The Bandwagon Effect

How Popularity Information Affects Electoral Expectations and Voting Behaviour

Matthew Barnfield

A dissertation submitted in partial fulfillment of the requirements of the degree of Doctor of Philosophy

September 2021

Statement of Originality

I, Matthew Barnfield, confirm that the research included within this thesis is my own work or that where it has been carried out in collaboration with, or supported by others, that this is duly acknowledged below and my contribution indicated. Previously published material is also acknowledged below. I attest that I have exercised reasonable care to ensure that the work is original, and does not to the best of my knowledge break any UK law, infringe any third party's copyright or other Intellectual Property Right, or contain any confidential material. I confirm that this thesis has not been previously submitted for the award of a degree by this or any other university. The copyright of this thesis rests with the author and no quotation from it or information derived from it may be published without the prior written consent of the author.

Signature: **Date:** 30th September 2021

Details of collaboration and publications:

Barnfield, M. (2020) Think twice before Jumping on the Bandwagon: Clarifying Concepts in Research on the Bandwagon Effect. *Political Studies Review*, 18(4), pp.552-574.

Abstract

This thesis studies whether and how information about the popularity of political candidates and parties, such as the results of opinion polls, influences what people think is going to happen at elections and how they intend to vote. It focuses on the 'bandwagon effect' – the idea that people will vote for candidates or parties because they are popular. This effect is consequential for debates about the regulation of opinion polls, as well as campaign strategy and the discipline of election forecasting. The thesis comes at this question theoretically, by rigorously defining and classifying the bandwagon effect and sketching a causal model of how it comes about. It also addresses methodological points related to the bandwagon effect, making the case for major improvements in how it is measured. Finally, it does significant empirical work, providing substantial new evidence on how popularity information affects people's attitudes and behaviours. Specifically, the key findings of this research suggest that voters who saw polls suggesting Biden was in the lead in their state in the 2020 US presidential election became less likely to vote for him, that people think parties that are growing in the polls are more likely to win elections even when they are not in the lead, and that voters in the 2017 UK general election who saw polls suggesting the Labour Party's vote share was

growing became more likely to vote Labour. In order to make these findings, I contribute new theoretical arguments and methodological approaches that show how to rigorously study the bandwagon effect. I conclude by clearly stating the scientific and political implications of my research and setting out fruitful avenues available to researchers for future study.

Acknowledgements

This thesis exists because of other people. Philip Cowley and Javier Sajuria are the reason it is any good. Peter Allen and Aurelien Mondon are the reason I even considered doing it in the first place. Tracey and Jim Barnfield are the reason I was in a position to make that choice. I am deeply grateful to all of these people.

I would also like to thank Rainbow Murray and Ksenia Northmore-Ball for their feedback, Jack Bailey and Sabrina Stöckli for methodological pointers, Tim Bale for the opportunities he has given me, Anthony Wells for his help with surveys, and Jason Reifler for allowing me the time to get this thesis finished. My fellow PhD students have all helped me in many ways, but especially Katy Brown and Lateesha Osbourne back at Bath, and Mirko Palestrino, Alex Stoffel, Ida Birkvad, Dan Gover, and Salomé Ietter at Queen Mary.

Finally, thank you to my long-suffering housemates James, Jacob and Will, my loving siblings Charley, Jamie, Harry and Samantha, and many other friends, for making these three great years.

This work was supported by the Economic and Social Research Council, grant number ES/P000703/1.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Brian: Look. You've got it all wrong. You don't need to follow me. You don't need to follow anybody! You've got to think for yourselves. You're all individuals!

Followers: Yes, we're all individuals!

Brian: You're all different!

Followers: Yes, we are all different!

Dennis: I'm not.

Arthur: Shhhh.

Followers: Shh. Shhhh. Shhh.

Monty Python's The Life of Brian

In the 1979 film *Life of Brian*, the eponymous hero acquires a large, unwanted following overnight. Following a badly delivered, unfinished, misunderstood speech about jobless lilies and birds that he only makes as a way of evading his Roman pursuers, Brian's (non-existent) cause comes to inspire the masses. Espoused by some and followed soon after by those who latch onto its growing popularity, the message that Brian is the ultimate saviour causes his newfound devotees to take leave of their senses, in great number and with great enthusiasm. The very reason Brian was even thought to be worthy of such devotion – which, he insists, he never was – is lost as his popularity becomes simply self-perpetuating.

Brian's predicament is a caricature of what is referred to in political science and social psychology as the *bandwagon effect*. Broadly defined, this refers to a situation in which people do something because lots of other people are doing that thing. In political science, and therefore this thesis, it is mostly thought of as what happens when people vote for candidates or parties because many, or increasingly many, others are doing so. This thesis tackles the key questions of what the bandwagon effect is, how to study it, and whether there is evidence of it in the real political world.

Concretely, the thesis studies how 'popularity information' about support for parties or candidates – conveyed most prominently by the polls – influences what people think is going to happen at elections and how they intend to vote. It comes at the question theoretically, by rigorously defining and classifying the bandwagon effect and sketching a causal model of how it comes about. This discussion draws on a comprehensive review of existing literature on the bandwagon effect and on insights from political psychology. It also reflects the emphasis this thesis puts on causal inference; this is a thesis about questions of causality. On this basis, the thesis also addresses methodological points about the bandwagon effect, making the case for major improvements in how it is measured as a causal effect. In doing so, it makes innovations in both experimental and observational methods. Finally, it does significant empirical work, providing three main new pieces of evidence on how popularity information affects people's attitudes and behaviours. First, that voters who saw polls suggesting Biden was in the lead in their state in the 2020 U.S. presidential election became less likely to vote for him. Second, that people think parties are more likely to win elections when they are growing in the polls, even when those parties are not in the lead. Third, that voters in the 2017 UK general election who received information suggesting the Labour Party's vote share was growing became more likely to vote Labour.

How This Thesis Thinks About the Bandwagon Effect

While for Brian the bandwagon effect is an inconvenience, it would be a dream scenario for a candidate vying for election. One ill-prepared speech given to a small audience sets the bandwagon rolling apace, and within a matter of hours a huge following has accrued as more and more people jump on board, for no apparent reason. Politics does not really look anything like this, but this has not stopped social scientists and political commentators from thinking that the so-called 'social influence' that we do observe in the real world is grounded in the kind of behaviour Brian suffers, and deriding the electorate for this. As Diana Mutz explains,

> There is a strong derogatory tone in most research on the influence of perceptions of mass opinion. Responding to perceptions of mass opinion is usually implied to be to the detriment of democracy; it is equated with sheeplike behaviour and blindly following a crowd (1998, 206).

Seeing social influence this way comes from thinking about it in terms of a desire to 'take a position that will lead others to have positive feelings toward a person' (Deutsch and Gerard 1955). Essentially, the thought is that the opinions of others tell people what they ought to think in order to be like those other people, because being like those other people will lead to things that feel good. This is called *normative* social influence. This view sits well with the historical thinking that voters are incapable of making choices that are guided by more relevant information. As Bowler and Donovan put it, many political scientists

> Have not been sympathetic to the idea that voters have the interest, sophistication, or information required to make policy-oriented decisions in highly visible, partisan elections (1998, 21).

In contrast to this is *informational* social influence, which involves a tendency to 'accept information obtained from another as evidence about reality' (Deutsch and Gerard 1955; Mutz 1998, 271). The idea here is that the opinions of others provide new information that people process like any other information, and this might end up with them adopting the same stance. This sits better with the view, now more commonly held amongst political scientists, that 'voters are able to reason and therefore are able to decide which side of an issue they support based on readily available information' (Bowler and Donovan 1998, 21).

In this thesis, I take an informational approach. There are three reasons for this choice. First, because the evidence for the informational perspective is stronger (Bowler and Donovan 1998; Key 1966; Mutz 1998; Page and Shapiro 1992). Second, because the informational approach provides clear ways to link the bandwagon effect to existing knowledge of how voters think (Lau and Redlawsk 2006; Lodge and Taber 2013). Chapter 2 explains both of these justifications further, while noting that although voters make choices based on information, they are nonetheless cognitively limited in *how* they do this.

But, third, I also take an informational approach because it makes sense from a historical perspective, when thinking about how the concepts at the heart of this thesis have developed over time. I briefly address this claim here, early, because it is useful not only in broadly justifying what is to come throughout the rest of this thesis (and this chapter), but also in situating this research in a longer-term view of politics, public opinion, and social psychology.

In Brian's day, polls did not exist. Instead, as shown in earlier scenes, 'oratory' and conversation – for example, in marketplaces – was the dominant mode of expressing opinion (Bauer 1930, 671). It was all about talking, meeting with others, listening to speeches, all together, as a *crowd*. Exposure to public opinion necessarily involved discussion, and being an active part of the creation and spread of that opinion. So-called 'crowd psychology' therefore mattered (Glynn et al. 2004). People in crowds are close to one another, and have 'shared emotions' (Price 1992). Gustave Le Bon (1948; see also Freud 1922) famously argued that crowds are 'suggestible' and ideas within them become 'contagious.' This is exactly what we see Brian's crowd doing. Once some of them are, quite easily, convinced he is somehow special (suggestibility), this opinion spreads rapidly through the crowd (contagion). This is *normative* social influence. Members of the crowd want to be part of what that crowd is experiencing together, and gain its approval – not become pariahs with different views. When the dominant way in which public opinion is created and expressed involves crowds, this means that to the extent that people come to hold the political beliefs that lots of other people do, there is highly likely to be a normative aspect to why this happens, which makes for irrational opinion formation and decision-making.

The modern context considered in research on the bandwagon effect looks nothing like this. In terms of the expression and diffusion of popular sentiment, the 'crowd' has largely been replaced by a 'public' – and 'crowd psychology' by 'public opinion' (Eisenstein 1979). Herbst (2021) argues that the 1930s represent a key turning point in this process, as 'the diffusion of radio, the devastating impact of the economic collapse on so many people, the appearance of professional pollsters' all spawned the idea that 'there was a public – whose opinions mattered.' Elsewhere, Herbst (1993, 51) also explains that 'in a public, compared to a crowd, individuals are connected through the communication of ideas and not by physical proximity.' The opinions of this public are measured through 'aggregation'; elections and polls count up opinions and take their sum as popular sentiment. The individuals constituting this public opinion are atomised and anonymous. There is no need for them to be actively involved in the creation of public opinion in order to learn about it. Even when their opinion is taken and added to the pile, this is essentially done privately, by filling out a web survey or casting a secret ballot. In this context, public opinion becomes a form of information, presented to individuals as a sum of expressions from mostly unknown others – or 'numbered voices' (Herbst 1993). This context is far less conducive to normative social influence – why care about the approval of others, most of whom will never meet you, nor be aware of your existence, nor ever know what your opinion is? To the extent that social influence takes place in this context, it is much more likely to be *informational*.[1]

This distinction matters enormously for research on the bandwagon effect, because whether the effect is of a normative or informational form significantly changes what its implications are. Mutz explains that

> While social influence of either variety has the capacity to corrupt individual decision making and constrain individual judgment, the nature of the *processes* underlying informational social influence makes them infinitely preferable. Rather than adopting others' views as a means of obtaining approval, people conforming on an informational basis are doing so for largely rational reasons (1998, 290).

When people undergo normative social influence, they are not reflecting rationally in making their political decisions. They are going with the crowd for the sake of it, or because it feels good somehow. When people undergo informational social influence, they are just using a particular type of information – which I will come to call 'popularity information' – about the levels of popular support for parties or candidates, as part of their reflection when making their political decisions. The latter is much more consistent with democratic ideals about how voters should behave, because it is fundamentally about making good decisions. The purpose of informational social influence is to improve decision-making. It represents rational behaviour, in which the decision-making process is not derailed by a desire to

[1] This does not necessarily apply to situations where people are influenced in close social settings, such as in protests or rallies, or even in family discussions, but these are not contexts that research on the bandwagon effect has taken any interest in. The bandwagon effect is instead about what Mutz (1998) calls processes of 'impersonal influence.'

conform to the crowd, but simply incorporates popularity information.

I leave a discussion of precisely what this decision-making process is for Chapter 2. The key point is that the bandwagon effect is not necessarily people making political decisions in apolitical or irrational ways, as is often thought (e.g. Bartels 1987, 19). Bandwagon effects are the result of people using a specific type of information to arrive at a vote decision. Like Brian, the observer of bandwagon effects who responds by imploring people to be individuals, to be different, and to think for themselves, does so in vain. But unlike Brian, it is not because the members of the public are already too blinded by their irrational devotion to heed such pleas that this response is so futile; it is because *they already are thinking for themselves*.

Why the Bandwagon Effect Matters

Taking the informational approach, though, means that it is less clear why research on the bandwagon effect, and thereby this thesis, matters. If bandwagon effects represented mass normative social influence, then given the prevailing belief that such irrational political behaviour is bad for politics and society (e.g. Achen and Bartels 2015), studying the effect would clearly be important as a way of establishing the severity of the problem. The informational approach suggests that such concerns might not be justified, so the importance of researching the bandwagon effect might have to come from elsewhere – but where? The answer to this is that the bandwagon effect is important because it has a meaningful impact on real-world debates about regulating opinion polls, on political and campaign strategy, and on electoral forecasting.

Bandwagon Effects and Regulation of Opinion Polling

There is a long-standing debate, in many democracies, about whether the publication of vote intention polls should be banned, restricted, or otherwise subject to tighter regulation. A recent report on polling regulations, spanning countries from every continent, found that polls are restricted before elections in around half of the countries covered, and that it is usually the government itself that enforces these restrictions (Frankovic, Johnson, and Stavrakantonaki 2018). Pre-election polling 'embargoes' or 'blackouts' ranged from one day (Norway, Portugal, Poland, Singapore) to over two weeks (Argentina, Greece, South Korea) – or, in the extreme case of Honduras, up to 45 days. By and large, policy-makers and regulating bodies implement restrictions of this sort because they believe polls influence the vote (Donsbach and Hartung 2008).

Likewise, calls for the implementation of restrictions tend to bring up the idea that polls influence voters. Danish political actors made such calls following the leading broadcasting service's publication of an exit poll, before polling stations closed, that 'substantially underestimated the Social Democrats' share of the vote' (Dahlgaard et al. 2017, 2). Social Democrat spokesperson Mogens Jensen claimed that there was 'no doubt that these polls can help to move and influence voters… either people react by not wanting to be on a losing team, or it has the opposite effect… voters must be able to decide without this massive media influence' (Albrechsten 2013). Importantly, the Danish Folketing were 'unanimous' in supporting the proposal to prevent media outlets publishing poll results from the time the polling stations open until they close (Albrechsten 2013). A similar situation sparked a similar reaction in Germany, in 2009 (Kirschbaum 2009). In Sweden, the Centre Party has frequently proposed legislation prohibiting the publication of poll results immediately before elections (Dahlgaard et al. 2017, 2), based on the idea that they might influence voters (Hernadi 2010).

Though not quite such an all-out call for greater regulation, a 2018 report by the United Kingdom House of Lords Select Committee on Political Polling and Digital Media expressed a ‘central concern’ that, given polls have been somewhat inaccurate in recent elections, ‘there is a significant risk that future elections will be affected by misleading information, potentially distorting the democratic process’ (2018, 3). The report notes that ‘there are a number of theories about how polls influence voters. These include the “bandwagon” effect, where it is suggested that information from polls can influence people to alter their opinion to accord with the majority view’ (2018, 72).

Indeed, calls for increased regulation of polls on the basis of their purported influence date back much further than these recent examples might suggest. As early as the 1930s, Representative from Oregon Walter M. Pierce railed against the use of early forms of polling, writing in a letter that

> It is my belief that much misinformation is carried to the public in so-called straw ballots. My objection is that they create a belief in the public that a certain issue or candidate is going to win. I am firmly convinced that the biggest single argument that can be made in a campaign is “you are going to win” (quoted in Allard 1941, 208).

Based on these concerns, Pierce made multiple attempts to curb straw polls, introducing bills to restrict and investigate their practice. He wrote about his thoughts on straw polls, and his travails attempting to legislate against them, in an article titled ‘Climbing on the Bandwagon’ (Pierce 1940).

These debates are also not new in the UK. In the Bermondsey by-election of 1982, many believed the polls played a large part in making people switch from the Labour candidate to the Liberal candidate in the week leading up to the election. As a result, Douglas Hoyle MP sponsored a bill to ban polling in pre-election periods (Marsh 1985, 72).

To some extent, such debates hinge on the normative/informational distinction above. Marsh explains that

> Hoyle and his supporters argued that poll results have a normative effect, placing pressure on people directly. The pundits… while agreeing that bandwagon effects occurred, viewed them as having an informational effect, allowing people to revise their view of majority opinion and so cast their vote more effectively (1985, 72).

Similarly, in Canada, the 2000 Canada Elections Act abolished previous legislation prohibiting the publication of poll results during the final three days of an election campaign, following a Supreme Court decision in the ‘Thomson Newspapers case.’ Bale (2002, 16) explains that ‘the majority judgement, arguing that voters should be credited with both maturity and intelligence, held that the ban was a serious invasion of freedom of expression.’ This could be read as a rejection of the idea of problematic normative social influence. Furthermore, in order to allay concerns about the influence of polls, academic contributors to the House of Lords report mentioned above explicitly noted that any influence of polls is likely to be informational (Select Committee on Political Polling and Digital Media 2018, 74).

Yet, whether the effect is normative or informational does not fully resolve the debate. This is because the polls can be *wrong*. Note that a large part of the outrage in the Danish case above comes down to the inaccuracy of the leaked poll. In the UK case, the House of Lords report is framed in terms of a concern about inaccurate polling. Walter M. Pierce saw straw polls as problematic because of the 'misinformation' they carried. Even if the influence polls might have on voters is informational, and grounded in rational responses, it is problematic if voters are basing their decisions on inaccurate information. Though there is little reason to think that polls are facing any crisis of inaccuracy (Jennings and Wlezien 2018; Prosser and Mellon 2018), uncertainty is built into their methodology. They can be wrong – in some elections more than others (Tudor and Wall 2021). As well as needing to establish how and why polls may fail to measure electoral support accurately, it is also important to know to what extent they affect voters. Indeed, the possibility that they are inaccurate makes the latter question all the more pressing. This is central to the case explored in Chapter 3.

Evidence suggesting polls induce major bandwagon effects in elections could be used to justify increased restrictions on their publication, or even an outright ban during certain periods. Given voters' appetite for polling as a substantial part of their information diet, such regulation could have a significant impact on what voters learn and how they think at election time (Iyengar, Norpoth, and Hahn 2004). It is therefore essential that the bandwagon effect is studied rigorously and with appropriate methods and measures. This thesis makes major advances on these points.

Bandwagon Effects and Political Strategy

Bandwagon effects also influence political strategy. As bandwagon-sceptic pollster George Gallup and Canadian diplomat Saul Rae lamented the best part of a century ago,

> Some politicians have spent so much effort in trying to swing doubtful voters to their side by prophesying ultimate victory that they have convinced themselves of the effectiveness under all conditions of appealing to the "bandwagon" psychology of the average voter. This is the background of the indictment that pre-election polls tend to handicap the "losing" side by influencing doubtful voters to vote for the "winning" candidate (1940, 244).

Politicians and political parties try to make themselves look like they have a good chance of winning elections, based on a belief in the bandwagon effect. Examples of this abound in political discourse. Donald Trump insisted that he would win the 2020 U.S. presidential election and 'four more years' as president – and perhaps even 'another four years' (Schwartz 2020). Marine Le Pen stated clearly in televised interviews that she would win the 2017 French presidential election (franceinfo 2017). In the 2019 Conservative Party leadership contest in the UK, various candidates went to great lengths to make themselves seem popular and viable, all while one candidate clearly dominated. Both Rory Stewart and Dominic Raab saw fit to highlight how the polls could be interpreted as showing them to be the only challenger to eventual landslide winner Boris Johnson. Raab (see Figure 1.1) put it that he was 'the main challenger to Boris with our members' so they should be given 'a real choice, and the contest they want.' Stewart, similarly, claimed that 'in message and in polling' it was '[him] or PM Boris' (see Figure 1.2). None of these candidates actually won the elections in question, so this is clearly not a matter of them just stating the obvious. It is likely, instead, that they saw it as a persuasive claim.

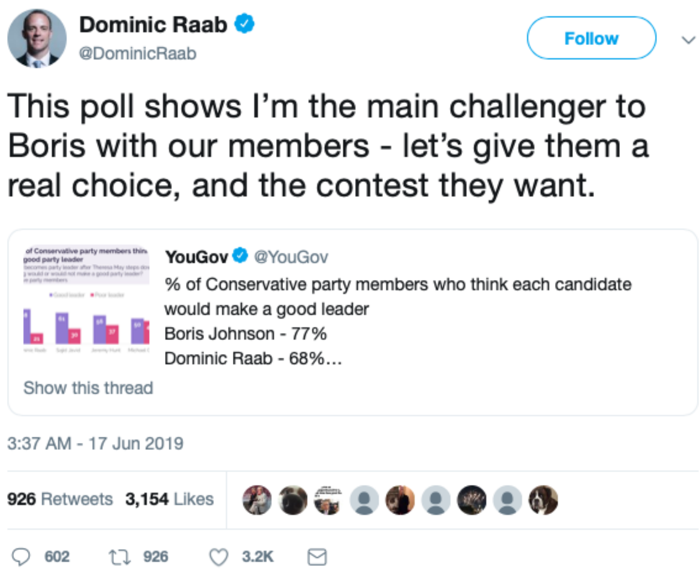


Figure 1.1: Tweet by @DominicRaab.

Research on the bandwagon effect matters because it serves to show whether or not this belief in the value of being seen as a winner – or, at least, a serious contender – has any basis in reality. Should research start to demonstrate clearly that there is no such thing as a bandwagon effect, for example, then this could permeate through into parties' campaign strategies. Again, this means that the bandwagon effect needs to be taken seriously and treated rigorously in research. Attention to the contributions of this thesis will help political scientists to do this well.

Figure 1.2: Tweet by @RoryStewartUK.

Bandwagon Effects and Election Forecasting

Finally, the bandwagon effect has implications for electoral forecasting. Not only do the predictions that forecasters generate have the potential to *produce* bandwagon effects, but (in doing so) they might also become invalid as a result of bandwagon effects. For example, if a forecast predicts that a candidate will win the election by three percentage points, this forecast might falsify itself by producing a bandwagon effect that leads the candidate to win by six points. In this case the main outcome – who wins – would still be called correctly, but this will not hold true in every case. Consider that in multiparty systems a bandwagon effect could change how many seats each party gains, and therefore determine its position in terms of coalition formation, leading to completely different governments. It is also possible for a second-placed candidate to benefit from a bandwagon effect

(Lanoue and Bowler 1998), which could tip her over the edge into first place. This is not to mention the possibility of 'dynamic bandwagon effects,' introduced in depth in Chapter 2 and demonstrated empirically in Chapter 5, which could benefit any party regardless of its current ranking in the polls, and thereby change that ranking. In sum, bandwagon effects imply that electoral forecasting is an attempt to predict an outcome it plays a part in producing (Henshel 1982; Simon 1954).

Forecasters may therefore do a better job of predicting election results if they are aware of the extent to which, and circumstances under which, bandwagon effects take place. Simon (1954, 248) claims that if bandwagon effects are large, it may be that the only way for election forecasts to be accurate is if forecasters privately adjust their results to account for the discrepancy. This would represent a major change in forecasting methodology. It could also have effects, downstream, on election betting markets and campaign donations.

Increasingly, political scientists propose that it could be more effective to forecast elections by asking the electorate who they think is going to win and aggregating these predictions, rather than just polling vote intention (e.g. Graefe 2014; Murr 2011; 2016; Murr, Stegmaier, and Lewis-Beck 2019b). The bandwagon effect also matters to such 'citizen forecasting,' which could begin to become more prominent beyond academia if it continues to prove an effective way of predicting election outcomes (though see Ganser and Riordan 2015). One important reason for this is that people strongly believe in the bandwagon effect on other people, even though they do not think they themselves are prone to it (Chung, Heo, and Moon 2018; Lang and Lang 1984). This implies that the individual predictions they give might include the sorts of 'private adjustment' recommended above by Simon (1954, 248); they might believe too strongly that a leading candidate is going to win, because they believe that leading candidate will benefit from a bandwagon effect amongst other voters. Beyond this, as I will draw out in Chapter 2, research on the bandwagon effect and knowledge about voters' electoral expectations are closely related and interlinked. In attempting to address the bandwagon effect, this thesis – in Chapters 4 and 5 – therefore ends up shedding new light on these expectations, which form the basis of citizen forecasts.

Bandwagons in the Real World

Beyond these specific real-world implications of the bandwagon effect, there are also more general examples of elections where it appears that the popularity of a candidate or party has influenced the course of the contest. The fact that these cases exist provides further reason to study the bandwagon effect, given that it is one explanation of how such influence could operate. I discuss a few of these examples here. Though not all of the examples strictly conform to a bandwagon effect as I will come to conceptualise it in Chapter 2, they all suggest that candidate or party's popularity plays a central role not only as a measure of their success, but also in altering these fortunes.

A 'Primary' Example

Arguably, presidential nomination contests in the USA are ripe for bandwagon effects (Abramowitz 1987, 50). A clear example is the case of Gary Hart, in 1984. Presidential nomination contests are sequential. Rather than voting all at the same time, voters in different states vote days, weeks, or even months apart. This means that voters in states that come later, such as California, have information about how likely candidates are to win and who is generally popular that those in earlier states, such as Iowa, do not have. Recognising this, Gary Hart made efforts to canvass early in New Hampshire. The so-called ‘Granite State’ is the second state to vote, but the first that does so through a ‘primary’ – a regular election in which individuals cast votes privately in polling places – rather than ‘caucuses’ – more communal decisions in which local people gather to discuss the candidates and cast a vote as a group. This early campaigning appeared to make no real impact on how Hart was polling nationally. He had still ‘barely registered in the polls’ (Bartels 1987, 2) and trailed far behind the leader Walter Mondale, prior to the Iowa caucuses, which are the first stage in the contest.

This changed dramatically when Hart first secured a surprisingly large share of the vote in Iowa, then won the New Hampshire primary. Following this, his performance in weekly voter surveys skyrocketed, ‘with Hart emerging as the front-runner and Mondale suddenly a distant second’ (Bartels 1987, 2). Super Tuesday – when multiple states all vote on the same day – reversed these fortunes again, as former Vice President Mondale won a series of states, regaining the lead in vote intentions, and eventually securing the Democratic nomination despite winning fewer contests overall than Gary Hart. Although Hart lost, his early performance appears to have given him a better chance than he could have hoped for had he not had good results in Iowa and New Hampshire, and he was not the first to benefit from this. As Bartels explains,

> Like George McGovern, Jimmy Carter, George Bush, and John Anderson before him, Hart capitalized on better-than-expected performances early in the caucus and primary season to break out of the pack, at-

> tracting remarkably broad support from an electorate that had barely recognised his name a few months earlier (1985, 804).

This image of going from zero to hero through the nomination campaign seems to have stuck with Gary Hart. He acquired a reputation as someone who was surprisingly successful, but about whom nobody knew anything. This popularity itself made journalists and the public want to find out more about him during his subsequent 1988 campaign. A New York Times Magazine article later referred to him as the ‘elusive front-runner’ (see Figure 1.3), claiming that

> He has been forced to reveal more about himself in the last four years than most politicians do in a lifetime. Yet many in politics and the press still insist that they don’t know him... In just 12 years, from 1972 to 1984, he went from being George McGovern’s campaign manager to a Presidential candidate who challenged almost everyone’s political calculations. He nearly took the Democratic nomination away from Walter F. Mondale, the man whom all the smart folks said could not be touched in his party (Dionne Jr 1987).

Unfortunately for Hart, when people did learn a lot about him, they discovered he was having an extramarital affair, a revelation that derailed his second attempt at securing the Democratic nomination in the 1988 primaries. His rise and fall inspired the 2018 film *The Front Runner*. It could be that, were it not for this scandal, Hart would have become president in 1988. Arguably, had he done so, he would have partly owed that victory to the bandwagon effect he experienced four years earlier.

This pattern, in which candidates seem to achieve broader support when they have

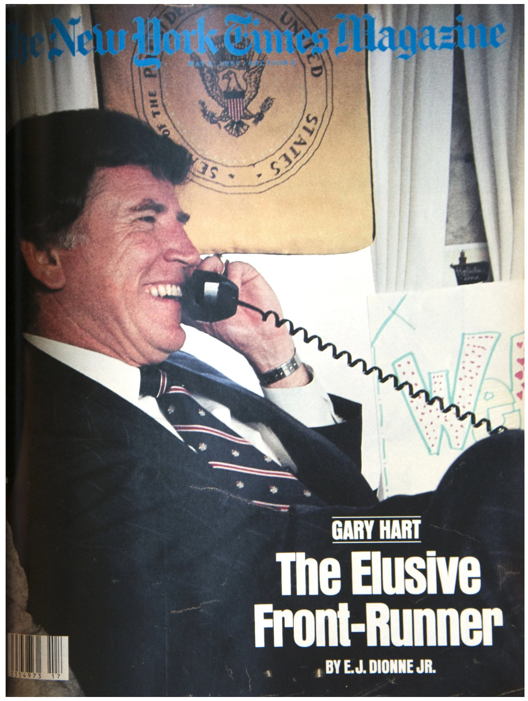


Figure 1.3: Cover of New York Times Magazine declaring Gary Hart the elusive frontrunner for the 1988 Democratic presidential nomination contest.

recently won contests, leads political scientists to think of presidential nomination campaigns as ‘dynamic’ (Abramowitz 1987; Bartels 1987, 1985; Callander 2007). The sequential nature of the election allows voters to use information about something they seem to care about: viability (Utych and Kam 2014). They learn, from recent results in other states, which candidates seem to have a chance of winning. They then become more likely to intend to vote for those candidates who have a better chance. In Chapter 2, I will argue that this is one key way a bandwagon effect can come about, in certain electoral contexts.

Yet, if bandwagon effects are so important here, why did Hart not go on to stand for president in 1984? Once he had momentum, why did it not carry him through to secure the Democratic nomination? Quite simply, because bandwagon effects do not tell the whole story. The first thing to notice here is that the 'momentum' itself does not drop out of the sky, or emerge from thin air. Hart probably did well in New Hampshire because he put so much effort into campaigning there, speaking to voters, paying the state attention. This victory, though, consolidated his popularity and served as a signal of his viability, showing he was popular and might be able to win elsewhere. This gave voters in other states pause for thought about whether, maybe, they might want to jump on the bandwagon too. But equally, these voters in other states also have other reasons to vote for certain candidates. This can interrupt the momentum and shift the dynamics of the contest. This means that, even in contexts such as presidential primaries that are seen as so conducive to bandwagon effects, it is difficult to disentangle the extent to which one took place, or observe clear evidence of it.

Oh, Jeremy Corbyn!

The case of Jeremy Corbyn and the Labour Party, in the United Kingdom, reinforces these points. Consider, first, Corbyn's election to the leadership of the party. Following Labour's disappointing performance in the 2015 general election, its leader Ed Miliband stood down, triggering an election. As candidates threw their hats in the ring, 'veteran' Corbyn put himself forward, arguing that the pool including fairly established party figures such as Yvette Cooper and Andy Burnham did not 'give Labour Party members a voice' (BBC 2015b). Pollsters expressed surprise as surveys conducted during the campaign began to suggest that Corbyn had the lead (Kellner 2015) and later extended it to over 50% of the vote (Dahlgreen

2015).

Some non-Labour supporters paid a small fee to have the right to vote in the election, precisely because Corbyn had a good level of support and they wanted to make sure he won as he seemed so clearly unelectable – with him at the helm, Labour would lose the next general election. In a sense, then, Corbyn's popularity concretely gained him more support, by bringing on board disingenuous voters. We could loosely construe this as a very unique and contingent sort of bandwagon effect. Some people even suspected that such an effect was the reason he was polling so well: those who had signed up to vote in order to doom the party at the next general election saw that Corbyn had a reasonable chance of doing well in the leadership contest, and inflated his vote share artificially. But the fact that Corbyn led in the polls not only among these '£3 sign-ups,' but also among trade unionists and regular members, went some way to refuting this view (Dahlgreen 2015).

Beyond this growth of disingenuous support, it is not enough to observe that Corbyn polled surprisingly well to assert that there was a bandwagon effect taking place. It is possible Corbyn's policies just appealed to a broad section of Labour members, so they voted for him. This is not a bandwagon effect. It is just people voting for the candidate whose policies they like best. Political commentator Owen Jones (2015) claimed, during the campaign, that Corbyn was 'making astounding headway – against the odds – because [he] offers a coherent, inspiring and, crucially, a hopeful vision.' Post-hoc explanations of Corbyn's victory also emphasised that his politics were well-aligned with the increasingly left-leaning membership of the Labour Party (BBC 2015a).

Yet, there appears to be more going on here. For one thing, the kinds of claims that the likes of Owen Jones made were only made because Corbyn was popular. It is unlikely that, as a fringe candidate, he would have had any meaningful impact on discussion about the campaign. This discussion gives people information about Corbyn which may then influence their vote decision, when they realise they like his politics. But it is only because he is a popular, seemingly viable candidate that this had the opportunity to happen. Beyond this, Corbyn's popularity influenced the behaviour of other actors in the party. For example, several MPs responded to a poll showing Corbyn leading on 53% of the vote by imploring former leader Gordon Brown to intervene to persuade members away from voting for him – to 'halt Jeremy Corbyn['s] bandwagon' (Stacey 2015). Supporters of other candidates even rejected the idea of uniting behind one candidate, having the others drop out of the race, because 'it could just make Jeremy look stronger' (Stacey 2015). It is also worth noting that of the two widely available polls of eligible voters in the election, Corbyn first polled at 43%, then at 53%. It is not possible to assert from this that the first poll had an effect on the second, but this would be consistent with his surprisingly strong performance gaining him increased support. Voters, still doubtful that his apparent rise to prominence represented a real growth in support rather than just a media buzz, may have seen Corbyn's runaway lead in the first poll and realised that it was worth paying attention to him, and ended up voting for him as a result.

Corbyn also surprised many observers in 2017, when he led the Labour Party in his first UK general election (see Allen 2020). First of all, it is worth noting that in large part, how likely people thought it was that an election would be held, and how loudly Conservatives called for one, was due to popularity information. According to Cowley and Kavanagh,

> As the Conservative lead in the opinion polls grew throughout late 2016 and early 2017, so did the calls for an early poll. In February 2017, the Conservatives won the Copeland by-election, taking the seat from Labour, the first time a governing party had gained a seat in a by-election since 1982 and increasing the pressure on May to call an early election yet further… in early April, some of the party's private polling… was reported as finding that the Liberal Democrats could make gains from the Conservatives… This was widely interpreted as yet more evidence that there would be no early election (2018, 4–5).

When Theresa May eventually did call an early election in 2017, the Conservative Party's lead over Labour in publicly available nationwide polling was substantial (Cowley and Kavanagh 2018, 15) – the polls 'predicted a Conservative landslide' (Strong 2018, 467). To many, the result looked like a 'foregone conclusion' (Ford 2017). Yet, in the event, the Conservative Party failed to secure an overall majority in the House of Commons. The polls, throughout the campaign, showed Labour doing better and better, steadily increasing its vote share, until eventually it received over 40% of the vote to the Conservatives' 42%.

One popular, but very contentious explanation of this was that there was a 'youthquake' in the 2017 election as the Labour Party benefited from an unforeseen rise in support from young people (Ehsan, Sloam, and Henn 2018; Prosser et al. 2020; Stewart et al. 2018; Whiteley and Clarke 2017). This was, after all, the election campaign during which the 'oh, Jeremy Corbyn' chant was popularised at mass events – culminating, after the election, in the famous events at Glastonbury festival where thousands of Stormzy fans sang the chant.[2] Explanations like this

[2] It is worth acknowledging that events like this might have involved the 'normative social

do not imply that a bandwagon effect necessarily took place. Rather, the focus here is on the idea that the election campaign gave Corbyn's message a chance to reach and appeal to a wider audience.

As I will demonstrate in Chapter 5 though, there is likely to have been a bandwagon effect here as well. The polls increasingly showed Labour gaining ground on the Conservatives throughout the campaign, and this information made a difference to those who saw it, making them more likely to go on to vote Labour. Rather than representing a distinct way in which people came to support Corbyn's Labour, this is likely to be related, as Chapter 2 will make clear, to the idea that Corbyn did increasingly well because his message was reaching more and more people. The very fact of becoming more popular puts political parties' messages on more people's radar. Popularity makes people pay attention. In 2017, Labour's 'surge' (Ford 2017) in popularity was uniquely evident. The context was conducive to a bandwagon effect, so the results of Chapter 5 may come as no surprise.

Dewey Defeats Truman

Back across the pond, though, it appears that sometimes, popularity might not do electoral candidates such favours. An infamous media debacle of the 1940s led some to believe that being too popular might jeopardise a candidate's chances. On the morning after the 1948 presidential election, the Chicago Daily Tribune ran the headline 'Dewey defeats Truman' (see Figure 1.4). Complications with the printing process at the Tribune meant that the newspaper had to commit to a

influence' discussed above. The point is that this is unlikely to explain a growth in support at the national level, even if it can happen at isolated events.

headline earlier than most, so its editors assumed Dewey would win based on a combination of conventional wisdom and the polls. The headline was incorrect; Harry Truman won the election. This was a major upset considering that polls had suggested, quite convincingly, that Dewey would win.

As intimated above, polls, nowadays, seem to be known largely for the fact that they get things wrong. This is the case despite the fact that research suggests there is no major problem or crisis in polling (Jennings and Wlezien 2018). What's more, 1948 was relatively early days for polling. The kinds of 'scientific' approaches we think of as how polls are done only came to the fore in the 1930s. It is quite possible that the Dewey defeats Truman debacle is simply one more case of the polls missing the mark.

Some have cast doubt on this explanation, offering an alternative version of events. Burns Roper, son of pioneering pollster Elmo Roper, put it that

> The labor vote was energized as Democrats worried about Dewey's strength in pre-election polls while Republicans felt their candidate would win so they played golf that day... the polls were not wrong in terms of measuring national sentiment although they clearly were wrong in determining the election. I think the 1948 polls were more accurate than the 1948 election (quoted in Goeree and Großer 2007, 52).

This implies, then, that it was actually detrimental to Dewey that he was so convincingly in the lead in the polls. Instead of a bandwagon effect in which people flocked to support him because he was clearly the winner, his dominance created complacency in his support base. Polls, on this view, can become 'self-defeating'

Figure 1.4: Victorious Democratic candidate Harry Truman holds up a copy of Chicago Daily Tribune incorrectly proclaiming his opponent, Thomas Dewey, victorious in the 1948 presidential election.

(Henshel 1982). They may measure vote intention accurately when they are conducted, but they alter those intentions by the very fact of being conducted and published.

But as Niall Ferguson has argued, there was more to this case than the fact that Dewey was the clear favourite, and that

> What should have worried Dewey was that his lead in the Gallup poll shrank from 12 points in late August to 9 points by mid-October. Gallup's final voter survey a fortnight before the election showed him ahead by just 6 points. Yet contemporaries still assumed that was a

> sufficient margin to assure him of victory. Dewey had summed up the prevailing view in 1944: “Never argue with the Gallup Poll. It has never been wrong and I very much doubt that it will ever be” (Ferguson 2020).

In other words, it looks like Dewey was in the lead, but Truman was gaining ground. Truman had some momentum going into polling day. Notice that this dynamic idea of momentum has come through in earlier examples too. It is used to describe what happens in presidential primaries, and is a good description of what happened for Labour in 2017, and even reflects some of the media coverage in the Conservative leadership contest mentioned earlier. It seems like the popularity people care about when making their vote decisions might be not only of a static form, but also a dynamic form. This is a crucial distinction that I draw out in the thesis. Here, it demonstrates that even in real-world cases that essentially look like the opposite of a bandwagon effect, considering different aspects of popularity information reveals that bandwagon effects still might have taken place. These bandwagon effects might just be based on the dynamic trends in, rather than the static state of, popularity.

Ferguson’s (2020) argument above was also part of a broader discussion about how Joe Biden might suffer a similar fate in the 2020 presidential election to Thomas Dewey in 1948. Both candidates were in a similar position. They both stood against a widely derided incumbent and were quite clearly dominant in the polls. What if, like Truman, Trump had a resurgence and Biden’s voters, like Dewey’s, were confidently complacent in the knowledge of his inevitable victory? Biden went on to win, but his lead over Trump was not quite as convincing as many observers expected it to be. In Chapter 3, I will demonstrate that, consistent with

Ferguson's and Roper's concerns, this might have been because of his popularity conveyed by the polls going into the election.

Time Zones and Late Voters

Beyond just pre-election polls, there has been some concern that U.S. presidential elections could also be affected by exit polls – surveys of voters after casting their ballots, asking how they just voted. Unlike the Danish case mentioned above, this is not because of pernicious leaks, but simply a matter of the timing of votes. Because the USA spans multiple time zones, exit polls from a state on the East Coast, where polls close earlier, could affect voting behaviour in a state on the West Coast, where people are yet to vote. Sudman imagines that

> A Californian plans to vote after work in what she believes to be a close presidential election. (She has little interest in the race for congressman for her district, although it is closer.) The day is rainy and as she approaches the polling place she sees a long line. On the radio she hears that one presidential candidate has a substantial lead in other states. She says why bother and turns her car around and drives home (1986, 332).

Sudman (1986, 332–33) adds that people have almost certainly behaved this way many times, notably in the 1980 election 'when a close race was expected… but the exit polls showed a Reagan sweep' (see Delli Carpini 1984; Jackson 1983) and likewise in 1960 'where the early returns from the East Coast gave President Kennedy a substantial lead.'

More recently, similar dynamics may have affected the 2000 election, contested between George W. Bush and incumbent Vice President Al Gore. On election night, as results came in, the electoral college count across the country was so close that it was apparent that whichever candidate won in Florida would win overall. Indeed, the contest was just as close in Florida itself. Relatively early though, it was reported that Al Gore had won the state of Florida. But after a drawn-out, 36-day process of recounts, it was confirmed that Bush had in fact won the state by a mere 537 votes – and with it, the election (Elving 2018).

Crucially, the timing of the initial, incorrect announcement – given that it was about Florida, one of the first states to close its polling places given its location on the East Coast – meant that people who had not yet voted were under the impression Gore had won. Bale (2002, 15) notes that, as a result of this, many were ‘convinced that the major networks’ exit poll predictions of a Gore win in Florida prior to the close of voting may have robbed Bush of votes’. Even some parts of Florida itself were yet to vote. As a report commissioned in the wake of the election explains,

> CNN projected Gore as the winner 10 minutes before polls closed in [Florida]’s western panhandle, which contains 5 per cent of the population and where, as in most polling places, voters in line at poll closing time may still cast their ballots. Viewers were not told that those polls still were open… Whether, and how much, such calls affect intrastate voter turnout in the remaining minutes is not known… But presidential elections are too sacred a part of our democratic system to take such a risk. And as Florida showed, a few votes can mean a great deal (Konner, Risser, and Wattenburg 2001, 3).

The USA is not the only country where time differences can matter in this way. Many French voters in the country's overseas territories are, indeed, much further away from the country's mainland than those on the USA's two coasts are from each other. In the past, this had similar implications. Voters in, for example, Martinique and French Guyana used to vote after those in mainland France, knowing how the candidates had performed there. In most cases, this was inconsequential, because the difference between the two candidates was already larger than the 1.5% of the voting population that OST voters represent (Morton et al. 2015, 70). However, in 1995 and 2002, it was mathematically plausible that voters in OSTs could sway the election outcome.

French constitutional law changed in 2005, such that western OSTs now vote on the day before the mainland, removing the possibility of exposure to mainland results before voting. Morton and colleagues (2015) show that this change lowered turnout significantly in these OSTs, resulting in larger margins of victory for leading candidates.[3] Even in elections since this change, many French voters have found ways to learn the election result before the polls have closed. In 2007, several Swiss and Belgian newspapers published exit polls for the French election, avoiding the legal restrictions imposed on this within France itself. Their websites crashed when they were overwhelmed by French voters seeking out this information. Similarly, in 2012, French election results were online while voting was still in progress (Sayare 2012). Popularity information, in the form of exit polls, matters in French presidential elections.

In the 2018 Brazilian election, rather than simply because of where they lived, many

[3] In Chapter 2 I explain why this is insufficient evidence to claim that a bandwagon effect took place.

voters went to the polls after election results became public due to an administrative error. The error, this time, was in the technology used for voting. As Araújo and Gatto explain,

> In 2018, for the first time in Brazilian history, fingerprints were used as the main form of identification for 73.6 million voters—more than half of the electorate of 147.3 million people. Technical glitches associated with the use of the newly introduced biometric identification technology caused delays in voting processes, leading some voters to cast ballots after the release of the first official vote tallies (2021, 3).

Again here, the leading candidate did significantly better among voters who cast their ballots after the result was known; Jair Bolsonaro's vote share was, on average, approximately 14% larger in voting machines that suffered delays owing to technical glitches (Araújo and Gatto 2021, 23). It appears that knowing how the election is likely to pan out makes a difference to how people vote, across a range of different geographical and electoral contexts.

Causal Inference and the Bandwagon Effect

These examples all serve to demonstrate the relevance of key questions in research on the bandwagon effect: does the bandwagon effect come about as a result of viability? How does information about popularity relate to other information about candidates when people make their vote choices? Does momentum matter, beyond simply who is in the lead? Is it good to be seen as a winner? These are all questions that I deal with in this thesis. The examples hint that there might be something to these questions, but alone they cannot answer any of them.

This is because it is difficult – if not impossible – just by observing these cases, to assert that people's votes were influenced by popularity information such as the results of polls or the results of earlier rounds of an election. It might seem like there is something going on that a bandwagon or related effect helps to explain, but it is not possible to assess this claim or measure it without something more. This also means, as I will address further in Chapter 2, that there are potentially many cases where bandwagon effects have occurred but where it is not clear, just from observing the events of the election, that this is the case. This is where 'causal inference' comes in.

In this thesis, I focus on causal relationships. I try to find *causal* effects of popularity information on people's attitudes and behaviours. This kind of evidence is by far the most useful for the purposes to which research on the bandwagon effect is applied. Debates around the regulation of opinion polls need to know what the effects of those polls are. Only this can reveal what happens when you let people see a poll versus what happens when you restrict its publication. Campaign strategists only learn whether it is useful to be seen as a winner, or as having momentum, through knowledge of what the effect of that information is on the vote. Electoral forecasters need to know if their forecasts will affect the outcome they are trying to predict. Research on the bandwagon effect is useful insofar as it can inform these concerns, so it is useful when it is causal.

This thesis presents evidence both from experimental and observational research. Rather than having one consistent methodological approach throughout, I use a range of methods that are united by this central focus on causal inference. In other words, the methods are tools to explore aspects of the theoretical causal model that I propose. This model is developed in Chapter 2, where I use 'directed acyclic graphs' (Pearl 2000) – causal diagrams – to present my assumptions transparently. Every empirical analysis carried out in the thesis, though differing in statistical method or research design, captures a part of this model. I choose methods themselves more pragmatically, and discuss and justify them in the relevant chapters. I also opt to use both experimental and observational approaches precisely in order to make improvements on how they have been used in the past, based on the lessons of my causal model.

In causal diagrams, researchers set out transparent assumptions about what causes what, based on theory. These are expressed in the form of a 'graph,' which is a series of nodes representing concepts, constructs or phenomena, connected to each other by arrows that point in one direction ('directed'). The arrows represent the flow from cause to effect, so cannot loop back on themselves to indicate 'reverse causality' because causality flows forwards in time ('acyclic'). Then, the assumptions captured by the diagram make it possible to establish whether or not a causal relationship between two phenomena or concepts can be measured, in principle, and what this would require the researcher to do in the analysis. If the researcher does those things and finds a relationship between the two phenomena, then within the causal model's assumptions – which readers may, of course, challenge theoretically – the effect can be interpreted as causal (Pearl 2000).

As an example, consider the 'toy model' (Pearl and MacKenzie 2018) in Figure 1.5, taken from Rohrer (2018, 29). Here, the aim is to establish whether educational attainment – doing well at school – makes you likely to earn more money in the future. However, more intelligent people do better at school, and are also more likely to earn more money in the future, for reasons unrelated to their educational attainment. This means that if we take a survey asking people how well they did at school and how much money they earn, then put the two into a statistical model, the resulting association will be larger than the real *causal effect* of educational attainment on income. It will also capture the extra association brought about by the 'confounding' effect of intelligence.

$$\begin{tikzcd} & Intelligence \arrow[dl] \arrow[dr] & \\ Educational\ Attainment \arrow[rr] & & Income \end{tikzcd}$$

Figure 1.5: Example causal model extending on Rohrer (2018, p.29).

