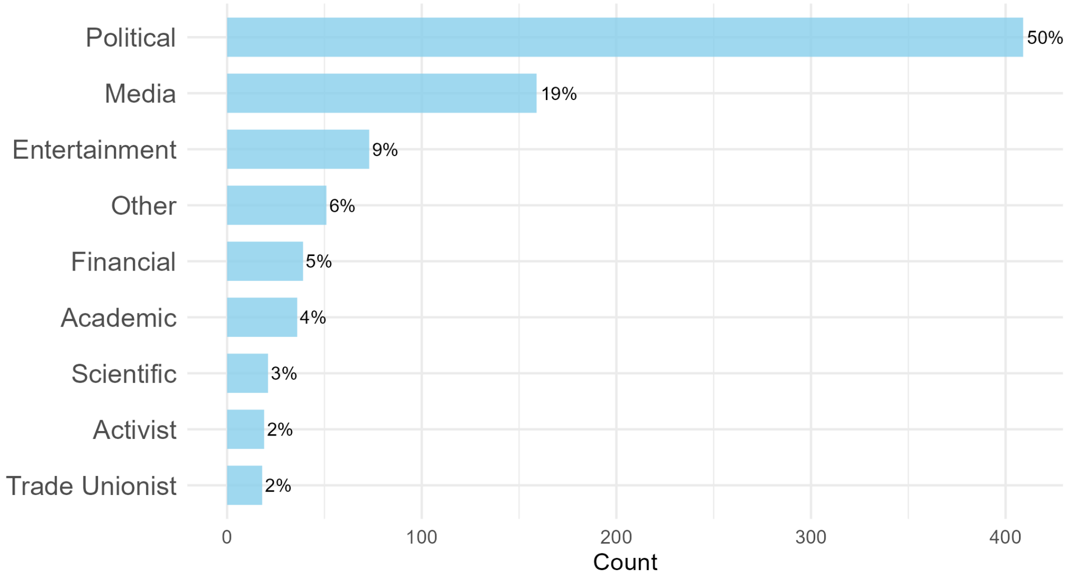


FIGURE 5.4: Political guests make up the majority of category types (50%), followed by media guests (19%) and entertainers (9%). The 'Other' category contains a range of guests who did not simply fit into any of the larger categories. This includes groups such as special advisers, lawyers, clerics, communication experts, and military officers.

individual actor-level ideal points to assess ideological representation is to allow researchers to compare not only between different organisations but also *within*.

One limitation of this dataset is missing ideal points for a number of guests on each show, as can be seen in **Table 5.1**. There are two reasons why a guest will be missing an ideal point: 1) the guest did not have an active Twitter/X account when the data was originally extracted, or 2) the guest had an active Twitter/X account, but did not follow a minimum of 10 UK MPs for an ideal point to be estimated. Failure to meet either of these two criteria is non-probabilistic, and examination of the types of accounts missing ideal points indicates that non-British guests were far less likely to have an ideal point (13%), as well as certain category types such as scientists (24%) and entertainers (19%) (See Appendix E for more detail). One possible solution for addressing the issue with missing data is to use the ideal point of the organisation they represent as a proxy, or the mean ideal point of accounts affiliated with the organisation. Addressing missing data in this way does increase ideal point availability from 77% of guests to 88% and, as is illustrated in Appendix E.4, these proxy ideal points do correlate strongly with the ideal points of individual guests. However, as is also shown in Appendix E.5, replacing missing data with proxies does not significantly alter the final results of this research and so does not warrant inclusion.

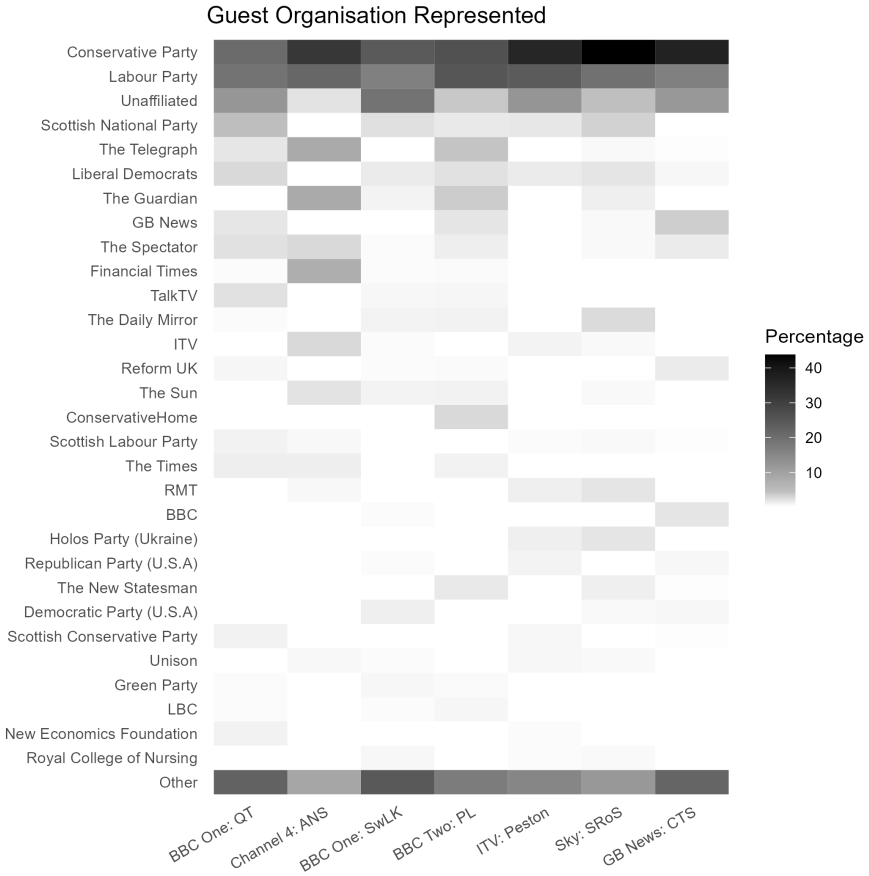


FIGURE 5.5: Organisations are ranked in descending order based on the overall number of times they appear across any of the seven shows combined. There were 235 organisations in total. This plot shows the top 30 and the remaining 205 were grouped into Other. Tiles are gradient shaded based on the percentage of each show that an organisation was represented by guests.

5.3 Results

Figure 5.6 illustrates the static ideal point distribution of guest appearances on each show at any point over the two year timeframe.

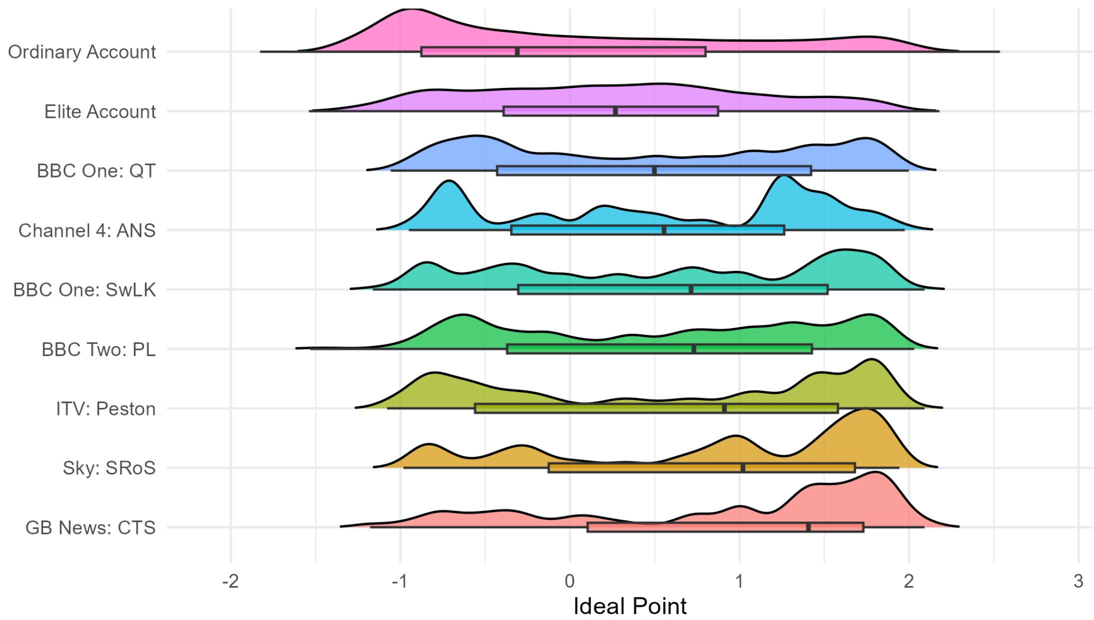


FIGURE 5.6: The ideal point distributions of guests on each programme, ranked in descending order by median ideal point. All seven programmes demonstrate median ideal points to the right of both ordinary and elite Twitter users, to varying extents. Density plots (smoothed using kernel density estimation with a bandwidth of 0.1) demonstrate relative bimodality in a number of cases.

Descriptive examination would suggest each programme drew their guests further to the right than the general Twitter user. The two non-PSB programmes selected for analysis demonstrate median ideal points furthest to the right , whilst BBC One: Question Time and Channel 4: The Andrew Neil Show have median ideal points closest to the centre. These programmes also demonstrate ideal point distributions notably dissimilar to that of ordinary Twitter user accounts which are disproportionality skewed towards the left of the axis, and to a lesser extent Twitter elites. Importantly, the shape of each distribution would suggest that guest ideal points for each programme are not normally distributed across the axis. The multimodality of many of the programmes suggests that, despite a slant towards the right, guests were drawn sizeably from multiple parts of the spectrum.

It may be intuitive that because broadcasters in the UK are expected to abide by due impartiality rules they deliberately select the most explicitly partisan actors from both sides of the axis to demonstrate balance in their coverage, thus creating a more bimodal distribution of ideal points. This may raise questions then about the ideological representativeness of these programmes if actors from the more centre ground are underrepresented as a result, which looks to especially be the case for ITV: Peston. Overall, what is most evident from comparisons of each distribution is that the spread of guests across the ideological spectrum is relatively similar between each programme (each show has, at least at one point or another, platformed a guest(s) from the further left across to the further right). However, the shape of these distributions vary greatly, as does their mean and median point.

In order to formally test the hypotheses, an ANOVA test is conducted with post-hoc Tukey HSD pairwise comparisons. **Figure 5.7** illustrates the results of the pairwise comparisons with 95% confidence intervals.

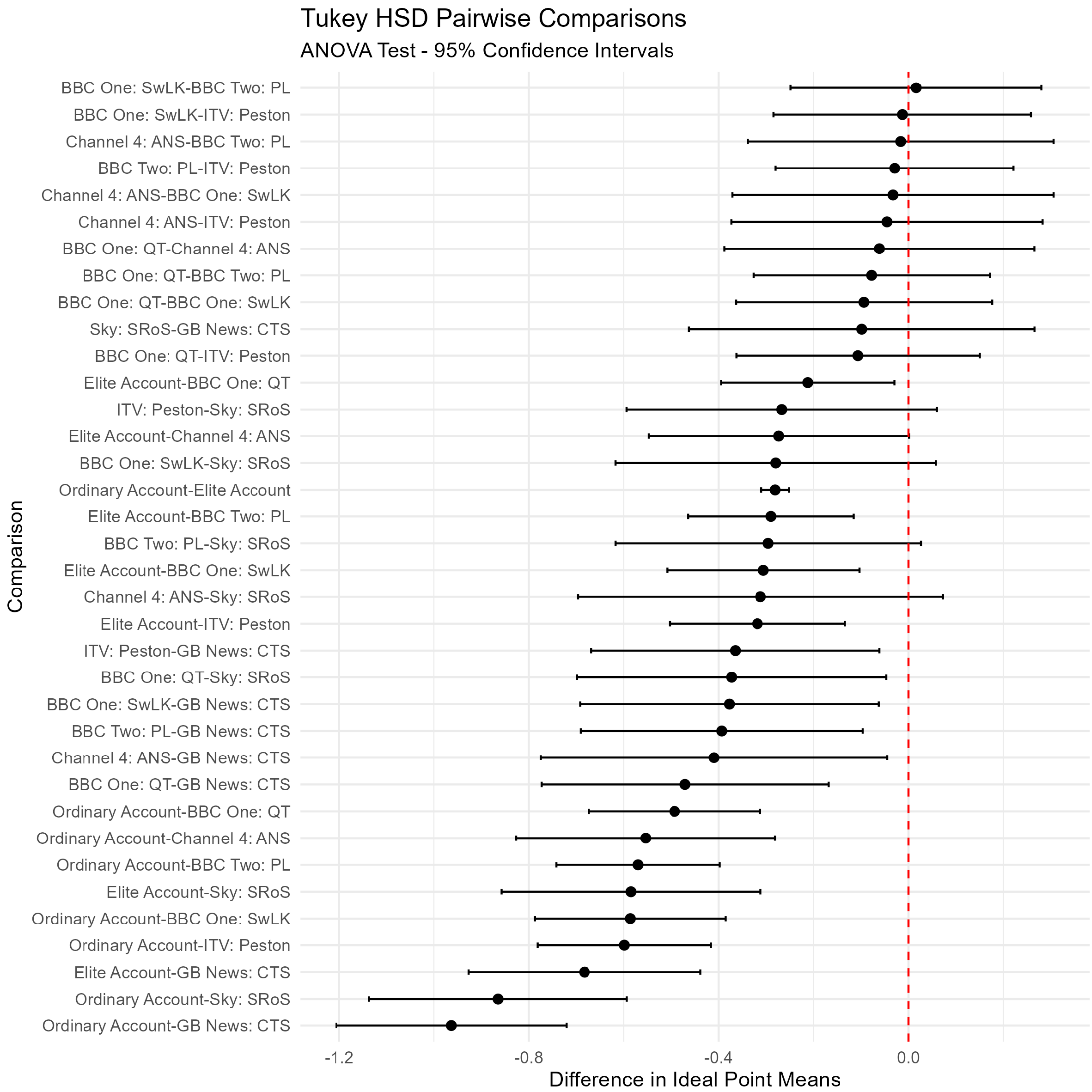


FIGURE 5.7: Pairwise comparisons of mean user ideal points across the seven programmes using Tukey's HSD test. Each line represents the estimated difference in mean ideal points between two programmes, with 95% confidence intervals. Comparisons are ranked by the size of the estimated difference. Intervals that do not cross zero indicate statistically significant differences in guest ideological positioning between programmes.

One can observe a number of key findings from the tests. Firstly, the mean ideal point of guest selections on all seven programmes differed significantly from both ordinary Twitter/X users and, more importantly, elite accounts. Thus, there is evidence to support H1.

Secondly, the results also demonstrates that significant differences exist in the mean ideal point of guests *between* each programme. Specifically, with the exception of BBC One: Question Time vs. Sky News: Sophy Ridge on Sunday, the ideological representation of guests across six of the seven shows *do not* differ significantly. However, with the exception of Sky News: SRoS, the guest selection of GB News: The Camilla Tominey Show differs significantly from every other programme. With a mean guest ideal point furthest to the right of any other programme, this provides evidence to suggest that GB News: CTS demonstrated the greatest degree of slant to the right of the axis and supports the rejection of H2a. Additionally, the two chosen non-PSBs had the highest mean ideal points to the right, indicating some support for H2b, though these differences were only consistently significant for GB News: CTS.

Along with testing for the significance in difference between mean ranks of each programme, Hartigan's dip test is used to assess the degree of multimodality of each distribution (Hartigan and Hartigan, 1985). The dip test checks if a data distribution has more than one mode by creating a unimodal distribution function that has the smallest value deviations from the empirical distribution function, the largest of which is the dip statistic. The larger the dip statistic, the greater the probability that empirical data has more than one mode (Stanbro, 2012). This particular method of assessing bimodality in an empirical distribution has the benefit of requiring no a priori knowledge about the underlying distribution of the data and computes a straightforward p-value to determine the presence of significant bimodality. The purpose of this test in this case is to statistically confirm whether each programme drew significantly from more than one area along the ideal point spectrum. Results are shown in **Table 5.2**.

Programme	**Complete *n***	**Median**	**Mean**	**S.D**	**Min.**	**Max**	**Dip Stat**
Ordinary Account	410,640	-0.31	-0.01	1	-1.83	2.54	0.0071***
Elite Account	11,504	0.27	0.27	0.82	-1.54	2.18	0.0031
BBC One: QT	295	0.5	0.49	0.95	-1.06	2.00	0.057***
BBC One: SwLK	238	0.71	0.58	0.22	-1.16	2.09	0.05***
BBC Two: PL	324	0.73	0.56	0.96	-1.53	2.03	0.049***
ITV: Peston	288	0.91	0.59	1.06	-1.08	2.09	0.077***
Channel 4: ANS	129	0.55	0.55	0.92	-0.95	1.97	0.082***
Sky News: SRoS	130	1.02	0.86	0.95	-0.98	1.94	0.05*
GB News: CTS	163	1.41	0.96	0.94	-1.18	2.09	0.036

TABLE 5.2: Summary statistics of the ideal point distributions for ordinary and elite Twitter users in the dataset, along with subsets of users who appeared as guests on each of the seven shows. Dip statistics include significance asterisks to indicate statistical significance of multimodality in each distribution. In both cases, statistical significance is measured at the 0.05 threshold. [$* * * \leq 0.001, ** \leq 0.01, * \leq 0.05$]

Hartigan's dip tests indicate evidence of multimodality in their guest selection. This is the case for all programmes with the exception of GB News: CTS. Analysis would indicate that not only did this programme demonstrate an average user ideal point furthest to the right, Hartigan's dip test would suggest that this was the only programme to not exhibit a multimodal distribution. This suggests that GB News: CTS drew their guests *almost exclusively* from the right side of the ideal point axis.

5.3.1 Additional Checks

Two additional checks are conducted to assess the robustness of these findings. Firstly, to check there is no significant variation in guest selection over time during the two year window that may bias the overall results, guest appearances are divided into quarterly time intervals. This is illustrated in **Table 5.8**.

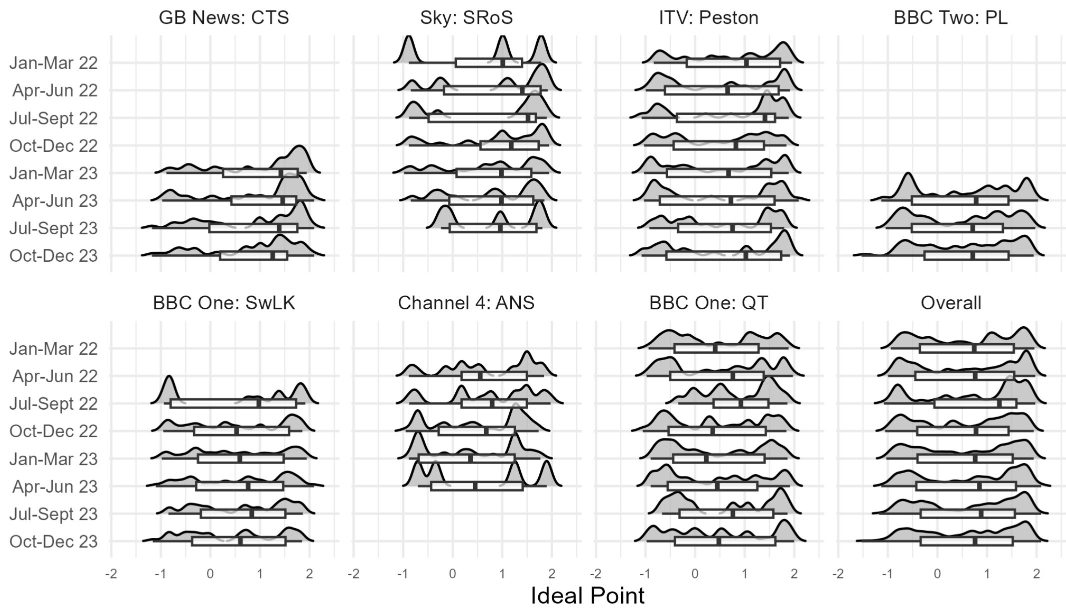


FIGURE 5.8: Ideal point distributions of guest appearances on each show between 2022 and 2024, divided into quarterly time intervals. Distributions vary between shows and between time intervals within shows, but the overall quarterly distributions remain relatively stable between intervals. The exception is the 3rd quarter of 2022 (Jul-Sept 2022) where the overall median guest ideal point shifts notably further to the right, though pairwise Tukey tests do not confirm any statistically significant differences between time intervals.

General examination would suggest that, whilst some time variance within the data is present (particularly within shows), the overall ideological representation of guests across all shows combined remains relatively stable between time intervals. This is with the exception of the 3rd quarter of 2022 between July and September, where the average guest ideal point shifts further to the right. This notable movement to the right is reflected in multiple programmes including BBC One: Question Time, BBC One: Sunday with Laura, and ITV: Peston. This finding is intuitive as this period was dominated by the Conservative Party leadership contest which ran from the 13th July to the 5th September. During this period it would be expected that more Conservative

MPs would be selected to appear on these programmes, including prospective leadership candidates as well as wider parliamentary party members. However, pairwise Tukey HSD tests could not confirm that any of the differences in distributions between time intervals were statistically significant. The relative stability of guest ideal point representation across time intervals is informative, and generally indicates that the distribution of guest ideal points remains fixed and evenly distributed across time points. Without more data, is difficult to know how far back or forward this trend can be extrapolated. Nonetheless, a lack of significant variance from time point to time point or any indication of a growing trend towards one side of the axis or the other would suggest that this spread is moderately resolute.

Secondly, the case has already been made that a slant towards the right in guest selection may be expected given the incumbency bonus afforded to Conservative MPs, increasing their relevancy and thus likelihood to be selected. To assess this effect, the same tests are conducted excluding MPs (who make up 46% of guest selection). Analysis indicates that, while all seven programmes still exhibit an average ideal point further to the right than elite Twitter accounts, distributional differences are now only significant for BBC One: Politics Live and GB News: The Camilla Tominey Show when compared to elite accounts more generally. This suggests that MP selection is a substantial driver of the representational differences in most cases.

Results of Hartigan's dip test indicate only Channel 4: The Andrew Neil Show demonstrated evidence of significant multimodality when removing MPs from the sample. This may be a function of how each programme selects their guests to show impartiality, especially in a two-party dominant system like the UK. The simplest way for a broadcaster to demonstrate balance is to select an MP from the major political party on one side of the political aisle and then directly balance it with an MP from the other side. Doing so inherently embeds bimodality into its guest selection by selecting the most strictly partisan actors on either side. This effect is likely reduced when MPs are removed from the sample.

5.4 Discussion & Conclusion

All in all, this article has demonstrated a novel application of social media network data to assess ideological representation on major TV broadcasters in the UK. Building on the earlier work and dataset from Gaughan (2024), this study has matched estimates of left-right position of UK elites elicited via their Twitter/X follow networks to the guest lists of seven flagship political programmes featured on six major broadcasters. These ideal points were formally validated using general public estimates of ideological position elicited via an original survey of 2,068 respondents, confirming a high degree of correlation (R^2 = 0.61). Where elite (mostly British)

Twitter/X users are evenly distributed across the political left-right axis, guests selected to appear on each of the seven programmes between 01/01/2022 and 01/01/2024 demonstrated average ideal points to the right of Twitter's elite. Pairwise Tukey HSD tests confirmed the statistical significance of differences in ideal point distributions between all seven programmes and elite Twitter users more generally, as well as between GB News: The Camilla Tominey Show and every other programme.

Despite the apparent slant of guest ideal points towards the right of the ideological spectrum, Hartigan's dip tests indicated that, with the exception of GB News: The Camilla Tominey Show, these programmes still drew their guests significantly from multiple parts of the political axis. Additional analysis confirmed that these patterns remained relatively stable between quarterly time intervals across the two-year period and that this slant was heavily driven by an over-representation of right-wing MPs. When removing MPs from the sample, the results of this study no longer hold with the exception of BBC Two: Politics Live and GB News: The Camilla Tominey Show.

The purpose of this study is to highlight an innovative way to assess actor-focused political representation where ideological position of the actor can be estimated independently from the organisation they represent. Where previous studies have relied upon manual coding of left-right position to assess media representation - which is fraught with issues regarding time and replicability - this study has demonstrated that social media networks can provide accurate measures of left-right position at the actor level. To the best of the author's knowledge, media representation has not yet been examined in such a way and there is hope that this study will encourage new avenues of research into media bias which makes greater use of social media network data given its efficacy. From a substantive perspective, the findings of this paper may have important implications for perceptions of ideological representation on TV broadcasters in the UK. The results of this research suggest that all of the UK's flagship TV political discussion programmes drew their average guest from significantly to the right of the average Twitter elite user since the start of 2022. The mechanisms for how and why different broadcasters select guests to appear on their shows are no doubt multifaceted. This is determined by a range of factors including the political context, media norms, relevancy, topics of discussion, and general interpretations of what impartial coverage should actually look like Hopmann et al. (2012). Additionally, this research is constrained by its evaluation of representation relative to the presence of elite users on Twitter only.

Subsequently, the methodology employed in this paper is by no means comprehensive enough to make any conclusive statements about the overall impartiality of UK broadcasters on their flagship political programmes over the last two years, nor can it make any casual claims about the drivers of guest selection. Additionally, if one accounts for the fact that the slant towards the right of the axis is predominantly driven by an over-selection of Conservative MPs, this may be a result of an incumbency bonus which boosts the relevancy of government ministers and members of the governing party. This is with the exception of BBC Two: Politics Live and GB News: The Camilla Tominey Show, which maintain statistical significance even when excluding MPs from the sample. GB News: CTS was also the only programme to not exhibit evidence of a multimodal distribution suggesting it platformed guests almost exclusively from the right of the average Twitter elite, MPs or otherwise.

Moreover, this paper also acknowledges the post-Brexit context this study has taken place in. While the observation period Jan 2022 - Jan 2024 occurs several years after the UK's exit from the European Union, there is no doubt that Brexit has played an influential role in realigning the conventional ideological axes of British politics. In order to establish a complete understanding of ideological representation on UK broadcasters, a one-dimensional approach to the left/right spectrum may not be comprehensive enough to capture this additional facet. In particular, if one considers the selection of guests on these programmes to be driven by some desire for political "balance", broadcasters may be driven as much by balancing the conventional left and right with an equal amount of Remainers and Leavers. It is well-established that the Remain-Leave divide cuts asymmetrically across traditional partisan and ideological lines and it is unclear how this is reflected in the research conducted here. Research has shown UK broadcasters pursued an even-handed coverage of the Leave and Remain campaigns in the lead-up to the referendum in 2016, but this leant in favour of right-wing sources over the left (Cushion and Lewis, 2017). The Remain-Leave divide does not map cleanly onto traditional conceptions of the the left-right axis and, as such, may warrant better incorporation into any future work.

In addition to this, a number of other missing details could be incorporate into future work to build on this preliminary analysis. For example, ideal points of guests could be compared against the length of time they are given to speak on each of these shows to assess the relationship between a guest's ideological position and general favourability shown towards them. Similarly, each programme's topic(s) of discussion could be categorised from week to week to assess the general relationship between issue coverage and the average ideological position of guests selected to discuss it. For instance, do certain issues encourage a more left/right leaning selection of guests, or are some particular issues more polarising than others? From a causal perspective, the relationship between public opinion polling and the mean ideal point of each show week to week might yield some interesting insight into how opinion polls can influence guest appearances on these programmes and whether or not a drop in public support for a political party can increase/decrease their general visibility. A recent change of government in the UK where the left-wing Labour Party were elected to government for the first time since before 2010 also provides an opportunity to put the incumbency effect to the test.

In conclusion, this paper has made a departure from conventional studies on media representation which typically rely on thematic content analysis by employing a novel method of ideal point estimation via elite social media networks. It has found evidence to suggest that the flagship political programmes of the UK's four major public service broadcasters (BBC One, BBC Two, ITV, Channel 4) and two major non-public service broadcasters (Sky News, GB News) selected guests disproportionately to the right of the average elite Twitter user between 2022 and 2024. Further research is required to investigate exactly what this means in substantive terms, but opens a range of new opportunities and methods for assessing media representation in future studies.

Chapter 6

Digital Dyadic Representation: Evidence from UK MPs on Twitter - Paper 3

Dyadic representation refers to the degree in which elected legislators represent the interests and opinions of their constituents (Weissberg, 1978). This paper seeks to assess how closely the online Twitter profiles of MPs in the UK align with their constituencies beyond party affiliation. Estimating local constituency position along two left/right ideological dimensions (economic and social) and their most important issue (MII) using multilevel regression with poststratification (MRP), this study tests the within-party responsiveness of MP Twitter profiles to their constituencies along three digital dimensions: (1) follower networks; (2) positive engagement networks; and (3) tweet content. Results indicate that when controlling for party affiliation, MP follower networks are positively associated with constituency left/right position along a social dimension but negatively associated along an economic one. However, no statistically significant association could be found with MP positive engagement networks or tweet content beyond party affiliation and time-period effects. The findings of this study reaffirm the staunchly partisan nature of Twitter usage by UK MPs and the prevailing ideological responsiveness of their follower networks beyond the dominance of the party line.

6.1 Introduction

In his Conservative Party conference speech in 2015, then UK Prime Minister David Cameron famously declared that Britain and Twitter were not the same thing (Williams, 2015). Studies since then have largely reaffirmed that assessment, illustrating that Twitter (now *X*) users in the UK are, on average, younger, wealthier, more highly educated, and more politically engaged than the general population (Blank and Lutz, 2016; Mellon and Prosser, 2017; Sloan, 2017). Broadly, Twitter has gained a reputation for being a platform for elites, not simply in terms of its average user but also in the volume and visibility of elite actors on the site relative to any other social network (Dagoula, 2019). Along with a large number of journalists and other media elites (Molyneux and McGregor, 2022), politicians over the last decade have increasingly made use of social media platforms like Twitter and Facebook to engage with the public and this has had significant implications for political representation. Whilst this uptake has not been without consequence - politicians on social media are routinely subjected to online abuse, hate speech, and threats of violence (Agarwal et al., 2021; Ward and McLoughlin, 2020) – studies have promisingly shown that it has helped facilitate new channels of engagement between citizens and MPs (Agarwal et al., 2019; Jackson and Lilleker, 2011). Further studies have also shown that this direct form of two-way engagement with elected officials can positively boost public trust and feelings of representation (Dvir-Gvirsman et al., 2022; Soo et al., 2020; Starke et al., 2020).

The term *dyadic representation* refers to the degree in which elected legislators represent the political preferences or interests of the geographical constituency from which they are elected (Weissberg, 1978). This form of substantive representation (Pitkin, 1969) is a fundamental tenet of single-member district (SMD) systems found in countries like the UK and the US, and the notion that SMDs promote stronger voter-member linkages is considered to be one of its prevailing strengths (Norris, 2001). The extent to which ordinary voters feel represented by their legislator has important consequences for perceptions of democracy and general satisfaction with the political system (ibid.). With this in mind, this paper seeks to assess the concept of dyadic representation through the digital profiles of elected MPs in the UK on Twitter. The UK offers a unique case study here, where, as a quintessential Westminster system, it's electoral culture is characterised by a parliamentary government with strong executive agenda-setting powers, encouraging stricter whips and highly disciplined parties (Bawn and Thies, 2003; Cox, 1987). Thus, within-party legislator responsiveness is typically much lower in these types of systems and makes the use of digital platforms such as Twitter even more valuable for politicians who wish to more closely represent their constituents. As such, the overall exploratory question here is straightforward: are the online Twitter profiles of UK MPs in some way responsive to their constituencies beyond that of their party position?

To answer this question, this study assesses the dyadic relationship between an MP's digital profile and the position of their respective constituency along three dimensions: (1) economic left/right position; (2) social left/right position; and (3) most important issue. Making use of a contemporary methodological technique for small area estimation, local opinion for 632 constituencies across mainland Britain (excludes Northern Ireland) are estimated at five stages across the majority of the previous parliamentary period (Dec 2019 – May 2023) using MRP. These estimates are generated using respondent data from five waves of the BESIP, augmented by constituency demographic and electoral data from the latest BES constituency results file and post-stratified using an existing poststratification frame built by Hanretty et al. (2018). Positions of corresponding MPs are estimated via three data types extracted from their Twitter profiles: (1) follower networks; (2) positive engagement networks; and (3) tweet timelines. Having done so, digital dyadic representation in the UK in subsequently examined through three separate studies:

Study 1 tests the association between constituency economic and social left/right positions and the left/right position of each MP estimated via their Twitter follower networks;

Study 2 tests the association between constituency economic and social left/right positions and the median left/right position of each MP's Twitter positive engagement network; and

Study 3 tests the association between constituency most important issues (MII) and the tweet topics of MPs.

To the best of the author's knowledge, the concept of digital dyadic representation has not yet been tested in such a way, and certainly not in the case of a Westminster system like the UK. Legislator-constituent relationship is an important concern in political research and social media platforms like Twitter and Facebook can provide politicians with an more direct medium for citizen engagement. The degree to which a legislator's digital profile (in this particular case, Twitter) reflects the general position of their constituency can have important implications for public perceptions of representation. This paper looks to examine this relationship and concurrently outline a novel methodological approach for doing so.

The following section of this article will first begin by summarising existing studies on traditional forms of dyadic representation, the limitations of such studies in Westminster systems thus far, and research that has previously analysed online political representation in the UK and elsewhere. The next section will then describe the methodological approach and data sources in detail before moving on to illustrate the findings of each of the three studies individually. Finally, the overall results and implications of this research are discussed and overall conclusions are drawn. It is important to note that this paper does not attempt to establish a causal mechanism driving constituency-legislator digital responsiveness. However, providing descriptive evidence that some form of dyadic relationship does in fact exist between constituents and the digital profiles of their representatives – even in strict legislative systems like the UK – will have implications for how we understand online political representation moving forward.

6.2 Literature Review

6.2.1 Dyadic Representation in Westminster Systems

Early research in the field of legislator-constituency relationship in the 1960s and 1970s (see: Achen, 1978; Erikson, 1978; Fenno, 1977; Miller and Stokes, 1963; Weisberg, 1978) laid the foundations for the concept of dyadic representation, and the idea that congressional behaviour could be in some way moved by individual constituency preferences. Since then, a range of newer studies have furthered our understanding of dyadic representation and largely confirmed that the association between legislator and constituency opinion beyond party affiliation is, at the very least, non-zero (Ansolabehere et al., 2001; Canes-Wrone et al., 2002; Kastellec et al., 2010; Krimmel et al., 2016; Rogers, 2017). However, to date, studies assessing dyadic representation in SMD systems have almost exclusively been conducted in the US, where its unusually weak party cohesion and individualistic electoral culture inherently favours a greater degree of within-party policy responsiveness (Bawn and Thies, 2003). This leaves open the question of whether or not such within-party legislator responsiveness would hold in stricter SMD systems, where the prospects for individual legislators breaking ranks with their party are much lower. One early study of dyadic representation in an SMD system outside of the US was conducted by Converse and Pierce (1986) looking at the French Fifth Republic, and their research found that the roll-call voting of French deputies *was* in fact responsive to constituency opinion. However, what their study failed to establish was whether or not this effect was a result of within-party responsiveness or simply a component of partisan competition more generally.

When looking to study Westminster systems in particular, these pose a uniquely difficult challenge. Legislative bodies which adopt Westminster systems are usually characterised by highly disciplined parties where members compete electorally on national policy platforms and voters (typically) select parties for government instead of local representatives (Bawn and Thies, 2003; Cox, 1987). Moreover, these types of systems have a tendency to produce “government vs. opposition” parliamentary cultures, making party rebellions far more uncommon (Dewan and Spirling, 2011; Kam, 2009; Slapin et al., 2018). From a methodological perspective, this makes obtaining unobtrusive measures of legislator position beyond their party position using roll-call votes extremely difficult (Franklin and Tappin, 1977). There are also methodological limitations to measuring constituency opinion in many of the countries that use this type of system: for instance, in the UK, which has 650 constituencies of circa 70,000 registered voters each, conducting representative opinion polls for each individual constituency would be too expensive, and disaggregating national polls by constituency would be too inaccurate (Hanretty, 2020, p.631). However, capitalising on modern methodological developments in small area estimation (Park et al., 2004) and a method for estimating legislator policy position via Early Day Motions (EDM) instead of roll-call votes (Kellermann, 2012), research by Hanretty et al. (2017) managed to circumvent both of these limitations and was able to directly model dyadic representation in a Westminster system. The study found that, across three separate policy dimensions, MPs in the UK were indeed weakly responsive to constituency opinion and this association held even when controlling for party affiliation. Consequently, despite the aforementioned constraints to testing dyadic representation in SMDs outside of the US, the possibility that within-party legislator responsiveness to constituency opinion can and does exist in these more restrictive systems is promising, and the methodological tools for assessing it are now much more advanced.

6.2.2 Alternative Mediums of Constituency Representation

With this in mind, and directly inspired by the original work of Hanretty et al. (2017), this paper looks to go one step further and assess dyadic representation in the UK House of Commons from a different perspective. The majority of work on dyadic representation understandably evaluates constituency-legislator relationships from a largely legislative standpoint, i.e., how closely does a legislator's roll-call vote mirror the corresponding position of their constituency. In restrictive legislatives where the opportunities for legislators to vote independently from their parties are limited, however, studies have shown that legislators will readily adopt alternative avenues to represent constituency interests (Mayhew, 1974; Proksch and Slapin, 2014). As Hall (1996, p.58) argues, representational acts extend far beyond that of voting behaviour alone and is as much a question of how far constituency influence impacts what, when and to what extent representatives will participate in particular matters. For instance, research by Blumenau and Damiani (2021) found that over the last 40 years MPs in the HoC have increasingly employed constituency-oriented language in their parliamentary speeches, and this is reaffirmed by the work of McKay (2020, p.5-6) that demonstrated an almost six-fold increase in the frequency of speeches mentioning constituency and constituent(s) between 1950 and 2019. Work by Soroka et al. (2009) illustrated how MPs in Canada will use question periods in parliament as an alternative venue outside of legislative voting to represent constituency interests. A study by Blumenau (2021) also found evidence of dyadic responsiveness between UK MPs and constituency online activism, where support for an e-petition (online) among an MP's constituents was positively associated with an increase in the probability that the MP advocates for the petition in parliamentary debate. Not only are MPs responsive to the preferences of their constituents in their legislative speeches, research by Lin and Osnabrügge (2018) found that MPs in the German Bundestag were even willing to simplify their language when representing constituents from poorer, less educated or immigrant backgrounds.

The incentive for reelection-seeking legislators to more closely represent the interests and policy preferences of their constituents is well-documented (Auel and Umit, 2018; Kam, 2009; Norris, 1997), and this advantage would be most sought after in more electorally vulnerable seats (Auel and Umit, 2018; Cain et al., 1987; Kellermann, 2016).[1]. Specific studies in the UK have also shown that MPs do place higher priority on their role as constituency representatives relative to other competing aspects of their job (Campbell and Lovenduski, 2015). This can of course also be conducted beyond the legislative, with time spent on constituency-oriented tasks such as conducting MP surgeries, responding to emails, phone calls and written letters from their constituents, and general activity within the local community, all of which have increased substantially in the UK since the 1950s (Gay, 2005, p.58).[2] Since the turn of the century, the role of the Web has increasingly factored into constituency engagement, originally starting with static one-way channels such as MP websites (Ward and Lusoli, 2005), e-newsletters (Jackson, 2006) and weblogs (Francoli and Ward, 2008; Jackson, 2008) in the early to mid-2000s, to using more interactive social networking sites like Facebook and Twitter from the late 2000s onwards (Jackson and Lilleker, 2009).

6.2.3 MPs on Social Media

Since then, the use of social networking sites by parliamentarians in the UK and around much of the world has grown exponentially, with Twitter specifically becoming the platform of choice for the majority of politicians and political journalists. Before the takeover of Twitter and subsequent conversion to X by Elon Musk in late 2022, 591 of the 650 UK MPs (91%) had active accounts on the platform (Gaughan, 2024). Many of the earlier studies looking at the way in which UK parliamentarians engage with their constituents and the wider public on Twitter was conducted in the early to mid-2010s, when social media usage by politicians was in its infancy and only a fraction of MPs were regular users (Agarwal et al., 2019, p.27) One such study by Jackson and Lilleker (2011) was one of the first to acknowledge the

[1] Though it is important to note that the results of Kellerman's study show that electorally vulnerable members uses parliamentary questions to signal their effort to constituents, rather than attention to constituency issues specifically.

[2] This increase is as much driven by a growing demand from constituents as it is by supply from MPs (Campbell and Lovenduski, 2015; McKay, 2020, p.5).

emerging importance of the “microbloggingsphere” to UK MPs and identified two primary functions that Twitter appeared to serve for the small group of MPs who were regular users of the site: impression management (Goffman, 1959) and, to a lesser extent, constituency service. However, they also note (p.14) that where MPs did use Twitter to enhance their constituency role, it was to promote their own activity within the community thus acting as a secondary form of impression management. A much more comprehensive and up to date analysis of how MPs in the UK use social media and for what purpose found that, when it comes to constituency representation, MPs use Twitter for more nationally related content and Facebook for more local issues (McLoughlin, 2019, p.172). This finding is reaffirmed by the results of semi-structured expert interviews conducted by Earl (2023) where all the MPs spoken to acknowledged that Facebook and Instagram were used far more for direct constituency engagement, but Twitter was where they spent a lot of their time. This also meant that their Facebook and Instagram accounts were often run by an MP’s team whereas Twitter was more often run by themselves.

The largely one-way “broadcasting” usage of Twitter by MPs is further highlighted by the fact that of all the tweets that MPs sent in the dataset collated by McLoughlin (2019), only 5% were comment replies (p.166). When an MP did respond to another user, it was most often to ‘elite’ accounts such as other MPs, journalists or prominent Twitter accounts with large numbers of followers. Additionally, less than 1% of tweets were used by MPs to directly ask for the opinions of their constituents (significantly lower than on Facebook) and only 3% of retweeted accounts were ordinary citizens (pp.167-168). This reemphasises the common notion that Twitter is a predominantly elite space (Dagoula, 2019) and that, ultimately, it is not the optimal platform for two-way constituency-to-legislator engagement. Nevertheless, directly assessing digital engagement between citizens and UK MPs on Twitter, research by Agarwal et al. (2019) found that, despite an information overload imposed on MPs during short focus periods on the platform, they were able to manage their Twitter presence strategically. By using selective replies and prioritisation of local or constituent concerns, they could successfully balance their role as party representatives with the role of hearing and responding to their citizen’s needs (p.36). Thus, Twitter does provide alternative avenues to facilitate direct MP-constituency engagement but these opportunities appear to be limited and highly selective, focused on particular groups of users within certain time windows.

6.2.4 Digital Dyadic Representation

In the end, the exact role of an MP still remains to be precisely defined, and specifically in relation to constituency service. While certain scholars have sought to offer theoretical categories for the functions of an MP (see: Norton, 1994; Searing, 1994), it remains to be seen exactly what representational and communicative purpose their use of social media is supposed to serve. When evaluating the impact of the emerging digitalisation of politics in the late 1990s, Campbell et al. (1999) considered MPs to represent one of four main groups: (1) their constituency; (2) their party; (3) Parliament itself; and (4) their nation. There is ongoing debate about in which order of priority these four groups should be (or are) ranked in the minds of an MP. Evidence from McLoughlin (2019) would suggest that, certainly in the case of UK MPs and Twitter, constituency representation is lower down on this list. Nevertheless, this paper believes that, while direct MP-to-constituent engagement on Twitter may be low, the digital profile that an MP curates for themselves on the site (whether intentional or otherwise) may still be in some way responsive to the position of their constituents beyond party affiliation. The justification for this belief is primarily rooted in Stanyer's early work on the presence of elected representatives on the Web, and the theory that they use online websites (in this case, their Twitter profile) as a space to present themselves to their constituents (Stanyer, 2008). They conclude that even in party-dominated systems like the UK, where party loyalties trump personal identity, evidence of constituency involvement could still be observed in the online self-presentation of a very high proportion of MPs. It is further hypothesised that this is most likely shaped by the desire to be re-elected, the likelihood of which can be improved by fostering a better constituency relationship.

Taking together the established knowledge that there is at least some electoral advantage to be gained by aligning more closely with constituents, and the idea that MPs may use their online profiles as a form of impression management to cultivate personal support (Jackson and Lilleker, 2011; Stanyer, 2008), there is good reason to suspect the possibility of some form of dyadic relationship between an MP's Twitter profile and their constituency. Methodologically, this is assessed along three dimensions, making the following assumptions:

1) **Follower Networks:** An MP's followership relative to other MPs acts as a signal about how they are perceived ideologically by ordinary users on the platform (Gaughan, 2024);

2) **Positive Engagement Networks:** The average ideological position of accounts that an MP positively engages with (retweets/@mentions/replies to) acts as a signal about their own ideological position;

3) **Tweet Topics:** The frequency of topics that an MP tweets about acts as a signal about what issues MPs are most concerned about.

These digital indicators are then assessed against three constituency-level dimensions:

1) **Economic Left/Right Position**

2) **Social Left/Right Position**

3) **Most Important Issue**

This is an exploratory study that does not look to establish the exact causal mechanisms that might drive digital dyadic representation on Twitter. Instead, it looks to assess whether any such association exists between an MP's Twitter profile and their constituency positions, beyond partisanship alone. The next section will now outline the methodological approach for doing so.

6.3 Estimating Constituency Opinion

In order to assess the association between constituency-level opinion and the digital behaviour of their elected MP, the concept of digital dyadic representation is explored through three studies. Study 1 investigates the link between constituencies' left/right positions and the left/right positions of MPs as revealed through their Twitter follower networks. These left/right positions in both cases are formally referred to as "ideal points" (Enelow and Hinich, 1984). Study 2 investigates the link between constituencies' left/right positions and the positive engagement networks of MPs on Twitter. These positive engagement networks are broken into two types: 1) retweet networks and 2) positive @mention/reply networks. Left/right positions of each MP's positive engagement networks are revealed through the median left/right position of the accounts that they have retweeted and/or have positively mentioned/replied to. Finally, Study 3 investigates the link between constituencies' MII and the topics most tweeted about by each MP on Twitter. These three studies are all assessed over the majority of the last UK parliament, starting from the first day after the date of the 2019 General Election (13/12/2019 00:00:00) up to the date of data extraction from Twitter (21/04/2023 00:00:00). Ideally, Twitter data would be examined across the entire parliamentary period up to the latest general election that took place on 4th July 2024 but major restrictions to data accessibility on the platform were implemented in spring of 2023, limiting data extraction after this point.

For each of the three studies, two sets of variables are estimated to assess the relationship between constituency and MP. In each case, estimated constituency-level positions are treated as the key independent variable and estimated MP positions are treated as the dependent variable. This follows the guiding assumption that MP digital behaviour is in some way responsive to their constituency. The estimation strategies for measuring constituency position and MP position rely on different data sources and methodologies, both of which will be discussed separately and in detail. Each of the three studies will then be addressed in turn, formally modelling the association between estimated constituency positions and the estimated positions of their corresponding MP.

6.3.1 Left/Right Ideal Points

Accurately estimating constituency opinion is conventionally a difficult task. When relying on nationally representative survey data to elicit public opinion, even suitably large sample sizes can become extremely small when disaggregated by individual constituency. As a result, attempting to derive estimates of constituency-level opinion from these data alone can be noisy and inaccurate. An increasingly popular method for circumventing this issue which has been frequently used to model constituency opinion over the last few years in the UK is MRP. MRP modelling is a modern technique that combines information from large national samples with census data about the demographic make-up of each constituency to provide more closely representative estimates (Hanretty, 2020). In order to estimate constituency-level opinion using this technique, three sources of information are required: 1) survey data with information on the respondent's opinion(s) of interest, as well as their demographic characteristics and the local area they belong to; 2) information on the relevant characteristics of the local areas in question; and 3) a poststratification frame containing information on the joint distribution of respondent demographic characteristics by local area (Hanretty, 2020, p.632). Combining this information together, an MRP model can be estimated in two stages. Stage 1 involves fitting a standard multilevel regression model using respondent measures from the survey data as individual predictors and measures from the constituency data as district-level predictors, estimating the chosen opinion of interest. In Stage 2, this model is then used to predict the variable of interest for every row in the poststratification frame, weighted by the proportion of that constituency made up of the combination of characteristics in each given row, and then summed to give the overall estimate for each constituency. The primary benefit of estimating constituency opinion in this way is that it can effectively reduce the amount of noise around constituency-level estimates by first "smoothing" individual predictions towards the national sample average and then subsequently smoothing predictions toward average opinion among respondents from characteristically similar constituencies (Hanretty et al., 2017, p.242)

Applying this method to this research, an MRP modelling strategy is used to estimate constituency-level opinion in the UK using 5 waves from the BESIP data (Fieldhouse et al., 2023). These waves are selected as they encompass the majority of the observation window (13/12/2019 – 21/04/2023) starting with Wave 20 (June 2020) up to the latest Wave 25 (May 2023). Wave 24 is excluded as it is a supplemental wave with a significantly lower sample size than the others. Ideally, Wave 19 (the post-2019 General Election survey) would have also been included but does not contain the required variables. On average, each wave of the BESIP has circa 30,000 respondents but only participants who are present in every wave are retained, leaving a sample of 7,100 respondents. For Study 1 and Study 2, estimates of left/right position for each constituency are required. To first generate left/right ideal point estimates for the

7,100 respondents across each of the 5 waves, two sets of value-based questions are used. An economic left/right ideal point is estimated based on individual responses to five statements relating to economic values (e.g. ”Government should redistribute income from the better off to those who are less well off”) and a social left/right ideal point is estimated based on individual responses to five statements relating to authoritarian-liberal values (e.g. ”Young people today don’t have enough respect for traditional British values”).[3] Responses to each statement are given on an ordinal scale ranging from 1 [Strongly Agree] to 5 [Strongly Disagree]. To generate individual left/right ideal point estimates along these two dimensions, a multidimensional scaling technique known as multiple correspondence analysis (MCA) is applied to the responses to these statements (Greenacre and Blasius, 2006). This method is an effective way of capturing underlying relationships between variables in a contingency table such that respondents with similar sets of responses will have ideal points closer together in the dimensional space and vice versa. Studies have shown that these methods of spatial estimation closely approximate more traditional ideal point models at a much reduced computational cost (Bonica, 2014, p.369). The ideal point distribution of respondents along the economic and social dimensions is illustrated in **Figure 6.1**.

Grouping ideal points of respondents by voting intention confirms that these estimates have a high degree of face validity.[4] Next, separate multilevel regression models are fitted for both sets of ideal points and for each of the five waves respectively. At the individual-level, five demographic predictors are included in the model: sex, age group, highest educational qualification, home ownership and social grade. At the constituency-level, a number of demographic, political and geographic predictors are included in the model such as age, gender, ethnic and religious compositions, party vote shares at the 2019 General Election, % Brexit leave vote, logged population density, and region. Constituency data is obtained from the latest BES constituency results file for 2019 (Fieldhouse et al., 2019). Only complete observations in every wave of the data are used, meaning any participants missing an ideal point or a recorded MII in any of the five waves are dropped, leaving a final sample of 3,357 observations in each wave. A Bayesian modelling strategy is adopted, estimating these models using Markov Chain Monte Carlo methods and implemented in R using the *brms* library (Bürkner, 2017). For ideal point prediction, linear multilevel regression models are fitted using the gaussian family combined with the identity link, and weakly informative priors of normal(0,1) are set for every parameter. Four chains are estimated in parallel for each model with 5000 iterations and 2500 burn-in. The *brms* library uses the Stan language on the back-end and makes use of the

[3]For further details about respondent ideal point estimation, see *Appendix C.1*.

[4]It is important to note that these ideal points are relative spatial coordinates along each latent dimension and have no meaning in absolute terms. For instance, Conservative voters are not necessarily right-wing economically, but simply further to the right on average than the other respondents.

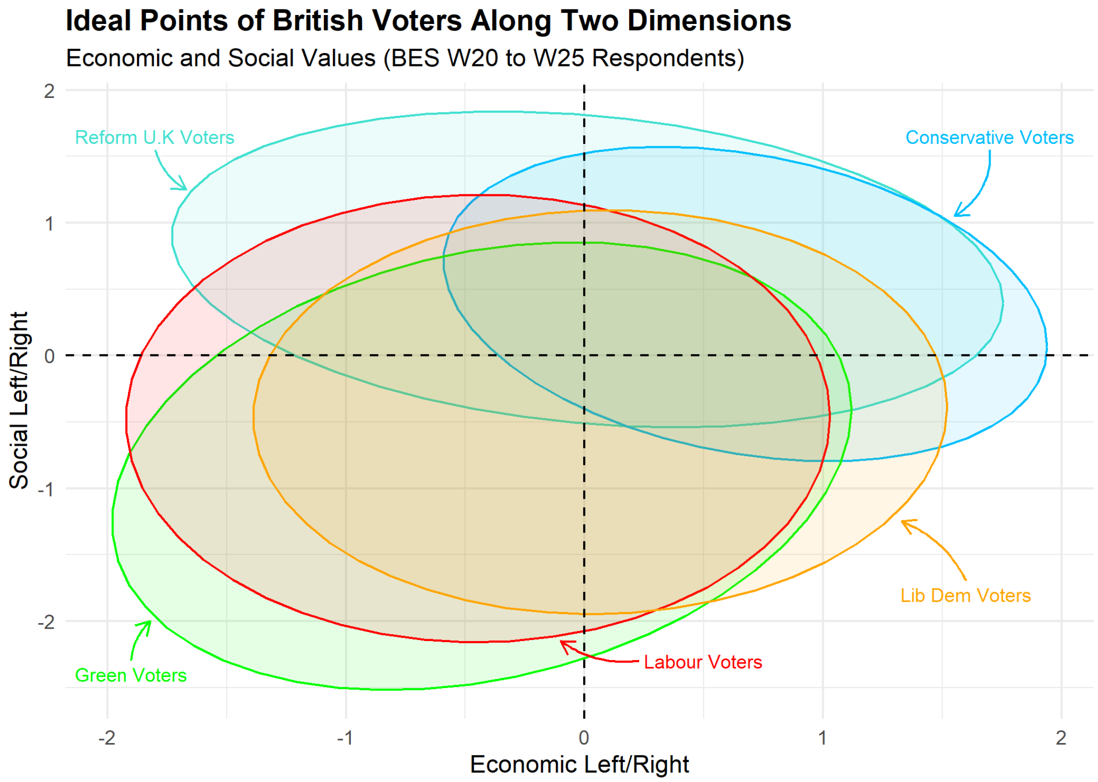


FIGURE 6.1: The two-dimensional left/right ideal point distributions of respondents across Wave 20 to Wave 25 of the British Election Study Internet Panel. Respondents are grouped by voting intention and excludes those who said they would vote for the SNP, Plaid Cymru, Other, wouldn't vote, didn't know or didn't respond (n = 19,564). Elipses represent 80% confidence levels for each group.

Hamilton Monte-Carlo algorithm for sampling. For more details about the modelling strategy, model variables and model diagnostics, see *Appendix C.2.1*.

Once the multilevel regression (MR) models are estimated, the final step in the MRP strategy is to conduct poststratification (P). Owing to the time and complexity of building a poststratification frame from scratch, an existing frame is taken from Hanretty (2020) which contains each possible combination of the five demographic predictors used in the regression models for each of the 632 constituencies in mainland Britain (the 18 Northern Irish constituencies are excluded as they are not included in the BESIP data). Each row in the frame also contains a weighting variable which represents the proportion of each constituency that is made up of each combination of demographic variables. To generate constituency-level estimates of left/right economic and social ideal points, the fitted models are used to predict the left/right ideal point of each row in the frame based on their combination of individual characteristics and the characteristics of their constituency. These predicted ideal points are multiplied by the weight variable and then summed together by constituency to generate a final constituency-level ideal point estimate (rescaled to have a mean of 0 and a SD of 1). While there is no "ground-truth" measure of constituency position that can be used to validate these estimates, **Figure 6.2** demonstrates that the MRP model has high face validity across both dimensions.

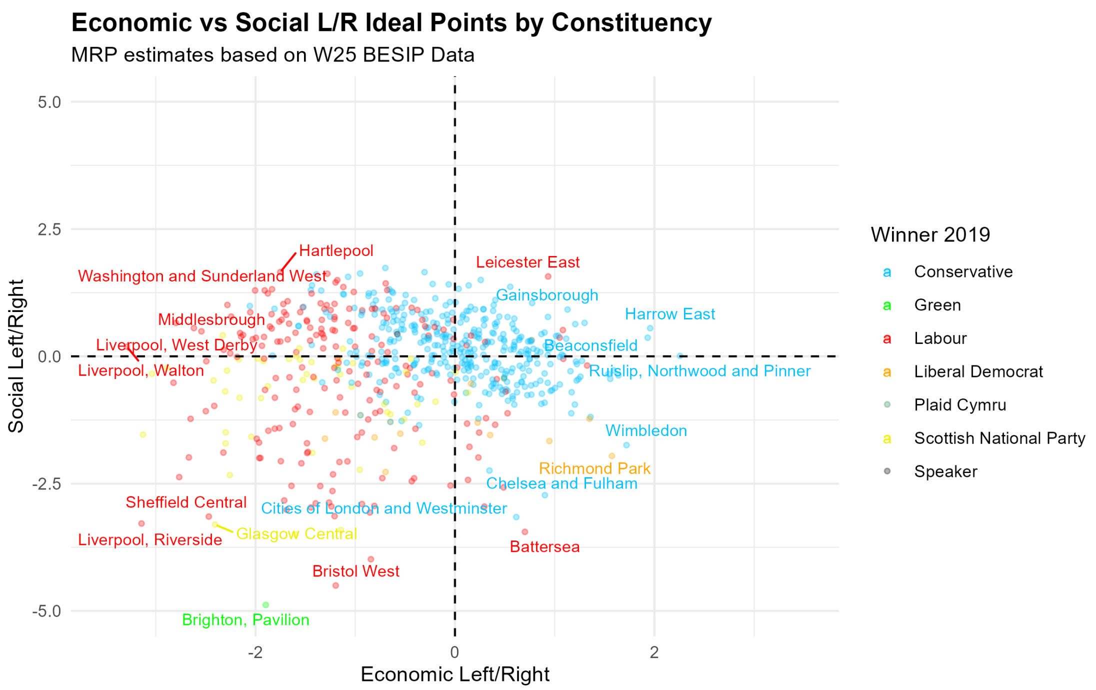


FIGURE 6.2: The two-dimensional left/right ideal points of 632 UK constituencies based on respondent data from Wave 25 of the British Election Study. Each point in the scatter plot represents a single constituency and are coloured by the party winner in the 2019 General Election. The top 5 constituencies in each quadrant are labelled by name.

According to the model based on the most recent Wave 25 data, the top 5 most economically left-wing and socially liberal constituencies are Brighton, Pavilion (a Green party stronghold since 2010) , Bristol West (now Bristol Central, a new Green party seat since GE:2024), and the major urban centres of Liverpool, Riverside, Sheffield Central, and Glasgow Central (characteristically Labour strongholds with large student populations, though Glasgow Central has been SNP since 2015). Conversely, the top 5 most economically right-wing and socially conservative seats include a Conservative stronghold in rural Lincolnshire (Gainsborough), three strongholds in affluent areas of Greater London (Harrow East; Ruislip, Northwood and Pinner) and Buckinghamshire (Beaconsfield), and Leicester East, a former Labour stronghold which became the Conservative Party's shock sole electoral gain in the 2024 General Election.

6.3.2 Most Important Issue

For Study 3, an estimation of the general issues that constituencies care about the most is required. As part of every wave of the BESIP, respondents are asked the following question: *"As far as you're concerned, what is the SINGLE MOST important issue facing the country at the present time?"*. They are then given the option to provide an open-ended response to the question which is manually recoded into one of 50 categories before being further collapsed into 12 broader groups. These 12 groups include issues such as the economy, Europe, the environment, health, and immigration. Owing to smaller numbers of observations, "Terrorism" is folded into the "Other lib-auth" category (which also includes issues relating to crime, war, and defence) and "Other left-right" is folded into the "Other" category. The same MRP strategy is then applied to estimate constituency opinion across each of the 5 waves, with the only difference fitting a categorical multilevel model with the logit link instead of gaussian. When making predictions on each row of the poststratification frame using a categorical model, the model returns a probability estimate for each individual category reflecting the likelihood that the category would be their most important issue. As before, these modelled probabilities in each row are then multiplied by the weight variable and summed by constituency to generate a set of constituency-level probability estimates for each MII. The modelled probabilities of each MII for each constituency across all 5 waves is illustrated in **Figure 6.3**

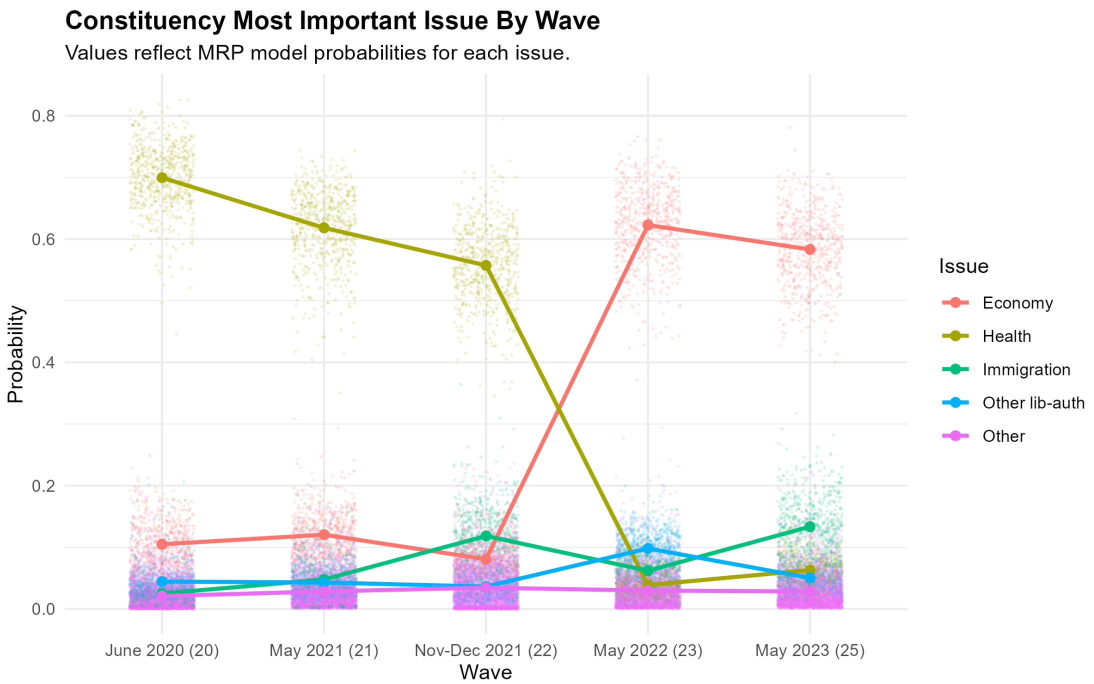


FIGURE 6.3: The modelled probability of each issue being a constituency's most important issue in each wave of the British Election Study. Each individual point represents a single constituency and are coloured by issue. Larger points connected by a line represent the mean probability for each issue from wave to wave. The six issues with the lowest mean probabilities across all the waves are folded into the Other category for easier readability.

In Waves 1 to 3, which spans June 2020 to December 2021, health was by the far the most important issue for constituencies across the UK, followed by the economy and immigration. This is intuitive given that the COVID-19 pandemic was at its height during this period of time. However, in Wave 4 and 5, May 2022 to May 2023, the importance of health drops significantly and is overtaken by the economy as the most important issue. Again, this is intuitive given that the prevalence of COVID-19 had fallen by May 2022 and the UK had begun to face a severe cost of living crisis by the end of 2021. In Wave 4, the Other auth-lib category becomes the second MII for the average constituency which likely reflects concern over Russia's invasion of Ukraine a few months prior while in Wave 5, immigration rises back to the second MII. For more details about the MII categories and model diagnostics, see *Appendix C.2.2*.

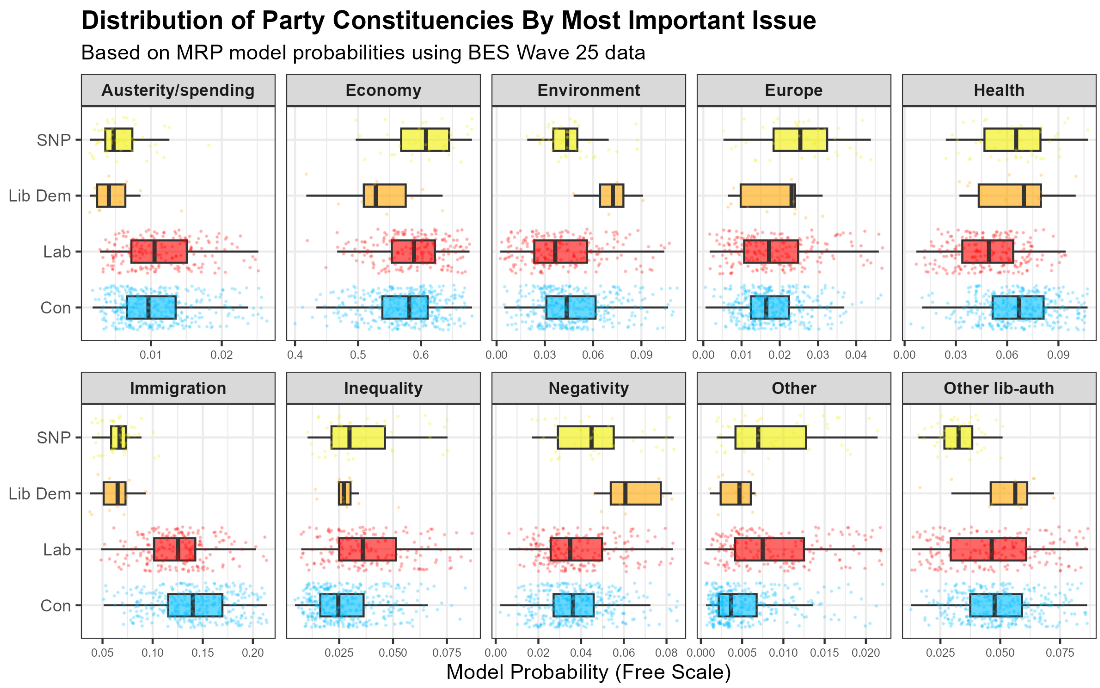


FIGURE 6.4: The modelled probability of each issue being a constituency's most important issue in Wave 25 of the British Election Study, grouped by party winner in the 2019 General Election. Individual points represent a single constituency and are coloured by winning party. Boxplots illustrate the distribution of constituency probabilities by party. Model probabilities along the x-axis are on a free scale, so group distributions should only be compared within each issue and not between. The top 5% of observations in each issue are filtered out for easier readability.

Notable discrepancies between constituencies held by the four largest parties can be identified for multiple MIIs. In regard to the two overall most prominent MIIs identified in **Figure 6.4**, health appears to be a more important issue on average for Conservative constituencies than compared to Labour, whilst the economy is a bigger concern on average for Labour and SNP constituencies. The environment is by far a more important issue on average for Lib Dem constituencies than any of the other three parties whilst Europe looks to be a more important issue for SNP constituencies. Conservative and Labour constituencies both demonstrate much higher levels of concern for immigration than the two smaller parties, though this is higher for Conservatives than for Labour. Other interesting discrepancies also exist within other MIIs, where concern for austerity/spending and inequality is moderately higher for Labour constituencies than Conservative, and negativity (when respondents are negative about specific political parties or politicians, the current state of politics more broadly, or societal divides) is higher among the SNP and Lib Dem constituencies on average than compared to the nation's two major parties. These observations affirm that important variations in MII do exist between characteristically different constituencies.

6.4 Testing Digital Dyadic Representation

Having described the methodology for generating constituency-level ideal point and MII estimates, each of the three studies will now be addressed in turn, explaining how MP positions are estimated in each case respectively, and modelling digital dyadic representation for each study:

6.4.1 Study 1: MP Twitter Follower Networks

Study 1 assesses digital dyadic representation of MPs to their constituencies based on the relationship between each constituency's estimated left/right ideal point across two dimensions (economic/social) and the left/right ideal point of their corresponding MP estimated via their Twitter follower network. Left/right ideal points for MPs are derived from previous analysis conducted by Gaughan (2024) which estimated left/right position of 591 UK MPs who had active Twitter accounts as of the 22/08/2022. To generate these estimates, complete sets of followers were harvested for each and every MP over a three-and-a-half-week period resulting in a combined total of over 34 million user accounts. These user accounts were initially filtered to exclude any users with less than 25 followers, less than 100 tweets, and following less than 10 UK MPs (p.3-4.) A follower adjacency matrix was subsequently constructed with the MPs forming the columns and the remaining 424,297 users forming the rows and, following an estimation strategy first established by Barberá et al. (2015), ideal points for each MP were generated by applying correspondence analysis to this matrix.

The general principle that underpins this method of ideal point estimation is that homophilic clusters that we know exist on social networking sites such as Twitter are heavily driven by an ideological component (Colleoni et al., 2014). When applying a dimensionality reduction technique like CA to such data, it can subsequently reveal this ideological component from which one can elicit left/right ideal point estimates for political elites. Specifically, given the assumption that ordinary users have a higher probability of following political elites they perceive to be closer to themselves ideologically, MPs with more similar follower networks on Twitter are more likely to sit in closer proximity to one another in the latent ideological space and vice versa (Barberá et al., 2015). As part of Gaughan's (2024) paper, these left/right ideal point estimates were validated against an expert survey conducted with 70 academics with specialised knowledge of British legislative politics and were confirmed to have both high between-party (R2 = 0.93) and within-party (Con: r = 0.84; Lab: r = 0.81) accuracy (p.6).

Consequently, these estimates can be treated as valid proxies for MP left/right position. From a substantive perspective, these ideal points can be considered a direct reflection of how they are perceived ideologically by politically engaged Twitter users.[5] Uniquely, unlike other sources of ideal point estimation such as roll-call data or legislative speeches, latent ideological position in this case is signalled passively via the conscious decision of wider Twitter users to follow (or not to follow) them. While MPs do have the ability to curate their follower networks to an extent by manually blocking/removing followers (or reviewing/accepting followers when using a private account), they cannot *force* users to follow them. As such, this method of ideal point estimation relies directly on the perception of hundreds of thousands of ordinary Twitter users to gauge ideological position of MPs and, given that British Twitter users are typically more highly educated and politically engaged than the general population (Blank and Lutz, 2016; Mellon and Prosser, 2017; Sloan, 2017), they can be *"treated as 'experts' who are 'rating' elites through their decisions of whom to follow"* (Barberá, 2015, p.77).

To model the relationship between constituency left/right ideal points and the left/right ideal point of their corresponding MP, OLS linear regression is conducted. Six models are fitted in total for all 575 MPs who had an ideal point available (and excluding Northern Irish MPs), with left/right position modelled as a function of (1) party affiliation alone; (2) constituency economic left/right position alone; (3) constituency economic left/right and party affiliation; (4) constituency social left/right position alone; (5) constituency social left/right and party affiliation; and (6)

[5] Gaughan's (2024) paper shows that the majority of Twitter users who follow at least one UK MP will only follow one. Thus, politically engaged users were classified as those who followed at least 10 or more.

constituency economic left/right , constituency social left/right and party affiliation . As the follower network data was extracted from Twitter at a fixed point in time (22/08/2022 – 15/09/2022), ideal points cannot be assessed temporally. Thus, only constituency left/right position for Wave 23 is used as this survey was conducted closest to the Twitter follower extraction date (May 2022). Results of these regression models are shown in **Table 6.1**.

	(1) Party Alone	(2) Econ L/R Alone	(3) Econ L/R + Party	(4) Social L/R Alone	(5) Social L/R + Party	(6) Econ L/R + Social L/R + Party
(Intercept)	1.263***	0.376***	1.283***	0.376***	1.252***	1.268***
	(0.015)	(0.034)	(0.017)	(0.04)	(0.015)	(0.017)
Const. Ideal Point (Econ)		0.579***	-0.035*			-0.025
		(0.034)	(0.015)			(0.015)
Const. Ideal Point (Social)				0.298***	0.042***	0.038**
				(0.04)	(0.012)	(0.012)
Party (relative to Con)						
Labour	-2.014***		-2.056***		-1.995***	-2.027***
	(0.024)		(0.03)		(0.024)	(0.031)
Lib Dem	-1.682***		-1.676***		-1.621***	-1.621***
	(0.072)		(0.072)		(0.073)	(0.073)
SNP	-1.741***		-1.804***		-1.710***	-1.759***
	(0.042)		(0.05)		(0.042)	(0.052)
Other	-1.743***		-1.783***		-1.710***	-1.742***
	(0.085)		(0.086)		(0.084)	(0.086)
R-squared	0.93	0.33	0.93	0.09	0.93	0.93
N	575	575	575	575	575	575

TABLE 6.1: Study 1: OLS Regression Results. *Note:* *p<0.05, **p<0.01, ***p<0.001

Results for models 1-5 indicate significance associations between constituency left/right positions along both ideological dimensions and the left/right ideal point of corresponding MPs. Interestingly however, the direction of association is reversed between the two dimensions, where constituency economic left/right is negatively associated with MP ideal point and social left/right is positive. These relationships remain statistically significant at the 0.05 confidence threshold when controlling for party affiliation, though the effect magnitudes drop substantially. However, when including both a constituency's economic *and* social left/right position in model 6, the relationship between the economic and MP position is no longer statistically significant.

Interpreting these findings then, they suggest at the broadest level that a positive dyadic relationship between an MP and their constituency does exist along a social (lib/auth) dimension (as estimated via MP follower networks), but this does not hold along an economic dimension. The simplest explanation for this is that individuals place more salience on a social dimension than an economic one. Post-Brexit realignment elevated the social dimension in British politics, where Conservative gains in former "red-wall" seats were built on social conservatism, and this may be reflected in the dominance of the social dimension over the economic. It may also be the case that economic policy is more party-centralised than social issues, where an MP's economic position is more aligned with their party but they demonstrate more freedom over their social position. In the context of Twitter as a platform, these findings may also shine some light on the nature of ideological networks on the site. Where ideological connections *are* present, these may be driven more by a social dimension than economic which is logical given that social positioning is generally the more visible of the two.

To assess how these relationships hold within individual parties, additional regression models were conducted on subsets of MPs from the four major parties, and these are visualised in **Figure 6.5**.

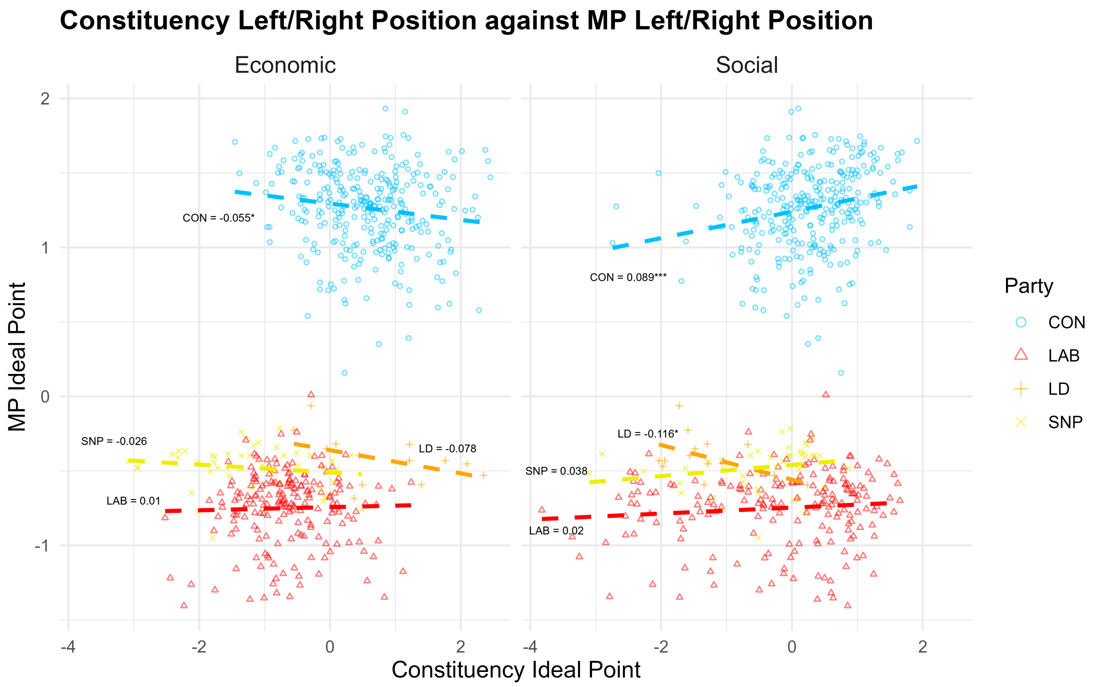


FIGURE 6.5: Scatter plot showing the relationship between 575 constituency left/right ideal points (x-axis) split across two dimensions (economic and social) and the left-/right ideal point of their corresponding MP (y-axis). Each individual point is a single constituency coloured and shaped by the party affiliation of their MP. The dashed lines are linear regression lines for each party group respectively.

When disaggregating these models by party affiliation, one can see that the associations indicated in the overall model trends are primarily driven by Conservative MPs. Along an economic dimension, the negative relationship between ideal points is present among all parties except Labour (though this is only statistically significant for the Conservatives) and along a social dimension, the positive relationship holds for all parties except the Lib Dems (again, this is only statistically significant for the Cons). A particularly interesting finding is the negative association along both dimensions for Lib Dem MPs, especially along the social dimension which goes against the general trend and is statistically significant. This indicates a direct mismatch between the left/right position of their constituencies and their own position as generated via their social media followership. Again, the post-Brexit context of this study may help to elucidate these findings. The contrasting associations between constituency economic and social positions and Conservative MPs may be a direct consequence of the fact that Conservative MPs won a sizeable number of seats in the 2019 General Election from the "red wall". Historically Labour-supporting areas, these are characteristically working class constituencies in the North and the Midlands where one might expect an alignment of constituency to Conservative MP along social ideology, but a disconnect economically.

6.4.2 Study 2: MP Twitter Positive Engagement Networks

Study 2 assesses digital dyadic representation based on MP Twitter positive engagement networks. These engagement networks are ascertained through two mechanisms: retweets and positive mentions/replies. Unlike Study 1 which tests the relationship between constituency position and the position of the MP themselves elicited via follower networks, this study is concerned with the relationship between constituency position and the average position of accounts that MPs *positively engage* with. Where Study 1 works on the assumption that follower networks that MPs organically curate over time will be ideologically associated with their constituency's position to some extent, this study assumes that MPs will positively engage, on average, with wider accounts more ideologically aligned with their own constituency. The method for measuring this relies on the same data used in Study 1. Alongside the estimation of ideal points for UK MPs via their Twitter follower networks, the model simultaneously estimates ideal points for the followers based on the MPs that they choose to *follow.* As such, the data generated by Gaughan's (2024) original study also provides ideal point estimates for 424,297 wider accounts that followed at least 10 MPs at the time of data extraction. As part of a follow-up study by Gaughan (nd, p.11), the ideal points of a sample of wider political and media elite accounts within the dataset were validated against a *YouGov* survey and found to have a high degree of accuracy against the ideological perception of the general public (R^2 = 0.61).

Thus, this data set provides valid ideal point estimates for a wide range of wider political and media elites (primarily based in the UK), many of whom will have been engaged with on a regularly basis by UK MPs on Twitter. To ascertain the average ideal point position of MP endorsement networks, the median ideal point of all accounts that each MP either retweeted and/or positively @mentioned/replied to on Twitter at any point during the observation timeframe (13/12/2019 00:00:00 - 21/04/2023 00:00:00) is taken.[6] While there is explicit assumption made here that retweets do, by default, equal positive endorsement (Metaxas et al., 2014), original tweets posted by MPs that @mention or reply to other users could be either positive or negative (or neutral), significantly changing the nature of that connection. Thus, to only use tweets that are positively coded, sentiment analysis is conducted to categorise

[6]The method for classifying tweets as replies and/or @mentions simply relies on the presence of any @mentioned username(s) embedded within the tweet. It is difficult to definitively separate tweets that are replies from tweets that just @mention another user(s) which is why they are grouped together. If a user replies to a thread, the username of every account who has posted in the thread is returned meaning that some replies might not be directly replying to the @mentioned user.

these tweets as either positive, negative or neutral using the *vader* library in R. Overall, the original retweet dataset contains 1,320,899 retweets by MPs over the observational timeframe and the original mentions/replies dataset contains 637,438 tweets.

Filtering the mentions/replies dataset to only include positive tweets, 414,245 tweets were left. Many of these tweets @mentioned multiple user accounts and so these tweets were unnested to a single observation for each individual user account referenced, leaving a final positive @mention/reply dataset containing 773,447 unique connections. To illustrate the clear ideological structure of these two engagement networks, MP retweet and positive mention/reply networks can be seen in **Figure 6.6**.

(A) **Retweet Network:** 12,813 Nodes and 148,086 Edges.

(B) **Positive Mention/Reply Network:** 10,676 Nodes and 116,230 Edges.

FIGURE 6.6: Positive engagement networks of 566 UK MPs between 13/12/2019 and 21/04/2023. For better clarity, plots are filtered to only include accounts that were engaged with at least 10 times and excludes connections between the MPs themselves. Nodes and labels are sized by weighted in-degree and edges are coloured by the party affiliation of the engaging MP. Networks were generated in *Gephi* and use a force-directed layout.