Focusing on causal inference highlights such problems, but also allows researchers to be pragmatic in how they resolve them. An experiment would deal with the problem in Figure 1.5 by randomly assigning people a level of educational attainment. By doing this, it would no longer be the case that more intelligent people have better educational attainment, because attainment was assigned completely randomly. This removes the arrow from intelligence to educational attainment, resolving the confounding problem. But of course, it is impractical, unethical, and probably impossible to randomly assign educational attainment. This would involve, for example, giving students random grades that do not actually reflect their performance in their exams. So instead of removing the confounding experimentally, it needs to be done statistically. For example, by taking the survey mentioned above and also measuring people's intelligence, these intelligence levels can be held constant in a statistical model, leaving the remaining association between educational attainment and income to capture the causal effect. Causal inferential considerations, combined with pragmatism about feasible research design and available data, drive the choice of method.

Such pragmatism is also particularly necessary in the case of this thesis because appropriate data to study the bandwagon effect are hard to come by. This is, indeed, why so many scholars have resorted to experimental approaches – in which they effectively create appropriate data themselves (Grillo 2017, 466) – as I also do in Chapters 3 and 4. As I demonstrate in Chapter 5, a commonly used form of observational data does not fit the bill. Rather, to study the bandwagon effect outside of experiments, it will typically be necessary to use longitudinal panel data, unless suitable natural or quasi-experiments become available (e.g. Araújo and Gatto 2021; Kiss and Simonovits 2014; Morton et al. 2015). But longitudinal panel datasets are somewhat rare because they are expensive to collect. Those that do exist rarely ask the right kinds of questions at frequent enough intervals to have any hope of identifying a bandwagon effect. This means that doing research on the bandwagon effect that is appropriately focused on causal inference is not always straightforward. This thesis makes multiple attempts to overcome this, but it remains a challenge for the field.

Chapter Overview and Research Questions

The aim of this introduction has been to establish, in broad terms, what this thesis is about, why that subject matters in the real world, and how I approach it. The following four chapters put this into practice. Over the course of these chapters, I address several key research questions, here denoted $RQ_{1\text{-}5}$.

Chapter 2 provides a concrete conceptual, theoretical and methodological framework for the thesis. It takes the view that ideally, prior to studying the bandwagon effect, we should establish what it is. As such, it asks the fundamental question: what do we mean when we talk about a bandwagon effect?

> **RQ$_1$**: What is the bandwagon effect?

To answer this question, Chapter 2 uses conceptual analysis, defining the effect by establishing what its necessary and sufficient conditions are and how these make it distinct from other phenomena, while accounting for how existing research has used the term. I define the bandwagon effect as

> A positive individual-level change in vote choice or turnout decision towards a more popular or an increasingly popular candidate or party, motivated initially by this popularity.

This definition leaves some leeway, along two dimensions: the effect can be based on which parties are more popular than others right now (static popularity information), or on which parties are becoming increasingly popular (dynamic popularity information); it can be expressed through a change in vote choice between available options (conversion), or through a change in intention to turn out to vote (mobilisation). As such, I propose a typology of bandwagon effects around these distinctions. This means that, in practice, there are four ideal-types of bandwagon effects, rather than one simple effect: static conversion effects, dynamic conversion effects, static mobilisation effects, and dynamic mobilisation effects. I show that some of these have received much more attention in research than others. This part of the chapter has been published in *Political Studies Review*

(Barnfield 2019). Following this, subsequent research on the bandwagon effect has employed the definition and typology (Araújo and Gatto 2021; Larsen and Fazekas 2021; Volckart 2021).

With the key concepts established, the chapter then asks how the bandwagon effect is likely to come about. This involves situating the bandwagon effect in broader, but also more detailed, theoretical terms.

> **RQ_2**: Why should the bandwagon effect happen, theoretically?

Here, the informational approach previewed above comes to the fore, as I combine the four types of effect with theories of political psychology, and existing knowledge of how voters decide, to build a causal model of the bandwagon effect. Consistent with this work, I argue that the way the bandwagon effect comes about is likely to be a function both of the electoral context and individual characteristics. These combine to determine how 'difficult' the vote decision is (Lau and Redlawsk 2006). When they make it difficult – for instance, when there are lots of candidates and voters do not have well-established, strong predispositions towards these candidates – the bandwagon effect can come about, as alluded to in the examples above, by pushing voters towards 'viable' candidates (Utych and Kam 2014). When things are easier – for instance, when there are only a few different parties and voters know which party they like – voters do not need these 'viability' cues so much, but they might still be subject to a bandwagon effect because the popularity of a party will motivate them to try to understand this popularity by reflecting on why other people like the party. These reasons may then convince them to support it too, in what is known as a 'cognitive response' (Mutz 1997). This is similar to how other types of information influence voters (Lodge and Taber 2013), thereby showing how thinking about the bandwagon effect in terms of information draws out links to broader political psychology. Importantly, and consistent with this broader theory, this makes people's preferences or predispositions very important in how we understand the bandwagon effect.

The last section of the chapter asks, with these theoretical points and the focus on causal inference in mind, how can we think about the ways in which the bandwagon effect has been measured in research? A key part of this discussion involves establishing where 'electoral expectations' – how well people think a party is going to perform at an election – fit into the bandwagon effect, because they are so commonly used as a way of studying the effect. Because expectations are opinions, I point out that they are based on (popularity) information: people think that the parties that information, such as the polls, suggests are popular, are likely to do better at elections. I note, here, that it is plausible that the relevant information here is again not only which parties are more popular than others right now, but also how the popularity of a party is changing over time. Opinions are also based on people's predispositions: people think that the parties they like are likely to do better at elections. This means that expectations, like voting behaviour according to the bandwagon effect, can be driven by popularity information and people's preferences or predispositions. This discussion feeds into RQ_4 and RQ_5, which I return to below.

Chapter 3 takes the main arguments from Chapter 2 and considers how they can be implemented empirically, to study the bandwagon effect directly.

> **RQ_3**: How can the concept and theory of the bandwagon effect be put into practice empirically?

This chapter recognises that no existing work has studied all four of the potential types of bandwagon effects together at once, nor has any existing work studied bandwagon effects in the case of state polls in presidential elections. It asks, theoretically, would it make sense for voters in US presidential elections to be influenced by the results of polls in their state? For example, would we expect the decision of a voter in Georgia potentially to be affected by how well the presidential candidates are polling among voters from Georgia? I argue that when considering both the viability and cognitive response explored in Chapter 2, it makes (potentially, especial) sense for state polls to produce bandwagon effects. But the conceptual arguments of Chapter 2 also show that the study should allow different types of bandwagon effects to emerge. I apply this thinking to the case of state polls in the 2020 US presidential election. Specifically, I randomly expose voters to information about the static *and* dynamic popularity of Trump and Biden in potential swing states (Georgia, North Carolina, Ohio, and Texas) in the lead up to the 2020 U.S. presidential election. Respondents have the choice to vote for a candidate or abstain. The results of this experiment are not promising for the bandwagon effect. They suggest that being portrayed as the leading candidate in a state poll tended to *decrease* Biden's support in that state, especially among those who did not already identify with the Democrats.

As well as providing important insights into the effects of state polls, and demonstrating how the typology in Chapter 2 can be applied to understand the role of polls in a given election, the findings of this study have implications for considering how these polls overstated Biden's dominance in 2020. Forecasters claim that a major reason things looked rosier than reality for the Democrats in the lead up to the election was that the polls in many states overestimated Biden's vote share

(Kurtzleben 2020). If Biden lost support when he was seen to be in the lead in state polls, then it could be that rather than simply over-estimating Biden's share, they may have partly *caused* it to decrease. While these effects are quite small, and therefore extremely unlikely to be the sole reason for the discrepancy between the polls and the outcome, this nonetheless suggests that the polls might not have been as wide of the mark as they appear to have been.

While hapter 3 therefore provides an example of how *experimental* research on the bandwagon effect could move towards more comprehensive, nuanced, and ethical measurement of the effect, this is of limited use when scholars are unable to, or do not wish to, conduct an experiment. It also does little to shed light on the complexities of the causal model set out in Chapter 2. Recognising this, Chapter 4 moves beyond simply looking for direct bandwagon effects of polls on voting behaviour experimentally, to think about other claims made by the theoretical model in Chapter 2. Although Chapter 4 itself continues with the experimental approach, it represents a step in the direction of asking whether and how the bandwagon effect can be studied outside of the experimental context. This is because the chapter considers an overlooked feature of the kinds of voter opinions scholars use to study the bandwagon effect in survey data. Specifically, it asks whether electoral expectations are affected by dynamic popularity information.

> **RQ$_4$**: Does dynamic popularity information affect electoral expectations?

My focus is specifically on dynamic effects because extensive research has already demonstrated that people's electoral expectations are affected by static popularity (e.g. Blais and Bodet 2006; Irwin 2002). Chapter 2 makes the claim that expectations should be responsive to *both*, but this goes untested until Chapter 4. This chapter carries out two studies designed to test the claim. As well as validating part of the causal model in general, this is useful because this claim has major implications for how justifiable it is to use data on people's expectations as a way of measuring the bandwagon effect.

Chapter 4 notes that when people answer questions about their electoral expectations, the election in question takes place, by definition, in the future. This means that people need to work out how the popularity of the parties might evolve between now and then. How popular the parties are right now is a good starting point, and should be increasingly accurate as the election draws near, but it is imperfect. In a sense, people need to mentally map this current, static popularity of the parties forwards in time, to work out how popular they *will become*. This makes 'dynamic popularity information' – change in popularity over time – relevant. If there is evidence that a party's share of vote intentions is currently growing, or has been growing, then this might make people think it is going to continue to grow and improve the party's eventual vote share, and thereby boost its chances of winning the election. Not only is this consistent with theory and evidence from social psychology, but also people think about 'momentum' in this way when predicting the outcomes of other contests, such as sports matches, so it could apply to political contests too.

The chapter presents experimental evidence of this effect in action. Across two studies with representative samples of UK respondents, Chapter 4 shows that people with information indicating that a second-placed party has recently gained ground in the polls think that this party has a better chance of winning, compared to those who just know that it is in second place. The greater this momentum, the larger the effect. The findings also suggest that in real-world contexts, depending on levels of information and the strength of voters' preferences about the outcome, the effect is reduced, but still significant overall.

These findings are important to Chapter 5, the final piece of the empirical puzzle in this thesis. The experimental approach demonstrated in Chapters 3 and 4 only goes so far. Scholars often want to, or can only, rely on survey data from the real world, rather than artificial contexts controlled by the researcher. In the case of the bandwagon effect, this usually involves taking data on people's expectations and relating it to questions about their vote choice, which raises the question of whether this approach works for making causal inferences.

> **RQ_5**: Can electoral expectations be used to study the bandwagon effect?

By scrutinising the causal model established in Chapter 2 and used throughout the thesis, Chapter 5 shows that what I call the 'expectations approach' typically sets itself up to fail. Essentially, this approach involves looking for associations between people's electoral expectations and their voting behaviour, in order to estimate the effect of the popularity information contained within those expectations on the vote – the bandwagon effect. In doing so, it has to deal with the fact that these expectations are also based on voters' predispositions – they will say the parties they like have a better chance, even beyond whether the information suggests this, because of 'wishful thinking.' It addresses this problem by controlling for measures of voters' partisan preferences in the analysis. But by doing so, the approach prevents the bandwagon effect from *passing through* these preferences. It assumes that people who undergo a bandwagon effect never start to like the parties that end up getting their vote more, along the way. This directly contravenes the model presented in Chapter 2, and goes against prominent literature on the bandwagon effect. It is wishful thinking to expect this approach to address the problem of wishful thinking. I take to calling the paradox at the heart of this chapter ‘double confounding.’

Moreover, the findings from Chapter 4 signal a different problem: with data on electoral expectations, there is no way to distinguish between *static* and *dynamic* popularity information. So, in principle, the expectations approach cannot claim to measure either static or dynamic bandwagon effects. More problematic than this though, it means that in situations where static and dynamic information might be at odds with each other (such as when Dewey was in the lead but Truman was gaining ground) the approach might rule out bandwagon effects where it should find them.

Chapter 5 resolves these problems by introducing time into the analysis. I demonstrate a method that uses longitudinal data on electoral expectations to study the bandwagon effect. This approach is based on recent recommendations for studying causal relationships through ‘general cross-lagged models’ (GCLM) (Zyphur, Allison, et al. 2019; Zyphur, Voelkle, et al. 2019). I explain the logic behind this method, and how this fits with bandwagon research, noting that drawing out the causal process over time means that the model’s more natural interpretation is as a measure of *dynamic* bandwagon effects. I apply the approach to three consecutive waves of the British Election Study Internet Panel (BESIP). The analysis finds evidence of a small dynamic bandwagon conversion effect in the Labour Party vote in the run up to the 2017 UK general election.

The conclusion, in Chapter 6, brings the work of these four chapters together to clarify and restate what they have to say about the research questions, and what their immediate implications are for debates about opinion polls, political and campaign strategy, and electoral forecasting. Following this, I move on to the question of what future research on the bandwagon effect and related questions could and should address.

Summary

Though at times it is critical of how things have been done in research on the bandwagon effect, my intention is that this thesis recognises the value of existing contributions and builds on them positively. With a combination of experimental and observational evidence, I try to demonstrate that much of what has been learned about the bandwagon effect so far can be adapted and reapplied to further our understanding. In order to do this, it is necessary first to go back to basics and more clearly establish what the bandwagon effect is, and why we should expect it to come about in the first place. This is the task to which I turn in Chapter 2.

Chapter 2

Think Twice Before Jumping on the Bandwagon: Concepts, Theory and Measurement

Much of the content of this chapter has been published as a peer-reviewed journal article in *Political Studies Review* (Barnfield 2019). The article focuses primarily on, and extends, the discussion presented here under 'A Definition of the Banwagon Effect' and 'A Typology of the Bandwagon Effect.'

Introduction

When studying the bandwagon effect, it is useful to know what exactly this term means and why we would expect the effect to happen. The existing literature is surprisingly inconsistent on these points. There is a wealth of insightful research on the bandwagon effect, but a lack of conceptual and theoretical consensus makes it difficult for these studies to speak to each other and build towards a coherent body of knowledge (Glynn and Mcleod 1985, 64). In this chapter, I lay the groundwork for my study of the bandwagon effect by attempting to address such conceptual and theoretical points. The chapter goes from conceptual analysis of the bandwagon effect, to a theoretical model, to finally introducing questions of measurement that are fleshed out in the empirical chapters. In doing so, the chapter addresses RQ_1 and RQ_2, as set out in Chapter 1:

> **RQ_1**: What is the bandwagon effect?
>
> **RQ_2**: Why should the bandwagon effect happen, theoretically?

The first section provides a definition and typology of bandwagon effects, in response to RQ_1. This discussion draws on principles of 'conceptual analysis' (see Dowding 2017, 1–5): it defines and delineates the bandwagon effect by establishing what its necessary and sufficient conditions are and how these make it distinct from other phenomena, while accounting for how existing research has used the term. Based on a review of 68 peer-reviewed journal articles, I define the bandwagon effect as: *a positive individual-level change in vote choice or turnout decision towards a more popular or an increasingly popular candidate or party, motivated initially by this popularity*.

This definition is the most succinct and coherent way of reconciling the existing evidence without being excessively proscriptive, while also clearly distinguishing the bandwagon effect from other phenomena with which it is often conflated. That is, it is an attempt to draw out as much consensus from the literature as possible while avoiding inconsistencies and confusion with other established concepts, such as strategic voting. New research on the bandwagon effect is beginning to adopt this definition (e.g. Volckart 2021).

This section also proposes a typology of bandwagon effects, drawing on distinctions made in this literature between ‘static’ and ‘dynamic’ popularity information, and between ‘conversion’ and ‘mobilisation’ as outcomes. I make the case for why these distinctions are useful and use this typology to frame existing research, considering and proposing explanations for the imbalance of conceptions of the bandwagon effect, while drawing on the definition set out in the first section. In doing so, I show that bandwagon research has focused heavily on ‘static bandwagon conversion effects’ (switching to support a leading candidate). This typology has informed new research on the bandwagon effect (e.g. Araújo and Gatto 2021; Larsen and Fazekas 2021).

The second section of the chapter then builds a theoretical model from this definition and typology, in response to RQ_2. The model formalises the proposal that the bandwagon effect is an effect of popularity on voting behaviours. However, this section also delves into *why* theoretically this effect might come about. I draw on existing knowledge and theory about how voters engage with and use information, both in general and in the specific case of what I call ‘popularity information’ – usually thought of as the results of opinion polls, but essentially any information about how much support electoral candidates or parties have and how this is developing. The consistency of these arguments with broader and deeper political and social psychology demonstrates that they represent a principled theoretical starting point.

Finally, the chapter discusses how these conceptual and theoretical points translate into measurement. I note that two approaches dominate the existing empirical literature – one observational and one experimental. The observational approach, based on measuring relationships between people's electoral expectations and their vote choice, creates the need to consider where expectations fit into a theoretical model of the bandwagon effect, in order to assess the validity and value of the approach. Incorporating expectations into the model presents challenges that are tackled later in the thesis. First, it raises the untested question of whether people's electoral expectations are responsive to what I will term 'dynamic popularity information' – information about how a party/candidate's popularity has changed over time. Second, it raises questions about the validity of using expectations data to study the bandwagon effect.

The second, experimental approach allows researchers to control directly what popularity information voters are exposed to, and assesses the differences between those who see that a candidate or party is popular and those who do not. While this guarantees internally valid causal inferences, the typology presented here demonstrates that the approach could be used to offer greater, more nuanced insight into the bandwagon effect than it has done so far.

A Definition of the Bandwagon Effect

The bandwagon effect refers to a positive individual-level change in vote choice or turnout decision towards a more or increasingly popular candidate or party, motivated initially by this popularity. This definition relies on three key features:

i that it is an individual-level phenomenon ***ii*** involving a positive change in vote choice or turnout decision towards the most, or increasingly, popular candidate, ***iii*** motivated at root by this popularity.

i The distinction between aggregate- and individual-level study of the bandwagon effect is stated clearly by Kiss and Simonovits, who explain that

> The individual-level bandwagon effect is the notion that some voters convert to become supporters of the leading candidate... The aggregate-level bandwagon effect is the notion that the leading candidate derives an electoral advantage of being perceived by voters to be in the lead (2014, 328).

The authors go on to adopt the latter approach on the basis that they are using aggregate data. They also argue that this makes sense because aggregate bandwagon effects are more consequential for policy debates. They admit, however, that an aggregate effect is consistent with 'many micro-level processes.' A simple reason why this is problematic is put forward by Chung, Heo, and Moon (2018, 422; see also Morwitz and Pluzinski 1996, 54): null effects at an aggregate level 'can either result from the null effect of the poll publications or from the simultaneous occurrence of both a bandwagon and an underdog effect.'[1] To the extent that it is safe to assume that individuals may be affected differently by information about popularity, it seems possible that one effect might simply cancel another out.

[1] The underdog effect can be seen as the exact opposite of the bandwagon effect, occurring when there is an equivalent individual-level change towards favouring the least popular, or increasingly unpopular, candidate. There is little evidence for the existence of such an effect and it is often only deemed a possible explanation for null results in studies of the bandwagon effect. This is especially true in aggregate contexts where it has been noted that candidates' leads tend to shrink as election day approaches (e.g. Gallup and Rae 1940).

While evidence of so-called underdog effects seems to appear very infrequently (Hardmeier 2008), it is possible to imagine a scenario in which both effects occur in tandem and produce no evidence of an overall aggregate effect. The aggregate approach has nonetheless been adopted in much scholarship, particularly earlier work (e.g. Beniger 1976; Gallup and Rae 1940), but also more recently in otherwise novel approaches relying on aggregate data (e.g. Hodgson and Maloney 2013; Kiss and Simonovits 2014; Morton et al. 2015).

Essentially then, it is not necessary for a bandwagon effect to result in a discernible aggregate-level increase in support for any popular candidate, because a bandwagon effect can still be occurring for some non-trivial number of individuals without any observable aggregate shift (Chung, Heo, and Moon 2018). Perhaps unsurprisingly, aggregate-level study has historically found little evidence of an effect (e.g. Gallup and Rae 1940; Mendelsohn and Crespi 1970). It is possible that in some of these cases bandwagon effects have been incorrectly ruled out on the basis of aggregate-level evidence – although a meaningful individual-level effect has nonetheless occurred – owing to the presence of countervailing effects. Where aggregate-level study has found effects, it has been difficult for scholars to determine the nature of these (e.g. Morton et al. 2015).

It is possible that such countervailing effects might mask the presence of bandwagon effects even in individual-level analyses. This is because, typically, scholars are interested to measure average effects. These averages might account for both bandwagon and underdog behaviours, which again cancel each other out. Crucially though, individual-level analysis is the only approach that enables researchers both to check whether this is the case, if necessary (see, e.g., Nadeau, Cloutier, and Guay 1993), and to link any 'net' average effect directly to the presence or absence of favourable popularity information through a causal model. That is, the individual-level approach facilitates causal inferences about individual voters.

ii Related to the point made by Kiss and Simonovits (2014) that aggregate appearance of a bandwagon effect is ‘consistent with many micro-level processes,’ the effect can also ostensibly occur in aggregate contexts without the individual-level characteristics of bandwagoning taking place. For this reason, the definition posits that a bandwagon effect has not occurred in those situations where there is no *positive* change of preference or turnout decision *in favour of* the most popular or an increasingly popular candidate – such that this candidate will a larger number of votes, *ceteris paribus*. As Hodgson and Maloney explain,

> If we see a leading party extending its lead during an election, or an election campaign, it does not necessarily mean that electors are flocking to the leading side because it is winning. Quite apart from the influence of events since the election began, it could be that the winning party’s (presumably more convincing) message has had more time to get through (2013, 79).

Henshel and Johnston (1987, 494) put it simply that ‘it is essential... that candidates gain or lose votes not merely after a poll but because of the poll.’ Clearly, it is important to distinguish a specific phenomenon like the bandwagon effect from general changes in support, which could stem from a vast range of voter considerations. It is not enough to claim every voter gained by more popular parties is the result of a bandwagon effect. The bandwagon effect only happens when they gain these voters as a result of this popularity. The need to identify this clarifies further why aggregate approaches fall short. Parties that appear to derive an electoral advantage from their perceived popularity may in fact be gaining voters for any number of reasons.

This is different from saying the bandwagon effect occurs when a more popular party or candidate benefits from other, less popular options *losing* support for their relative *un*popularity. Aggregate-level study allows these to be conflated, because the effect on vote shares can be the same, but the individual-level process differs importantly. If a loyal Labour voter in the 2017 UK General Election decided not to bother voting because the polls told her Labour was not going to win, this would have increased the Conservative vote share without signalling a bandwagon effect. This voter has not made a positive switch towards another party, nor a positive decision to turn out, and the party her decision benefits only increases – through this behaviour alone – its relative *share* of the vote, and not overall *number* of votes. In other words, mass behaviour of this nature would give the illusion of a bandwagon effect at the aggregate level, but does not involve anyone jumping on another party's bandwagon.

A particularly clear example of this can be found in Kiss and Simonovits (2014, 339), who note that at least part of their effect comes from districts in which 'parties under-performed relative to what could be expected based on the first-round result, but the party expected to lose under-performed by 1 to 1.5 percentage points more.' In other words, the party for whom their conclusions suggest there was a bandwagon effect was in many cases *losing* votes, just at a (marginally) lesser rate than its opponent. Claiming that this party benefited from a bandwagon effect quite clearly misrepresents the evidence. This general point is made more important by the fact that scholars – and, indeed, journalists – often speak about bandwagon effects directionally, as occurring for or in favour of a party or candidate.

iii A subtly different point is that a bandwagon effect requires this change to be motivated *initially* by the popularity of the candidate or party. To repeat what was said above, it is not enough to claim every voter gained by more popular parties is the result of a bandwagon effect. This becomes increasingly complicated, though, when noting that voters can be motivated to support more popular candidates ostensibly for their popularity, in a decision that is entirely strategic. Yet strategic votes do not *inherently* go towards such candidates, whereas bandwagon effects do. This creates a problem.

Intuitively, the idea that the bandwagon effect and strategic voting are different is quite clear. Strategic voting is defined by Bartels (1987, 21) as an ‘attempt’ by a voter to ‘maximize their favourable impact upon the outcome of the election, if necessary by voting for a second-best candidate in order to forestall the election of a less-attractive alternative’ (see also Alvarez, Boehmke, and Nagler 2006; Blais, Gidengil, and Nevitte 2006; Blais and Nadeau 1996; Evrenk and Sher 2015; Fisher 2014, 2004). These specific requirements are obviously not the same as the broader meaning of the bandwagon effect. Yet, finding a clear conceptual point of distinction without inconsistencies arising is somewhat challenging.

In attempting to make such a distinction, Evrenk and Sher (2015, 406) focus on whether the voter believes ‘his [sic] vote will be decisive’; only if this is the case has strategic voting occurred, otherwise a ‘misaligned vote’ might be a bandwagon effect. Blais et al. (2006, 270) claim that strategic voting has taken place when a voter switches to support an expected winner without their evaluations of the parties/candidates changing, whereas when these change as well this is a bandwagon – or ‘contagion,’ see below – effect. There is a conflict between these two interpretations, in that the former deems the bandwagon effect another type of ‘misaligned’ voting alongside strategic behaviour, whereas the latter appears to distinguish between the two along the lines of whether or not they constitute a misaligned vote. That is, for Blais and colleagues, strategic voting is misaligned with preferences, whereas the bandwagon effect supposedly involves realigning these preferences themselves. This misalignment/realignment contradiction speaks not only to the difficulty in distinguishing the bandwagon effect from strategic voting, but also arguably to the fact that neither of these are appropriate ways of doing so.

Neither distinction appears to hold water when properly scrutinised. Regardless of the vanishingly small probability of any individual vote being ‘decisive,’ it seems entirely possible that a bandwagon voter could perceive her vote as such, without her vote therefore becoming strategic. There are also doubtless many strategic voters who think their strategic behaviour is unlikely to make any difference but engage in it anyway, simply because voting for their preferred party makes even less sense, in ‘purposive’ terms (Alvarez, Beohmke, and Nagler 2006). Distinguishing by whether party evaluations change is equally questionable – though, as discussed further below, this is theoretically likely to happen in the bandwagon effect. It contradicts the assumption many make, that bandwagon effects can come about when people dispassionately back winners – as Lazarsfeld, Berelson, and Gaudet’s (1948) respondent put it: ‘Didn’t make any difference to me who won, but I wanted to vote for the winner.’ The requirement of an adjustment of party evaluations is also absent from the literature in general. That is, much of what Blais and colleagues claim to be strategic voting might actually be considered bandwagon effects in other studies where expectations determining vote choice is deemed sufficient for a bandwagon effect (e.g. Granberg and Brent 1983; Kiss and Simonovits

2014; McAllister and Studlar 1991). The existing literature seems to suggest that bandwagon effects can be both a matter of misaligned voting (e.g. Evrenk and Sher 2015; Lanoue and Bowler 1998) and a matter of realignment of preferences (Blais, Gidengil, and Nevitte 2006; Goidel and Shields 1994). Arguably, defining the bandwagon effect in either of these contradictory ways would be unhelpful because it would mean arbitrarily discarding the insights of one side in favour of the other. The definition proposed here instead accepts both as valid.

I propose that the distinction between strategic voting and the bandwagon effect – where the two are otherwise 'observationally equivalent' (Araújo and Gatto 2021; Little 2021) – lies in the initial motivation. The bandwagon effect occurs when a more or increasingly popular option is favoured because of its popularity, as the *primary factor*. This comes very close to a distinction made by Lanoue and Bowler (1998, 373). It distinguishes the bandwagon effect from strategic voting because the latter is initially motivated by the relative unpopularity of the initial preference – voters try to avoid 'wasting' their vote (Alvarez, Boehmke, and Nagler 2006). If preferences align with popularity, then this vote-wasting logic need not lead to strategic voting – although some argue that strategic and sincere votes can be aligned (Riambau 2016). If they do not align, then voters might use the information available to them to attempt to affect the election in a favourable way, and this can mean voting for a more popular party or candidate. The distinction made here cuts off this channel by stating that a bandwagon effect has to be initially motivated by the popularity of the vote beneficiary.[2]

[2] First-past-the-post systems with single member districts, as in the UK, complicate this slightly. Strategic voting is typically construed, in such contexts, as being based on the inability of a given candidate to win in the constituency, and thereby provide constituents a means of influencing the election outcome at the national level. There are therefore two levels to consider: party and candidate. In such cases, it might make sense to assume the popularity on which a bandwagon

It could be argued, of course, that distinguishing between strategic voting and the bandwagon effect is unnecessary. Some research has explicitly stated that strategic voting is a type of bandwagon effect (Kenney and Rice 1994). As intimated above, a simple reason this distinction is necessary is that strategic voting does not *necessarily* benefit more popular candidates or parties – especially in proportional representation systems where strategic voters might flock towards smaller parties who are likely to be included in a victorious coalition (Gschwend 2007; Riambau 2015, 2016) – whereas bandwagoning does necessarily benefit more, or increasingly, popular options.[3] Smaller parties might also benefit from strategic voting in first-past-the-post single member district systems, because a given candidate has a good chance of winning a seat.

effect is based is that of the party whereas strategic voting is primarily based on the (un)popularity of the candidate. This is another reason why the effects are not to be conflated. Quite feasibly, constituency-level bandwagon effects could occur, although these would seem less likely. In this case Lanoue and Bowler's (1998, 373) simple distinction would essentially apply, in that strategic voting would be triggered by candidate unpopularity and a bandwagon effect by their popularity. Both of these cases still fit with the popularity criterion proposed here.

[3] In such cases, it might not be true that strategic voting is motivated by how unpopular one's preferred party is, because voters might even forgo voting for a leading party in order to vote for a smaller party in a 'coalition insurance strategy' (Gschwend 2007). However, this is of little consequence here because this kind of strategic voting does not risk confusion with bandwagon voting. My claim is that considering what the initial motivation is allows for a clear distinction between these behaviours where they *might otherwise be confused*, and not that strategic voting is always of the 'wasted vote' form. In precise terms, as already noted, my distinction applies when bandwagon effects and strategic voting are 'observationally equivalent' (Araújo and Gatto 2021; Little 2021). In coalition insurance strategies, or 'rental voting' (Gschwend, Stoetzer, and Zittlau 2016), this is not the case.

A Typology of the Bandwagon Effect

The bandwagon effect can be motivated by how popular an option is relative to other options, or how that option's popularity has changed over time – in other words, it is either static or dynamic (Irwin and Van Holsteyn 2000; Marsh 1985; Stolwijk, Schuck, and de Vreese 2017; van der Meer, Hakhverdian, and Aaldering 2016). The effect can be felt through a positive change towards a new candidate, or towards turning out to support the already-preferred candidate – it is either a conversion or a mobilisation effect (Agranov et al. 2018; Ansolabehere and Iyengar 1994; de Bock 1976; Morton et al. 2015; Moy and Rinke 2012). Table 2.1 summarises a typology of bandwagon effects around these distinctions.[4]

This typology provides a way of classifying existing research. Table 2.2 is an attempt to do this, placing each study in the cell to which its conception of the bandwagon effect seems to conform best. Some studies appear twice, because they have explicitly considered different types of bandwagon effect. In total, 68 peer-reviewed journal articles from English-language publications are classified here, selected by consulting a previous review, for older literature (Irwin and Van Holsteyn 2000), and a Web of Science search with snowball sampling, for more

[4] Consistent with the thesis, this typology is framed in electoral terms. This also matches the focus on voting behaviour within the literature and notes that it is in this context that the risk of conflating different effects seems high. It is nonetheless possible to adapt this typology to other public opinion contexts. The static/dynamic dimension would remain, referring to bandwagon effects towards adopting the majority opinion, or an increasingly popular opinion, respectively (see Marsh 1985). Bandwagon conversion would become simply about switching from one opinion to another, based on either static or dynamic popularity of that opinion. Bandwagon mobilisation would become about expressing or mobilising in support of an unchanging opinion, where one wouldn't have done so before (spiral of silence theory would likely be relevant here: Noelle-Neumann 1977). Recognising that this translation is quite straightforward, I classify studies on issue opinions in the typology.

Table 2.1: A typology of bandwagon effects.

	Static	**Dynamic**
Conversion	Switching to vote for a candidate or party on the basis of their/its popularity at a given time, relative to the popularity of other options.	Switching to vote for a candidate or party on the basis of their/its growth in popularity over time.
Mobilisation	Turning out to vote for an already-preferred candidate or party on the basis of their/its popularity at a given time, relative to the popularity of other options.	Turning out to vote for an already-preferred candidate or party on the basis of their/its growth in popularity over time.

recent studies.

Bandwagon Conversion and Mobilisation Effects

A bandwagon effect is possible both when an individual switches to vote for a more or increasingly popular candidate and when she decides to turn out to vote for that candidate, having not previously intended to do so.[5] This is widely asserted (e.g. Agranov et al. 2018; Ansolabehere and Iyengar 1994; de Bock 1976; Morton et al. 2015; Moy and Rinke 2012), but not always clearly expressed. In their natural

[5] There are certain possible scenarios that do not seem to fit this binary distinction. To clarify these, a non-voter without a preference undergoes a bandwagon effect and votes for a candidate, this could still a bandwagon conversion effect. If this non-voter has a preference and then such information leads them to vote for a candidate other than the original preference, this also is a bandwagon conversion effect. A bandwagon mobilisation effect only occurs when the preference is constant, but the new information leads to a positive decision to turn out to vote in accordance with this preference. Whenever the preference itself changes, even when this change is from 'none of the above' to a substantive option, this is a bandwagon conversion effect.

experiment, Morton et al. argue that

> Individuals may receive utility from voting for the winner such that they disregard their private information or preferences and vote for the candidate who feel ex ante and has more support [sic] (bandwagon vote switching effect) or (2) individuals receive utility from voting for the winner such that they are [more] willing to turn out to vote for their most preferred candidate when he or she is expected to win than when he or she is expected to lose (bandwagon turnout effect) (Morton et al. 2015, 76).

However, implicitly, this explanation allows for the possibility that deciding not to turn out when one's party is expected to lose constitutes a bandwagon effect, which it does not, because there is no movement towards another candidate. In other words, the leading candidate will not receive a greater number, as opposed to just a larger share, of votes as a result of this behaviour. The bandwagon effect is not to be confused, as this implies and was noted above, with what has been dubbed a 'Titanic effect': instances of individuals abandoning a 'sinking ship' by deciding not to turn out to vote because their party has no chance of winning (Irwin and Van Holsteyn 2000, 16–17). This confusion is common, however, with similar studies that allow the bandwagon effect to be about mobilisation frequently blurring this distinction (e.g. Grillo 2017; Hodgson and Maloney 2013; Kiss and Simonovits 2014). This is reflected in a related study which turns to calling the effect 'bandwagon abstention' (Morton and Ou 2015). Großer and Schram explain that

> If a poll indicates that a vast majority of the electorate supports either

> of the two candidates, some voters may assume that the outcome of the election is obvious with or without their vote and choose to abstain... Other voters who support the strong candidate may decide to jump on the bandwagon and vote where they would otherwise have abstained (2010, 700).

The precise terms adopted here – conversion (over 'vote switching') and mobilisation (over 'turnout') -– not only better avoid this confusion, but also fit better with broader literatures on persuasion effects in public opinion (e.g. Dilliplane 2013, 2011; Kaplan, Park, and Gelman 2012) and allow easier adaptation of the concepts to non-electoral contexts.

To clarify this distinction with an example, take the case of the 2017 French Presidential Election and Emmanuel Macron's *En Marche!* movement. Someone in late 2016 might have wanted to vote for Macron above all the other options, but had no intention to bother turning out because it was overall not a very popular option.[6] Another individual might have been intending to vote for the then more popular Republican candidate, François Fillon. If the later popularity of Macron led these people to vote for him, then the first person underwent a bandwagon mobilisation effect, and the second a bandwagon conversion effect. One was *mobilised* by Macron's popularity, the other was *converted*.

The relevance of this distinction stems largely from how it enables the bandwagon effect to speak to broader trends in political behaviour. If it appears that popularity has the potential to mobilise but not to convert, this potentially demonstrates

[6] Macron polled well through most of 2016, but dipped to around 11% in November of that year (https://pollofpolls.eu/FR/25/presidential-election-2017-1st-round). His percentage of projected vote share had doubled by March 2017.

Table 2.2: Situating the literature in the typology.

	Static	**Dynamic**
Conversion	Abramowitz (1989); Allard (1941); Ansolabehere and Iyengar (1994); Cook and Welch (1940); Dizney and Roskens (1962); Areni, Ferrell, and Wilcox (1998); Arnesen et al. (2017); Atkin (1969); Beniger (1976); Bischoff and Egbert (2013); Ceci and Kain (1982); Cloutier, Nadeau, and Guay (1989); Faas, Mackenrodt, and Schmitt-Beck (2008); Farjam (2020); Gallup and Rae (1940); Glynn and Mcleod (1985); Goidel and Shields (1994); Granberg (1983); Grillo (2017); Henshel and Johnston (1987); Hodgson and Maloney (2013); Hong and Konrad (1998); Kiss and Simonovits (2014); Lanoue and Bowler (1998); Lee (2011); McAllister and Studlar (1991); Meffert et al. (2011); Mehrabian (1998); Morton and Ou (2015); Morton et al. (2015); Morwitz and Pluzinski (1996); Mutz (1992); Myers, Wojcicki, and Aardema (1977); Nadeau, Cloutier, and Guay (1993); Nadeau, Niemi, and Amato (1994); Navazio (1977); Obermaier, Koch, and Baden (2015); Riambau (2015); Riambau (2016); Rothschild and Malhotra (2014); Roy et al. (2015); Schmitt-Beck (1996); Simon (1954); Skalaban (1988); Straffin (1977); Supadhiloke (2015); Toff (2018); Waddell (2017); West (1991); Zech (1975)	Bartels (1987); Bartels (1985); Baumol (1957); Callander (2007); Cloutier, Nadeau, and Guay (1989); Dahlgaard et al. (2017); de Bock (1976); Fleitas (1971); Glynn and Mcleod (1982); Kenney and Rice (1994); Laponce (1966); Marsh (1985); Nadeau, Cloutier, and Guay (1993); Stolwijk, Schuck, and de Vreese (2017); van der Meer, Hakhverdian, and Aaldering (2016)
Mobilisation	Agranov et al. (2018); Ansolabehere and Iyengar (1994); Gartner (1976); Grillo (2017); Henshel and Johnston (1987); Kiss and Simonovits (2014); Morton et al. (2015); Morton and Ou (2015); Panova (2015)	de Bock (1976)

the strength of partisanship or voting habit formation. This might be expected given increased polarisation and 'sorting' of party support (Mason 2018) and the observation that voter volatility is 'bounded' (Tom W. G. van der Meer et al. 2012). However, work correcting for methodological limitations in past studies of conversion has found that information can sway voters to the opposing side more than previously thought (Dilliplane 2013). This could speak to work on partisan dealignment (e.g. Dalton 2013; Dassonneville 2016). The distinction also enables bandwagon research to clarify and frame its theoretical contributions. For example, research demonstrating bandwagon mobilisation effects occurring clearly speaks to rational choice theory on costly voting (Blais 2000). These models would predict that larger candidate leads actually suppress turnout, which should be highest when elections are closest and each voter's possibility of casting a 'pivotal' vote is maximised. The fact that the bandwagon effect, particularly the bandwagon mobilisation effect, appears to hold water in empirical study poses a threat to these models (Agranov et al. 2018). This has even led some to introduce additional terms to explain bandwagon behaviour (e.g. Grillo 2017).

Clearly, the literature has so far favoured the study of bandwagon conversion effects. Of the reviewed articles, 60 have at least some conceptual or empirical focus on conversion while only ten discuss mobilisation (or turnout more generally) – and only three do so *exclusively*. Only one of these discusses mobilisation in dynamic terms (de Bock 1976).[7] In other words, research on the bandwagon effect has broadly not considered the question of whether a party or candidate's increasing popularity over time motivates latent supporters to turn out to vote.

[7] Even here, the bandwagon effect studied is only 'dynamic' in the sense that it is based on comparing the results of a poll with people's pre-existing election predictions.

Among the nine that have considered mobilisation statically, seven are published in economics journals, reflecting a theoretical concern with rational choice and game-theoretic approaches to understanding voter turnout. Only one political science study has empirically tested the possibility of static bandwagon mobilisation effects (Ansolabehere and Iyengar 1994).

This dominance of conversion effects is somewhat at odds with mounting evidence that popularity information matters for turnout. Promising research from economics suggests that this is the case. This is an important finding in the economic literature on public choice due to how it speaks to rational choice models of voting. These generally predict that the *closer* a race is, the more likely people are to turn out, and are therefore directly contradicted by bandwagon mobilisation effects. One economic experiment found that 'the availability of information reduces the probability of minority participation and increases the probability of majority participation' (Agranov et al. 2018, 839). Such effects also emerge in variously specified formal models or games (Gartner 1976; Grillo 2017; Panova 2015). Some recent evidence also suggests that the availability of polling information can stimulate turnout among young voters (Stolwijk and Schuck 2019). In contrast though, more complex forecasting probabilities might depress turnout (Westwood, Messing, and Lelkes 2020).

Static and Dynamic Bandwagon Effects

The second distinction is that the bandwagon effect can be static or dynamic. Someone can be influenced to vote for a party based on its position as the largest party or its improving performance over time (Marsh 1985; Stolwijk, Schuck, and de Vreese 2017; van der Meer, Hakhverdian, and Aaldering 2016). Irwin and Van Holsteyn (2000, 12) note that the size of support for the parties and the change in their support are the two pieces of information typically made available to readers of electoral polls, so disentangling the differences between how voters make use of such pieces of information is of significance to debates about poll publication and regulation. Given concern over the way in which polls are reported and how this might influence voters, the ability to inform regulators about which aspects might have an effect, when emphasised, is arguably important.

As a simple example, consider the 2016 US presidential election. A hypothetical bandwagon voter would likely make two different decisions depending on the type of bandwagon effect she underwent. A static bandwagon effect would see this voter opt for Hillary Clinton; she was consistently projected to obtain more votes than her opponent, Donald Trump. Trump, however, appeared to be making gains on Clinton throughout much of 2016, meaning a dynamic bandwagon effect would lead this voter to opt for him.[8]

Essentially, then, the static/dynamic distinction centres around the *comparisons* voters are taken to be making. In the case of static bandwagon effects, voters are comparing the leading party to all the others, and supporting this party or candidate as a result. In the case of dynamic bandwagon effects, this comparison is instead between the same party or candidate's popularity at two or more time points.

There are reasons to think the prevalence of each type of effect might vary by electoral context. Recent applications of bandwagon research in proportional

[8] The trend line for Trump would appear to be generally positive from mid-2015 onwards (https://www.realclearpolitics.com/epolls/2016/president/us/general_election_trump_vs_clinton-5491.html).

representation contexts have noted, for example, the ambiguity of 'winning and losing... in multiparty systems with coalition governments' (Meffert et al. 2011, 808). This is supported by burgeoning research on voters' senses of 'winning.' This research has demonstrated that prior expectations about election outcomes moderate the otherwise large effect of 'objective performance' (Plescia 2018, 1) and that 'supporters of smaller parties also feel their party won, when and if their party gained votes and seats compared to the previous election' (Stiers, Daoust, and Blais 2018, 21). The key point is that winning is not only defined by securing the largest share of the votes, or the largest share of seats, partly because neither of these factors guarantee a party will even be in government. Van der Meer et al. explain that this point is relevant to bandwagon research, because

> The definition of success in the opinion polls is a point of discussion by itself, especially in a multiparty context... The label "winner" might refer to having support from a majority of the electorate, to having the most support (from a plurality, i.e., being larger than any of the other parties), or to having momentum (i.e., gaining support the fastest). The definition of momentum depends, in turn, on the reference point: A candidate may have lost support over several months, but (re-) gained it more recently over several weeks... In the media cycle leading up to elections, the image of a winner generally relates to this momentum rather than to static size, as changes carry more news value (van der Meer, Hakhverdian, and Aaldering 2016, 49).

This argument, considered elsewhere in proportional representation contexts (Dahlgaard et al. 2017; Irwin and Van Holsteyn 2000; Stolwijk, Schuck, and de Vreese 2017) is worth comparing to the dominance of the static approach, which is arguably more applicable in 'winner-takes-all' contexts (Irwin and Van Holsteyn 2000, 9). The distinction between static and dynamic bandwagon effects is therefore also worth making in that it provides a way of classifying how bandwagon effects operate in different electoral systems, and the extent to which this conforms to this simple proportional/winner-takes-all distinction made so far. However, in order to address this question fully, it is also necessary to measure dynamic bandwagon effects in majoritarian/plurality contexts. This is particularly true given that van der Meer and colleagues' (2016) description above evokes the example, explored extensively in Chapter 1, and returned to in Chapter 4, of the Labour Party's performance in the 2017 UK general election – in a decidedly non-proportional election context, with the vote overwhelmingly dominated by two parties.

There has long been confusion in the literature over whether dynamic bandwagon effects are, indeed, bandwagon effects. This is slowly changing as the work in proportional representation systems cited above becomes more common, but certain studies have rejected the bandwagon hypothesis on the basis of evidence suggestive of a dynamic bandwagon effect (e.g. Atkin 1969). Very similar evidence is presented by other scholars, meanwhile, as supporting the bandwagon hypothesis (e.g. de Bock 1976). A typology that expressly states that dynamic as well as static bandwagon effects are possible can potentially resolve such confusion.

Irwin and Van Holsteyn (2000, 13) have even claimed that 'the momentum-affect [i.e., dynamic] bandwagon is undoubtedly a more useful concept than the size-affect [i.e., static] bandwagon,' because the sheer volume of polls in modern election campaigns, even when public opinion is fairly static, means that the most relevant information contained in them becomes how they vary from each other. Consider, for example, a voter in the UK who sees the the Conservatives are in the lead. How likely is it that this will have an effect on her vote choice, given how likely it is that she already knew the Conservatives were the largest party? In this vein, Atkin (1969, 516) notes that 'any new perception must be considered relative to the voter's previous estimate… as the magnitude of the electorate's support for a political object may be larger or smaller than expected.' From this perspective it seems more likely that, if voters are making decisions based on popularity, they will be more likely to respond to dynamic changes in the polls, rather than the static state of affairs. This is bolstered by the fact that polls suggesting a party's performance has changed over time get more attention from journalists and commentators (Larsen and Fazekas 2021). These considerations make it worth incorporating dynamic effects explicitly into how we think about the bandwagon effect.

A Theoretical Model of the Bandwagon Effect

Having analysed the bandwagon effect conceptually (RQ_1), I now turn to how it can be represented as a causal model and how the parts of this model can be explained in theoretical terms (RQ_2). I use 'directed acyclic graphs' (McElreath 2020; Pearl 2000; Rohrer 2018) – causal diagrams – to present this model, as introduced in Chapter 1. The discussion breaks the model down into component relationships and layers these onto each other one-by-one, drawing extensively on theory and existing knowledge in order to justify each of these additions and explain why they might arise in voting behaviour. Throughout the discussion I also explicitly define any potentially ambiguous terms that were not already outlined in the conceptual discussion above, but will be used throughout the rest of the thesis.

Popularity Information → Vote

The basic theoretical claim at the heart of the bandwagon effect can be depicted as in Figure 2.1. The bandwagon effect requires that voters are *motivated by* the popularity of a party or candidate. This just means that it *causes* a change in their voting behaviour. But voters cannot be motivated by something that they never know about. In this sense, then, popularity immediately reveals itself as simply a type of information. As established in Chapter 1, thinking about the bandwagon effect in such informational terms makes sense in historical perspective. I make this informational approach clear by referring to effects of ‘popularity information’:

> **Popularity information**: any information to which a person might be exposed that tells her how much electoral support there is for candidates or parties, and how this is developing over time. It has both a static – the level of support for each option, right now, relative to the others – and a dynamic – the change in level of support for a given option relative to the past – dimension.

We usually think of popularity information in terms of the results of opinion polls, because they offer clear information about the popularity of opinions, candidates or parties. The bandwagon effect happens when polls tell someone that a party is popular and then she switches to vote for that party, for example. Yet people learn about the popularity of parties or candidates through other means, such as general news media or the balance of opinion among their friends and family (see Zerback, Reinemann, and Nienierza 2015). Popularity information could also encompass new types of cues in the modern media environment: numbers of likes, shares, and follows (Larsen and Fazekas 2021). Any information that has a direct bearing on a rational understanding of how much support there is for candidates or parties is popularity information.[9] The bandwagon effect happens when such information brings about changes in voting behaviour, in favour of popular parties or candidates.

$$Popularity\ information \longrightarrow Vote$$

Figure 2.1: Basic causal model of the bandwagon effect: popularity information affects voting behaviour.