Ideological clustering can be clearly observed in both networks, affirming that MP retweet and positive mention/reply networks are driven to a large extent by an ideological component. To directly quantify this at the account level, these two sets of networks are merged together and each referenced user is matched to their corresponding ideal point generated as part of Gaughan's (2024) study. Any referenced user accounts that do not have a corresponding ideal point are filtered out, leaving a final positive engagement network with 1,398,322 connections from MP to user. Finally, these connections are grouped into five time intervals which correspond to the date in which each BES survey wave is conducted, with each new interval starting from the first day of the first month after the survey was fielded. This is to help account for any movement in constituency position or MP digital behaviour over the study timeframe. The median ideal point of user accounts that each MP positively engaged is calculated for each of the five time intervals respectively and matched with their corresponding constituency's economic and social ideal point for that given time interval (wave). A similar modelling strategy is taken as in Study 1 but to account for serial correlation between observations of the same MP-constituency at different time intervals, generalised least squares (GLS) regression is used instead (Rao, 1982). These models are fitted using the *gls* function from the *nlme* library in R with a specified AR(1) correlation structure. Results are shown in **Table 6.2**.

	(1) Party Alone	(2) Econ L/R Alone	(3) Econ L/R + Party	(4) Social L/R Alone	(5) Social L/R + Party
(Intercept)	1.354***	0.471***	1.349***	0.476***	1.349***
	(0.013)	(0.04)	(0.014)	(0.041)	(0.014)
Const. Ideal Point		0.064***	0.007	0.053**	0.017
		(0.012)	(0.009)	(0.019)	(0.01)
Party (relative to Con)					
Labour	-2.047***		-2.039***		-2.037***
	(0.021)		(0.024)		(0.022)
Lib Dem	-1.616***		-1.616***		-1.591***
	(0.069)		(0.07)		(0.075)
SNP	-1.404***		1.394***		-1.391***
	(0.36)		(0.039)		(0.037)
Other	-1.601***		-1.595***		-1.584***
	(0.074)		(0.075)		(0.075)
AIC	-130.161	1538.899	-124.046	1557.769	-126.737
Log Lik.	72.08	-765.445	70.023	-774.884	71.369
N	2,750	2,750	2,750	2,750	2,750

TABLE 6.2: Study 2: GLS Regression Results. *Note:* *p<0.05, **p<0.01, ***p<0.001

Model results indicate that no statistically significant association between constituency position and the median ideal point of MP positive engagement networks exists at the 0.05 confidence threshold when controlling for party affiliation, along either an economic or social dimension. All five models do confirm that party affiliation is a significant predictor of the median ideal point of MP engagement networks, suggesting that the party clustering observed in Figure 6 is statistically significant (at the very least, along partisan lines if not ideological). When modelling the relationship between constituency economic and social left/right positions and MP engagement networks alone, both models show a statistically significant positive association but these effects are not significant when controlling for party affiliation. These results suggest that, while a between-party relationship between constituency left/right ideal point (econ and social) and the median left/right ideal point of MP positive engagement networks does exist, there is no evidence of within-party responsiveness. In substantive terms, this implies that MPs do engage more commonly with accounts that are ideologically closer to their political party thus indicating that MP positive engagement networks are partisan driven but beyond that, there is no significant relationship between these networks and the constituencies that MPs represent.

6.4.3 Study 3: MP Twitter Tweet Topics

Moving away from networks, Study 3 assesses the dyadic relationship between the issues constituents care about most and the general topics that MPs tweet about. Between the 13/12/2019 to 21/04/2023, 1,232,692 original tweets (incl. replies) were posted by 567 different MPs. Aligning as closely as possible with the BES's issue coding schema for respondent MIIs (see pp.13-23 of the BESIP codebook), which coded responses into one of 50 sub-issues that folded into 12 parent issues, MP tweets were coded into the same sub-categories using a manually curated key term dictionary. 43 of the 50 sub-categories were used for categorising MP tweets and, as previously mentioned, two of the 12 parent categories (Terrorism and Other left-right) were folded into sub-categories of Other lib-auth and Other respectively. The key term dictionary contains 436 terms relating to different sub-categories and closely followed the criteria stated in the BESIP codebook along with some additional knowledge of terms that directly related to each issue. Examples might include "brussels", "eu" and "brexit" for tweets relating to Europe, "covid", "virus" and "pandemic" for tweets relating to Coronavirus, and "environment", "climate" and "global warming" for tweets relating to the Environment. A partial matching strategy was adopted in most cases, allowing for word stems to be used (e.g. "migra" could be used to capture different lexical variations relating to immigration, or "covid" for different variations of COVID-19). However, in certain cases an exact match was used where partial matching would have been too broad (e.g: "eu" or "gp"). Standard pre-processing steps were taken prior to matching: (1) tweets were converted to lower case; (2) @mentioned accounts were removed; (3) punctuation, URL links and emojis were stripped; and (4) stopwords were removed (excluding any that were essential to a key term: e.g. "bank OF england" or "defund THE bbc").

When matching tweets, exception terms were used in certain cases where the meaning of a word changes when used in conjunction with another word or term. For instance, tweets containing the word "war" were coded into the sub-category War except when preceded by the word "culture" which would move it into the Pol Values – Auth sub-category. Exception terms were also used to reduce incorrect matches when using stem words (e.g: tweets containing the stem "migra" would code tweets as Immigration except when using the word "migraine"). Tweets could be coded into more than one sub-category but in cases where a tweet had multiple sub-categories in the same parent category (e.g: Taxation and Unemployment both fall under Economy), this would only count once towards the parent category. Finally, the BESIP's coding schema mentions certain instances where respondents would talk about multiple issues and how these would be categorised and this is matched in the tweet categorisation. For instance, where a tweet talks about the economy-general (Economy) but in relation to Brexit (Europe) this would only be coded as Europe, or talks about homelessness (Housing) but in relation to poverty, this would only be coded as Poverty. More details about the coding strategy, key term dictionary and tweet category descriptive statistics can be found in *Appendix C.3*.

Following this strategy, 507,622 (41%) of unique tweets were coded into at least one sub-category. Aggregating sub-categories up to their parent categories, there were 693,494 overall tweets split across the 10 individual issue groups. For validation, a random sample of 100 tweets were taken from each of the 10 parent categories (proportionally stratified by % of tweets per party) and hand-coded as either in some way discussing (1) or not discussing (0) the given sub-topic to assess the false-positive rate. A random sample of 100 tweets from the non-coded group were also hand-coded as either in some way discussing (1) or not discussing (0) any of the sub-topics to assess false-negatives. Overall, 84.1% of tweets were correctly coded and 76% of tweets in the non-coded group were correctly identified as non-topical. To ascertain a measure of tweet allocation to each topic per MP, all coded tweets were grouped by MP and wave interval (tweets grouped by time in the same way as Study 2) and split by proportion of tweets per topic. Where an MP did not reference a particular category at all in a certain time interval, this was coded as 0. The distribution of tweet categories by MP per wave is illustrated in **Figure 6.7**.

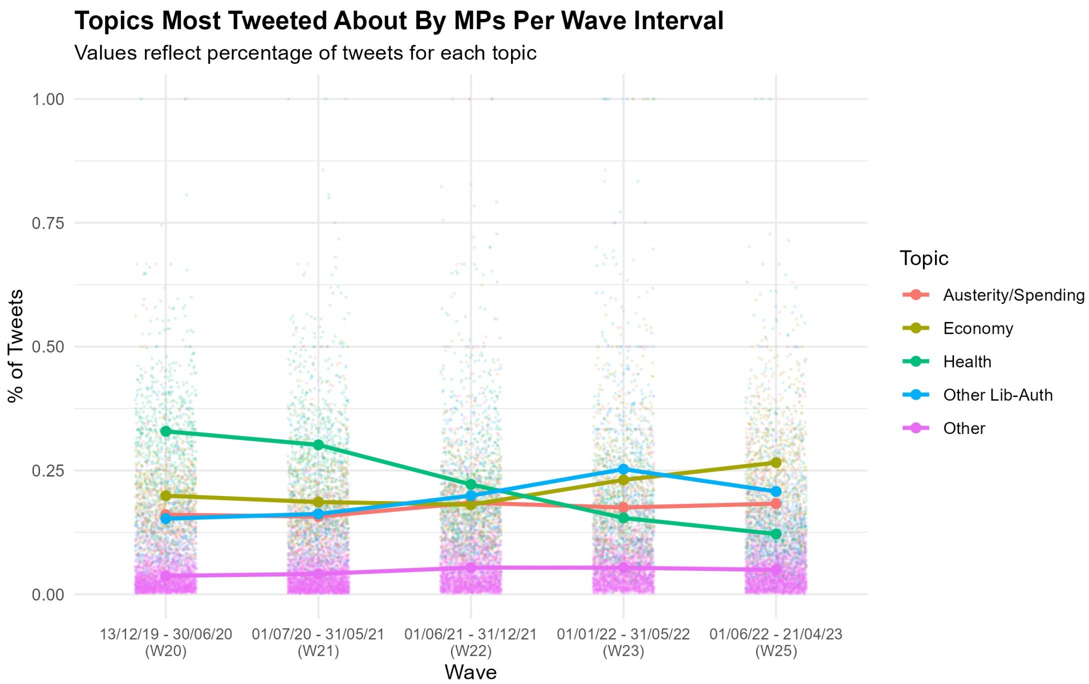


FIGURE 6.7: The proportion of each topic discussed by MPs in their tweets per time interval (wave). Each individual point represents a single MP and are coloured by topic. Larger points connected by a line represent the mean proportion for each topic from wave to wave. The six topics with the lowest mean proportions across all the waves are folded into the Other category for easier readability.

The general proportion of MP tweets related to each topic wave to wave follows a similar trend to that which is observed in the modelled constituency MIIs in **Figure 6.3**. Health is the most tweeted about topic in the first three time intervals but drops below Economy and Other lib-auth in the final two intervals, both of which show an incline in the last three waves. This would suggest a broad association between what the general public care most about and what MPs are more likely to tweet about. The major difference appears to be in the fourth most prominent issue behind health, the economy, and lib-auth topics (war, defence, terrorism, crime etc.). For constituencies it would appear to be immigration, yet MPs tweet far more about austerity/spending (transport, education, welfare, austerity, pensions etc.) and this remains relatively consistent across the five waves. Interestingly, immigration is in fact the least tweeted about topic of the 10, comprising only 2% of all categorised tweets. **Figure 6.8** illustrates how the distribution of MP tweet topics vary by political party, highlighting some degree of partisan variation in tweeting patterns. For instance, Labour MPs on average tweet more about topics relating to inequality, whereas the environment looks to be tweeted about more about Liberal Democrats. The topic of Europe is a far bigger issue for SNP and Lib Dem MPs than for the two major parties, with Labour in particular tweeting far less about the subject. The Other category is more commonly tweeted about by SNP MPs which is intuitive given that this topic contains issues relating to devolution and Scottish independence, whilst Negativity tweets are notably lower for Conservative MPs than the other three parties (this involves negative attitudes towards specific parties, politicians, or the current state of politics/politicians in general).

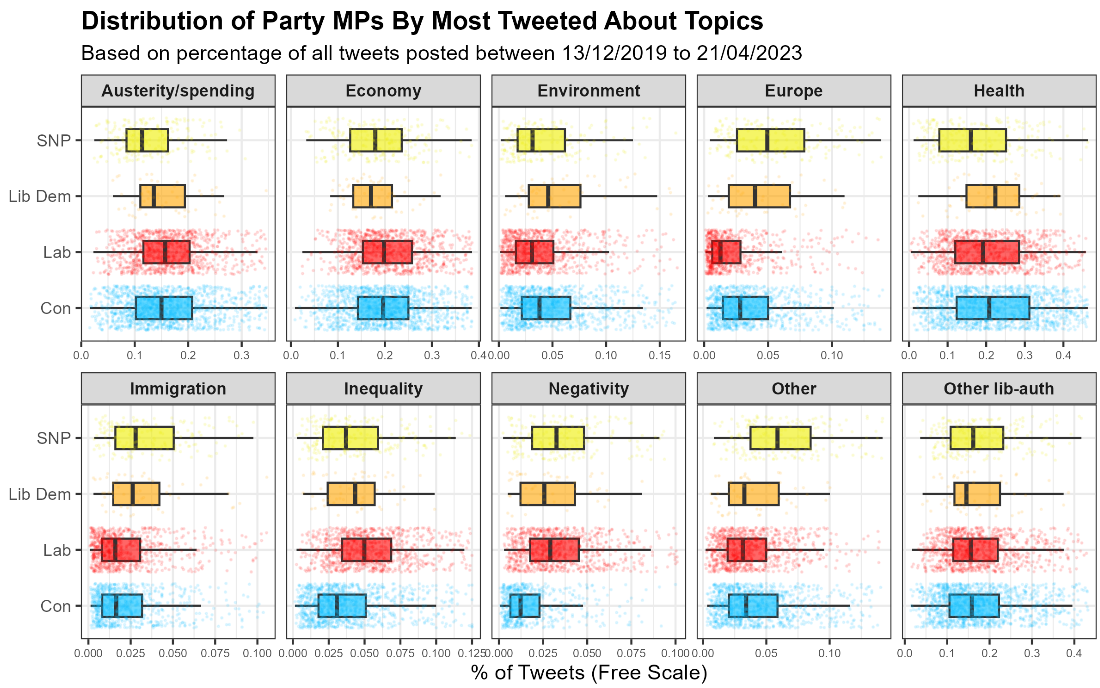


FIGURE 6.8: The proportion of tweets from MPs about each topic, grouped by party affiliation. Individual points represent a single MP and are coloured by party. Boxplots illustrate the distribution of topic proportions by party. Tweet proportions along the x-axis are on a free scale, so group distributions should only be compared within each issue and not between. The top 5% of observations in each topic are filtered out for easier readability.

To formally model the dyadic relationship between constituency most important issues and MP tweet topics, a similar modelling strategy is adopted as in Study 2. The association between MP and constituency across all five waves is again modelled using GLS regression. Each row in the data set is stacked so that each observation is a unique combination of MP-Constituency x Wave x Topic. With 567 unique MPs tweeting during the observation timeframe across five waves, split into ten seperate topics, this results in a stacked data set with 28,350 observations. MPs with less than 10 coded tweets in a specific wave interval are filtered out to account for low sample sizes. This leaves a final data set with 26,290 observations. Ten GLS models modelled are fitted, one per topic, predicting the proportion of tweets about that topic by each MP as a function of the modelled probability that said topic is their constituency's MII. To assess within-party responsiveness, both party affiliation and wave are included as controls. Regression results are shown in **Table 6.3**:

	(1) Austerity/Spending	(2) Economy	(3) Environment	(4) Europe	(5) Health	(6) Immigration	(7) Other Lib-Auth	(8) Inequality	(9) Negativity	(10) Other
(Intercept)	0.17***	0.19***	0.03***	0.04***	0.3***	0.01***	0.15***	0.03***	0.00	0.04***
	(0.005)	(0.006)	(0.003)	(0.002)	(0.003)	(0.002)	(0.007)	(0.002)	(0.002)	(0.005)
Const. Prob	-0.02	-0.03	0.04	0.06	0.05	0	0.09	0.01	0.00	0.04
	(0.12)	(0.03)	(0.04)	(0.038)	(0.038)	(0.02)	(0.06)	(0.03)	(0.025)	(0.062)
Party (relative to Con)										
Labour	-0.01	0.02***	0.00	-0.01***	-0.02**	0	-0.02	0.03***	0.02***	-0.01*
	(0.006)	(0.006)	(0.004)	(0.003)	(0.003)	(0.003)	(0.008)	(0.003)	(0.002)	(0.003)
Lib Dem	-0.02	-0.02	0.01	0.01	0	0.01	-0.03	0.01	0.03***	0
	(0.02)	(0.02)	(0.013)	(0.009)	(0.009)	(0.008)	(0.03)	(0.008)	(0.006)	(0.009)
SNP	-0.05	-0.01	-0.01	0.04***	-0.04***	0.02***	-0.01	0.00	0.03***	0.03***
	(0.01)	(0.009)	(0.007)	(0.005)	(0.005)	(0.005)	(0.014)	(0.005)	(0.004)	(0.005)
Other	-0.05*	-0.02	0.01	0.02*	-0.05*	0.01	0.03	0.02	0.01*	0.01
	(0.02)	(0.02)	(0.014)	(0.01)	(0.01)	(0.009)	(0.028)	(0.009)	(0.007)	(0.01)
Wave (relative to Wave 1)										
Wave 2	-0.01	-0.01***	0.01***	0.01***	-0.02***	0	0.01*	0.00	0.01***	0.01***
	(0.003)	(0.004)	(0.002)	(0.002)	(0.002)	(0.002)	(0.004)	(0.002)	(0.001)	(0.002)
Wave 3	0.02***	-0.02***	0.05***	-0.01***	-0.10***	0.01**	0.04***	0.01***	0.01***	0
	(0.004)	(0.004)	(0.003)	(0.002)	(0.002)	(0.003)	(0.004)	(0.002)	(0.002)	(0.003)
Wave 4	0.01	0.04*	0.01***	-0.01***	-0.15***	0.02***	0.09***	0.01**	0.02***	0
	(0.005)	(0.017)	(0.003)	(0.002)	(0.002)	(0.002)	(0.006)	(0.002)	(0.002)	(0.003)
Wave 5	0.02***	0.08***	0.02***	-0.01**	-0.18***	0.01***	0.05***	0.02***	0.01***	0.01
	(0.005)	(0.012)	(0.003)	(0.003)	(0.003)	(0.003)	(0.006)	(0.003)	(0.002)	(0.003)
AIC	-5992.119	-5977.104	-8386.597	-9824.824	-5329.936	-10164.86	-5396.114	-9530.794	-11039.14	-8850.217
Log Lik.	3008.06	3000.552	4205.298	4924.412	2676.968	5094.428	2710.057	4777.397	5531.57	4437.108
N	2559	2559	2559	2559	2559	2559	2559	2559	2559	2559

TABLE 6.3: Study 3: GLS Regression Results. *Note:* *p<0.05, **p<0.01, ***p<0.001

Results across all ten models indicate no statistically significant association between the modelled MII probability of each constituency and the proportion of tweets posted about that topic by their MP, when controlling for party and time effects. Some interesting effects do exist between parties, however: Labour MPs were significantly more likely to tweet about the economy and inequality than the Conservatives, but less likely to tweet about Europe. With the exception of the Lib Dems, all opposition parties were less likely than the Conservatives to tweet about health. However, they were all significantly more likely to tweet negatively about specific parties, politicians and/or the current state of politics (negativity) than the Conservatives. The SNP were significantly more likely to tweet about Europe and immigration than the Conservatives.[7] SNP MPs also tweet significantly more about "other" topics, which includes issues relating to Scottish independence and devolution. All ten models also

[7]The suprising immigration effect may be an unintended consequence of the keyword coding schema where the SNP's focus on the Anglo-Scottish border may overlap with UK borders.

identify significant time effects, where statistically significant fluctuations in topic prevalence exist between waves.

6.5 Discussion & Conclusion

All in all, this paper has made a novel attempt to explore the dyadic relationship between elected legislators and their constituencies via the digital Twitter profiles of UK MPs. Building on the idea that MPs may use their online profiles as a form of impression management to cultivate personal support for electoral advantage (Jackson and Lilleker, 2011; Stanyer, 2008), it has sought to provide evidence of a within-party responsive of MP Twitter profiles to their constituents along three digital dimensions: 1) follower networks; 2) positive engagement networks; and 3) tweet topics. Leveraging a contemporary technique for small-area estimation in the form of MRP to measure constituency left/right ideal points and their most important issues, this study can confirm a number of findings:

1) UK MP Twitter follower networks are individually responsive to their constituency left/right positions beyond party affiliation, and this effect is negative along an economic dimension and positive along a social one;

2) Neither MP Twitter positive engagement networks nor their tweet topics are individually responsive to their constituency positions beyond party affiliation or time-period effects;

3) Partisanship is a strong and consistently significant predictor of MP position across all three digital dimensions, reaffirming the "party broadcasting" nature of Twitter usage by MPs over constituency engagement (Earl, 2023; Jackson and Lilleker, 2011; McLoughlin, 2019);

4) At an overall level, MP tweet topic patterns are broadly aligned with general public concerns and these shift in tandem over time, though the exact causality of this relationship is unclear.

In regards to the specific research question of digital dyadic representation in a Westminster system, finding 1 is perhaps the most illuminating. Of the three digital dimensions explored, MP follower networks was the only one that was responsive to constituency opinion beyond party affiliation. Most interestingly, the results indicate that MP Twitter follower networks are weakly attuned to the left/right positions of their constituencies along a social dimension but are disconnected along an economic one. The exact reason for this is unclear and may warrant further investigation in future research, but could help to decompose the ideological nature of social media networks that we know exist on platforms like Twitter. If the follower networks of political elites on Twitter can be exploited to estimate their ideal points (Barberá, 2015; Gaughan, 2024), the findings of this research may indicate that these estimates are built more on user's perceptions of social (lib-auth) left/right positions than economic. This would align with the fact that political identities are more strongly determined by social attitudes than economic and that the former tends to be more ideologically divisive than the latter (Crawford et al., 2017). It is also important to note that this dyadic relationship appears to be driven primarily by Conservative MPs more so than Labour, Lib Dem or SNP. A straightforward account for this could be that the digital profiles of Conservative MPs are simply more responsive to their constituents than other parties. Alternatively, it may speak to a deeper mechanism behind how MPs from different parties make use of their digital platforms or how differently politicians on the left and the right are perceived by ordinary users.

The lack of within-party responsiveness beyond follower networks alone supports the strength of using followers as an indicator of ideological position compared to other more idiosyncratic signals like retweets, @mentions, or tweet content. These measures are far more prone to low sample sizes and the unpredictability of individual behaviour, whereas follower networks can benefit from the wisdom of the crowd. Nevertheless, MPs still appear to fall in line with their party at a general level and this is observed through the similarities of their positive engagement networks and tweeting patterns. Despite the lack of evidence for within-party responsiveness, the overall relationship between MP tweet topics and the concerns of the general public over time is still important. It is unclear whether this is because MPs are responding to the concerns of the public, whether public opinion shifts in response to the issues MPs bring attention to, or that both move in tandem with the shifting socio-economic landscape. Regardless, it provides some evidence that MPs are at least in tune with the issues that the public care about the most at the broadest level, and as concerns among the public shifted away from health towards the economy and security from 2020 to 2023, the tweet focus of MPs moved with it. It is also important to contextualise these findings within the restrictions of a Westminster system, where within-party responsiveness can be difficult to establish at any level. Thus, it may reemphasise the value of Gaughan's (2024) research, where social media followership is an extremely effective source for disentangling legislators in restrictive parliamentary systems like the UK from their party position.

Overall, the contribution of this research is both methodological and substantive. Methodologically, it combines the informative power of social media "big data" in the form of MP Twitter networks and tweet timelines with a contemporary computational technique for small-area estimation to assess the concept of digital dyadic representation. In doing so, it finds evidence of within-party responsive of Twitter follower networks to an MP's constituency along a social left/right dimension, as well as reaffirming the staunchly partisan nature of Twitter usage by UK MPs more broadly.

Chapter 7

Conclusions

The last decade and a half has seen a seismic transformation in political communication and campaigning, where digital technologies have profoundly changed the way in which politicians and media elites consume and disseminate information. A central part of this digital revolution has been the growth of social media sites where, for the many reasons discussed throughout this thesis, Twitter emerged as the platform of choice for the majority of political and media elites in the Western world. By the time of the last parliamentary period (2019-2024) in the UK, nine in ten MPs had an active Twitter account (Gaughan, 2024) and tens of thousands of wider political and media elites were also present on the site (Gaughan, nd). Never before had an environment existed on such a scale where both elites and ordinary citizens existed together in tandem and their behaviours and social networks could be directly mapped for political research. In its infancy, Twitter had sparked genuine hope of a digital town square: an unrestricted space for open debate and deliberation, with equality in access and participation, and free from elite intervention. Yet, as was discussed in detail in **Chapter 2**, the pervasive effects of offline digital divides, self-selection biases, and inequalities in participation on the platform has left this promise unrealised. Instead, what emerged was a highly unrepresentative digital space packed with users who were disproportionately from more socially privileged groups. This raised a number of concerns about early research studies that had attempted to use data from the platform to make offline observations where a failure to adjust for these biases would lead to erroneous results.

This thesis not only acknowledged these biases but has directly sought to exploit them to improve our understanding of three important concepts within legislative and communicative politics: 1) intra-party competition; 2) media representation; and 3) dyadic representation. This was explored through three distinct papers that formed the body of this research, the results of which are succinctly summarised in the following section.

7.1 Paper Summaries

7.1.1 Paper 1: Modelling the Ideological Component to the Sept 2022 Conservative Party Leadership Contest

Alongside acting as the empirical backbone of the thesis by outlining the methodology for estimating ideological maps of elite actors using social media networks, this paper sought to model intra-party competition in a strict legislative system like Westminster's HoC. Extracting the complete Twitter follower networks of 591 sitting UK MPs in the HoC as of the 22/08/2022, left/right ideal point estimates of each MP were derived using multidimensional scaling in the form of correspondence analysis. These estimates were validated against ideological estimates for a subset of 30 MPs derived from a survey of 70 experts in the field of British politics and confirmed to have a high degree of both between-party (R^2 = 0.93) and within-party (Con: r = 0.84: Lab r = 0.81) accuracy. To demonstrate the informativeness of these ideal points generated via social media followership to disentangle intra-party competition, this paper used these ideal points to model candidate endorsement in the September 2022 Conservative Party leadership contest.

Having conducted face and formal validity checks of the ideal points for all UK MPs with active Twitter accounts, these ideal points were employed as the key independent variable in a logistic regression model estimating the association between each Conservative MP's ideal point and their endorsement of Liz Truss or Rishi Sunak for party leader. Three models were fitted, the first of which only includes an MP's ideal point, the second incorporated a number of demographic control variables such as an MP's age, gender, ethnicity, and educational background, and the third also included a number of political controls such as length of time in office and seat marginality. The results of the model found conclusive evidence that an MP's left/right ideal point *was* a statistically significant predictor of candidate endorsement, where the further to the right (higher) an MP's ideal point was, the greater the likelihood of them endorsing Liz Truss over Rishi Sunak. Additionally, the model also confirmed that gender was a statistically significant predictor of candidate endorsement with females more likely to endorse Truss over Sunak than their male counterparts.

7.1.2 Paper 2: Assessing the Ideological Representativeness of Major UK Television Broadcasters Between 2022 and 2024

Using the left/right ideal points of follower accounts generated in tandem with MP estimates in **Paper 1**, **Paper 2** conducted a descriptive assessment of the ideological representation of guest selection on the flagship political programmes of the UK's major T.V broadcasters between January 2022 and January 2024. Using a *YouGov* survey of 2,068 adult members of the British public commissioned exclusively for this thesis, these wider user ideal points were validated against the mean ideological estimate of 6 UK political parties, 24 media organisations and 30 individual media personalities and confirmed to also have a high degree of accuracy (R^2 = 0.61). The guest lists of seven flagship programmes across six major T.V broadcasters between January 2022 and January 2024 were collated from a number of external sources and their ideal point distributions were examined to assess ideological representation. Two key statistical tests were conducted to assess both the significance of the differences between distributions on each of the seven shows compared to general elite accounts on Twitter (Tukey HSD test), as well as the significance of multimodality within each distribution (Hartigan's dip test).

The results found that all seven political programmes demonstrated an average ideal point to the right of general elite accounts on Twitter to varying extents, and these differences were all statistically significant. Of the seven programmes, the UK's two selected non-public service broadcasters (Sky News and GB News) demonstrated average ideal points furthest to the right of the spectrum. Furthermore, tests of multimodality confirmed that despite the slant in guest selection towards the right on each programme, multimodality was present in six of the seven shows indicating that guests were drawn significantly from multiple parts of the spectrum. This is with the exception of *GB News: The Camilla Tominey Show* where results failed to confirm the presence of multimodality suggesting that this programme drew *almost exclusively* from the right side of general Twitter elites for its selection of guests between 2022 and 2024. Additionally, guest selection across the two year period was also broken into quarterly three month intervals to assess any variation in distributions over time, where results failed to indicate to any statistically significant differences.

7.1.3 Paper 3: Modelling the Within-Party Responsiveness of UK MPs on Twitter to Their Constituency Positions

Employing a contemporary technique for small area estimation in the form of MRP and drawing on five waves of data from the BESIP, **Paper 3** estimated the left/right ideal point positions of UK constituencies as well as their most important issues and assessed their relationship with their MP via their Twitter behaviour. First estimating left/right ideal points of BESIP respondents who appeared in each of the five selected waves between June 2020 and May 2023 by employing MCA on responses two sets of value-based statements (economic and social), constituency-level positions were modelled from these data using MRP. The same strategy was subsequently applied to open-ended responses given to the question of what respondents believed to be the most important issue facing the nation at the present time to model constituency-level most important issue probabilities. Drawing on the left/right ideal points of MPs derived from Twitter networks in **Paper 1** and **Paper 2** as well as the topics of content discussed in their Twitter timelines, **Paper 3** modelled digital dyadic representation across three studies and yielded a number of key findings.

Firstly, Study 1 found that left/right ideal points of MPs estimated via their Twitter follower networks are responsive to constituency left/right ideal points along a social left/right dimension but disconnected along an economic one, and this relationship holds when controlling for party affiliation. Disaggregating this effect by party affiliation showed that this relationship is largely driven by Conservative MPs. Secondly, results of Study 2 found that neither MP Twitter positive engagement networks nor their tweet topics are individually responsive to their constituency positions beyond party affiliation or time-period effects. Finally, Study 3 suggests that, while no evidence of within-party responsiveness of MP tweet content to the most important issues of their constituents could be shown, MP tweet topic patterns are broadly aligned with general public concerns at an overall level with both shifting in tandem over time. One significant effect that held across all three studies is partisanship where party affiliation of an MP was a strong and consistent predictor of constituency position along all dimensions, reaffirming the ideologically embedded nature of MP Twitter usage. Overall, **Paper 3** confirmed that a weak but statistically significant dyadic relationship exists between the Twitter follower networks of UK MPs and the social left/right positions of their respective constituencies beyond party affiliation alone.

7.2 Key Findings of this Research

Taken together as a whole, from **Chapter 2** to **Chapter 6**, this thesis has produced a number of high-level key findings to take away from this research. Summarised concisely and in the order they appear in the body of the thesis, these key findings are as follows:

Key Finding 1: British Twitter users continue to be disproportionately representative of elite groups in wider society.

As is outlined in detail in **Chapter 2**, the reality of British Twitter users as disproportionately drawn from more socially privileged groups in society is not a novel finding. It has been well-established in a number of studies that British Twitter users are, on average, younger, wealthier, more highly educated, more likely to be from a higher socio-economic background, more likely to be male, more politically engaged, and more socially liberal (Blank, 2017; Mellon and Prosser, 2017; Sloan, 2017). However, many of these studies are from the mid-to-late 2010s and the social media landscape is dynamic and ever-changing. Nevertheless, descriptive analysis of Twitter users from much more recent data from the BESIP in June-July 2024 confirms that this bias still persists on the platform even after conversion from Twitter to X. With the exception of age, where TikTok and Instagram users are younger on average, Twitter users are still the most highly educated of any other user group (incl. non-users or the general public), as well as from the highest social grade and with the highest proportion of male users. They are also the most likely of any other group to turnout to vote, they pay the greatest amount of attention to politics, and are the most socially liberal. They are also the least likely to find themselves trapped inside an informational echo chamber when taking into account their wider news diet, with Twitter users getting their political information from a higher average number of sources than any other group.

Key Finding 2: The Twitter follower networks of British MPs can be successfully used to generate valid estimates of their left/right ideal point positions, as well as for the wider accounts that follow at least ten of them.

One of the major implications of this research is the confirmation that the social media followership of UK MPs can be successfully leveraged to estimate valid left/right ideal points for both the MPs themselves as well as for accounts that follow a substantial number of them (circa. 10). Again, knowledge that Twitter networks can be leveraged to estimate ideal points is not strictly novel (Barberá, 2015; Barberá et al., 2015), but as of yet such a method has not been applied comprehensively to the British case. This is particularly important in the UK context owing to its restrictive parliamentary and broadcasting culture and demonstrates in detail the ability that social media offers to circumvent these limitations. Most importantly, these ideal points are both formally validated and successfully applied, completing a proof of concept element to the research which can provide the foundation for further applications of this type of analysis in future studies of UK legislative and communicative politics.

Key Finding 3: Liz Truss drew her support primarily from the further right of the Conservative Party during the September 2022 leadership election, as well as from more female MPs.

The ideological component of the September 2022 Conservative Party leadership contest is relatively well-covered (Jeffery et al., 2023), though many studies still struggle to settle on the best method of quantifying this where roll-call data is uninformative. Whilst Truss' appeal among Conservative MPs from the further right of the party is not a novel finding, the fact that an MP's Twitter follower network was acutely attuned to this is notable and emphasises the informativeness of these data. A point of note is that these follower network ideal points were used to retroactively model candidate endorsement in that leadership contest. However, if such networks are acutely attuned to the intricate differences between MPs from within the same party, it may promote the *proactive* use of these data to predict outcomes of intra-party conflicts and contests in future. In particular, one might assume that these networks are not static and will fluctuate in their general structure over time as users follow (and unfollow) MPs. These changes may provide further insights into the changing patterns of MPs that may be difficult to identify person to person, but could be drawn out when crowdsourcing the behaviour of thousands of highly engaged users.

Key Finding 4: Ordinary user accounts that follow a significant number of British MPs are ideologically skewed to the left of the political spectrum, while elite accounts are much more moderate and evenly spread.

Again, while it is common knowledge that British Twitter users are ideological to the left of the political spectrum along both economic and social dimensions (as is reaffirmed by Key Finding 1), this has not been directly examined on the platform itself. Validated estimates of hundreds of thousands of politically engaged user accounts illustrating a left-wing bias is a novel contribution and serves to reaffirm Key Finding 1 of this of thesis. It also confirms, however, that this distribution is distinctly bimodal, reaffirming the polarised nature of political Twitter. Comparatively, the ideological distribution of official public figures (most of which will be British in this case) who are politically engaged (following at least ten MPs) is much more evenly spread across the axis

Key Finding 5: Major UK T.V broadcasters have selected guests disproportionately to the right of Twitter elites on their flagship political programmes in recent years, and this is most extreme on non-PSBs.

Conventional studies into media representation tend to rely on thematic content analysis (Cushion et al., 2018; Cushion and Lewis, 2017; Morani et al., 2022) and **Paper 2** makes a departure from this literature by using social media networks to quantify political position. Descriptive analysis of representation shows evidence of a slant towards the right of spectrum relative to general Twitter elites in the distribution of guests selected to appear on the flagship political programmes of the UK's six major broadcasters. Among them, Sky News and GB News showed the greatest degree of skew to the right, with GB News specifically drawing their guests almost exclusively from the right of the political axis on *The Camilla Tominey Show*.

Key Finding 6: Between mid-2020 and mid-2023, health, the economy, immigration, and liberal-authoritarian issues were the four biggest concerns for constituents, though the prioritisation of these issues fluctuated over time.

Although not specifically the main focus of the empirical analysis in Paper 3, constituency-level estimates of the British public's most important issues across five waves of the BESIP between June 2020 and May 2023 illustrated that health, the economy, immigration and liberal-authoritarian issues were the four biggest concerns. Intuitively, health was the most important issue for constituencies between June 2020 and December 2021 when the COVID-19 pandemic was at its height, but drops significantly in the final two waves between May 2022 and May 2023. During this time, concerns around the economy shift to the top of the priority list when Britain began to face a severe cost-of-living crisis amid major inflationary pressure. This was followed by immigration and liberal-authoritarian issues (which includes war, crime and defence), the latter of which likely would have increased as a concern following Russia's invasion of Ukraine in February 2022.

Key Finding 7: The left/right positions of MPs as estimated via their Twitter follower networks were dyadically responsive to their constituencies' left/right positions beyond party affiliation along a social dimension, but not economic.

One of the key exploratory research objectives of **Paper 3** was to establish a within-party responsiveness of an MP's digital Twitter profile to the positions of their constituency along any one of a number of dimensions. Although partisanship was the overwhelmingly dominant segregator of MP accounts across followership, positive engagement, and tweet content, a mild association between follower networks and social left/right constituency position was established beyond party affiliation. This is a significant and novel finding and indicates that MP Twitter follower networks are dyadically correlated to the social positions of their constituents beyond the dominance of their party position alone. The exact casual nature of this relationship is uncertain but it connects well with Key Finding 3: MP Twitter follower networks are finely attuned to minor differences between MPs within the same party, however mild the effect. This has major implications for strict legislative systems like Westminster where researchers have historically struggled to disentangle MPs from their party.

Key Finding 8: The retweet and positive mention/reply networks of MPs are highly polarised along partisan lines, though no evidence of within-party constituency responsiveness was found.

Along with follower networks, the positive engagement networks of MPs on Twitter are highly polarised and driven largely by partisanship. There is clear evidence of party clustering among MPs and this is evident in the accounts that MPs will retweet and/or positive @mention or reply to. Unlike follower networks, these networks are not attuned to the positions of their constituents but does reaffirm the party broadcasting nature of MP Twitter usage (Earl, 2023; Jackson and Lilleker, 2011; McLoughlin, 2019).

Key Finding 9: MP tweeting patterns and the most important concerns of the general public are broadly aligned and shift in tandem over time, although no within-party constituency responsiveness could be identified.

It is valuable to understand the primary focus of MP tweet topics and, while no dyadic relationship could be established between MP tweet topics and their constituency MIIs, they were broadly aligned with general public concerns. Matching constituency-level issues, health, the economy and liberal-authoritarian topics dominated the tweeting focus of MPs between 13/12/2019 and 21/04/2023, and their order of priority matched the general public as they evolved over time. Interestingly, however, immigration, while a top four concern for the general public, was in fact the least tweeted about topic for MPs of the ten categories, comprising only 2% of all tweets. This potentially suggests a disconnect between the public and MPs along this particular issue, where constituents view this is a much more important issue than MPs would care to tweet about. The exact cause of this disconnect is uncertain and could warrant further investigation in future work.

7.3 Intellectual Contribution

Taken together, this thesis makes a number of novel contributions to the fields of legislative politics, political communication and digital research. At the very broadest level, the primary purpose of this thesis is demonstrate across three areas of study the informative power of social media data (and digital data more generally). Primairly, this thesis strives to encourage a greater incorporation of social media and digital data into mainstream political study. If harnessed correctly, online data can augment offline data in a myriad of ways, from helping legislative researchers better understand intra-party conflict to aiding media scholars studying representation and bias. A substantial contribution made in **Paper 1** of this thesis is the availability of left/right ideal points of the majority (91%) of UK MPs sat in the HoC during the last parliamentary period (2019-2024) which are formally validated using an expert survey. While the estimation of left/right ideal points of UK MPs is not novel in and of itself (Kellermann, 2012; Hanretty et al., 2017), one strength of this data is in the simultaneous estimation of ideal points for tens of thousands of wider elites and ordinary users which will be of significant interest to researchers. Moreover, at the time of conducting this research, substantial changes to the Twitter platform has significantly restricted the ability to gather large-scale data from this site again in the future, making this thesis one of the last of its kind to conduct a mass study of UK elites on the site.

Paper 2 makes a number of novel contributions of its own, beyond the substantive findings regarding media representation of guest selection on major UK broadcasters. Firstly, it presents a clear illustration of the ideological distribution of both ordinary and elite (primarily British) Twitter users. While it is common knowledge that British Twitter users are disproportionately to the left of the political spectrum (Blank, 2017; Mellon and Prosser, 2017), this is confirmed from a sample of almost half a million users on the site. Additionally, it demonstrates and compares the average position of Twitter's elite users in the UK, which is relatively evenly spread across the spectrum, affirming a disconnect between the general ideological position of ordinary users and elite accounts in the UK. Furthermore, having validly quantified the ideological position of hundreds of thousands of ordinary and predominantly British user accounts, the paper descriptively illustrates the type of language and emojis used in the profile descriptions of accounts on either side of the spectrum. This provides an interesting insight into how the traditional left/right ideological axis in the British context manifests itself on Twitter. This paper also includes a *YouGov* primary survey exclusively commissioned for this thesis which confirms that the left/right ideal points estimated via social media networks aligns closely with the average perception of the general public. This provides a novel understanding of how well the general public conceives of the ideological positions of media and political elites in the UK and how this marries up to the ideal points that can be estimated via social media followership.

Finally, **Paper 3** makes a novel contribution to the field of legislative study in the UK by testing an important concept for electoral systems with SMDs: dyadic representation. As is discussed in detail in the literature review of the paper, the study of dyadic representation has been almost exclusively conducted in the US owing to its atypically loose legislative culture (Bawn and Thies, 2003). Only a handful of studies have attempted to analyse MP-to-constituent responsiveness outside of the U.S, one of which was conducted in the UK by Hanretty and colleagues in 2017 (Hanretty et al., 2017). As is discussed in detail throughout the chapters of this thesis, the UK poses a unique challenge when it comes to ideal point estimation of its legislators owing to its idiosyncratic parliamentary culture, and this is highlighted by Hanretty's study. Following in a similar vein to this earlier research, **Paper 3** further reaffirms the informativeness of social media platforms like Twitter for political research and makes a novel contribution to our understanding of how MPs curate their online profiles and how this may reflect the positions of their constituents. To the best of the author's knowledge, no research to date has attempted to assess dyadic representation in such a way and makes an original effort to combine the scale of social media data with the contemporary technique of small area estimation via MRP modelling. While the high degree of both between and within-party accuracy of the ideal points generated for MPs in **Paper 1** is notable, the fact that the ideological dimension to follower networks is attuned (however mildly) to the individual left/right social positions of each constituency is a fascinating addition to this research space. At a broader level, with the use of digital technology in political communication and electoral campaigning growing ever more central to modern politics, this paper hopes to encourage further study of MP-to-constituency relationships along a digital dimension.