Yet, as already discussed, the popularity information voters receive can be both static and dynamic. They learn not only about how popular the parties are right now, relative to each other, but often also how a given party's popularity is changing over time. Moreover, the actual change in voting behaviour that signals a bandwagon effect might be either a switch to vote for a different party, or a decision to turn out to vote for a party one already supports – voters can be converted or mobilised. In line with the discussion above, then, Figure 2.2 recognises these two distinctions.

To understand why these effects of popularity information are possible, I draw on political psychological knowledge about how voters use information in deciding how to vote. In their pioneering study of these questions, Lau and Redlawsk (2006) state that voters cannot possibly find, absorb, and recall all of the information available to them in making a voting decision. As human beings with competing demands on their time, voters cannot be omniscient. How much information they are likely to gather, and how they engage with it, depends instead on how difficult

[9] By this I mean that a perfect 'Bayesian' learner (Hill 2017) would update her beliefs – however marginally – about the level of support for the party or candidate in response to hearing this information. In other words, a rational response to hearing the information is to incorporate it into how much support you think there is for the party or candidate.

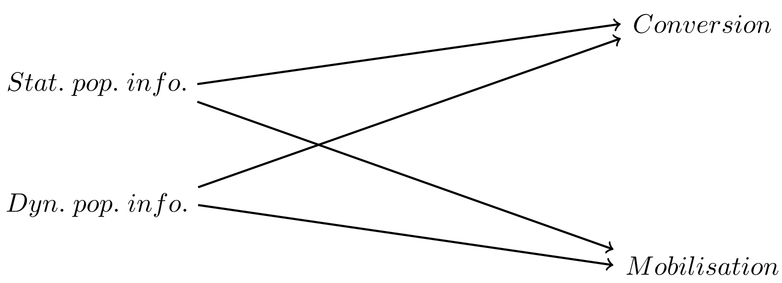


Figure 2.2: Simplified causal model of the relationship between static and dynamic popularity, and vote choice and turnout.

the task of making the decision seems to be. Contextual and individual factors determine this perceived difficulty, as

> Voter characteristics combine with the decision environment… and the level of sophistication to generate a perceived sense of the ease or difficulty of the choice. This perceived nature of the decision task then influences the information processing strategies the voters use to make sense of a potentially chaotic information environment. In general, when the task is perceived to be especially difficult – with multiple candidates, little differentiation between choices, and/or disproportionate resource allocations – many voters employ simplifying strategies that result in unbalanced and/or shallow search… When the task at hand is simple, with only two candidates, for example, it becomes a more reasonable proposition for voters to study all the options in depth (Lau and Redlawsk 2006, 257).

The ‘simplifying strategies’ used in difficult tasks are also known as ‘heuristics’ and some of the earliest comprehensive study of voting behaviour noted that voters use them (Berelson, Lazarsfeld, and McPhee 1954). Page and Shapiro (1992, 28) explain that such heuristics 'permit voters to act as if they had all the available information.' Lau and Redlawsk refer to popularity information as part of one such heuristic, known as 'viability,' which emerges when

> Poll results tell voters which candidates are ahead in a campaign and which are hopelessly behind and could never win… Reducing the choice set from four candidates to two, say, immediately provides a 50% reduction in the amount of information that must be processed. Moreover, seeing a candidate leading in the polls provides a type of 'consensus information' that could motivate a voter who had previously rejected or ignored a candidate to more closely consider that alternative (2006, 257).

Popularity information provides a simple cue indicating which candidates are worth considering, by revealing which candidates other people are seriously considering. This would indeed be expected to push people towards favouring more popular options in response to popularity information, creating a bandwagon effect.

Lau and Redlawsk demonstrate that voters rely on such heuristics in intraparty contexts, such as presidential primaries. Here, voters cannot easily be guided by their party preferences that would otherwise give them their 'default' option. As Mutz (1998, 194) puts it, 'the absence of other cues – such as political parties – also encourages people to turn to mass opinion for guidance.' This, arguably, makes such contexts ripe for bandwagon effects (Abramowitz 1987, 50; Gimpel and Harvey 1997, 158). These contexts are particularly difficult when there are lots of candidates. In the 2020 Democratic primaries in the USA, voters in many states were faced with a large pool of candidates. Learning lots of information about all of them is a very demanding task. The viability heuristic provides a simple way to set bounds on how many of these candidates a voter needs or bothers to think about, by focusing attention only on those popular enough to stand any chance of winning – in this case, the likes of Joe Biden and Bernie Sanders. As such, candidates who are polling well are likely to attract voters who employ this heuristic. This leads to a clustering of voters among the most popular candidates, with their distribution likely determined by the other information they consider about these candidates. For example, while a Democratic voter might employ the viability heuristic to make her vote decision easier in the primary election, limiting her pool of potential options to Joe Biden and Bernie Sanders, she will then actually engage with enough further information about this smaller pool of candidates to work out which of the two she wants to back. Utych and Kam (2014) demonstrate this tendency only to learn about 'viable' candidates, both in experiments and in observational data based on linking Republican primary candidates' polling performance to Google searches of their names. This idea has also been likened to the so-called 'paradox of choice' (Cunow et al. 2021, 9), linking these observations to broader research and theory on human behaviour (e.g. Efferson et al. 2008; Muthukrishna and Henrich 2019).

It is worth noting that the viability heuristic could explain what happens in any type of bandwagon effect that a voter undergoes in these difficult elections: it can be applied to static and dynamic, and conversion and mobilisation effects. Viability is typically thought to be based on static popularity, and this is consistent with how Lau and Redlawsk suggest voters use polls above. But there is nothing stopping these viability considerations from being dynamic. Perhaps dynamic popularity can put candidates on people's radar by suggesting that those candidates are on their way to becoming viable, or are becoming more viable. This would be consistent with how presidential nomination campaigns are often seen to be dynamic (Bartels 1987; Callander 2007). It is also consistent with the findings of Chapter 4. As well as bringing about conversion by pointing voters towards certain candidates rather than others, as suggested by the framing above, viability heuristics might also bring about mobilisation. A supporter who has enough information to know which candidate she prefers, but is not intending to vote for whatever reason, might change her mind about turning out as a result of paying attention to the viable candidates, one of whom is her preferred option. This might not have happened if her preferred option stood no chance. The viability heuristic can explain more than just static bandwagon conversion effects.

This contrasts with what can happen in elections that are easier for voters. Lau and Redlawsk (2006) use the example of a US presidential election, in which there are two candidates belonging to different parties. Voters are likely to have relatively well-formed ideas about these parties. They 'typically know – or quickly learn – that there are two candidates who stake out positions more-or-less consistent with the history of their respective parties' (Bowler and Donovan 1998, 24). The small number of options, with clear partisan cues, makes the decision task relatively easy. In this context, it should not be beyond voters' capacities to engage with quite a lot of information about the candidates, and polls – or popularity information more generally – are one form of information among many that voters consult.

The perceived difficulty of the decision task, and thereby the way in which voters engage with information, also depends on characteristics of the individual voter. In particular, more politically knowledgeable or sophisticated voters will generally tend to be more likely to engage more deeply with information, whereas less politically sophisticated voters will be more likely to find the decision task difficult and therefore employ heuristics. Mutz (1998) explains this in more depth. She argues that 'an elite, highly politically involved segment of the citizenry' (Mutz 1998, 216) might change their voting behaviour *strategically* in response to popularity information.[10] As argued above, it is important to recognise that strategic voting and bandwagon effects are two different things. For most other people, though, if their attitudes or behaviours change in response to popularity information, this will follow the logic set out above by Lau and Redlawsk (2006), in that

> The highly unknowledgeable... are likely to rely on consensus information as a heuristic cue under conditions of low knowledge and high uncertainty. Still others may be influenced by representations of mass opinion because this information prompts them to think about the reasons that may have led all those other people to hold that particular view. The segment of the population most susceptible to this mechanism is the group in the middle with respect to political knowledge and involvement. These people will possess the minimal levels of information necessary to rehearse cognitive responses, yet they will not be so involved as to be precommitted to particular views (Mutz 1998, 216)

It is not immediately obvious why deeper, more sophisticated engagement with popularity information would lead anyone to vote for the candidates that are doing

[10] Mutz (1997, 105) cites Abramson et al. (1992) to suggest that strategic voting occurs 'around 10% of the time.' More recent evidence would however suggest that, in certain contexts, around a third of people vote strategically *when they have the chance to do so* (Alvarez and Nagler 2000; Eggers and Vivyan 2020).

well. Here though, Mutz has hinted at a mechanism through which this can happen. She further explains, directly applying this logic to the case of vote choice, that

> When people hear that more and more people are rushing to support a particular candidate, they may mentally rehearse the possible reasons that people might have that would lead them to support such a candidate. If a person is not strongly committed to a particular view, the thoughts that then pass through the individual's mind are most likely to be supportive of the candidate. By means of the positive thoughts generated by information about mass candidate support, information about mass opinion influences the development of the person's own viewpoint; arguments in favor of the candidate are now more salient and well rehearsed in the person's mind than those in opposition (Mutz 1998, 212).

Those who go beyond simple heuristics in their decision-making, when they encounter popularity information such as poll results, should be stimulated to reflect further on the information they aware of that would justify these results. This may result in changes in voting behaviour if the reflection gives voters reason to change their mind. This is called the cognitive response mechanism. Importantly, it serves to emphasise a point alluded to in Chapter 1: popularity does not come about in a vacuum. To gain electoral traction, candidates obviously need political arguments that appeal to some voters. The popularity this gains them can itself win them further support, through a bandwagon effect, but this again depends on the arguments themselves. Below I discuss how this fact ends up meaning that people's pre-existing preferences significantly constrain, or moderate, the bandwagon effect.

Where Mutz emphasises that people would ‘rehearse’ the arguments they are already aware of, contemporary models of political psychology suggest that exposure to information that is incongruent with a voter’s beliefs – for example, a poll showing a different party in the lead – can stimulate an external, as well as internal, information *search* in order to understand this unwanted information (Lodge and Taber 2013, 220; see also Marcus, Neuman, and MacKuen 2000). When popularity information affects voters’ decision-making, this is not only because it encourages them to reflect more on the information they already have to explain why certain parties or candidates enjoy a lot of support, but also because it stimulates them actively to acquire, or pay greater attention to, more such information.

Again, it is worth establishing that this process could potentially come into play with any type of bandwagon effect. It could arise with static bandwagon effects where people engage in a cognitive response in order to understand why a candidate is in the lead. It could arise with a dynamic bandwagon effect where, consistent with Mutz’s phrasing above, people undergo a cognitive response in order to understand why ‘more and more people’ are supporting a candidate. Bandwagon conversion effects happen when such responses lead to a voter changing the candidate they are going to vote for. Bandwagon mobilisation effects happen when the cognitive response instead involves a person reflecting on the arguments they and/or others have for supporting a candidate and being motivated by these considerations to turn out and vote for the candidate. In every case, popularity information is creating a cognitive response, which ends up with a person switching or turning out to vote for a more or increasingly popular candidate or party on the basis of its popularity.

Preferences → Vote

But of course, voters are not constantly switching their voting intentions on the basis of things like what the polls say. On the contrary, it can take a lot to convert or mobilise voters. The processes described above already reveal a big part of the reason for this: people's existing preferences and predispositions. These, of course, have a broader meaning in terms of political attitudes, but here I define them specifically with reference to electoral candidates or parties:

> **Preferences**: a voter's (pre-existing) attitudes towards political parties or candidates.

For the most part here, I use 'predispositions' interchangeably with 'preferences.' This avoids over-complicating my argument by constantly distinguishing between the two concepts in ways that do not alter my conclusions. Later, when discussing the use of electoral expectations to measure the bandwagon effect, I distinguish between predispositions and preferences as follows: predispositions are underlying, unobserved pre-existing attitudes, whereas preferences are measured attitudes that are designed to capture these predispositions, or key aspects of them. For example, a voter is predisposed in that she is motivated to draw certain conclusions about politics (Zaller 1992, 6). These *predispositions* are not necessarily tied to any single measurable sentiment or attitude, but capture a general disposition or motivation. A key part of this sentiment, though, might be captured by her evaluations of the parties. Her answers to questions about these are part of her (partisan) *preferences* – convenient measurements that capture aspects of her predispositions that we care about. I use the term 'preferences' more in the present discussion because it is more widespread in bandwagon research.

My non-prescriptive definition of preferences also reflects the fact that the concept is not consistently tied to any particular measure or survey item in bandwagon research (e.g. Bartels 1985; Granberg and Brent 1983). For most measures of how much people 'like' or how well they 'rate' parties – such as feeling thermometers or other evaluation scales – the particular way in which preferences are operationalised is unlikely to change how applicable my arguments are, providing the measure is itself valid. However, arguably, my arguments might be weaker when party preferences are measured as 'party identification.' I deal with this objection further below, and again in Chapter 5 where it becomes most relevant.

It is worth clarifying the distinction between preferences and popularity information. Imagine a voter who is going about her life during an election campaign. She has attitudes about the parties contesting that election: she likes some more than others. These are her preferences. She sees a poll telling her the estimated vote shares of the parties. This is popularity information. The causal model proposed here is at this individual level, which serves to separate preferences and popularity in this way. Popularity information could in many cases be seen as reflecting an aggregation of preferences. But, at the individual level, one (preferences) is something the voter brings to the table, while the other (popularity) is information that she finds there.

The discussion of viability and the cognitive response above describes the situation in which voters' choices *do* change. Taking preferences into account reveals how such changes are constrained. Voters rarely go into an election as a blank slate, ready to be convinced to vote for any candidate, with no preconceived ideas about which ones they like or dislike. Indeed, it is precisely in the cases where they lack such preferences that popularity information might provide a useful heuristic, because lacking pre-existing preferences makes decisions hard. It is also worth noting that high levels of political sophistication are associated with holding strong preferences, meaning that those who lack preferences at a given election are also likely to be less politically sophisticated, making the task even more difficult (Lau and Redlawsk 2006, 123; Zaller 1992). But this also means that when voters consult popularity information beyond using it heuristically, they typically do so with predefined ideas about which party or parties they like. This drastically limits the extent to which we should expect them to be susceptible to change their vote choice through a cognitive response. In sum, preferences *moderate* the bandwagon effect.

To see why this is, recall Mutz's (1998, 212) point above that when someone learns that a candidate is polling well, 'if a person is not strongly committed to a particular view, the thoughts that then pass through the individual's mind are most likely to be supportive of the candidate.' This means in contrast that if someone comes with predispositions, these thoughts might not be so positive. Zaller (1992) offers perhaps the clearest explanation of why this is the case. When making political decisions or expressing opinions, people 'call to mind a sample of... ideas, including an oversample of ideas made salient by the questionnaire and other recent events, and use them to choose among the options offered' (Zaller and Feldman 1992, 580). His Receive-Accept-Sample (RAS) model (1992, 58) includes the 'resistance axiom,' which states that 'people tend to resist arguments that are inconsistent with their political predispositions.'

This process has become known as motivated reasoning: people have a 'tendency to seek out information that confirms prior beliefs... to view evidence consistent with prior opinions as more relevant and stronger... and to spend more time resisting arguments inconsistent with prior opinions regardless of their objective merit' (Druckman 2014, 471). Lodge and Taber (2013, 149) put it that 'citizens are prone to accept those facts and arguments they agree with and discount or actively counterargue those that challenge their preconvictions.' Arguably, these kinds of tendencies may have become increasingly significant in a new media environment in which the flow of information to individuals is increasingly 'curated' both by them, and by the platforms used by those attempting to persuade them (Thorson and Wells 2015). Crucially, the result is that opinions are a 'marriage of information and predisposition' (Zaller 1992, 6). The considerations and ideas people draw on in order to form opinions are a biased selection of the information that could have presented itself to them in their lives, because they factor in more of the information that is congenial to their preferences.

Therefore, when people are confronted with a poll that tells them a candidate they are not currently intending to vote for is in the lead, the arguments they call to mind when thinking about why this candidate is doing so well might either be insufficient or unconvincing, if they have engaged in motivated reasoning. Some evidence even suggests that these arguments can at times be so unconvincing as to further entrench a person's existing view (Petty and Cacioppo 1996). This is a vicious (or perhaps virtuous) circle, because this very reaction is itself a form of motivated reasoning. Motivated reasoning also means that people are less likely to encounter information that is incongruent with their preferences in the first place – for example, people will generally seek out polls that are likely to report positive results for their preferred candidates or parties. This makes it highly unlikely that they would jump on the candidate's bandwagon. The weaker their preferences are, or the more abundant and convincing the information in support of other candidates is, the more likely a bandwagon effect would overwhelm this resistance.

Political preferences therefore have a crucial place in the bandwagon effect. On the one hand, they are likely to play a role in determining whether people use popularity information as a heuristic in their decision-making. Those with strong party preferences should find decision tasks easier, giving them less reason to fall back on heuristics.[11] On the other hand, they should also be crucial in shaping the considerations in people's minds when they reflect on why a certain candidate or party is popular – the very reflections that might determine whether or not they switch to support that option. Whichever way the bandwagon effect comes about, preferences matter, as a moderator. Figure 2.3 therefore builds these preferences into the theoretical model explicitly.[12]

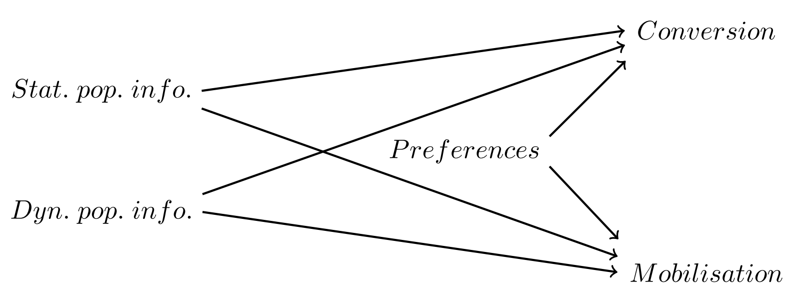


Figure 2.3: Simplified causal model of the relationship between static and dynamic popularity, preferences, and vote choice and turnout.

[11] Another way to frame this is that they fall back on what is essentially a partisanship heuristic, rather than falling back on other heuristics (Bowler and Donovan 1998, 29).

[12] The moderating relationship between popularity information and political preferences could be described as an 'interaction': the effect of one depends on the other. This is not displayed specifically in the causal model, because causal diagrams 'do not specify how variables combine to influence other variables' (McElreath 2020, 240). They are 'non-parametric,' simply telling us that there is some function of the variables that influences other variables (Pearl 2010).

Popularity Information → Preferences

Preferences have a further role to play though, in that as well as having an important effect on vote choice, they should also be seen as affected by popularity information in the bandwagon framework. If the bandwagon effect assumes that voters' decision-making could be affected by, for example, what the polls say, it would be questionable to assume that this popularity information cannot also affect which parties or candidates they like, or how they rate them when asked. That is, the bandwagon effect is consistent with an effect of popularity information on preferences. This comes back to a distinction made in the conceptual discussion above between misaligned voting and realignment. On the one hand, Blais and colleagues, as discussed above, suggest that this effect on preferences is the defining factor in the bandwagon effect, because

> If there is a contagion effect, voters should come to evaluate the parties and the leaders who are doing well in the polls more positively and those who are not doing well more negatively. If all of the polls' effect is accounted for by contagion, we should find that polls had an impact on ratings of parties and of leaders and that it is only because of these evaluations that they affected the vote, that is, the polls had no independent effect on the vote once preferences were taken into account (2006, 270).[13]

Bandwagon effects, on this view, are entirely mediated by changes in voters' evaluations of political parties and candidates. They necessarily involve a realignment

[13] Even though the authors refer here to a 'contagion effect,' elsewhere in the chapter they draw an equivalence between this and the bandwagon effect (e.g. Blais, Gidengil, and Nevitte 2006, 264).

of preferences. In contrast, some work either theoretically labels the bandwagon effect 'misaligned voting' (Evrenk and Sher 2015) or, empirically, operationalises the bandwagon effect as a vote for a non-preferred option (Lanoue and Bowler 1998). I noted above that a simple solution to this disagreement is to allow both of these to be a bandwagon effect. What matters, in the definition of a bandwagon effect, is that popularity information persuades people to *vote* for a more popular option, regardless of whether this is a misalignment or realignment.

This, however, means that when modelling the bandwagon effect, we must allow for both of these possibilities, rather than ruling either out. It is therefore necessary to factor in the effect of popularity information on preferences, because otherwise the model does not capture the possibility of realignment, which suggests that the preferences *mediate* the bandwagon effect (VanderWeele 2015). This raises the question of what the theoretical basis is for assuming that bandwagon effects can be mediated by preferences, beyond just sticking to earlier conceptual commitments. The answer to this is that both of the theoretical causal mechanisms discussed above – the viability heuristic and the cognitive response – suggest that preferences can mediate the bandwagon effect.

Consider first the case of viability. A voter, facing the difficult choice presented by the 2020 Democratic primary, restricts her range of options to the viable candidates, Biden and Sanders. For all intents and purposes, she only learns about these two candidates. She ends up voting for one of them – say, Biden. It is unreasonable to assume that what she learnt about Biden definitely did not make her like him more or prefer him over the other options, given that she ended up voting for him. Popularity information led to her learning about Biden. What she learned convinced her he was the best option. Then she voted for him. This looks like her

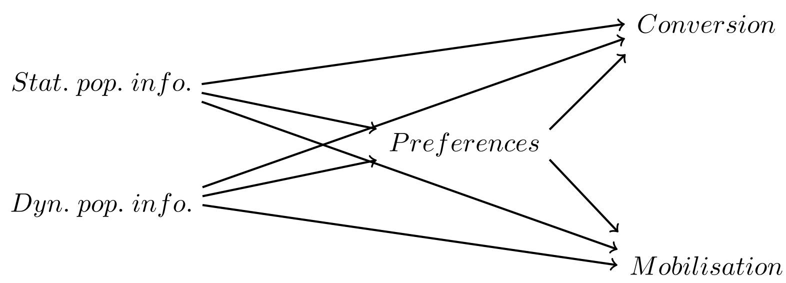


Figure 2.4: Causal model of the relationship between static and dynamic popularity, preferences, and vote choice and turnout.

preferences have mediated the bandwagon effect. This might not have happened, but it is implausible to rule it out *a priori*. Another way to see this is to imagine the reverse scenario, in which the viability mechanism makes this voter consider Biden as a candidate, but then everything she learns about him she does not like – we would not expect her then to vote for him, but we would expect her to lower her evaluation of him. This is just another, negative, instance of preferences mediating the bandwagon effect. A causal account of the bandwagon effect that wants to allow for such viability processes should therefore incorporate the possibility that the bandwagon effect is mediated by preferences.

Second, consider the cognitive response mechanism. Recall that this involves voters receiving popularity information, and then reflecting on, or searching for, reasons people might have for supporting those candidates or parties that are doing well in the polls. In many cases, it is likely that should these reasons convince the voter to support this candidate or party, it will be because they convinced her that this candidate is better than she previously thought. Perhaps in some instances the reasons will be more instrumental or purposive, and not serve to persuade her about the candidate's political merits, but this is unlikely to represent the majority of cases. Building the mediated bandwagon effect into the model, through preferences, simply means refusing to rule out the possibility that the arguments people consider in order to understand new popularity information – arguments which then convince them to vote for a popular candidate or party – could possibly convince them to like that candidate or party more, identify with it more, or evaluate it higher, along the way. Again, the negative case helps to clarify this: if a voter engages in an information search in order to understand why a party is so popular, she might end up encountering arguments that are very negative and speak to issues she cares about. We would not expect her to vote for the party, but we would expect her to revise her estimations of it. This, again, would be preferences mediating the bandwagon effect, and is a direct implication of the cognitive response.

The diagram in Figure 2.4 therefore incorporates the mediating role of preferences in the bandwagon effect. This means that popularity information can affect preferences, which in turn can affect vote choice. The direct arrows from popularity information to vote choice mean, however, that the mediated part of the effect does not necessarily tell the whole story. Maybe, for some, a bandwagon vote is truly misaligned. Both options count.

It is worth acknowledging, however, that a large body of research would suggest that if preferences are defined as 'party identification' – a measure of the party people feel closest to, or identify with – then my model suggests preferences are more labile than they really are. Since the publication of *The American Voter* (Campbell et al. 1960), many in political science have maintained that party identification develops at a young age and remains quite stable throughout voters' lives (Achen and Bartels 2015; Green, Palmquist, and Schickler 2002; Huddy, Mason, and Aarøe 2015) with some even considering it similar to demographic characteristics in its stability (Lewis-Beck et al. 2008). On this view, it is unlikely to be swayed by 'short-term forces' (Howell 1981) such as recent polling. However, if this understanding of party identity is correct on the whole, then rather than showing that the model presented here is false, it just helps to explain why bandwagon effects are small. It is quite widely acknowledged – and consistent with the claim that party preferences moderate and mediate the bandwagon effect – that bandwagon effects are likely to be more common among those with the loosest party attachments (Bartels 1987; Hardmeier 2008; Mehrabian 1998). Those who have strong partisan identities are not the kinds of people who tend to undergo bandwagon effects. This does not mean that parties people identify with most are not able to change as part of the bandwagon effect; it means that this is very unlikely for those for whom this identity is strong. The implication of this is that when modelling and studying the bandwagon effect, we might want to allow for the possibility that this is mediated by changes in party identity – even though such identities are stable on average – because party identity might mediate the effect among those to whom the bandwagon effect is most likely to actually happen. These are likely to be the people to whom the argument that party identification is a stable, unwavering force is least applicable. Chapter 5 returns to this discussion.

A final point to re-emphasise, before turning to questions of measurement, is that this theoretical discussion is very zoomed-in. It focuses on a relationship between information and voting behaviour. This is important to note, first, because this relationship does not capture the whole picture, in terms of a voter's information environment and decision-making. For example, the model does not account for where the information comes from. But of course, in order for popularity information to present itself, something has to happen to make a party or candidate popular in the first place – as noted in some of the examples in Chapter 1. Voters also have to receive this information somehow, a part of the process that is not automatic, especially considering that people pay more attention to information that says what they want to hear. In many cases, even when they do receive the information, I have tried to clarify above that this does not necessarily mean they will undergo a bandwagon effect in anything like a deterministic sense. In short, the model mostly tells the story of what happens when a voter *does* receive popularity information and it *does* motivate them to change their voting behaviour. It is not the whole causal story of how voters decide.

The second reason it is important to emphasise the scope of this model is because it leaves out other ways in which popularity information can end up affecting the vote. In particular, polls that voters never even see are known to play a role in determining election outcomes by influencing political donations, or guiding campaign spending and resource allocation. Henshel and Johnston (1987) summarise these potential 'indirect' processes, which I return to in the conclusion of this thesis. The reason I do not address them here is because they are not part of the bandwagon effect *per se*. The bandwagon effect refers to the specific process of voters responding to popularity information that they receive by voting for popular parties or candidates. It is nonetheless worth noting that I do not mean to rule out this additional, likely very consequential, role of popularity information in the democratic process. Indeed, I will note in Chapter 6 that future research on the bandwagon effect would do well to consider it alongside or as a specific part of this broader phenomenon.

Approaches to Studying the Bandwagon Effect

Given the causal structure set out above, how is it possible to measure the bandwagon effect? Abstractly, Figure 2.4 tells us that for each type of bandwagon effect, the total effect can be estimated by measuring the statistical association between a given type of popularity information and the outcome. For example, if we measured the association between voters' static popularity information and their vote-switching, we would estimate the total static bandwagon conversion effect captured in the bold arrows in Figure 2.5.

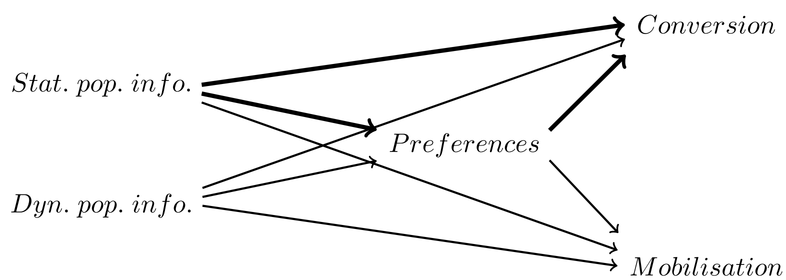


Figure 2.5: Causal model of the relationship between static and dynamic popularity, preferences, and vote choice and turnout, highlighting static bandwagon conversion effect.

The problem is that there is no obvious way to capture a given person's popularity information. Short of engaging in very invasive observation in which, for example, a sample of voters' internet usage is tracked throughout an election campaign to obtain fine-grained coverage of all the information they have come across online, it is very difficult to know with any accuracy what information someone has at her disposal. For example, how can a researcher know with any certainty which polls every survey respondent has seen?[14] For all intents and purposes, popularity information is typically *unobserved*. This means that researchers cannot simply go out into the world and observe relationships between popularity information and vote choice.

Faced with this difficulty, two dominant approaches have emerged in the empirical literature: one observational (using normal survey questions), the other experimental (using randomised trials). Briefly, the observational approach involves taking data on people's electoral expectations and attempting to isolate a statistical association between the information contained within these opinions and their vote choice. The experimental approach uses randomised controlled trials to control directly what popularity information people see, and measuring the differences in vote choice between those with and without polling information. I explore the logic of these approaches more thoroughly below, directly setting up the empirical chapters (3, 4 and 5) of the thesis in the process.

Expectations and the Bandwagon Effect

Although we do not know exactly what information voters have access to, we can attempt to get at this indirectly. One way to do this is via their attitudes, opinions, or beliefs. We can attempt to glean what information they have in mind when answering questions about these. In the case of the bandwagon effect, we could

[14] Some have tried to measure the bandwagon effect by isolating which polls a given voter is likely to have seen based either on the timing of the interview (Blais, Gidengil, and Nevitte 2006) or self-reported media consumption combined with content analysis (Stolwijk, Schuck, and de Vreese 2017), but this is complex and potentially introduces many forms of unobserved confounding. Beyond this, self-reported media consumption correlates only modestly with real, 'logged' media use patterns, casting doubt on the validity of such approaches (Parry et al. 2021).

ask people about their *electoral expectations*:

> **(Electoral) expectations**: a voter's opinions about the likely performance of political parties or candidates at a future election.

This definition of expectations shows that they are likely to have a close affinity with popularity information. Candidates win elections by being popular. When working out who we expect to win elections, we would therefore be wise to pay attention to information about that popularity. This means that when we ask people their expectations, their answers reveal something about the popularity information they are likely to have encountered. Consider that the polls are typically directed towards some election, either hypothetical or real. They might be conducted during a campaign, about the upcoming election. Outside of campaigns, they will typically ask voters how they would vote if an election were held tomorrow. The relevance of this information is usually in its implications for what will or would happen in such an election – even though forecasters are keen to resist this immediate association (Gelman 2013). It is therefore quite natural to think of expectations as a kind of attitudinal counterpart to popularity information. Popularity information tells *us* how parties are likely to perform at an election. Expectations are how *we* tell *others* how they are likely to perform – in our opinion.[15]

Specific instances of expectations will diverge from this close affinity with popularity information, because they rely on other informational considerations. For

[15] Note that this definition does not imply any specific survey item or operationalisation. Expectations could be tapped into with a question about which prime minister will lead the government after an election, how many votes each party will receive, which party will be the largest in government, etc. For this reason, approaches to measuring expectations that capture their uncertainty and potentially multiple dimensions may be the gold-standard for measuring this concept (Leemann, Stoetzer, and Traunmüller 2021).

example, coalition expectations in multiparty systems also draw on cues such as which parties are willing to negotiate with each other (Gschwend, Meffert, and Stoetzer 2017). Yet even in these cases, popularity information affects expectations (Bowler, McElroy, and Müller 2021). Because the bandwagon effect is concerned with how popularity information in particular influences voting behaviour, the role of other informational cues is not of central interest here, so it is left out of my models.

As noted above, while they take information into account, opinions are also based on ‘predispositions’ (Zaller 1992). Recall that each opinion, expressed in a survey or interview, is the result of the respondent calling to mind relevant ‘considerations’ from which to form an answer. These ‘considerations’ are biased, owing to motivated reasoning. As such, the opinion constructed from them is also biased. Opinions are not an accurate reflection of voters’ full information environment, but instead reflect their biased reading of that environment. Different people will form different opinions when given the same information.

This model can quite easily be applied to electoral expectations. Expectations are opinions. In the formation of these opinions, the process above becomes what is known as ‘wishful thinking’: voters express expectations that are more in line with their party preferences (Babad 1997; Bartels 1985; Bowler, McElroy, and Müller 2021; Lanoue and Bowler 1998; Lazarsfeld, Berelson, and Gaudet 1948; Meffert and Gschwend 2011; Stiers and Dassonneville 2018). This means that, in reality, asking a voter questions about the likely performance of parties will not lead to an answer which reflects her information environment accurately, but rather she will generally give more weight to any information suggesting her preferred party is more popular, and the answer will reflect this. Some even claim that wishful thinking is the *main* driver of expectations (Mongrain 2021).

> **Wishful thinking**: the tendency of voters to think or say they think, *ceteris paribus*, that their (less-) preferred parties are going to perform better (worse) at an election.

These relationships between party preferences, popularity information and expectations indicate that expectations could be introduced to the causal model presented above, as in Figure 2.6.

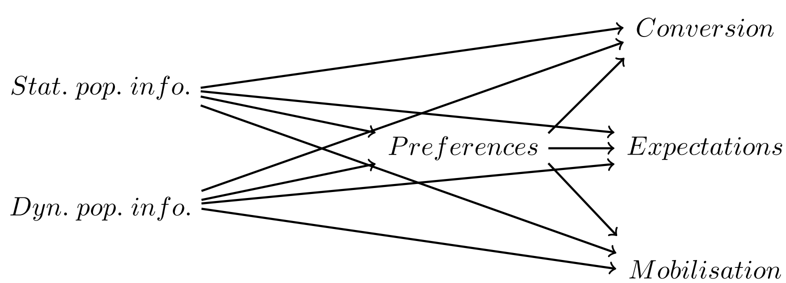


Figure 2.6: Causal model of the relationship between static and dynamic popularity, preferences, expectations and vote choice and turnout.

A first point to note here is that I have specified that both static and dynamic popularity information affect expectations. Expectations are about the *future* (Zerback, Reinemann, and Nienierza 2015). Scholars have demonstrated that static information such as what the polls are saying right now influences these expectations (e.g. Blais and Bodet 2006; Irwin 2002). However, dynamic popularity information – for instance, cues in the polls about how parties' popularity is changing over time – provides a basis on which to link this present state of affairs to the future. This kind of reasoning can be seen in sports (Arkes 2011), including in the form of the so-called 'hot hand fallacy' (Ayton and Fischer 2004). Recent social psychological research also finds that people apply dynamic considerations to how they predict future behaviours (Mortensen et al. 2019; Sparkman and Walton 2019). This dynamic effect has, nonetheless, not been explored seriously in empirical research on electoral expectations. I therefore incorporate it here on the basis of its theoretical plausibility, and provide extensive experimental evidence of the relationship in Chapter 4.

Beyond this, Figure 2.6 reveals the logic behind the 'expectations approach' to studying the bandwagon effect. Because expectations are in a sense 'downstream' of popularity information, they serve as a proxy for it. Voters' expectations contain clues about what popularity information they are likely to have encountered, because they factor this information into their opinion. However, the relationship between expectations and voting outcomes is 'confounded.' That is, they share a common cause. With both conversion and mobilisation, expectations share the causal impact of preferences. People vote for parties they like, *and* they have higher expectations for those parties. This means that the relationship between expectations and these outcomes will be strongly biased upwards, giving the spurious impression of a large effect that is not at all due to the bandwagon effect it is meant to measure. This demonstrates the problem wishful thinking poses for this approach and why researchers using this approach always make adjustments for party preferences, in order to remove this confounding bias.

In Chapter 5, I argue that this approach, in its common form, does not typically work. The reason it does not work is already apparent on proper scrutiny of Figure 2.6. Briefly, the act of adjusting for preferences here closes off the causal path flowing from popularity information to the outcomes of conversion and mobilisation, meaning that these analyses will never capture the total effect of this information. They remove the possibility of preferences mediating the bandwagon effect. They fail to measure the full bandwagon effect by design. The approach also risks confusing static and dynamic effects, or suppressing either, because expectations are affected by both static and dynamic popularity information. I explain this in more depth in Chapter 5, and propose a new approach that uses panel data to overcome these problems in measuring the bandwagon effect.

Studying the Bandwagon Effect Experimentally

Alternatively, experiments or 'randomised controlled trials' offer a way to vary popularity exogenously, such that its effect on voting behaviour can be estimated without bias. Rather than relying on an attitudinal proxy that introduces problems of confounding, voters can simply receive polling information at random, and differences in their behaviour can be observed. If people who randomly receive information suggesting a candidate is in the lead are more likely to vote for the candidate than those without the information, this would be evidence of a bandwagon effect. This demonstrates how experimental designs enhance internal validity. Indeed, researchers employing this method to study the bandwagon effect, by presenting respondents with randomly varying news reports, have noted that

> Because the news program was identical in all respects except for the poll-related story, and because participants were randomly assigned to experimental conditions, any differences between conditions in the dependent measures can be attributed only to the poll results (Ansolabehere and Iyengar 1994, 415–16).

This point is central to much bandwagon research. It has even been asserted that proof of the effect's existence *must* be based on an experimental schema (Arnesen et al. 2017, 741; Nadeau, Cloutier, and Guay 1993, 204). While this is likely an overstatement, experiments have nonetheless been the most common approach to studying the bandwagon effect in the empirical literature.

The main limitation to experimental research is that it has focused almost entirely, as noted in the discussion of the typology above, on static bandwagon conversion effects. This is not necessarily problematic – this is a type of bandwagon effect that deserves study. However, the experimental approach is underplaying its hand by not simultaneously measuring static and dynamic, and conversion and mobilisation effects. In Chapter 3, I provide an example of how to study all types of bandwagon effect in tandem, experimentally. I also use the experimental approach to unpack the effects of different types of popularity on expectations in Chapter 4.

Conclusion

There is a lot of research on the bandwagon effect, but as yet the concept has rarely been formulated in a consistent and coherent way that allows researchers a) to distinguish it from other phenomena in voting behaviour, b) to clearly situate it theoretically, and c) to question methods for studying it critically. In this chapter, I have provided a framework in which to read the empirical work of this thesis by addressing each of these points.

First, I answered the question of what we mean when we say 'bandwagon effect,' through conceptual analysis. Reviewing the way the concept has been treated in the literature over approximately the past 80 years, the first part of the chapter attempted to define the effect so as to draw out the common focus of these studies and clearly demarcate the bandwagon effect. This definition describes the bandwagon effect as *a positive individual-level change in vote choice or turnout decision towards a more or increasingly popular candidate or party, motivated initially by this popularity*. The room for manoeuvre left by ‘vote choice or turnout decision’ and ‘more or increasingly popular’ sets up a typology of bandwagon effects, which can be static or dynamic, and conversion or mobilisation effects.

Second, I asked what the broader theoretical implications are of such an effect. In other words, what would we expect this effect to look like as a causal process? How might people end up being motivated to vote for more or increasingly popular parties? Why does it make sense to imagine such an effect taking place? The second part of the chapter addresses this by linking the bandwagon effect to broader political psychological theory on how information influences voters’ decisions. To summarise, the bandwagon effect might come from a tendency to use popularity information such as the polls as a ‘heuristic’ to circumvent difficult vote decisions, and from a deeper engagement with the information about the pool of candidates when vote decisions are easier. Because this is consistent with how voters are known to treat information in general, it is apparent there is a place for the bandwagon effect within these models. This place is very similar to how Mutz (1998) situates her theory of ‘impersonal influence.’ Recognising this, I draw on her work on the ‘cognitive response’ to explain the psychological mechanisms underlying the effect when it comes from ‘deeper engagement,’ and further draw out the similarities between this model and more recent, prominent models of political psychology (e.g. Lodge and Taber 2013; Marcus, Neuman, and MacKuen 2000).

This section builds a simple theoretical model of the bandwagon effect – of every form – presented through causal diagrams. This model and the attendant theoretical discussion reveal that voters' predispositions should considerably constrain the bandwagon effect, as both mediator and moderator.

Finally, I outlined the two dominant approaches to studying the bandwagon effect that work around the difficulties this causal model presents in different ways. A key problem is that the popularity information voters consume is hard to pin down. For all intents and purposes, it is unobserved. The first, observational approach, involves inferring this information from people's 'electoral expectations.' I set out the assumptions behind this approach, setting the stage for large parts of the rest of the thesis. One assumption – that dynamic popularity information affects electoral expectations – is as yet untested, so in Chapter 4 I test it. More broadly, and partly because of this finding, the causal inferential assumptions behind this 'expectations approach' mean that it typically falls short of measuring the bandwagon effect. This is a point that Chapter 5 draws out fully, with potentially broad implications for the study of a whole class of questions in political science, beyond the bandwagon effect.

Before progressing to these discussions, though, it is worth asking how the arguments above can be put into practice. Arguably, the conceptual and theoretical arguments are only useful insofar as they can drive research design. The discussion of approaches to measurement above describes how this is typically done in broad terms, but I have noted that there are limitations to such existing approaches. So if a researcher has full control over the design and implementation of a study, how would he (that is, I) apply these arguments in order to learn as much as possible about the bandwagon effect? How could the potential of the experimental approach to studying the effect be maximised? These are the questions I turn to next, in Chapter 3.

Chapter 3

Off the Biden Bandwagon: An Experimental Study of the Bandwagon Effect in the 2020 US Presidential Election

Introduction

This chapter takes the conceptual and theoretical points made in Chapter 2 and directly implements them to design and conduct empirical research. In doing so, it addresses RQ_3:

> **RQ_3**: How can the concept and theory of the bandwagon effect be put into practice empirically?

The chapter presents an experimental study of the effect of state polls on voting behaviour at the 2020 US presidential election. I note, first, that research on the bandwagon effect has overlooked the potential impact of state-level polls on voting behaviour (in favour of national polls). I use the theoretical discussion from Chapter 2 to establish that these polls could be likely – perhaps especially likely – to induce bandwagon effects. I then design an experiment to capture all four types of bandwagon effect set out in the typology. Contrary to predictions of the bandwagon effect, voters seem to emerge as *less* likely to support Biden when learning he is in the lead in polling in their state. Beyond this, the dominant story the results tell is that state polls had minimal net overall effects.

It makes sense to study the bandwagon effect in state polls for two reasons. First, because existing research has not done so, leaving a gap in our knowledge about the bandwagon effect in major presidential elections. Second, because the theoretical mechanisms underlying the bandwagon effect, set out in Chapter 2, suggest the effect could happen in state polls. Research on the bandwagon effect in presidential primaries and nomination campaigns suggests that *state* contest results stimulate the viability mechanism. More relevantly, the cognitive response could also operate particularly strongly at this state level. The first section of the chapter attempts to draw out these points.

The 2020 US presidential election also serves as an important case study of the bandwagon effect. In this election, polls overestimated Joe Biden's support in multiple states. This contributed to the inaccuracy of election forecasts, which nonetheless correctly predicted that Biden would win overall (Keeter et al. 2021; Kurtzleben 2020). As Hardmeier (2008, 504) notes, 'polls are accompanied by two constants: the debate about their quality… and the debate about their alleged effects in the run-up to elections.' As drawn out in Chapter 1 though, these two questions are also related: if the polls were wrong, the question of whether or not they influenced people is arguably more pressing. If pollsters and forecasters are able to give candidates a significant electoral boost by incorrectly reporting that they are in the lead, regulators and policy-makers may have good reason to seek to restrict publication of these reports. This is certainly what the framing of discussions about such regulation seem to suggest. Alongside inquests into why the polls 'missed' what was to happen in a lot of states in 2020, then, it is natural to ask what the effects of these inaccurate state-level polls might have been on voters.

As Chapter 2 explains, though, there are multiple dimensions to the bandwagon effect: voters can be influenced not only by which candidate is in the lead (static popularity), but also how the candidates' popularity appears to be changing over time (dynamic popularity), and the effect can be expressed through changes in vote choice between the candidates (conversion), or changes in turnout (mobilisation). To understand the effects of state polls in this election fully, it is necessary to account for all of these potential effects.

In order to apply these arguments, I conducted an experiment in which I randomly exposed voters to information about the static and dynamic popularity of Trump and Biden in potential swing states (Georgia, North Carolina, Ohio, and Texas) in the lead up to the 2020 US presidential election – using data from real, publicly available state-level polls. Respondents had the choice to vote for a candidate or abstain. The results of this experiment suggest, contrary to the bandwagon effect, that being portrayed as the leading candidate in a state poll tended to *decrease* Biden's support in that state, especially among those who did not already identify with the Democrats. There is no convincing evidence of any significant dynamic or mobilisation effects.

These observations point to the possibility that polls missed the outcome partly *because* they affected the outcome. Polls suggesting Biden was in the lead may have actually driven down support for Biden. Perhaps these polls were not as wide of the mark as it seems. They likely still substantially overstated Biden's popularity, but publication of these very overestimates may have itself amplified the disparity between the polls and the eventual outcome in many states. As such, as well as demonstrating how to apply the insights of Chapter 2 when designing research addressing the bandwagon effect, and providing an empirical insight into the effects of state polls, this chapter potentially contributes to resolving the puzzle of how the polls overestimated Biden's dominance.

The chapter proceeds by exploring the findings of existing research on the bandwagon effect, focusing on experimental studies, and explaining how these findings relate to the theory behind the effect, in order to demonstrate the space this leaves specifically for the study of effects of state-level presidential polls. Throughout this discussion, I propose hypotheses that cover the full range of bandwagon effects from Chapter 2. I then explain the experimental design and data collection process, before presenting my results. The Discussion and Conclusion section explores the implications of these findings and provides a note of caution on over-interpretation, acknowledging the chapter's limitations.

Bandwagon Effects and State Polls

Experimental study of the bandwagon effect dates back as far as the 1940s, following the introduction of 'scientific' polling methods (Cook and Welch 1940), based on the concern that publication of such polls might create the effect. While some experimental studies have either found very nuanced or limited evidence (Navazio 1977), or none at all (Dizney and Roskens 1962), most have found support for bandwagon effects, whether in hypothetical (Farjam 2020; Goidel and Shields 1994; Roy et al. 2015) or real election scenarios (Mehrabian 1998). Despite this long history, to date, no research has studied whether or not *state*-level polling can produce bandwagon effects in presidential contests.

Chapter 2 sets out two ways in which a bandwagon effect might emerge. Scrutinising these theoretical underpinnings of the effect helps to demonstrate why state-level polling offers a potentially interesting case. First, in the US context, scholars believe that the viability mechanism has an effect on presidential primary elections. Kenney and Rice (1994, 927) argue that George H. W. Bush enjoyed a 'substantial' bandwagon effect in 1988 owing to his momentum in early primary contests. Bartels (1985, 804; see also Bartels 1987) presents Gary Hart's 'meteoric rise' as evidence that 'bandwagons play an important role in nominating campaigns.' In these studies, the 'momentum' behind a bandwagon effect is signaled by success in individual states throughout the sequential election, rather than by overall vote intention polls of all voters. It is believed that winning states in the primaries, particularly early ones, determines a candidate's viability, and creates a bandwagon effect. This same logic was used in 2020, when success in early contests led some to claim Bernie Sanders had 'momentum' (Burgis 2020), but this changed after Joe Biden won South Carolina and set his ‘bandwagon’ rolling (Veit 2020).

It therefore seems that candidates’ popularity and viability in *other* states, as revealed through victories in earlier contests, has the potential to influence voting behaviour. This raises the question of whether people within a given state might partially base their vote choice on how well a candidate is likely to do *in their state*. There is no clear reason why voters would privilege popularity elsewhere over popularity closer to home. The difference could of course be brought about by the sequential nature of the election. It may be that the clear, concrete signal of recent electoral victories (in other states) is more persuasive than the uncertain information conveyed in polls (in the voter’s state). As the effects of state-level polls on the vote have not received scholarly attention, little is known about whether this distinction is consequential.

Arguably, though, presidential elections are not so difficult for voters that they should have to rely substantially on the viability heuristic. Chapter 2 explains that the ‘cognitive response’ (Mutz 1998) could lead to a bandwagon effect in easier contests such as the general presidential election. This happens when voters receive popularity information and, in response, consider the arguments that people might have to justify their support of the popular candidates. These arguments might end up convincing the individual to jump on the bandwagon too.