Overall, this thesis strives to highlight the continued value of using digital data in political research. Alongside the many substantive contributions to our knowledge of British politics and media this thesis makes, it has primarily attempted to employ novel ways to assess established concepts in political research using large-scale data derived from the Twitter platform. As previously stated, recent changes to the Twitter platform that took place during this research likely means these studies will be some of the last conducted using data from the site. However, not only does this increase the impact of this research now – such data was extracted at the peak of Twitter's maturity before its subsequent reconfiguration and restriction – it hopefully provides a template for such research to be replicated and built upon if a future social media platform can take Twitter's place.

7.4 Research Limitations and Future Recommendations

Throughout the thesis, a number of limitations to this research have been acknowledged and addressed ranging from the methodological to the theoretical. Perhaps the most glaring of these limitations is addressed in **Chapter 3**, where the use of DTD is fraught with methodological concerns about the stability and representativeness of data sets extracted from the Internet. Chiefly, studies using Twitter data, like many DTD, suffer from major issues around **reproducibility.** Owing to the dynamism and instability of social media data, where profiles and data are edited, suspended, deleted, removed and/or privatised on a daily basis, this can severely impede the ability of researchers to replicate studies where a re-extraction of data from the same platform using the same search query at a different time point can yield different results. The likelihood of extracting the exact same data from Twitter using the replication data and code supplied alongside this thesis is practically null, and thus so is the likelihood of reproducing identical results.

Moreover, the noted changed to the Twitter Developer API which took place during the undertaking of this research have severely restricted data access for researchers, further limiting the ability to reproduce these findings. To a large extent, this is an accepted limitation of digital research and where the dynamism and vastness of social media data is one of its many prevailing strengths, it is also one of its major drawbacks. Nonetheless, the data sets used in this thesis and the ideal points generated from them remain fixed and should be treated at the very least as a snapshot of the Twitter platform at the time of extraction. It is highly likely that the follower networks and tweet timelines of the UK MPs studied in this thesis are already markedly different from when they were originally harvested through Twitter's API. This should not, however, detract from what they represented at that specific time point and what we can extrapolate from that moving forward.

On that note, the difficulty in extracting new follower data from Twitter also limits the ability of researchers to update left/right ideal points of UK MPs for the new parliamentary session (2024-2029). During the completion of this research, the UK held a general election on July 4^{th} 2024 where more than half of the elected MPs (335) were new to the HoC. This now leaves the MP left/right ideal points generated as part of **Paper 1** somewhat dated, although the wider user ideal points validated as part of **Paper 2** will still be of use to researchers. The same can also be said for both Data Set 1 and Data Set 2 more generally, where the follower networks and tweet timelines are reflective of many MPs who are now no longer in parliament. However, these may still prove useful for researchers who might be interested in studies of historical MP online behaviour.

Another overall limitation of this research is in its single platform focus. While the many strengths of using Twitter for the purposes of the thesis are made clear in **Chapter 2**, there is a valid case to be made for a diversification of study of MP behaviour across multiple platforms. Although no other social media platform has the level of MP/elite usage that Twitter has, other sites such as Facebook and Instagram still have a sizeable number of MPs and media users. A more diverse evaluation of Britain's digital elite across multiple platforms would have provided a more well-rounded portrayal of Westminster's online presence and this would have been particularly beneficial to **Paper 3**. When exploring digital dyadic representation, an understanding of how this relationship sustains from platform to platform would have been insightful, and the results of all three studies ultimately succumb to the specific biases that Twitter's users and affordances generate. While it has been extensively argued that Twitter's unique biases made it an ample choice for this thesis, evaluating whether these results remain consistent across different platforms would have added credence to these findings. Thus, one recommendation for future research is to adopt a multi-platform approach to social media study where possible, though the growing restrictions to data access on many of these sites is severely limiting the options presently available to researchers.

Substantively, there a number of limitations that can be identified in the specific cases of **Paper 2** and **Paper 3**. Firstly, for **Paper 2**, the major limitation to this study is that it focuses solely on descriptive representation when evaluating ideological selection. As is noted in the paper, there are multiple different ways to assess media representation beyond the descriptive including not just visibility but also favourability and issue coverage. Paper 2 quantifies the ideological representation of the seven selected programmes as a simple visibility distribution: what are the individual ideal points of guests who *appear* on each show? It does not account in any way for how they are treated when they are on the show, nor the issues that they discuss. While this paper never claims to address the ideological representativeness of these programmes in such a manner, it does provide a slightly more simplistic view of political representation than perhaps would be necessary for a more definitive assessment. It is also constrained to the distribution of ideal points among active Twitter users only, which is disconnected from the ideological axis one might observe in the offline world. This leaves open a lot of potential to build on this research in future work, where the informativeness of social media networks for quantifying actor positions could be combined with techniques for quantifying other important aspects of media representation like programme transcripts and time afforded to each guest.

For **Paper 3**, one of the main limitations is arguably in the conception of digital dyadic representation in and of itself. **Paper 3** opts to assess digital dyadic representation as the within-party association between constituency positions and MP Twitter follower networks, positive engagement networks, and tweet topics. A case could be made that this is not the only way to conceive of constituency representation in this regard, where other important factors such as direct MP-to-constituency engagement on the platform itself should be considered, or how often MPs tweet about local issues or their constituency directly. Furthermore, this paper is fundamentally exploratory and does not offer any casual argument for the relationship between MP Twitter patterns and their constituency positions. While this paper does clarify from the outset that it is purely exploratory, a deeper understanding of exactly how MP online behaviour is influenced, if at all, by their individual constituencies is an important question and may warrant further investigation.

Finally, with the changes made to the Twitter API and the growing exodus of British users specifically from the site over the last year, there are lingering questions about how one might build on this research in the future. With the increasing importance of digital campaigning and online representation for politicians, the value of a concept like digital dyadic representation is growing. For over a decade and a half, Twitter became the natural home to the overwhelming majority of UK MPs making it the obvious platform of choice for studying such a concept. Now, with Twitter's future uncertain, there runs the risk of many politicians (and other important political elites) becoming digital nomads with their online presence fragmented across multiple platforms. Not only does this reaffirm the importance of adopting multi-platform strategies in future research, if we do not see a re-concentration of political elites to an alternative platform akin to what was observed on Twitter, these types of studies may be the last of their kind.

7.5 Concluding Remarks

This thesis has strived to provide a comprehensive quantitative analysis of the UK's online elite via data extracted from the platform formerly known as Twitter. During the time of conducting the analysis for this thesis, significant changes were made to the platform that has rendered it substantially different from what it was when this research began. Following its acquisition by Elon Musk in October 2022, the platform has been rebranded to the name X, paid subscription tiers have been introduced for both users and developers, and sizeable reductions in connect moderation strategies have been instituted. Owing to the paid subscription tiers introduced for developers who wish to extract data via the X Developer API, harvesting data from the platform at the scale used in this thesis would now be extremely expensive. Additionally, there has been a growing exodus of British users from the site over the last year which is increasing at an exponential rate. Ofcom's *Online Nation* report for 2024, which annually maps the online behaviour of the British public, has shown that, while X still remains the highest-reaching microblogging service in the UK, its average monthly adult reach continues to decline year on year – from 27.9 million in 2021 to 22.2 million as of August 2024 (Ofcom, 2024b, p.4).

At the same time, social media start-up Bluesky Social is now experiencing a simultaneous uptake in users, many of whom are members of the UK commentariat (Ittimani, 2024). Since the US election on November 5th, 2024, US and UK usage on the site has skyrocketed by almost 300 per cent to 3.5m users and the site is now proving a genuine challenger to other rival microblogging services such as Meta's *Threads* (Murphy and Burn-Murdoch, 2024). There is a hope that much of the Westminster digital bubble that has been traced and mapped across this thesis can reemerge on this new platform in the coming months and years, providing an entirely new frontier for study of online elite networks in the UK and elsewhere. If we can observe a reconcentration of elite users on this site to a degree comparable to Twitter then this can open up new avenues for political research and can potentially address some of the limitations noted in the previous section.

For now, the future of Twitter/X research remains uncertain, as does the future of online communication and campaigning by political elites in the UK. It may be the case that the era of concentrated elite activity onto a single site is over and elite networks become scattered across multiple platforms. Or it may be that an alternative challenger like Bluesky can take its place. Either way, as the findings of this thesis show, the applications of studying the online behaviour of elites for offline political and communicative research are endless.

Appendix A

Chapter 4 - Supplementary Material (Paper 1)

A.1 Sample Filter Robustness Checks

This appendix consists of a robustness check undertaken to assess the performance of the correspondence model using different sample subsets. As is explained in the data collection section of the main paper, the set of ordinary follower profiles is filtered to only include those who follow at least 10 MPs. This is to ensure that an especially informative subset of users is used to estimate ideal points of MPs, reducing noise and optimising model performance. The choice of setting the sample filter at a minimum of 10 was informed by Barberá's (2015) original paper but in order to test the robustness of the model estimates in relation to this threshold, the CA model was ran using subsets of the follower data at multiple different thresholds. The ideal points of MPs generated by the model at each of these thresholds was validated against the same expert survey estimates to assess how well the model performed at each threshold, and how much the estimates varied in relation to this filtering. Model validation was conducted using Pearson's *r* correlation coefficients of the model estimates against the expert validation estimates for both the overall model (between-party accuracy) and for Conservative and Labour MPs respectively (within-party accuracy). 20 iterations of the correspondence model were ran, increasing the sample filter threshold at each point starting at all users who followed at least 1 MP (all profiles) increasing by increments of 1 up to 5, then by increments of 5 up to 20, then increments of 10 up to 50, and finally increments of 50 up to 500. The details of the sample threshold for each iteration can be found in **Table 1**:

Using the entire sample of followers with a minimum threshold of MPs followed = 1, the model performs relatively poorly, particularly when discriminating between the ideal points of Conservative MPs. There is a sizeable improvement in both between

Iteration	Threshold	MPs	Users	Connections	Between r	CON r	LAB r
1	1	591	4,460,657	20,048,554	0.83	0.06	0.78
2	2	591	2,161,097	17,748,994	0.97	0.88	0.86
3	3	591	1,450,688	16,328,176	0.97	0.91	0.86
4	4	591	1,094,788	15,260,476	0.97	0.90	0.86
5	5	591	878,815	14,396,584	0.97	0.89	0.85
6	10	591	424,297	11,443,165	0.97	0.84	0.81
7	15	591	260,849	9,532,641	0.96	0.78	0.77
8	20	591	178,909	8,158,385	0.96	0.73	0.74
9	30	591	99,736	6,274,876	0.95	0.66	0.68
10	40	591	62,735	5,017,414	0.95	0.61	0.64
11	50	591	42,351	4,120,608	0.94	0.58	0.59
12	100	591	12,079	2,098,538	0.94	0.51	0.47
13	150	590	5,496	1,309,591	0.93	0.49	0.46
14	200	590	2,856	857,265	0.93	0.43	0.56
15	250	590	1,721	606,195	0.91	0.27	0.44
16	300	590	1,080	430,673	0.68	0.13	0.16
17	350	590	694	306,490	0.06	0.07	0.35
18	400	590	434	209,697	0.19	0.04	0.45
19	450	590	278	144,063	0.05	0.09	0.43
20	500	590	169	92,439	0.20	0.01	0.27

TABLE A.1: Details of each sample subset filtering at different MP following thresholds. Between r shows the Pearson's correlation coefficient of the overall model estimates against the expert validation estimates. CON and LAB r illustrate the within-party correlation coefficients.

and within-party accuracy when increasing the minimum threshold to 2, with between-party accuracy remaining reasonably stable until iteration 16 when it declines dramatically. After iteration 5, the within-party accuracy of the model begins to decline for both Conservative and Labour MPs. This trend is intuitive given that one would expect there to be an optimal "Goldilocks" point when subsetting the follower data to include only the most informative users. Setting the threshold too low will include users who are not especially politically informed and thus have less ability to discriminate ideologically between MPs, and setting the threshold too high will only include users who follow a significant proportion of MPs from across the spectrum, flattening out the ideological component altogether. In this case, the model estimates are most closely correlated with the expert estimates (both between and within) when setting the threshold at a minimum of 3 MPs, notably lower than Barberá's original threshold of 10. However, there are also efficiency considerations to take into account. At lower thresholds, the follower adjacency matrices are far larger and, as such, dimensionality reduction can be far more computationally intensive and time consuming. Thus, one could justifiably trade-off some degree of model accuracy in favour of a less time consuming and computationally demanding scaling procedure.

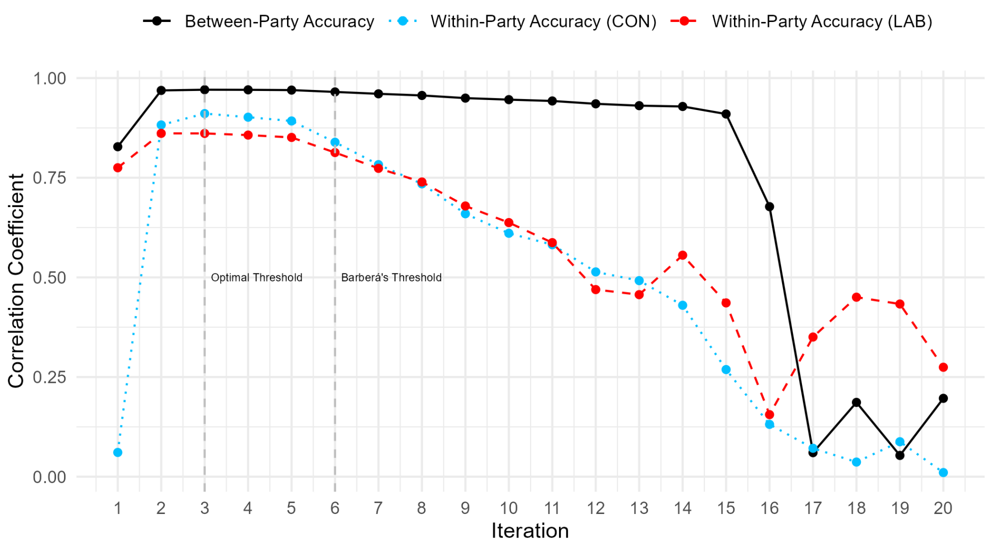


FIGURE A.1: Correlation coefficient represents the Pearson's *r* between the model estimates and the expert validation estimates at each iteration sample threshold. Optimal model performance is achieved in iteration 3 at a threshold of 3 and Barberá's threshold of 10 was set in iteration 6.

A.2 CA Model Dimensions

This appendix contains details of the multiple dimensions generated from the correspondence analysis conducted on the follower adjacency matrix.

A.2.1 Variance Capture By Dimension

Figure A.2 illustrates the individual variance capture of each of the first 10 dimensions as a Scree plot:

Dimension 1 captures the largest amount of the variance in the Twitter follower adjacency matrix, although this overall proportion is still low (3.6%). There is a steep drop in the amount of variance capture in dimension 2 (2.1%) which gradually reduces further for each additional dimension in the model. Although the first dimension has been shown in the paper to approximate an ideological component, the fact that it only accounts for 3.6% of the overall variance in the follower matrix indicates that much of the variance in the data cannot be solely explained by ideology. Nevertheless, using only the first dimension in the model still produces a highly accurate measure of MP ideal points.

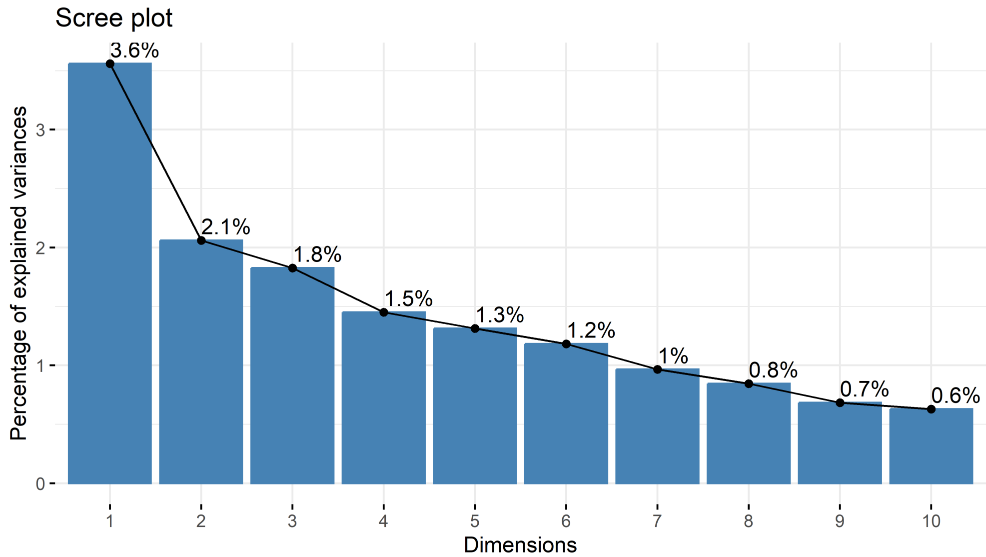


FIGURE A.2: Scree plot demonstrating the percentage of explained variance captured by each of the first 10 dimensions of the CA model.

A.2.2 Treating Nationalist MPs as Supplementary Columns

When conducting correspondence analysis on the follower adjacency matrix, it is explained in the paper that MPs representing national parties are initially excluded from the scaling process and treated as supplementary columns. Consequently, users who exclusively follow national MPs are also excluded from the initial scaling procedure and treated as supplementary rows. These supplementary columns and rows in the matrix and then retroactively projected onto the dimensional space after it has been scaled to ensure that the component being scaled is ideology. This is because the regional component to nationalist parties overwhelms their ideological one. **Figure A.3** illustrates a 2-dimensional scatterplot comparing the original model estimates when treating the nationalist MPs as supplementary columns against model estimates when including all MPs in the scaling procedure.

The estimates for non-nationalist MPs stay approximately equal in both models, but the estimates for nationalist MPs are skewed significantly. This is particularly the case for SNP and Alba Party MPs. There is also a notable skew for MPs of non-nationalist parties who represent constituencies based outside of England. This suggests that including nationalist parties in the scaling process and not treating them as supplementary columns overwhelms the ideological component captured by the first dimension of the model, where nationalist MPs and MPs representing non-English constituencies will instead cluster around a nationalist element. In order to remove this effect, it is subsequently justifiable to treat nationalist MPs as supplementary columns in the initial scaling process and retroactively project them onto the dimensional space.

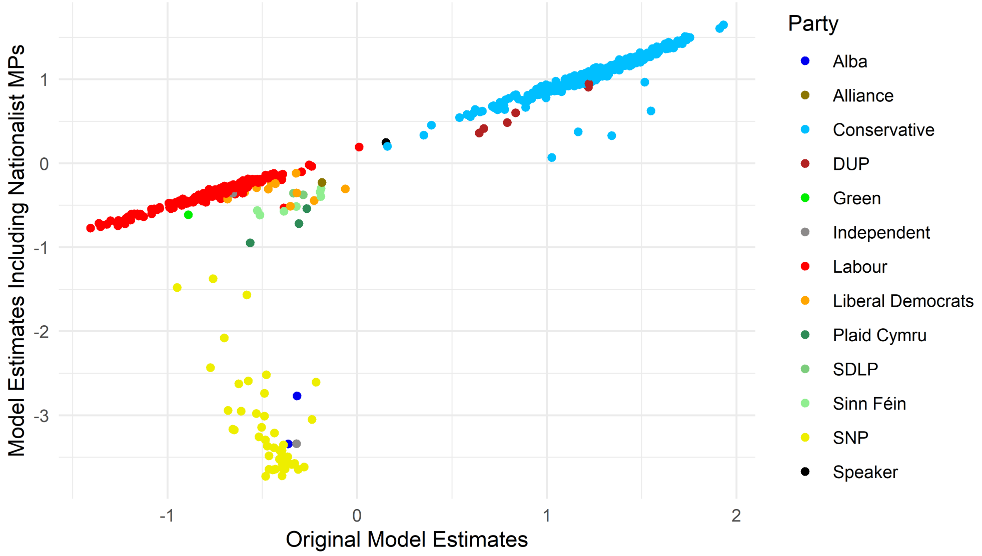


FIGURE A.3: Original CA Model Estimates vs. Model Estimates Including National MPs During Scaling

A.2.3 Interpreting Additional Dimensions

It has been confirmed in the main paper that the first dimension of the CA follower model approximates left/right ideology. It is uncertain what components the additional model dimensions beyond the first capture, pending deeper investigation. **Figure A.4** illustrates the first two dimensions of the CA follower model plotted on a 2-dimensional scatterplot:

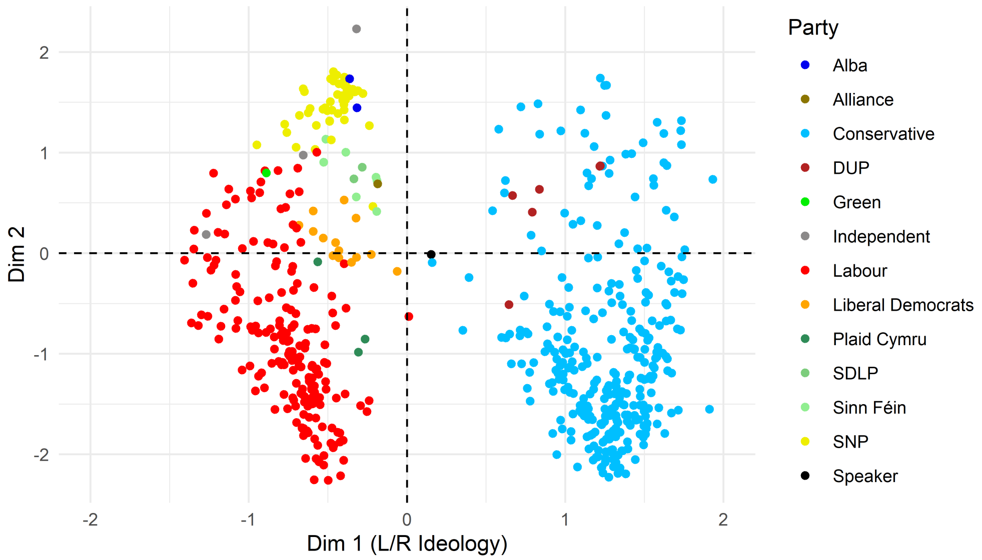


FIGURE A.4: CA Model Dimension 1 Estimates vs. CA Model Dimension 2 Estimates

It is clear that the second model dimension cuts across partisan and ideological boundaries, although nationalist party MPs do still appear to cluster relatively closely together. Initial investigation into the second and third dimensions have so far proven difficult to understand, and further analysis will be needed to ascertain what these components capture. For the purpose of this paper, interpretation of these additional dimensions is not essential given that the first dimension provides an adequate approximation of left/right ideology. It is important to note that these additional dimensions do not necessarily represent additional dimensions of ideology *per se*; rather, they capture whatever additional dimensions drive the follower networks of MPs. An intuitive assumption might be a geographical component to MP follower networks, where ordinary users are more likely to follow MPs who represent constituencies closer to where they live, or a seniority component which divides MPs between more senior ministers and party backbenchers. One possibility for identifying what these additional dimensions represent may be to conduct cluster analysis of MPs using their first three dimensional coordinates and inspect the underlying characteristics that MPs within the same clusters share. This way, a more comprehensive understanding of the primary drivers of MP follower networks can be established, incorporating additional dimensions beyond the initial left/right ideological component. This analysis would be complex and beyond the requirements of this paper, but can encourage future work that can build on this initial research.

A.3 Expert Survey Format & Responses

A short survey was circulated to 133 experts in British politics and political science on the 08/11/2022. The survey was generated through Qualtrics and made accessible to participants via a link attached to an email that was distributed to each of them individually. All responses were anonymous. The experts chosen to participate in the survey were all academics with substantive research expertise in electoral, legislative, or parliamentary politics and/or public opinion with a specific focus on the UK. In order to source these academics and their contact information, the politics faculty websites of all UK universities who are either members of the Russell Group and/or ranked in the top 30 UK universities for politics in any of the Times Higher Education World University Rankings, the Guardian University Guide, or the Complete University Guide were used. Although this sampling method is slightly restrictive, it helped to narrow down the process as efficiently as possible and provide a straightforward way of sourcing an appropriate number of individuals who could be justifiably considered credible sources of expert opinion on British politics. Of the 133 experts contacted and invited to participate, 70 participants completed the survey in full indicating a 53% response rate.

The survey asked participants to place a sub-sample of 30 UK MPs, 12 UK political parties, and 13 UK media organisations on an 11-point scale between 0-10 based on where they believed each MP/party/organisation sat on the left/right ideological spectrum. Following the same format and questioning as the British Election Study (BES), 0 represented the furthest left and 10 represented the furthest right on the spectrum. All responses were optional and a Don't Know option was also provided. A starting question also asked participants to rank their level of knowledge of British parliamentary politics on a 6-point likert scale. This question was purely exploratory and was not used to filter the responses in any way. Of the 30 MPs, 13 were members of the Conservative Party, 13 were members of the Labour Party, 2 were members of the Liberal Democrats, and 1 was a member of the Green Party. The final MP chosen was an Independent but was formerly a member of the Labour Party. There was a deliberate attempt to balance the sample of MPs as much as possible across the left/right spectrum (based on personal knowledge) as well as between more well-known MPs such as party leaders and cabinet ministers and lesser-known backbenchers.
The complete survey format distributed to participants as follows:
Q1: What would best describe your level of knowledge of British parliamentary politics?
- None
- Know a limited amount
- Know a moderate amount
- Know a fair amount

- Know a great deal
- Expert

Q2: In politics people sometimes talk of left and right. Where would you place each of these MPs on the following scale?

0 - Left
1
2
3
4
5
6
7
8
9
10 - Right
Don’t Know

Q3: Where would you place each of these political parties on the following scale?

0 - Left
1
2
3
4
5
6
7
8
9
10 - Right
Don’t Know

Q4: Where would you place each of these media organisations on the following scale?

0 - Left
1
2
3
4
5
6
7
8
9

10 - Right

Don't Know

Details of the 30 MPs chosen for the survey along with the results of the expert survey can be found in **Table 2** overleaf:

Name	Party	Min.	Max.	Mean	Std.D	Variance	Count
Anneliese Dodds	Labour	3	6	4.25	0.73	0.54	63
Bell Ribeiro-Addy	Labour	1	5	2.62	1.03	1.06	34
Ben Wallace	Conservative	7	10	8.38	0.70	0.49	63
Boris Johnson	Conservative	6	10	8.38	0.82	0.68	68
Caroline Lucas	Green	1	9	2.69	1.19	1.42	68
Chris Bryant	Labour	2	6	4.37	0.87	0.77	60
Desmond Swayne	Conservative	8	11	9.76	0.85	0.72	45
Diane Abbott	Labour	1	7	2.25	1.05	1.10	68
Ed Miliband	Labour	2	7	3.90	0.90	0.81	67
Edward Davey	Lib Dem	4	7	5.78	0.75	0.57	65
Hilary Benn	Labour	3	7	4.41	0.92	0.85	66
Ian Lavery	Labour	1	5	2.35	0.99	0.98	40
Jacob Rees-Mogg	Conservative	9	11	10.04	0.71	0.51	69
Jeremy Corbyn	Independent	1	4	1.67	0.76	0.57	69
John McDonnell	Labour	1	3	1.84	0.72	0.53	67
John Redwood	Conservative	7	11	10.11	0.72	0.51	63
Keir Starmer	Labour	3	8	4.81	0.89	0.79	69
Kemi Badenoch	Conservative	8	11	9.58	0.79	0.63	67
Lisa Nandy	Labour	3	7	4.17	0.81	0.66	65
Matthew Hancock	Conservative	7	10	8.06	0.81	0.65	67
Priti Patel	Conservative	8	11	9.85	0.73	0.54	68
Rishi Sunak	Conservative	7	11	8.58	0.89	0.79	69
Steven Baker	Conservative	4	11	9.63	1.03	1.06	63
Stuart Andrew	Conservative	4	11	8.00	1.65	2.73	11
Suella Braverman	Conservative	9	11	10.22	0.68	0.47	68
Theresa May	Conservative	7	10	8.07	0.71	0.51	68
Thomas Tugendhat	Conservative	7	10	7.82	0.81	0.66	62
Tim Farron	Lib Dem	4	8	5.79	1.02	1.03	67
Yvette Cooper	Labour	3	7	4.62	0.79	0.63	66
Zarah Sultana	Labour	1	5	2.07	0.98	0.95	45

TABLE A.2: Summary statistics of the expert survey ideology estimates of 30 MPs. Although the choice ranges from 0 - 10, values in the table reflect their raw value on the 11-point scale. Thus, a chosen value of 0 (furthest left) is point 1 on the scale and so on.

Additionally, details of the 12 political parties and 13 media organisations can also be found in **Table 3**:

Name	Group	Min.	Max.	Mean	Std.D	Variance	Count
Conservative	Party	7	11	8.71	0.71	0.50	68
Labour	Party	2	7	4.50	0.78	0.60	68
SNP	Party	2	8	4.29	0.97	0.94	68
Liberal Democrats	Party	4	7	5.74	0.74	0.55	68
Green	Party	1	6	2.96	0.97	0.94	67
Reform	Party	5	11	10.05	0.94	0.89	55
Plaid Cymru	Party	2	8	3.95	1.10	1.20	64
Alba	Party	2	11	5.24	2.38	5.65	38
SDLP	Party	1	6	4.31	0.92	0.84	48
DUP	Party	6	11	9.67	1.05	1.09	55
Sinn Fein	Party	1	5	2.96	0.86	0.74	51
Alliance	Party	3	6	5.38	0.70	0.48	48
BBC	Media	4	9	5.98	0.80	0.64	64
ITV	Media	4	8	6.17	0.73	0.54	60
Sky	Media	4	9	6.59	0.88	0.77	61
The Daily Express	Media	7	11	9.74	0.88	0.78	65
The Daily Mail	Media	8	11	9.91	0.72	0.51	65
The Daily Mirror	Media	1	9	3.88	1.16	1.34	65
The Daily Telegraph	Media	7	11	9.53	0.84	0.70	66
The Financial Times	Media	4	10	6.71	1.19	1.41	63
The Guardian	Media	2	6	3.86	0.80	0.63	66
The Huffington Post	Media	2	8	4.59	1.13	1.27	41
The Independent	Media	3	7	4.96	1.08	1.16	57
The Sun	Media	6	11	9.02	0.91	0.83	64
The Times	Media	6	11	8.02	1.07	1.14	64

TABLE A.3: Summary statistics of the expert survey ideology estimates of 12 political parties and 13 media organisations. Although the choice ranges from 0 - 10, values in the table reflect their raw value on the 11-point scale. Thus, a chosen value of 0 (furthest left) is point 1 on the scale and so on.

A.4 MP Control Variables By Candidate Endorsed - Summary Statistics

	Sunak		Truss	
	N	**%**	**N**	**%**
Vote Share	130	-	150	-
Social Variables				
Gender				
Male	106	82%	106	71%
Female	24	18%	44	29%
Ethnic Minority				
No	124	95%	141	94%
Yes	6	5%	9	6%
Year of Birth (Median)	52.1	-	51.5	-
School Type				
Private	56	49%	51	41%
Grammar	19	17%	19	15%
State	39	34%	53	43%
University Type				
Oxbridge	40	33%	28	21%
Russell Group	42	35%	46	35%
Non-Russell Group	33	27%	41	31%
None	6	5%	17	13%
Political Variables				
Cohort				
Pre-1997	4	3%	6	4%
1997 - 2010	29	22%	21	14%
2010 - 2015	37	28%	37	25%
2015 - 2019	28	22%	40	27%
2019	32	25%	46	31%
Ministerial Position				
Current	29	22%	46	31%
Former	41	32%	33	22%
Never	60	46%	71	4&%
Majority (%)(Median)	27.1	-	26.7	-

TABLE A.4: % is calculated as a percentage of complete observations for each variable by candidate.

A.5 Ordinary User Ideal Points

This appendix consists of details about the ideal points generated by the CA follower model for the row data (ordinary users). When applying correspondence analysis to the MP follower adjacency matrix, the CA model will scale both the columns *and* the rows meaning that ideal points are also generated for the ordinary users who follow the MPs. **Figure A.5** illustrates the ideal point distribution of these users (who will follow 10 or more MPs):

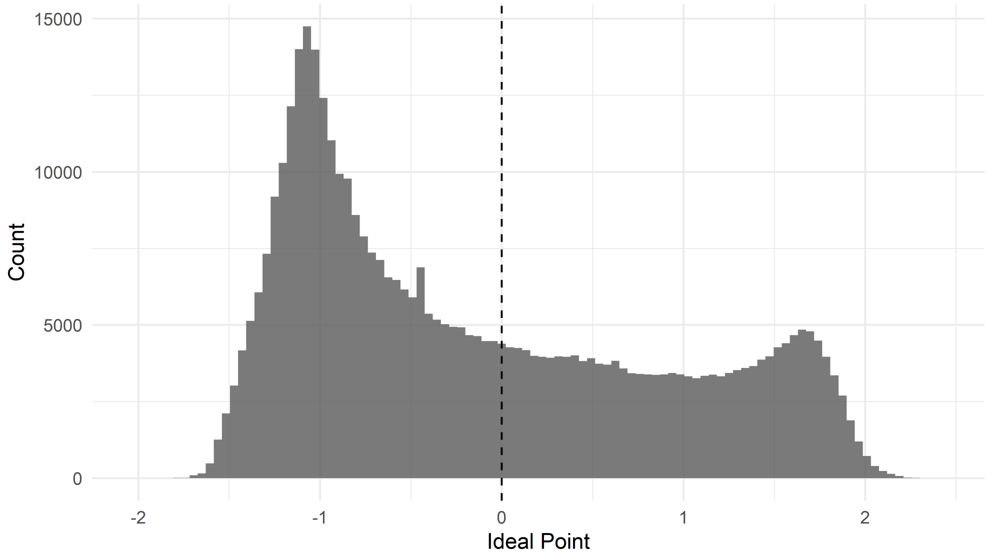


FIGURE A.5: Histogram distribution of ordinary user ideal points (who follow at least 10 MPs). Bins = 100.

Along the distribution, it is interesting to note an imbalance between users to the left compared to the right. Given that the liberal skew of Twitter is well-documented, this is perhaps an intuitive finding and likely also explains the imbalance between the number of followers of MPs from left-wing parties compared to right within the dataset. Although these users represent an especially informative subset of profiles (follows at least 10 MPs), they can still be regarded as ordinary users. To test the general ideological distribution of wider elite users on the platform (e.g: journalists, media organisations, political commentators etc.), this set of followers is subsetted to only include 'elite' accounts. This is done by filtering the follower profiles to only include accounts which are either verified (before the creation of *Twitter Blue*, where users can now pay a monthly subscription to receive verified status, verification was an official signal of authenticity) and/or have at least 30,000 followers (This is in accordance with ruling by the Advertising Standards Authority (ASA) - https://www.asa.org.uk/rulings/sanofi-uk-A19-557609.html). This included 11,525 users and **Figure A.6** illustrates their ideal point distribution:

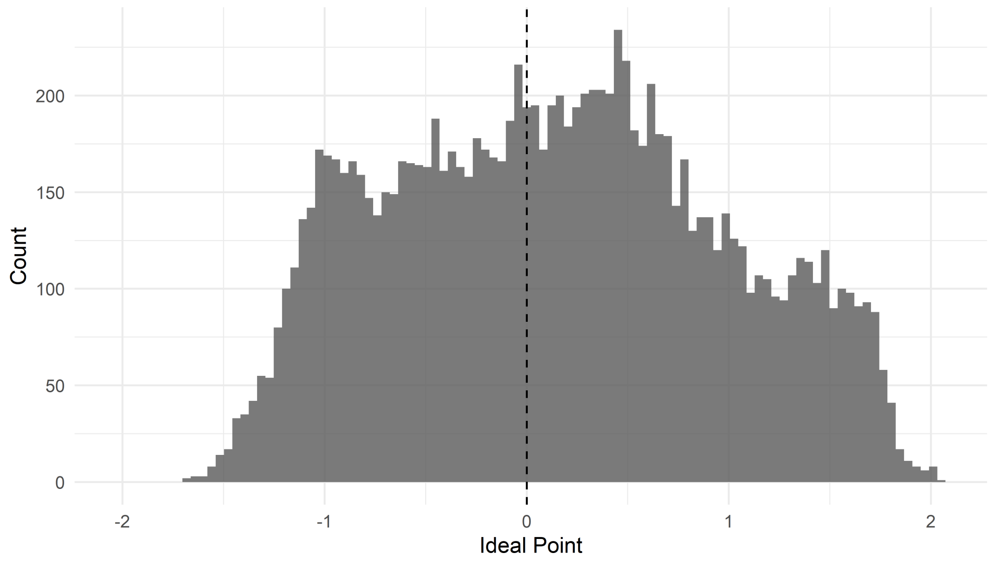


FIGURE A.6: Histogram distribution of wider elite user ideal points (verified and/or have a min. 30,000 followers). Bins = 100

This subset of elite users demonstrate a more balanced distribution of ideal points across the spectrum, with a slight remaining over-representation of left-wing elite users on the site.

Appendix B

Chapter 5 - Supplementary Material (Paper 2)

B.1 Twitter User Data

B.1.1 Summary Statistics of Twitter User Profile Metadata

Users			Median Count					
Account Type	*n*	Verified (%)	MPs Followed	Followers	Following	Tweets	Listed	Ideal Point
Overall	424,297	2%	17	347	1,163	2,490	2	-0.29
Ordinary	412,772	0%	17	330	1,140	2,379	2	-0.31
Eite	11,525	86%	24	16,892	2,425	12,191	183	0.27
Guest	511	81%	88	33,869	1,834	10,941	444	0.71

TABLE B.1: Summary statistics of the Twitter profile metadata of users from each sub-set.

B.2 YouGov Survey Data - General Public Estimates

B.2.1 Summary Statistics - Political Parties

Name	Type	*n*	Don't Know (%)	Mean	Median	Mode	S.D	Min	Max	Username	Ideal Point
Conservative Party	Political Party	1,652	20%	7.42	8	10	2.42	0	10	Conservatives	1.87
Green Party	Political Party	1,440	30%	2.91	3	3	2.14	0	10	TheGreenParty	-0.47
Labour Party	Political Party	1,649	20%	3.37	3	3	2.44	0	10	UKLabour	-0.87
Liberal Democrats	Political Party	1,525	26%	4.23	5	5	2.04	0	10	LibDems	-0.35
Reform UK	Political Party	1,300	37%	7.73	8	10	2.58	0	10	reformparty_uk	1.30*
Scottish National Party	Political Party	1,228	41%	3.67	4	3	2.47	0	10	theSNP	0.41

TABLE B.2: Summary statistics of general public estimates for ideological position of UK political parties. Representative sample of UK adults (18+). n = 2,068.

B.2.2 Summary Statistics - Media Outlets

Name	Type	*n*	Don't Know (%)	Mean	Median	Mode	S.D	Min	Max	Username	Ideal Point
BBC	Media Outlet	1,501	27%	4.91	5	5	2.57	0	10	BBCNews	0.27*
Byline Times	Media Outlet	242	88%	4.04	4	5	2.38	0	10	BylineTimes	-0.92
Channel 4	Media Outlet	1,263	39%	4.15	4	5	1.97	0	10	Channel4News	0.51
GB News	Media Outlet	1,173	43%	7.8	8	10	2.36	0	10	GBNEWS	1.10*
ITV	Media Outlet	1,261	39%	4.95	5	5	1.82	0	10	itvnews	0.24
LBC	Media Outlet	735	64%	4.95	5	5	2.23	0	10	LBC	0.63
Novara Media	Media Mutlet	227	89%	3.37	3	0	2.9	0	10	novaramedia	-1.32
Sky News	Media Outlet	1,197	42%	5.55	5	5	2.1	0	10	SkyNews	0.49*
TalkTV	Media Outlet	635	69%	6.94	7	10	2.46	0	10	TalkTV	1.13*
The Daily Express	Media Outlet	1,114	46%	7.04	7	10	2.4	0	10	Daily_Express	1.47
The Daily Mail	Media Outlet	1,356	34%	7.63	8	10	2.36	0	10	MailOnline	0.33*
The Daily Mirror	Media Outlet	1,216	41%	4.39	4	3	2.77	0	10	DailyMirror	-0.70
The Daily Star	Media Outlet	959	54%	5.13	5	5	2.87	0	10	dailystar	0.44*
The Daily Telegraph	Media Outlet	1,189	42%	7.19	7	10	2.33	0	10	Telegraph	0.55*
The Economist	Media Outlet	887	57%	5.89	6	5	2.08	0	10	TheEconomist	0.07*
The Financial Times	Media Outlet	1,139	45%	6.49	6	5	2.08	0	10	FinancialTimes	0.33*
The Guardian	Media Outlet	1,299	37%	3.95	3	3	2.69	0	10	guardian	-0.26*
The Huffington Post	Media Outlet	613	70%	4.86	5	5	2.34	0	10	HuffPost	-0.31*
The i	Media Outlet	754	73%	4.67	5	5	1.87	0	10	theipaper	-0.29
The Independent	Media Outlet	1,130	45%	4.96	5	5	2.06	0	10	Independent	-0.20*
The New Statesman	Media Outlet	618	70%	5.13	5	5	2.7	0	10	NewStatesman	-0.57
The Spectator	Media Outlet	718	65%	6.21	7	5	2.56	0	10	spectator	1.15
The Sun	Media Outlet	1,316	36%	6.62	7	8	2.76	0	10	TheSun	0.49*
The Times	Media Outlet	1,240	40%	6.72	7	7	2.15	0	10	thetimes	0.28*

TABLE B.3: Summary statistics of general public estimates for ideological position of UK media outlets. Representative sample of UK adults (18+). n = 2,068.

B.2.3 Summary Statistics - Media Personalities

Name	Type	*n*	Don't Know (%)	Mean	Median	Mode	S.D	Min	Max	Username	Ideal Point
Alan Sugar	Media Personality	1,214	41%	6.81	7	7	2.24	0	10	Lord_Sugar	0.35
Alastair Campbell	Media Personality	1,018	51%	4.13	4	4	2.58	0	10	campbellclaret	0.10
Andrew Marr	Media Personality	933	55%	5.1	5	5	2.12	0	10	AndrewMarr9	0.32
Bill Bailey	Media Personality	832	60%	3.75	4	4	1.99	0	10	BillBailey	-0.67
Carol Vorderman	Media Personality	1,050	49%	3.67	3	3	2.51	0	10	carolvorders	0.29
Darren Grimes	Media Personality	313	85%	7.2	8	10	2.59	0	10	darrengrimes_	1.98
David Baddiel	Media Personality	841	59%	3.69	4	4	2.2	0	10	Baddiel	-0.53
Deborah Meaden	Media Personality	751	64%	5.68	6	5	2.22	0	10	DeborahMeaden	-0.10
Denise Welch	Media Personality	542	74%	4.68	5	5	2.52	0	10	RealDeniseWelch	0.93
Frankie Boyle	Media Personality	898	57%	3.31	3	3	2.52	0	10	frankieboyle	-1.15
Gary Lineker	Media Personality	1,222	41%	3.34	3	3	2.49	0	10	GaryLineker	-0.21
Gary Neville	Media Personality	742	64%	3.91	4	4	2.35	0	10	GNev2	-0.07
Isabel Oakeshott	Media Personality	358	83%	7.24	7.5	10	2.48	0	10	IsabelOakeshott	1.10
James O'Brien	Media Personality	495	76%	3.87	4	5	2.55	0	10	mrjamesob	-0.31
Jamie Oliver	Media Personality	956	54%	4.75	5	5	2.24	0	10	jamieoliver	0.75
Jeremy Vine	Media Personality	931	55%	4.97	5	5	2.48	0	10	theJeremyVine	0.57
Jon Snow	Media Personality	821	60%	4.68	5	5	2.12	0	10	jonsnowC4	0.06
Laura Kuenssberg	Media Personality	931	55%	5.38	5	5	2.63	0	10	bbclaurak	0.87
Laurence Fox	Media Personality	1,014	51%	7.94	9	10	2.64	0	10	LozzaFox	1.54
Martin Lewis	Media Personality	1,153	44%	4.56	5	5	1.82	0	10	MartinSLewis	0.38
Nick Robinson	Media Personality	691	67%	5.4	5	5	2.24	0	10	bbcnickrobinson	0.76
Nigel Farage	Media Personality	1,503	27%	8.2	9	10	2.5	0	10	Nigel_Farage	2.09
Owen Jones	Media Personality	603	71%	2.54	2	0	2.66	0	10	OwenJones84	-0.87
Piers Morgan	Media Personality	1,374	34%	7.3	8	10	2.4	0	10	piersmorgan	0.61
Prof. Brian Cox	Media Personality	840	59%	4.34	4	5	2.06	0	10	ProfBrianCox	-0.33
Rachel Riley	Media Personality	702	66%	5.26	5	5	2.44	0	10	RachelRileyRR	0.16
Richard Osman	Media Personality	758	63%	4.79	5	5	2.05	0	10	richardosman	-0.21
Robert Peston	Media Personality	774	63%	4.86	5	5	2.27	0	10	Peston	0.47
Stephen Fry	Media Personality	1,136	45%	4.33	4	3	2.36	0	10	stephenfry	-0.52
Susanna Reid	Media Personality	698	66%	5.05	5	5	2.09	0	10	susannareid100	0.26

TABLE B.4: Summary statistics of general public estimates for ideological position of UK media personalities. Representative sample of UK adults (18+). n = 2,068.

B.3 Guest Category Details

B.3.1 Overall Categories and Subcategories Summary

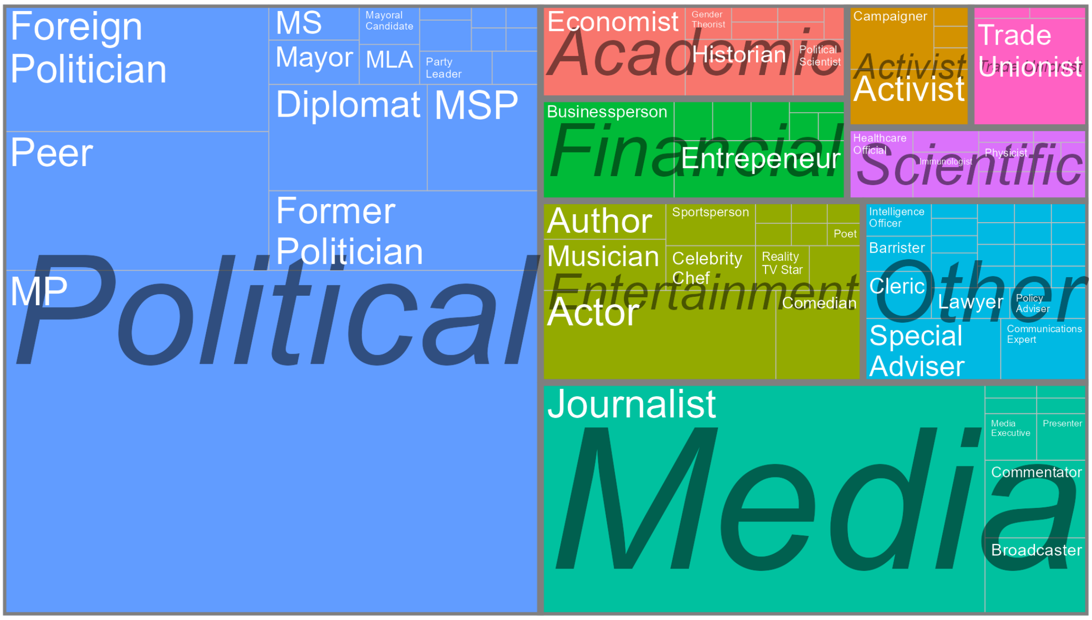


FIGURE B.1: Treemap which illutsrates the breakdown of all 825 unique guests by category type. Tiles are sized by proportion of the data.

B.3.2 Ideal Point Distribution by Category

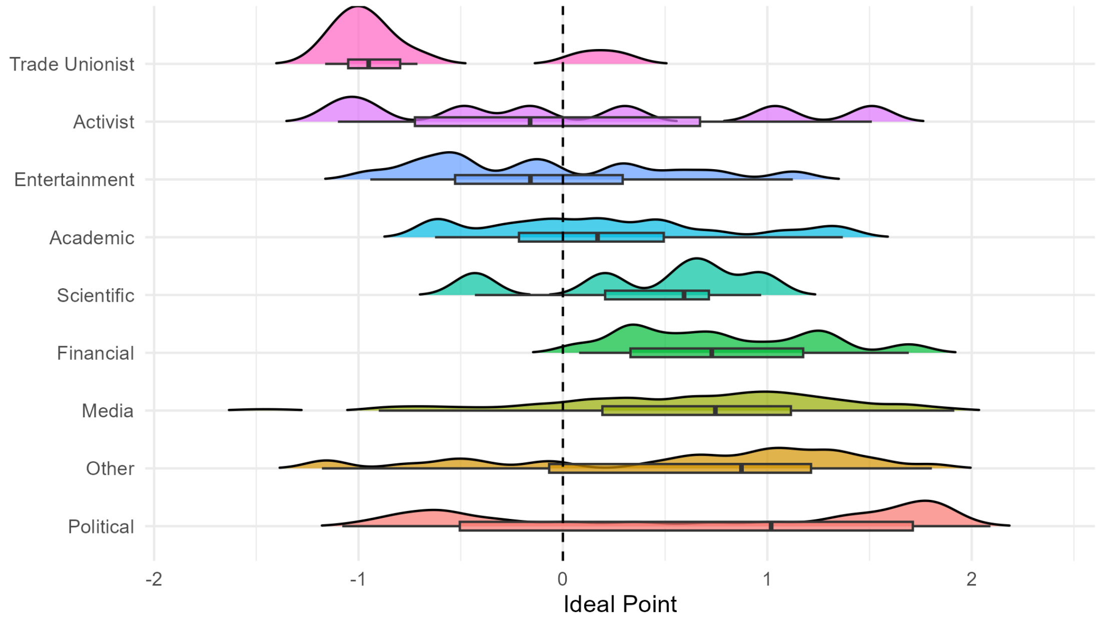


FIGURE B.2: Ideal point distributions of guests grouped by their primary category type.

B.4 Missing Data

B.4.1 Missing Data By Show

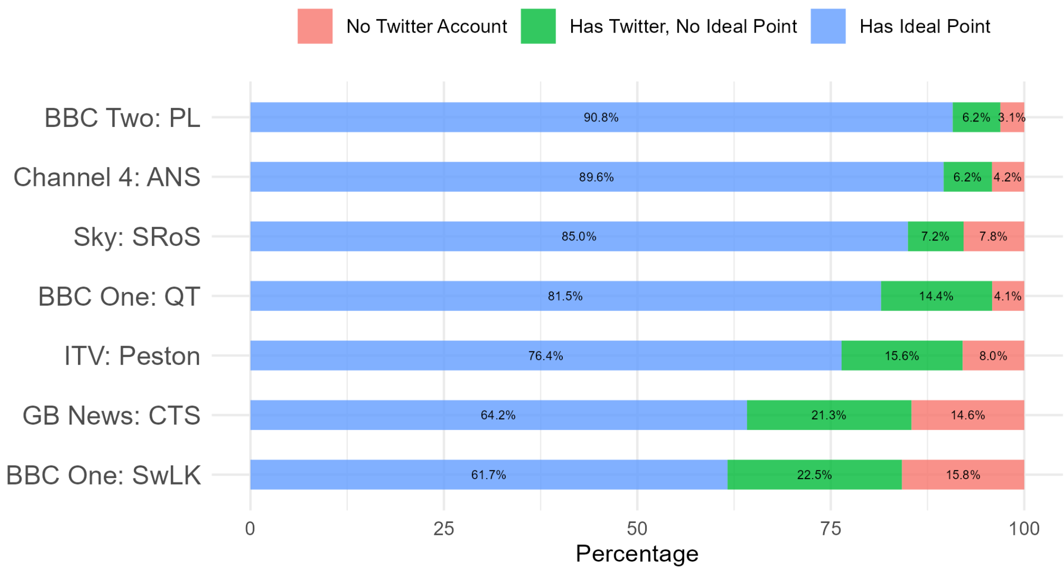


FIGURE B.3: Breakdown of proportion of guests missing ideal points by each show.

B.4.2 Missing Data by Guest Category and Nationality

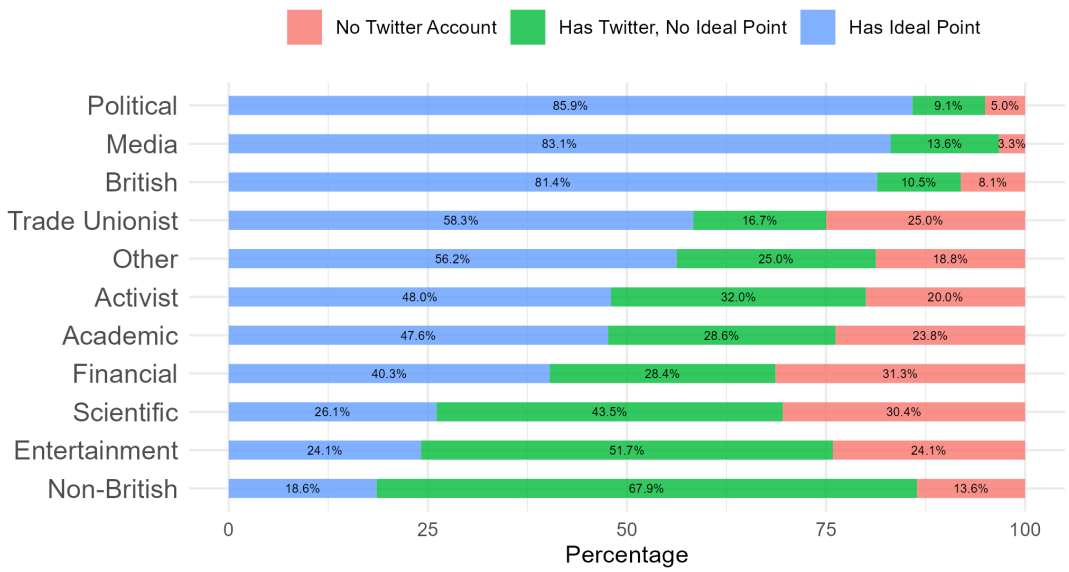


FIGURE B.4: Breakdown of proportion of guests missing ideal points by each categroy type and nationality.

B.4.3 Breakdown of Category and Organisation Leaning of Missing Data by Show

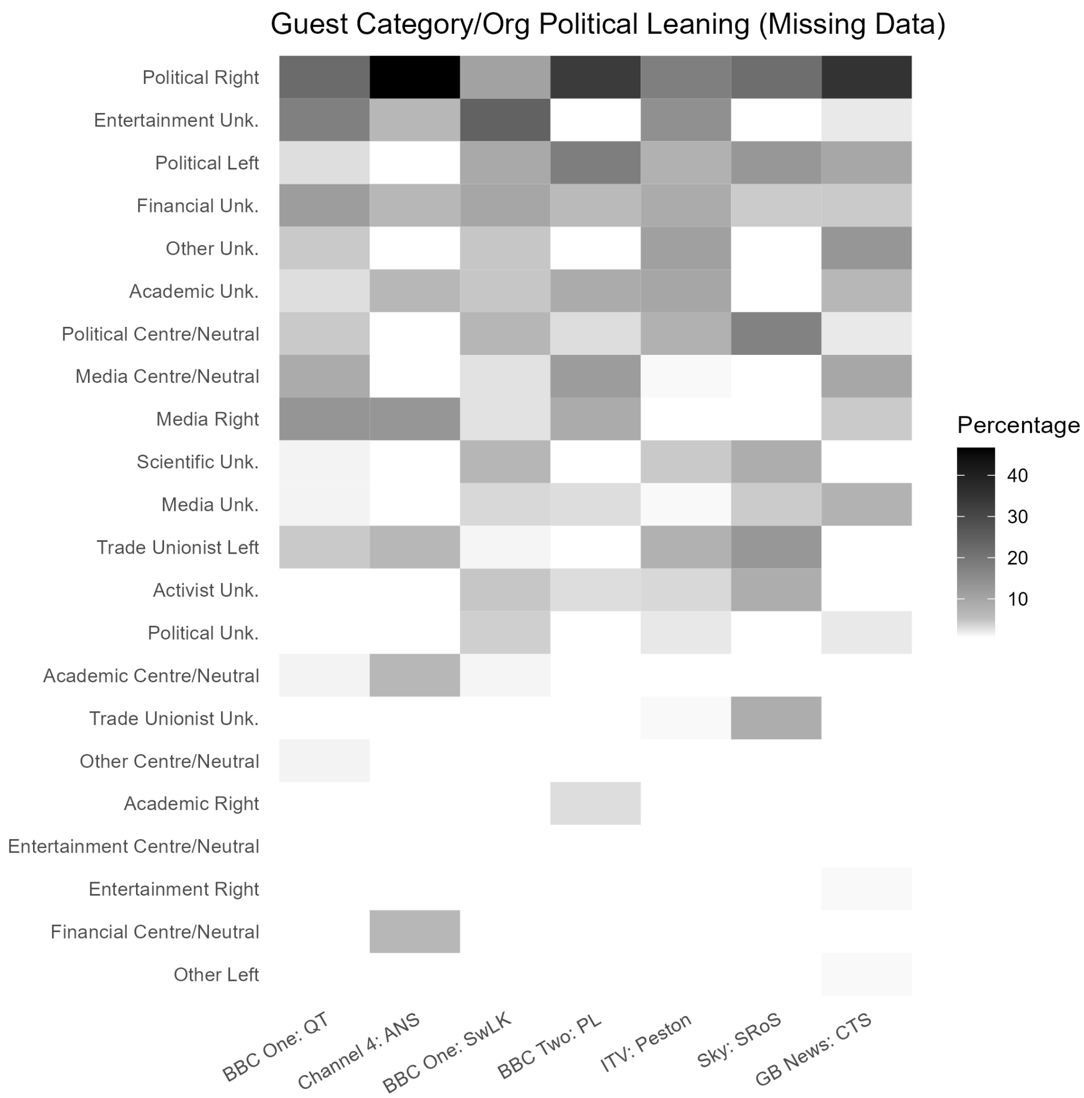


FIGURE B.5: Proportion of guests on each show who did not have an ideal point from each category type and political leaning. This demonstrates that for each show except BBC One: SwLK, the majority of guests who were missing ideal points were from the Political Right, followed by Entertainers without political leaning.

B.4.4 Correlation Matrix - User Ideal Points and Their Proxies

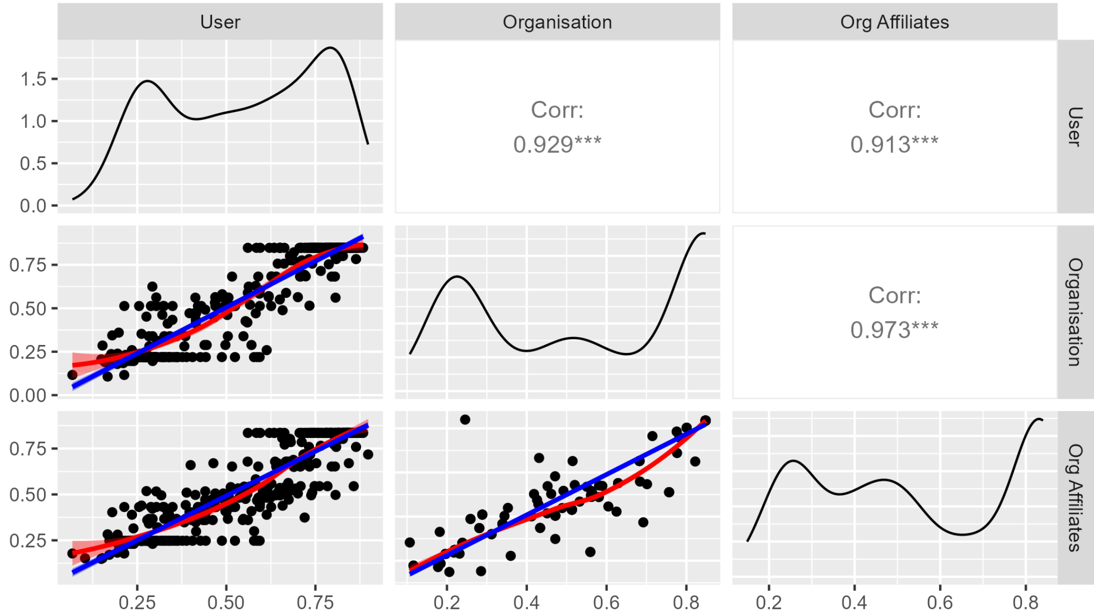


FIGURE B.6: Correlations between a guest's individual ideal point, the ideal point of their organisation, and the mean ideal point of organisation affiliates all demonstrate extremely high correlation statistically signifcant at the 0.05 level. [$*** \leq 0.001, ** \leq 0.01, * \leq 0.05$]

B.4.5 Results with Missing Data Imputed With Proxies

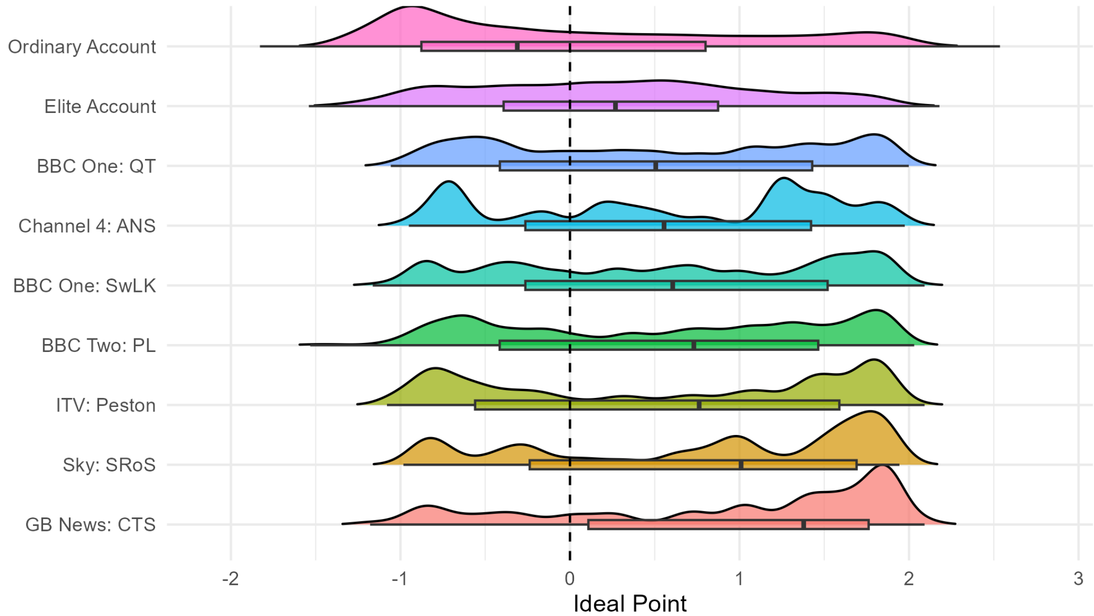


FIGURE B.7: Ideal point distributions of guests who appeared on each show, imputing missing ideal points for some guests by using their organisation's ideal point or mean ideal point of affiliated accounts. Results indicate the same general pattern as results without imputation.

Programme	Complete n	Median	Mean	S.D	Min.	Max	Dip Stat
Ordinary Account	410,640	-0.31	0	1	-1.83	2.54	0.0071***
Elite Account	11,504	0.27	0.27	0.83	-1.54	2.18	0.0031
BBC One: QT	332	0.51	0.53	0.95	-1.05	2	0.054***
BBC One: SwLK	284	0.61	0.58	0.95	-1.16	2.09	0.04**
BBC Two: PL	351	0.73	0.57	0.97	-1.53	2.03	0.049***
ITV: Peston	321	0.76	0.57	1.06	-1.08	2.09	0.077***
Channel 4: ANS	139	0.56	0.55	0.93	-0.95	1.97	0.077***
Sky News: SRoS	144	1	0.81	0.97	-0.98	1.94	0.055**
GB News: CTS	208	1.38	0.94	0.97	-1.18	2.09	0.03

TABLE B.5: Summary statistics of the ideal point distributions for ordinary and elite Twitter users in the dataset, along with subsets of users who appeared as guests on each of the 7 shows. Median ideal points include significance asterisks to indicate statistical significance of pairwise Mann-Whitney U tests compared with the Elite Account distribution. Dip statistics also include significance asterisks to indicate statistical significance of multimodality in each distribution. In both cases, statistical significance is measured at the 0.05 threshold. [$*** \leq 0.001, ** \leq 0.01, * \leq 0.05$]

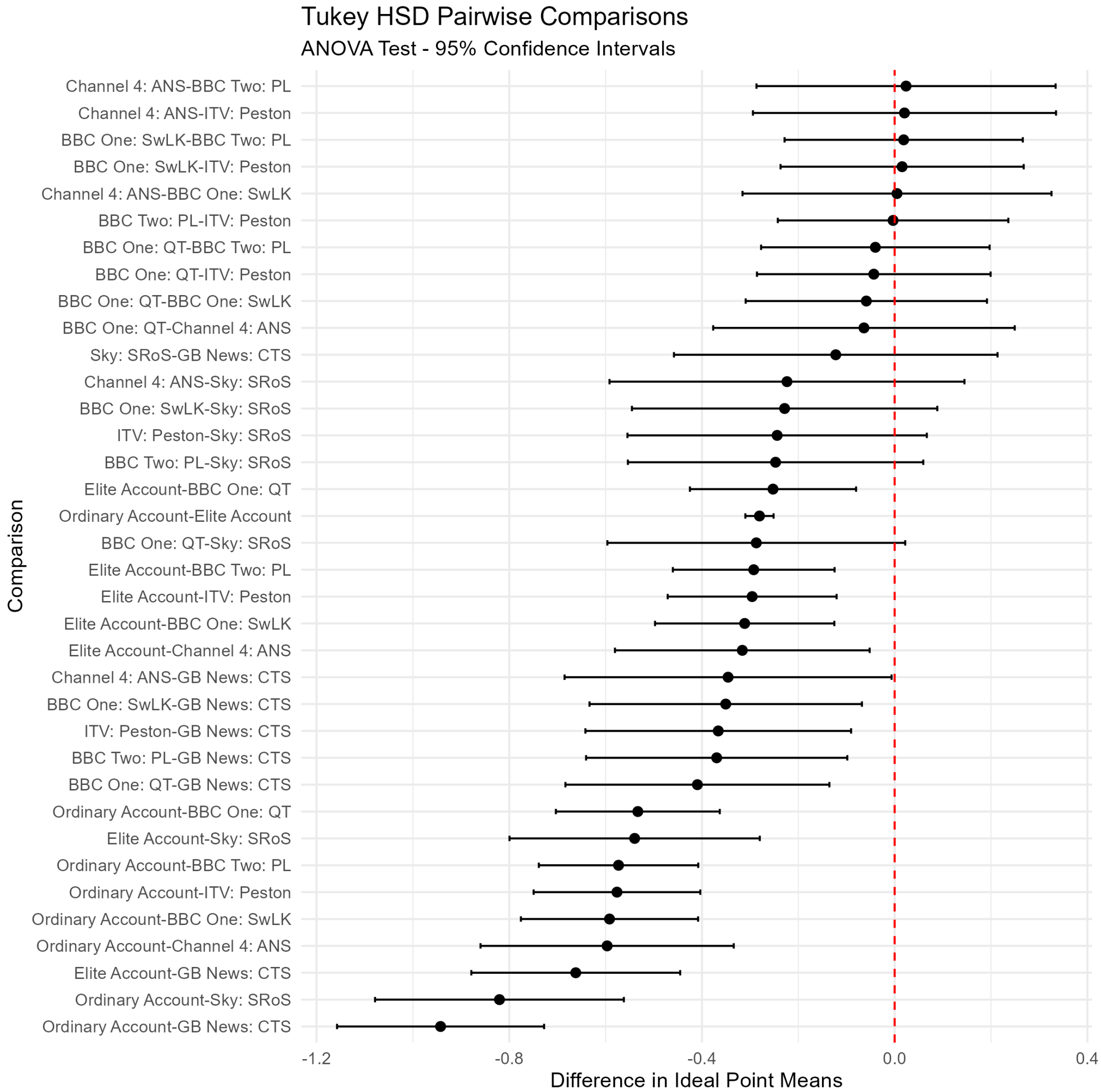


FIGURE B.8: Pairwise comparisons of mean user ideal points across the seven programmes using Tukey's HSD test. Each line represents the estimated difference in mean ideal points between two programmes, with 95% confidence intervals. Comparisons are ranked by the size of the estimated difference. Intervals that do not cross zero indicate statistically significant differences in guest ideological positioning between programmes. With imputed ideal points for missing data.

Appendix C

Chapter 6 - Supplementary Material (Paper 3)

C.1 BESIP Respondent Ideal Point Estimation

The British Election Study is the longest running social science survey in the UK and is central to the study of public opinion and electoral research in the country. General public respondent data used in this study is drawn from five individual waves of the British Election Study Internet Panel (BESIP) - (Fieldhouse et al., 2023).

These online panel studies follow the same survey respondents over time to study within-person change and can also be treated as individual cross-sectional surveys of the population at a single point in time. The five waves used in this study, their sample sizes, wave-to-wave retention rates, and data collection dates can be found in **Table C.1**:

Wave	Data Collection Period	*n*	Retention %
20	June 2020 - June 2020	31,468	55.3%
21	May 2021 - May 2021	30,281	52.5%
22	November 2021 - December 2021	28,113	63.5%
23	May 2022 - May 2022	30,949	72%
24	**December 2022 - December 2022**	**15,439**	41.7%
25	May 2023 - May 2023	30,407	74%

TABLE C.1: BESIP Wave-to-Wave details. Retention rate is the % of respondents retained from the previous wave. Wave 24 (in bold) was not used in the study as it was a supplemental wave with a reduced sample size and variables.

Across the five waves used in the study, 7,100 respondents appeared in every one. For left/right ideal point estimates along both an economic and a social dimension, new variables were created. These were generated using a multidimensional scaling technique known as a multiple correspondence analysis (MCA). It is an extension of simple correspondence analysis and works much in the same way as other forms of dimensionality reduction such as principal components analysis or factor analysis but can work with categorical data. It is an appropriate technique for scaling the relationship between ordinal responses in a two-way contingency table. It is commonly applied to generate spatial estimates of relative ’distances’ between columns and rows in a set of scaled latent dimensions, and has been used in a number of fields ranging from the social sciences to medicine, ecology and marketing (Barberá et al., 2015; Bendixen, 1996; Gaughan, 2024; Greenacre and Blasius, 2006; Palmer, 1993; Sourial et al., 2010).Typically, a model like ordinal item-response theory (IRT) might be used to generate estimates of left/right ideal points such as in the original study by Hanretty and colleagues (Hanretty et al., 2017). However, research has shown that a method like correspondence analysis can produce estimates that closely approximate traditional spatial models at a much reduced computational cost (Barberá et al., 2015; Bonica, 2014).

To generate estimates of left/right ideal points using this method, two sets of value-based statements are used from which respondents can give one of five responses [1 - Strongly agree, 2 - Agree, 3 - Neither agree not disagree, 4 - Disagree, 5 - Strongly disagree] (excl. Don’t know). For economic ideal points, the five left/right value-based statements are as follows:

1) ”Government should redistribute income from the better off to those who are less well off”

2) ”Big business takes advantage of ordinary people”

3) ”Ordinary working people do not get their fair share of the nation’s wealth”

4) ”There is one law for the rich and one for the poor”

5) ”Management will always try to get the better of employees if it gets the chance”

For social ideal points, these are based on five authoritarian-liberal value-based statements and are as follows:

1) ”Young people today don’t have enough respect for traditional British values”

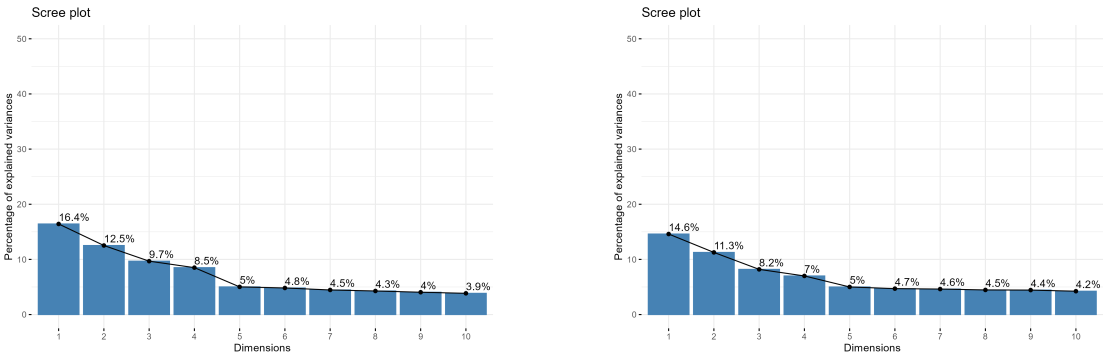
2) ”For some crimes, the death penalty is the most appropriate sentence”

3) ”Schools should teach children to obey authority”

4) ”Censorship of films and magazines is necessary to uphold moral standards”

5) ”People who break the law should be given stiffer sentences”

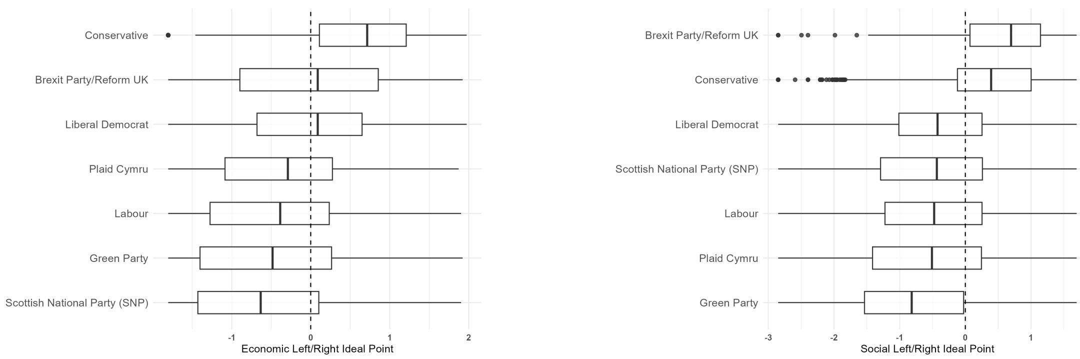
As the ideal point estimates produced using this procedure are scaled relative to one another, the MCA models for each dimension are fit to all responses across all five waves together to allow for comparison over time (n = 35,500). The MCA models are fit in R using the *MCA* function from the *FactoMineR* library. Don't know responses are recoded to NA. Scree plots for both MCA models which show the variance capture in each dimension are shown in **Figure C.2**.



(A) Economic Left/Right Dimensions (B) Social Left/Right Dimensions

FIGURE C.1: Scree plot showing the % of variance explained in each latent dimension scaled by the MCA model, for economic (left) and social (right) values.

The first dimension captures the most amount of variance in both models and look to approximate left/right ideal points. This can be inferred from exploratory analysis of the variance in ideal points between voters from different parties. After re-scaling the first dimensions to have a mean of 0 and SD of 1, **Figure ??** illustrates boxplots of respondent economic and social ideal points across all five waves, grouped by vote intention.



(A) Economic Left/Right Ideal Points (B) Social Left/Right Ideal Points

FIGURE C.2: Boxplots of economic (left) and social (right) ideal points of respondents split by general election voting intention.

Both dimensions align generally with what one would expect of ideal points of different party supporters in the UK. Economically, Conservative supporters are furthest to the right followed by Brexit/Reform UK and the Liberal Democrats. SNP supporters are furthest to the left, followed by the Green Party and then Labour. Socially, Brexit/Reform UK supporters are the most conservative, followed by the

Conservatives, while the Greens are the most socially liberal. The Liberal Democrats, the SNP, and Labour all occupy a relatively similar space along this dimension.

C.2 Constituency Position Estimation

As outlined in the main paper, this study makes use of a contemporary technique for small-area estimation known as multilevel regression with poststratification (MRP). The goal of this method is to reduce the amount of noise around local area opinion when disaggregating large national survey data such as the BESIP down to 632 individual constituencies. Fifteen seperate models were estimated in total, five for each wave of the BESIP data split across three positions to be estimated: 1) economic left/right ideal point; 2) social left/right ideal point; 3) most important issue. Predictions for constituency position in each wave are based on Bayesian multilevel regression models which leverage data on both individuals and constituencies, subsequently post-stratified using census data about population compositions within each constituency. Crucially, these models are built on three key data sources: 1) individual respondent data from the BESIP; 2) auxiliary constituency data from the BES 2019 General Election results files (Fieldhouse et al., 2019); and 3) a pre-built poststratification frame based on 2011 Census data (Hanretty et al., 2018). Thorough outlines of the technical mechanics of this method can be found in Hanretty et al. (2018) and Hanretty (2020).

C.2.1 Constituency Left/Right Ideal Points

For estimating left/right ideal points of each UK constituency in each wave, the general modelling strategy follows thus:

The political position y in this case takes the form of a continuous variable, where our goal is to estimate the average score of y for citizens in each constituency $j \in \{1, ..., 632\}$. Taking our national survey sample of 7,100 respondents per wave, filtering out those with missing observations leaving a realised sample of 3,357, for each respondent i in this data we have measures of their constituency location (GSS Code), their estimated political position y_i, and a set of categorical demographic variables $\{1, ..., K\}$. Additionally, we have a matrix of auxiliary data X_j containing a combination of demographic, geographic and political variables for each constituency j. Based on this information, we can model political position y_i as follows:

$$y_i = \beta_0 + \sum_j \beta_j X_{ij} + \sum_k \gamma_k Z_{ik} + u_i + \epsilon_i$$

where β_0 is the grand intercept, β_j are the coefficients for the fixed effects, X_{ij} are the fixed effect covariates (constituency-level), γ_k are the coefficients for the random effects, Z_{ik} are the random effect covariates (respondent-level), u_i are the random intercepts, and ϵ_i is the residual error term. Predictor variables are shown in **Table C.2**:

Individual Vars.	Values	Constituency Vars.	Values
Sex	1 = "Male", 2 = "Female	% Female	Cont. %
Age Group	1 = "16-19", 2 = "20-24", 3 = "25-29", 4 = "30-44", 5 = "45-59", 6 = "60-64", 7 = "65-74", 8 = "75+"	% Aged 16-19, % Aged 65+	Cont. %
Education Level	1 = "No Quals.", 2 = "Level 1", 3 = "Level 2", 4 = "Level 3", 5 = "Level 4/5", 6 = "Other"	% Level 4 Quals., % FT Students	Cont. %
Social Grade	1 = "DE", 2 = "C2", 3 = "C1", 4 = "AB"	% Deprived	Cont. %
Housing Status	1 = "Owns", 2 = "Rents"	% Home Owners	Cont. %
GSS Code	Geographic Area Codes: 632 unique categories	-	-
		% Household Married	Cont. %
		% Non-white	Cont. %
		% Christian, % Non-Christian, % No Religion	Cont. %
		% Unemployed, % Self-Employed, % Economically Inactive	Cont. %
		Log Population Density	Cont.
		Region	1 = "East Midlands", 2 = "East of England", 3 = "London", 4 = "North East", 5 = "North West", 6 = "Scotland", 7 = "South East", 8 = "South West", 9 = "Wales", 10 = "West Midlands", 11 = "Yorkshire & Humber"
		% Conservative Vote, % Labour Vote, % Lib Dem Vote, % SNP Vote, % Plaid Cymru Vote, % Green Party Vote, % Brexit Party Vote	Cont. %
		% E.U Leave Vote	Cont. %

TABLE C.2: Individual and Constituency-Level Predictors.