Again, here, the role of candidates’ popularity in a given voter’s state is largely overlooked. Yet the outcome of the election is determined by the number of states in which a candidate wins a majority, rather than whether a candidate wins a majority of votes across the USA.[1] Recall that, now (in)famously, in 2016, Hillary Clinton received more votes than Donald Trump overall, but Trump still won the election. This makes the state race important. Theoretically, there is reason to believe that the cognitive response could also operate particularly strongly at this state level. Voters are likely to feel a closer affinity, or mutual interest, with others in their state. This would make understanding the reasons which, for example, Georgians have for voting for Biden more pertinent to a given Georgian who finds out Biden is leading in Georgia state polls. For similar reasons, a voter in a given state is likely to have a deeper understanding of the concerns and motivations of other voters in her state, making these considerations more accessible to her for mental 'rehearsal.' Both of these factors should combine to strengthen the cognitive response mechanism.

Taking this body of evidence and arguments together, it is highly plausible that state polls could create bandwagon effects. This is notable given that, in the most recent presidential election, many state polls suggested Joe Biden had a larger vote share than he ended up receiving (Keeter et al. 2021). Some voters may have been motivated by this potentially false popularity information to vote for him where they otherwise wouldn't have. There may have been a Biden bandwagon at the state level. Importantly though, this could also mean that in states where Trump was dominant in the polls, or in which some polls gave Biden the lead, and others gave it to Trump, Trump may have derived a similar advantage. These differential effects would have been unlikely if people only followed national polls, which

[1] The states of Maine and Nebraska allocate two electoral votes to the state popular vote winner but also give one electoral vote to the popular vote winner in each of their 'Congressional districts.' This means that these states sometimes spit their electoral college vote, providing support to both candidates.

universally favoured Biden.

Static and Dynamic Bandwagon Effects

As well as varying across different levels of competition, the picture painted by the polls also depends on both 'static' and 'dynamic' information. As the overview of the literature in Chapter 2 indicated, little attention has historically been paid to the possibility that voters might also be influenced by the momentum a party or candidate has *over time* (dynamic popularity), rather than just which party or candidate is most popular (static popularity), even though poll results reveal both of these things to voters (Irwin and Van Holsteyn 2000). Where experiments have studied dynamic and static effects together, this has been in research on issue opinions, and typically without allowing static and dynamic popularity to vary independently. For example, experimental participants will be told that public opinion is both in a certain direction, and moving increasingly towards that direction, on an issue such as abortion (Nadeau, Cloutier, and Guay 1993) or free trade (Cloutier, Nadeau, and Guay 1989). Only Marsh (1985) has clearly treated static and dynamic bandwagon effects distinctly, with one group in her experiment seeing static information and another seeing dynamic information about abortion opinions. These studies find evidence suggestive of dynamic effects in attitude formation, but this does not readily or necessarily generalise to voting behaviour.

In general, it is important to consider the role of dynamic bandwagon effects. Recent research even suggests that polls are selectively reported and published based on the dynamic consideration of how much they suggest parties' vote shares have changed over time (Larsen and Fazekas 2021). This could mean that voters are more likely to see polls that have the potential to produce dynamic bandwagon effects. By overlooking the potential for such effects, it is therefore possible that we miss a large part of how these polls affect election outcomes.

Experimental voting behaviour research has only studied dynamic bandwagon effects in proportional, multiparty systems. Arguably here, because being in first place in the election is not necessarily the only way to 'win' (Plescia 2018; Stiers, Daoust, and Blais 2018), dynamic bandwagon effects should be more relevant than static effects. Van der Meer et al.'s (2016, 49) choice to 'focus exclusively on the bandwagon vote related to momentum,' finding evidence of a small bandwagon effect for the Dutch PvdA, reflects this. Similarly, Dahlgaard et al. (2017) find that Danish voters with fake polling information are significantly more likely to vote for the party with momentum. These findings raise the question of whether such 'dynamic' considerations might carry over into majoritarian, presidential contests. This is an overlooked possibility, particularly in experimental research, despite the fact that the idea of dynamic bandwagon effects originates in research on the US (primary) context (Abramowitz 1987, 1989; Bartels 1985, 1987; Callander 2007).

There are theoretical reasons to expect that dynamic bandwagon effects would happen in presidential contests. First, note that Mutz (1998, 212) frames the 'cognitive response' in fundamentally dynamic terms, as a process that might be triggered when 'more and more people are rushing to support a particular candidate.' Second, because presidential contests are 'easier' decision tasks for voters, they are naturally expected to encounter more information about the candidates in the race, including polling information. In this context, it is unlikely that most polls will be surprising to many voters, limiting the extent to which they affect their behaviour. As such, dynamic changes in the polls should exert *more* of an influence on voting behaviour than the static state of the race (Atkin 1969, 516; Irwin and Van Holsteyn 2000, 13).

Recognising this possibility and reflecting the distinction drawn out in Chapter 2 – as well as the purpose of this chapter as a demonstration of how to operationalise these arguments experimentally – I propose the following hypotheses:

H_1 *Static bandwagon conversion effect*: polls suggesting a candidate (Trump/Biden) was in the lead in their state made voters more likely to vote for him over other options.

H_2 *Dynamic bandwagon conversion effect*: polls suggesting a candidate (Trump/Biden) was becoming more popular in their state made voters more likely to vote for him over other options.

Bandwagon Conversion and Mobilisation Effects

In calling these 'conversion' effects, I distinguish them from 'mobilisation,' consistent with Chapter 2. To recap, it is possible for bandwagon effects to come about through changes not only in which party someone intends to vote for (conversion) but also whether or not they intend to turn out to vote for their preferred party (mobilisation). Promising research from economics suggests that popularity might motivate turnout (e.g. Agranov et al. 2018; Grillo 2017; Panova 2015). This is an important finding in the economic literature on public choice due to how it speaks to rational choice models of voting. These generally predict that the *closer* a race is, the more likely people are to turn out, and are directly contradict bandwagon mobilisation effects (Grillo 2017, 466). One economic experiment found that 'the availability of information reduces the probability of minority participation and increases the probability of majority participation' (Agranov et al. 2018, 839). Such effects also emerge in variously specified formal models or games (Gartner 1976; Grillo 2017; Panova 2015). Beyond this, scholars conducting quasi-experiments have also claimed to find evidence of bandwagon 'turnout' effects (Kiss and Simonovits 2014; Morton et al. 2015) though this largely captures distinctly non-bandwagon behaviours, as noted in Chapter 2. It is important to account for this effect, as well as the far more widely studied bandwagon conversion effect, in order to capture the full range of decisions that can be motivated by polls.

Again, there are also theoretical reasons to expect bandwagon mobilisation effects in presidential elections. As Chapter 2 explained, the cognitive response mechanism depends on convincing arguments motivating changes in voting behaviour. Insofar as the cognitive response can be applied to mobilisation effects, which Chapter 2 claims it can, this suggests that there is plenty of opportunity for mobilisation – it might even be easier for popularity information to stimulate mobilisation than conversion. To the extent that an individual has well-developed party preferences, which most voters tend to have in presidential contexts (Abramowitz 1987, 50; Gimpel and Harvey 1997, 158), she is likely to engage in 'motivated reasoning' (Leeper and Slothuus 2014) which will make the considerations that come to mind, when she undergoes the cognitive response, less likely to convince her of the merits of the other party and therefore convert (Zaller 1992). However, when the poll or popularity information she receives is favourable to the party she *already prefers*, the arguments she considers to understand why people are voting for the party are likely to be more favourable towards that party and motivate her to turn out. She understands the reasons behind supporting the party, she just needs to be covinced to go and cast a ballot.

In light of this argument and the goal of this chapter as a demonstration of how to operationalise all the distinctions made in Chapter 2, I propose the following hypotheses:

H_3 *Static bandwagon mobilisation effect*: polls suggesting a candidate (Trump/Biden) was in the lead in their state made voters more likely to vote for him rather than abstaining.

H_4 *Dynamic bandwagon mobilisation effect*: polls suggesting a candidate (Trump/Biden) was becoming more popular in their state made voters more likely to vote for him rather than abstaining.

Data and Method

Experimental Design

I conducted an experiment during the 2020 US presidential election campaign, in order to assess these hypotheses. I received ethical approval for the experiment from the Queen Mary Ethics of Research Committee (QMERC2500).

Voters from four states in the USA were randomly assigned to receive different polling information from within their state, taken from real, recently conducted polls. They could see a text describing a poll in which Donald Trump was in the lead or one in which Joe Biden was in the lead. In both cases, these were real polling results that voters could have encountered outside of the experimental context. Randomised independently of this, but shown at the same time, they could be told that Donald Trump was gaining ground in the polls or that Joe Biden was gaining ground. The other candidate, in each case, had stable polling (either no change, or at most an increase or decrease of one percentage point since a previous poll). This, again, was information the voter could have gleaned outside of the experimental context, because this change over time is based on a comparison to another real recent poll.[2] An additional layer in the design randomly assigns one of these features to be 'qualitatively emphasised,' with an additional sentence drawing attention to it.[3] Respondents choose between four outcomes. They can either abstain, vote for 'other,' or vote for Biden or Trump. This experiment was fielded on October 1st to a total of 2,661 participants during the late stages of the campaign for the 2020 US presidential election.[4]

The use of real polling information in the experimental treatments arguably makes not only for a more ethical, but also more internally and externally valid design. The main priority in designing this experiment was to ensure it was ethical and justifiable, particularly given the salience of the election campaign into which I was intervening. As such, I wanted to minimise as far as possible any risk of deception. Rather than dwelling on this here though, I reflect at length on why such considerations matter in the essay on experimental design in Appendix A. That discussion also sets out why, arguably, minimising deception can be seen as a strategy that is beneficial to causal inference as well as ethical justifiability,

[2] All of the polls used here were taken from FiveThirtyEight's state poll aggregators, e.g.: https://projects.fivethirtyeight.com/polls/president-general/ohio/.

[3] See Appendix B.9 for the wording of these sentences.

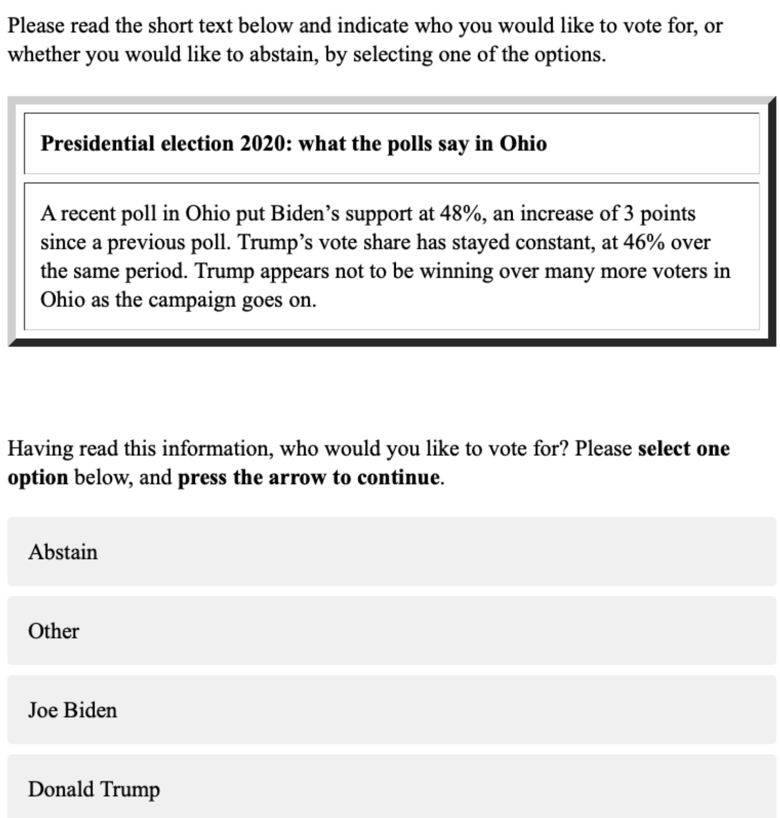
[4] The experiment was designed in Qualtrics and fielded on the Prolific platform (Palan and Schitter 2018). Based on the power analysis reported in Appendix B.1, I targeted a sample of over 4,000 voters but received large rates of non-response. State sample sizes: Georgia 481, North Carolina 559, Ohio 519, Texas 1102.

contrary to intuitive assumptions about what deception does.

Figure 3.1 shows a real example of a vignette a respondent in Ohio might have seen. Here, the design randomly assigns static popularity and dynamic popularity to Joe Biden, and the qualitative emphasis is set to Donald Trump's dynamic popularity (which is stagnant). Figure 3.1 is exactly what respondents saw – the technical wording of 'static' and 'dynamic popularity' is just what I use here for shorthand and was not itself presented to respondents. The recent poll the vote shares are drawn from is real, and the same for both candidates. The dynamic popularity is based on linking this to a real previous poll which is again the same for both candidates, and calculating the change in vote shares between the two. I provided a full list of all the polls used in the experiment to respondents in the debrief.[5] This vignette is just one example that a voter in Ohio could have seen. Appendix B.9 contains examples of the other options, along with the possible emphasis sentences that can be combined with these in Appendix B.4. An equivalent range of options was available in each of the other states, *mutatis mutandis*. In order to minimise 'profile-order' effects (Leeper, Hobolt, and Tilley 2019), the treatment information could be presented mentioning either candidate first (i.e., it was equally as likely that the trailing candidate would be mentioned first as the winning candidate). In order to avoid 'response-order' effects – similar to ballot order effects (Ho and Imai 2006) – the response options were shown in a completely random order. That is, the order in Figure 3.1 was one of 24 ($4 \times 3 \times 2 \times 1$) possible orderings, all randomly determined for each respondent and shown with equal probability.

I chose the four states – Georgia, North Carolina, Ohio and Texas – based on

[5] Respondents were shown all of the possible combinations they could have seen, and links to information on the polls were provided. This information is available in Appendix B.3.


Please read the short text below and indicate who you would like to vote for, or whether you would like to abstain, by selecting one of the options.

Presidential election 2020: what the polls say in Ohio

A recent poll in Ohio put Biden's support at 48%, an increase of 3 points since a previous poll. Trump's vote share has stayed constant, at 46% over the same period. Trump appears not to be winning over many more voters in Ohio as the campaign goes on.

Having read this information, who would you like to vote for? Please **select one option** below, and **press the arrow to continue**.

Abstain

Other

Joe Biden

Donald Trump

Figure 3.1: Example experimental task from Ohio.

closely related theoretical and practical considerations.[6] In all four of these states it was possible to find recent polls showing a lead for either candidate, and as a result it was also possible to find previous polls relative to which either candidate's more recent performance was an improvement. For example, in Georgia, one poll gave Trump a 47%-46% lead. A previous poll gave Trump 44% to Biden's 46%. Combining this with the recent poll means that Trump's performance improved (+3%) while Biden's stayed constant. The same poll could however be combined

[6] In the event, North Carolina, Ohio and Texas were all won by Donald Trump, but Joe Biden won in Georgia.

with a previous poll in which Trump had 48% to Biden's 41%, meaning Trump's performance stayed relatively constant (-1%) while Biden's improved (+5%).[7] This variability meant that it was possible to randomise static and dynamic popularity independently, without recourse to fake data. This is, of course, also what makes these states of theoretical interest – they were those where the winner seemed hardest to predict, and therefore where a bandwagon effect would potentially be most consequential for the overall election outcome.

In the case of Texas, it was not possible to have a combination of polls such that Trump had dynamic popularity, while Biden had static popularity. There was no recent enough combination of poll results available. Owing to the emphasis placed here on realism and avoiding deception, I therefore did not include this combination. As a robustness check, in Appendix B.5 I carry out a version of my analysis in which the data from Texas is removed and reproduce the main findings. Similarly, I reproduce the main findings only among respondents who passed attention checks (Appendix B.6) and only among those who had not already voted when the experiment was conducted (Appendix B.7).

The independently randomised sentence emphasising one of the candidates' static or dynamic popularity reflects the finding that bandwagon effects only emerge when dry poll results are given a qualitative spin (Fleitas 1971; van der Meer, Hakhverdian, and Aaldering 2016). Experimental designs should incorporate this because in the real world it is very rare for voters to see polls without any qualitative interpretation (Dahlgaard et al. 2017). This part of the design works by simply randomising between four options – Biden static popularity, Trump

[7] Some of this flexibility is enabled by differing methodologies across polling agencies and the distinction between registered voter and likely voter polls.

static popularity, Biden dynamic popularity, Trump dynamic popularity – and selecting a corresponding sentence that is added to the short text, emphasising this aspect of the poll results. If Biden has the larger vote share, and the qualitative emphasis is 'Biden static popularity,' this sentence will draw attention to the fact he is in the lead. If Trump has not increased his vote share since the last poll, and the qualitative emphasis is 'Trump dynamic popularity,' the sentence will draw attention to the fact that he is not winning people over as the campaign goes on.[8]

Results

Overall Effects

Figure 3.2 presents the overall picture of how participants in the experiment responded to the main static and dynamic popularity information. The effects visualised here are drawn from regression models in which the outcome is either a binary indicator of whether or not someone voted for Biden (left), or a corresponding binary indicator of whether someone voted for Trump (right). These effects describe the change in the probability of voting for the candidate brought about by switching from that candidate not having static/dynamic popularity (No static/No dynamic) to having static/dynamic popularity (Yes static/Yes dynamic). Again, this wording is just shorthand for the more clearly phrased vignette text

[8] I include this 'emphasis' feature in the regression models reported below in order to improve precision , but it is not of central interest. I originally planned to conduct analyses with an interaction term between this emphasis variable and the 'popularity' factors, in order to detect whether the effect of a specific type of 'popularity' depended on it being emphasised in the vignette, but I do not report any such analysis because – as the power analysis in Appendix B.1 shows – the sample collected here is too small to detect these effects with sufficient power.

the respondents saw. For example, the 'Yes static' effect in the left-hand plot describes the change in the probability of voting for Biden associated with Biden being the most popular option in the poll described in the vignette, as opposed to not being the most popular option (the baseline level). These models include all of the participants, partially pooled by state, with random state-level intercepts.[9] For increased precision, both models control for which factor was 'emphasised.' The circular points represent the average effect estimate, and the horizontal bars represent 95% confidence intervals. For each factor, the level used as the baseline has its estimate fixed to zero.

The results in Figure 3.2 immediately contradict the idea of static bandwagon effects (H_1/H_3): when they were in the lead rather than being behind, both candidates on average *lost* votes. However, in neither case are these effects statistically significant. The results are different for dynamic popularity. Support for both candidates was somewhat more likely when they had momentum relative to a previous poll, pointing to a potential dynamic bandwagon effect (H_2/H_4). Although, again, this effect is not statistically significant.

Conversion Effects

As the above results use data on the whole sample, they cannot distinguish between conversion and mobilisation. The inclusion of those who already identify with the relevant party in each case (e.g., including Democrats in Biden's models)

[9] This just means that the linear regression model has a separate 'intercept' for each state. Rather than just modelling the average probability of voting for each candidate across the whole sample, expecting this not to vary by state, the average probability is measured in each state (and the variance amongst these averages), and effects are relative to these averages. This is an appropriate way of measuring data that is hierarchical, as is the case here (McElreath 2020).

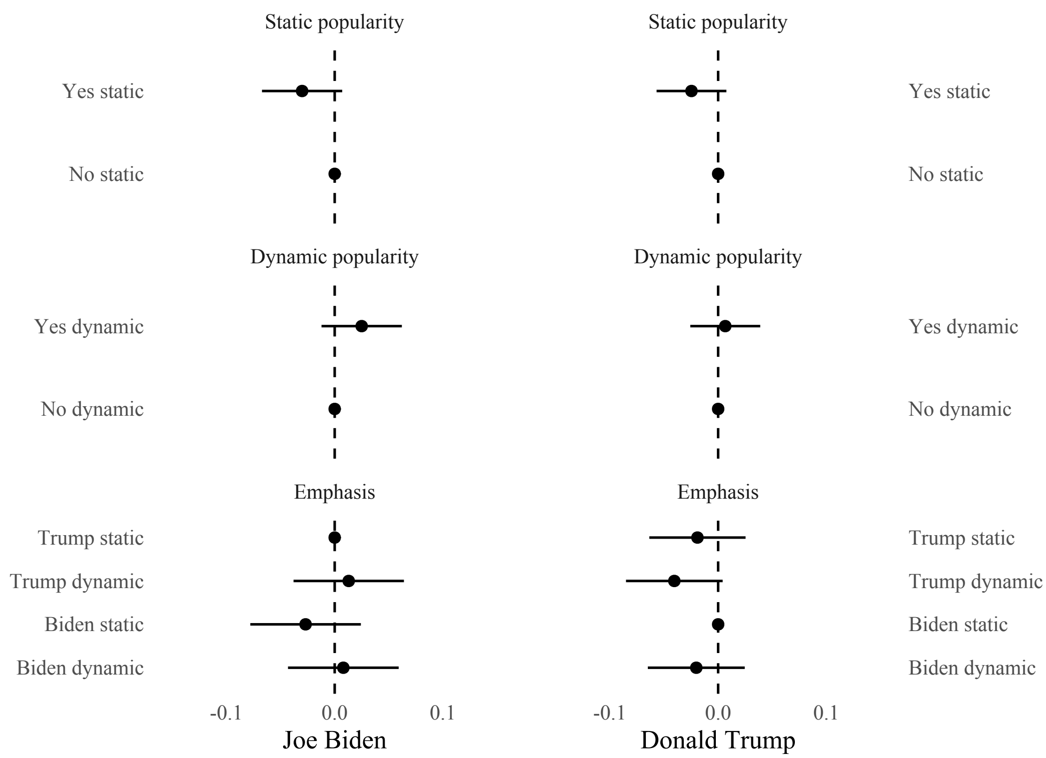


Figure 3.2: Effects of popularity information on voting for Biden (left) and Trump (right), across whole sample, partially pooled with varying intercepts by state.

might contribute to the null results. The results presented in Figure 3.3 are instead based only on those who did *not* say, prior to the experiment, that they identified with the party of interest.[10] Therefore, these models measure conversion effects

[10] I operationalised conversion this way partly because many respondents were too young to have voted in 2016, providing no basis of past vote choice on which to define conversion. In Appendix B.8, I consider ways of operationalising conversion using 2016 vote choice. These change the results slightly. If conversion is operationalised as bringing onside people who voted for the other main party at the last election, then the effect of Biden's static popularity reported here is still significant at the 95% confidence level. If it is operationalised as bringing onside people who voted for any other candidate at the last election, then the effect of Biden's static popularity reported here is only significant at the 90% confidence level. By conventional standards, this is not statistically significant, but note that recent, highly prominent experimental research on the bandwagon effect uses a 90% confidence level (see van der Meer, Hakhverdian, and Aaldering

only. The plots on the left-hand side show the effects of Biden's popularity on non-Democrats' probability of voting for him ($N = 1482$), and those on the right show the effects of Trump's popularity on non-Republicans' probability of voting for him ($N = 2204$).[11]

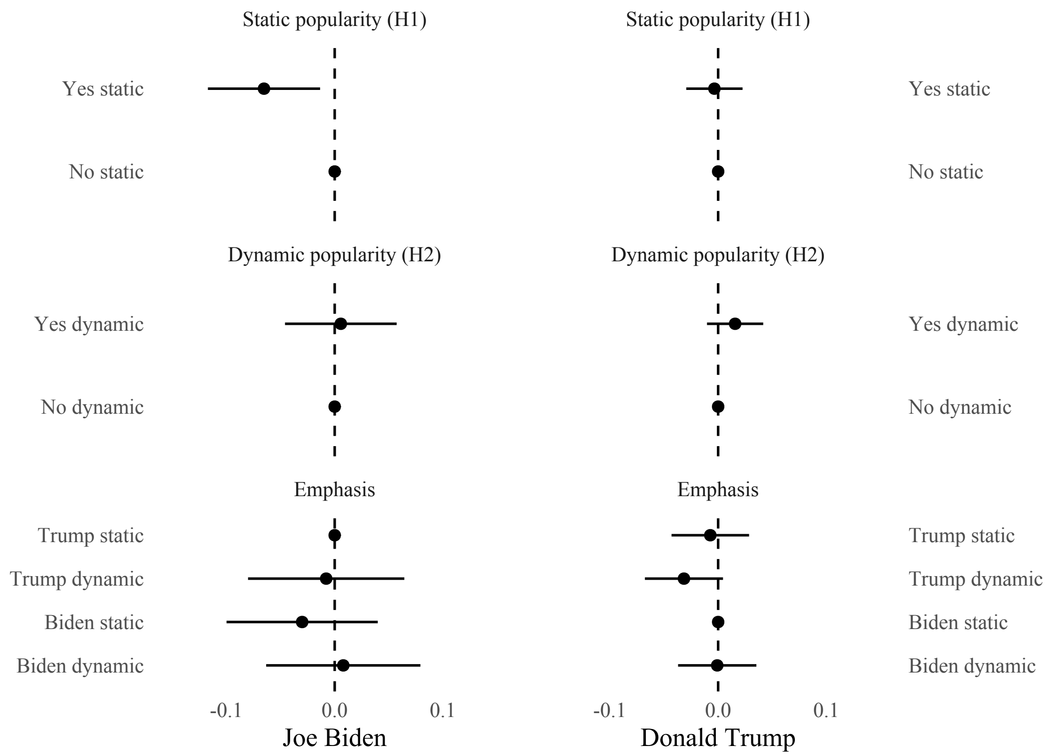


Figure 3.3: Effects of (emphasised) popularity information on voting for Biden (left) and Trump (right), for those who do not identify with Democrats (left) or Republicans (right), partially pooled with varying intercepts by state.

Looking first at the left-hand plots, Biden's tendency to lose support when he is in the lead is now significant at the 95% level. When non-Democrats saw a

2016). Recent research suggests that promising alternative to these approaches to operationalising conversion is to have a repeated measure of vote choice (Clifford, Sheagley, and Piston 2021).

[11] In each case, these sub-samples include those who identify with the opposing party (1177 Democrats, 455 Republicans), and those who deem themselves independents (709), or supporting 'none' (174) or 'other' (144).

poll, from their state, suggesting Biden was going to win that state, they became between approximately 1.35% and 11.71% (mean estimate 6.53%) less likely to vote for him. The relative imprecision of this estimate owes to the fact that the analysis draws on such a reduced sample, having removed those who identify with the Democrats. Among non-Republicans, there is no corresponding evidence of any effect of Trump's popularity on voting for him. Therefore, neither candidate provides any support for static bandwagon conversion effects (H_1), and the results even suggest the opposite of this for Joe Biden.

Superficially, the negative effect of Biden's popularity might appear to support the idea of an 'underdog effect' – as each candidate was either in the lead or trailing, this estimate also captures a positive effect of being *behind* in the poll. However, Trump does not enjoy an equivalent effect. If a pure underdog effect were in operation, Trump might also be expected to gain voters when seen to be behind. This raises the possibility that this negative effect might be better explained by 'oppositional reactivity' – a tendency to support trailing candidates as a reaction against a leading candidate (Ceci and Kain 1982). Rather than a general tendency simply to back an underdog across both non-Democrats and non-Republicans, non-Democrats specifically are more likely to vote for Joe Biden when he is behind, i.e. when Trump is in the lead. It is possible that polls favourable to Trump were enough to persuade people to get behind Biden, especially considering Trump was such a divisive candidate.

For neither candidate does dynamic popularity appear to have a significant conversion effect, although the estimate for both is positive. H_2 finds no convincing support.

Mobilisation Effects

In order to assess mobilisation effects, I take a sub-sample of only those who identify with each of the two parties. That is, I measure the effects of each type of popularity on voting for Biden, only among Democrats ($N = 1177$) and the same for voting for Trump among Republicans ($N = 455$). The main results are shown in Figure 3.4. Again, the left-hand plots show the effects for Biden, and those on the right-hand side for Trump.

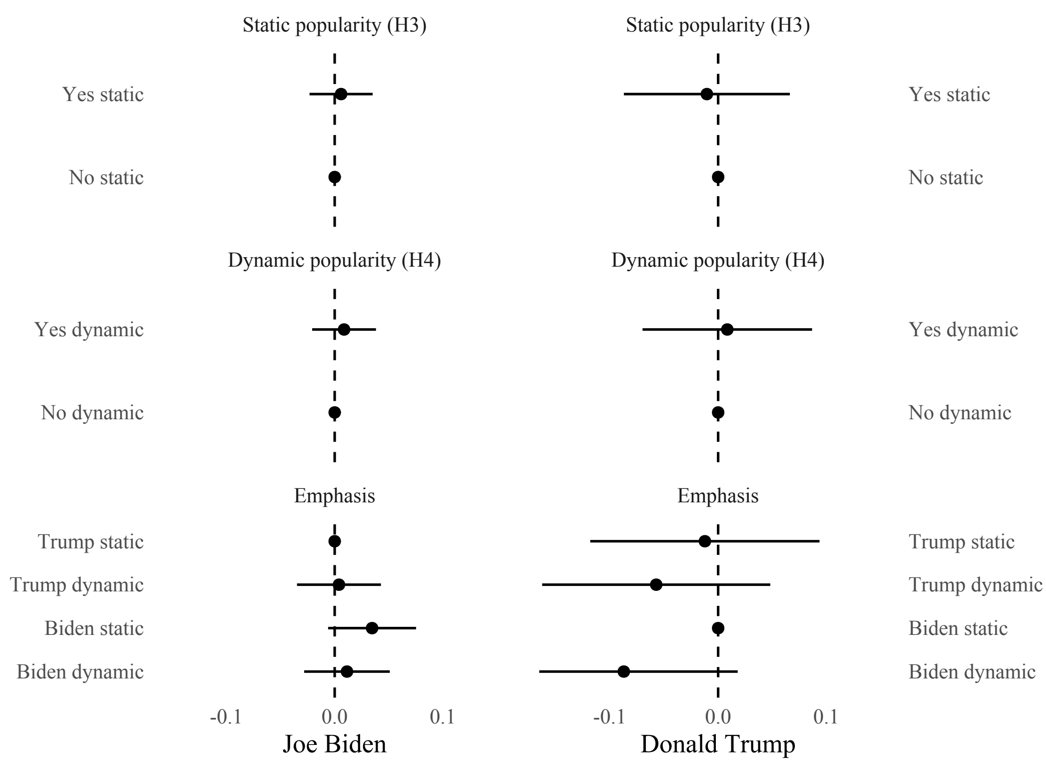


Figure 3.4: Effects of (emphasised) popularity information on voting for Biden (left) and Trump (right), for those who identify with Democrats (left) or Republicans (right), partially pooled with varying intercepts by state.

For neither candidate do static or dynamic popularity appear to have any discernible effect on support. There is no evidence supportive of static or dynamic bandwagon mobilisation effects (H_3/H_4). However, these results should be read cautiously. As is apparent from the large standard errors indicated by the wide confidence intervals in the right-hand plot, the small sub-samples on which these effects are based makes it very difficult to detect anything other than very large effects. Compounding this, mobilisation effects are necessarily small. Very few people report not intending to vote in surveys, limiting the effect size, and this is borne out here too: only 27 Republicans abstained in the experiment, and only 32 Democrats. Others have encountered this difficulty (Ansolabehere and Iyengar 1994), which speaks to a wider problem of attempting to measure turnout in survey research (Karp and Lühiste 2016). This might go some way to explaining why little research has measured bandwagon mobilisation effects.

Discussion and Conclusion

This chapter set out to demonstrate how the conceptual and theoretical arguments developed in Chapter 2 can be implemented throughout the research design process. It points to a gap in the literature on bandwagon effects, where state-level poll effects have not been considered, and argues that theoretical mechanisms behind the bandwagon effect are consistent with state polls having such an effect, including typically overlooked dynamic and mobilisation effects. This is, moreover, important to understand given controversy around such polls in the 2020 US presidential election, which in many cases overstated the extent of Biden's victory. If these polls had the potential to induce 'bandwagon effects,' then inaccurate polls would have unduly biased the outcome of the election. To test this, in four of the states where the outcome was most difficult to predict, I exposed a sample of 2,661 people to the results of real recent polls from their state, randomly varying whether they saw a poll in which Trump was in the lead or one in which Biden was in the lead. This was also connected to a randomly chosen previous poll which gave either of the candidates 'momentum' over the intervening period. This, combined with respondents' options of abstaining or voting for any candidate, means that the experiment tests for the full range of bandwagon effects conceptualised in Chapter 2.

The results suggest that in the 2020 election, for Joe Biden, being portrayed as the leading candidate in swing states tended to *decrease* his support, particularly among those who did not identify with the Democrats. In other words, people were *more* likely to switch to vote for him when he was trailing. In an experiment designed from the bottom up to test for the bandwagon effect, I found the opposite. This could be interpreted as an 'underdog effect' – such effects have been observed before but are usually dominated by bandwagon effects (Hardmeier 2008). The specific context, however, provides a potentially better way to understand this. Given the strong negative reaction many people had to Trump's presidency, it is possible that polls favourable to him were enough to persuade people to support Biden. Ceci and Kain (1982) call this process 'oppositional reactivity.' When Biden was in the lead, this was not a concern, so these people did not vote for him. Strengthening this interpretation, Trump did not similarly benefit when he was behind. The effect was only significant for Biden. An additional implication of these findings is that Trump did not appear to gain or lose significant numbers of voters based on his popularity in the polls, casting doubt on claims of 'suppression polls' (Campanile 2020).

It is important to note that these findings, though they provide no convincing evidence of bandwagon effects, should not be seen as absolutely ruling out such effects. The sample used here is too small to detect effects of only a few percentage points with precision, as the power analysis in Appendix B.1 shows. This is more problematic when splitting the sample into partisan sub-groups. Based on these considerations, as I have noted, I originally targeted a sample of over 4,000 when fielding the experiments. Large rates of non-response meant, however, that the sample fell short of this number. Any future experimental research on the bandwagon effect reapplying the experimental framework proposed here should seek to gather larger samples (see, e.g., van der Meer, Hakhverdian, and Aaldering 2016). Note, for example, that the effect of Donald Trump's dynamic popularity suggests a non-statistically significant dynamic bandwagon conversion effect of 2%. This effect may have been significant if similar behaviour took place in a larger sample capable of detecting such small effects (see Lanoue and Bowler 1998, 374–75). Indeed, Chapter 5 finds evidence of such small dynamic bandwagon effects in the UK context.

Another reason that this experiment may fail to detect effects is that it may have exposed respondents to polls they had already seen. As the treatments are based on real polls – a decision made primarily for reasons of experimental ethics – some respondents may already have come across these results, and therefore been unsurprised by them and not reacted (Atkin 1969). Indeed, it is even possible that some voters' decisions may have already been influenced by these polls, when they previously heard about them. This all limits the power of the experimental treatment to produce an effect. Beyond this, the unrepresentative nature of the sample, and the small number of states in which the experiments were run, both limit the extent to which these findings can be generalised across the US electorate. It could well be that Biden benefited from a bandwagon effect in other states where he was more dominant in the polls.

Despite these limitations – many of which are common in experimental research on the bandwagon effect – as well as providing important insights into the effects of state polls, the findings of this study have implications for the three key real-world areas to which bandwagon research speaks. For instance, they may have implications for our understanding of how these polls overstated Biden's dominance in the 2020 election – and thereby for electoral forecasting as a whole. Forecasters claim that a major reason things looked rosier than reality for the Democrats in the lead up to the election was that the polls in many states overestimated Biden's vote share (Kurtzleben 2020). If, as the findings presented here suggest, Biden lost support when he was seen to be in the lead in state polls, then it could be that rather than simply over-estimating Biden's share, they may have partly *caused* it to decrease. While these effects are quite small, and therefore extremely unlikely to be the sole reason for the discrepancy between the polls and the outcome, this nonetheless suggests that the polls might not have been as wide of the mark as they appear to have been. Polls may have the potential to change the very outcome they measure, but not necessarily in the way we usually expect. Rather than just becoming 'self-fulfilling,' they might become 'self-defeating' (Henshel 1982).

This last point, in turn, suggests that if future research replicates the finding of negative static effects, this could have some implications on debates about poll regulation. Such debates might have too narrow a focus if they concentrate on the idea that doing well in the polls might unduly *benefit* electoral candidates. Polls might actually have a negative effect on the vote for popular candidates. Assuming the main reason a candidate would be popular is that his or her message chimes with broad swathes of the electorate, this could mean that polls play a minor role in jeopardising the chances of would-be or deserving winners.

Finally, campaign strategists might do well to remember that looking like the inevitable winner might not always be a good thing. Most people thought Joe Biden was going to win the 2020 election. He did win it of course, but this image of likely victory may have lost him some votes along the way, in states that had the potential to be consequential for the outcome.

In terms of implementing the theoretical arguments of this thesis, this chapter has dealt with what is essentially an ideal case. The experimental approach affords the researcher complete control to operationalise and measure the bandwagon effect. The chapter has provided an example of how to move towards doing this to its full potential. However, what about when we do not have the opportunity to conduct an experiment, or do not want to? In these cases researchers often resort to what I have termed the 'expectations approach.' This approach is more complicated, and raises more puzzling questions and challenges for researchers. It is to these challenges that I now turn, starting with the question of how expectations are formed, in Chapter 4.

Chapter 4

Momentum Matters: The Effect of Dynamic Popularity Information on Electoral Expectations

Introduction

This chapter assesses a relationship that the theoretical model at the heart of this thesis posits, but which has not been tested as of yet in political science research. Specifically, I address RQ_4:

> **RQ_4**: Does dynamic popularity information affect electoral expectations?

The chapter shows, through multiple experiments, that how much a party's vote share has increased over time (dynamic popularity) has a significant effect on how likely people think that party is to win the election. This is important to address because it has implications for the 'expectations approach' to studying the bandwagon effect. When researchers cannot or do not want to conduct experimental research on the bandwagon effect, this is the approach to which they most commonly resort. Researchers look for associations between who people think is going to win an election and which party they are going to vote for, in order to draw conclusions about whether there is a bandwagon effect, which they typically interpret in *static* terms. As Chapter 5 will flesh out further, the fact that dynamic popularity information influences people's expectations means that this approach does not necessarily measure what it claims to measure and may in many cases be unable to measure the bandwagon effect at all.

Research addressing the link between popularity information and expectations has only asked whether people's expectations align with what I have called their 'static popularity information.' In forming their expectations, voters should take into account information about the relative numbers of people currently saying they intend to vote for each party. For example, in general, people paying attention to the polls in the UK at the time of writing should mostly recognise that the Conservative Party would have the best chance of winning if an election were held tomorrow. Several studies have found a relationship of this kind (e.g. Blais and Bodet 2006; Irwin 2002; Zerback, Reinemann, and Nienierza 2015). I review this and other relevant literature in the first section of the chapter.

Yet, when people answer questions about their electoral expectations, the election in question takes place – by definition – in the future. This means that people need to work out how the popularity of the parties might evolve between now and then. How popular the parties are right now is a good starting point, and should be increasingly accurate as the election draws near, but it is imperfect. In a sense, people need to mentally map this current, static popularity of the parties forwards in time, to work out how popular they *will become*. This makes 'dynamic popularity information' – change in popularity over time – relevant. If there is evidence that a party's share of vote intentions is currently growing, or has been growing, then this might make people think it is going to continue to grow and improve the party's eventual vote share, and thereby boost its chances of winning the election. This is consistent with social psychological research demonstrating that people predict that behaviours that are becoming more common will continue to increase in prevalence (Mortensen et al. 2019). People also think in a similar way to this when predicting the outcomes of other contests, such as sports matches, so it could be expected to apply to political contests too. The relationship is captured by the bold arrows in Figure 4.1, which reproduces the theoretical model developed in Chapter 2.[1] I discuss this effect, and the limited existing evidence of it, in the second section.

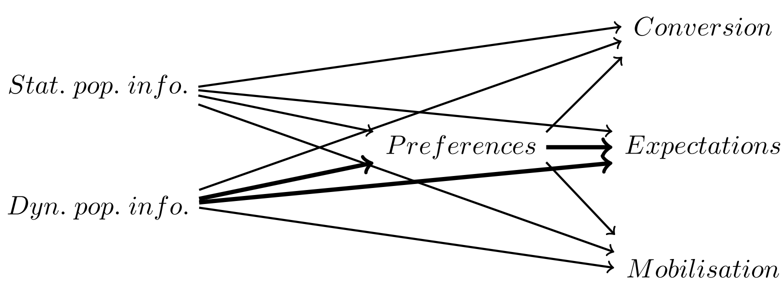


Figure 4.1: Causal model of the bandwagon effect, with the relationship studied in this Chapter in bold.

[1] Note that the bold arrows capture the 'total effect' of dynamic popularity information on expectations. The total effect is what I measure here. The mediating role of preferences is not tested directly, but it cannot be ruled out as part of the overall effect. I do however factor in the direct effect of preferences on expectations in Study Two.

The chapter then presents novel empirical evidence of this effect in action. Across two experimental studies with representative samples of UK respondents, I show that people with information indicating that a second-placed party has recently gained ground in the polls think that this party has a better chance of winning, compared to those who just know that it is in second place. Generally, people appear to see the leading party as more likely to win, but above and beyond the current standings of the parties, dynamic popularity information affects expectations. Study One establishes this in an abstract, stylised election scenario and also suggests that the greater this momentum, the larger the effect.

Study Two replicates the finding that dynamic popularity affects expectations, but provides some nuance by showing that in 'real-world' contexts, these effects are still significant but less pronounced. This, I argue, is because people are able to factor in their existing information and predispositions – the two components of opinions, as discussed in Chapter 2 – when processing the new popularity information. When it is a familiar (UK) context, the effect is slightly reduced. When it is an unfamiliar (Canadian) context, the effect is considerably reduced. I interpret these differences in terms of levels of information. People who voted for different parties in the last general election also report substantially different expectations from each other, in line with wishful thinking. These differences are far more apparent in a UK election, where people are likely to care strongly about the outcome, than in a Canadian election, where they are likely to care much less. I therefore interpret these differences in terms of the strength of predispositions.

Static Popularity Information and Expectations

When trying to work out how well a party is going to perform at an election, it seems reasonable to imagine that people will think about how popular that party is – how many people want to vote for it? This perception of how popular a party is comes largely from information. Maybe a voter has been following vote intention polls, or knows who lots of her friends are going to vote for, or knows which candidate is creating a buzz on Twitter, or has heard political commentators make bold predictions on talk shows. All of this is popularity information. As discussed in Chapter 2, a key assumption in much bandwagon research is that people's expectations are indeed based on popularity information. Consistent with this, scholars have shown that voters' expectations are influenced by a range of popularity cues, such as the slant of media coverage (Zerback, Reinemann, and Nienierza 2015), how well parties have performed at past elections (Irwin 2002), current candidate performance in an ongoing nomination contest (Abramowitz 1989; Abramson et al. 1992; Brady and Johnston 1987), and – above all – what the polls say (Blais and Bodet 2006; Faas, Mackenrodt, and Schmitt-Beck 2008; Irwin 2002; Schmitt-Beck 1996). The claim made in Chapter 2, captured as $Static\ popularity\ information \rightarrow Expectations$, holds water.

More specifically, Irwin and Van Holsteyn (2002, 101) demonstrate that, in the 1994 Dutch election, greater attention to politics and elections brought expectations 'into the range set by the polls.' Similarly, Blais and Bodet (2006) find that expectations align more closely with the polls for those who are more involved and interested in a campaign. This receives further support from Meffert, Huber, Gschwend and Pappi (2011, 809), who find a 'strong positive effect of political knowledge on the quality of expectations,' while also finding that engagement with polls 'improved' expectations. Lavrakas, Holley, and Miller (1991) further found that 95 percent of voters knew who was leading in the 1988 U.S. presidential election and attributed this knowledge to pre-election polling.

Multiparty and proportional representation contexts make this slightly more complex. In these cases, it is rare for any party to form a government on its own, so there is not such a clear link between the parties' standings in the polls and which are likely to be in government. For example, recently, in the immediate aftermath of the 2021 German federal election, it was not entirely clear which of the two leading parties (the SPD and the CDU) would lead a government coalition – confusion that was complicated further by the fact that one party appeared to have more votes, while the other had more seats. In general, it is nonetheless possible to predict which coalitions are likely to be formed by using the polls (Stoetzer and Orlowski 2020), raising the question of whether voters use the information in this way.

Recent research demonstrates that the availability of polling information is indeed a key driver of accurate expectations, helping voters to work out which coalitions will be formed, because it indicates which particular combinations of parties could collectively accumulate a majority of seats (Bowler, McElroy, and Müller 2021, 4–5). Explaining this link between the availability of polling information and the accuracy of expectations, recent evidence shows that voters in these contexts – though particularly those who are more politically knowledgeable – are good at recalling how well parties are polling and this improves slightly with more exposure to polls (Zerback, Reinemann, and Barnfield 2021). In multiparty contests, popularity information is, however, 'only part of the relevant informational context' given the importance of 'pre-electoral coalition signals' (Bowler, McElroy, and Müller 2021, 5; see also Gschwend, Meffert, and Stoetzer 2017).

However, often, it also appears that voters are swayed more by how they feel towards candidates or parties than by any such objective information. Granberg and Brent (1983, 478), for example, find that 'by a ratio of about 4:1, with relatively little overall variation across years, people tend to expect that their preferred candidate will win.' This fits in with the well-established empirical finding, across electoral contexts, that voters engage in 'wishful thinking' (Babad 1995; Babad and Yacobos 1993; Bartels 1985; McAllister and Studlar 1991; Meffert et al. 2011; Stiers and Dassonneville 2018), as introduced in Chapter 2. Indeed, this finding dates back at least as far as the 1930s (Hayes 1932). It is echoed as recently as by Mongrain (2021), according to whom wishful thinking is the *main* determinant of electoral expectations. Bowler, McElroy, and Müller (2021) also find robust evidence of wishful thinking in coalition expectations – voters are more likely to predict that coalitions they want in power are more likely to end up in power. This wishful thinking effect has consistently proven a major obstacle to observational research on the bandwagon effect, complicating the link between popularity information and expectations – a point explored in depth in Chapter 5. Study Two of this chapter also provides evidence of wishful thinking in action.

Dynamic Popularity Information and Expectations

A further complication, and my focus in this chapter, is that there is a difference between polls, popularity information and opinions about what is going on *right* *now*, and the likely outcome of an election at some date in the *future*. Information and expectations change over time. One study addresses this fact in the context of the Californian public ballot on the legalization of cannabis for recreational use – Proposition 19 (Krizan and Sweeny 2013). Although polling initially suggested the ballot might pass, support dropped substantially as the vote approached, eventually getting 46.2% of the vote and thus failing to pass. People 'varied greatly not only in their initial electoral forecasts, but also in the pattern of change in these forecasts' and while wishful thinking largely determined the initial, static state of these expectations, there was a 'general trend for people on both sides to relinquish optimism about their desired outcome as the election drew near' (Krizan and Sweeny 2013, 712). Particularly relevant here is the observation that 'well-informed voters were likely to lower their expectations regarding the measure's passage as the vote neared, *in line with polling results*' (Krizan and Sweeny 2013, 706, emphasis added). This suggests that changes over time matter for the development of expectations.

Expectations are opinions about the likelihood of *future* electoral performance (Zerback, Reinemann, and Nienierza 2015). The current, static popularity of different options in a given contest right now is a sensible basis for these expectations, absent other information, but time still has to pass between this present time point and the future event. The trends, over time, in the popularity of these options therefore become salient. Krizan and Sweeny (2013) track changes in expectations over a time period in which the trend of declining support for cannabis legalisation took it from net approval, to net disapproval. This means that the influence of momentum cannot be distinguished from the switch of legalisation going from being the *most* popular to the *least* popular option. Dynamic and static popularity are indistinguishable here. It could be that people's expectations changed not just because the proposition was becoming less popular relative to its own previous popularity, but because of this switch that changes its position relative to the other available option. Yet it is reasonable to expect that the momentum, in itself, without depending on a change in the relative ranking of different options, would alter voters' expectations too. Dynamic popularity information provides a basis on which people can in a sense 'map forward' the current popularity of the available options, inferring how this popularity is likely to change between now and the future time on which their expectations are based. If a party has momentum now, we might imagine that this momentum will carry it through, in the future, to overtake its rivals and outperform them on polling day.

Notably, social psychological research demonstrates that people think in this way when predicting how behavioural norms will develop in the future. For example, Mortensen et al. (2019, 205) find that information suggesting a minority of people were willing to donate to an environmental organisation, but that this minority was growing, 'led participants to project that financial donations would be more common in the future.' Psychologists argue that this could occur because the growth in popularity of a behaviour leads 'observers to infer that whatever factors had loomed large as barriers to change do not, in fact, prevent change' (Sparkman and Walton 2019, 238). When minority support for something is growing, this changes people's perceptions of what is possible and likely in the future.

This reasoning can also be seen elsewhere, for example in sports prediction. Imagine a football match is taking place in which Team A is leading 2-1 at half time. Do you think Team A is likely to win? Does that prediction change somewhat if you learn that Team B scored just before half time? Team B has the momentum. The commentators will tell you that this completely changes the outlook of the match. Football fans endorse this view even though research finds that there appears to be no specific benefit to scoring just before half time (Gauriot and Page 2018). In such cases, it does not appear to be the simple fact that the score is 2-1, rather than 2-0, that matters. Instead, it matters that Team B scored *most recently* – Team B's change in performance over time, not its number of goals relative to Team A, tells us that it has momentum.

In tennis, breaking an opponent's serve can signal momentum and change people's predictions of the end result.[2] Here though, psychologists have found that 'the breaking players' probability of winning a game increases after converting a break point, which provides evidence for momentum' (Meier et al. 2020, 1). So in this case, recent performance does seem concretely to change players' fortunes in the match as a whole. Crucially, this research shows that this happens not because players' positions relative to each other on the scorecard change, but because of 'psychological momentum' – the player who has 'momentum' gets a psychological boost. The flow of performance over time works independently of how many points the players have, in total, relative to each other. This can be demonstrated by the fact that the change in a player's probability of winning the game owed to winning a break point is lost when play is subsequently interrupted, due to weather conditions (Meier et al. 2020). The break in play draws a halt to the momentum a player has, even though when they return to the court they still have just as many points as before.

[2] In tennis, it is somewhat rare for players to 'break' others' serves (win the game when the other player gets to take the first shot). So, when they do, this is a signal that things might be moving in their favour.

Moreover, sports bettors have a tendency to support basketball teams on 'winning streaks' (Arkes 2011) – even though these 'streaks' are usually no different from what would happen if match outcomes were completely random (Vergin 2000). During these matches themselves, the 'hot hand fallacy' also dictates that people see basketball players who have just scored multiple times as more likely to score on their next attempts (Ayton and Fischer 2004).[3] This affects all levels of observers of basketball – players, coaches, and fans alike (Cohen 2020). It is not how many points players have scored that really matters, but whether or not they are currently on a scoring 'streak.' These examples all demonstrate that we readily incorporate considerations about momentum into our thinking about what is likely to happen in a contest, both in cases where that momentum appears to matter empirically and where it appears not to.

A simple example of how this could apply to electoral politics can be seen in the Labour Party's performance in the 2017 UK general election campaign, discussed in Chapter 1. Theresa May, then Prime Minister and leader of the Conservative Party, which held a majority of seats in the House of Commons, called the election at a time when the polls gave her a commanding lead over Jeremy Corbyn's Labour (Hobolt 2018; Prosser 2018). Yet the Labour Party, throughout the campaign, saw its share in the polls grow considerably and quite rapidly without ever convincingly polling as more popular than the Conservatives (Burn-Murdoch, Stabe, and Leach 2017).[4] If people's expectations about the likelihood of a Labour victory changed over the campaign, this could very feasibly be due to its dynamic popularity. Its

[3] Whether or not the hot hand really exists is the subject of debate among psychologists and mathematicians (Daks, Desai, and Goldberg 2018; Gilovich, Vallone, and Tversky 1985; Miller and Sanjurjo 2018).

[4] One poll by the non-British Polling Council affiliated firm Qriously gave Labour the lead on the eve of the election (Medeiros 2017).

static popularity necessarily changes, so this could of course be what drives changes in expectations. For example, someone might be asked how likely they think it is that Labour wins the upcoming election, on a day on which it is polling at around 38%, and this 38% could be what causes her to say that it has a 4/10 chance, rather than the fact that it has grown six points over the past week. But given that the relative ranking of the parties never changed – Labour was always trailing behind the Conservatives overall – and reporting on the election emphasised Labour's *growth* (Ford 2017), it is very plausible that dynamic popularity had an effect either in addition to, or instead of, any static effect (van der Meer, Hakhverdian, and Aaldering 2016). That is, in the above example, it seems likely that the fact Labour had grown six points would influence this voter's answer.

Indeed, as Figure 4.2 shows, the average perceived probability of a Labour victory in the 2017 general election increased noticeably throughout the campaign. Those who were interviewed on the eve of the election tended to rank Labour's chances up to 1.5-2 points higher on a 0-10 scale than those interviewed in the early days of the campaign. There is also a clear tendency for Labour identifiers to rate its chances higher than those who do not identify with the party – especially Conservative identifiers – by a constant margin of at least one point, indicating significant wishful thinking.

The argument that dynamic popularity matters in this way leads to a principal hypothesis:

H_1 Dynamic popularity information has a positive effect on electoral expectations: knowing a party's popularity is increasing in the polls makes people see it as more likely to win an upcoming election.

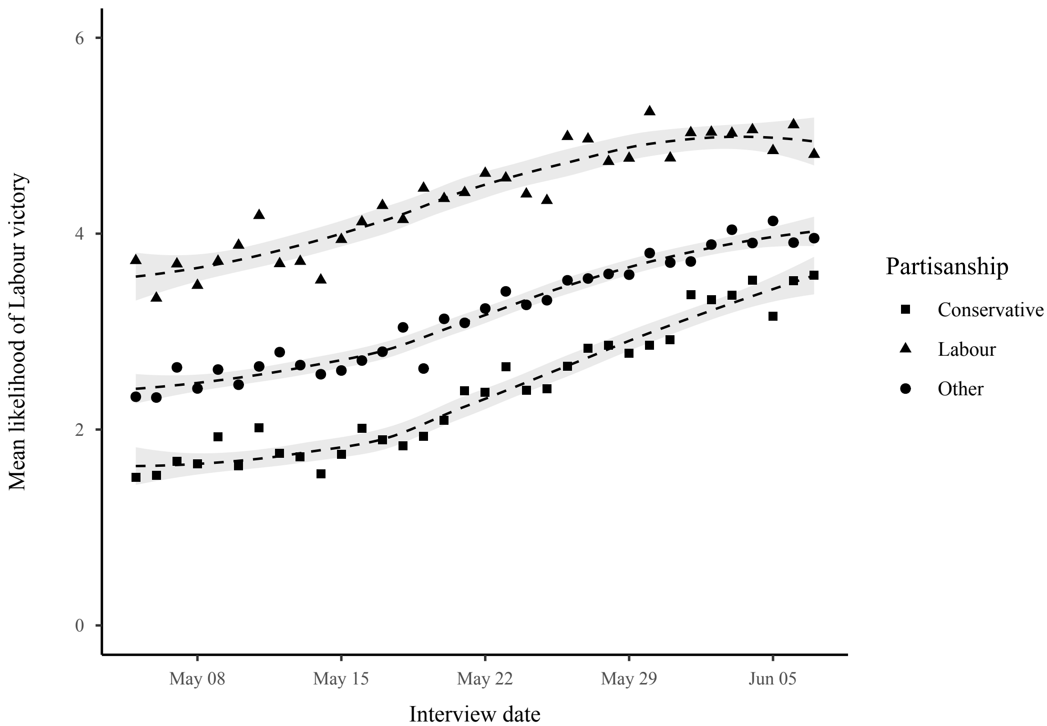


Figure 4.2: Mean expected likelihood of Labour Party victory (0-10) during the 2017 UK general election, among Labour partisans and non-partisans, with loess smoothed trends. Data: British Election Study Internet Panel 2014-18, Wave 12.

If this argument holds, then it should not only be the case that learning a party has momentum makes people think it has a better chance of winning. It should also mean that *more* momentum implies a greater chance of winning. The greater a party's momentum, the easier we might think people will find it to imagine 'mapping forward' its current, static popularity, to a greater popularity in the future. Perhaps this will even make them imagine the current rankings of the parties will change by election day, such that a given party with momentum will overtake an opponent. This leads to a second hypothesis:

H_2 The more a party's popularity is increasing in the polls, the greater the effect of this increase will be on its perceived chances of winning an upcoming election.

Empirical Analysis

I conducted two studies to test these hypotheses. I received ethical approval for the experiments in these studies from the Queen Mary Ethics of Research Committee (QMERC2158 and QMERC2494).

In Study One, three separate randomised controlled trials (RCTs) were fielded to a representative sample of UK respondents via YouGov in 2019 and 2020. In these experiments, I randomly split people between a control condition in which they had no information about 'momentum' and a treatment condition in which, depending on which experiment they took, they were told that a hypothetical second-placed party had grown by either 3, 6 or 9 points in the polls. Study One therefore provides a clean test of both H_1 and H_2.

In Study Two, a two-way factorial experiment was fielded in June 2021 to a large, again representative sample of UK respondents, also via YouGov. This followed the same format as Study One, except that in addition to randomisation between treatment conditions, people were also randomised to see hypothetical parties or parties with real-world labels (Conservative/Labour or Conservative Party of Canada/Liberal Party of Canada). As explained below, Study Two therefore takes H_1 and tests the extent to which it is robust to voters' use of their own information and predispositions. These studies are summarised in Table 4.1. I describe these designs in detail and provide a rationale for them below.

Study One

Experimental Design

In Study One, respondents answered the following question. Text in bold was only included for the treatment group – the control group saw all of the text presented here except for the text in bold. That is, everything was the same for both groups except for whether or not they received the text in bold.[5]

> Imagine a general election, contested between two parties, A and B, which is to be held in the next few weeks. Recent opinion polling has shown the two main parties on the following vote shares **(with the change over the past month in brackets)**.
>
> Party A - 44% **(-1)**
>
> Party B - 40% **(+*M*)**
>
> How likely is it that Party B wins the election?
>
> Very unlikely to win 0-10 Very likely to win

I carried out three versions of the experiment, with the figure indicated by M above varying in each. In the first experiment, the M figure was set at 3, in the second 6, and in the third 9. It is the effect of this M figure – which, again, was only available to the treatment group – that I am looking to estimate, because it represents dynamic popularity information. These experiments were fielded in YouGov's daily omnibus surveys to a sample of British respondents in 2019 and 2020. They were answered by 1632 (treatment 817, control 815), 1868 (treatment

[5] The text displayed in bold here was not in bold in the experiment itself.

Table 4.1: Summary of studies.

	Type	N	Date
Study 1			
+3 Experiment	Standard RCT	1632	April 2020
+6 Experiment	Standard RCT	1868	October 2019
+9 Experiment	Standard RCT	1659	April 2020
Study 2			
Factorial Experiment	Two-way Factorial	6724	June 2021

914, control 954) and 1659 (treatment 837, control 822) respondents, respectively.

The guiding principle behind the design was to provide a clean, stripped-back test of the hypotheses above, which was entirely accessible to respondents. The simple idea is that respondents in the control group only have static popularity information to go by when deciding how likely it is that Party B will win, so this should be all they use to do so. Those in the treatment group also have dynamic popularity information. Because this is the only thing that differs across the two groups, any significant or systematic differences in their responses should be down to this.

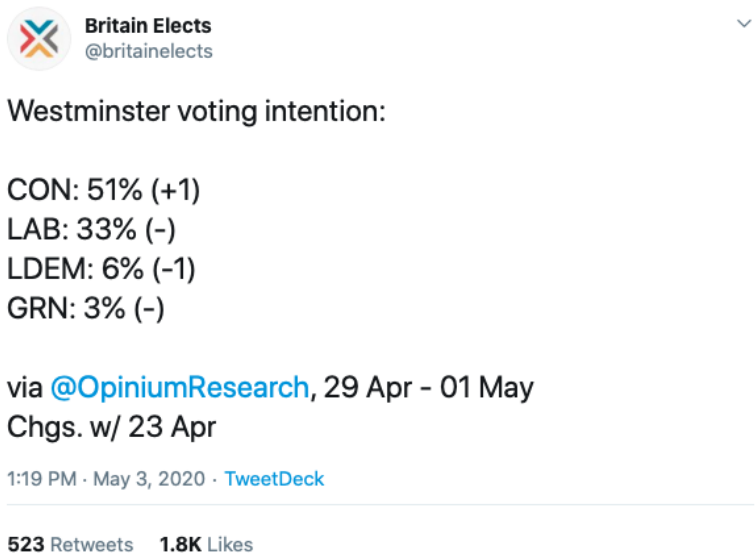
The general presentation of the poll results is similar to how polls are typically presented in tweets, as in Figure 4.3. This presentation is simple while also therefore offering the benefit of being familiar to many respondents. The party with momentum, B, was fixed at a current share of 40%, in order to hold static popularity constant and allow the effect of dynamic popularity to stand alone. This figure and Party A's share of 44% were chosen to remove any ambiguity about any other parties having a chance of winning, while still representing attainable vote shares in a UK general election.[6]

The dynamic popularity figure was varied across the three experiments in order to establish whether *more* momentum rather than *mere* momentum has an influence on the size of the effect of dynamic popularity – as H_2 says. Party A's dynamic popularity is set at (-1) in order to keep it as close to constant as possible without introducing any confusion. For example, polls presented in this format often use (-) to indicate no change (see Figure 4.3), but this is ambiguous and might have made respondents think something was missing from the experiment. Awkwardly setting the figure to (+0) might have had a similar effect, while +1 would indicate that Party A *also* has momentum, albeit very limited. The (-1) level gives the impression of stagnation without this seeming severe.

The treatment values were set at +3, +6, and +9 for two reasons. First, this gives a good range of strength in the experimental stimuli. While +3 is not an uncommon fluctuation in poll results, +9 is noticeably strong and unexpected in most cases. The +6 value sits halfway between these. This range of treatments therefore allows me to get a meaningful idea of how the effect of momentum varies depending on how much there is. Second, and related, it is commonly said that the typical 'margin of error' of a poll is +/-3% (Mercer 2016).[7] Using multiples of three

[6] In the 2017 UK general election, the Conservatives received approximately 42% of the vote to Labour's 40%. In 2019, the Conservatives received approximately 44% of the vote.

[7] Sturgis and Kuha (2018) explain that this is a 'heuristic based on an 'as if' assumption that the poll was carried out using a simple random sample of 1000 respondents and an estimate of 50%. Were this an accurate description of the poll then the sampling variability would indeed be close to +/- 3% at the standard level of statistical confidence'. In practice, the margin of error of a poll is nearly always larger than this. The error in the difference between two parties' level of support is


Britain Elects
@britainelects

Westminster voting intention:

CON: 51% (+1)
LAB: 33% (-)
LDEM: 6% (-1)
GRN: 3% (-)

via @OpiniumResearch, 29 Apr - 01 May
Chgs. w/ 23 Apr

1:19 PM · May 3, 2020 · TweetDeck

523 Retweets 1.8K Likes

Figure 4.3: Tweet by @britainelects showing typical Twitter presentation of poll results.

therefore means that the treatments can be read in terms of heuristic margins of error. While +3 is very feasibly just noise, +6 is towards the upper limit of change that could be brought about by such noise, and +9 is large enough that it almost certainly means a meaningful change has occurred in the electorate.[8] This not only allows my findings to speak more to the terms in which developments in the polls are discussed, but also to comment on how responsive people are to potentially statistically meaningless, versus substantively significant, dynamics (Bailey and Barnfield 2021).

Appendix C provides additional information on the design of Study One, including

larger still.

[8] For example, if a party polls at 40%, this has a margin of error of 37-43%. This implies that it could feasibly have polled at 37%. But next time it could poll at 43%, even though the underlying true vote share has not changed. Across two polls, it could go from 37% to 43% – an increase of six points – while always having a true vote share of 40% – because of sampling variability.

a power analysis (Appendix C.1) and a randomisation check (Appendix C.3). Appendix C.6 also presents the results of Bayesian ordinal regression models in order to verify that the results presented here are not due to misspecification of the outcome variable (a 0-10 ordinal scale) as a continuous or 'metric' measure (Bürkner and Vuorre 2019; Liddell and Kruschke 2018).

Results

Table 4.2 presents the raw results of linear regression analyses. In the first three columns, the 'Intercept' captures the average answer given in the control group in the specified experiment, and the 'Treatment' effect captures the change in this average response associated with being in the treatment group, in that same experiment.

All three experimental treatments had a substantive and statistically significant effect on average expectations for Party B's performance. Looking at the treatment effects in the first three columns of Table 4.2, knowing Party B had momentum made people, on average, rate its chances of victory approximately half a point higher, or more, on a 0-10 scale. This is reflected visually in Figure 4.4, where expectations are noticeably higher in the treatment group, in every experiment. The faint horizontal line in Figure 4.4 represents the midpoint of the scale (5/10). On average, in the +6 and +9 experiments, the treatment was sufficiently convincing to push people, on average, over this midpoint. This represents clear support for H_1.

The effect grows in line with the strength of the experimental stimulus. While, on average, those who knew Party B had grown by three percentage points in the polls rated it approximately 0.5 points higher, the equivalent effect in the six

Table 4.2: Linear regression results, Study One.

	+3 experiment	+6 experiment	+9 experiment	Pooled
Intercept	4.338***	4.702***	4.558***	4.338***
	(0.069)	(0.056)	(0.064)	(0.065)
Treatment	0.499***	0.673***	0.839***	0.499***
	(0.098)	(0.081)	(0.090)	(0.091)
+6 experiment				0.364***
				(0.088)
+9 experiment				0.221**
				(0.091)
Treatment:+6				0.174
				(0.125)
Treatment:+9				0.340***
				(0.129)
Observations	1,632	1,868	1,659	5,159
Adjusted R^2	0.015	0.035	0.050	0.043

Note: *p<0.1; **p<0.05; ***p<0.01

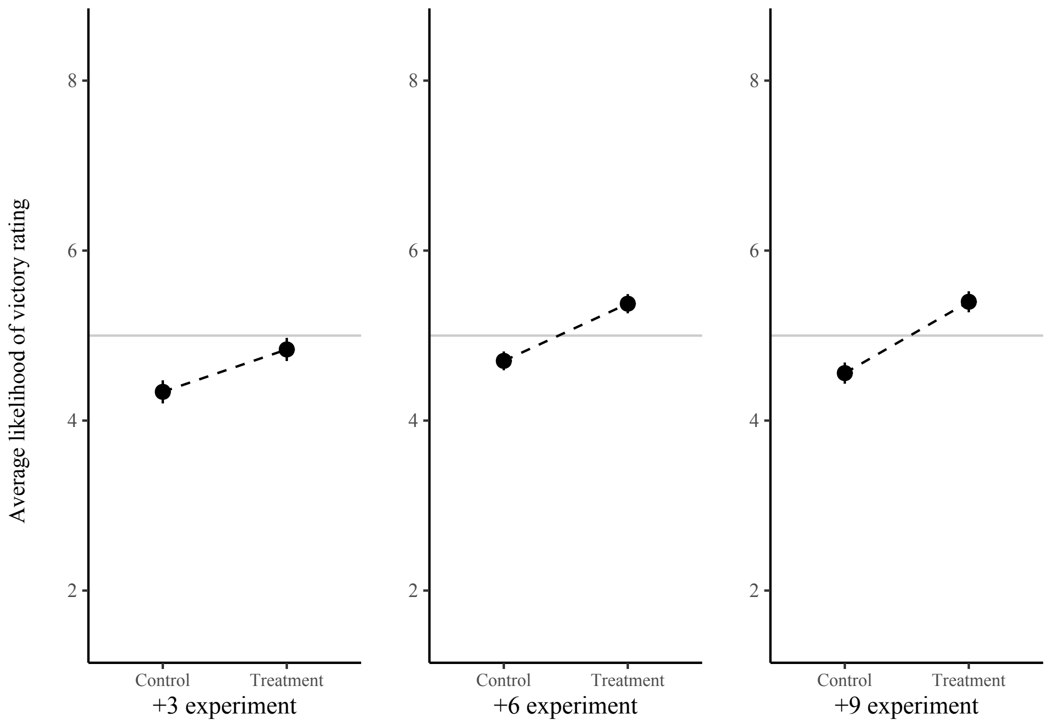


Figure 4.4: Treatment effect visualised as change in model-predicted average response in treatment and control. Vertical bars represent 95 per cent confidence intervals.

point treatment was just under 0.7, and just over 0.8 in the nine point treatment. Again, this is reflected visually in Figure 4.4, where the stronger the stimulus, the steeper the slopes capturing the treatment effect. However, just because two effects are statistically significant and different, they are not necessarily statistically significantly different (Gelman, Hill, and Vehtari 2021, 58–59). So in order to assess whether these differences are significant, the fourth column in Table 4.2 pools all three experiments together and interacts the treatment effect with an indicator of which experiment it is drawn from. This means that 'Treatment' now captures the effect in the +3 experiment – which is why the coefficient is the same as in the first column – while ‘Treatment:+6’ and ‘Treatment:+9’ measure the difference in this effect across the other two experiments.[9] This reveals that there is a highly statistically significant difference between the treatment effect in the +9 experiment and that in the +3 experiment. This provides clear support to H_2. However, I cannot rule out that there is no systematic difference in the +6 experiment compared to the +3 experiment.

Study Two

It is possible that these clear, consistent effects only emerge because of the artificial experimental context. The fact that the election in Study One is highly stylised and takes place between two undifferentiated parties means that respondents have no background knowledge with which to contextualise the treatment information. They also do not have any reason to *want* Party B to win, which means that wishful thinking should not operate here. However, in the real world, when people see the results of polls, they usually do have background knowledge and they usually do want certain parties to do better than others. As discussed in Chapter 2, opinions are a ‘marriage of information and predisposition’ (Zaller 1992). When voters form opinions on the basis of new information, they rarely do so as a blank slate. They bring their information and predispositions to the table. Study Two allows them to do this.

If respondents had background information about Party A and B, this might have an impact on their responses in the experiment. For example, consider a scenario

[9] The ‘+6 experiment’ and ‘+9 experiment’ variables measure the difference in the intercepts – average expectations in the treatment group – across the experiments.

in which Party B has a very geographically inefficient distribution of support. Lots of its voter base is concentrated in a few constituencies or voting districts. Election after election, Party B secures large shares of the vote which are not transformed into large numbers of seats in parliament. A large vote share does not necessarily indicate that it has a particularly good chance of winning, and this changes little if it has momentum in the polls. While this is an extreme example, it serves to demonstrate that a minimal-information context such as that used in Study One could produce effects that are larger than they would be in the real world. Study Two therefore introduces real-world contexts in order to account for this.

Also, as explained above, and explored further in Chapter 5, ‘wishful thinking’ is a key driver of people’s electoral expectations – they tend to say that the parties they like have a better chance. This might have two effects relevant here. First, those who support a party will have higher expectations for its performance, on average, than those who do not support that party. This would correspond, above, to people who support Party B giving it a higher likelihood of winning the election, regardless of treatment status. Second, those who support a party might be more responsive to information that is positive for that party. This would correspond to people who support Party B being more strongly affected by treatment information about its momentum. This could mean that for large segments of the population, who do not support a party that has momentum, the effects observed above do not hold in many cases. Study Two therefore splits respondents by their party support in order to account for this.

Experimental Design

Study Two mirrors Study One but introduces an additional dimension. As well as being randomly assigned to a treatment or control condition, respondents are also randomly split between three different groups: one in which the parties are again labelled A and B (A/B condition), one in which they are labelled Conservative and Labour (UK condition), and one in which they are labelled Conservative and Liberal (Canada condition). This gives a total of six different conditions, summarised in Table 4.3. The experiment was taken by 6724 respondents from the YouGov panel.

The purpose of including the two additional conditions (UK and Canada) is to factor in both of the concerns just discussed. Firstly, it allows me to examine how levels of information influence the outcome. In the A/B condition, respondents know that all of the relevant information is given to them. The election scenario is completely imaginary, and they are aware that Parties A and B do not exist outside of this made-up polling environment. This means there is no extra information for them to bring into their opinion. In the UK condition however, it is clear that there is more to it. The parties really exist, and there is potentially a lot of information out there that further explains how likely it is that the Labour Party would win the election, other than what I am telling participants in the experimental design. As a sample of UK voters, respondents have access to a lot of this information, so they are able to use it. In the Canada condition, respondents should again recognise that there is more to it, but they are not expected to have much knowledge about real Canadian politics, so there is considerable uncertainty. As such, this design considers how people form their expectations in response to new information in

Table 4.3: Summary of factorial conditions.

	Control	**Treatment**
A/B		
Text	Imagine a general election is going to be held in the next few weeks, and recent opinion polling shows the two main parties, Party A and Party B, on the following vote shares.	Imagine a general election is going to be held in the next few weeks, and recent opinion polling shows the two main parties, Party A and Party B, on the following vote shares (with the change over the past month in brackets).
Popularity	Party A - 44% Party B - 40%	Party A - 44% (-1) Party B - 40% (+6)
Question	How likely is it that Party B would win the election?	How likely is it that Party B would win the election?
UK		
Text	Imagine a general election is going to be held in the next few weeks, and recent opinion polling shows the two main parties, the Conservative Party and the Labour Party, on the following vote shares.	Imagine a general election is going to be held in the next few weeks, and recent opinion polling shows the two main parties, the Conservative Party and the Labour Party, on the following vote shares (with the change over the past month in brackets).
Popularity	Conservative Party - 44% Labour Party - 40%	Conservative Party - 44% (-1) Labour Party - 40% (+6)
Question	How likely is it that the Labour Party would win the election?	How likely is it that the Labour Party would win the election?
Canada		
Text	Imagine a general election is going to be held in the next few weeks in Canada, and recent opinion polling shows the two main parties, the right-wing Conservative Party and the left-wing Liberal Party, on the following vote shares.	Imagine a general election is going to be held in the next few weeks in Canada, and recent opinion polling shows the two main parties, the right-wing Conservative Party and the left-wing Liberal Party, on the following vote shares (with the change over the past month in brackets).
Popularity	Conservative Party of Canada - 44% Liberal Party of Canada - 40%	Conservative Party of Canada - 44% (-1) Liberal Party of Canada - 40% (+6)
Question	How likely is it that the Liberal Party of Canada would win the election?	How likely is it that the Liberal Party of Canada would win the election?

three different contexts: when there is no other relevant information (A/B), when there is other relevant information and they have lots of it (UK) and when there is other relevant information but they have very little of it (Canada). This, in turn, compares the abstract scenario in Study One to how people make sense of polling information in two realistic scenarios: when it pertains to elections in their own polity, and when it pertains to elections in other polities.

Secondly, including the UK and Canada conditions enables me to examine how voters' predispositions influence the outcome. In the A/B condition, voters of different parties have no clear reason to differ in their responses. In the UK condition however, the sample of UK voters is likely to have clear preferences about who wins. Labour voters should want Labour to win and Conservative voters should want the Conservatives to win, for example. Wishful thinking holds that these predispositions will influence their expectations of how likely it is Labour wins. Splitting respondents by their vote choice and observing how they differ across the conditions allows me to assess this and the extent to which the effect of dynamic popularity is robust to these biases. The Canada condition shows whether wishful thinking also has this effect in a situation where different party voters might still have preferences about the outcome, but these preferences should be somewhat weaker. In order to clarify the ideological positioning of the parties, the wording (see Table 4.3) explicitly points out to respondents that the Conservative Party is more right-wing and the Liberal Party is more left-wing.

I chose to use Canada as the additional condition in the design primarily because very few British voters are likely to have a good level of knowledge about Canadian politics, but its two main parties are broadly similar in ideological terms to those who dominate UK politics. This allows the distinctions above to be made: British voters should have much lower information than they do in a UK context, and their preferences about which party wins should be predictable but weaker than in the UK. That is, an average Labour or Conservative voter might not be expected to know anything about Canadian politics, but should know which party she would rather see win an election, when told their names and ideological position. In addition to this, Canada has a very similar electoral system to the UK, facilitating a comparison across the two. It would be more difficult to compare effects between the UK and a proportional system such as the Netherlands, for example, owing to the ambiguity of what 'winning' means across contexts (Stiers, Daoust, and Blais 2018).

The wording reported in Table 4.3 is slightly different from the wording in Study One, even in the A/B condition. This is to ensure that the wording in the UK and Canada conditions makes it as clear as possible that these election contexts are purely imaginary. I discuss the importance of such concerns in the essay on the ethics of experimental deception in Appendix A. As the wording was therefore adjusted in the UK and Canada conditions, for consistency across the study, the A/B wording was also adjusted accordingly.

In each case, the numbers used for both the static and dynamic popularity of the two parties is the same as in the +6 experiment from Study One. That is, the leading party (A/Conservative/Conservative) has 44% of the vote, while the trailing party (B/Labour/Liberal) has 40%. Under treatment, the leading party has changed by -1, and the trailing party has changed by +6. This is based on three

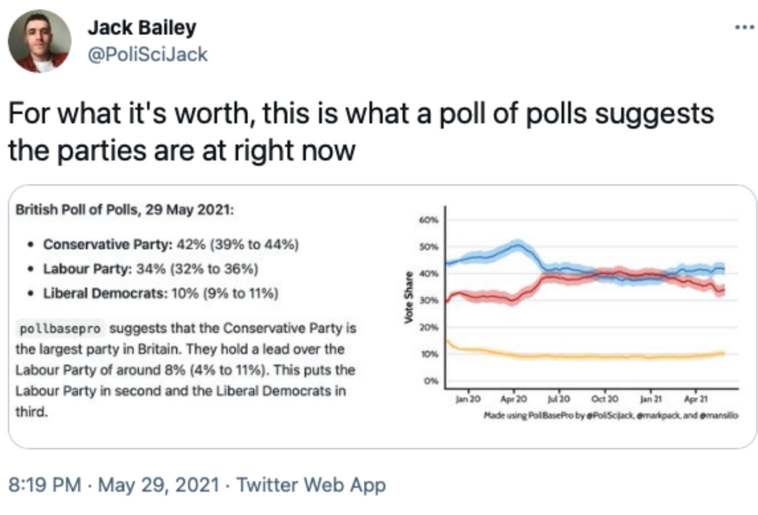


Figure 4.5: Tweet by @PoliSciJack showing most recent poll of polls prior to experiment field dates.

key considerations. First, it provides a direct comparison with Study One. This is useful not only for the purpose of verifying that the effect can be replicated, but also for establishing whether the effect holds with the slightly adjusted wording (discussed just above). Second, Study One showed that the +6 treatment had quite a large effect on responses. For the purposes of experimental power, this is preferable to using a treatment (e.g., +3) shown to have a smaller effect, given the increased difficulty of detecting effects in this factorial design (see Appendix C.2 for power analysis). Yet, it is arguably also preferable to using the +9 treatment, which may have looked implausible to respondents in, for example, the UK condition. Third, poll aggregator estimates prior to the conduction of the experiment suggested that the Labour Party had approximately 34% support – as shown in Figure 4.5.[10] This implies that, for it to poll at 40%, the party would need to increase its share by six

[10] These estimates are based on 'PollBasePro' data (see Bailey, Pack, and Mansillo 2021).

points, in line with the +6 treatment. The Liberal Party of Canada had also recently polled in this region, with poll aggregators suggesting it had approximately 35% support (Grenier 2021). This was therefore also the most 'experimentally realistic' treatment option (McDermott 2011).

In the Canadian case, this realism is compromised slightly by the fact that the experimental design flips the ranking of the parties according to polls at the time. That is, in reality, the Conservative Party of Canada is trailing behind the Liberal Party in vote intention polls, at the time of writing. I flip this here in order that in both cases, the trailing party is a more 'left-wing' party, such that wishful thinking would be expected to operate in the same direction. This facilitates analysis and inference. Given that respondents are likely to know relatively little about Canadian electoral politics, the loss of realism should be minimal.

Again, Appendix C provides additional information on the design of Study Two, including a power analysis (Appendix C.2) and a randomisation check (Appendix C.4), and the Bayesian ordinal regression models reported in Appendix C.6 verify that the results presented here are not due to statistical misspecification.

Results

Table 4.4 reports the overall results of the factorial experiment, using the full sample. This model uses an interaction between treatment and condition in order to capture the differences in behaviour across the conditions.[11] Here, 'Intercept'

[11] Note that this is just equivalent to ANOVA, the standard psychometric approach to analysing factorial experiments like this one. ANOVA just refers to linear regression where all of the predictors are categorical, and their interactions are measured as well as their main effects, all of which is the case here (Platt 1998). I use the language of linear regression to avoid introducing unnecessary

captures the average likelihood of victory rating given to Party B in the control group in the A/B condition. This is very close to the rating given to Party B under control across all three experiments in Study One. 'Treatment' captures the change in this rating brought about by being in the treatment group. This effect, at just over 0.8 on the 0-10 scale, is larger than that observed in the +6 experiment in Study One, but the standard errors indicate that the two effects are of a comparable size. Both of these things indicate that Study Two successfully replicates Study One, strengthening support for H_1.

Beyond this, Table 4.4 also crucially shows how these effects varied across the conditions. First, 'UK condition' captures the difference in ratings for the Labour Party versus ratings for Party B, in the control group. That is, in the UK condition, respondents in the control group (no dynamic popularity information), on average, gave the Labour Party a chance of approximately 1.2 points lower than comparable people gave Party B, on a 0-10 scale. 'Treatment:UK condition' then captures how the treatment effect differs in the UK condition, compared to the A/B condition. That is, in the UK condition, the effect of knowing the Labour Party had grown six points in the polls had a smaller effect, by approximately 0.26 (Table 4.5 below shows that is effect is approximately 0.6), than the same information had in the A/B condition. Essentially, people both see Labour as substantially less likely to win at a 40% vote share than a hypothetical party in a minimal-information election would be, and are slightly but significantly less swayed in this judgment by information that it is growing in the polls.

In the Canadian condition, the results are different. 'Canada condition' captures the difference in ratings for the Liberal Party versus ratings for Party B, in the

terminology that makes no meaningful methodological distinction.

Table 4.4: Linear regression results, Study Two.

Intercept	4.565***
	(0.061)
Treatment	0.828***
	(0.086)
UK condition	-1.206^{***}
	(0.086)
Canada condition	0.252***
	(0.085)
Treatment:UK condition	-0.267^{**}
	(0.122)
Treatment:Canada condition	-0.531^{***}
	(0.120)
Observations	6,724
Adjusted R^2	0.106
Note:	*p<0.1; **p<0.05; ***p<0.01

control group. This shows that people were inclined to rate the trailing party's chances higher when it was the Liberal Party of Canada, rather than Party B, by approximately 0.25 points. 'Treatment:Canada condition' describes the change in the treatment effect, showing that the information about momentum is substantially less influential in the Canadian condition than in the A/B condition, by approximately 0.5 points (Table 4.5 below shows this effect is just below 0.3).

These findings suggest that information levels are indeed significant in moderating how people use popularity information to form their expectations. Whereas the minimal but complete information of the A/B condition means that voters place a strong emphasis on popularity information, in the UK condition they are significantly less strongly affected by this information, seemingly because they have pre-existing knowledge about the Labour Party with which to contextualise the experimental treatment. This affects both their reliance on the static popularity information – they rate Labour's chances lower, regardless of whether it has momentum – and on the dynamic popularity information – the momentum treatment has a smaller effect. In the Canada condition meanwhile, people are less inclined to write off the Liberal Party's chances based on the fact it is trailing – they give it a significantly higher likelihood of winning, closer to 5/10. They are also much less swayed in this judgment by information that the Liberal Party has momentum, likely reflecting an awareness of the fact that there is a bigger picture out there about which respondents know very little.

Beyond this, the fact that dynamic popularity has a significant effect on expectations across all three conditions – see Table 4.5 – further reinforces H_1. On average, the finding that people's expectations for a trailing party are significantly higher when that party has momentum, is robust when tested with reference to 'real-world' parties.

In order to understand how wishful thinking and predispositions play into these findings, Table 4.5 reports 'average marginal effects' (AMEs; see Leeper 2018) drawn from a regression analysis in which treatment status is interacted with condition and respondents' 2019 vote choice. Specifically, respondents are broken down by whether they voted Conservative (N = 2328), Labour (N = 1729), or for another party (N = 1150) in the 2019 UK general election.[12] Owing to some missingness and invalid responses on this variable, this is a reduced sample of 5,207 respondents. In order to verify that behaviour in this reduced sample does not differ from the full sample, I also conducted a version of the first analysis, with no interaction with partisanship, using this reduced sample, reported in Appendix C.5. The results are very close to those reported in Table 4.4. I focus here on AMEs, and a visual representation of the effects in Figure 4.6, for ease of interpretation.

In addition to providing AMEs for the effects discussed above, Table 4.5 describes how they vary by party support. Each of these effects is captured by a dashed slope in Figure 4.6. In the A/B condition, treatment had substantial and highly statistically significant effects on everyone, increasing Conservative voters' ratings for Party B by just over 0.6, Labour voters' by just under 0.9, and supporters of other parties by approximately 1.25 points on the 0-10 scale.

In the UK condition, treatment again has a strong and statistically significant effect on everyone, increasing Conservative voters' ratings for Party B by over 0.6, Labour voters' by just under 0.5, and supporters of other parties by approximately 0.8 points. Figure 4.6 shows, however, that there are substantial differences in base-

[12] Each party group is almost exactly equally split across treatment and control in the reduced sample (Conservative: 1163 C, 1165 T; Labour 858 C, 871 T; Other: 575 C, 575 T).

Table 4.5: Average marginal effects of treatment.

Treatment Effect	AME
Average in A/B (pooled)	0.841∗∗∗(0.096)
Average in Canada (pooled)	0.288∗∗∗(0.093)
Average in UK (pooled)	0.617∗∗∗(0.095)
Average (pooled)	0.577∗∗∗(0.055)
Conservative in A/B	0.603∗∗∗(0.144)
Conservative in Canada	0.196 (0.137)
Conservative in UK	0.64∗∗∗(0.143)
Other in A/B	1.252∗∗∗(0.197)
Other in Canada	0.363∗(0.201)
Other in UK	0.772∗∗∗(0.205)
Labour in A/B	0.889∗∗∗(0.168)
Labour in Canada	0.363∗∗(0.161)
Labour in UK	0.482∗∗∗(0.163)

Note: $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

line levels of expectations for the Labour Party, net of these treatment effects. For example, Conservative voters give the Labour Party a considerably lower chance under both treatment and control than Labour voters do. The difference is approximately 1.5 points, which is larger than the difference made to either group by the treatment information. Supporters of other parties give it a chance approximately half way between the two. These overall differences paint a remarkably similar picture to Figure 4.2, which tracks real developments in electoral expectations for the Labour Party, on the same scale, when it had momentum in the 2017 UK general election campaign.

These results are in line with a wishful thinking effect, but one that does not primarily affect how responsive people are to the treatment information. Instead, Labour and Conservative voters adjust their expectations accordingly when learning that the Labour Party has momentum – they roughly 'update in parallel' (Bailey 2019b) – while other voters seem to be slightly more responsive to the information. However, the starting points from which people are adjusting are very different. For Labour and Conservative voters, this means that the dynamic popularity information does not cause their expectations to converge, but to remain distinct. It does, however, appear to cause other party supporters' expectations to converge closer to those of Labour voters.

The results are different in the Canada condition. First, only for Labour voters is the treatment effect statistically significant at the 95% level – estimated to be approximately 0.36, smaller than the effect in either the A/B or UK condition.[13] Across treatment and control, as shown in Figure 4.6, everyone's expectations

[13] The effect is estimated to be the same size for supporters of other parties, but as they are fewer in number, the standard error is larger and the effect is not statistically significant at the 95% level.

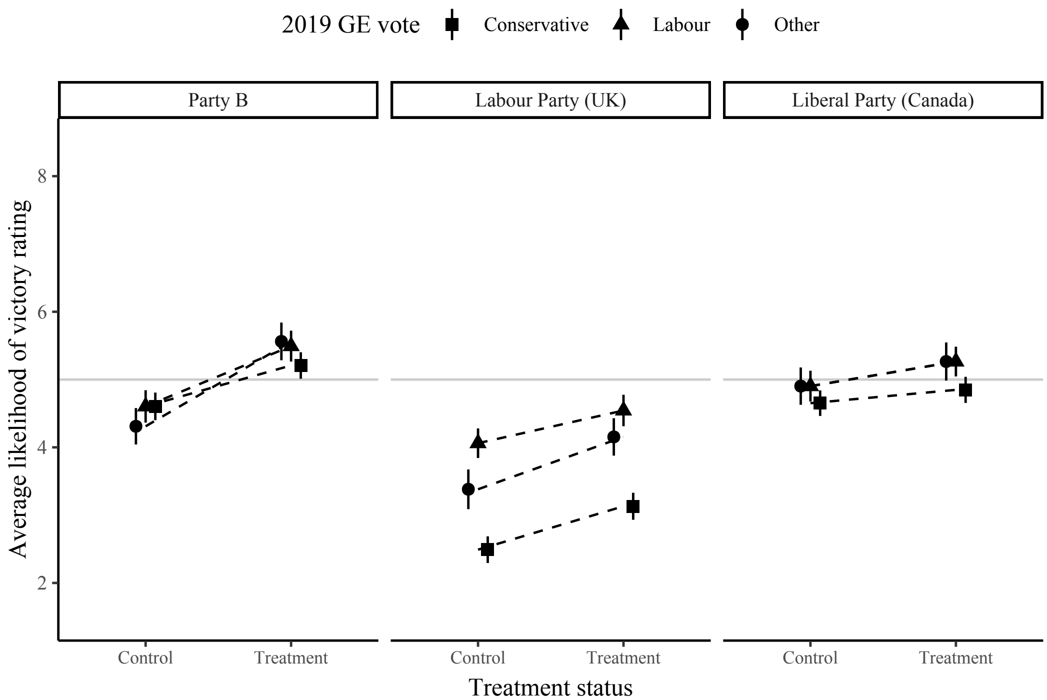


Figure 4.6: Treatment effects by party support in each condition, visualised as change in model-predicted average response in treatment and control. Vertical bars represent 95 per cent confidence intervals.

remain clustered very close to 5/10 (represented by the faint horizontal line), with Conservatives' expectations barely higher on average under treatment than under control. Conservatives also report slightly lower expectations overall, net of treatment.

These results appear to indicate that wishful thinking might be playing a minor role in affecting both overall responses and responsiveness to the treatment effect. Only Labour voters are significantly affected by information telling them that an explicitly 'left-wing Liberal Party' has momentum in the polls, and their expectations are also higher on average than those of Conservative voters.[14] Yet they do not appear to be set apart from other party supporters in the same way.[15] The fact that these differences are not particularly large is likely due to the fact that UK voters' preferences about the outcome of a Canadian election are relatively weak. Labour voters might rather the Liberal Party of Canada win than the Conservative Party of Canada, but it is not necessarily a major concern.

Discussion and Conclusion

Expectations are inherently about the future. Using information about how (statically) popular parties or candidates are right now is useful in forming these expectations, but inevitably gives only an imperfect picture. Dynamic popularity provides additional information by telling people how this support is changing, and can therefore be used to form expectations by indicating how it might *continue* to change. This claim emerged from the causal model explored in Chapter 2, when considering how electoral expectations fit into the overall picture of the bandwagon effect. It also reflects psychological tendencies in the way people predict future changes in behavioural norms and outcomes in sports matches. Before now, though, it had not been put to the test.

[14] The unresponsiveness of Conservative voters could also be related to the 'asymmetry hypothesis,' which holds that Conservatives are less open to new information that conflicts with their political identity due to a greater 'need for certainty' (Jost 2017). I do not focus on this interpretation because it requires more complex measurement of psychological traits, and recent research casts considerable doubt on its validity (Guay and Johnston 2021).

[15] Speculatively, this could reflect the fact that many 'other' respondents would be more supportive of a Liberal Party than a Conservative Party – for instance, many voted Liberal Democrat in 2019.

Across two studies designed to test this claim directly, I find robust support for it. Knowing a trailing party has more momentum than the leader makes people, on average, significantly more likely to rate its chances of winning at higher levels. This is true both in stylised, hypothetical election contests (Study One) and when invoking real electoral contexts (Study Two). Moreover, my findings suggest that this effect is larger the more momentum the trailing party has. This strongly suggests that dynamic popularity information has a causal effect on electoral expectations, providing support to this aspect – the only aspect without any previous direct empirical support – of the theoretical model proposed in Chapter 2.

The findings of Study Two provide some nuance to this overall picture. This study was designed to account for the fact that people, in the real world, bring their own information and predispositions to the table when encountering popularity information (see Chapter 2). The results show that when people have the ability to capitalise on existing knowledge of UK politics – i.e., when they are reporting expectations about an electoral context they are familiar with – the effect of dynamic popularity information is suppressed. Overall expectations also seem to be adjusted. Here, they were adjusted downwards. This arguably makes sense given that the party people were reporting expectations for was the Labour Party. Not only did the Labour Party fail to win the 2017 UK general election when securing the same vote share that it is given in the experiment, but it also suffered disappointing results both in the 2019 general election and the more recent local elections just over a month prior to the experiment. In this context, it makes sense for people to doubt its chances of winning even a hypothetical future election. It also seems that, in a (Canadian) context where there is clearly more information out there that people do not have – i.e., when they are reporting expectations about an electoral context they are unfamiliar with – the effect of dynamic popularity is suppressed even further. I argued that this is likely due to respondents' uncertainty about the outcome, and thereby their unwillingness to commit to an election prediction. In order to establish whether these interpretations are valid, future research would need to study the effects in a broader range of contexts, while also directly measuring levels of political knowledge in each of these contexts.

In addition, breaking this down by partisanship reveals the extent to which wishful thinking – the effect of predispositions – is at play. There is evidence of considerable wishful thinking effects in the UK context, in an election that people have reason to care strongly about. Here, people's 2019 vote choice is substantially more influential on their expectations than dynamic popularity information. But wishful thinking has much more of an effect on *overall* expectations for the Labour Party's performance than on how responsive people are to information about its momentum in the polls. In Canada, in an election that people have reason to care less about, there is some evidence of such effects in line with the parties' left-right political ideology, but they are much smaller.

All of this contributes new knowledge to our understandings of electoral expectations. While we know that momentum can help to shape electoral campaigns (Abramowitz 1987, 1989; Bartels 1987), that people use objective cues to help inform their expectations (Blais and Bodet 2006; Irwin 2002), that people learn quite well from the polls (Zerback, Reinemann, and Barnfield 2021), even using them to inform complex expectations about likely future coalitions (Bowler, McElroy, and Müller 2021), and that voters' expectations tend to become less optimistic when polls indicate they should (Krizan and Sweeny 2013), the evidence that dynamic popularity information in the polls has an independent effect on expectations is novel. Some work has studied how what the polls say compares to other information in its effect on expectations (Zerback, Reinemann, and Nienierza 2015), but this chapter provides the first insight into momentum's effect on expectations across situations in which people have different levels of information overall. Finally, though some have found that wishful thinking affects expectations above and beyond static popularity information (Granberg and Brent 1983), this is the first time any evidence has been presented that assesses how wishful thinking behaves in light of dynamic popularity information about momentum over time in the polls.

As noted in Chapter 2 and in the theoretical discussion in this chapter, there are forms of expectations – such as coalition predictions – that rely on much more than just popularity information (Bowler, McElroy, and Müller 2021). Future research could fruitfully build on my arguments by investigating whether these more complex expectations are also responsive to dynamic popularity information. Arguments suggesting that dynamic popularity is more pertinent than static popularity in proportional systems would suggest that this is likely to be the case (van der Meer, Hakhverdian, and Aaldering 2016).

Of the three real-world problems, stated in Chapter 1, to which research on the bandwagon effect speaks (polling regulation, campaign strategy, and electoral forecasting) these findings have the most clear and direct implications for the discipline of forecasting. It has frequently been demonstrated that aggregating citizens' expectations is an effective way to forecast elections (Lewis-Beck and Stegmaier 2011; Mongrain 2021; Murr 2011; 2016). This study suggests that voters might use more of the information available to them than previously thought in forming the expectations that constitute these forecasts. Knowing this could therefore help to enhance the accuracy and diversity of information used in election forecasting. In this sense, the findings above might also be cause for concern. Because polls have fairly large margins of error, it is quite possible for ‘changes’ over time, which serve as dynamic popularity information, simply to come about as a result of random fluctuations in the polls (Bailey and Barnfield 2021). Voters might therefore form opinions, with the apparent allure of scientific objectivity, on the basis of meaningless random noise. This could mean that, in some cases, voters’ use of dynamic popularity in forming their expectations compromises these citizen forecasts. I draw these implications out further when concluding in Chapter 6.

In terms of how expectations fit into the bandwagon effect as discussed back in Chapter 2, the work of this chapter is meaningful because it demonstrates that when using expectations to measure the bandwagon effect (e.g. Meffert et al. 2011), any effect could potentially be a result of both the static and the dynamic popularity information people have. In short, it means that (cross-sectional) expectations data alone often cannot distinguish between static and dynamic bandwagon effects. I flesh out why this is potentially far from a trivial concern next in Chapter 5, where I engage more deeply with the ‘expectations approach’ to measuring the bandwagon effect.