As stated in the main paper, each model is fitted in R using the *brm* function from the *brms* library (Bürkner, 2017). Due to the computational demand of MCMC, this study made use of the IRIDIS High Performance Computing Facility (HPC) at the University of Southampton for the running of these models.[1]. The code template used is posted below:

```
#### ---- FIT BRM ---- ####

options(scipen = 999)

#### ---- LOAD LIBRARIES ---- ####

library(dplyr)
library(data.table)
library(readxl)
library(parallel)
library(brms)
library(cmdstanr)

#### ---- COMMAND LINE ARGUMENTS ---- ####

# Set command line arguments
args <- commandArgs(trailingOnly = TRUE)

# Stop the script if no command line argument given
if(length(args)==0){
  stop("Requires command line argument!")
}

# Wave number: Options = (1,2,3,4,5)
# Dependent variable: Options = (econ_ideal_point,
                                 social_ideal_point,
                                 mii_grouped)

#### ---- LOAD DATA ---- ####

# BES data subset with estimated ideal points
bes_data <- fread("bes_data_with_ideal_points_formatted.csv")
```

[1]https://www.southampton.ac.uk/isolutions/staff/iridis.page

```
# Constituency data
const_data <- fread("bes_constituency_data_formatted.csv")

#### ---- FIT BRM MODEL ---- ####

# Subset data by wave
bes_data_subset <- bes_data %>% filter(wave == args[1])

# Merge individual-level and constituency-level datasets
data_merged <- bes_data_subset %>%

  # Select the necessary individual level predictors and the
  specified dependent variable
  select(GSSCode,sex,age0,housing,hrsocgrd,education,args[2]) %>%

  # Merge with constituency data
  left_join(const_data, by = "GSSCode") %>%

  # Select final columns for model
  select(GSSCode,sex,age0,housing,hrsocgrd,education,args[2],
         c11Female,c11Age18to24,c11Age65plus,c11QualLevel4,
         c11HouseholdMarried,
         c11NonWhite,c11Christian,c11NonChristianReligion,
         c11NoReligion,
         c11Unemployed,c11SelfEmployed,c11EconomicInactive,
         c11FulltimeStudent,
         c11Deprived,c11HouseOwned,c11PopulationDensityLog,Region,Lab19,
         Con19,LD19,SNP19,PC19,Green19,Brexit19,leaveHanretty)

# Specify number of model chains
num_chains <- 4

# Specify number of cores available
total_cores <- parallel::detectCores()

# Specify formula
formula_string <- paste(
  args[2], "~ (1|sex) + (1|age0) + (1|housing) +
  (1|hrsocgrd) + (1|education) + (1|GSSCode) +",
  "c11Female + c11Age18to24 + c11Age65plus +
  c11QualLevel4 + c11HouseholdMarried +",
```

```
  "c11NonWhite + c11Christian +
  c11NonChristianReligion + c11NoReligion + c11Unemployed +",
  "c11SelfEmployed + c11EconomicInactive + c11FulltimeStudent +
  c11Deprived +", "c11HouseOwned + c11PopulationDensityLog +
  Region + Lab19 + Con19 + LD19 +",
  "SNP19 + PC19 + Green19 + Brexit19 + leaveHanretty"
)

formula <- as.formula(formula_string)

# Specify priors and response family

if (args[2] == "mii_grouped") {

  family <- categorical(link = "logit", refcat = "Health")

  priors <- default_prior(formula, data = data_merged, family = family)

  priors$prior <- "normal(0,1)"

} else {

  family <- gaussian()

  priors <- set_prior("normal(0, 1)", class = "b")

}

# Fit BRM
bayes_model <- brm(formula = formula,
                   data = data_merged,
                   family = family,
                   chains = num_chains,
                   cores = total_cores,
                   prior = priors,
                   iter = 5000,
                   warmup = 2500,
                   backend = "cmdstanr",
                   seed = 12345,
                   refresh = 10,
                   control = list(adapt_delta = 0.95,
```

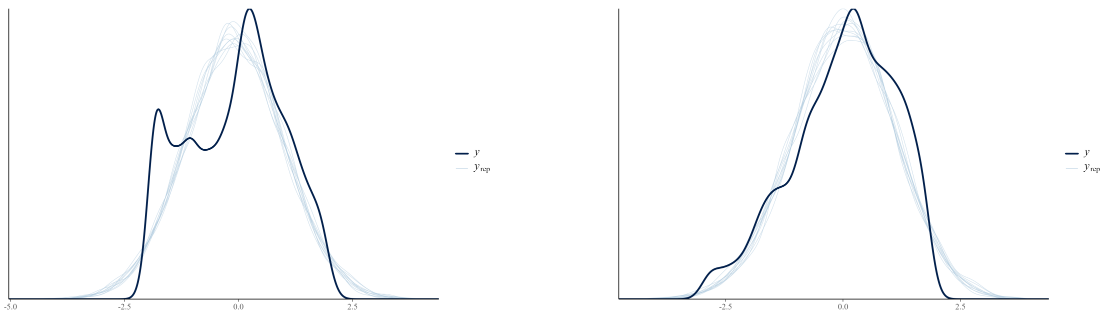
```
                                max_treedepth = 12))
```

In this code block, args[1] takes a numerical command line argument when running the script between 1 and 5 reflecting the wave of BESIP data to model, and args[2] specifies the chosen political position of interest (econ/social ideal point or MII). Convergence statistics for model parameters across the four chains are reported in **Table C.3**.

	Economic Ideal Point			Social Ideal Point		
	Rhat Max.	**ESS Bulk Min.**	**ESS Tail Min.**	**Rhat Max.**	**ESS Bulk Min.**	**ESS Tail Min.**
Wave 1	1.001	4097.592	5609.739	1.001	3333.775	5420.193
Wave 2	1.001	4381.225	6127.271	1.001	3543.787	5678.947
Wave 3	1.001	4217.332	5661.389	1.002	4264.785	6099.31
Wave 4	1.001	4424.407	6019.557	1.003	3339.439	5108.265
Wave 5	1.002	3709.431	5586.376	1.001	3427.45	5426.418

TABLE C.3: Convergence statistics for the ideal point BRM models.

Chain convergence and efficiency diagnostics would indicate that all five model waves for both ideal point dimensions have converged effectively, with $\hat{R}$ values for all parameters below 1.1 and with suitably high effective sample size (ESS) bulk and tail values ($>$ 1,000). Posterior prediction plots for Wave 5 ideal point models suggest good model fits, particularly the social model. These plots are relatively the same across all waves between the two dimensions. These can be seen in **Figure C.3**



(A) Economic Ideal Point Model

(B) Social Ideal Point Model

FIGURE C.3: Posterior predictions for Wave 5 models compared to 10 random draws.

C.2.2 Constituency Most Important Issue

Respondent most important issues are also derived from the BESIP data and the same question is asked to respondents in every wave of the survey: *"As far as you're concerned, what is the SINGLE MOST important issue facing the country at the present time?"* Any response provided is then manually code into a range of 50 codes, before being further collapsed into 12 parent categories and then 4 general categories. These are hand-coded by members of the BES team and a full description of their coding criteria can be found on pages 12-22 of the 2014-2023 BESIP Combined Waves 1-25

Internet Panel Codebook https://www.britishelectionstudy.com/data-object/british-election-study-combined-wave-1-21-internet-panel/. Broadly, these issue categories are shown in **Table C.4**:

Parent Issue	Sub-Issue
Europe	Europe
Immigration	Immigration
	Asylum
Economy	Economy-general
	Economy-personal
	Unemployment
	Taxation
	Debt/deficit
	Inflation
	Living costs
Health	Health
	Coronavirus
	COVID-economy
Inequality	Poverty
	Inequality
	Housing
Environment	Environment
Austerity/spending	Education
	Welfare
	Austerity
	Social care
	Pensions/ageing
	Transport/infrastructure
Negativity	Pol-neg
	Partisan-neg
	Societal divides
Other lib-auth	Morals
	Nat ident, goals loss
	Racism/discrimination
	Crime
	Foreign affairs
	War
	Defence
	Pol values-auth
	Pol values-liberal
	Gender/sexuality/family
	Terrorism
Other	Pol values-right
	Pol values-left
	Constitutional
	International trade
	Devolution
	Scot-ind
	Foreign emergency
	Domestic emergency
	Election outcome
	Other
	Referendum unspecified
NA	Uncoded

TABLE C.4: Parent Issues and their corresponding sub-issues.

Estimating constituency MII is principally the same as with estimating constituency left/right ideal points, but fitting a categorical logit model with a nominal outcome variable instead of continuous. In this case, we model political position y_i as:

$$y_i \sim \text{Multinomial}(\pi_{ij})$$

where y_i can take on multiple nominal categories and the probabilities π_{ij} of each issue category can be modeled as:

$$\log\left(\frac{\pi_{ij}}{\pi_{iK}}\right) = \beta_0 + \sum_j \beta_j X_{ij} + \sum_k \gamma_k Z_{ik} + u_j$$

where π_{ij} represents the probability that respondent i belongs to issue category j over the probability that respondent i belongs to the reference issue category K (health). All other model parameters remain the same. Convergence statistics for model parameters across the four chains are reported in **Table C.5**.

	Most Important Issue		
	Rhat Max.	**ESS Bulk Min.**	**ESS Tail Min.**
Wave 1	1.003	2320.428	2087.395
Wave 2	1.002	3399.059	2850.434
Wave 3	1.002	3779.703	3932.518
Wave 4	1.007	451.906	437.425
Wave 5	1.007	599.726	964.335

TABLE C.5: Convergence statistics for the most important issue BRM models.

These statistics would again suggest good model convergence (with $\hat{R}$ for all parameters below 1.1) and effective sample sizes are suitably large in waves 1-3 ($>$ 1,000) and slightly lower in waves 4 and 5. A posterior plot for the MII model in Wave 5 is illustrated in **Figure C.4** and would suggest a very good model fit .

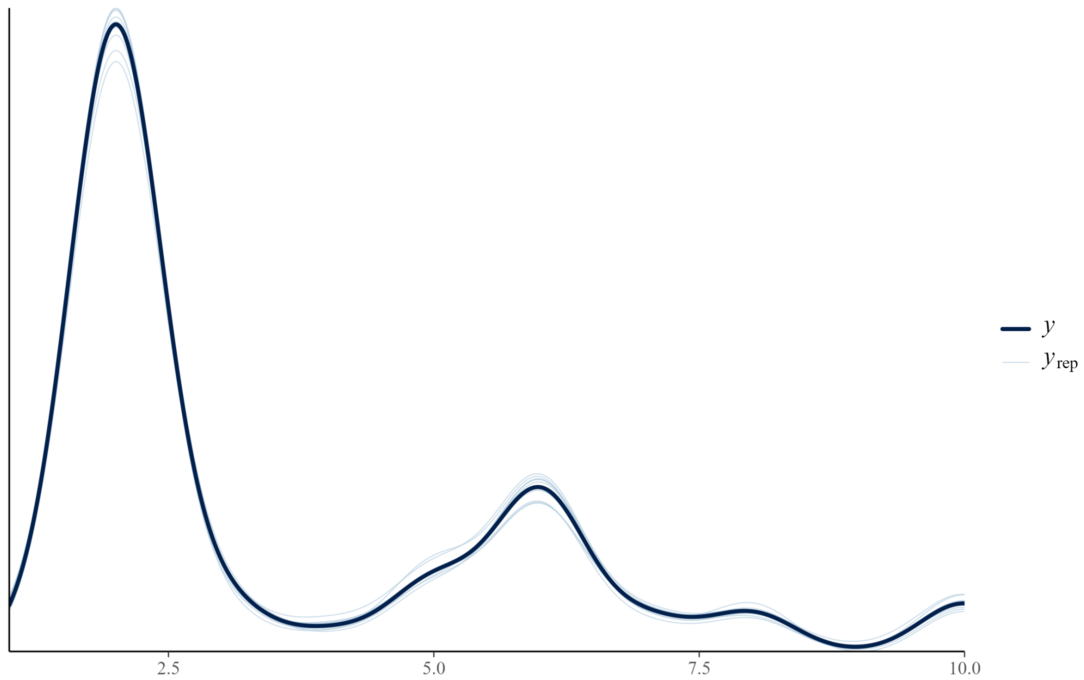


FIGURE C.4: Posterior predictions for Wave 5 MII model compared to 10 random draws.

C.3 MP Tweet Topics

Aligning as closely as possible with the BESIP's issue coding schema for respondent MIIs, MP tweets were coded into the same sub-issues using a manually curated key term dictionary. Of the 50 sub-issues, 43 were used to categorise MP tweets. Pol-neg and partisan-neg and foreign emergency and domestic emergency were both folded into single categories respectively, while referendum unspecified and other were not used as they are too vague. Election outcome (anything that refers to a recent or upcoming election) was deliberately not used as this would likely be too dominate in the tweet data and would risk over-inflating the Other parent category. One category in the BESIP data (constitution) was previously folded into the constitutional category. The list of keywords used to code tweets by topic are shown in **Table C.6**

Parent Topic	Sub-Topic	Keywords
Europe	Europe	brexit,european union,european commission, european parliament,european referendum, eu,withdrawal agreement,remainer, leaver,single market,eurozone, transition period,free trade agreement, customs union,northern ireland protocol, northern irish protocol,ni protocol, windsor framework,take back control, brussels,mep,meps
Immigration	Immigration	migra,border,hostile environment
	Asylum	small boat,refugee,asylum, people smuggl,channel crossing, crossing the channel,cross the channel, legal route,rwanda plan,rwanda policy, rwanda proposal,rwanda scheme
Economy	Economy-general	econom,recession,financial stability, gdp,supply chain,business, bank of england,budget,debt,bills
	Unemployment	employ,jobs,worker,staff shortage, strik,trade union,zero hour, wage,wages,income,pay rise, pay increase,pay decrease,pay freeze, pay and condition,fire and rehire, job rentention,picket line

	Taxation	tax,national insurance,licence fee
	Debt/deficit	deficit
	Inflation	inflation,interest rate
	Living costs	living cost,cost of living,costofliving, prices,living standard, household support fund,rising costs
Health	Health	health,nhs,doctor,medic,nurs,gp,gps, hospital,pharma,patients,dentist, surgeon,disease,treatments,
	Coronavirus	covid,virus,pandemic,infect,vaccin, lockdown,social distanc,self isolat test and trace,track and trace,stay home
	Covid-economy	furlough,job retention
Inequality	Poverty	poverty,foodbank,food bank,breadline
	Inequality	inequalit,wealthy,wealthie,poore,riche, super rich,mega rich,fair society, fairer society millionaire,billionaire, level up,levelling up
	Housing	housing,home own,rent,renters,rented, house price,homeless,rough sleep, sleeping rough,vagran,mortgage,evict, tenan,rental

Environment	Environment	environment,climate,global warming,net zero, pollut,fly tip,flytip,sewage,carbon,co2,emission, green agenda,green recovery,green policy,green deal, greendeal,green new deal,greennewdeal,greener, greenhouse gas,green belt,greenbelt,green job, green invest,sustainability,renewable,fracking, fossil fuel,cop1,cop2,ipcc,ecological
Austerity/Spending	Education	education,teacher,teaching,school, universit,student,gcse,a level, btec,pupil,ofsted,tuition, tutor,graduate,college,
	Welfare	welfare state,universal credit,uc benefit claim,free school meal, job seekers allowance,winter fuel, legacy benefit,allowance
	Austerity	austerity,cuts,cutting
	Social Care	social care,socialcare,carers, care home,child care,childcare
	Pensions/Ageing	pension,waspi,ageing population, triple lock

	Transport/Infrastructure	transport,trains,hs2,bus,buses,rail, roads,pothole,pot hole,airport,airline, motor,cyclist,cycling,tram,trams,tramline, vehicle,cars,aviation,infrastructure, energy generation,energy provision
Negativity	Pol/Partisan-neg	corrupt,sleaze,lying,liar,unelected, untrustworthy,dishonest,incompeten, useless,scandal,debacle,chaos,indecis, dysfunction,bias,biased,misinformation, disinformation,fake news,fakenews, defund the bbc,defundthebbc,partygate
	Societal Divides	social div,societal div,regional div
Other Lib-Auth	Morals	moral,ethic,selfish, self interest,greed,greedy
	Nat Ident, Goals-loss	national identity
	Racism/Discrimination	racis,discriminat,sexis,misogyn,ableis, ageis,classis,bigot,prejudice,hatred, hate crime,hate speech,xenophob, islamophob,transphob,homophob, biphob,antisemit,black lives matter, blacklivesmatter,blm,windrush,intoleran, holocaust,george floyd,sectarianis

Crime	crime,criminal,anti social,antisocial, police,policing,drugs,cannabis,gangs, violence,rape,sexual assault,physical abuse, domestic abuse,child abuse,sexual abuse, grooming,harassment,law and order, justice system,forced marriage,child marriage, illegal,exploitation,trafficking,prison
Terrorism	terror,extremis,isis,islamic state, ira,taliban,supremac,neonazi
Foreign Affairs	foreign,overseas,global, international,allies
War	war,conflict,warhead,warfare, warplane,invasion,bomb,missile, air strike,drone strike,ceasefire
Defence	defence,security,army,military, navy,air force,raf,armed force, trident,weapons,nato,arms,veteran
Pol Values - Auth	victim culture,cancel culture,politically correct, political correctness,woke,wokeness,wokeism, wokerati,wokery,snowflake,identity politics
Pol Values - Liberal	culture war,fascis,jingoi,far right, alt right,populis,nationalis,authoritarian, tolerance,civil libert,free speech, freedom of speech,human right

	Gender/Sexuality/Family	gender,sexuality,feminis,abortion, lgbt,patriarchy,trans,gay,lesbian, bisexual,rape culture,marriage, parental,conversion therapy
Other	Pol Values - Right	red tape,far left,marx communis,commie,statism
	Pol Values - Left	worker rights,workers rights,privatis, social justice,corporatis, neoliberal,neo liberal,public ownership
	Constitutional	constitution,voting system,voting reform, electoral system,electoral change,electoral reform, monarchy,welsh independence,good friday agreement, proportional representation,first past the post, firstpastthepost,fptp
	International Trade	trade,trading,imports,exports,protectionis
	Devolution	devolv,devo max,unionis
	Scot-ind	scottish independence,indyref,independent scotland
	Foreign/Domestic Emergency	genocide,earthquake,tornado,famine, volcan,flood,drought,energy shortage, food shortage,fuel shortage

TABLE C.6: Tweet Topic Keyword List - 436 keywords across 10 parent topics

A partial match was used in most cases except in certain cases where this may have been too broad, in which case an exact match was used. After coding of the tweets, a post-processing step was taken to de-code certain tweets where it contained a keyword in a particular category along with an additional term that would change its meaning such as "war" with "culture" or "deficit" with "democratic" and so on. Commonly recurring unigrams, bigrams, and trigrams were qualitatively examined by sub-topic to identify any commonly recurring terms and phrases that were missed, added into the keyword dictionary and then re-ran to update the tweet coding.
Figure C.5 illustrates the breakdown of MP tweets by topic and sub-topics:

FIGURE C.5: Treemap breakdown of MP tweet topics. The larger coloured groups represent parent topics and the groups inside them represent sub-topics. Area sizes are proportional to the number of tweets.

After taking a proportionally stratified random sample of 100 tweets from each parent topic of tweets, 84.1% of tweets were correctly coded and 76% of tweets in the non-coded group were correctly identified as non-topical. The breakdown of false-positive rates by topic are illustrated in **Figure C.6**.

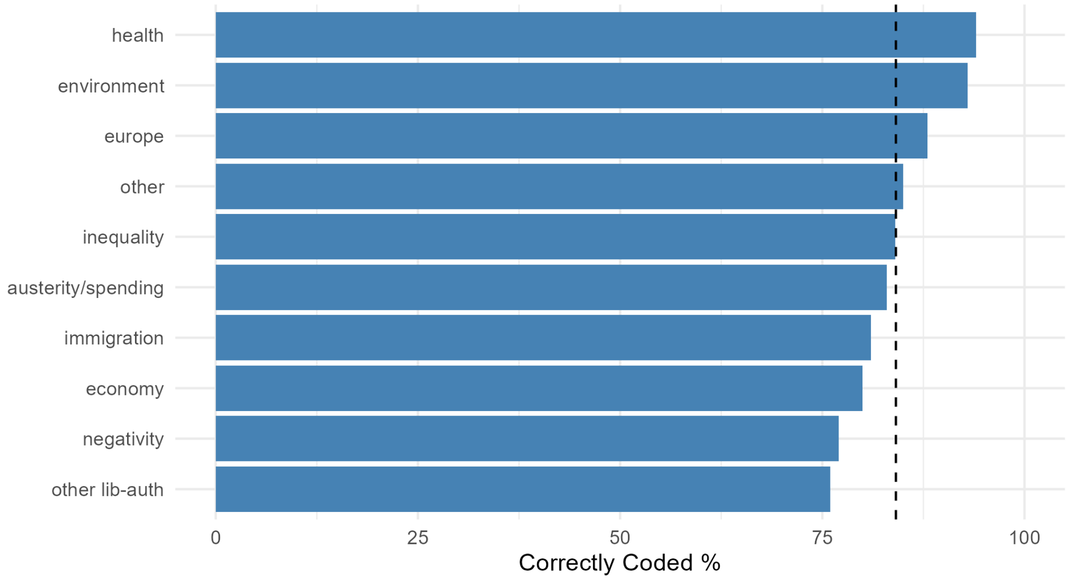


FIGURE C.6: False positive rate. The dashed line indicates the mean score.

References

Achen, C. H. (1978). Measuring representation. *American Journal of Political Science*, pages 475–510. doi: https://doi.org/10.2307/2110458.

Agarwal, P., Hawkins, O., Amaxopoulou, M., Dempsey, N., Sastry, N., and Wood, E. (2021). Hate speech in political discourse: A case study of uk mps on twitter. In *Proceedings of the 32nd ACM conference on hypertext and social media*, pages 5–16. doi: https://doi.org/10.1145/3465336.3475113.

Agarwal, P., Sastry, N., and Wood, E. (2019). Tweeting mps: Digital engagement between citizens and members of parliament in the uk. In *Proceedings of the International AAAI Conference on Web and Social Media*, volume 13, pages 26–37. doi: https://doi.org/10.1609/icwsm.v13i01.3359.

Agirre, E., Banea, C., Cer, D., Diab, M., Gonzalez Agirre, A., Mihalcea, R., Rigau Claramunt, G., and Wiebe, J. (2016). Semeval-2016 task 1: Semantic textual similarity, monolingual and cross-lingual evaluation. In *SemEval-2016. 10th International Workshop on Semantic Evaluation; 2016 Jun 16-17; San Diego, CA. Stroudsburg (PA): ACL; 2016. p. 497-511.* ACL (Association for Computational Linguistics). doi: https://doi.org/10.18653/v1/S16-1081.

Aiello, L. M., Barrat, A., Schifanella, R., Cattuto, C., Markines, B., and Menczer, F. (2012). Friendship prediction and homophily in social media. *ACM Transactions on the Web (TWEB)*, 6(2):1–33. doi: https://doi.org/10.1145/2180861.2180866.

Almog, A. and Shmueli, E. (2019). Structural entropy: monitoring correlation-based networks over time with application to financial markets. *Scientific reports*, 9(10832). doi: https://doi.org/10.1038/s41598-019-47210-8.

Almuhimedi, H., Wilson, S., Liu, B., Sadeh, N., and Acquisti, A. (2013). Tweets are forever: a large-scale quantitative analysis of deleted tweets. In *Proceedings of the 2013 conference on Computer supported cooperative work*, pages 897–908. doi: https://doi.org/10.1145/2441776.2441878.

Andrews, S., Brewster, B., and Day, T. (2018). Organised crime and social media: a system for detecting, corroborating and visualising weak signals of organised crime online. *Security Informatics*, 7(1):3. doi: https://doi.org/10.1186/s13388-018-0032-8.

Ansolabehere, S., Snyder Jr, J. M., and Stewart III, C. (2001). Candidate positioning in us house elections. *American Journal of Political Science*, pages 136–159. doi: https://doi.org/10.2307/2669364.

Arnold, E. (2023). Labour urges guarantee that channel 4 privatisation is 'off agenda for good'. *The Independent*. Online; accessed 3rd April 2024 at https://www.independent.co.uk/news/uk/lucy-powell-government-labour-secretary-of-state-caroline-dinenage-b2258880.html.

Auel, K. and Umit, R. (2018). Explaining mps' communication to their constituents: Evidence from the uk house of commons. *The British Journal of Politics and International Relations*, 20(3):731–752. doi: https://doi.org/10.1177/1369148118762280.

Ausserhofer, J. and Maireder, A. (2013). National politics on twitter: Structures and topics of a networked public sphere. *Information, Communication & Society*, 16(3):291–314. doi: https://doi.org/10.1080/1369118X.2012.756050.

Back, M. D., Stopfer, J. M., Vazire, S., Gaddis, S., Schmukle, S. C., Egloff, B., and Gosling, S. D. (2010). Facebook profiles reflect actual personality, not self-idealization. *Psychological Science*, 21:372–374. doi: https://doi.org/10.1177/0956797609360756.

Badawi, S. (2021). Covid-19 has exposed britain's client journalism problem. *Tribune*. Online; accessed 3rd April 2024 at https://tribunemag.co.uk/2021/01/covid-19-has-exposed-britains-client-journalism-problem.

Bail, C. (n.d). Strengths weaknesses of digital trace data. Online; accessed 4th Nov 2024 at https://sicss.io/2020/materials/day2-digital-trace-data/strengths-weaknesses/rmarkdown/Strengths_and_Weaknesses.html.

Bail, C. A., Argyle, L. P., Brown, T. W., Bumpus, J. P., Chen, H., Hunzaker, M. F., Lee, J., Mann, M., Merhout, F., and Volfovsky, A. (2018). Exposure to opposing views on social media can increase political polarization. *Proceedings of the National Academy of Sciences*, 115(37):9216–9221. doi: https://doi.org/10.1073/pnas.1804840115.

Bakker, R., Edwards, E., Jolly, S., Polk, J., Rovny, J., and Steenbergen, M. (2014). Anchoring the experts: Using vignettes to compare party ideology across countries. *Research Politics*, 1(3). doi: https://doi.org/10.1177/2053168014553502.

Barberá, P. (2015). Birds of the same feather tweet together: Bayesian ideal point estimation using twitter data. *Political analysis*, 23(1):76–91. doi: https://doi.org/10.1093/pan/mpu011.

Barberá, P., Jost, J. T., Nagler, J., Tucker, J. A., and Bonneau, R. (2015). Tweeting from left to right: Is online political communication more than an echo chamber?

Psychological science, 26(10):1531–1542. doi: https://doi.org/10.1177/0956797615594620.

Barnett, S. (2019). Journalists must not allow themselves to be used by unscrupulous politicians. *The Conversation*. Online; accessed 3rd April 2024 at https://theconversation.com/journalists-must-not-allow-themselves-to-be-used-by-unscrupulous-politicians-125883.

Barnett, S. (2022). Channel 4: why selling the broadcaster is a risky move for the uk government. *LSE Blog*. Online; accessed 3rd April 2024 at https://blogs.lse.ac.uk/medialse/2022/04/22/channel-4-why-selling-the-broadcaster-is-a-risky-move-for-the-uk-government/.

Barnett, S. and Petley, J. (2023). Come off it, ofcom. *British Journalism Review*, 34(2):17–23. doi: https://doi.org/10.1177/09564748231179343.

Barrett, E. J. and Cook, F. L. (1991). Congressional attitudes and voting behavior: An examination of support for social welfare. *Legislative Studies Quarterly*, pages 375–392. doi: https://doi.org/10.2307/440103.

Barrett, P. M. and Sims, J. G. (2021). *False Accusation: The Unfounded Claim that Social Media Companies Censor Conservatives*. NYU Stern Center for Business and Human Rights. Accessed at: https://bhr.stern.nyu.edu/publication/false-accusation-the-unfounded-claim-that-social-media-companies-censor-conservatives/.

Bartley, N., Abeliuk, A., Ferrara, E., and Lerman, K. (2021). Auditing algorithmic bias on twitter. In *Proceedings of the 13th ACM Web Science Conference 2021*, pages 65–73. doi: https://doi.org/10.1145/3447535.3462491.

Bastos, M., Mercea, D., and Baronchelli, A. (2018). The geographic embedding of online echo chambers: Evidence from the brexit campaign. *PloS one*, 13(11):e0206841.

Bawn, K. and Thies, M. F. (2003). A comparative theory of electoral incentives: representing the unorganized under pr, plurality and mixed-member electoral systems. *Journal of Theoretical Politics*, 15(1):5–32. doi: https://doi.org/10.1177/0951692803151001.

Bayram, U., Pestian, J., Santel, D., and Minai, A. A. (2019). What's in a word? detecting partisan affiliation from word use in congressional speeches. *In 2019 International Joint Conference on Neural Networks (IJCNN)*, pages 1–8. doi: https://doi.org/10.1109/IJCNN.2019.8851739.

BBC (n.da). About the bbc. global news services. Online; accessed 28th Jan 2025 at https://www.bbc.com/aboutthebbc/whatwedo/worldservice.

BBC (n.db). About the bbc. learn more about what we do. Online; accessed 3rd April 2024 at https://www.bbc.com/aboutthebbc.

BBCTrust (2011). Bbc trust review of impartiality and accuracy of the bbc's coverage of science. Report. Accessed at: https://downloads.bbc.co.uk/bbctrust/assets/files/pdf/our$_{w}ork/science_{i}mpartiality/science_{i}mpartiality.$

Beauchamp, N. (2011). Using text to scale legislatures with uninformative voting. *Unpublished transcript.New York University*. Accessed at: https://democraticwriting.com/work/Beauchamp$_{s}caling_{c}urrent.pdf$.

Bendixen, M. (1996). A practical guide to the use of correspondence analysis in marketing research. *Marketing Research On-Line*, 1(1):16–36.

Benedetto, G. and Hix, S. (2007). The rejected, the ejected, and the dejected: Explaining government rebels in the 2001-2005 british house of commons. *Comparative Political Studies*, 40(7):755–781. doi: https://doi.org/10.1177/0010414006299095.

Benigni, M. C., Joseph, K., and Carley, K. M. (2017). Online extremism and the communities that sustain it: Detecting the isis supporting community on twitter. *PloS one*, 12(12):e0181405. doi: https://doi.org/10.1371/journal.pone.0181405.

Benoit, K. and Laver, M. (2006). *Party policy in modern democracies.* Routledge.

Berrington, H. (1973). *Backbench opinion in the House of Commons 1945-55*. Oxford: Pergamon Press.

Berry, M. (2016). No alternative to austerity: how bbc broadcast news reported the deficit debate. *Media, Culture & Society*, 38(6):844–863. doi: https://doi.org/10.1177/0163443715620931.

Berry, W. D., Ringquist, E. J., Fording, R. C., and Hanson, R. L. (1998). Measuring citizen and government ideology in the american states, 1960- 93. *American Journal of Political Science*, 42(1):327–348. doi: https://doi.org/10.2307/2991759.

Binder, M. and Payton, A. L. (2022). With frenemies like these: Rising power voting behavior in the un general assembly. *British Journal of Political Science*, 52(1):381–398. doi: https://doi.org/10.1017/S0007123420000538.

Bisbee, J., Larson, J., and Munger, K. (2022). # polisci twitter: A descriptive analysis of how political scientists use twitter in 2019. *Perspectives on Politics*, 20(3):879–900. doi: https://doi.org/10.1017/S1537592720003643.

Black, D. (1958). On the rationale of group decision-making. *Journal of Political Economy*, 56(1):23–24.

Blank, G. (2017). The digital divide among twitter users and its implications for social research. *Social Science Computer Review*, 35(6):679–697. doi: https://doi.org/10.1177/0894439316671698.

Blank, G. and Lutz, C. (2016). The social structuration of six major social media platforms in the united kingdom: Facebook, linkedin, twitter, instagram, google+ and pinterest. In *Proceedings of the 7th 2016 International Conference on Social Media & Society*, pages 1–10. doi: https://doi.org/10.1145/2930971.2930979.

Bleakley, P. (2023). Panic, pizza and mainstreaming the alt-right: A social media analysis of pizzagate and the rise of the qanon conspiracy. *Current sociology*, 71(3):509–525. doi: https://doi.org/10.1177/00113921211034896.

Blidook, K. and Kerby, M. (2014). Constituency influence on 'constituency members': The adaptability of roles to electoral realities in the canadian case. In *The roles and function of parliamentary questions*, pages 69–81. Routledge.

Blumenau, J. (2021). Online activism and dyadic representation: Evidence from the uk e-petition system. *Legislative Studies Quarterly*, 46(4):889–920. doi: https://doi.org/10.1111/lsq.12291.

Blumenau, J. and Damiani, R. (2021). *Legislative Debates in the British House of Commons. In: Bäck, H and Debus, M and Fernandes, JM, (eds.) The Politics of Legislative Debate*. Oxford University Press: Oxford, UK. Accessed at: https://discovery.ucl.ac.uk/id/eprint/10109456/.

Bohdanova, T. (2014). Unexpected revolution: the role of social media in ukraine's euromaidan uprising. *European View*, 13(1):133–142. doi: https://doi.org/10.1007/s12290-014-0296-4.

Bond, R. and Messing, S. (2015). Quantifying social media's political space: Estimating ideology from publicly revealed preferences on facebook. *American Political Science Review*, 109(1):62–78. doi: https://doi.org/10.1017/S0003055414000525.

Bonica, A. (2014). Mapping the ideological marketplace. *American Journal of Political Science*, 58(2):367–386. doi: https://doi.org/10.1111/ajps.12062.

Booth, O., Butler, C., Jeffery, D., and Roe-Crines, A. (2023). Selecting sunak: Conservative mps' nomination preferences in the (second) british conservative party leadership election. *Parliamentary Affairs*. doi: https://doi.org/10.1093/pa/gsad010.

Booth, P. (2020). The future of public service broadcasting and the funding and ownership of the bbc. *Economic Affairs*, 40(3):324–343. doi: https://doi.org/10.1111/ecaf.12419.

Bouko, C. and Garcia, D. (2020). Patterns of emotional tweets: The case of brexit after the referendum results. In *Twitter, the Public Sphere, and the Chaos of Online Deliberation*. Palgrave Macmillan, Cham. doi: https://doi.org/10.1007/978-3-030-41421-4$_8$.

Boulianne, S. (2011). Stimulating or reinforcing political interest: Using panel data to examine reciprocal effects between news media and political interest. *Political communication*, 28(2):147–162. doi: https://doi.org/10.1080/10584609.2010.540305.

Bourgeois, D., Rappaz, J., and Aberer, K. (2018). Selection bias in news coverage: learning it, fighting it. In *Companion Proceedings of the The Web Conference 2018*, pages 535–543. doi: https://doi.org/10.1145/3184558.3188724.

Boutet, A., Kim, H., and Yoneki, E. (2013). What's in twitter, i know what parties are popular and who you are supporting now! *Social Network Analysis and Mining*, 3(4):1379–1391. doi: https://doi.org/10.1007/s13278-013-0120-1.

Boyd, D. and Crawford, K. (2012). Critical questions for big data: Provocations for a cultural, technological, and scholarly phenomenon. *Information, communication & society*, 15(5):662–679. doi: https://doi.org/10.1080/1369118X.2012.678878.

Brandenburg, H., Boyle, B. P., and Lemesheva, Y. (2024). Brexit and the iraq war on bbc question time: demographic and political issue representation in uk public participation broadcasting. *The International Journal of Press/Politics*. doi: https://doi.org/10.1177/19401612241292716.

Briatte, F. and Gallic, E. (2015). Recovering the french party space from twitter data. *In Science Po Quanti*. Accessed at: https://shs.hal.science/halshs-01511384v1.

Bright, J. (2016). Explaining the emergence of echo chambers on social media: the role of ideology and extremism. *arXiv preprint arXiv:1609.05003.* doi: https://doi.org/10.48550/arXiv.1609.05003.

Brito, K., Silva Filho, R. L. C., and Adeodato, P. J. L. (2024). Stop trying to predict elections only with twitter–there are other data sources and technical issues to be improved. *Government Information Quarterly*, 41(1). doi: https://doi.org/10.1016/j.giq.2023.101899.

Brito, K. D. S., Silva Filho, R. L. C., and Adeodato, P. J. L. (2021). A systematic review of predicting elections based on social media data: research challenges and future directions. *IEEE Transactions on Computational Social Systems*, 8(4):819–843. doi: https://doi.org/10.1109/TCSS.2021.3063660.

Bruns, A. (2017). Echo chamber? what echo chamber? reviewing the evidence. In *6th Biennial Future of Journalism Conference (FOJ17)*. Accessed at: https://eprints.qut.edu.au/113937/.

Bruns, A. (2019). *Are filter bubbles real?* John Wiley & Sons.

Bruns, A. (2021). Echo chambers? filter bubbles? the misleading metaphors that obscure the real problem. In *Hate speech and polarization in participatory society*, pages 33–48. Routledge.

Bruns, A. and Burgess, J. (2015). Twitter hashtags from ad hoc to calculated publics. *Hashtag publics: The power and politics of discursive networks*, pages 13–28.

Bruns, A. and Enli, G. (2018). The norwegian twittersphere: structure and dynamics. *Nordicom Review*, 39(1):129–148. doi: https://doi.org/10.2478/nor-2018-0006.

Bruns, A. and Highfield, T. (2015). Is habermas on twitter?: Social media and the public sphere. In *The Routledge companion to social media and politics*, pages 56–73. Routledge.

Budak, C., Goel, S., and Rao, J. M. (2016). Fair and balanced? quantifying media bias through crowdsourced content analysis. *Public Opinion Quarterly*, 80(S1):250–271. doi: https://doi.org/10.1093/poq/nfw007.

Budge, I. (2001). Validating party policy placements. *British Journal of Political Science*, 31(1):179–223. doi: https://doi.org/10.1017/S0007123401230087.

Bullock, J. G. (2011). Elite influence on public opinion in an informed electorate. *American Political Science Review*, 105(3):496–515. doi: https://doi.org/10.1017/S0003055411000165.

Bürkner, P.-C. (2017). brms: An r package for bayesian multilevel models using stan. *Journal of statistical software*, 80:1–28. doi: https://doi.org/10.18637/jss.v080.i01.

Burnap, P., Gibson, R., Sloan, L., Southern, R., and Williams, M. (2016). 140 characters to victory?: Using twitter to predict the uk 2015 general election. *Electoral Studies*, 41:230–233. doi: https://doi.org/10.1016/j.electstud.2015.11.017.

Burrell, I. (2023). How the bbc is undermining public trust in its self-harming quest for 'impartiality'. *inews*. Online; accessed 3rd April 2024 at https://inews.co.uk/news/media/how-the-bbc-is-undermining-public-trust-in-its-self-harming-quest-for-impartiality-2218581.

Buyens, W. and Van Aelst, P. (2022). Alternative media, alternative voices? a quantitative analysis of actor diversity in alternative and mainstream news outlets. *Digital Journalism*, 10(2):337–359. doi: https://doi.org/10.1080/21670811.2021.1929366.

Cain, B., Ferejohn, J., and Fiorina, M. (1987). *The personal vote: Constituency service and electoral independence*. Harvard University Press.

Caldarelli, G., Chessa, A., Pammolli, F., Pompa, G., Puliga, M., Riccaboni, M., and Riotta, G. (2014). A multi-level geographical study of italian political elections from twitter data. *PloS one*, 9(5):e95809. doi: https://doi.org/10.1371/journal.pone.0095809.

Cameron, C. M. and Park, J. K. (2009). How will they vote? predicting the future behavior of supreme court nominees, 1937–2006. *Journal of Empirical Legal Studies*, 6(3):485–511. doi: https://doi.org/10.1111/j.1740-1461.2009.01151.x.

Campbell, A., Harrop, A., and Thompson, B. (1999). Towards the virtual parliament – what computers can do for mps. *Parliamentary Affairs*, 52(3). doi: https://doi.org/10.1093/pa/52.3.388.

Campbell, R. and Lovenduski, J. (2015). What should mps do? public and parliamentarians' views compared. *Parliamentary Affairs*, 68(4):690–708. doi: https://doi.org/10.1093/pa/gsu020.

Canes-Wrone, B., Brady, D. W., and Cogan, J. F. (2002). Out of step, out of office: Electoral accountability and house members' voting. *American Political Science Review*, 96(1):127–140. doi: https://doi.org/10.1017/S0003055402004276.

Carson, R. T. and Oppenheimer, J. A. (1984). A method of estimating the personal ideology of political representatives. *American Political Science Review*, 78(1):163–178. doi: https://doi.org/10.2307/1961255.

CCDH (2024). Musk's x helps tommy robinson rack up 434 million views during uk riots. *Center for Countering Digital Hate*. Online; accessed 28th Oct 2024 at https://counterhate.com/research/musks-x-helps-tommy-robinson-rack-up-434-million-views-during-uk-riots/.

Çela, E. (2015). Social media as a new form of public sphere. *European Journal of Social Science Education and Research*, 2(3):195–200.

Cherepnalkoski, D. and Mozetič, I. (2016). Retweet networks of the european parliament: Evaluation of the community structure. *Applied Network Science*, 1(1):1–20. doi: https://doi.org/10.1007/s41109-016-0001-4.

Christensen, C. (2011a). Discourses of technology and liberation: State aid to net activists in an era of "twitter revolutions". *The Communication Review*, 14(3):233–253. doi: https://doi.org/10.1080/10714421.2011.597263.

Christensen, C. (2011b). Twitter revolutions? addressing social media and dissent. *The Communication Review*, 14(3):155–157. doi: https://doi.org/10.1080/10714421.2011.597235.

Chugh, R., Grose, R., and Macht, S. A. (2021). Social media usage by higher education academics: A scoping review of the literature. *Education and information technologies*, 26(1):983–999. doi: https://doi.org/10.1007/s10639-020-10288-z.

Cinelli, M., Morales, G. D. F., Galeazzi, A., Quattrociocchi, W., and Starnini, M. (2021). The echo chamber effect on social media. *Proceedings of the National Academy of Sciences*, 118(9). doi: https://doi.org/10.1073/pnas.2023301118.

Clark, J. K. and Evans, A. T. (2014). Source credibility and persuasion: The role of message position in self-validation. *Personality and Social Psychology Bulletin*, 40(8):1024–1036. doi: https://doi.org/10.1177/0146167214534733.

Clemence, M. and King, L. (2023). Trust in politicians reaches its lowest score in 40 years. *Ipsos Mori*. Online; accessed 27/11/2024 at https://www.ipsos.com/en-uk/ipsos-trust-in-professions-veracity-index-2023.

Clerici, P. (2021). Legislative territorialization: The impact of a decentralized party system on individual legislative behavior in argentina. *Publius: The Journal of Federalism*, 51(1):104–130. doi: https://doi.org/10.1093/publius/pjaa036.