Chapter 5

Great Expectations? How (Not) to Use Electoral Expectations to Study the Bandwagon Effect

Introduction

In this chapter I bring together parts of Chapters 2 and 4 in order to appraise the 'expectations approach' to studying the bandwagon effect. In doing so, I take up RQ_5:

> **RQ_5**: Can electoral expectations be used to study the bandwagon effect?

I argue, as alluded to in Chapter 2, that using electoral expectations to study the bandwagon effect requires longitudinal data. The cross-sectional 'expectations approach' – looking for associations between electoral expectations and vote choice, while 'controlling for' preferences, in cross-sectional data – does not make sense in causal inferential terms given the model at the heart of this thesis. Moreover, the insights of Chapter 4 demonstrate that even if it did make sense, it would give a confused estimate of the bandwagon effect. To better address both of these flaws, I explain and demonstrate how to use expectations data in longitudinal panel analyses, finding small (probably dynamic) bandwagon effects in the 2017 UK general election in the process.

The common approach of studying the bandwagon effect by looking for an association between voters' electoral expectations and their voting behaviour is (or should be) based on the assumption that expectations contain popularity information in some form. How people think a given party is going to perform at a future election is closely linked to how well that party is polling, for example. As the overview of the literature in Chapter 4 made clear, this assumption is very likely to hold. Yet this introduces a simple problem: people are more likely to report higher expectations for the parties or candidates that they prefer – they engage in wishful thinking (Babad 1995, 1997; Meffert et al. 2011; Stiers and Dassonneville 2018). People also tend to vote for the parties they like. So the relationship between expectations and vote choice is confounded and estimates of it will give biased, inflated results. These results will not accurately reflect the effect of the popularity information contained in people's expectations on their vote choice. The bold arrows in Figure 5.1, which again reproduces the causal model from Chapter 2, illustrate a simplified version of this problem. Presented in this way, the solution seems simple: control for voter preferences when modelling the effect of expectations on vote choice in cross-sectional data.

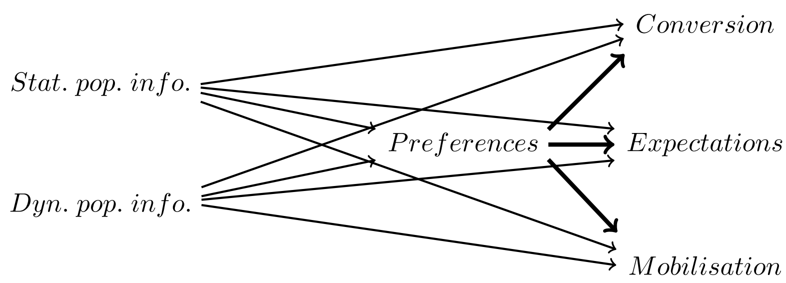


Figure 5.1: Causal model of the bandwagon effect, with the main problem addressed by this Chapter in bold.

This chapter attempts to dispel the notion that the problem is so easily resolved. It does so primarily by drawing on – and extending – the theoretical work done in Chapter 2. Scholars make a leap from studying expectations to studying popularity information by removing the effect of ‘predispositions’ on these expectations, through statistical controls for measures of partisan preferences, in vote intention models. However, the theory of the bandwagon effect suggests that the popularity information will itself have an effect on these preferences, as an indirect route to affecting vote choice. The causal model in Figure 5.1 shows this, and Chapter 2 justified it extensively with reference both to some existing research on the bandwagon effect, and to the mechanisms through which the effect is likely to come about. This problem, which I call *double confounding*, makes statistical adjustment a non-starter. Controlling for preferences will remove any effect of popularity on them, but this should be part of the total bandwagon effect that researchers want to allow their models to measure. I argue that this even applies when party preferences are operationalised as party identification, a supposedly stable trait. So it is itself wishful thinking to expect statistical control to resolve the problem of wishful thinking. The first part of the chapter draws out this key argument and discusses its implications, which potentially go far beyond study of the bandwagon effect. I also compare and contrast this to the 'mutual causation' understanding of the wishful thinking problem.

In addition to this key concern drawn from the work of Chapter 2, the findings of Chapter 4 demonstrate a further problem for the expectations approach. Expectations are not only based on static popularity information, but also dynamic popularity information. This is what Chapter 4 shows. Taking a snapshot, cross-sectional survey means that expectations across the sample could be based on either static or dynamic information, and likely some combination of both, making it impossible to distinguish static and dynamic bandwagon effects. Crucially, this not only creates a problem in working out which type of effect is taking place, but even potentially prevents measurement of *either* type of effect in some cases. As such, larger statistical associations do not necessarily indicate larger static bandwagon conversion effects, as is commonly thought. In fact, it is not at all clear what these associations mean.

In light of these discussions, the second part of the chapter describes and demonstrates a method that instead uses *longitudinal* data on electoral expectations to study the bandwagon effect. This approach is based on recent recommendations for studying causal relationships through 'general cross-lagged models' (GCLM) (Zyphur, Allison, et al. 2019; Zyphur, Voelkle, et al. 2019). I explain the logic behind this method, and how this fits with bandwagon research, noting that drawing out the causal process over time means that the model is more naturally interpreted as capturing *dynamic* bandwagon effects. I apply the approach to three consecutive waves of the British Election Study Internet Panel (BESIP). Here, the distinction between conversion and mobilisation – omitted from the preceding discussion for the sake of simplicity and clarity – is reintroduced to note that it is only feasible to study *conversion* in this case. The analysis finds some evidence of a very small dynamic bandwagon conversion effect in the Labour Party vote in the run up to the 2017 UK general election.

The Discussion and Conclusion section clarifies what the work of this chapter means for research on the bandwagon effect, while noting its limitations. Quite simply, it means that existing studies using the expectations approach cannot and do not measure the bandwagon effect in its totality. If scholars want to continue using data on electoral expectations as a convenient way of studying the bandwagon effect, they should seek out or conduct suitable panel surveys. Longitudinal data is not a silver bullet, but goes a long way to addressing the key concerns set out here.

The Expectations Approach

In Chapter 2, I provided an overview of the logic underpinning the ‘expectations approach’ to studying the bandwagon effect. To summarise, people’s expectations can serve as a proxy for the popularity information they have seen, heard, or otherwise been exposed to. People who have seen polls suggesting a big lead for the Conservative Party should have higher expectations for its performance at an upcoming election. For example, they might answer a survey question about the party’s probability of winning by giving a large probability. Because there is no easy or obvious way to work out what popularity information people have received, we can instead estimate this as something captured by their answers to such expectations questions. We cannot observe the popularity information a voter possesses or has access to, but we can observe its assumed effects on their answers. Then, by looking at the association or correlation between these answers and their vote choice, we can begin to observe whether there might have been a bandwagon effect. This must be done with care though, because people's vote choice and expectations are *bound* to be correlated to some extent even if there is no bandwagon effect, because of wishful thinking: people have higher expectations for parties they like, and they are obviously also more likely to vote for those parties. Account for this, and we will get close to measuring the bandwagon effect.

A first point to note is that this strategy will be insufficient to isolate bandwagon effects in cases where expectations variables are likely to capture much more information than just popularity information. As intimated in Chapter 2 and again in Chapter 4, *coalition* expectations rely on other informational considerations, such as coalition signals (Bowler, McElroy, and Müller 2021). There is no clear precedent or theoretical justification for treating such other cues as the basis of a bandwagon effect. So if researchers want to use the expectations approach, they either have to i) do so with expectations variables whose informational substance is almost all popularity information, or ii) find ways of controlling for other information that could be affecting those expectations, as well as controlling for wishful thinking, or finally iii) claim that it makes theoretical and conceptual sense to talk about the other information bringing about a bandwagon effect. Existing research has tended to employ the first of these three options, as I do below.

Indeed, the 'expectations approach' has long been used to study the bandwagon effect. Purportedly, by including control variables in regression analyses that capture people's partisan preferences, it is possible to isolate the bandwagon effect.

As Skalaban (1988, 138) puts it, 'in order to assess the effect of polls on vote choice, one must be careful to select control variables to partial out projection effects' – 'projection' being another term for wishful thinking.

In an attempt to do this, Abramowitz fits models to data from an exit poll of presidential primary voters in 1988, in which

> The dependent variable is the voter's candidate preference (Choice); the independent variables are the voter's overall evaluations of the major candidates (Candidate Evaluation), the voter's perceptions of the candidates' chances of receiving their party's nomination (Viability), and the voter's perceptions of the candidates' chances of winning the November election (Electability) (Abramowitz 1989, 980–81).

These viability and electability variables are just specifically targeted measures of expectations: viability is expectations targeted at the party's nomination, electability is expectations targeted at the subsequent presidential race. By including candidate evaluation alongside these expectations in the model, this should remove any association between vote choice and expectations that is only down to wishful thinking. The results are slightly more complicated, because these variables are combined into a path model: they suggest that there was no direct statistically significant association between viability and vote choice, but that electability was largely driven by viability, and there was a significant association between electability and vote choice. Abramowitz argues that this is evidence against a bandwagon effect in presidential primaries, and in favour of an 'expected utility' model of primary voting seeing voters as rational actors – they prefer presidential nomination candidates who seem likely to win the presidential election. This finding is, however, broadly consistent with the idea that vote choices are being slightly influenced by popularity information. It is not clear from the conceptual and theoretical discussion in Chapter 2 that using 'expected utility' means such an effect cannot be seen as a bandwagon effect.

Lanoue and Bowler (1998, 366), relatedly, claim that by taking the residuals from a regression analysis of expectations on party identification, they derive 'an indicator of viability that is free from the effects of wishful thinking.' The residuals are the part of expectations that are unexplained by party preferences here. This indicator is then used as a predictor in further analyses. Their *district*-level viability variable was significantly associated with vote choice, whereas their *national*-level measure was not. Similarly to Abramowitz (1989), Lanoue and Bowler (1998) treat this as evidence that voters are behaving in a sophisticated, rational way, to best maximise the possibility of a favourable election outcome. Again though, the observation that district-level popularity has an effect on vote choice is, in itself, consistent with the idea of a bandwagon effect – just one at the district level. Looking more closely and considering whether voters opted for their favourite, second-favourite, or third-favourite party however, the authors explain that

> If voters were supporting their number-two party as a means of climbing aboard the bandwagon, we would also expect to see a significant coefficient for our measure of that party's perceived viability. However, we do not. Instead, we find voters who are willing to abandon their top choice when its chances in the district appear to be low, in favor of their next preferred alternative (Lanoue and Bowler 1998, 374).[1]

[1] In distinguishing between strategic and bandwagon voting, Lanoue and Bowler (1998, 373) apply a distinction which comes very close to the one proposed back in Chapter 2, as also noted

Quite convincingly then, when applying the expectations approach, Lanoue and Bowler (1998) find evidence of *strategic voting* but not of district-level bandwagon effects. They note, though, that the coefficient that would suggest a bandwagon effect if it were significant, is 'strong and in the expected… direction,' suggesting 'some voters' might be might be 'climb[ing] on the… bandwagon,' but that this is a claim they cannot affirm (Lanoue and Bowler 1998, 374–75).

More recently, Meffert and colleagues (2011, 812) have measured an effect of expectations on voting behaviour while controlling for 'dichotomous party preference, party evaluation, strength of party identification, and the evaluation of party leader.' The results 'suggest that the expectation of the ÖVP winning the election had indeed a small but significant additional effect on the vote intention for the ÖVP, above and beyond various preference measures' (Meffert et al. 2011, 813) – a small bandwagon effect.

Variants of this, albeit following this logic less directly, are also proposed and tested by Nadeau, Niemi and Amato (1994, 376–77), who find a bandwagon effect for the Conservative Party in Britain which is 'small in magnitude, [but] may be of no small consequence,' and by Riambau, who finds 'small but non-negligible' bandwagon effects in Austria, Germany and the Netherlands (2016, 4). Quite consistently then, results from the expectations approach to studying the bandwagon effect suggest that popularity information of some form has a small effect on the vote – but in some cases there is no strict evidence of bandwagon effects.

there.

The Bandwagon Effect Is Not an Expectations Effect

It is important to establish that insofar as these studies set themselves up in terms of the bandwagon effect, they are not, or at least should not be, interested in an effect of *expectations* on the vote – although their wording suggests otherwise at times. They are studying the effect of *popularity information* on the vote. There are two ways to see this.

Firstly, we cannot *intervene* on people's expectations, but proper causal reasoning involves imagining counterfactual interventions (Pearl and Mackenzie 2018). Because of the now well empirically established effect of wishful thinking, to imagine manipulating people's expectations directly, we have to accept that this might entail manipulating their predispositions. This is both unfeasible and unethical. You cannot turn a Republican into a Democrat with the flick of a switch. Even if we could do this, or want to imagine we could, this would just reveal what the effect of supporting a party is on voting for it. So talking about an effect of these expectations on people's voting behaviour is not particularly informative. For example, it cannot inform policy debates through a meaningful estimate of what the effect of restricting poll publication would be. To inform such debates in a meaningful and useful way, it is necessary to study the effect of this information itself, as established back in Chapter 1.

To see the relevance of these points more clearly, consider that regarding the bandwagon effect as the effect of expectations leads to nonsensical claims. Imagine a voter A who supports Party X and intends to vote for it. However, the polls start to suggest that its only opponent, Party Y, is becoming more popular. Voter A's expectations of the likely election outcome change as a result. Voter B, also intending to vote for Party X, sees Party Y's leader deliver a campaign speech, which she finds convincing. Voter B starts to feel more sympathetic to Party Y, liking it more than Party X. Her expectations of the likely election outcome change as a result, because of wishful thinking. Both of these people end up voting for Party Y. Yet only the first underwent a bandwagon effect. If we think about the bandwagon effect as expectations affecting voting behaviour, we could capture both of these as bandwagon effects, even though the second voter's initial motivations were nothing to do with how well Party Y was performing in purely electoral, or popular, terms. Both voters' expectations for Party Y increased, but voter A's increased because of a change in *popularity information*, whereas voter B's increased because of a change in *preferences*.

The second way to see that the effect of expectations is not really of interest is by thinking about what researchers are doing when they control for measures of party preferences. For all intents and purposes, expectations are inherently 'motivated,' or 'biased,' or 'wishful.' This is part of their nature. So if you want to learn what their effect is on some outcome, why would you remove this part of them? The very fact that researchers see fit to account for wishful thinking shows that really they are studying the effect of the information that drives expectations: popularity information. Or, at least, it shows that they are not studying some component part of expectations – not expectations themselves.

$$
\begin{tikzcd}
 & Preferences \arrow[dl] \arrow[dr] & \\
Expectations \arrow[rr] & & Vote
\end{tikzcd}
$$

Figure 5.2: Causal model according to which preferences confound the effect of expectations on vote choice.

Double Confounding

Clarifying this distinction matters because confusion about it means that the expectations approach ends up treating the question of wishful thinking like the model presented in Figure 5.2. This model implies that political preferences simply confound the relationship between expectations and vote choice. This is an over-simplification. As explained in Chapter 2, political opinions are a 'marriage of information and predisposition' (Zaller 1992, 6): people tend to resist arguments that are inconsistent with their political predispositions, such that when they form their opinions they end up drawing on a biased sample of considerations that are in line with what they *want* to believe as well as what their information suggests they *should* believe.

Political preferences are a measure of key aspects of these predispositions. They are variables that should predict wishful thinking – for example, a party feeling thermometer – and thereby account for the role that predispositions play in shaping expectations. These predispositions are a constituent part of expectations, along with popularity information. The bandwagon effect is the effect of the information on voting behaviour. Yet this information, which may have an effect on vote choice, is also likely to affect the preferences variables themselves. As explored at length in Chapter 2, preferences to a large extent mediate the bandwagon effect. Figure 5.3 is a simplified version of the model from Chapter 2, with static and dynamic information collapsed together, and the same for conversion and mobilisation (here, $Vote$). This simplified model serves to draw attention to the specific problem in question here.[2]

$$
\begin{tikzcd}
Expectations & Preferences \arrow[l] \arrow[dr] & \\
Popularity\ Information \arrow[u] \arrow[ur] \arrow[rr] & & Vote
\end{tikzcd}
$$

Figure 5.3: Causal model of the relationship between expectations, preferences, popularity, and vote choice.

A statistical model cannot tell the difference between the effect of preferences on expectations and the effect of popularity information on *both* expectations and preferences. Both are just statistical associations that we account for by bringing preferences into the equation. By controlling for preferences, scholars therefore remove the possibility that the bandwagon effect might involve people becoming more positive about those parties they end up becoming more likely to vote for. Any effect on vote choice that passes through a change in preferences will be removed. In the terms used at times in Chapter 2, any 'realignment' effect will be missed, in favour of only observing 'misaligned' voting. This is what happens, by

[2] Note that this model is also simplified because it says that *preferences* affect expectations. Really, these preferences are a proxy measure for the unobserved predisposition within the expectations. However, building this into the causal model explicitly would introduce complications while leading to exactly the same conclusion. See Appendix D.1 for more nuanced, complex versions of the model.

design, when making statistical adjustments for a variable that lies on the 'causal chain' between the 'exposure' and the 'outcome' (McElreath 2020, 185; Pearl and MacKenzie 2018, 113–14). Yet the assumption this implies – that the bandwagon effect is never mediated by preferences – is rarely, if ever, justified theoretically.[3]

The challenge this observation presents arises from the fact that the independent variable, popularity information, is unobserved – as explained in Chapter 2. Instead there is an observed variable – expectations – which should theoretically have two components: information (about popularity) and predisposition. This predisposition is also unobserved, so it cannot be controlled for directly as something that is truly exogenous. If it could, this would remove the wishful thinking problem. Instead, we can account for predispositions by proxy, using measures of preferences. Yet, the informational component of expectations should itself be considered capable of having an effect on these preferences, given the theoretical account of the bandwagon defended in Chapter 2. So on the one hand, researchers need to control for preferences, in order to remove the confounding effect of wishful thinking. On the other hand, they must not control for them, in order to avoid closing off the path from popularity information to vote choice that passes through preferences.

This paradox could be referred to as *double confounding*. 'Confounding' means that something is present which biases or complicates the ability to measure an effect (McElreath 2020, 183). This confounding is 'double' because the information and

[3] Arguably, Lanoue and Bowler's (1998) theoretical framing assumes that the bandwagon effect involves people *abandoning* their preferred political parties at the ballot box, without their preferences changing. If this assumption is correct, then the problem explored here does not necessarily affect their analysis. My claim is that, theoretically, it is unlikely that this assumption holds, and conceptually, it is unclear why the term 'bandwagon effect' should be limited to the situation in which the assumption does hold. Chapter 2 already covers these points at length.

the predisposition interact with the causal model in two different, contradictory ways. This phenomenon could have serious implications beyond this specific research question, in cases where researchers attempt to account for partisan motivated reasoning by controlling for party preferences. I pay more attention to this in the Discussion and Conclusion section.

As an example of double confounding, consider again a voter who is generally predisposed to want Party X in power. In a survey, we ask her 'how likely is it that Party X wins the upcoming election?' This is a measure of her expectations. It is based largely on information about how much support there is for Party X. But also, whenever we ask her this question, her answer is biased by her predispositions. She overstates Party X's chances. We want to know if she is engaging in this kind of wishful thinking, so we also ask her, 'how much do you like or dislike Party X?' This is a measure of her preferences. We know that if she is a wishful thinker, these preferences will be strongly associated with her expectations, because they capture a key way in which her predispositions affect them. That is, if her predispositions form part of her expectations, then this will show up in an association between her preferences and her expectations. It begins to become apparent to this voter that Party Y is gaining ground, and becoming increasingly popular. She might see polls suggesting this, read it in the news, or hear that many of her friends are planning to vote for Party Y. This is all popularity information. The bandwagon effect would predict that this increased popularity might make her more likely to switch to support Party Y. But why would this happen? The theory behind the bandwagon effect suggests that, in order to understand and rationalise why Party Y is becoming more popular on receiving this information, the voter will reflect, search for, or pay greater attention to reasons people have for supporting it. These arguments might convince her that Party Y is actually not all that bad, leading her to jump on the bandwagon. In order to work out whether the bandwagon effect happened to this voter, we need to remove the effect of wishful thinking from her expectations, because these expectations are the only insight we have into the popularity information she is receiving, but are skewed from accurately representing this by her predispositions. However, we also need to recognise that this information we are trying to isolate might have an effect on her answer to the question about her preferences, because the way the bandwagon effect came about is by revising her preferences about Party Y. This effect is part of what we are trying to measure. It is part of the bandwagon effect. But by removing any association between expectations and preferences, we remove the possibility of observing this effect. To the extent that anyone undergoes a bandwagon effect of this mediated form, we will not measure it. If all bandwagon effects are of this form – as some insist (e.g. Blais, Gidengil, and Nevitte 2006) – and the behaviour is widespread enough to be statistically detectable, then the expectations approach will result in a null effect of popularity information when in fact this information is at the root of a significant number of vote choices. It will give rise to a Type 2 error, or false negative.

For any given sample of voters, we simply do not know at what point we are catching them in this causal process, so we cannot work out the direction of these effects between the different forces at work. This is because the data used to answer the question are cross-sectional. Such data take a snapshot of an electorate and cannot assert which variables they measure have affected others. If somehow we knew what affected what, such that we could assert causal ordering between the variables without giving rise to double confounding, then there would be no problem. But theory contradicts this. Causal inferences can only be made with cross-sectional data when a theoretical causal model permits it. The causal model in this thesis does not. Adjusting for preferences in a statistical model of cross-sectional data will always risk distorting estimates of the bandwagon effect, because of double confounding.

It is worth establishing that this remains true even if the preferences variable scholars 'control for' in isolating the bandwagon effect is party identification (e.g. Lanoue and Bowler 1998). As Chapter 2 noted, it is widely argued across political science that party identification serves as a stable trait guiding people's political actions and choices (Achen and Bartels 2015; Campbell et al. 1960; Lewis-Beck et al. 2008). In this sense, arguably party identification is an exogenous force which could not feasibly be affected by popularity information. As such, using it to control for wishful thinking isolates the bandwagon effect while avoiding double confounding. Note though that even if controlling for party identification did resolve double confounding, it would not resolve the problem discussed further below of distinguishing between static and dynamic popularity. This further point alone would be enough to call into question the cross-sectional expectations approach. But if the argument that party identification is a stable force is correct, then it may indeed be true that controlling for party identification does a better job of avoiding double confounding than controlling for other, more labile party preference measures.

However, there are three factors that complicate this. The first is that this account of party identification as a stable, exogenous force might not be quite right (Fiorina 1981; Popkin 1991). It might in many cases represent more of a powerful heuristic providing a 'default decision' (Bowler and Donovan 1998, 29; Lau and Redlawsk

2006, 257–58), rather than an overriding and unwavering force. The psychology of party identification may also depend on the nature of the electoral system (Bowler, Lanoue, and Savoie 1994). Even some of the strongest proponents of party identification as a stable force insist that 'it would be foolish to take the position that party attachments are altogether unwavering' (Green, Palmquist, and Schickler 1998, 896).

Relatedly, as argued in Chapter 2, the very people who are most likely to undergo bandwagon effects are those for whom party attachments are weakest.[4] This is a widespread claim in research on the bandwagon effect (Bartels 1987; Hardmeier 2008; Mehrabian 1998), as well as being entirely consistent with the theoretical arguments of this thesis. A large part of the reason these people are more prone to bandwagon effects is because their party attachments are weaker and therefore more labile. Combined with the fact that, if someone's party identification changes, then we would expect their voting behaviour to be likely to change in accordance (Bonneau and Cann 2015), this means that in the instances where a bandwagon effect happens, it could be mediated by party identification. People who undergo bandwagon effects are likely to have relatively weak party attachments, but the party to which they are most attached could still change as part of the bandwagon effect – it is affected by popularity information and then, in turn, affects their vote. The considerable average stability of party identification across the population just implies that these instances are likely to be few and far between. This is another way of saying that the bandwagon effect is small. This is, in turn, precisely why it is important not to rule out observing it in the instances where it involves a change

[4] In other words, (strength of) party identification is unequally distributed across voters, and this distribution is likely to be correlated with the distribution of bandwagon effects.

in people's answers to questions about which party they identify most with. This might not happen often, but given that the bandwagon effect does not seem to happen often either (Hardmeier 2008), controlling away party identification could easily be the difference between detecting it and not detecting it.

Finally, relying on party identification alone as a way of accounting for wishful thinking is likely to be insufficient. Consider the following example. There is a UK general election taking place. A voter identifies with the Green Party. As usual, though, the only parties with a chance of winning the election are Labour and the Conservatives. She would rather Labour win than the Conservatives. In many cases, if we ask this voter how likely it is that the Labour Party will win the election, her answer will show signs of wishful thinking. Controlling for whether or not she identifies with the Labour Party will do nothing to account for this. Indeed, Figure 4.2 in Chapter 4 showed signs of exactly this: 'other' party identifiers were consistently more optimistic about the Labour Party's chances than Conservative supporters were, in the 2017 election. In order to account for this extra wishful thinking not explained by party identification, we would need to introduce other variables, such as party evaluations, which are more labile and thereby more prone to double confounding (Cowden and McDermott 2000). So whether or not party identification is truly 'stable' does not even resolve the issue – the cross-sectional expectations approach, upon scrutiny, necessarily entails double confounding regardless.

Technically, double confounding means that the expectations approach measures the *direct* effect of popularity information on voting behaviour, rather than the *total* effect. This is, *prima facie*, consistent with the fact that studies using the expectations approach have tended to find such small or insignificant effects. These studies might be finding small or null effects because, by design, they are suppressing the effect. This might not seem particularly problematic, and appropriately conservative – at least we are not over-estimating the effect. But in fact, research should avoid biases in either direction: 'underestimating a substantively significant phenomenon could be just as problematic as overestimating a substantively trivial phenomenon' (Fowler 2019). Consider that an election could take place in which polls – perhaps incorrect polls – bring about non-trivial bandwagon effects, but controlling for preferences in models of the vote yields results indicating that this effect was either tiny or non-existent. This would represent a major failure on the part of political science to inform policy decisions on the regulation of opinion polls – a key reason bandwagon research is useful, as noted in Chapter 1. The research would suggest that the polls played no role in the election outcome, accordingly but falsely suggesting to regulators that there is no cause for concern. The expectations approach could even create a negative estimate of the bandwagon effect, where there should be a positive one. Researchers will never know if they created these effects by controlling for preferences, or if there is just no bandwagon effect – or, indeed, an underdog effect. We simply should not adopt methods that we know to induce 'Type M' (magnitude) and 'Type S' (sign – negative or positive) errors of any form (Gelman and Carlin 2014).

Distinction From 'Mutual Causation' Approach

Superficially, this argument comes close to one made historically in research on expectations and preferences. Scholars have frequently understood wishful thinking and the bandwagon effect as mutual causation between expectations and preferences: they cause each other in a kind of feedback loop. Figure 5.4 is a representation of the model behind any claims of such a relationship. Bartels (1985) presents a causal model that explicitly stipulates a mutually causative relationship between expectations and preferences. The problem, as these approaches see it, is that the two arrows going back and forth between $Expectations$ and $Preferences$ are indistinguishable in statistical analyses.

$$Expectations \; \substack{\longrightarrow \\ \dashleftarrow} \; Preferences$$

Figure 5.4: Causal model of non-recursive relationship between electoral expectations and preferences. Dashed arrow indicates potential reverse causality.

Of course, expectations and preferences cannot be simultaneously mutually causative in this sense. Causality occurs *over time*. Bartels makes this explicit when describing the problem that

> Potential voters may first decide which candidate they prefer, and then project these preferences onto their expectations about the outcome of the nominating contest, using selective perception or wishful thinking to convince themselves that their favorite is doing well (1985, 805).

These models want to capture the problem depicted in Figure 5.5, where they seek to isolate $Expectations_{T1} \rightarrow Preferences_{T2}$ but cannot rule out that they are instead measuring $Preferences_{T2} \rightarrow Expectations_{T3}$, because they are using cross-sectional data in which T is constant. This kind of thinking motivates the use of panel data. For example, Lazarsfeld, Berelson and Gaudet (1948, 105), Campbell (1963) and Granberg and Brent (1983, 483–84) have carried out panel analyses on the interplay between expectations and preferences, drawing mixed conclusions from the same data. Granberg and Brent (1983) also introduced their own panel data, analysing which they found no effect of expectations at Time 1 on preferences at Time 2, when controlling for the stability of preferences. The reverse effect, however, was found to be significant. The simplicity of this method unfortunately makes it susceptible to several biases, however (Hamaker, Kuiper, and Grasman 2015; Kropko and Kubinec 2020; Zyphur, Voelkle, et al. 2019). In short, even when panel data has been used – the crux of my recommendation below – this has been done in a way that cannot reliably measure the bandwagon effect. Longitudinal data, as I explain below, is necessary, but it is not sufficient. The method of analysis is critical too.

$$Exp._{T1} \longrightarrow Pref._{T2} \dashrightarrow Exp._{T3}$$

Figure 5.5: Causal model of the relationship between electoral expectations and party support over time.

This body of research also falls short of studying the bandwagon effect in two key theoretical ways. First, its focus is on the effect of expectations on ‘preferences.’ The implication of this is that there are two variables at play, expectations and preferences, and the latter can be thought of as anything from vote choice to party evaluations. Research then considers changes in which of these two variables appears to bring about changes in the other. This creates a problem. If ‘preferences’ are operationalised as party evaluations, then this setup learns nothing about the bandwagon effect as an effect on vote choice. If ‘preferences’ are operationalised as vote choice, then this setup assumes that wishful thinking is the effect of changing vote choice on expectations. But we should consider expectations always to be biased by how much people like each party, without them needing to change which of these they vote for.

The second, related problem is that expectations should not be thought of as bringing about changes in preferences. As the discussion above suggests, it is popularity information that might bring about changes in preferences, as part of the bandwagon effect. Understanding expectations and preferences as mutually causative is too simplistic, and ignores what political psychology teaches us about the structure of opinions and how wishful thinking is a direct implication of this theory.

Static and Dynamic Popularity in Expectations

When considering these relationships between popularity information, expectations, and the vote, a second problem rears its head: the expectations approach cannot distinguish between static and dynamic popularity information. This much is demonstrated by Chapter 4. People's expectations are informed by popularity information, such as the polls. They have higher expectations for parties leading in the polls. However, they also take into account how much momentum parties have, expressing higher expectations when a party is increasing its popularity over time. As such, when measuring voters' expectations, there is no way to know the extent to which these are based on static and dynamic popularity information.

This, in turn, means that the association between these expectations and vote choice – assuming wishful thinking is accounted for – is typically the effect of some unknowable combination of static and dynamic popularity information. Both might be driving expectations upwards, in tandem, if a party is in the lead and growing, or they may be working against each other if a party is trailing but gaining ground, for example. This might seem trivial – we may not know whether we are studying the static or dynamic bandwagon effect, but at least we know we are studying the bandwagon effect – but it becomes problematic in the latter case, when considering that these two different types of information can work against each other, as well as working together.

For instance, people could be undergoing a dynamic bandwagon conversion effect, switching to vote for a party because it is gaining ground in the polls – say, it has increased its vote share by 6%. But, using their expectations to measure this, we do not capture the effect because that party is not polling well in *static* terms – say, it is polling in third place. The expectations might be based on both of these considerations, cancelling each other out, which then means that the association between them and vote choice does not capture this dynamic effect. The evidence presented in Chapter 4 would suggest that the reverse is also true – a lack of dynamic popularity could mask static bandwagon effects when we try to measure them using expectations data.

However, in some cases, it is possible to approach this concern by taking into account broader contextual factors. Knowledge of the electoral context and the information that vote intention polls portrayed throughout a campaign can give researchers a stronger basis on which to assert what kind of effect they observe. For example, think back to the context given by the experiment in Chapter 4, in which Party A led with 44% of the vote but this vote share was stagnating, while Party B trailed on 40% but was gaining ground in the polls over time. Using the expectations approach to study the bandwagon effect for Party B in such a context, by comparing vote choices and expectations between the treatment and control group, we could be quite certain that any effect we capture would be a dynamic bandwagon effect. This implies that in real-world electoral contexts, research can focus on specific cases where it is known that a given party or candidate was, for example, constantly in the lead but never gained ground, or gained ground but never overtook its opponents, etc. In these cases, we can quite safely assert that expectations reflect either static or dynamic popularity information, because we know whether such information was available or not.[5] In other words, rather than just interpreting effects in the expectations approach as static bandwagon effects, researchers can leverage the very factor that complicates this assumption (the fact that static and dynamic popularity can affect expectations independently) to their advantage in interpreting their models.

This does not go all the way to resolving the problem, however, because the fact that a certain type of popularity information was more abundant in a given case does not rule out the possibility that some people still had access to the other type. For example, some voters might undergo a static bandwagon effect because they hear a biased pundit claim that a party is in the lead, even though really it is in second place and gaining ground, suggesting any bandwagon effect they should undergo would really be dynamic. The cross-sectional approach to using expectations data makes this problem unavoidable

[5] This somewhat overlooks the fact that static popularity does not only refer to rank order, but more broadly to the current standing of each party in terms of its vote share. People could, in the example here, technically undergo a static bandwagon effect for Party B based on its large share of the vote (40%). However as discussed in Chapter 2 the distinction between static and dynamic bandwagon effects can be thought of as a difference of *comparisons* that voters are making: between parties at one time point (static) and across time within one party (dynamic). A dynamic comparison is much more favourable for Party B.

The solution I propose below addresses this question of interpretation because, by looking at expectations data longitudinally, it further encourages a dynamic interpretation. By focusing on a specific case – the UK Labour Party in 2017 – in which the party gained ground in the polls but never moved from second to first place, it makes sense for people to be subject to a dynamic bandwagon effect, but this becomes by far the more natural interpretation in my longitudinal framework. I explain why this is the case below. Combined though, these two points also mean that the method I propose is not a silver bullet. It encourages a dynamic interpretation by design, so is less useful if we are interested in static bandwagon effects. It also works best if applied to contexts where the circumstances suggest a dynamic interpretation makes more sense, which limits its applications. Nonetheless, this approach demonstrates that there are rigorous ways to avoid double confounding and study the bandwagon effect using expectations data in a way that is consistent with the concepts, theory and causal assumptions set out in Chapter 2. I now turn to describe this method and application.

Using Panel Data to Study the Bandwagon Effect

In the remainder of this chapter, I set out a blueprint for how better to measure the bandwagon effect using expectations data, and apply this method empirically. To resolve the issues above, it is necessary to track changes in both expectations and preferences, as well as in vote intention, over time in the same individuals. McAllister and Studlar (1991, 728) were correct in claiming that ‘the only data which could address this implied question of causality are panel data, which interview the same respondents at two or more points in time.'[6] Only in this way can preferences be allowed to bring about changes in expectations while also being responsive to changes in popularity information, with both affecting voting behaviour.

Unfortunately, simply using panel data is not enough on its own to solve the problem. Existing panel analyses on the bandwagon effect fall short of resolving the problem of wishful thinking and double confounding, as just noted above (Campbell 1963; Granberg and Brent 1983; Lazarsfeld, Berelson, and Gaudet 1948). Indeed, most approaches to analysing effects in panel analyses fail when there is some form of 'reverse causality' involved (Leszczensky and Wolbring 2019), including existing research on the bandwagon effect. While I have refuted the reverse causality framing above in theoretical terms, the problem of double confounding has similar empirical implications. The two phenomena are similar enough that methods which fail due to reverse causality will necessarily fail in the presence of double confounding. Any time-series cross-sectional approaches with 'fixed effects' would necessarily fail, because they just carry out repeated versions of the 'statistical control' approach above, at multiple time points (Kropko and Kubinec 2020; Zyphur, Voelkle, et al. 2019). Leszczensky and Wolbring (2019) find that panel models using structural equations typically work most effectively in these situations. Accordingly, Zyphur et al (2019, 652) propose 'general cross-lagged models' (GCLM), which build from typical cross-lagged panel models to 'a more general structural equation model' – an attempt to 'expand the toolkits of researchers who regularly use panel data to make causal inferences' over time.

[6] McAllister and Studlar (1991) were describing the mutual causation problem, not double confounding.

This is the method I suggest and make use of here.

The General Cross-Lagged Panel Model (GCLM)

A typical cross-lagged model is summarised in Figure 5.6. Focusing on expectations ($Exp.$), this graph shows firstly that at each time point, they are allowed to be affected by their past state. This is called the 'auto-regressive' component. It captures the fact that if your expectations were at a certain level in the past, in general this will carry forward. Where this presents a challenge for many approaches to panel analysis, such as random- and fixed-effects models, owing to serial correlation of error terms, in cross-lagged models (and, by extension, the GCLM) it is accounted for directly by modelling the dependency of each observation on the past. In models with several ($\geqslant 3$) waves, this also captures the extent to which past effects on a variable spill over into the future of that variable.

Secondly, Figure 5.6 shows that expectations are affected by the past state of preferences and vote choice. The is called the 'cross-lagged' effect. It is designed to capture the idea that if your preferences or voting intentions change, your future expectations should change. This is what wishful thinking would predict.

Thirdly, because these effects are measured in regression models, each time point of expectations by default also has its own intercept or constant, which captures the general average level of expectations at that point in time. This directly accounts for the extent of cross-sectional dependence, another problem faced in typical time series cross-sectional models. This could be extended to account for more nuanced claims about 'spatial correlation' by allowing the intercept to vary by certain group characteristics, in a multi-level regression model. Essentially, this means that instead of having one intercept that captures average expectations, there are several intercepts that capture average expectations among certain groups – people in the same voting district, for example. This would represent the fact that some people, based on some characteristics, are more similar to each other than others.[7]

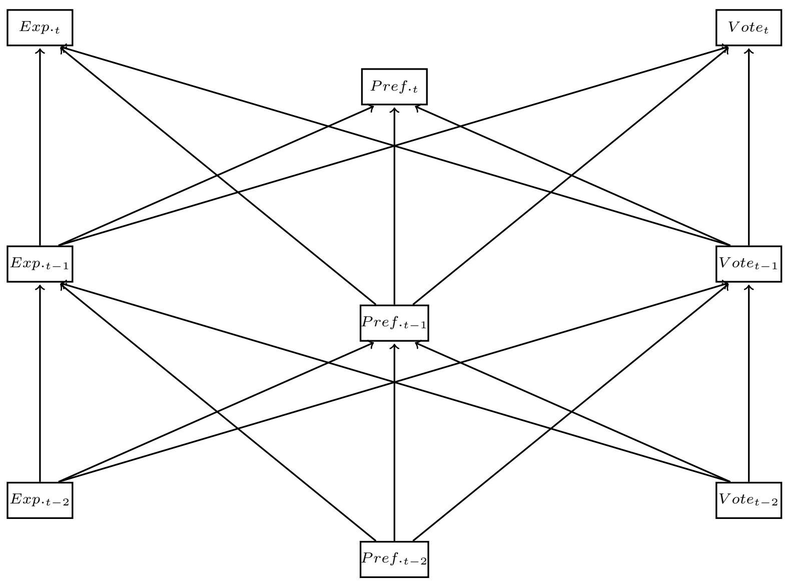


Figure 5.6: Basic cross-lagged model of possible relationships between expectations, preferences and vote choice. This is an extension of the model used by Granberg and Brent (1983) to cover three time periods and introduce vote choice. Preferences are staggered to prevent arrows and shapes overlapping, for clarity.

The 'general cross-lagged model' described by Zyphur et al. (2019) extends and improves this framework in two ways. First, it highlights that the regressions also have a residual, and that this residual can be thought of as an 'impulse.' That is, each person's expectations at a given point in time will differ from what the regression

[7] Currently, the software I use here does not permit this.

model including the autoregressive component, the cross-lagged effect, and the intercept would predict. This is the residual, or left-over variance. The co-variance between these different residuals (e.g., the residuals of Exp_{t-1} and $Pref_{t-1}$) is also modelled, in order to further account for cross-sectional dependence.

These residuals have a specific interpretation. Something changes, at time t, which makes expectations change, but which is not accounted for by past levels of expectations, preferences and vote intention, or more general things that are unique about the current time point. These *impulses* are unpredictable, and therefore as-if-randomly assigned 'shocks' to the system (Stock and Watson 2001, 102; Zyphur, Allison, et al. 2019, 657–58).

Impulses are allowed to affect future states of other variables. For example, an impulse on expectations at $t-1$ can affect expectations at t or indeed preferences and vote choice at t. This introduces the idea that unpredictable changes in a respondent's environment at a given time point can be associated with her future behaviour, and it is these effects which give the model a causal interpretation, and in which we should be interested. The logic of this is crucial to applying the GCLM to bandwagon research, as discussed below.

Zyphur, Allison, et al. (2019) also extend the cross-lagged model by explicitly accounting for stable variation *between people*. As they note, 'between-unit differences' are 'akin to unit-specific trends (e.g., long-run averages) that systematically differentiate units over time' and should not confound the other effects described above because these 'represent perturbations around any such trends… stable factors are constant by definition and thus do not have a clear role in models of causality *over time*' (2019, 659, emphasis in original). This is crucial to addressing a key criticism of cross-lagged models: 'if stability of the constructs is to some extent of a trait-like, time-invariant nature, the inclusion of autoregressive parameters will fail to adequately control for this' (Hamaker, Kuiper, and Grasman 2015, 102). That is, 'vector autoregression' (Stock and Watson 2001) does not often fully account for serial correlation. Between-person differences are therefore captured by a 'unit' effect, which is created as a latent variable. This means that it is estimated according to a calculation of the level of common variance across the same variable, for a given person, over time.

Each respondent's expectations will differ from others' expectations in systematic ways, owing to individual differences. By loading their expectations at each time point onto a single latent variable, these systematic differences are factored in and controlled for. This ensures that serial correlation does not distort effect estimates by modelling this serial correlation as representing some stable underlying traits of each individual. These unit effects' co-variance is also factored into the model, representing that a given person's stability on one variable is likely to be related to their stability on another variable. Importantly, this approach avoids the 'dynamic panel bias' introduced by more typical approaches, such as fixed effects, group mean centering, or 'first differences' (Nickell 1981; Zyphur, Voelkle, et al. 2019).

A full GCLM, with all of these aspects included, is demonstrated graphically in Figure 5.7 for the three-variable case. Applying the GCLM, in general, therefore involves fitting a structural equation model to three or more waves of panel data, in which there are measures of the relevant observed variables. The only thing not explicitly captured in Figure 5.7 is the covariance between 'impulses' and between 'unit effects.' As explained above, these are *not* constrained to be uncorrelated.

In practice, this model is just a combination of regression models and latent variable modelling (Sturgis 2016). Statistical software such as MPlus and the 'lavaan' and 'blavaan' packages in R are easily capable of doing this. Zyphur et al. (2019) provide code in their supplementary material for the general case in MPlus, as well as the means of adapting this for use in R.

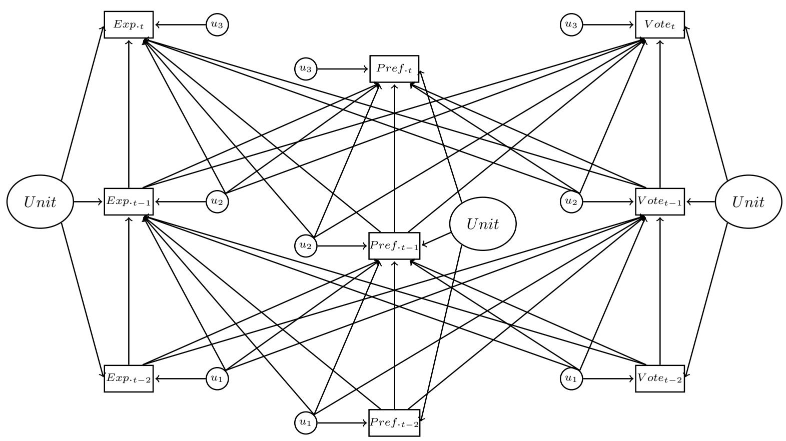


Figure 5.7: Full general cross-lagged model. u represents impulses, Unit represents unit effects. Rectangles represent observed variables, circles/ellipses represent latent variables. Variables are staggered purely to prevent arrows and shapes overlapping.

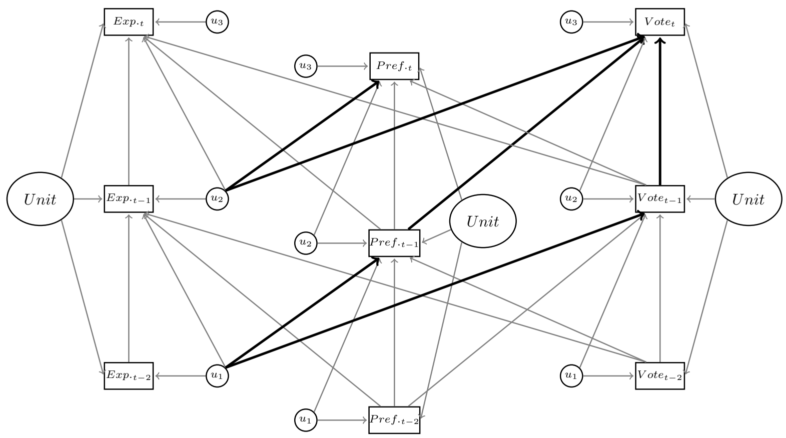


Figure 5.8: Reduced general cross-lagged model. All paths coming out of expectations, and out of vote/preference impulses, removed. The bandwagon effect is captured by the bolded black paths.

Adapting the GCLM to Study the Bandwagon Effect

The bandwagon effect can be studied through a reduced form of this GCLM, depicted in Figure 5.8. To reduce the model, I remove any causal path coming out of each observed 'expectations' variable, except the autoregressive paths. This is consistent with the discussion above and the model in Figure 5.3. Expectations do not affect preferences or vote choice – at least, that is not what the bandwagon effect refers to. It also capitalises on the fact that 'conservative models are typically best for out-of-sample generalizations, wherein conservatism means simpler models that rely on theory, past findings, and contextual information' (Zyphur, Allison, et al. 2019, 668). Just because the GCLM framework allows all of the relationships in Figure 5.7 to be modelled does not mean they should be. Theory tells us when to remove certain relationships, and parsimony is useful not only for making models more 'generalisable' but also for preventing overfitting or saturated models.

Given this preference for parsimony, and to emphasise interpretability, I also remove from the model any direct paths going from the preferences and vote impulses to the other variables in the model. These 'CLMA' paths can easily be included in order to expand the range of short-term and long-term dynamics that the model can detect, 'making each unit's standing on an observed variable a direct function of other variables' past impulses' (Zyphur, Allison, et al. 2019, 662). However, allowing for such dynamics is questionable when the data cover inconsistent time lags. Their inclusion is also not essential in order to estimate the bandwagon effect. Indeed, in the example below, their inclusion would not change its estimate, while confusing its interpretation.

Instead of including paths from expectations, and therefore allowing expectations to have effects on other variables in the model, I exploit the special place the GCLM framework affords to 'impulses' (u). Given that the model explicitly treats each expectations observation, at each time point, as a function of the previous observation, stable individual-level factors, voter preferences, and vote choice, the only thing left for each 'impulse' on these expectations to capture is some change in the respondent's environment that makes her respond in a different way to an expectations question (Stock and Watson 2001, 102). In other words, given the theoretical discussion above, the impulse should capture changes in *information*.[8] The cross-lagged effects from past levels of preferences

[8] To some extent, these residuals will also capture truly random errors and measurement error. Such errors could potentially be accounted for by taking multiple measures of each variable at each time point (for example, asking more than one expectations question) and then loading them onto a single latent variable, to remove item-specific errors. Zyphur, Allison, et al. (2019) address this point in more depth. Here, it would also mean that the slightly vague concept of 'expectations' would not have to be contained simply within one specific question about the probability of *winning*,

and vote choice into expectations carry forward wishful thinking, removing the predisposition component from these expectations. What is left over should be determined by information.

Given that the bandwagon effect is the effect of this information on vote choice, the model captures the bandwagon effect in the bolded arrows in Figure 5.8. The information might directly affect vote choice, but equally this effect might flow through, or be 'mediated' by, preferences. As the information in a voter's environment starts to suggest a party is more popular, this may affect vote choices directly, or it may first make her like this party more, which in turn has an effect on her vote choice. This information should also make her have higher expectations for the party, and the model factors this in.

This model therefore has a very specific interpretation. As Kropko and Kubinec (2020, 20) explain, approaches that account for both within-person and between-person trends, like the approach I adopt here,

> Can only be understood as a generalization of the effect of deviations from the case-means at a particular point in time, or equivalently, as a generalization of the effect of deviations from the time-means for each particular case.