Clinton, J., Jackman, S., and Rivers, D. (2004). The statistical analysis of roll call data. *American Political Science Review*, 98(2):355–370. doi: https://doi.org/10.1017/S0003055404001194.

Colleoni, E., Rozza, A., and Arvidsson, A. (2014). Echo chamber or public sphere? predicting political orientation and measuring political homophily in twitter using big data. *Journal of communication*, 64(2):317–332. doi: https://doi.org/10.1111/jcom.12084.

Conover, M. D., Gon¸calves, B., Ratkiewicz, J., Flammini, A., and Menczer, F. (2011). Predicting the political alignment of twitter users. *In 2011 IEEE third international conference on privacy, security, risk and trust and 2011 IEEE third international conference on social computing*, pages 192–199. doi: https://doi.org/10.1109/PASSAT/SocialCom.2011.34.

Converse, P. (1964). The nature of mass beliefs in mass publics. *Ideology and discontent*, pages 206–261.

Converse, P. and Pierce, R. (1986). *Political Representation in France*. Cambridge, MA: Harvard University Press.

Coughlan, L. (2024). *The Commentariat: A Demographic, Ideological and Rhetorical Analysis of Twitter Power-Users*. PhD thesis, United Kingdom: Royal Holloway, University of London.

Covert, T. J. A. and Wasburn, P. C. (2007). Measuring media bias: A content analysis of time and newsweek coverage of domestic social issues, 1975–2000. *Social science quarterly*, 88(3):690–706. doi: https://doi.org/10.1111/j.1540-6237.2007.00478.x.

Cowley, P. (2002). *Revolts and Rebellions: Parliamentary Voting Under Blair*. London: Politico's.

Cox, G. (1987). *The Efficient Secret: The Cabinet and the Development of Political Parties in Victorian England*. Cambridge: Cambridge University Press.

Crawford, J. T., Brandt, M. J., Inbar, Y., Chambers, J. R., and Motyl, M. (2017). Social and economic ideologies differentially predict prejudice across the political spectrum, but social issues are most divisive. *Journal of personality and social psychology*, 112(3):383. doi: https://psycnet.apa.org/doi/10.1037/pspa0000074.

Crines, A. S., Jeffery, D., and Heppell, T. (2018). The british labour party and leadership election mandate (s) of jeremy corbyn: Patterns of opinion and opposition within the parliamentary labour party. *Journal of Elections, Public Opinion and Parties*, 28(3):361–379. doi: https://doi.org/10.1080/17457289.2017.1403921.

Crossley, A. M. (1937). Straw polls in 1936. *Public Opinion Quarterly*, 1(1):24–35. doi: https://doi.org/10.1086/265035.

Curini, L. (2010). Experts' political preferences and their impact on ideological bias: An unfolding analysis based on a benoit-laver expert survey. *Party Politics*, 16(3):299–321. doi: https://doi.org/10.1177/1354068809341051.

Cushion, S., Kilby, A., Thomas, R., Morani, M., and Sambrook, R. (2018). Newspapers, impartiality and television news: Intermedia agenda-setting during the 2015 uk general election campaign. *Journalism Studies*, 19(2):162–181. doi: https://doi.org/10.1080/1461670X.2016.1171163.

Cushion, S. and Lewis, J. (2009). Towards afoxification'of 24-hour news channels in britain? an analysis of market-driven and publicly funded news coverage. *Journalism*, 10(2):131–153. doi: https://doi.org/10.1177/1464884908100598.

Cushion, S. and Lewis, J. (2017). Impartiality, statistical tit-for-tats and the construction of balance: Uk television news reporting of the 2016 eu referendum campaign. *European journal of communication*, 32(3):208–223. doi: https://doi.org/10.1177/0267323117695736.

Cygan, A., Lynch, P., and Whitaker, R. (2021). *MPs' Positions on the UK's Referendum on European Union Membership, 2016. [Data Collection].* UK Data Service. doi: https://doi.org/10.5255/UKDA-SN-854015.

Dagoula, C. (2019). Mapping political discussions on twitter: Where the elites remain elites. *Media and Communication*, 7(1):225–234. doi: https://doi.org/10.17645/mac.v7i1.1764.

Darius, P. (2022). Who polarizes twitter? ideological polarization, partisan groups and strategic networked campaigning on twitter during the 2017 and 2021 german federal elections' bundestagswahlen'. *Social Network Analysis and Mining*, 12(1):151. doi: https://doi.org/10.1007/s13278-022-00958-w.

Denham, A. and Dorey, P. (2018). Tales of the unexpected: The selection of british party leaders since 1963. *British Politics*, 13(2):171–194. doi: https://doi.org/10.1057/s41293-017-0044-x.

Dewan, T. and Spirling, A. (2011). Strategic opposition and government cohesion in westminster democracies. *American Political Science Review*, 105(2):337–358. doi: https://doi.org/10.1017/S0003055411000050.

Di Marco, N., Cinelli, M., and Quattrociocchi, W. (2021). Infodemics on youtube: Reliability of content and echo chambers on covid-19. *arXiv preprint arXiv:2106.08684*.

Divine, G. W., Norton, H. J., Barón, A. E., and Juarez-Colunga, E. (2018). The wilcoxon–mann–whitney procedure fails as a test of medians. *The American Statistician*, 72(3):278–286. doi: https://doi.org/10.1080/00031305.2017.1305291.

Djemili, S., Longhi, J., Marinica, C., Kotzinos, D., and Sarfati, G. E. (2014). What does twitter have to say about ideology?. *In NLP 4 CMC: Natural Language Processing for Computer-Mediated Communication/Social Media-Pre-conference workshop at Konvens 2014*. Accessed at: https://shs.hal.science/halshs-01058867v2.

Downs, A. (1957). *An Economic Theory of Democracy.* Harper.

Drury, B., Drury, S. M., Rahman, M. A., and Ullah, I. (2022). A social network of crime: A review of the use of social networks for crime and the detection of crime. *Online Social Networks and Media*, 30:100211. doi: https://doi.org/10.1016/j.osnem.2022.100211.

Dubois, E. and Blank, G. (2018). The echo chamber is overstated: the moderating effect of political interest and diverse media. *Information, communication & society*, 21(5):729–745. doi: https://doi.org/10.1080/1369118X.2018.1428656.

Dujeancourt, E. and Garz, M. (2023). The effects of algorithmic content selection on user engagement with news on twitter. *The Information Society*, 39(5):263–281. doi: https://doi.org/10.1080/01972243.2023.2230471.

Dutton, W. H., Reisdorf, B., Dubois, E., and Blank, G. (2017). *Social shaping of the politics of internet search and networking: Moving beyond filter bubbles, echo chambers, and fake news*. Quello Center Working Paper No. 2944191. doi: https://dx.doi.org/10.2139/ssrn.2944191.

Dvir-Gvirsman, S., Tsuriel, K., Sheafer, T., Shenhav, S., Zoizner, A., Lavi, L., Shamir, M., and Waismel-Manor, I. (2022). Mediated representation in the age of social media: How connection with politicians contributes to citizens' feelings of representation. evidence from a longitudinal study. *Political Communication*, 39(6):779–800. doi: https://doi.org/10.1080/10584609.2022.2124335.

Eagleton, T. (1991). *Ideology: an introduction*. London: Verso.

Earl, B. (2023). Social media and the british mp. Online; accessed 5th Sept 2024 at: https://review.brunswickgroup.com/article/social-media-british-mp/.

Eberl, J.-M., Boomgaarden, H. G., and Wagner, M. (2017). One bias fits all? three types of media bias and their effects on party preferences. *Communication Research*, 44(8):1125–1148. doi: https://doi.org/10.1177/0093650215614364.

Elmas, T. (2023). The impact of data persistence bias on social media studies. In *Proceedings of the 15th ACM web science conference 2023*, pages 196–207. doi: https://doi.org/10.1145/3578503.3583630.

Enelow, J. and Hinich, M. (1984). *The Spatial Theory of Voting: An Introduction*. Boston: Cambridge Univ. Press.

Erikson, R. S. (1978). Constituency opinion and congressional behavior: A reexamination of the miller-stokes representation data. *American Journal of Political Science*, pages 511–535. doi: https://doi.org/10.2307/2110459.

Esteve Del Valle, M., Broersma, M., and Ponsioen, A. (2022). Political interaction beyond party lines: Communication ties and party polarization in parliamentary twitter networks. *Social Science Computer Review*, 40(3):736–755. doi: https://doi.org/10.1177/0894439320987569.

Fagni, T. and Cresci, S. (2022). Fine-grained prediction of political leaning on social media with unsupervised deep learning. *Journal of Artificial Intelligence Research*, 73:633–672. doi: https://doi.org/10.1613/jair.1.13112.

Farkas, J. and Maloney, M. (2024). *Digital Media Metapahors: A Critical Introduction*. London: Routledge. doi: https://doi.org/10.4324/9781032674612.

Fazekas, Z., Popa, S. A., Schmitt, H., Barberá, P., and Theocharis, Y. (2021). Elite-public interaction on twitter: Eu issue expansion in the campaign. *European Journal of Political Research*, 60(2):376–396. doi: https://doi.org/10.1111/1475-6765.12402.

Fenno, R. F. (1977). Us house members in their constituencies: An exploration. *American Political Science Review*, 71(3):883–917. doi: https://doi.org/10.2307/1960097.

Ferrara, E., Wang, W.-Q., Varol, O., Flammini, A., and Galstyan, A. (2016). Predicting online extremism, content adopters, and interaction reciprocity. In *Social Informatics: 8th International Conference, SocInfo 2016, Bellevue, WA, USA, November 11-14, 2016, Proceedings, Part II 8*, pages 22–39. Springer. doi: https://doi.org/10.1007/978-3-319-47874-6$_3$.

Festinger, L. (1957). *A theory of cognitive dissonance*. Stanford University Press.

Fieldhouse, E., Green, J., Evans, G., Mellon, J., and Prosser, C. (2019). British election study 2019 constituency results file, version 1.1. doi: https://doi.org/10.48420/20278599.

Fieldhouse, E., Green, J., Evans, G., Mellon, J., Prosser, C., Bailey, J., de Geus, R., Schmitt, H., and van der Eijk, C. (2023). British election study internet panel waves 1-25. doi: https://doi.org/10.5255/UKDA-SN-8202-2.

Finer, S., Berrington, H., and Bartholomew, D. (1961). *Backbench Opinion in the House of Commons, 1955–59*. Oxford: Pergamon Press.

Fischer, P., Greitemeyer, T., and Frey, D. (2008). Self-regulation and selective exposure: the impact of depleted self-regulation resources on confirmatory information processing. *Journal of personality and social psychology*, 94(3):382. doi: https://psycnet.apa.org/doi/10.1037/0022-3514.94.3.382.

Flaxman, S., Goel, S., and Rao, J. M. (2016). Filter bubbles, echo chambers, and online news consumption. *Public opinion quarterly*, 80(S1):298–320. doi: https://doi.org/10.1093/poq/nfw006.

Fletcher, R., Kalogeropoulos, A., and Nielsen, R. K. (2023). More diverse, more politically varied: How social media, search engines and aggregators shape news repertoires in the united kingdom. *New Media & Society*, 25(8):2118–2139. doi: https://doi.org/10.1177/14614448211027393.

Fletcher, R. and Nielsen, R. K. (2018). Are people incidentally exposed to news on social media? a comparative analysis. *New media & society*, 20(7):2450–2468. doi: https://doi.org/10.1177/1461444817724170.

Francoli, M. and Ward, S. (2008). 21st century soapboxes? mps and their blogs. *Information Polity*, 13(1-2):21–39. doi: https://doi.org/10.3233/IP-2008-0143.

Franklin, M. N. and Tappin, M. (1977). Early day motions as unobtrusive measures of backbench opinion in britain. *British Journal of Political Science*, 7(1):49–69. doi: https://doi.org/10.1017/S0007123400003963.

Freeder, S., Lenz, G. S., and Turney, S. (2019). The importance of knowing "what goes with what": Reinterpreting the evidence on policy attitude stability. *The Journal of Politics*, 81(1):274–290. doi: https://doi.org/10.1086/700005.

Freeman, J. B. and Dale, R. (2013). Assessing bimodality to detect the presence of a dual cognitive process. *Behavior research methods*, 45:83–97. doi: https://doi.org/10.3758/s13428-012-0225-x.

Fuchs, C. (2012). The political economy of privacy on facebook. *Television & New Media*, 13(2):139–159. doi: https://doi.org/10.1177/1527476411415699.

Fuchs, C. (2015). Social media and the public sphere. In *Culture and economy in the age of social media*, pages 315–372. Routledge.

Gabel, M. and Scheve, K. (2007). Estimating the effect of elite communications on public opinion using instrumental variables. *American Journal of Political Science*, 51(4):1013–1028. doi: https://doi.org/10.1111/j.1540-5907.2007.00294.x.

Gao, H., Barbier, G., and Goolsby, R. (2011). Harnessing the crowdsourcing power of social media for disaster relief. *IEEE intelligent systems*, 26(3):10–14. doi: https://doi.org/10.1109/MIS.2011.52.

Garimella, V. R. K. and Weber, I. (2017). A long-term analysis of polarization on twitter. In *Proceedings of the International AAAI Conference on Web and social media*, volume 11, pages 528–531. doi: https://doi.org/10.1609/icwsm.v11i1.14918.

Garrett, R. K. (2013). Selective exposure: New methods and new directions. *Communication Methods and Measures*, 7(3-4):247–256. doi: https://doi.org/10.1080/19312458.2013.835796.

Gaughan, C. (2024). Estimating ideal points of british mps through their social media followership. *British Journal of Political Science*, pages 1–13. doi: https://doi.org/10.1017/S0007123424000450.

Gaughan, C. (n.d.). Assessing media balance through the social media networks of political elites. Unpublished.

Gay, O. (2005). Mps go back to their constituencies. *The Political Quarterly*, 76(1):57–66. doi: https://doi.org/10.1111/j.1467-923X.2005.00656.x.

Gentzkow, M. and Shapiro, J. M. (2010). What drives media slant? evidence from us daily newspapers. *Econometrica,*, 78(1):35–71. doi: https://doi.org/10.3982/ECTA7195.

Gentzkow, M. and Shapiro, J. M. (2011). Ideological segregation online and offline. *The Quarterly Journal of Economics*, 126(4):1799–1839. doi: https://doi.org/10.1093/qje/qjr044.

Gerring, J. (1997). Ideology: A definitional analysis. *Political Research Quarterly*, 50(4):957–994. doi: https://doi.org/10.1177/106591299705000412.

Gil de Zúñiga, H., Diehl, T., and Ardèvol-Abreu, A. (2018). When citizens and journalists interact on twitter: Expectations of journalists' performance on social media and perceptions of media bias. *Journalism studies*, 19(2):227–246. doi: https://doi.org/10.1080/1461670X.2016.1178593.

Gleason, B. (2013). # occupy wall street: Exploring informal learning about a social movement on twitter. *American Behavioral Scientist*, 57(7):966–982. doi: https://doi.org/10.1177/0002764213479372.

Goet, N. D. (2019). Measuring polarization with text analysis: Evidence from the uk house of commons, 1811–2015. *Political Analysis*, 27(4):518–539.

Goffman, E. (1959). *Presentation of Self in Everyday Life*. Carden City, New York: Doubleday Anchor.

Gold, N. (2022). Using twitter data in research guidance for researchers and ethics reviewers. Technical report, University College London. Accessed at: https://findocnet.fi/pluginfile.php/83578/mod$_l$*abel/intro/using − twitter − research − v*1.0.*pdf*.

Goodhind, W. (2023). Mps are boycotting early day motions. here is why, and how to fix it. *The Political Quarterly. Government & Parliament Blog.* Online; accessed at: https://politicalquarterly.org.uk/blog/mps-are-boycotting-early-day-motions-here-is-why-and-how-to-fix-it/.

Grčar, M., Cherepnalkoski, D., Mozetič, I., and Kralj Novak, P. (2017). Stance and influence of twitter users regarding the brexit referendum. *Computational social networks*, 4:1–25. doi: https://doi.org/10.1186/s40649-017-0042-6.

Greenacre, M. and Blasius, J. (2006). *Multiple Correspondence Analysis and Related Methods*. Boca Raton: Chapman Hall/CRC.

Greenacre, M. J. (2010). Correspondence analysis. *Wiley Interdisciplinary Reviews: Computational Statistics*, 2(5):613–619. doi: https://doi.org/10.1002/wics.114.

Grevet, C., Terveen, L. G., and Gilbert, E. (2014). Managing political differences in social media. *In Proceedings of the 17th ACM conference on Computer supported cooperative work social computing*, pages 1400–1408. doi: https://doi.org/10.1145/2531602.2531676.

Griffin, A. (2024). How a few twitter posts on elon musk's x helped fan the flames of unrest and rioting across the uk. *The Independent*. Online; accessed 28th Oct 2024 at https://www.independent.co.uk/tech/elon-musk-uk-riots-southport-twitter-x-b2591725.html.

Grimmer, J. and Stewart, B. M. (2013). Text as data: The promise and pitfalls of automatic content analysis methods for political texts. *Political analysis*, 21(3):267–297. doi: https://doi.org/10.1093/pan/mps028.

Groseclose, T. and Milyo, J. (2005). A measure of media bias. *The Quarterly Journal of Economics*, 120(4):1191–1237. doi: https://doi.org/10.1162/003355305775097542.

Gruzd, A. and Roy, J. (2014). Investigating political polarization on twitter: A canadian perspective. *Policy & internet*, 6(1):28–45. doi: https://doi.org/10.1002/1944-2866.POI354.

Gruzd, A., Wellman, B., and Takhteyev, Y. (2011). Imagining twitter as an imagined community. *American Behavioral Scientist*, 55(10):1294–1318. doi: https://doi.org/10.1177/0002764211409378.

Guardian Correspondents (2020). 'a giant': what other countries make of the bbc and how their media compare. *The Guardian*. Online; accessed 28th Jan, 2025 at https://www.theguardian.com/media/2020/jan/25/a-giant-what-other-countries-make-of-the-bbc-and-how-their-media-compare.

Guess, A. M. (2016). Media choice and moderation: Evidence from online tracking data. Unpublished manuscript. Accessed at: https://www.scribd.com/document/323356298/Media-Choice-and-Moderation-Evidence-from-Online-Tracking-Data.

Guess, A. M. (2021). (almost) everything in moderation: new evidence on americans' online media diets. *American Journal of Political Science*, 65(4):1007–1022. doi: https://doi.org/10.1111/ajps.12589.

Gunter, B. (2024). How is the bbc regarded around the world. In *The BBC and the Public*, pages 147–169. Springer. doi: https://doi.org/10.1007/978-3-031-67814-17.

Habermas, J. (1974). The public sphere: An encyclopedia article. *New German Critique*, 3:49–55.

Habermas, J. (1989). *The structural transformation of the public sphere, An inquiry into a category of bourgeois society*. Cambridge, MA: MIT Press.

Habermas, J. (2006). Political communication in media society: Does democracy still enjoy an epistemic dimension? the impact of normative theory on empirical research. *Communication theory*, 16(4):411–426. doi: https://doi.org/10.1111/j.1468-2885.2006.00280.x.

Hahn, C. (2024). The voice of the absent? the link between descriptive and substantive representation of the working class in western europe. *Political Studies*, 72(2):527–550. doi: https://doi.org/10.1177/00323217221126668.

Hain, P. (2012). Time to prick the westminster bubble. *British Journalism Review*, 23(2):74–77. doi: https://doi.org/10.1177/0956474812450674.

Hall, R. L. (1996). *Participation in congress*. New Haven: Yale University Press.

Hall, S. (2003). Howard to step outside the bubble. *The Guardian*. Online; accessed 27/11/2024 at https://www.theguardian.com/politics/2003/nov/06/uk.conservatives.

Hamborg, F., Donnay, K., and Gipp, B. (2019). Automated identification of media bias in news articles: an interdisciplinary literature review. *International Journal on Digital Libraries*, 20(4):391–415. doi: https://doi.org/10.1007/s00799-018-0261-y.

Hanretty, C. (2007). The gospel truths of italian media bias. *Comunicazione politica. Fascicolo 1, 2007*, pages 1000–1018. doi: https://doi.org/10.1400/94697.

Hanretty, C. (2020). An introduction to multilevel regression and post-stratification for estimating constituency opinion. *Political Studies Review*, 18(4):630–645. doi: https://doi.org/10.1177/1478929919864773.

Hanretty, C., Lauderdale, B. E., and Vivyan, N. (2017). Dyadic representation in a westminster system. *Legislative Studies Quarterly*, 42(2):235–267. doi: https://doi.org/10.1111/lsq.12148.

Hanretty, C., Lauderdale, B. E., and Vivyan, N. (2018). Comparing strategies for estimating constituency opinion from national survey samples. *Political Science Research and Methods*, 6(3):571–591. doi: https://doi.org/10.1017/psrm.2015.79.

Hartigan, J. A. and Hartigan, P. M. (1985). The dip test of unimodality. *The annals of Statistics*, pages 70–84. Accessed at: https://www.jstor.org/stable/2241144.

Hasan, M. (2009). Bias and the bbc. *The New Statesman*. Online; accessed 3rd April 2024 at https://www.newstatesman.com/long-reads/2009/08/mehdi-hasan-bbc-wing-bias-corporation.

Hassell, H. J., Holbein, J. B., and Miles, M. R. (2020). There is no liberal media bias in which news stories political journalists choose to cover. *Science Advances*, 6(14). doi: https://doi.org/10.1126/sciadv.aay9344.

Heppell, T. (2008). *Choosing the Tory Leader: Conservative Party Leadership Election form Heath to Cameron*. Lonodon: IB Tauris.

Heppell, T. (2010). *Choosing the Labour Leader: Labour Party Leadership Elections from Wilson to Brown*. London: Bloomsbury Publishing.

Heppell, T. (2022). The labour party leadership election: The stark model and the selection of keir starmer. *British Politics*, 17(4):369–385. doi: https://doi.org/10.1057/s41293-021-00164-w.

Hermida, A. (2015). Power plays on social media. *Social Media + Society*, 1(1). doi: https://doi.org/10.1177/2056305115580340.

Hern, A. (2022). Elon musk claims he has acquired twitter 'to help humanity'. *The Guardian*. Online; accessed 28th Oct 2024 at https://www.theguardian.com/technology/2022/oct/27/elon-musk-twitter-advertisers-donald-trump.

Herrick, R. and Moore, M. K. (1993). Political ambition's effect on legislative behavior: Schlesinger's typology reconsidered and revisited. *The Journal of Politics*, 55(3):765–776. doi: https://doi.org/10.2307/2132000.

Hickey, D., Schmitz, M., Fessler, D., Smaldino, P. E., Muric, G., and Burghardt, K. (2023). Auditing elon musk's impact on hate speech and bots. In *Proceedings of the international AAAI conference on web and social media*, volume 17, pages 1133–1137. doi: https://doi.org/10.1609/icwsm.v17i1.22222.

Hix, S. (2001). Legislative behaviour and party competition in the european parliament: An application of nominate to the eu. *Journal of Common Market Studies*, 39(4):663–688. doi: https://doi.org/10.1111/1468-5965.00326.

Hix, S. and Noury, A. (2010). Scaling the commons: Using mps' left–right self-placement and voting divisions to map the british parliament, 1997–2005. *Paper prepared for presentation at the annual meeting of the American Political Science Association in Washington, DC, September 2–5, 2010.*

Ho, S. S., Binder, A. R., Becker, A. B., Moy, P., Scheufele, D. A., Brossard, D., and Gunther, A. C. (2011). The role of perceptions of media bias in general and issue-specific political participation. *Mass Communication and Society*, 14(3):343–374. doi: https://doi.org/10.1080/15205436.2010.491933.

Hoffman, A. J. and Wallach, J. (2007). The effects of media bias. *Journal of Applied Social Psychology*, 37(3):616–630. doi: https://doi.org/10.1111/j.1559-1816.2007.00177.x.

Holmwood, L. (2009). Bbc news needs more tories, says shadow culture secretary jeremy hunt. *The Guardian*. Online; accessed 3rd April 2024 at https://www.theguardian.com/media/2009/sep/24/bbc-news-tories-jeremy-hunt.

Hong, S. and Kim, S. H. (2016). Political polarization on twitter: Implications for the use of social media in digital governments. *Government Information Quarterly*, 33(4):777–782. doi: https://doi.org/10.1016/j.giq.2016.04.007.

Hopmann, D. N., De Vreese, C. H., and Albæk, E. (2011). Incumbency bonus in election news coverage explained: The logics of political power and the media market. *Journal of Communication*, 61(2):264–282. doi: https://doi.org/10.1111/j.1460-2466.2011.01540.x.

Hopmann, D. N., Van Aelst, P., and Legnante, G. (2012). Political balance in the news: A review of concepts, operationalizations and key findings. *Journalism*, 13(2):240–257. doi: https://doi.org/10.1177/1464884911427804.

Hotelling, H. (1929). Stability in competition. *Economic Journal*, 39:41–57.

Houston, J. B., Hawthorne, J., Perreault, M. F., Park, E. H., Goldstein Hode, M., Halliwell, M. R., Turner McGowen, S. E., Davis, R., Vaid, S., McElderry, J. A., et al. (2015). Social media and disasters: a functional framework for social media use in disaster planning, response, and research. *Disasters*, 39(1):1–22. doi https://doi.org/10.1111/disa.12092.

Howard, P. N. and Hussain, M. M. (2013). *Democracy's fourth wave?: digital media and the Arab Spring*. Oxford: Oxford University Press.

Howison, J., Wiggins, A., and Crowston, K. (2011). Validity issues in the use of social network analysis with digital trace data. *Journal of the Association for Information Systems*, 12(12):2. doi: https://doi.org/10.17705/1jais.00282.

Huber, J. and Inglehart, R. (1984). Expert interpretations of party space and party locations in 42 societies. *Party Politics*, 1(1):73–111. doi: https://doi.org/10.1177/1354068895001001004.

Hughes, C., Morani, M., Cushion, S., and Kyriakidou, M. (2023). Does the political context shape how "due impartiality" is interpreted? an analysis of bbc reporting of the 2019 uk and 2020 us election campaigns. *Journalism Studies*, 24(14):1715–1733. doi: https://doi.org/10.1080/1461670X.2023.2173956.

Huszár, F., Ktena, S. I., O'Brien, C., Belli, L., Schlaikjer, A., and Hardt, M. (2022). Algorithmic amplification of politics on twitter. *Proceedings of the National Academy of Sciences*, 119(1). doi: https://doi.org/10.1073/pnas.2025334119.

Ittimani, L. (2024). Bluesky adds 1m new members as users flee x after the us election. *The Guardian*. Online; accessed 29/11/2024 at https://www.theguardian.com/technology/2024/nov/12/us-election-bluesky-users-flee-x-twitter-trump-musk.

Iyyer, M., Enns, P., Boyd-Graber, J., and Resnik, P. (2014). Political ideology detection using recursive neural networks. *In Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics*, 1:1113–1122. doi: https://doi.org/10.1177/1940161220935058.

Jackson, N. (2006). An mp's role in the internet era–the impact of e-newsletters. *The Journal of Legislative Studies*, 12(2):223–242. doi: https://doi.org/10.1080/13572330600739561.

Jackson, N. (2008). Representation in the blogosphere: Mps and their new constituents. *Parliamentary Affairs*, 61(4):642–660. doi: https://doi.org/10.1093/pa/gsn021.

Jackson, N. and Lilleker, D. G. (2009). Mps and e-representation: Me, myspace and i. *British Politics*, 4:236–264. doi: https://doi.org/10.1057/bp.2009.2.

Jackson, N. and Lilleker, D. G. (2011). Microblogging, constituency service and impression management: Uk mps and the use of twitter. *Journal of Legislative Studies*, 17(1):86–105. doi: https://doi.org/10.1080/13572334.2011.545181.

Jackson, S. J. and Foucault Welles, B. (2016). # ferguson is everywhere: Initiators in emerging counterpublic networks. *Information, Communication & Society*, 19(3):397–418. doi: https://doi.org/10.1080/1369118X.2015.1106571.

Jamieson, K. H. and Cappella, J. N. (2008). *Echo chamber: Rush limbaugh and the conservative media establishment*. Oxford University Press.

Jasny, L., Waggle, J., and Fisher, D. R. (2015). An empirical examination of echo chambers in us climate policy networks. *Nature Climate Change*, 5(8):782–786. doi: https://doi.org/10.1038/nclimate2666.

Jeffery, D., Heppell, T., Hayton, R., and Crines, A. (2018). The conservative party leadership election of 2016: an analysis of the voting motivations of conservative parliamentarians. *Parliamentary Affairs*, 71(2):263–282. doi: https://doi.org/10.1093/pa/gsx027.

Jeffery, D., Heppell, T., and Roe-Crines, A. (2022). The conservative party leadership election of 2019: An analysis of the voting motivations of conservative parliamentarians. *Parliamentary Affairs*, 75(1):113–134. doi: https://doi.org/10.1093/pa/gsaa046.

Jeffery, D., Heppell, T., Roe-Crines, A., and Butler, C. (2023). Trusting truss: Conservative mps' voting preferences in the (first) british conservative party leadership election of 2022. *Representation*, 59(4):555–572. doi: https://doi.org/10.1080/00344893.2023.2231469.

Jennings, W. and Stoker, G. (2016). The bifurcation of politics: Two englands. *The Political Quarterly*, 87(3):372–382. doi: https://doi.org/10.1111/1467-923X.12228.

Jennings, W. and Stoker, G. (2017). Tilting towards the cosmopolitan axis? political change in england and the 2017 general election. *The Political Quarterly*, 88(3):359–369. doi: https://doi.org/10.1111/1467-923X.12403.

Jennings, W., Stoker, G., Willis, H., Valgardsson, V., Gaskell, J., Devine, D., McKay, L., and Mills, M. C. (2021). Lack of trust and social media echo chambers predict covid-19 vaccine hesitancy. *MedRxiv*, pages 2021–01. doi: https://doi.org/10.1101/2021.01.26.21250246.

Jiang, J., Ren, X., Ferrara, E., et al. (2021). Social media polarization and echo chambers in the context of covid-19: Case study. *JMIRx med*, 2(3:e29570). doi: https://doi.org/10.2196/29570.

Jin, F., Wang, W., Chakraborty, P., Self, N., Chen, F., and Ramakrishnan, N. (2017). Tracking multiple social media for stock market event prediction. In *Advances in Data Mining. Applications and Theoretical Aspects: 17th Industrial Conference, ICDM 2017, New York, NY, USA, July 12-13, 2017, Proceedings 17*, pages 16–30. Springer. doi: https://doi.org/10.1007/978-3-319-62701-4_2.

Jolly, S., Bakker, R., Hooghe, L., Marks, G., Polk, J., Rovny, J., Steenbergen, M., and Vachudova, M. A. (2022). Chapel hill expert survey trend file, 1999–2019. *Electoral studies*, 75:102420. doi: https://doi.org/10.1016/j.electstud.2021.102420.

Jones, O. (2014). It's the bbc's rightwing bias that is the threat to democracy and journalism. *The Guardian*. Online; accessed 3rd April 2024 at https://www.theguardian.com/commentisfree/2014/mar/17/bbc-leftwing-bias-non-existent-myth.

Jurkowitz, M., Mitchell, A., Shearer, E., and Walker, M. (2020). U.s media polarization and the 2020 election: A nation divided. *Pew Research Center*. Online; accessed 3rd April 2024 at https://www.pewresearch.org/journalism/2020/01/24/u-s-media-polarization-and-the-2020-election-a-nation-divided/.

Kahan, D. M., Braman, D., Cohen, G. L., Gastil, J., and Slovic, P. (2010). Who fears the hpv vaccine, who doesn't, and why? an experimental study of the mechanisms of cultural cognition. *Law and human behavior*, 34:501–516. doi: https://doi.org/10.1007/s10979-009-9201-0.

Kam, C. (2001). Do ideological preferences explain parliamentary behaviour? evidence from great britain and canada. *Journal of Legislative Studies*, 7(4):89–126. doi: https://doi.org/10.1080/714003894.

Kam, C., Bianco, W. T., Sened, I., and Smyth, R. (2010). Ministerial selection and intraparty organization in the contemporary british parliament. *American Political Science Review*, 104(2):289–306. doi: https://doi.org/10.1017/S0003055410000080.

Kam, C. J. (2009). *Party discipline and parliamentary politics*. Cambridge University Press.

Kampakis, S. and Adamides, A. (2014). Using twitter to predict football outcomes. *arXiv preprint arXiv:1411.1243*. doi: https://doi.org/10.48550/arXiv.1411.1243.

Kaneko, T., Asano, T. A., and Miwa, H. (2021). Estimating ideal points of newspapers from editorial texts. *The International Journal of Press/Politics*, 26(3):719–742. doi: https://doi.org/10.1177/1940161220935058.

Kastellec, J. P., Lax, J. R., and Phillips, J. H. (2010). Public opinion and senate confirmation of supreme court nominees. *The Journal of Politics*, 72(3):767–784. doi: https://doi.org/10.1017/S0022381610000150.

Keller, T. (2020). To whom do politicians talk and listen?: Mapping swiss politicians' public sphere on twitter. *Computational Communication Research*, 2(2):175–202. doi: https://doi.org/10.5117/CCR2020.2.003.KELL.

Kellermann, M. (2012). Estimating ideal points in the british house of commons using early day motions. *American Journal of Political Science*, 56(3):757–771. doi: https://doi.org/10.1111/j.1540-5907.2012.00587.x.

Kellermann, M. (2016). Electoral vulnerability, constituency focus, and parliamentary questions in the house of commons. *The British Journal of Politics and International Relations*, 18(1):90–106. doi: https://doi.org/10.1111/1467-856X.12075.

Kemp, S. (2024). *Digital 2024 July Global Statshot Report*. DataReportal. Online; accessed 24th Sept 2024 at https://datareportal.com/social-media-users.

Kim, Y., Kim, Y., Lee, J. S., Oh, J., and Lee, N. Y. (2015). Tweeting the public: journalists' twitter use, attitudes toward the public's tweets, and the relationship with the public. *Information, Communication & Society*, 18(4):443–458. doi: https://doi.org/10.1080/1369118X.2014.967267.

King, A. (2002). Do leaders' personalities really matter? In *King A. (Ed.), Leaders' personalities and the outcomes of democratic elections*, pages 1–42. Oxford: Oxford University Press.

Klapper, J. T. (1960). *The effects of mass communication*. New York: Free Press.

Knobloch-Westerwick, S. and Kleinman, S. B. (2012). Preelection selective exposure: Confirmation bias versus informational utility. *Communication research*, 39(2):170–193. doi: https://doi.org/10.1177/0093650211400597.

Korda, H. and Itani, Z. (2013). Harnessing social media for health promotion and behavior change. *Health promotion practice*, 14(1):15–23. doi: https://doi.org/10.1177/1524839911405850.

Krehbiel, K. (1986). A technique for estimating congressmen's ideal points. *The Journal of Politics*, 48(1):97–115. doi: https://doi.org/10.2307/2130927.

Krimmel, K., Lax, J. R., and Phillips, J. H. (2016). Gay rights in congress: Public opinion and (mis) representation. *Public Opinion Quarterly*, 80(4):888–913. doi: https://doi.org/10.1093/poq/nfw026.

Kruse, L. M., Norris, D. R., and Flinchum, J. R. (2018). Social media as a public sphere? politics on social media. *The Sociological Quarterly*, 59(1):62–84. doi: https://doi.org/10.1080/00380253.2017.1383143.

Küpfer, A. (2024). Nonrandom tweet mortality and data access restrictions: Compromising the replication of sensitive twitter studies. *Political Analysis*, 32(4):493–506. doi: https://doi.org/10.1017/pan.2024.7.

Laor, T. (2022). Twitter as a clique: journalists' patterns of twitter use in israel. *Online Information Review*, 46(1):40–58. doi: https://doi.org/10.1108/OIR-07-2020-0324.

Larsson, A. O., Kalsnes, B., and Christensen, C. (2017). Elite interaction: Public service broadcasters' use of twitter during national elections in norway and sweden. *Journalism Practice*, 11(9):1137–1157. doi: https://doi.org/10.1080/17512786.2016.1234943.

Larsson, A. O. and Moe, H. (2012). Studying political microblogging: Twitter users in the 2010 swedish election campaign. *New media & society*, 14(5):729–747. doi: https://doi.org/10.1177/1461444811422894.

Lauderdale, B. E. and Herzog, A. (2016). Measuring political positions from legislative speech. *Political Analysis*, 24(3):374–394. doi: https://doi.org/10.1093/pan/mpw017.

Lax, J. R. and Phillips, J. H. (2009). How should we estimate public opinion in the states? *American Journal of Political Science*, 53(1):107–121. doi: https://doi.org/10.1111/j.1540-5907.2008.00360.x.

Leece, J. and Berrington, H. (1977). Measurements of backbench attitudes by guttman scaling of early day motions: A pilot study, labour, 1968–69. *British Journal of Political Science*, 7(4):529–541. doi: https://doi.org/10.1017/S0007123400001150.

Lewis, J. and Cushion, S. (2019). Think tanks, television news and impartiality: The ideological balance of sources in bbc programming. *Journalism Studies*, 20(4):480–499. doi: https://doi.org/10.1080/1461670X.2017.1389295.

Lewis, J. and Speers, T. (2003). Misleading media reporting? the mmr story. *Nature Reviews Immunology*, 3(11):913–918. doi: https://doi.org/10.1038/nri1228.

Liang, H. (2018). Broadcast versus viral spreading: the structure of diffusion cascades and selective sharing on social media. *Journal of Communication*, 68(3):525–546. doi: https://doi.org/10.1093/joc/jqy006.

Lin, J., Tremblay-Taylor, G., Mou, G., You, D., and Lee, K. (2019). Detecting fake news articles. In *2019 IEEE International Conference on Big Data (Big Data)*, pages 3021–3025. IEEE. doi: https://doi.org/10.1109/BigData47090.2019.9005980.

Lin, N. and Osnabrügge, M. (2018). Making comprehensible speeches when your constituents need it. *Research & Politics*, 5(3). doi: https://doi.org/10.1177/2053168018795598.

Lo, J. (2018). Dynamic ideal point estimation for the european parliament, 1980–2009. *Public Choice*, 176(1):229–246. doi: https://doi.org/10.1007/s11127-018-0551-3.

Loader, B. D. and Mercea, D. (2011). Networking democracy? social media innovations and participatory politics. *Information, communication & society*, 14(6):757–769. doi: https://doi.org/10.1080/1369118X.2011.592648.

Lohmann, S. and Zagheni, E. (2023). Diversity of social media use: Self-selection explains associations between using many platforms and well-being. *PLOS Digital Health*, 2(7):e0000292. doi: https://doi.org/10.1371/journal.pdig.0000292.

Lowe, W. (2008). Understanding wordscores. *Political Analysis*, 16(4):356–371. doi: https://doi.org/10.1093/pan/mpn004.

Mair, P. and Castles, F. (1984). Left-right political scales: some 'expert' judgements. *European Journal of Political Research*, 12(1):73 – 88.

Mansbridge, J. (1999). Should blacks represent blacks and women represent women? a contingent" yes". *The Journal of politics*, 61(3):628–657. doi: https://doi.org/10.2307/2647821.

Marks, G. and Steenbergen, M. R. (1984). European integration and political conflict. *Party Politics*, 1(1):73–111. doi: https://doi.org/10.1177/1354068895001001004.

Martin, A. R. (2021). Party group collapse and strategic switching in the european parliament. *European Union Politics*, 22(3):521–544. doi: https://doi.org/10.1177/1465116521999718.

Maurer, M., Haßler, J., Kruschinski, S., and Jost, P. (2022). Looking over the channel: The balance of media coverage about the "refugee crisis" in germany and the uk. *Communications*, 47(2):219–240. doi: https://doi.org/10.1515/commun-2020-0016.

Mayhew, D. (1974). *Congress: The Electoral Connection*. New Haven: Yale University Press.

McClain, C., Widjaya, R., Rivero, G., and Smith, A. (2021). The behaviors and attitudes of u.s. adults on twitter. *Pew Research Center*. Online; accessed 28th Oct 2024 https://www.pewresearch.org/internet/2021/11/15/the-behaviors-and-attitudes-of-u-s-adults-on-twitter/.

McDonnell, A. (2024). How britain voted in the 2024 general election. *YouGov UK*. Online; accessed 24th Sept 2024 at https://yougov.co.uk/politics/articles/49978-how-britain-voted-in-the-2024-general-election.

McDonnell, A. and Curtis, C. (2019). How britain voted in the 2019 general election. *YouGov UK*. Online; accessed 24th Sept 2024 at https://yougov.co.uk/politics/articles/26925-how-britain-voted-2019-general-election.

McKay, L. (2020). Does constituency focus improve attitudes to mps? a test for the uk. *The Journal of legislative studies*, 26(1):1–26. doi: https://doi.org/10.1080/13572334.2020.1726635.