More precisely still, when unit effects are introduced and effects are lagged rather than cross-sectional, it is the effect of 'the individual's temporal deviations from their expected scores' (Hamaker, Kuiper, and Grasman 2015, 104), rather than from group means, that is measured. These deviations are the 'impulses' in a GCLM. So the GCLM measures the average effect of individuals' deviations from

but could also encompass other things like predicted vote share.

their expected scores at a given time point on the future. Where such a deviation occurs in a voter's expectations, this should be due to a change in information about the popularity of the parties. If this leads to future change in vote intention, then this should signal a bandwagon effect – providing the relationship is positive.

This means that the GCLM measures a between-person lagged bandwagon effect. That is, it measures a bandwagon effect in terms of whether some people were more likely to switch to support a party than others, owing to how much they previously deviated from their predicted level of expectations. It asks whether those who deviated more were more likely to switch.[9] Consider a situation in which a party very clearly becomes more popular, and *everyone's* expectations increase because of a universal change in information, and some of those people switch to support the party as a result. The fact that expectations have, on average, increased will be accounted for as a time-varying factor, meaning that these bandwagon effects are unlikely to be picked up. This highlights that the GCLM approach resolves the problem of double confounding by lagging the effects between the variables, but *otherwise works in the same way as the cross-sectional expectations approach*. That is, the standard approach of using expectations is also prone to this same limitation, if indeed it is deemed a limitation. It, too, depends on between-person variation in expectations. It, too, will therefore miss bandwagon effects of the sort just discussed, in which average expectations increase because of a universally recognised change in popularity information. Only the GCLM, though, studies the bandwagon effect in a theoretically reliable way. Nonetheless, it must be recognised that the bandwagon effect it studies is of this between-person type.

[9] It does not necessarily have to measure switching, or conversion. As noted below, this approach could be reapplied to study mobilisation.

Finally, note that this structure also gives the bandwagon effect studied in the GCLM a more natural *dynamic* interpretation. The impulses that capture popularity information become relevant when something changes that was unpredictable based on the past. In this case, this should be that a party or candidate has *become more popular* than it was before. Such information about change over time is what I have called dynamic popularity information. Although, just as was the case with the example in Krizan and Sweeny (2013) discussed in Chapter 4, the impulses could also respond only when a party or candidate has gained so much popularity that it overtakes a rival. In this case, respondents' expectations may only be changing based on this new static popularity information. It would therefore be more demanding to argue convincingly that it is only either a dynamic or static effect at play, and the problems of interpretation discussed earlier may again rear their heads.[10] The case I now turn to discuss avoids this problem by focusing on a second-place party that gained ground over time but never took the lead.

Data

To demonstrate this method, I apply it to data from three waves of the British Election Study Internet Panel (BESIP) (Fieldhouse et al. 2020). These are the campaign wave of the 2017 UK general election, the pre-campaign wave, and the wave immediately prior to that. These three waves are the only three consecutive waves of the BESIP to ask people questions about their electoral expectations for the same upcoming election. This election is also a relevant case because

[10] Panel data collected at a sufficient number of close intervals might be able to distinguish between the two effects by identifying the two waves between which the change of 'rank order' in the polls occurred.

the Labour Party went from polling very badly in the run-up to the campaign, to performing increasingly well during the campaign, and securing an unexpectedly large share of the vote on polling day. This raises the possibility that the bandwagon effect may have been present in the 2017 election, either as an explanation for, or consequence of, this change in fortunes. Because of the theoretical relevance of the Labour Party case, and for simplicity in the analyses, I focus only on the Labour Party. As just noted above, in this particular circumstance in which the party was always polling in second place but increased its vote share, any effects are likely to be *dynamic* bandwagon effects.

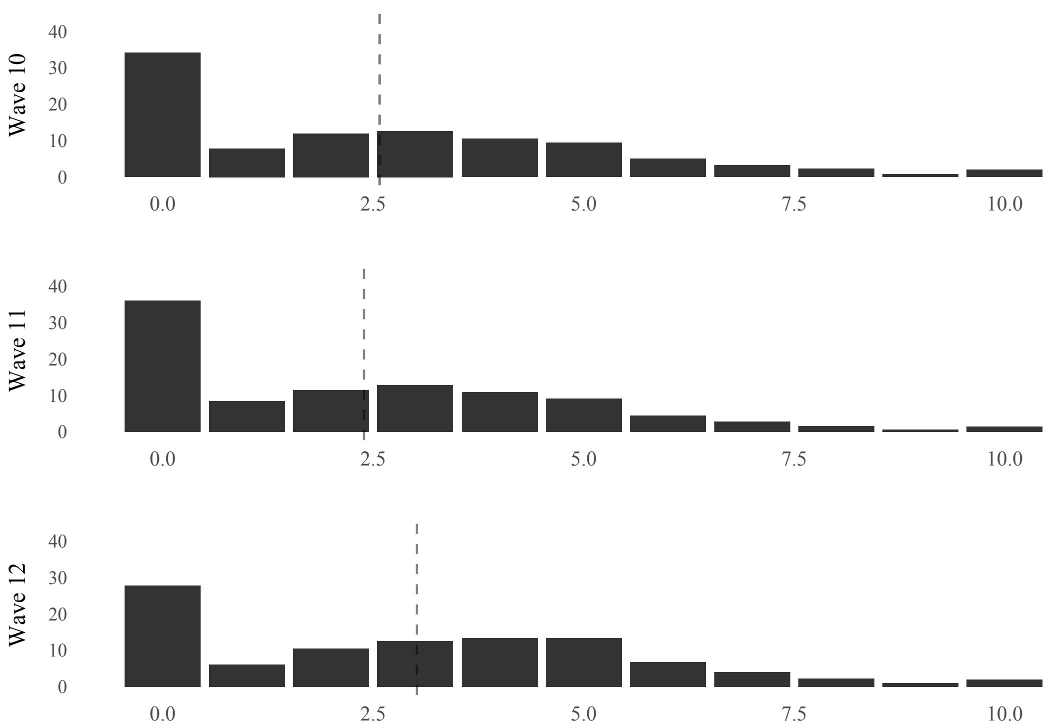


Figure 5.9: Percentage of respondents reporting each level of expectations in wave 10 (late 2016), wave 11 (pre-campaign) and wave 12 (2017 election campaign) of the BESIP. Dashed line represents the mean estimate for a given wave.

The following question measures expectations:

> How likely do you think it is that either of these parties will win more than half of the seats in the General Election so it will be able to form a government on its own? (0-10)

This is the only question gauging expectations about the overall election in these waves of the BESIP, but it is worth noting that it is unlikely to be an optimal measure, particularly in this case. It was always very unlikely that the Labour Party would win enough seats to be able to form a majority government. That changed little with its growth in the polls. This means that people's answers to this question are likely to be constrained to lower values, possibly making it more difficult to observe effects on or of this variable. Indeed, the distribution of this variable in Figure 5.9 shows that responses were skewed towards values suggesting a probability of winning below 50%, across all three waves, and approximately 30% of respondents gave it zero (actually labeled 'very unlikely').

The following question measures party preferences:

> How much do you like or dislike each of the following parties? (0-10)

Other measures of partisan preferences are available, such as party identification, but again such variables are more likely to have constrained variance, suppressing effects. The extent of like or dislike of a party on a scale – especially combined with a measure of vote choice – should capture more natural variance in opinions. As discussed above, it may also do a better job of capturing the full extent of wishful thinking because, for example, people who identify with other parties may still engage in wishful thinking for parties they deem to be more ideologically proximate to their own. The distribution of this variable in Figure 5.10 shows that,

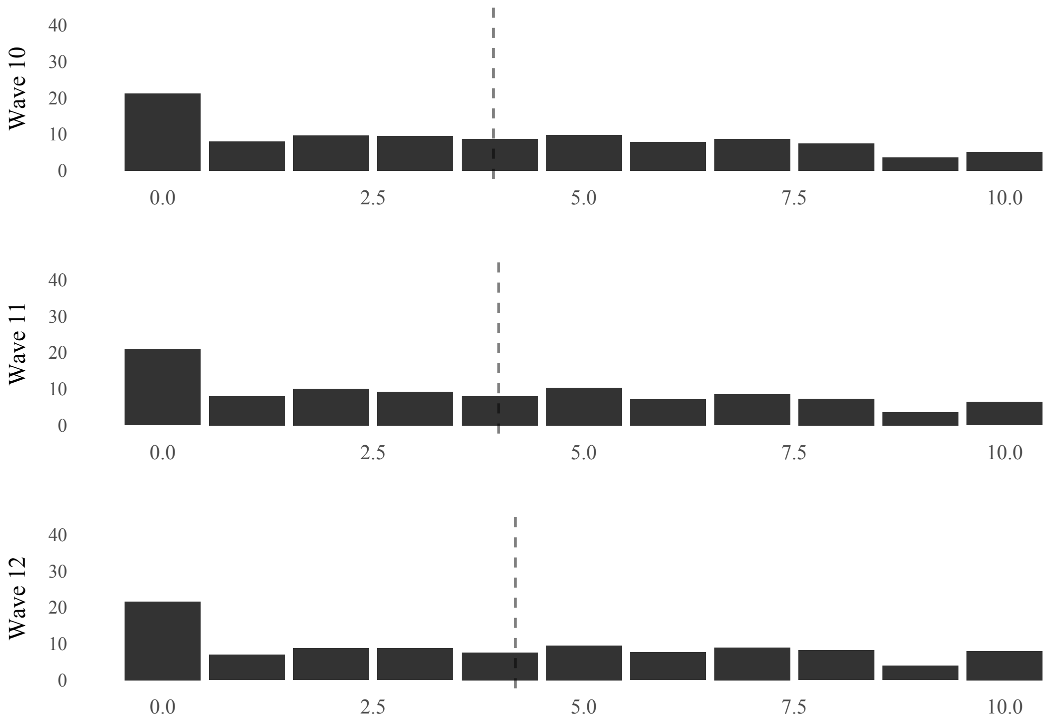


Figure 5.10: Percentage of respondents reporting each level of preferences in wave 10 (late 2016), wave 11 (pre-campaign) and wave 12 (2017 election campaign) of the BESIP. Dashed line represents the mean estimate for a given wave.

although about 20% of people in every wave rate the Labour Party 0/10, its ratings are otherwise fairly uniformly distributed 1-10. Opinions towards the party become slightly more positive over time.

Responses to a standard vote choice question are recoded to a dummy variable representing whether or not the respondent intends to vote for the Labour Party. Its distribution in Table 5.1 shows that in the aggregate there is some suggestion that a bandwagon effect might have happened – because Labour's vote share grows over time – but Chapter 2 discussed at length why this measure is not particularly useful.

Table 5.1: Distribution of Labour vote.

Vote Labour	Percentage
Wave 10	
No	76.09%
Yes	23.91%
Wave 11	
No	75.69%
Yes	24.31%
Wave 12	
No	71.81%
Yes	28.19%

Each of the variables is measured at all three time points. I restrict the sample to those who provided valid responses to all of these questions in every wave. This leaves a large total sample of 11,565 respondents. I fit the GCLM to these data using the R package 'lavaan' (Rosseel 2012). This is based on adapting the example code provided by Zyphur et al. (2019) in their supplementary materials. Note that, in the GCLM, 'unit effects' are designed to capture all person-specific traits that could confound estimates. For this reason, there is no need to control, for example, for gender, ethnicity, etc, as is common in vote intention models. This is similar to the role played by 'person fixed effects' in many multilevel model approaches to panel data. But, as noted above, such time series cross-sectional models introduce bias (Nickell 1981; Zyphur, Voelkle, et al. 2019) – as well as being unable to assess causal order.

Finally, the discussion in this chapter has not mentioned the conversion/mobilisation distinction made in Chapter 2. This is because this distinction pertains to the dependent variable (voting behaviour) and not the independent variable (expectations/popularity information), which is the subject of the chapter. There is nothing inherent in the expectations approach that makes it difficult to distinguish between, or study either one of, conversion and mobilisation. Indeed, the variables in the dataset used here would allow me to explore both, in principle. In practice though, studying bandwagon mobilisation here is not feasible because so few people state that they are not intending to vote.[11] Without observing sufficient abstention, it is not possible to measure mobilisation with enough power. As such, I focus here on conversion effects. The vote variable captures whether or not people intend to vote *for the Labour Party*. When it goes from 0 to 1 over time, people have switched to vote Labour. Combined with the natural dynamic interpretation of these analyses, the estimates should therefore correspond best to dynamic bandwagon conversion effects.

Results

Table 5.2 displays key fit statistics for the reduced GCLM. All of these are within typically desirable ranges, suggesting the model fits the data well (Lai and Green 2016). The Comparative Fit Index (CFI) and Tucker-Lewis Index (TLI), which

[11] This would involve reducing the sample down to those who have been most supportive of Labour, of the parties, throughout all three waves, and measure the relationships between expectations, preferences, and turnout in these respondents. There are vanishingly few respondents who claim not to intend to turn out while also consistently supporting Labour. This speaks to the broader difficulty of measuring turnout in political surveys: people who respond to surveys are typically more engaged in politics than the average voter, and more likely to participate in elections (Karp and Lühiste 2016). It also perhaps provides an explanation for Chapter 2's observation that very few studies have considered bandwagon mobilisation effects.

compare the correlations in the GCLM to a 'null model' in which the variables are unrelated, are both above their usual 'cut-off' value of 0.95. The Root Mean Squared Error of Approximation (RMSEA), which measures the strength of correlations while adjusting for sample size, is below 0.06, deemed the maximum level for good fit (Shi, Lee, and Maydeu-Olivares 2018). These fit statistics therefore suggest that the reduced GCLM fits the data much better than a null hypothesis in which the variables are unrelated, and that this is not simply due to the large sample size. Appendix D.2 reports the full range of fit statistics reported by the 'lavaan' software.

Table 5.3 displays the estimated regression coefficients. As is standard, each regression is constrained so that, for example, $Exp._{t-1} \rightarrow Exp._{t}$ does not have a separate coefficient from $Exp._{t-2} \rightarrow Exp._{t-1}$. Therefore, t and t-1 in Table 5.3 refer to any time point and the previous time point. For each effect, the first variable is the dependent variable or outcome, and the variable to the right of the tilde (~) is the independent variable or predictor. Each of these effects can be mapped back onto an arrow in Figure 5.8. For clarity, Figure 5.11 visualises the main effects of interest (the total bandwagon effect and wishful thinking). The circles mark the average effect estimate, while the horizontal bars coever the range of the 95% confidence interval around these estimates.[12]

Of primary interest here are the effects of popularity information, contained within the 'expectations impulse.' For ease of interpretation, the top panel of Figure 5.11 zooms in on these effects. Note, first of all, that its direct effect on vote

[12] Note that some of these confidence intervals are so narrow that they are not visible. This just means that the standard error of the effect estimate is very small – i.e., the effect is estimated with a lot of statistical precision.

Table 5.2: Fit statistics for reduced bandwagon GCLM.

Fit Measure	Estimate
CFI	0.997
TLI	0.988
RMSEA	0.051

choice appears to be positive and significant, if substantively small. The results suggest that for a one-unit change in how much someone's Labour expectations deviate from their expected value, they become approximately 1% (95% CI: 0.007-0.014) more likely to vote Labour in the future. Recall that this expected value is determined by individual differences from other respondents, the past state of their expectations, and their past preferences/vote intention, leaving only exogenous changes in the popularity of the parties to determine their residual differences in expectations. This result is consistent with a very small dynamic bandwagon conversion effect.

Likewise, popularity information (*expectations impulse*) also has a statistically significant – and larger – effect on preferences. For each one unit increase in how much someone's Labour expectations deviate from their expected value, their future response as to how much they like that party increases by approximately 0.07 (0.050-0.098) points on a 0-10 scale. This is a substantively small effect. Note, however, that previous attempts to study such an effect in panel data have concluded that there was no such effect (e.g. Granberg and Brent 1983). It is unsurprising that the effect is small, given the points made in Chapter 2 that it is difficult for information to override people's preferences. Changes in preferences are, in turn, significantly associated with voting Labour in the future, with a one point increase

Table 5.3: Regression estimates for bandwagon GCLM.

Effect	Estimate	SE	p	CI
Effects on expectations				
Expectations (t) ~ Expectations (t-1)	0.133	0.016	0.000	0.101-0.165
Expectations (t) ~ Preferences (t-1)	0.318	0.011	0.000	0.295-0.34
Expectations (t) ~ Vote Labour (t-1)	0.607	0.082	0.000	0.447-0.767
Expectations (t) ~ Expectations impulse (t-1)	-0.046	0.014	0.001	-0.073–0.019
Effects on preferences				
Preferences (t) ~ Preferences (t-1)	1.064	0.008	0.000	1.049-1.08
Preferences (t) ~ Expectations impulse (t-1)	0.074	0.012	0.000	0.05-0.098
Preferences (t) ~ Preferences Impulse (t-1)	-0.177	0.010	0.000	-0.196–0.158
Preferences (t) ~ Vote Labour (t-1)	0.171	0.056	0.002	0.06-0.282
Effects on vote				
Vote Labour (t) ~ Preferences (t-1)	0.017	0.001	0.000	0.014-0.02
Vote Labour (t) ~ Vote Labour (t-1)	0.919	0.013	0.000	0.895-0.944
Vote Labour (t) ~ Expectations impulse (t-1)	0.011	0.002	0.000	0.007-0.014
Vote Labour (t) ~ Vote Impulse (t-1)	-0.193	0.010	0.000	-0.214–0.173

in preferences bringing about an approximate increase of 2% (0.014-0.020) in this probability. This might seem like an astonishingly small effect, but note that the model controls for past vote intention, which carries forward substantially: those who said they were intending to vote Labour in the past were approximately 92% (0.895-0.944) more likely to say they would in the future. On the whole, these findings suggest that a total bandwagon effect, of the sort proposed in Figure 5.3, is likely to be at work.

The results also suggest that wishful thinking is substantial and significant, and much more so than any bandwagon effect. The bottom panel of Figure 5.11 focuses on these wishful thinking effects. Voters' Labour expectations at a given point in time are strongly affected by how much they previously said they liked Labour. A one point increase in Labour preferences in the past is associated with a 0.32 (0.295-0.340) point increase in Labour expectations in the future. Beyond this, having expressed the intention to vote Labour in the past also increases future Labour expectations by approximately 0.61 points on a 0-10 scale, though this effect is less precise (0.447-0.767). This therefore demonstrates the importance of effectively accounting for wishful thinking in measuring the bandwagon effect – as well as providing further evidence of how important the wishful thinking phenomenon is in its own right.

Discussion and Conclusion

This chapter has appraised the widespread use of data on people's electoral expectations to study the bandwagon effect. Its main argument is that simple assumptions

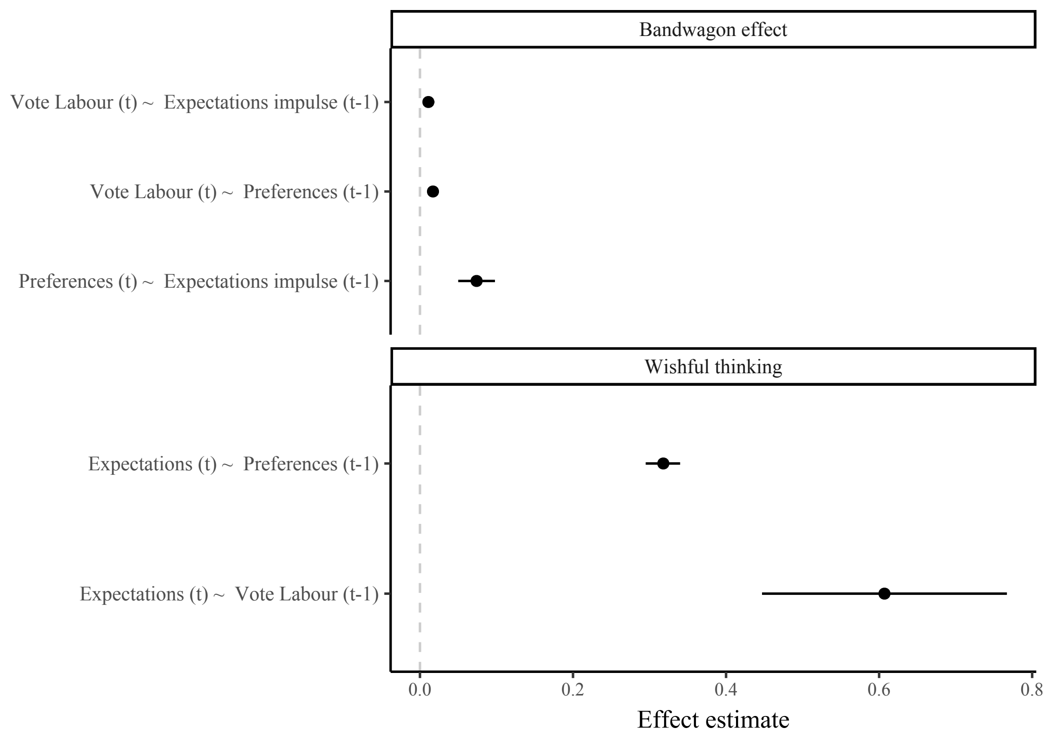


Figure 5.11: SEM regression estimates.

derived from the theoretical discussion in Chapter 2 demonstrate that the common practice of 'controlling' for preferences in cross-sectional models in order to prevent wishful thinking from biasing the bandwagon effect gives rise to 'double confounding.' Data on expectations are a way of measuring voters' underlying understandings of popularity based on their information environments. Scholars get at this popularity information by removing the effect of preferences on the expectations. However, the bandwagon effect should imply that the popularity information will itself have an effect on these preferences, as an indirect, mediated route to affecting vote choice. In this sense the relationship is doubly confounded, ruling out using preferences as a control variable. I also noted that while controlling for party identification as a specific measure of preferences might improve matters, it is unlikely to resolve the problem of double confounding.

This problem of double confounding could potentially arise in any situation involving measuring the effect of an opinion on vote choice. These opinions are used to study causal effects like the bandwagon effect, but the approach of removing 'bias' from the opinions by controlling for party preferences introduces double confounding. This practice is widespread. One very prominent example is research on the economic vote, where those who support the incumbent government are more likely to have more positive economic perceptions than those supporting the challenger party (Anderson, Mendes, and Tverdova 2004; Bailey 2019a; Pickup and Evans 2013). Working out the effect of information about the economy on vote choice – which a similar argument to that presented here would reveal to be the underlying effect of interest – involves removing this predisposition from people's opinions on the state of the economy. Cross-sectional statistical models which therefore control for party preferences in estimating the economic vote in this way would introduce double confounding.

I also argued that the findings of Chapter 4 reveal a further difficulty with using the expectations approach: it cannot distinguish between static and dynamic bandwagon effects, because expectations are driven by information about individual parties' momentum as well as the current popularity of parties relative to each other. This not only precludes the nuance introduced by distinguishing between the two, but also means that bandwagon effects can be difficult to observe, or artificially suppressed, when studying them using expectations data.

Both of these main arguments, then, show that the expectations approach sets itself up to under-estimate the bandwagon effect – or, perhaps, fail to find it when it is there. This is notable in light of the fact that studies adopting this approach have either found very small effects, or no significant bandwagon effect. Sometimes it is argued that approaches to measurement that 'work against' finding an effect are good, because they are conservative, but it is important to resist this claim (Fowler 2019). Consider a situation in which a policy debate hinges on whether or not the bandwagon effect takes place, and how large it is. If the effect happens and is substantial, policy-makers want to know this. The methods we use to measure the effect should not underestimate it.

To address these limitations, I proposed an adapted version of 'general cross-lagged models' (GCLM) (Zyphur, Allison, et al. 2019). Using panel data, this approach models the effect of popularity information on future vote intention, as an 'impulse' – left-over variation in a voter's expectations at a given point in time that is not explained by her past expectations, her past preferences, or her past vote intention. As such, it effectively applies the same between-person logic that is already used by scholars when studying the bandwagon effect using expectations, but while crucially resolving the double confounding problem by lagging these effects over time. These effects therefore have a more natural *dynamic* interpretation, especially when applied to a case study in which the party of interest had momentum over time (dynamic popularity) but never changed its overall standing relative to other parties (static popularity). I demonstrated this approach to such a case, using data from the British Election Study, and found some evidence suggestive of small dynamic bandwagon conversion effects in the Labour Party vote in the lead up to the 2017 UK general election. This could be significant for debates about the regulation of opinion polls, as it suggests that they might be missing part of the picture by focusing only on static effects. The findings also have potential implications for political strategy, in that they appear to demonstrate that parties and candidates can benefit from being seen to have momentum. Forecasting, too, could learn from these findings, as under certain circumstances they may do better at predicting election outcomes when factoring in that parties or candidates with momentum might attract more votes. I return to all of these points in Chapter 6.

Given that I find small effects, arguably the expectations approach has not yielded particularly problematic results when it has found small effects in other studies. This is not sufficient evidence to justify continuing to take the approach. Existing analyses may still have considerably mis-estimated the bandwagon effect *in the elections they studied*. Crucially, their estimates do not capture any indirect component of the bandwagon effect, which could well have been large in those electoral contexts. The analysis presented above found a small indirect bandwagon effect passing through party preferences, which a cross-sectional approach would remove by design. The only way to justify using the cross-sectional expectations approach is to claim that this part of the effect is not of interest. Chapter 2 explained why this does not fit with a theory-driven conceptual understanding of the bandwagon effect.

Nonetheless, while the GCLM approach is more theoretically justified, this comes at a cost. Firstly, appropriate panel data for applying the GCLM to the bandwagon effect are very rare. One reason for this may be that they are simply more costly to collect. Secondly, scholars who are unfamiliar with structural equation modelling may struggle to interpret or reapply the method, whereas the popular ‘statistical control’ approach is relatively straightforward. Finally, the use of panel data potentially lures scholars into thinking that their causal inferences do not still depend fundamentally on a between-person logic, when they do. This dependence on between-person variance also means the GCLM approach cannot resolve the fact that using expectations to study the bandwagon effect potentially misses part of the effect that may occur when *everyone’s* expectations increase in response to a strong exogenous shock.

There are other general limitations to using expectations data to study the bandwagon effect that the GCLM approach does not resolve. One is simply that the estimation of the popularity information contained within expectations is very uncertain and prone to error, especially when only based on responses to one survey question. As such, this could be improved upon by gathering responses to several expectations questions and modelling a latent expectations variable, or by using the recently proposed method of Manski question formats (Leemann, Stoetzer, and Traunmüller 2021). A second general limitation is that expectations data could feasibly give the illusion of bandwagon effects that are partially brought about by strategic voting behaviours – from which they should be distinguished.

In the application here, however, it would typically be perceptions about parties' chances *in the voter's constituency* that would drive strategic decisions, rather than the overall national race (e.g. Eggers and Vivyan 2020). Changes in information about Labour's overall, national-level popularity would not represent a primary concern to a strategic voter (Alvarez, Boehmke, and Nagler 2006). In order to more closely scrutinise this distinction, future research could potentially combine aspects of the approach taken here with aspects of Lanoue and Bowler's (1998) method for distinguishing between bandwagon and strategic voting – especially considering that this is the closest existing scholarship has come to distinguishing the two concepts in line with how I did so back in Chapter 2.

In summary, if bandwagon research is to continue to use data on expectations, then as a minimum this chapter has shown that efforts should be made to do so using a longitudinal approach to causal inference. Though not a silver bullet, this will address the problems created by wishful thinking and double confounding, making for less-biased estimates of the bandwagon effect.

Chapter 6

Conclusion

Popularity matters in elections because elections are popularity contests. The popular parties and candidates are those most people vote for, and getting votes is how you win elections. But this thesis has shown that popularity matters even more than this. Elections and polls measure popularity, but in doing so they also provide this information to voters. This information then has an impact. It can change people's expectations of electoral outcomes, and even how they themselves will vote.

Political science has mostly thought about these phenomena in terms of the 'bandwagon effect.' In this thesis, I have unpacked this effect, examining how it can be understood in a broader political psychological framework, and how it can be studied as a result of these theoretical claims. I have then used this analysis to inform empirical studies of how popularity information affects, or has affected, attitudes and behaviours in both hypothetical and real recent electoral contexts.

This concluding chapter proceeds in three parts. The first part addresses these contributions in turn and restates why they matter for political science, summarising the work of each of the preceding chapters in the process, with reference to the research questions set out in Chapter 1. The second part returns to the three real-world areas to which Chapter 1 claimed this thesis speaks – polling regulation, campaign strategy and electoral forecasting – and draws out the direct implications of my findings for each of them. The third part shifts gears, looking forward and suggesting avenues for future research.

What This Thesis Has Done and Why It Matters

RQ$_1$: What Is the Bandwagon Effect?

The first contribution this thesis makes is to provide a full conceptual account of the bandwagon effect. Although it has been studied a lot, little has been said about what exactly the bandwagon effect is. This has led to confusion. Taking the view that a better approach is to start from a clear understanding of what the effect is meant to be, and only *then* study it, Chapter 2 defined and provided a typology of the bandwagon effect. Specifically, I defined the bandwagon effect as

> A positive individual-level change in vote choice or turnout decision towards a more or increasingly popular candidate or party, motivated initially by this popularity.

This definition is grounded in conceptual analysis, building on a review of over 60 peer-reviewed journal articles. It captures what the majority of these articles study, while drawing strict lines between the bandwagon effect and other phenomena with which it could be confused.

There is flexibility in this definition. Chapter 2 also formulates a typology of the bandwagon effect around this flexibility. The bandwagon effect can be about *conversion* – switching to support a more or increasingly popular party/candidate – or it can be about *mobilisation* – turning out to vote for a candidate/party one already prefers. The bandwagon effect can be *static* – based on information about how popular the candidates/parties are relative to each other, right now – or *dynamic* – based on information about how a candidate/party's popularity is changing over time.

This therefore sets clear terms and boundaries for scholarly debate. Until now, studies operationalising the bandwagon effect in different ways have been taken as a basis for claims about the existence of the effect, and some that find would-be bandwagon effects are treated as evidence of something else entirely (e.g. Atkin 1969). A clear definition and typology resolves this by clarifying what counts as a bandwagon effect and how to distinguish between different types. Had such a definition and typology been proposed and accepted half a century ago, research on the bandwagon effect could have made much greater advances. Indeed, research is now beginning to apply the definition and typology set out here (Araújo and Gatto 2021; Larsen and Fazekas 2021; Volckart 2021).

RQ_2: Why Should the Bandwagon Effect Happen, Theoretically?

The second thing this thesis does is situate the bandwagon effect theoretically. Understandings of why the bandwagon effect happens are quite diffuse, with some relying on notions of irrationality and mass herd-like behaviour (Callander 2007), and others suggesting more rational information processing (Roy et al. 2015). From the outset, I have argued that the former of these categories has severe limitations for studying the bandwagon effect in modern politics. Such normative accounts of the bandwagon effect do not make sense in historical perspective, nor are they borne out in any reliable evidence; they also do not mesh well with contemporary understandings of voter psychology (Bowler and Donovan 1998; Lodge and Taber 2013).

Chapter 2 argues instead that if approached from an 'informational' perspective, the bandwagon effect makes perfect sense in terms of such contemporary political psychology. To do this, it integrates Mutz's (1998) important work on 'impersonal influence' with Zaller's (1992) path-breaking model of opinion formation and Lau and Redlawsk's (2006) comprehensive account of how voters make decisions. This suggests that bandwagon effects might come about as a result of voters using popularity information as a 'heuristic' telling them which candidates are most viable, and therefore worth considering, or through a 'cognitive response' in which voters respond to popularity information by considering the arguments other people might have for supporting a candidate. Both cases can lead to the voters supporting these candidates where they otherwise wouldn't have, and they do so because those candidates are popular. By scrutinising how these processes work, I argue that people's pre-existing political preferences play a central role in moderating and mediating the bandwagon effect. This means that the bandwagon effect is unlikely to be large – insofar as it is measured in real political contexts where voters care to some extent about the outcome. This, in turn, makes concerns about how to measure the bandwagon effect all the more important, because it is difficult to measure small effects precisely.

The broader reason all of these arguments matter works in two directions. First, they show that work on political psychology can enrich our understanding of the bandwagon effect. They pave the way for future work to use insights from other areas of voter behaviour research to better understand the bandwagon effect and why it emerges in some contexts but not others. Second, working in the other direction, these points matter because they show that new knowledge about the bandwagon effect can make far-reaching contributions to political psychological theory.

RQ_3: How Can These Concepts and Theory Be Put Into Practice Empirically?

These conceptual and theoretical points are all well and good, but would be of little use if they could not be implemented in empirical research. Chapter 3 shows how to do precisely that. It takes a previously unexplored potential cause of bandwagon effects: state polls. It asks, theoretically, would it make sense for voters in US presidential elections to be influenced by the results of polls in their state? For example, would we expect the decision of a voter in Georgia potentially to be affected by how well the presidential candidates are polling among voters from

Georgia? With reference to the work of Chapter 2, I argue that the answer is yes. Yet Chapter 2 also claims that there are multiple potential dimensions of such an effect, because the bandwagon effect can be static or dynamic and can be felt through conversion or mobilisation. So the study should ideally allow all of these effects to emerge. As such, Chapter 3 takes the conceptual and theoretical insights of Chapter 2 and uses them to inform empirical research design. The key result of the experiment is that people who saw polls from their state suggesting Joe Biden was in the lead were subsequently *less* likely to report intending to vote for him in the 2020 US presidential election. This goes directly against the predictions of the (static) bandwagon (conversion) effect.

This contribution therefore has two key elements. The first is methodological. Chapter 3 demonstrates how to operationalise a full theoretical and conceptual understanding of the bandwagon effect in order to study the effects of polls in a real, ongoing election. Crucially, by using real polling data, it shows that this can be done ethically and realistically. These issues are discussed at more length in the essay in Appendix A. Along with other similar recent, deception-free, realistic experimental research (van der Meer, Hakhverdian, and Aaldering 2016), this therefore lays the groundwork for future experiments on the bandwagon effect.

The second element is empirical. I show that it is possible voters in the 2020 election, in swing states, were made less likely to vote for Biden by polls suggesting he was leading there. This has important implications for how we see what happened at this election, particularly in terms of the fact that Biden did not dominate the vote quite as much as some forecasts suggested he would. If Biden lost support when he was seen to be in the lead in state polls, then it could be that rather than (or as well as) simply over-estimating Biden's share, they may have partly *caused* it to decrease. Polls have the potential to change the very outcome they measure, but not necessarily in the way we might expect. Rather than just becoming self-fulfilling, they might become 'self-defeating' (Henshel 1982).

RQ$_4$: Does Dynamic Popularity Information Affect Electoral Expectations?

Next, the thesis moves beyond simple measurement of the direct relationship between polls and voting behaviour. Specifically, I note that 'electoral expectations' – how well people expect parties/candidates to perform at an election – seem to represent a part of the bigger picture of the bandwagon effect. Not only do they appear to be the other outcome, alongside vote choice, that many people think popularity information will affect (Irwin and Van Holsteyn 2000), but they are sometimes used to study the bandwagon effect itself (Meffert and Gschwend 2011). The justifications for both of these approaches are typically under-developed.

The first contribution of the thesis to these questions is to demonstrate in Chapter 4 that, just as the bandwagon effect can be static or dynamic, so can the effect of popularity information on electoral expectations. Where existing research has mostly demonstrated that people tend to think a candidate/party in the lead in the polls is most likely to win (e.g. Irwin and Van Holsteyn 2002), I show that these expectations also depend on whether a party/candidate has 'momentum' – whether it is growing in the polls. This finding is robust across multiple experiments, though the relationship is weaker when these experiments refer to real-world parties, in line with how much people are able to incorporate other information and predispositions. This is the first time it has been shown that parties' growth in the polls has an independent effect on people's expectations.

This is a useful psychological contribution, showing that voters are using information about momentum to understand elections, just as they use it to understand other phenomena, such as the development of behavioural norms, and the results of sports matches. It also matters because pollsters regularly report not only the current standings of the parties, but also how that compares to a previous poll. This is likely consequential for how people predict what is going to happen at an upcoming election. This observation may even be concerning when considering that such changes between polls are purely the result of random noise rather than meaningful shifts (Bailey and Barnfield 2021). Another way of forecasting elections is to predict the outcome based on aggregating individual voters' expectations (e.g. Murr 2016), so my findings are also of broader consequence for the discipline of election forecasting, because they help to understand the basis of such expectations. I return to these practical implications below.

RQ$_5$: Can Electoral Expectations Be Used to Study the Bandwagon Effect?

Another reason it is important to understand electoral expectations is because they are used as a way to study the bandwagon effect. The final contribution of the thesis is to appraise, improve and reapply this approach. Chapter 5 uses the model developed in Chapter 2 to demonstrate that the common practice of 'controlling for' measures of party preferences in regression models – to estimate the bandwagon effect via the cross-sectional relationship between expectations and vote choice – is flawed. This is because it gives rise to what I call 'double confounding.' A secondary problem comes from the findings of Chapter 4: expectations not only contain information about 'static popularity' but also about 'dynamic popularity.' This means that they are unreliable in measuring how either of these types of information affect people's behaviour. The only clear solution to these problems is to use longitudinal data in which the variables are measured at multiple time points within the same individuals, and observe how they change through appropriate panel models (e.g. Zyphur, Allison, et al. 2019).

I apply this improved approach to the case of the 2017 UK general election, in order to study whether there was a bandwagon effect for the Labour Party. The findings show that people whose information suggested that the Labour Party were doing better over time became more likely (than those whose information suggested this less) to say they would vote for the party as the election neared. In other words, it looks like the Labour Party might have benefited from a relatively small dynamic bandwagon conversion effect. This is an important finding as it helps to explain the events of this election, which saw the Labour Party go from performing very badly in the polls to winning over 40% of the vote.

The more significant contribution, though, is the 'double confounding' argument, because it speaks to a whole class of questions in political science. In a nutshell, this argument is about a paradox: the theory behind questions where an opinion or perception is taken to have an influence on vote choice demands that researchers make adjustments for confounding by party preferences; it also demands that they do not. Because of how well-established it is that party preferences influence attitudes and behaviours in politics (Bolsen, Druckman, and Cook 2014; Leeper and Slothuus 2014), and how common it is to ask about effects of perceptions or opinions on vote choice, this paradox calls into question much political science research, beyond the bandwagon effect. Attention to this argument will make for better causal inference in questions of this type.

What This Thesis Says About Polls, Strategy and Forecasting

The theoretical and methodological insights summarised above are of indirect importance in the real world in the sense that they pave the way for rigorous study of the bandwagon effect, which in turn informs debates about the regulation of polling, campaign strategy and electoral forecasting. But the empirical findings also have more direct implications for these three areas. I summarise these implications here.

Implications for the Regulation of Opinion Polling

As covered in Chapter 1, debates about the regulation of opinion polls draw on research on the bandwagon effect. In many cases, this debate hinges on whether the bandwagon effect is ‘informational’ or ‘normative’ (Marsh 1985, 72) – those who seek greater regulation are more likely to claim it is normative, while those who defend polls claim their influence is informational. But regardless of this point of contention, it is typically thought that if vote intention polls are capable of producing significant bandwagon effects, it might be wise to restrict their publication (Donsbach and Hartung 2008). A key part of this is that polls can be wrong. Their influence may be informational, but they may *misinform* voters. If significant numbers of voters are swayed by incorrect information, they may not end up making good electoral choices.

My main direct contribution to this debate is that in a recent election, information that the polls contain likely had an effect on the vote. This effect was not beneficial to the runaway leader – the Conservative Party – but its main rival – the Labour Party. The momentum that the polls and other popularity information conveyed made people exposed to this information more likely to vote Labour. This effect was quite small. The implications of this are threefold. First, that policy debates need to recognise the potential for the polls to generate such *dynamic* bandwagon effects, rather than just *static* ones. Second, that they should recognise that these bandwagon effects are unlikely to cause major electoral shifts in and of themselves.

Third and most boldly though, partly because of the first two points, I would argue that there is a sense in which these effects are beneficial, or at least not particularly problematic – on the condition that the polls are not too wide of the mark. Bandwagon effects have the potential to perpetuate momentum for political campaigns that appeal to broad swathes of the electorate. The straightforward examples of this are campaigns that have gone against the grain of mainstream political discourse and are said to have engaged young voters – such as those of Jeremy Corbyn and Bernie Sanders – but this is potentially true of any political movement, of any stripes. It is not necessarily the case that bandwagon effects perniciously give certain electoral competitors an unfair advantage. Maybe they just give those who are gaining traction a small 'leg-up,' helping them to reach the audiences for whom they are the best ideological fit, allowing electoral democracy to function as it supposedly should.[1]

[1] This is somewhat similar to Mutz's (1998) defense of 'impersonal influence,' which empha-

By the same token though, it is perhaps important to note in these debates that polling has the potential to *jeopardise* the chances of would-be successful candidates. This is what Chapter 3 demonstrated: US citizens were less likely to vote for Joe Biden, the runaway favourite in the 2020 election, when they saw that he was leading in a poll in their state. What's more, it is unlikely that this small effect completely accounts for the inaccuracy of state polls in this election. This means that any effect state polls had was potentially an effect of *inaccurate* information. This also chimes with some interpretations of the famous 'Dewey Defeats Truman' case, as explored in Chapter 1 (Goeree and Großer 2007). Policy debates need to recognise that polls might not only benefit those who are doing well, but also harm their chances.

I would argue that this is a more significant consideration for such debates than findings of small bandwagon effects. It is reasonable to assume that the main reason a candidate would be in the lead in election polls is because that candidate, by and large, appeals to a majority (or plurality) of the electorate. Assume that this is because she is the closest match to most voters' political ideology, as a rational choice approach would hope for (Downs 1957). On this view, this candidate *should* win. If polls contribute to preventing this from happening, they are pushing the electorate away from sensible or 'correct' political choices (Lau and Redlawsk 2006). Despite the importance of this point, it goes largely overlooked in policy debates, because of the prevailing logic of the bandwagon effect.

On balance, then, it is difficult to take one policy recommendation from the work of this thesis. I do not see an overwhelming argument in favour of further regulation

sises its role as a mechanism of political accountability, by exposing people to (and allowing them to scrutinise) more viewpoints and political programmes than they would otherwise encounter.

of opinion polls, nor of the status quo, nor of less regulation. My findings are not consistent enough across contexts to point clearly in any direction. My contribution is more to suggest a reorientation of the debate. If bandwagon effects are small, benefit candidates that voters like, and are not a sign of irrational behaviour, are they really a major democratic concern around which to centre debates about polls? Might it not also be worth considering that (potentially incorrect) polls could sometimes play a part in making it more *difficult* for the popular candidates to win?

Implications for Campaign Strategy

The implications for political strategy are similarly nuanced. The overall logic of the bandwagon effect certainly suggests that it is good to look like you are in the lead. But this thesis presents no concrete findings in support of this. Rather, Chapter 3's finding that Biden suffered when seen to be in the lead suggests that candidates might be wise to exercise restraint in cultivating the image of certain victory. This is not news to campaign strategists, but it is likely a wise move to remind voters that elections are not a foregone conclusion just because of what the polls say: they need to vote the right way to make it happen.

Portraying oneself as a candidate with 'momentum' might, however, be beneficial. Chapter 5 demonstrates that Labour's dynamic popularity in 2017 likely brought more voters on board. Chapter 4 also shows that people see second-placed parties with momentum as having a better chance at an election than they would without momentum, but crucially not a better chance than the leading party. It seems likely, then, that dynamic popularity strikes a useful balance for candidates. It puts them on the radar and increases their viability without necessarily making them look anything like inevitable winners. It also stimulates more people to think about why the candidate is gaining ground, and those arguments might convince them to jump on the bandwagon. The slant in media coverage also plays a part in this, increasing the salience of these arguments, driven by and in turn perpetuating the momentum. The perception of dynamic popularity can create a bandwagon effect, but it is not immediately clear why it would create much of any countervailing effect. At the very least, I have not found any convincing evidence of such backfire.

In summary, then, if this thesis has something to say to campaign strategists, it comes in two parts. If they are working for candidates that are trailing in the polls, they should cultivate the image of momentum where possible. If they are working for candidates that are in the lead, they should be careful not to make victory look like a foregone conclusion. It is worth emphasising, though, that more evidence is still needed on both of these points. They are also probably of little to no use on their own. Popularity will only get a candidate so far without some political substance. Indeed, to be popular, you need to become popular in the first place. This part of the process is certainly likely to require some sort of convincing political platform, and that platform is likely to be more consequential than any effect of the popularity it brings about. Furthermore, note that in both the UK and US cases studied here, the leading candidate (Biden) or party (Conservative) still won the election. The effects that worked against these candidates were not large enough to prevent them from winning. Just because the image of likely victory might lose a candidate some support, or the image of momentum might benefit a rival, does not change the fact that, on balance, electoral candidates should want to be in the lead. As I discuss below, there are also likely to be other advantages to this that go beyond the bandwagon effect.

Implications for Electoral Forecasting

Finally, both static and dynamic popularity are also likely to have an impact on electoral forecasting. Consider, first, the case of Chapter 3 in which I argue that state polls might have been so over-optimistic regarding Biden's chances partly because their publication actually harmed his chances slightly. Given that it is widely deemed that part of the reason overall electoral forecasts suggested Biden would dominate the election more than he did was because of the inaccuracy of state polls, this suggests that the polls themselves may have had some impact on how over-confident people were that Biden would win. Of course, he did win, so perhaps this is not technically over-confidence. Many would have predicted, though, that his *margin* of victory would be much larger. The polls did not simply over-estimate his support in terms of their measurement of it; they actually probably decreased it going forwards, turning their (over-)estimation into a (larger) over-estimation.

The upshot of this for electoral forecasting is, first, the frequently repeated point that forecasters need to make sure they do a good job of reflecting uncertainty (Gelman et al. 2020). Second, it might be worth considering the possibility that runaway leads suggested by forecasts that rely heavily on polling data will have a tendency not to be reproduced quite so clearly in the eventual election outcome. This could lead forecasters to adjust their models towards expecting closer-run races than raw polling suggests. It would, of course, be hasty to implement this at this stage. The point is tentative, and based on very limited evidence, but may at least be worth studying in more depth. It is worth noting, though, that it tallies with the earliest work on the bandwagon effect in the aggregate: Gallup and Rae (1940, 245) claimed that 'support for the leading candidate is usually flat, and in fact, generally goes down toward the end of a campaign.'

Forecasting might also benefit from recognising the importance of dynamic popularity. When a candidate has momentum in the polls, say, in the weeks preceding an election, forecasts conducted at that point considering only how the candidates are polling *right now* might be less accurate than those that account for the fact that this candidate is gaining ground, and allow for the possibility that this momentum might carry forward to a better performance on polling day. Again, this can be seen in Chapter 5 where I provide evidence of dynamic bandwagon conversion effects for the Labour Party in 2017.

To flip this on its head, though, *citizen* forecasts might also do well to adjust for the fact that individuals reporting their electoral expectations might over-adjust for dynamic popularity. Chapter 4 demonstrated that voters' electoral expectations are significantly influenced not only by the parties' rankings relative to each other, but also by how a given party is growing over time. As such, citizen forecasts – which produce election predictions by asking voters which party is going to win and aggregating these following a 'wisdom of the crowds' logic (Mongrain 2021; Murr, Stegmaier, and Lewis-Beck 2019a; 2019b) – are partly based on information about momentum. They naturally incorporate dynamic popularity, by virtue of the fact that voters factor it into their expectations. Yet, often information about changes between vote shares in different polls over time, which serves as dynamic popularity information, is actually just the result of random statistical noise (Bailey and Barnfield 2021). We also know that voters do learn from the polls (Zerback, Reinemann, and Barnfield 2021). Chapter 4 showed that changes in the polls that could easily be produced by such randomness are sufficient to significantly increase voters' expectations for a given party. This may, then, compromise the accuracy of citizen forecasts, because they aggregate predictions that are partly based on meaningless noise. Dynamic popularity seems to be both a boon and a curse for citizen forecasting.

As a result, citizen forecasters should remain aware of the broader information environment into which they intervene. For example, if a citizen forecast is conducted at a time in which a presidential candidate B has had a slight – but potentially meaningless – uptick in her poll performance, and the citizen forecast suggests this candidate is going to win, despite being ten points behind candidate A in the polls overall, this may be because the people doing the predicting are adjusting for the dynamic popularity. This could be a good thing, because perhaps the dynamic popularity is real and is going to create a bandwagon effect that leads candidate B to her eventual victory. This could feasibly be assessed, though, by waiting a little while to see how the situation in the polls develops. If it turns out that the gap between the candidates stays roughly the same, or even increases back to its previous size, it could be that over-adjustment for dynamic popularity compromised the validity of the citizen forecast. If candidate B continues to gain ground on A, the citizen forecast may have spotted something that the polls could not have told us alone.