McKelvey, K., DiGrazia, J., and Rojas, F. (2016). Twitter publics: How online political communities signaled electoral outcomes in the 2010 us house election. In *Current Research on Information Technologies and Society*, pages 46–60. Routledge.

McKnight, P. E. and Najab, J. (2010). Mann-whitney u test. *The Corsini encyclopedia of psychology*. doi: https://doi.org/10.1002/9780470479216.corpsy0524.

McLellan, D. (1986). Ideology: Concepts in social thought. *Minneapolis: U of Minnesota Press*.

McLoughlin, L. (2019). *How social media is changing political representation in the United Kingdom*. PhD thesis, United Kingdom: University of Salford.

McPherson, M., Smith-Lovin, L., and Cook, J. M. (2001). Birds of a feather: Homophily in social networks. *Annual Review of Sociology*, 27(1):415–444. doi: https://doi.org/10.1146/annurev.soc.27.1.415.

McQuail, D. (1992). *Media Performance*. London: Sage.

Meeks, L. (2018). Tweeted, deleted: theoretical, methodological, and ethical considerations for examining politicians' deleted tweets. *Information, Communication & Society*, 21(1):1–13. doi: https://doi.org/10.1080/1369118X.2016.1257041.

Mellon, J. and Prosser, C. (2017). Twitter and facebook are not representative of the general population: Political attitudes and demographics of british social media users. *Research & Politics*, 4(3). doi: https://doi.org/10.1177/2053168017720008.

Meraz, S. and Papacharissi, Z. (2013). Networked gatekeeping and networked framing on# egypt. *The international journal of press/politics*, 18(2):138–166. doi: https://doi.org/10.1177/1940161212474472.

Mercier, A. (2022). The 2022 state of journalism on twitter. *Muck Rack*. Online; accessed 28th Oct 2024 at: https://muckrack.com/blog/2022/05/18/2022-state-of-journalism-on-twitter.

Metaxas, P. T., Mustafaraj, E., Wong, K., Zeng, L., O'Keefe, M., and Finn, S. (2014). Do retweets indicate interest, trust, agreement? *arXiv preprint arXiv:1411.3555*. doi: https://doi.org/10.48550/arXiv.1411.3555.

Metzger, M. J., Hartsell, E. H., and Flanagin, A. J. (2020). Cognitive dissonance or credibility? a comparison of two theoretical explanations for selective exposure to partisan news. *Communication Research*, 47(1):3–28. doi: https://doi.org/10.1177/0093650215613136.

Miller, C., Weir, D., Ring, S., Marsh, O., Inskip, C., and Chavana, N. (2023). Antisemitism on twitter before and after elon musk's acquisition. Technical report, CASM Technology and the Institute for Strategic Dialogue. Accessed at: https://www.isdglobal.org/isd-publications/antisemitism-on-twitter-before-and-after-elon-musks-acquisition/.

Miller, W. E. and Stokes, D. E. (1963). Constituency influence in congress. *American political science review*, 57(1):45–56. doi: https://doi.org/10.2307/1952717.

Mills, T. (2020). *The BBC: Myth of a public service*. Verso Books.

Molyneux, L. and McGregor, S. C. (2022). Legitimating a platform: Evidence of journalists' role in transferring authority to twitter. *Information, Communication & Society*, 25(11):1577–1595. doi: https://doi.org/10.1080/1369118X.2021.1874037.

Monroe, B. L., Colaresi, M. P., and Quinn, K. M. (2008). Fightin'words: Lexical feature selection and evaluation for identifying the content of political conflict. *Political Analysis*, 16(4):372–403. doi: https://doi.org/10.1093/pan/mpn018.

Moore, M. (2024). Keeping democracies alive: The role of public service media. *The Political Quarterly*, 95(1):108–112. doi: https://doi.org/10.1111/1467-923X.13359.

Mora-Cantallops, M., Sánchez-Alonso, S., and Visvizi, A. (2021). The influence of external political events on social networks: The case of the brexit twitter network. *Journal of Ambient Intelligence and Humanized Computing*, 12(4):4363–4375. doi: https://doi.org/10.1007/s12652-019-01273-7.

Morani, M., Cushion, S., Kyriakidou, M., and Soo, N. (2022). Expert voices in the news reporting of the coronavirus pandemic: A study of uk television news bulletins and their audiences. *Journalism*, 23(12):2513–2532. doi: https://doi.org/10.1177/14648849221127629.

Morris, J. S. (2007). Slanted objectivity? perceived media bias, cable news exposure, and political attitudes. *Social science quarterly*, 88(3):707–728. doi: https://doi.org/10.1111/j.1540-6237.2007.00479.x.

Mosleh, M., Martel, C., Eckles, D., and Rand, D. G. (2021). Shared partisanship dramatically increases social tie formation in a twitter field experiment. *Proceedings of the National Academy of Sciences*, 118(7:e2022761118). doi: https://doi.org/10.1073/pnas.2022761118.

Mukerjee, S., Jaidka, K., and Lelkes, Y. (2022). The political landscape of the us twitterverse. *Political Communication*, 39(5):565–588. doi: https://doi.org/10.1080/10584609.2022.2075061.

Murphy, H. and Burn-Murdoch, J. (2024). Meta loses ground to bluesky as users abandon elon musk's x. *Financial Times*. Online: accessed 29/11/2024 at https://www.ft.com/content/e1b52147-c171-4902-8bce-204ba0905912.

Nelson, J. L. and Webster, J. G. (2017). The myth of partisan selective exposure: A portrait of the online political news audience. *Social media + society*, 3(3). doi: https://doi.org/10.1177/2056305117729314.

Németh, R. (2023). A scoping review on the use of natural language processing in research on political polarization: trends and research prospects. *Journal of Computational Social Science*, 6(1):289–313. doi: https://doi.org/10.1007/s42001-022-00196-2.

Nguyen, C. T. (2020). Echo chambers and epistemic bubbles. *Episteme*, 17(2):141–161. doi: https://doi.org/10.1017/epi.2018.32.

Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of general psychology*, 2(2):175–220. doi: https://doi.org/10.1037/1089-2680.2.2.175.

Nielsen, R. K. and Schrøder, K. C. (2014). The relative importance of social media for accessing, finding, and engaging with news: An eight-country cross-media comparison. *Digital journalism*, 2(4):472–489. doi: https://doi.org/10.1080/21670811.2013.872420.

Nielsen, R. K., Schulz, A., and Fletcher, R. (2023). The bbc is under scrutiny. here's what research tells about its role in the uk. *Reuters Institute for the Study of Journalism*. Online; accessed 4th April 2024 at https://reutersinstitute.politics.ox.ac.uk/news/bbc-under-scrutiny-heres-what-research-tells-about-its-role-uk.

Noonan, J. (2022). "where did their tweets go?": A quantitative analysis of parliamentarians "missing tweets" in western europe. Master's thesis, Uppsala University, Department of Government. Accssed at:

Norris, P. (1997). The puzzle of constituency service. *The Journal of Legislative Studies*, 3(2):29–49. doi: https://doi.org/10.1080/13572339708420508.

Norris, P. (2001). The twilight of westminster? electoral reform and its consequences. *Political Studies*, 49(5):877–900. doi: https://doi.org/10.1111/1467-9248.00345.

Norris, P. (2009). Comparative political communications: Common frameworks or babelian confusion? *Government and opposition*, 44(3):321–340. doi: doi:10.1111/j.1477-7053.2009.01290.x.

Norris, P. and Lovenduski, J. (1992). British candidate study, 1992. Online; accessed at http://www.pippanorris.com/.

Norris, P. and Lovenduski, J. (1997). British candidate study, 1997. Online; accessed at http://www.pippanorris.com/.

Norris, P. and Lovenduski, J. (2001). British candidate study, 2001. Online; accessed at http://www.pippanorris.com/.

North, S., Piwek, L., and Joinson, A. (2021). Battle for britain: Analyzing events as drivers of political tribalism in twitter discussions of brexit. *Policy & Internet*, 13(2):185–208. doi: https://doi.org/10.1002/poi3.247.

Norton, P. (1994). The growth of the constituency role of the mp. *Parliamentary Affairs*, 47(4):705–720.

Oborne, P. (2019). In its election coverage, the bbc has let down the people who believe in it. *The Guardian*. Online; accessed 3rd April 2024 at https://www.theguardian.com/commentisfree/2019/dec/03/election-coverage-bbc-tories.

Ofcom (2004). Psb review - a plain english summary. ofcom review of public service television broadcasting. Online; accessed 3rd April 2024 at https://www.ofcom.org.uk/consultations-and-statements/category-2/psb.

Ofcom (2021). The ofcom broadcasting code. section five: Due impartiality and due accuracy. Online; accessed 3rd April 2024 at https://www.ofcom.org.uk/tv-radio-and-on-demand/broadcast-codes/broadcast-code/section-five-due-impartiality-accuracy.

Ofcom (2024a). News consumption in the uk: 2024. *Research Report*. Accessed at: https://www.ofcom.org.uk/media-use-and-attitudes/attitudes-to-news/news-consumption/.

Ofcom (2024b). Online nation: 2024 report. *Ofcom*. Online: accessed 30/11/2024 at https://www.ofcom.org.uk/media-use-and-attitudes/online-habits/online-nation.

Ofcom (n.d). Public service broadcasting (psb). Online; accessed 3rd April 2024 at https://www.ofcom.org.uk/tv-radio-and-on-demand/information-for-industry/psb.

O'Malley, J. (2019). The westminster twitter bubble — an illustrated guide. *Politico*. Online; accessed 27/11/2024 at https://www.politico.eu/article/westminster-twitter-bubble-illustrated-guide-graphics/.

Pack, M. (2018). Political views of twitter users do not match the wider population, but that matters less than you might think. Online; accessed 21st Sept 2024 at: https://www.markpack.org.uk/153955/twitter-users-political-views/.

Palmer, M. W. (1993). Putting things in even better order: the advantages of canonical correspondence analysis. *Ecology*, 74(8):2215–2230. doi: https://doi.org/10.2307/1939575.

Papacharissi, Z. (2008). The virtual sphere 2.0: The internet, the public sphere, and beyond. In *Routledge handbook of Internet politics*, pages 230–245. Routledge.

Papacharissi, Z. (2010). *A private sphere: Democracy in a digital age*. Cambridge: Polity Press.

Papacharissi, Z. and de Fatima Oliveira, M. (2008). News frames terrorism: A comparative analysis of frames employed in terrorism coverage in us and uk newspapers. *The international journal of press/politics*, 13(1):52–74. doi: https://doi.org/10.1177/1940161207312676.

Pariser, E. (2011). *The filter bubble. what the internet is hiding from you*. Penguin Group.

Park, D. K., Gelman, A., and Bafumi, J. (2004). Bayesian multilevel estimation with poststratification: State-level estimates from national polls. *Political Analysis*, 12(4):375–385. doi: https://doi.org/10.1093/pan/mph024.

Park, S., Ko, M., Kim, J., Choi, H.-J., and Song, J. (2011). Newscube 2.0: an exploratory design of a social news website for media bias mitigation. In *Workshop on Social Recommender Systems*, pages 1–5. url: https://nclab.kaist.ac.kr/files/papers/Conference/NC2.pdf.

Patel, P. and Quilter-Pinner, H. (2022). Road to renewal: Elections, parties and the case for renewing democracy. institute for public policy research. *Institute for Public Policy Research*. Online; accessed 27/11/2024 at https://ippr-org.files.svdcdn.com/production/Downloads/road-to-renewal-april22.pdf.

Paul, M. and Dredze, M. (2011). You are what you tweet: Analyzing twitter for public health. In *Proceedings of the international AAAI conference on web and social media*, volume 5, pages 265–272. doi: https://doi.org/10.1609/icwsm.v5i1.14137.

Peat, J. (2020). Laura kuenssberg's "client journalism" called out after she posts unnamed attacks on starmer. *The London Economic*. Online; accessed 3rd April 2024 at https://www.thelondoneconomic.com/news/laura-kuenssbergs-client-journalism-called-out-after-she-posts-unnamed-attacks-on-starmer-205486/.

Pennycook, G., Epstein, Z., Mosleh, M., Arechar, A. A., Eckles, D., and Rand, D. G. (2021). Shifting attention to accuracy can reduce misinformation online. *Nature*, 592(7855):590–595. doi: https://doi.org/10.1038/s41586-021-03344-2.

Petley, J. (2020). Foxification: Just what a divided country doesn't need. 7:116–123. in Mair, J. (ed.) The BBC. A winter of Discontent?

Pfeffer, J., Mooseder, A., Lasser, J., Hammer, L., Stritzel, O., and Garcia, D. (2023). This sample seems to be good enough! assessing coverage and temporal reliability of twitter's academic api. In *Proceedings of the International AAAI Conference on Web and Social Media*, volume 17, pages 720–729. doi: https://doi.org/10.1609/icwsm.v17i1.22182.

Pitkin, H. F. (1969). *The Concept of Representation*. Berkeley: University of California Press.

Poole, K. T. and Rosenthal, H. (1985). A spatial model for legislative roll call analysis. *American Journal of Political Science*, 29(2):357–384. doi: https://doi.org/10.2307/2111172.

Poole, K. T. and Rosenthal, H. (1991). Patterns of congressional voting. *American Journal of Political Science*, 35(1):228–278. doi: https://doi.org/10.2307/2111445.

Poole, K. T. and Rosenthal, H. (1997). *Congress: A Political Economic History of Rollcall Voting*. Oxford University Press.

Potter, S. J. (2012). *Broadcasting empire: The BBC and the British world, 1922-1970*. Oxford University Press, USA. Available at: http://ukcatalogue.oup.com/product/9780199568963.do.

Preoţiuc-Pietro, D., Liu, Y., Hopkins, D., and Ungar, L. (2017). Beyond binary labels: political ideology prediction of twitter users. *In Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics*, Volume 1: Long Papers:729–740. doi: https://doi.org/10.18653/v1/P17-1068.

Priniski, J. H., McClay, M., and Holyoak, K. J. (2021). Rise of qanon: A mental model of good and evil stews in an echochamber. *arXiv preprint arXiv:2105.04632*. doi: https://doi.org/10.48550/arXiv.2105.04632.

Proksch, S.-O. and Slapin, J. (2014). *The Politics of Parliamentary Debate*. Cambridge: Cambridge University Press.

Quinn, B. and Mason, R. (2024). Gb news is being allowed to break impartiality rules, say senior tv figures. *The Guardian*. Online; accessed 3rd April 2024 at https://www.theguardian.com/media/2024/feb/24/gb-news-is-being-allowed-to-break-impartiality-rules-say-senior-tv-figures.

Rao, U. G. (1982). Generalized least squares procedure for regression with autocorrelated errors. In *Developments in Water Science*, volume 17, pages 100–109. Elsevier. doi: https://doi.org/10.1016/S0167-5648(08)70705-6.

Ray, L. (1999). Measuring party orientations towards european integration: Results from an expert survey. *European Journal of Political Research*, 36(2):283–306. doi: https://doi.org/10.1023/A:1007009106184.

Rayner, G. (2022). Bbc dramas 'warping modern britain' by 'allowing left-wing bias'. *The Telegraph*. Online; accessed 3rd April 2024 at https://www.telegraph.co.uk/news/2022/12/29/bbc-dramas-warping-modern-britain-allowing-left-wing-bias/.

Reece, A. G., Reagan, A. J., Lix, K. L., Dodds, P. S., Danforth, C. M., and Langer, E. J. (2017). Forecasting the onset and course of mental illness with twitter data. *Scientific reports*, 7(1):13006. doi: https://doi.org/10.1038/s41598-017-12961-9.

Reher, S. (2022). Do disabled candidates represent disabled citizens? *British Journal of Political Science*, 52(2):520–534. doi: https://doi.org/10.1017/S0007123420000733.

Reinecke, L. and Trepte, S. (2014). Authenticity and well-being on social network sites: A two-wave longitudinal study on the effects of online authenticity and the positivity bias in sns communication. *Computers in Human Behavior*, 30:95–102. doi: https://doi.org/10.1016/j.chb.2013.07.030.

Resende de Mendonça, R., Felix de Brito, D., de Franco Rosa, F., dos Reis, J. C., and Bonacin, R. (2020). A framework for detecting intentions of criminal acts in social media: A case study on twitter. *Information*, 11(3):154. doi: https://doi.org/10.3390/info11030154.

Reynolds, A. (2013). Representation and rights: The impact of lgbt legislators in comparative perspective. *American Political Science Review*, 107(2):259–274. doi:10.1017/S0003055413000051.

Ringe, N., Victor, J. N., and Gross, J. H. (2013). Keeping your friends close and your enemies closer? information networks in legislative politics. *British journal of political science*, 43(3):601–628. doi: https://doi.org/10.1017/S0007123412000518.

Roberts, A. (2022). Channel 4 has become a left-wing frankenstein. *The Telegraph*. Online; accessed 3rd April 2024 at https://www.telegraph.co.uk/politics/2022/04/05/channel-4-has-become-left-wing-frankenstein/.

Robertson, C. T. (2023). Here's what our research says about news audiences on twitter, the platform now known as x. Technical report, Reuters Institute. University of Oxford. Online; accessed 4th Nov 2024 at https://reutersinstitute.politics.ox.ac.uk/news/heres-what-our-research-says-about-news-audiences-twitter-platform-now-known-x.

Rodrigo-Ginés, F.-J., Carrillo-de Albornoz, J., and Plaza, L. (2024). A systematic review on media bias detection: What is media bias, how it is expressed, and how to detect it. *Expert Systems with Applications*, 237:121641. doi: https://doi.org/10.1016/j.eswa.2023.121641.

Roe-Crines, A., Heppell, T., and Jeffery, D. (2021). Theresa may and the conservative party leadership confidence motion of 2018: Analysing the voting behaviour of conservative parliamentarians. *British Politics*, 16:317–335. doi: https://doi.org/10.1057/s41293-020-00138-4.

Rogers, E. M. (2001). The digital divide. *Convergence*, 7(4):96–111. doi: https://doi.org/10.1177/135485650100700406.

Rogers, S. (2017). Electoral accountability for state legislative roll calls and ideological representation. *American Political Science Review*, 111(3):555–571. doi: https://doi.org/10.1017/S0003055417000156.

Rosenthal, H. and Voeten, E. (2004). Analyzing roll calls with perfect spatial voting: France 1946–1958. *American Journal of Political Science*, 48(3):620–632. doi: https://doi.org/10.1111/j.0092-5853.2004.00091.x.

Saalfeld, T. (2014). Parliamentary questions as instruments of substantive representation: Visible minorities in the uk house of commons, 2005–10. In *The roles and function of parliamentary questions*, pages 13–31. Routledge.

Saez-Trumper, D., Castillo, C., and Lalmas, M. (2013). Social media news communities: gatekeeping, coverage, and statement bias. In *Proceedings of the 22nd ACM international conference on Information & Knowledge Management*, pages 1679–1684. doi: https://doi.org/10.1145/2505515.2505623.

Sanders, L. (2022). Trust in media 2022: Where americans get their news and who they trust for information. *YouGov*. Onlines; accessed 28th Jan 2025 at https://today.yougov.com/politics/articles/41957-trust-media-2022-where-americans-get-news-poll.

Sang, E. T. K. and Bos, J. (2012). Predicting the 2011 dutch senate election results with twitter. In *Proceedings of the workshop on semantic analysis in social media*, pages 53–60. Accessed at: https://aclanthology.org/W12-0607.pdf.

Schoenbach, K., De Ridder, J., and Lauf, E. (2001). Politicians on tv news: Getting attention in dutch and german election campaigns. *European Journal of Political Research*, 39(4):519–531. doi: https://doi.org/10.1111/1475-6765.00586.

Schreurs, L. and Vandenbosch, L. (2021). Introducing the social media literacy (smile) model with the case of the positivity bias on social media. *Journal of Children and Media*, 15(3):320–337. doi: https://doi.org/10.1080/17482798.2020.1809481.

Schwarz, D., Traber, D., and Benoit, K. (2017). Estimating intra-party preferences: comparing speeches to votes. *Political Science Research and Methods*, 5(2):379–396. doi: https://doi.org/10.1017/psrm.2015.77.

Schweiger, S. and Cress, U. (2019). Attitude confidence and source credibility in information foraging with social tags. *PloS one*, 14(1: e0210423). doi: https://doi.org/10.1371/journal.pone.0210423.

Searing, D. (1994). *Westminster's World: Understanding Political Roles*. London: Harvard University Press.

Sears, D. O. and Freedman, J. L. (1967). Selective exposure to information: A critical review. *Public opinion quarterly*, 31(2):194–213. doi: https://doi.org/10.1086/267513.

Semetko, H. (2003). Political bias in the media. In *Johnston D (ed.) Encyclopedia of International Media and Communications 3*, pages 517–25. Amsterdam: Academic Press.

Semetko, H. A. and Canel, M. J. (1997). Agenda-senders versus agenda-setters: Television in spain's 1996 election campaign. *Political Communication*, 14(4):459–479. doi: https://doi.org/10.1080/105846097199263.

Sen, I., Flöck, F., Weller, K., Weiß, B., and Wagner, C. (2021). A total error framework for digital traces of human behavior on online platforms. *Public Opinion Quarterly*, 85(S1):399–422. doi: https://doi.org/10.1093/poq/nfab018.

Serafino, P. (2019). Exploring the uk's digital divide. *Office for National Statistics*, 2019:1–24. Online; available at: https://www.ons.gov.uk/peoplepopulationandcommunity/householdcharacteristics/homeinternetandsocialmediausage/articles/exploringtheuksdigitaldivide/2019-03-04.

Sheafer, T. and Weimann, G. (2018). An empirical analysis of the issue of media bias in israeli elections, 1996-2003. In *The Elections in Israel 2003*, pages 123–142. Routledge.

Sheather, S. J. and Jones, M. C. (1991). A reliable data-based bandwidth selection method for kernel density estimation. *Journal of the Royal Statistical Society: Series B (Methodological)*, 53(3):683–690. doi: https://doi.org/10.1111/j.2517-6161.1991.tb01857.x.

Shore, J., Baek, J., and Dellarocas, C. (2018). Twitter is not the echo chamber we think it is. *MIT Sloan Management Review*, 60(1):1–5. Accessed at: https://www.proquest.com/scholarly-journals/twitter-is-not-echo-chamber-we-think/docview/2137400837/se-2.

Shu, K., Sliva, A., Wang, S., Tang, J., and Liu, H. (2017). Fake news detection on social media: A data mining perspective. *ACM SIGKDD explorations newsletter*, 19(1):22–36. doi: https://doi.org/10.1145/3137597.3137600.

Sim, Y., Acree, B. D., Gross, J. H., and Smith, N. A. (2013). Measuring ideological proportions in political speeches. *In Proceedings of the 2013 conference on empirical methods in natural language processing*, pages 91–101. Accessed at: https://aclanthology.org/D13-1010.pdf.

Singh, M. (2018). *We knew social media wasn't representative, but these numbers are incredible*. Number Cruncher Politics. Online; accessed 24th Sept 2024 at https://www.ncpolitics.uk/2018/01/we-knew-social-media-wasnt-representative-but-these-numbers-are-incredible/.

Skoric, M., Poor, N., Achananuparp, P., Lim, E.-P., and Jiang, J. (2012). Tweets and votes: A study of the 2011 singapore general election. In *2012 45th hawaii international conference on system sciences*, pages 2583–2591. IEEE. doi: https://doi.org/10.1109/HICSS.2012.607.

Slapin, J. B., Kirkland, J. H., Lazzaro, J. A., Leslie, P. A., and O'grady, T. (2018). Ideology, grandstanding, and strategic party disloyalty in the british parliament. *American Political Science Review*, 112(1):15–30. doi: https://doi.org/10.1017/S0003055417000375.

Slapin, J. B. and Proksch, S.-O. (2008). A scaling model for estimating time-series party positions from texts. *American Journal of Political Science*, 52(3):705–722. doi: https://doi.org/10.1111/j.1540-5907.2008.00338.x.

Sloan, L. (2017). Who tweets in the united kingdom? profiling the twitter population using the british social attitudes survey 2015. *Social Media + Society*, 3(1):1–11. doi: https://doi.org/10.1177/2056305117698981.

Smith, M. (2023). Which media outlets do britons trust in 2023? *YouGov UK*. Online; accessed 4th April 2024 at https://yougov.co.uk/politics/articles/45744-which-media-outlets-do-britons-trust-2023.

Sobolewska, M., McKee, R., and Campbell, R. (2018). Explaining motivation to represent: how does descriptive representation lead to substantive representation of racial and ethnic minorities? *West European Politics*, 41(6):1237–1261. doi: https://doi.org/10.1080/01402382.2018.1455408.

Soo, N., Weinberg, J., and Dommett, K. (2020). One moment, please: Can the speed and quality of political contact affect democratic health? *The British Journal of Politics and International Relations*, 22(3):460–484. doi: https://doi.org/10.1177/1369148120919683.

Soroka, S., Penner, E., and Blidook, K. (2009). Constituency influence in parliament. *Canadian Journal of Political Science/Revue canadienne de science politique*, 42(3):563–591. doi: https://doi.org/10.1017/S0008423909990059.

Sourial, N., Wolfson, C., Zhu, B., Quail, J., Fletcher, J., Karunananthan, S., Bandeen-Roche, K., Béland, F., and Bergman, H. (2010). Correspondence analysis is a useful tool to uncover the relationships among categorical variables. *Journal of clinical epidemiology*, 63(6):638–646. doi: https://doi.org/10.1016/j.jclinepi.2009.08.008.

Souza, R. M. D., Graça, L. F. G. D., and Silva, R. D. S. (2017). Politics on the web: using twitter to estimate the ideological positions of brazilian representatives. *Brazilian Political Science Review*, 11(3). doi: https://doi.org/10.1590/1981-3821201700030003.

Spirling, A. and McLean, I. (2007). Uk oc ok? interpreting optimal classification scores for the uk house of commons. *Political Analysis*, 15(1):85–96. doi: https://doi.org/10.1093/pan/mpl009.

Spirling, A. and Quinn, K. (2010). Identifying intraparty voting blocs in the uk house of commons. *Journal of the American Statistical Association*, 105(490):447–457. doi: https://doi.org/10.1198/jasa.2009.ap07115.

Spirling, A. P. and Quinn, K. (2009). Identifying intra-party voting blocs in uk house of commons. *Journal of the American Statistical Association*, 105(409):447–457. doi: https://doi.org/10.1198/jasa.2009.ap07115.

Stanbro, M. (2012). Hartigan's dip test of unimodality applied on terrestrial gamma-ray flashes. honors thesis. Available at: https://louis.uah.edu/honors-capstones/601/.

Stanyer, J. (2008). Elected representatives, online self-preservation and the personal vote: party, personality and webstyles in the united states and the united kingdom. *Information, Communication and Society*, 11(3):414–432. doi: https://doi.org/10.1080/13691180802025681.

Stark, J. P. (1996). *Choosing a Leader: Party Leadership Contests in Britain from Macmillan to Blair.* Macmillan.

Starke, C., Marcinkowski, F., and Wintterlin, F. (2020). Social networking sites, personalization, and trust in government: Empirical evidence for a mediation model. *Social Media+ Society*, 6(2):1–11. doi: https://doi.org/10.1177/2056305120913885.

Stiers, D. (2022). Spatial and valence models of voting: The effects of the political context. *Electoral Studies*, 80:102549. doi: https://doi.org/10.1016/j.electstud.2022.102549.

Stone, J. (2021). Boris johnson dismisses hancock scandal as 'stuff going on in westminster bubble', sparking row. *The Independent*. Online; accessed 27/11/2024 at https://www.independent.co.uk/news/uk/politics/boris-johnson-hancock-westminster-bubble-b1875437.html.

Strömbäck, J., Djerf-Pierre, M., and Shehata, A. (2013). The dynamics of political interest and news media consumption: A longitudinal perspective. *International journal of public opinion research*, 25(4):414–435. doi: https://doi.org/10.1093/ijpor/eds018.

Stroud, N. J. (2008). Media use and political predispositions: Revisiting the concept of selective exposure. *Political behavior*, 30:341–366. doi: https://doi.org/10.1007/s11109-007-9050-9.

Sturgis, P., Baker, N., M., C., Fisher, S., Green, J., Jennings, W., J., K., Lauderdale, B., and Smith, P. (2016). *Report of the Inquiry into the 2015 British general election opinion polls*. London: Market Research Society and British Polling Council. Online; accessed at https://eprints.ncrm.ac.uk/id/eprint/3789/.

Surzhko-Harned, L. and Zahuranec, A. J. (2017). Framing the revolution: the role of social media in ukraine's euromaidan movement. *Nationalities Papers*, 45(5):758–779. doi: https://doi.org/10.1080/00905992.2017.1289162.

Sweney, M. (2021). The great british broadcasting shake-up - all you need to know. *The Guardian*. Online; accessed 3rd April 2024 at https://www.theguardian.com/media/2021/jun/23/the-great-british-broadcasting-shake-up-all-you-need-to-know.

Swinford, S. (2015). Nigel farage says douglas carswell is in 'westminster bubble' in hiv row. *The Telegraph*. Online; accessed 27/11/2024 at https://www.telegraph.co.uk/news/politics/nigel-farage/11643343/Nigel-Farage-says-Douglas-Carswell-is-in-Westminster-bubble-in-HIV-row.html.

Taylor, J. and Pagliari, C. (2018). Mining social media data: How are research sponsors and researchers addressing the ethical challenges? *Research Ethics*, 14(2):1–39. doi: https://doi.org/10.1177/1747016117738559.

Temporão, M., Kerckhove, C. V., van der Linden, C., Dufresne, Y., and Hendrickx, J. M. (2018). Ideological scaling of social media users: a dynamic lexicon approach. *Political Analysis*, 26(4):457–473.

Theocharis, Y., Lowe, W., Van Deth, J. W., and García-Albacete, G. (2015). Using twitter to mobilize protest action: online mobilization patterns and action repertoires in the occupy wall street, indignados, and aganaktismenoi movements. *Information, Communication & Society*, 18(2):202–220. doi: https://doi.org/10.1080/1369118X.2014.948035.

Thomas, M., Pang, B., and Lee, L. (2006). Get out the vote: Determining support or opposition from congressional floor-debate transcripts. *arXiv preprint cs/0607062*. doi: https://doi.org/10.48550/arXiv.cs/0607062.

Toledano, C. A. (2013). Web 2.0: the origin of the word that has changed the way we understand public relations. In *Representing PR: Images, identities and innovations. Barcelona international PR conference, Barcelona, Spain*. Accessed at: https://www.academia.edu/download/46906637/CristinaAcedToledano$_v$6.pdf.

Tong, J. and Zuo, L. (2018). Elite domination in the asymmetrical twitter space. In *Tweeting the Environment# Brexit*, pages 69–83. Emerald Publishing Limited. doi: https://doi.org/10.1108/978-1-78756-499-220181008.

Townsend, L. and Wallace, C. (2016). Social media research: A guide to ethics. Technical report, University of Aberdeen. Accessed at: https://www.gla.ac.uk/media/Media$_4$87729$_s$$mxx.pdf$.

Treen, K. M. d., Williams, H. T., and O'Neill, S. J. (2020). Online misinformation about climate change. *Wiley Interdisciplinary Reviews: Climate Change*, 11(5):e665. doi: https://doi.org/10.1002/wcc.665.

Tremayne, M. (2016). Anatomy of protest in the digital era: A network analysis of twitter and occupy wall street. In *Social Networks and Social Movements*, pages 110–126. Routledge.

Tromble, R. (2018). Thanks for (actually) responding! how citizen demand shapes politicians' interactive practices on twitter. *New media & society*, 20(2):676–697. doi: https://doi.org/10.1177/1461444816669158.

Tumasjan, A., Sprenger, T., Sandner, P., and Welpe, I. (2010). Predicting elections with twitter: What 140 characters reveal about political sentiment. In *Proceedings of the international AAAI conference on web and social media*, volume 4, pages 178–185. doi: https://doi.org/10.1609/icwsm.v4i1.14009.

Urman, A. (2020). Context matters: political polarization on twitter from a comparative perspective. *Media, culture & society*, 42(6):857–879. doi: https://doi.org/10.1177/0163443719876541.

Urman, A. and Katz, S. (2022). What they do in the shadows: examining the far-right networks on telegram. *Information, communication & society*, 25(7):904–923. doi: https://doi.org/10.1080/1369118X.2020.1803946.

Van Dalen, A. (2012). Structural bias in cross-national perspective: How political systems and journalism cultures influence government dominance in the news. *The International Journal of Press/Politics*, 17(1):32–55. doi: https://doi.org/10.1177/1940161211411087.

Van Dijck, J., Poell, T., and De Waal, M. (2018). *The platform society: Public values in a connective world*. Oxford: Oxford university press.

van Dijk, T. A. (2006). *Politics, ideology, and discourse*. Elsevier.

Van Vliet, L., Törnberg, P., and Uitermark, J. (2020). The twitter parliamentarian database: Analyzing twitter politics across 26 countries. *PLoS one*, 15(9):e0237073. doi: https://doi.org/10.1371/journal.pone.0237073.

Vishnuprasad, P. S., Nogara, G., Cardoso, F., Cresci, S., Giordano, S., and Luceri, L. (2024). Tracking fringe and coordinated activity on twitter leading up to the us capitol attack. In *Proceedings of the international AAAI conference on web and social media*, volume 18, pages 1557–1570. doi: https://doi.org/10.1609/icwsm.v18i1.31409.

Wahl-Jorgensen, K., Berry, M., Garcia-Blanco, I., Bennett, L., and Cable, J. (2017). Rethinking balance and impartiality in journalism? how the bbc attempted and failed to change the paradigm. *Journalism*, 18(7):781–800. doi: https://doi.org/10.1177/1464884916648094.

Wakefield, K. (2016). Using fan passion to predict attendance, media consumption, and social media behaviors. *Journal of Sport Management*, 30(3):229–247. doi: https://doi.org/10.1123/jsm.2015-0039.

Walker, M. and Matsa, K. E. (2021). News consumption across social media in 2021. *Pew Research Center: Report*. Accessed at: https://www.jstor.org/stable/resrep63516.

Walsh, M. (2024). Bbc question time: analysis of guests over nine years suggests an overuse of rightwing voices. *The Conversation*. Online; accessed 31st Jan 2025 at https://theconversation.com/bbc-question-time-analysis-of-guests-over-nine-years-suggests-an-overuse-of-rightwing-voices-232315.

Walter, S., Brüggemann, M., and Engesser, S. (2018). Echo chambers of denial: Explaining user comments on climate change. *Environmental Communication*, 12(2):204–217. doi: https://doi.org/10.1080/17524032.2017.1394893.

Ward, B. (2012). The bbc is sacrificing objectivity for impartiality in its coverage of climate change. *LSE Blog*. Online; accessed 3rd April 2024 at https://blogs.lse.ac.uk/politicsandpolicy/bbc-climate-change-ward/.

Ward, S. and Lusoli, W. (2005). 'from weird to wired': Mps, the internet and representative politics in the uk. *The Journal of Legislative Studies*, 11(1):57–81. doi: https://doi.org/10.1080/13572330500158276.

Ward, S. and McLoughlin, L. (2020). Turds, traitors and tossers: the abuse of uk mps via twitter. *The Journal of Legislative Studies*, 26(1):47–73. doi: https://doi.org/10.1080/13572334.2020.1730502.

Waterson, J. (2020). Uk's news channels shouldn't have to be impartial, says ex-bbc chair. *The Guardian*. Online; accessed 3rd April 2024 at https://www.theguardian.com/media/2020/nov/11/uk-news-channels-should-not-have-to-be-impartial.

Waterston, J. (2022). Emily maitlis says 'active tory party agent' shaping bbc news output. *The Guardian*. Online; accessed 3rd April 2024 at https://www.theguardian.com/media/2022/aug/24/emily-maitlis-says-active-tory-party-agent-shaping-bbc-news-output.

Waterston, J. (2023a). Ofcom to investigate gb news and talktv over impartiality rules. *The Guardian*. Online; accessed 3rd April 2024 at https://www.theguardian.com/media/2023/jul/03/ofcom-to-investigate-gb-news-and-talktv-over-impartiality-rules.

Waterston, J. (2023b). Piers morgan says 'due impartiality' rules for british tv out of date. *The Guardian*. Online; accessed 3rd April 2024 at https://www.theguardian.com/media/2023/sep/21/piers-morgan-tv-news-impartiality-rules-out-of-date.

Weaver, I. S., Williams, H., Cioroianu, I., Williams, M., Coan, T., and Banducci, S. (2018). Dynamic social media affiliations among uk politicians. *Social Networks*, 54:132–144. doi: https://doi.org/10.1016/j.socnet.2018.01.008.

Webster, J. G. and Ksiazek, T. B. (2012). The dynamics of audience fragmentation: Public attention in an age of digital media. *Journal of communication*, 62(1):39–56. doi: https://doi.org/10.1111/j.1460-2466.2011.01616.x.

Weeks, B. E., Lane, D. S., Kim, D. H., Lee, S. S., and Kwak, N. (2017). Incidental exposure, selective exposure, and political information sharing: Integrating online exposure patterns and expression on social media. *Journal of computer-mediated communication*, 22(6):363–379. doi: https://doi.org/10.1111/jcc4.12199.

Weisberg, H. F. (1978). Evaluating theories of congressional roll-call voting. *American Journal of Political Science*, pages 554–577. doi: https://doi.org/10.2307/2110461.

Weissberg, R. (1978). Collective vs. dyadic representation in congress. *American Political Science Review*, 72(2):535–547. doi: https://doi.org/10.2307/1954109.

Wellings, T., Hänggli Fricker, R., Pournaras, E., and Fu, W. (2024). Brexit on twitter: Unraveling the dynamics of polarization over time. In *Proceedings of the 25th Annual International Conference on Digital Government Research*, pages 855–866. doi: https://doi.org/10.1145/3657054.3657269.

Whitefield, S., Vachudova, M. A., Steenbergen, M. R., Rohrschneider, R., Marks, G., Loveless, M. P., and Hooghe, L. (2007). Do expert surveys produce consistent estimates of party stances on european integration? comparing expert surveys in the difficult case of central and eastern europe. *Electoral Studies*, 26(1):50–61. doi: https://doi.org/10.1016/j.electstud.2006.04.006.

Wilke, J. and Reinemann, C. (2006). Die normalisierung des sonderfalls? die wahlkampfberichterstattung der presse 2005 im langzeitvergleich. *Die Massenmedien im Wahlkampf: Die Bundestagswahl 2005*, pages 306–337. doi: https://doi.org/10.1007/978-3-531-90383-511.

Williams, R. (2015). David cameron suggests twitter users not 'decent, sensible or reasonable'. *The Independent*. Online; accessed April 3rd 2024 at https://www.independent.co.uk/news/uk/politics/david-cameron-suggests-twitter-users-not-decent-sensible-or-reasonable-a6684521.html.

Wojcieszak, M., Casas, A., Yu, X., Nagler, J., and Tucker, J. A. (2021). Echo chambers revisited: The (overwhelming) sharing of in-group politicians, pundits and media on twitter. *OSF Preprints*. doi: https://doi.org/10.31219/osf.io/xwc79.

Wojcieszak, M., Casas, A., Yu, X., Nagler, J., and Tucker, J. A. (2022). Most users do not follow political elites on twitter; those who do show overwhelming preferences for ideological congruity. *Science Advances*, 8(39). doi: https://doi.org/10.1126/sciadv.abn9418.

Wojcik, S. and Hughes, A. (2019). Sizing up twitter users. *Pew Research Center*. Online; accessed 28th Oct 2024 at https://www.pewresearch.org/internet/2019/04/24/sizing-up-twitter-users/.

Wolfsfeld, G., Segev, E., and Sheafer, T. (2013). Social media and the arab spring: Politics comes first. *The International Journal of Press/Politics*, 18(2):115–137. doi: https://doi.org/10.1177/1940161212471716.

Woolley, S. C. and Howard, P. N. (2018). *Computational propaganda: Political parties, politicians, and political manipulation on social media*. Oxford: Oxford University Press.

Yang, K.-C., Hui, P.-M., and Menczer, F. (2022). How twitter data sampling biases us voter behavior characterizations. *PeerJ Computer Science*, 8:e1025. doi: https://doi.org/10.7717/peerj-cs.1025.

Yeung, E. S. and Quek, K. (2024). Self-reported political ideology. *Political Science Research and Methods*, pages 1–22. doi: https://doi.org/10.1017/psrm.2024.2.

Zaller, J. (1992). *The Nature and Origins of Mass Opinion*. New York: Cambridge University Press.

Zeldes, G. A., Fico, F., Carpenter, S., and Diddi, A. (2008). Partisan balance and bias in network coverage of the 2000 and 2004 presidential elections. *Journal of broadcasting & electronic media*, 52(4):563–580. doi: https://doi.org/10.1080/08838150802437354.

Zhang, A. X. and Counts, S. (2015). Modeling ideology and predicting policy change with social media: Case of same-sex marriage. *In Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, pages 2603–2612. doi: https://doi.org/10.1145/2702123.2702193.

Zubiaga, A. (2018). A longitudinal assessment of the persistence of twitter datasets. *Journal of the Association for Information Science and Technology*, 69(8):974–984. doi: https://doi.org/10.1017/pan.2024.7.