But the point of forecasting is to forecast, not wait around just in case. So this is just a demonstration, suggesting that forecasters should openly recognise the potential for these discrepancies, and the fact that they can come from the psychology of the formation of expectations. Again, this should encourage a reorientation of debates. Rather than asking whether citizen forecasting or traditional forecasting based on polls is better – a comparison proponents of the former like to make in order to extol its virtues (e.g. Murr, Stegmaier, and Lewis-Beck 2019a)– forecasters can instead ask: how can a combination of the polling environment and citizens' predictions be used to improve election forecasting?

What Remains for Future Research to Do

While the above shows that this thesis has made several significant contributions – both to scholarly thinking on the bandwagon effect and related phenomena, and to immediate real-world concerns in the areas of policy, campaign strategy and electoral forecasting – it leaves open several potentially fruitful avenues for future research. I address some of the most pressing, and potentially enlightening, of these here.

Test the Theoretical Mechanisms

Each of the preceding empirical chapters put a theoretical relationship to the test. However, they did not rigorously test the mechanisms that underlie these relationships: the viability heuristic and the cognitive response. The viability route to bandwagon effects is, however, the subject of extensive existing research in the context of US presidential nomination campaigns. Research has consistently supported the operation of this mechanism, in both experimental and observational evidence (Abramowitz 1989; Bartels 1987; Lanoue and Bowler 1998; Lau and Redlawsk 2006; Utych and Kam 2014). Recent research has also found this mechanism in operation experimentally in Brazil (Cunow et al. 2021). Future research could study whether this process takes place in other 'difficult' election contexts, such as multiparty contexts (Roy et al. 2015) or new democracies.

Less well-established is the case for the cognitive response mechanism. The cognitive response, as highlighted in Chapter 2, has the benefit of being a direct implication of theories of political opinion formation (Lodge and Taber 2013; Marcus, Neuman, and MacKuen 2000) and dual-process theories of social psychology (Chaiken 1980; Petty and Cacioppo 1986). It has also been tested qualitatively by Mutz (1997, 1998) in the case of 'impersonal influence.' This nonetheless leaves substantial room for future research to put the cognitive response mechanism to the test using quantitative approaches to mediation analysis (Acharya, Blackwell, and Sen 2018a; Bullock and Ha 2011; Bullock, Green, and Ha 2010; VanderWeele 2015). This would go a long way to validating the arguments of Chapter 2 and to understanding when and where bandwagon effects are more or less likely to occur.

Study the Bandwagon Effect and Expectations Longitudinally

Chapter 5 explains why it is difficult – if not impossible – to study the bandwagon effect cross-sectionally, outside of experimental data. As such, there, I study it longitudinally in the case of the Labour Party at the 2017 UK general election. But this is just one party, in one election, in one country. Moreover, it is a case in which *prima facie* it is easy to imagine that a bandwagon effect took place. So this limited scope not only limits the generalisability of the findings, but potentially also indicates 'selection on the dependent variable' (Geddes 1990) – only studying a case where the outcome of interest is likely to have occurred. The longitudinal approach would be more useful if extended to other contexts, with data collected specifically for this purpose, and without any *a priori* knowledge of how likely bandwagon effects are. Such data would cover multiple waves, ask multiple questions about electoral expectations, and cover multiple parties. Doing this would provide clear, nuanced, comparative evidence as to whether popularity information has an observable, real-world effect on the vote. It would also facilitate comparisons across time and space, shedding new light on the conditions in which bandwagon effects are more likely to emerge and how contingent this is.

Just as measuring the bandwagon effect through expectations data is possible through a longitudinal approach, so too could we learn more about how expectations themselves respond to popularity information by measuring the relationship over time. Thankfully, this potentially comes for free in the approach set out in Chapter 5. In order to establish whether expectations are responsive to changes in polls specifically, a more complex approach might be necessary in which the panel data are linked to poll aggregator estimates over time. This would complicate causal inference, but could be combined with the experimental approach demonstrated in Chapter 4 to draw conclusions about expectations balancing internal and external validity.

Study All Types of Bandwagon Effect

Chapter 2 develops a typology of bandwagon effects, which can be static or dynamic and based on conversion or mobilisation. I note that the vast majority of bandwagon research is on static bandwagon conversion effects. Yet, Chapter 3 casts doubt on the validity of this effect in a major election, and both Chapter 4 and Chapter 5 demonstrate that dynamic popularity has important attitudinal and behavioural effects. There is certainly ample room to explore dynamic bandwagon effects further. This is increasingly common in proportional representation contexts

(Dahlgaard et al. 2017; Stolwijk, Schuck, and de Vreese 2017; van der Meer, Hakhverdian, and Aaldering 2016) but also needs to be done elsewhere, in order to establish the extent to which such effects are contextually dependent. Chapters 3 and 5 provide blueprints of how to study these effects in experimental and observational contexts, respectively.

I have been less able to shed light on the significance of bandwagon mobilisation effects, whether static or dynamic. This is because of a familiar problem in social science: it is difficult to measure turnout in surveys, because people who take political surveys either (a) are the type of people who always turn out to vote or (b) do not want to admit they (intend to) abstain (Karp and Lühiste 2016; Prosser and Mellon 2018; Prosser et al. 2020). This problem plagues both Chapter 3 and Chapter 5. Nevertheless, very large sample sizes could permit the application of the methods in Chapters 3 and 5 to the study of mobilisation. Failing this, it may be that future evidence on the prevalence of such effects has to come from experiments in the economic tradition, which measure turnout through abstract, incentivised games (e.g. Agranov et al. 2018).

Expand the Scope of the Question – 'Indirect' Bandwagon Effects

Finally, all of the contributions I have made, and all of the above suggestions for future research, remain within the specific question of what happens when voters are exposed to popularity information – how does this affect their attitudes and behaviours? But the ways in which popularity information shapes electoral fortunes are not limited to this direct route. Popularity information, particularly from the polls, shapes the behaviour of political actors, and this behaviour can have important effects on politics. Even just remaining in the realm of the bandwagon effect, there is enormous potential for indirect influence. That is, it is possible to imagine many ways in which the polls might influence political actors such that their subsequent behaviour benefits parties that are doing well in the polls. This is not to mention wider influences on politics writ large.[2]

In the 1997 UK general election, for example, it was a widely held – though likely false – belief that polls published by *The Observer* had a (deliberate) direct impact on the vote. Cowley points out, though, that this is not the only reason these polls would have mattered, because

> At a grassroots level, their publication almost certainly affected the nature of the campaign in the seats covered... Even where the results were not widely circulated, publication may still have affected the nature of the campaign. One victorious Labour candidate believes that the polls were not 'widely read, or widely publicized in the local area or indeed had any effect whatsoever on the final outcome.' But he admits that 'it might just be that it spurred us on to put in the extra effort as we approached the final fence' (2001, 966).

More generally, Crewe (1992, 478) argues that 'opinion polls are not... the mere fluff of elections. They influence the timing, strategy and course of election campaigns and, to that extent, the result.' These kinds of considerations are what led Henshel and Johnston to formalise the idea of 'indirect' bandwagon effects,

[2] Consider, for example, that many claim the popularity of the United Kingdom Independence Party largely led David Cameron to call a vote on Britain's membership of the EU (Goodwin 2015). Also, scholars of 'mainstreaming' have highlighted substantial shifts in the rhetoric of mainstream political parties following the growth of those at the extremes (Brown, Mondon, and Winter 2021).

proposing that

> Existing models [of the bandwagon effect]... miss a major alternative: *indirect causes* in which the forecast first operates selectively on certain key individuals, influencing their decisions in ways that, in turn, influence the election outcome. Specifically... (1) through alteration of the flow of *financial contributions* to candidates, (2) through alteration in *volunteerism* (campaign support), and (3) through alteration in the quality and quantity of *endorsements* (1987, 495–96, emphasis in original).

This theoretical innovation has never been taken up in research on the bandwagon effect since. This thesis is no exception to that. But these indirect processes are certainly a part of why it could be good for a party or candidate to poll well. Such effects may even be so substantial as to override any of the more direct processes. Perhaps the fact that voters might undergo a slight (direct) bandwagon effect when they see a poll pales in comparison to the part these polls play in determining how a campaign is conducted, via increased donations, volunteerism, or endorsements – or just 'spurring on' the candidates. It is relevant to this argument that much of the vote intention polling survey companies conduct is never even seen by voters. It is undertaken privately, for parties, candidates, and campaigns themselves. These polls *cannot* induce the kinds of bandwagon effect that have been my focus. But they have the potential to affect the vote indirectly.

It is, of course, useful to know about these indirect processes simply in order to understand politics better. This would, after all, be major knowledge about 'who gets what, when, how' (Lasswell 1936). But this would also be useful in order to contextualise the bandwagon effect, as it is currently understood, as part of the broader political significance of popularity information such as the polls. This would enable bandwagon research to speak better, and with more nuance, to what exactly the place of polls is in the democratic process.

Summary

This research project set out with the goal of answering a universal political question: is there a bandwagon effect? The answer to this question takes the form of: maybe, sometimes. Yet what this question has done is guide the thesis through solving important puzzles along the way to understanding it. In pursuing the somewhat elusive bandwagon effect, I hope to have made substantial contributions to knowledge, and how to acquire such knowledge, about the role of popularity information in modern electoral democracies. Beyond this, as set out in this final chapter, my findings have more immediate implications for politics in the real world. They also pave the way for future research that both fills the gaps I have left and extends the insights I have provided.

Appendix A

Telling Lies to Tell the Truth? Reflections on Experimental Design

> When I pretend or engage in make-believe, I close my eyes to the world around me, sometimes literally, the better to imagine a world that isn't actually there.
>
> **Agnes Callard**, *Aspiration: The Agency of Becoming* (2018, 84)

> If you're honestly trying to tell what happened, you find facts are very obstinate things to deal with. But if you begin to fake them, to pretend things happened in a way that makes a nice neat story, you're misusing imagination. You're passing invention off as fact, which is, among children at least, called lying.
>
> **Ursula K. Le Guin**, *Making Up Stories* (2019, 108)

New recommendations adopted by the American Political Science Association (APSA) in 2020 state that deception in research with human subjects can take at least four forms (see Landgrave 2020, 494). They assert that all of these should be carefully considered and justified. The first three forms are, collectively, the subject of extensive commentary in academic work on experimental ethics:

> Identity deception: Deception about who you are (a researcher in political science) or with whom you are working.
>
> Activity deception: Deception about what you are doing (e.g. research for social science) or the situation confronting research participants.
>
> Motivation deception: Deception about the reasons for the research or the use to which the research or data will be put.

In these cases, false information of a specific type is given to respondents: false information about the way the research project is being conducted. It is widely argued that this is an unethical practice (e.g. Berghmans 2007; Kelman 1967). For example, in political science, there was recently some uproar about a study conducted by researchers from King's College London and the London School of Economics that 'involved sending emails from fictitious constituents claiming they were concerned about financial support during the coronavirus lockdown' (Allegretti 2021; cf. Cowley 2021). People saw this as a problematic instance of all three types of deception above.

But a different type of false information can also be given to respondents: a false *treatment* that says nothing about the conduction of the experiment. For example, I could make up a poll result to give to the treatment group in an experiment on the bandwagon effect (e.g. Dahlgaard et al. 2017), without engaging in identity, acitivity or motivation deception. APSA dubs this 'misinformation':

> Misinformation: Providing false information about the state of the world – e.g., by providing unreliable or inaccurate information about political candidates.

While this does not necessarily mean that misinformation takes the form of false treatments, this is most commonly where misinformation is communicated in research on the bandwagon effect and related questions (e.g. Dahlgaard et al. 2017; Dizney and Roskens 1962; Madson and Hillygus 2019; Mehrabian 1998), and likely elsewhere, so for simplicity I will treat misinformation and false treatments interchangeably here.

When designing the research carried out in this thesis, I had to grapple with the ethical implications of 'misinformation' and whether or not I could justify engaging in it. Notwithstanding its inclusion in political science's foremost ethical guidelines, scientific and philosophical attention to the implications of this type of deception is scarce. This essay sets out what these implications are, the conditions under which they arise, and how they complicate 'utilitarian' perspectives on research design. I also explain how I have attempted to account for all of these considerations in the experiments conducted in this thesis.

First, I ask whether it is always unethical, or problematic, to invent treatment information. This section makes a distinction between what will be called 'pretend' and 'pretence' experiments. The difference between these ideal-types is in whether or not they invent an artificial context, to go with the invented treatment. 'Pretend' experiments are those where treatment information is made up in a made-up context; 'pretence' experiments are those where treatment information is made up in a *real* context. Linking this to philosophical work on 'lying' (Carson 2006), I argue that this contextual point matters. Pretend experiments are of little concern because the artificial context means that they do not constitute a lie; pretence experiments, by the same token, do constitute a lie. I explore why these lies matter ethically in both deontological (Bok 1995; Kant 1978 (1798)) and consequentialist terms (Kelman 1967). The argument of this first section is that when experimental designs invoke a real-world context, researchers incur an ethical cost if they make up false treatment information about that context – whichever way we look at it.

Arguably, such costs are acceptable if they are outweighed by the benefits of the research – because the costs are small, or the benefits are large. If experimental research on the bandwagon effect teaches us things we want to know about the political world, as repeatedly pointed out in this thesis, then it may matter relatively little that a few survey-takers were lied to along the way. But there is an important point that this 'utilitarian' account has tended to miss: research with false treatment information not only has a higher ethical 'cost,' but also usually a lower practical 'benefit.' That is, although it is usually thought that invoking a real-world context improves causal inference, false treatment information about that context potentially compromises this gain – in terms of external and internal validity. So-called 'debriefing' does not necessarily resolve this. This means that balancing the utilitarian equation is often not as simple as typically thought.

The final section discusses how to navigate, and how I have navigated, experimental design in light of these complications. First is the option of sticking to 'pretend' experiments, as indeed I did in Study One in Chapter 4. Another option is to move towards the 'pretence' approach, but attempt to account for its potential downsides in various ways. This is what I did in Study Two in Chapter 4. Finally, researchers can simply attempt to tell participants the truth. For example, respondents can be given real poll results from a real election; neither context nor treatment is made-up. In a pure form, these 'true treatments' would represent an ideal case minimising ethical implications (costs) while also maximising internal and external validity (benefits). The experiment in Chapter 3 of this thesis approximates this. Each of these approaches has limitations, and they are not an exhaustive list. They simply demonstrate the ways in which ethical and inferential considerations can interact to inform experimental research with popularity information as a treatment.

Pretend and Pretence Experiments

In experimental research, the treatment is the thing whose effect we want to estimate. Experiments work by giving (some) people this thing, and then seeing what they do in response. In much political science research, experiments involve treating people with information of some form. In research on the bandwagon effect, this is usually the result of a poll. As I explained in Chapter 2, the logic of this is that, if people who see a poll are subsequently more likely to say they will vote for the front-runner, this is evidence that popularity information influences the vote, in line with the bandwagon effect. Researchers have an opportunity here to give respondents any poll results they want, to answer whatever specific question they have: what if X candidate were in the lead? What if Y candidate had a larger lead? A wealth of such opportunities present themselves because the researchers can simply make up the information, if they so choose. In the terms employed by

Table A.1: Pretend and pretence experiment examples.

Pretending to be on Mars	**Voting for President under False Pretences**
Imagine you inhabit the planet Mars, and as a Martian you are eligible to vote in an upcoming presidential election between the candidates Marvin and Martin. A survey of your fellow Martians suggests they are going to vote as follows:	A recent poll of vote intentions at the upcoming presidential election, conducted in your state of New York, put Joe Biden and Donald Trump on the following vote shares:
Marvin 54% - 46% Martin	Trump 54% - 46% Biden
Who would you vote like to for?	Who would you vote like to for?

APSA, researchers can engage in 'misinformation.'

Consider the two stylised examples in Table A.1. Both present the participant with polling information. But in both, the poll is made-up. The information is false. Yet the two examples are very different. One invites the participant to imagine she is in an entirely made-up situation and then also invites her to imagine made-up information about that context. The second intervenes in an election that is taking place in the real world and gives respondents false information about that election. What differs here is not really the treatment information and whether or not it is true – in both it is the same, and false – but the *context* in which this treatment is delivered, and whether or not this context is real. Treatments are always given within a context, so researchers have to make a decision about both of these dimensions.

I will call 'pretend experiments' those cases in which both the context and the treatment are made-up. The experiment is hypothetical, or make-believe. This is like the Martian case in Table A.1. Everything is pretend. I will term 'pretence experiments' those cases in which the context is real, but the treatment is made-up.

Given a real, usually ongoing context, researchers elicit respondents' attitudes or behaviours by false pretences. This is what the presidential election case in Table A.1 does. There are plenty of examples of both pretend (e.g. Fleitas 1971; Goidel and Shields 1994) and pretence experiments (e.g. Dizney and Roskens 1962; Madson and Hillygus 2019; Mehrabian 1998) that use made-up polling information as their treatment.

Beyond this, the context of a treatment matters because context determines whether a false statement is a lie. As Carson (2006, 298) argues, a false statement is a lie when it is delivered in a context that 'warrants the truth' of the statement. Statements are assumed to be true unless there is good reason to expect otherwise. Context is what guides this assumption. That is, in order to be able to deliver a false statement without that statement being a lie, one needs to establish a context in which the default assumption of 'warranting the truth' no longer holds. By either making the context something real or something invented, researchers affect whether their treatments – which are, here, untrue statements – can be seen as lies or not. This is consistent with the emphasis APSA places on 'the state of the world.' False information *about the state of the world* is misinformation.

In pretend experiments, the context does not warrant that statements about it be true. Indeed, there is no clear standard against which the veracity of any such statements could even be assessed. Who can say whether Marvin or Martin is really polling better in the Martian election? A useful analogy is to consider the logic by which fiction – especially science fiction – operates, in that

> You can read… a lot of… science fiction, as a thought experiment.
> Let's say (says Mary Shelley) that a young doctor creates a human

> being in his laboratory; let's say (says Philip K. Dick) that the Allies lost the Second World War; let's say this or that is such and so, and see what happens... In a story so conceived, the moral complexity proper of the modern novel need not be sacrificed, nor is there any built-in dead end; thought and intuition can move freely within bounds set only by the terms of the experiment, which may be very large indeed (Le Guin 2017, xiv).

Elsewhere, Le Guin (2019, 108) uses this logic to point out that 'fiction is invention, but it is not lies.' By inventing and imagining, science fiction authors have licence to see where this imagination leads – and they cannot lie about that, in any meaningful sense.[1] In pretend experiments, the participants play a role in deciding where the imagination leads. They are not being lied to any more than the reader of science fiction is being lied to, or the characters are being lied about. Both, as Le Guin puts it, resemble thought experiments. I ask you to imagine that you are on Mars. Once I do this, I can make up pretty much anything about that context. But rather than writing a book about it in which I tell you what happens next, I give you something I am interested in and ask you to play along in *imagining* what happens next.[2]

By the same token then, pretence experiments must involve lies. The real-world context they invoke provides a reference point of true information that the treatment

[1] Hannah Arendt (1972, 5) appears to blur this distinction somewhat when arguing that 'the ability to lie,' like the ability to 'act,' owes its existence to 'imagination.' Yet she crucially defines this 'ability to lie' as 'the deliberate denial of truth.' Fiction and lies may both be grounded in the capacity to imagine alternatives to the truth, but it is when one states these alternatives about a factual reality, rather than 'invented' reality, that one is telling a lie.

[2] Note that philosophers such as David Lewis (1978) have provided more complex discussions of how we can define 'truth in fiction' that complicate to some extent what I have said here, by questioning on a more fundamental, metaphysical level what it means for something to be 'true.' I avoid dwelling on this discussion because I am only using fiction as an analogy.

does not meet. Donald Trump never polled anywhere near as well as 54% in New York. By telling you that I am talking about the (upcoming) presidential election, I establish a context that warrants that my statement be true. False information about that context is therefore a lie. If pretend experiments mirror science fiction, pretence experiments mirror false biography – or fictitious non-fiction. In Le Guin's terms, they 'misuse imagination' by 'passing invention off as fact' (2019, 108).

Moreover, Bok (1995) draws on Iris Murdoch's (1992) distinction between *truth* and *truthfulness* to note that, just because it cannot be said that any particular fact or statement about the world is necessarily true, it is not therefore impossible to lie to, deceive or deliberately mislead others when talking about that world. That is, regardless of whether there are any objective truths, people still make a choice about whether to be truthful in their interactions with others (see Frankfurt 1998, 133). Pretence experiments are a choice not to be truthful. In APSA's terms, they misinform respondents about the state of the world.

The Ethics of Lying

Whether or not an experiment involves lies matters because lying in experiments is arguably unethical. There are two ways to look at this. The *deontological* view sees lying should as unethical full stop, so lying in experiments is unethical by extension. The *consequentialist* view sees lying in experiments as unethical because it entails undesirable political and scientific consequences.

First, the Kantian deontological view. Here, lying in experiments is troublesome because it treats people as means to some other end. This violates Kant's 'highest maxim, or subjective rule of life' (Bok 1995, 14), of 'uninhibited truthfulness toward oneself as well as in the behaviour toward everyone else' (Kant 1978 (1798), 207). Telling the truth, in this framework, is the most fundamental way in which people maintain dignity in themselves, and respect for others. This extends through to the research context, in which experimentalists' own dignity is compromised when they lie to, and thereby disrespect, participants.

Some criticise this view, though, on the basis that it opens a can of worms. Berghmans writes that

> The major problem with deontological absolutism is that it involves a high moral ideal of truth and truthfulness and as a result of this would frustrate any research project in which the researchers do not live up to this high ideal, or in which the research participant is not fully aware of all the ins and outs of the study in which he or she is participating (2007, 14).

In the case of pretence experiments, it could be argued that false treatment information is just one more way in which participants are deceived, which is inevitable in an experiment. Kant's 'uninhibited truthfulness' would compromise experiments in many cases. For example, in the pretence experiment above, if I told respondents 'this experiment is designed to test whether you will vote for a candidate just because he is in the lead,' this could compromise the findings by giving rise to 'demand effects' (Zizzo 2010). On this basis, few experiments will ever reach Kant's bar. Indeed, APSA's revised guidelines note that 'research integrity' often depends on some level of deception. Short of rejecting the experimental approach in its entirety, this arguably means that a deontological critique cannot be used to guide experimental design. At least, it would seem that it provides little grounds on which to criticise pretence experiments as a special case. This is, of course, without mentioning the fact that researchers are human beings who therefore, to a number, will all go about their regular lives telling fibs or otherwise being economical with the truth.

To stop here would be to ignore the fact that pretence experiments are capable of having non-trivial negative effects. They can be seen as problematic from a consequentialist standpoint. A prominent concern in discussions of deception (of the identity, activity and motivation forms), is its effect on trust. Deceiving people gives them less reason to trust you. Where Bok (1995) emphasises this because trust is a social good that is to be valued in its own right, others point out that compromising research participants' trust in researchers could have several negative downstream effects on future research, because

> Deception contaminates the participant pool. Whereas in sociology it was suggested that a likely outcome of deceptive practices is participants' future resistance to other research efforts... psychologists and economists have expressed concern that the expectation of being deceived produces suspicion and second-guessing and that these reactions—rather than the experimenter's scenario and instructions—guide and ultimately distort experimental behavior (Hertwig and Ortmann 2008, 63).

It is also possible that deceptive experimental practices have an effect on the trust of those who never even participate in experiments, through their encounters with participants, or otherwise hearing about deceptive experiments (Orne 1962). Given rising rates of survey non-response and the 'overextraction' of human respondents by social research (Leeper 2019), which already place increasing strain on scholars' ability to conduct good survey research, it is better to avoid putting more people off participating. This might also be particularly concerning in the case of 'elite field experiments' or 'audit experiments,' where potential participants are more savvy to the deceptive practices used by researchers (Landgrave 2020, 490).

These concerns arguably extend to the case of pretence experiments – deception of the 'misinformation' form. It is easy to imagine someone being put off by participating in an experiment which would have her believe Donald Trump has a clear lead in the polls in New York, which is patently untrue. It is even easy to imagine her passing this on to her friends. This is bad for social science, so researchers should seek to avoid it.[3]

Beyond this, consider that the treatment itself is expected to have an effect on participants, and that this effect is potentially problematic if it is based on a falsehood. In the pretend experiment above, the treatment is designed to influence people's vote choice in a made-up election. This behaviour is not founded on anything true, but it is also not directed towards anything real. Nothing happens in the real world as a result. In the pretence experiment, however, the fact that the context is real means that it is typically likely the respondent will already have some behavioural intention in this context, which the treatment seeks to change.[4]

[3] Arguably, regular survey respondents – perhaps in contrast to 'elite' participants – are unlikely to have sufficient awareness of the ways in which social science is conducted to be able to catch onto the fact that an experiment is deceiving them. As I will note below, though, this is less likely to hold the more unrealistic the treatment is, and is almost guaranteed not to hold if 'debriefing' tells them they were just lied to.

[4] If the respondent has no such intention or pre-existing attitude, then the experiment may (seek

Participants probably have an idea which candidate they are more likely to vote for – Biden or Trump. The treatment is designed to change this. Research intervenes in the real world and manipulates this choice by false pretences. Participants might end up voting for Trump where they would have voted for Biden, and this choice is based on information that they could not have received in the real world, because it is false.

A first objection to this might be that this simply is not the intention of such experiments. Rather, the intention is to affect how people behave in the experiment without this spilling over into the real world. But while this may be the intention, that does not guarantee spill-over will not happen. This potential counterargument also casts doubt on how useful an experimental approach even is in the first place. It plays into the hands of those who criticise the experimental approach precisely because it does not tell us much about real-world behaviour: there is a leap in assuming how people say they will vote under experimental conditions reflects how they will really vote (Green and Gerber 2003, 101). It is therefore not a particularly useful point to make when considering what the negative consequences of experiments are on the condition that experiments are useful in the first place.

A second objection could be that the example I have provided is too stark – treatments can be 'made up' and much more closely approximate reality. For example, I could flip the candidates in the pretence experiment in Table A.1 and the treatment would be much closer to the truth.[5] This is certainly a grey area, and ultimately

to) form one.

[5] The fact that researchers could feasibly make up treatment information that *accidentally* approximates the truth raises the question of whether they are 'lying' or instead 'bullshitting.' Frankfurt (1998, 129) explains that bullshit is 'produced without concern with the truth' but 'need not be false,' whereas lies are knowingly false. In response to this, it is first worth noting that this would not somehow make misinformation more acceptable: 'bullshit is a greater enemy of

comes down to the fact that pretend and pretence experiments are meant to be ideal-types which, in practice, experiments *approximate* rather than fully embody. It is possible that researchers using pretence experiments could reduce ethical concerns by using more realistic treatments – a point I return to when discussing my experimental approaches below. Yet, arguably, this is not always possible, depending on the type of treatment information. In bandwagon research, the treatment is typically numerical and how close it comes to the 'truth' can therefore be assessed – for example, by comparing it to real-world polls. It is less clear how we would do this for more categorical information – for example, reporting a candidate's ethnicity. Having to use treatments that approximate the truth also limits the type and number of questions pretence experiments can attempt to tackle, by putting more extreme counterfactuals off limits.

A third objection could be that although experiments may alter people's behaviour in the real world, they will not alter *many* people's behaviour – the effect will be trivial. In the bandwagon case, an experiment with 2,000 respondents, assuming everyone is entirely unaffected by treatment except ten percent of the treatment group (itself half of the full sample), might change the vote choice of 100 people. This is vanishingly unlikely to affect the outcome of the election, so why should we care? Yet this implies that we also should have less reason care about the bandwagon effect itself in most cases, given that it is nearly always also a small effect (see Hardmeier 2008). So again, claiming that small effects are trivial is not

the truth than lies are' (Frankfurt 1998, 132). Second though, it is unreasonable to assume that political scientists who are experts in the subject matter of their experiments are unaware, or pay no heed to, the ground truth of the information they give to respondents. Indeed, it is arguably the ability to make up information that contradicts reality that gives pretence experiments their appeal to researchers, as already noted and discussed further below. It seems unlikely that experimentalists see their research as a chance to make up 'any old bullshit' and see what happens. Instead, they have hypotheses about specific counterfactuals that they *know* are counterfactual.

a particularly useful argument to make on the assumption that a whole political science literature showing the bandwagon effect to be small is itself non-trivial. This could be a matter of degree though – if the bandwagon effect is small in the population, then the bandwagon effect in an experiment with a sample from that population is much smaller. Yet this would imply that the larger the sample, the more we should care about the size of the effect, because the less trivial its real-world impact could be. But of course larger effects, and larger sample sizes, are also less scientifically trivial, making us want to study them more. So the number of people affected – in terms of both effect size and sample size – looks like something we should want both to maximise and to minimise. This hints at the kind of utilitarian thinking that I explore further below.

The triviality objection also implies that some cut-off point is required at which the effect would become non-trivial. Perhaps it is the point at which the experiment changes the outcome of an election. Yet it is impossible to know what this point will be *a priori*, so this cannot be the basis for the design of a pretence experiment. For a given election that has already taken place, we can say what it would have taken for the balance to be tipped, but we cannot know this ahead of time. It is also not necessarily the case that experiments just have one, immediate spill-over effect. Consider, for example, that voting is known to be habit-forming (Coppock and Green 2016; Cutts, Fieldhouse, and John 2009) and voting a certain way changes how people subsequently think about politics (Acharya, Blackwell, and Sen 2018b; Bølstad, Dinas, and Riera 2013). Experiments on the bandwagon effect that change people's behaviour in one election could have unmeasurable knock-on effects for their future behaviour. Remember that all of this would be on the basis of false pretences.

Note that none of this is to suggest that experimental research has ever substantially changed election outcomes or produced major realignments. But it is entirely possible that many of the things political science investigates have never, in themselves, determined election results or changed the course of politics. Voting is mostly stable and predictable (Gelman and King 1993), and voting behaviour experiments largely investigate influences at the margins. My point is that it is not straightforward to apply the triviality criterion in designing an experiment or deciding whether a study should be conducted.

The False Promise of Debriefing

Many propose 'debriefing' as a way to resolve these potential issues (Berghmans 2007). In the pretence experiment example, suppose respondents are told after the experiment that the poll results they saw were fake and the purpose was to see how they would behave if they were real. This should remove spill-over or otherwise downstream effects, because people should see that there was no basis to their change in behaviour or attitudes that they went through in the experiment, cueing them to reverse that change.

This only holds if the new behaviour observed in the experiment is entirely and directly brought about by the information contained in the treatment. Yet in this thesis, I have noted the importance of the broader information environment in which new 'popularity information' is received. The cognitive response mechanism explored in Chapter 2 points to the fact that when people change their mind in response to poll results, to vote for the leading candidate, this is likely because they reflect on other arguments supportive of that candidate, as a reaction to the new information (Mutz 1998, 212). This suggests, therefore, that the behaviour observed in the pretence experiment would be the product of reflection on arguments supportive of candidates other than just their popularity, brought about by false information about their popularity. These arguments then become more salient to the respondent, and debriefing cannot guarantee that they will cease to be so, because the debrief cannot hope to address all of these additional considerations. Indeed, in many cases these considerations will be based on valid information and therefore it would be questionable for the researcher to attempt to reverse them. The way in which the respondent thinks about the candidates in the election might be fundamentally changed by participating in an experiment in which she is misinformed about these candidates.[6] Importantly, as noted in Chapter 2, this proposed 'mechanism' is consistent with broader political psychological theory about how new information leads to attitude or behavioural changes (e.g. Lodge and Taber 2013; Marcus, Neuman, and MacKuen 2000; Zaller 1992). The problem is not unique to this example.[7]

Some of the negative consequences of untrue treatments are possibly even more likely to occur *because of* debriefing. Consider, for example, the trust problem noted above. If respondents do not find out after the experiment that the information

[6] Paul and Healy (2018) have proposed the framework of 'transformative treatments' to describe examples in which experimental treatments cause shifts in people's preferences or values. Under this framework, the person who receives a debriefing message could be a fundamentally different person from the one who entered the experiment.

[7] Some of these alternative theories put the emphasis on 'hot' or otherwise 'affectively charged' cognition (Lodge and Taber 2013; Marcus, Neuman, and MacKuen 2000). In the case of Affective Intelligence Theory (Marcus, Neuman, and MacKuen 2000), central to how persuasion works is by inducing 'anxiety' in people, which motivates a 'search' for information not unlike Mutz's 'rehearsal' of arguments. Debriefing may reassure people that they do not need to have been anxious, but it cannot undo the anxiety they experienced – which was, recall, based on false information. Kelman (1967) questions the researcher's right to affect the psychology of participants in such ways at length, albeit employing more extreme examples.

they were given was untrue, they may see no reason to believe it was, meaning their trust in researchers is unaffected. Debriefing, as such, cannot necessarily undo the effects of experimental treatments on either politics or science. In some ways it might even make them worse.

Weighing up the Costs and Benefits: The Utilitarian Case

Arguably, the benefits of experimental research can be weighed up against these ethical costs. To some extent, I have already alluded to this idea in discussing objections to my arguments above. This utilitarian perspective holds that lying 'is prima facie morally wrong, but can be justified in the light of the good consequences which may flow from the experiment in the interest (actual or future) of others' (Berghmans 2007, 14). The idea is that researchers can justify untruth because it helps them to reach greater truth. In Bok's (1995) and Murdoch's (1992) terms, the truth justifies the lack of truthfulness. This greater truth might then also be applied to benefit society, at the small cost of an initial lie. Researchers tell lies in order to tell the truth.

It is worth noting that this thinking can explain why researchers often avoid using pretend experiments. It is questionable whether the results of pretend experiments are particularly useful in the real world, given that they are about imaginary worlds. Pretence experiments are supposed to be more externally valid, by virtue of their real-world context. Because they combine this with the reasearcher's freedom to make up the treatment information, they enable externally valid causal inferences about a whole range of effects that could not be studied as successfully either in observational evidence or in pretend experiments. For example, the pretence experiment above enables a researcher to answer the question of what would happen if a poll were published in New York suggesting that Donald Trump is in the lead there. This is the benefit they offer that is meant to justify their ethical cost.

This implies that the apparent benefits of the research are entirely unaffected by the fact that it is based on a lie. The assumption is that made-up treatment information creates ethical problems, while only producing benefits in terms of inference. But this is unlikely to be the case. The first way to see this problem is to reflect on the factors that affect external validity. It not only matters that the *context* reflects the real world; it also matters that *treatments* do.[8] In the pretence experiment example, we can think about there being a population of polls out in the world. The experimental design needs to maximise the representativeness of the sample of polls used in the experiment with respect to this population, if its results are to be generalisable to the relevant context: in my example, the real US presidential election (de la Cuesta, Egami, and Imai 2021). If the design and analysis of the experiment implicitly treats Trump's lead in New York as no more or less likely than Biden being in the lead (which the polls would really have suggested), then the effect of this scenario can be seriously miscalculated when mapping it onto the real world (see de la Cuesta, Egami, and Imai 2021). Treatments that are not reflective of reality harm inference by compromising external validity.

A useful, additional way to see this is to think about what it would mean for

[8] Typically, discussions of external validity also focus on the representativeness of the respondent sample.

Trump to be ahead in the polls in New York. It is likely that this would require a fundamentally different political context that would greatly alter the behaviour of respondents in the experiment above and beyond the effect of treatment. The findings of the experiment will be based on the *real*, current political context in which these alterations do not apply. So this study would not really tell us what would happen to voters in New York if Trump were shown to be in the lead in the polls, because it is so far from capturing what this situation would actually entail. Pretence experiments are explicitly about the real world, and their untrue treatments are not from that world.

It is possible to imagine that this might encourage many respondents to make assumptions beyond what the treatment is trying to measure. A fake poll might make them assume broader differences in the election context. They might think, for example, that the economy must be doing well if the incumbent Trump is polling so well in New York. This could be what is of interest to them, and not the treatment. This is similar to the discussion of the 'cognitive response' above, but differs in that it suggests people are *inventing* rationalisations for the treatment, not calling upon what they already know to rationalise it. Importantly, if the treatment is false, so are these rationalisations likely to be. This means the experiment will end up measuring the effects of these uncontrolled, unobserved, false imaginings about the world, rather than just the treatment itself.

With this in mind, pretence experiments also have implications for internal validity, because they compromise the tenet of 'experimental realism' which holds that

> To the extent that subjects become psychologically engaged in the process they confront, internal validity intensifies. Similarly, internal

> validity diminishes in the face of subject disengagement… If subjects approach a task with skepticism or detachment, then genuine responses fade… This raises the possibility that measures obtained do not accurately reflect the process being manipulated, but rather manifest a different underlying construct altogether (McDermott 2011, 28).

Consider what might happen when providing respondents with false information. As a result of 'skepticism,' respondents might not look at the experimental task as something to take seriously. Rather than getting them to think about the election and how they would behave, instead the false treatment leads them to behave in arbitrary ways. They might also see the false treatment as deliberately pushing them in a certain direction, and then do what they think the researcher 'wants' them to do (Orne 1962; Zizzo 2010). In both of these cases, the effect observed in the experiment does not fully reflect what people would do in the real-world scenario on which the experiment is designed to comment. The benefit of the realistic context is compromised.

Arguably, this concern about experimental realism is only valid if we assume participants *know* they are being lied to. If they are really being deceived then they should not realise that the treatment is unrealistic. This is a valid critique, but it also means that the ethical problems discussed above presumably must be taken more seriously. Essentially, if a treatment is deceptive, it has to pass a threshold of realism. If it is not realistic, it must not be deceptive. If an experiment is successfully duping participants, then it may be more realistic than if they recognised the treatment as untrue, but it is therefore also more concretely deceptive. Pretence experiments, seen in this way, depend on being deceptive in order to maximise their scientific value. To the extent that they are effective, ethical concerns about deception become more pressing. But, as I have just shown, as well as being ethically problematic, this deception also creates other problems for how valuable they are. This interplay of realism and deception, therefore, further complicates the utilitarian case.

The key point, then, is that the weighing up of costs and benefits in pretence experiments is not as simple as it might first appear. The lies researchers tell participants in these experiments do not simply produce ethical costs that are outweighed by the unique combination of external and internal validity afforded by experimental designs in real contexts. Rather, in some ways, these lies compromise external and internal validity.

Resolving the Dilemma: The Experiments in This Thesis

Faced with this challenge, this thesis features three partial solutions. The first is to try to extract as much potential as possible from pretend experiments, to overcome their limitations, without needing to lie. This is what I did in Study One in Chapter 4. The second is to move more towards pretence experiments, while attempting to minimise the ethical and inferential problems set out above. This is what I did in Study Two in Chapter 4. The third is to avoid using false information in experimental treatments. This is what I tried to do in Chapter 3.

Maximising the Potential of Pretend Experiments

As noted above, pretence experiments are used in order to do better than pretend experiments can offer, particularly in terms of external validity. However, in some cases, pretend experiments still serve as a good starting point. Because they come with negligible ethical concerns, the utilitarian view should hold that they are worthwhile providing they offer some benefit, even if this benefit is small.

Arguably, pretend experiments are useful as an approach to causal questions about political behaviour about which we know very little, or for which we have little to no evidence. They create an abstract context in which respondents have little useful information to go on other than the treatment. If their decision is not affected by that treatment, and we fail to reject the null hypothesis, then this gives us reason to be very suspicious as to whether the treatment would have any effect outside of this stylised, pretend context. This means that pretend experiments provide a conservative test of the null hypothesis. This is a good place to start when the dominant way experiments are used is as a form of null hypothesis significance testing. It means that pretend experiments can be a good starting point in working out whether a theory is worth pursuing with more rigorous, externally valid study.

Consider the experimental design used in Study One in Chapter 4. This is the first attempt to experimentally assess whether people's electoral expectations are influenced by dynamic popularity, or 'momentum' in the polls over time. Because this is the first time this has been tested, it is useful to know whether it is even possible to reject the null hypothesis in the abstract. If the treatment had no effect in these experiments, it would have been very unlikely that it would have an effect in non-imaginary circumstances, where it is not the only information people have, and they have a real reason to care about the election outcome. The broader point is that the ethical implications of such experiments are negligible, so on a utilitarian basis, the possibility that they are beneficial, however minimally so, is likely to justify conducting them.

Minimising the Problems With Pretence Experiments

Once the relationship has shown up in a pretend experiment like this, it is then useful to implement changes and see whether the effect holds up. However, it is difficult to enable people to use more information, and encourage them to care about the election outcome, without introducing a real-world context. Or, at least, this is the easiest and most obvious way to do so.[9] This is what I did in Study Two in Chapter 4. The introduction of a real-world context, I argued, gave people chance to use their existing information and predispositions in order to react to the treatment information as they would react to it in the real world. However, in doing so, I moved towards the pretence experiment format, and risked the problems set out above.

In order to minimise these concerns, I made a series of design choices.[10] First was that, although the political contexts (UK and Canada) and parties (Conservative, Labour, and Liberal) were real, they were not really currently competing in an election. I did not intervene in a real ongoing electoral process. To some extent,

[9] This is how 'psychology' experiments tend to try to increase theoretical interest, whereas 'economics' experiments tend to do so using highly stylised games in which preferences and information are controlled using incentives (Dickson 2011).

[10] These are all design choices that I discuss in Chapter 4 itself, which is why I do not explain them in detail here. The situation is slightly more complicated in the Canadian case, but that is again explained in more depth in Chapter 4.

this keeps things closer to the realm of pretend experiments. This is an example of a case where the lines become somewhat blurry between pretend and pretence experiments. Respondents are imagining there is an election taking place. It just so happens that they are imagining an election taking place in real countries.[11] To encourage and emphasise this, the wording of the experimental vignette asks respondents to 'imagine' there is an election taking place. In one condition, this election was taking place in Canada, where respondents (from the UK) are unable to vote. It therefore could not have any downstream consequences for their vote choice in that electoral context. Moreover, although I made up the vote shares the parties were on and this did not mirror current polling at the time of the experiment, they were nonetheless realistic vote shares that the parties have achieved in recent elections (at least in the UK case). The treatment itself was also designed to reflect how much the parties' vote shares in the polls would have to change in order to produce the figures given in the experiment. Again, this demonstrates a grey area in pretence experiments which I noted above: treatments can be made-up, but still approximate the truth, even accidentally. It is likely that consequentialist ethical concerns, and inferential problems, diminish the closer the treatments approximate reality.

In sum, I aimed to minimise ethical concerns by establishing a more 'imaginary' rather than truly 'real' context, and I aimed to maximise realism by tailoring treatment information to make sense in terms of the real-world polling environment. Arguably, these are the kinds of changes that are necessary in order to account for the complicated cost-benefit analysis that comes with pretence experiments.

[11] This is quite a common approach (see, e.g., Campbell and Cowley 2014; Campbell et al. 2016; Cunow et al. 2021; Roy et al. 2015).

Applying the utilitarian approach, I believe the added value of this experiment, given these adjustments, justifies any (reduced) ethical cost. This is, of course, up for debate.

True Treatments

It is possible to avoid these concerns altogether, though, by simply using true information in experimental treatments. This would involve taking data from the real world and deploying it in a real-world context. In experiments on the bandwagon effect, this means taking real polling data and exposing respondents in the treatment group to it in a real ongoing election (Cook and Welch 1940; Tom W. G. van der Meer, Hakhverdian, and Aaldering 2016).[12] This enables researchers to intervene in real elections and conduct an experiment without facing the charge of unduly influencing it by false pretences. They might still affect the election outcome in some small way, but they will do so by simply giving people information that they could have encountered in the real world. In the case of polling, it could even be argued that they are *likely* to encounter such information, because of the ubiquity of poll results in election coverage. In such cases, any change they bring about may have been likely to happen anyway. And given what Chapters 1 and 2 pointed out about people's rational, informational use of poll results, such influence can mostly be seen as unproblematic because it pushes people towards appropriate political choices.

This is the approach I took in the experiment in Chapter 3. I randomly assigned

[12] There is also the option to use real polling information in an imaginary context, but this would represent a wasted opportunity to leverage a real-world context to improve external validity.

voters in the US presidential election to see *real* poll results. To the extent that this has effects on their voting behaviour, this just reflects the fact that once they saw that information in their real life, they would have changed their behaviour in the same way.

Nonetheless, the experiment in Chapter 3 is not perfect in terms of the arguments set out above. The key reason for this is that it does not completely reflect the probability with which people would have seen the different polling results in their day-to-day lives. In most cases, my experiment implicitly treats polls in which Donald Trump was in the lead as just as likely as those in which Joe Biden was in the lead, for example. Even though both were true polling results, and I conducted the experiment in states where it was most likely that respondents could see polls suggesting either candidate was in the lead, a more externally valid design would weight the probability with which respondents saw each of them in accordance with how common these outcomes were in real polling data. This could have a meaningful impact on the estimation of effects, as noted above (de la Cuesta, Egami, and Imai 2021). So while Chapter 3 nonetheless takes a major step in the direction of using true treatments, it does not avoid the problems set out above altogether.

Even a fully-fledged version of the true treatments approach might be criticised on the grounds that it does not do what experiments are supposed to do. Essentially, experiments allow us to answer counterfactual questions: what would happen if we changed X to Y? Keeping X within the range of values it takes within the real world, currently, misses the potential to change it to Y and observe this counterfactual. This is likely part of the reason why researchers do not use true treatments more often. Arguably though, this is not quite right. The types of experiment I have considered here, including those on the bandwagon effect, are *informational*. The counterfactual is about exposure to information. What would happen if people received information suggesting X? True treatments facilitate answers to this question while only considering the real Xs that the information could suggest. This avoids the ethical and inferential problems that come with making up convenient falsehoods.

Arguably, it is not because experiments allow researchers to deceive respondents, and thereby estimate things that they cannot observe in the real world, that an experimental approach should be taken. This short-cut to causal inference compromises research ethics *and* validity. Rather, scholars should opt for an experimental approach because it allows them to estimate things they can observe in the real world but from which it is impractical to draw casual inferences. The true treatment approach still allows for this, but simply does so without telling lies. It still allows for random assignment, and thereby for estimates of causal effects of information, but it simply depends on this information being *true*.

Conclusion

The purpose of this essay is mostly reflective. It orders the considerations I had to make in designing large parts of the research done in this thesis, to make sure I did my work in justifiable, ethical ways. These concerns are particularly salient given new APSA guidelines on research ethics, which recognise that 'misinformation' matters as a way in which researchers deceive participants. The essay has noted that when conducting experimental research with made-up treatment information, context matters. Context alters whether this invention is a lie. Lying, in turn, matters, because it is morally wrong to at least some degree. And crucially, balancing this cost of moral wrong against the benefits of research is not as straightforward as it seems, because these benefits are compromised to some extent by the made-up treatment information. I explained how my experimental designs – fleshed out in Chapters 3 and 4 – represent different ways of navigating these challenges.

These reflections also have the potential to bring increased attention to the idea of 'misinformation' as a form of deception in experimental research. As noted from the outset, debates about deception have mostly focused on whether researchers mislead people about the purpose of an experiment, or the way in which it is being conducted, as a specific type of false information conveyed to participants. My arguments demonstrate that this is not the only way in which researchers can lie to respondents that is potentially problematic, substantiating APSA's recent recognition of this. Discussions of experimental ethics need to recognise the fact that researchers routinely lie to respondents through false *treatment* information.

To this end, future work on the ethics of deception in experiments could engage in more depth with the philosophical points raised above. For example, I have primarily based my argument around 'lying' and used this term more or less interchangeably with 'deception' and 'misinformation.' But prominent discussions of what it means to lie argue that deception does not even come into it (Carson 2006; Sneddon 2021), and 'misinformation' appears to be a more neutral term than either of the others. Alongside this, the ways in which I navigated the problem of lying in the experiments used in this thesis demonstrated various grey areas where my arguments are weaker, but still seem relevant. This raises the question of whether and how we can conceive of experimental misinformation in different types, or on a spectrum. Conceptual analysis of this kind could bring greater nuance to debates about research ethics than I have offered here, or than is currently available in political science guidelines.

Appendix B

B.1 Power Analysis

Figure B.1 shows the probability of detecting statistically significant effects, by effect size, at a range of sample sizes, in the experiment in Chapter 3. This is based on 100 simulations at each of 500 different randomly generated combinations of effect sizes. For example, one simulation might have generated a dynamic bandwagon effect of 5% (0.05), a static bandwagon effect of 2% (0.02) and an emphasis effect of 1% (0.01). A random sample of voters is then generated, and their responses are randomised with these probabilities, 100 times. Power is defined by the proportion of those times that each of these effects is statistically significant. It is higher for larger effects and for larger sample sizes. The simulation-based approach is more suitable to a design like this that has a lot of moving parts.

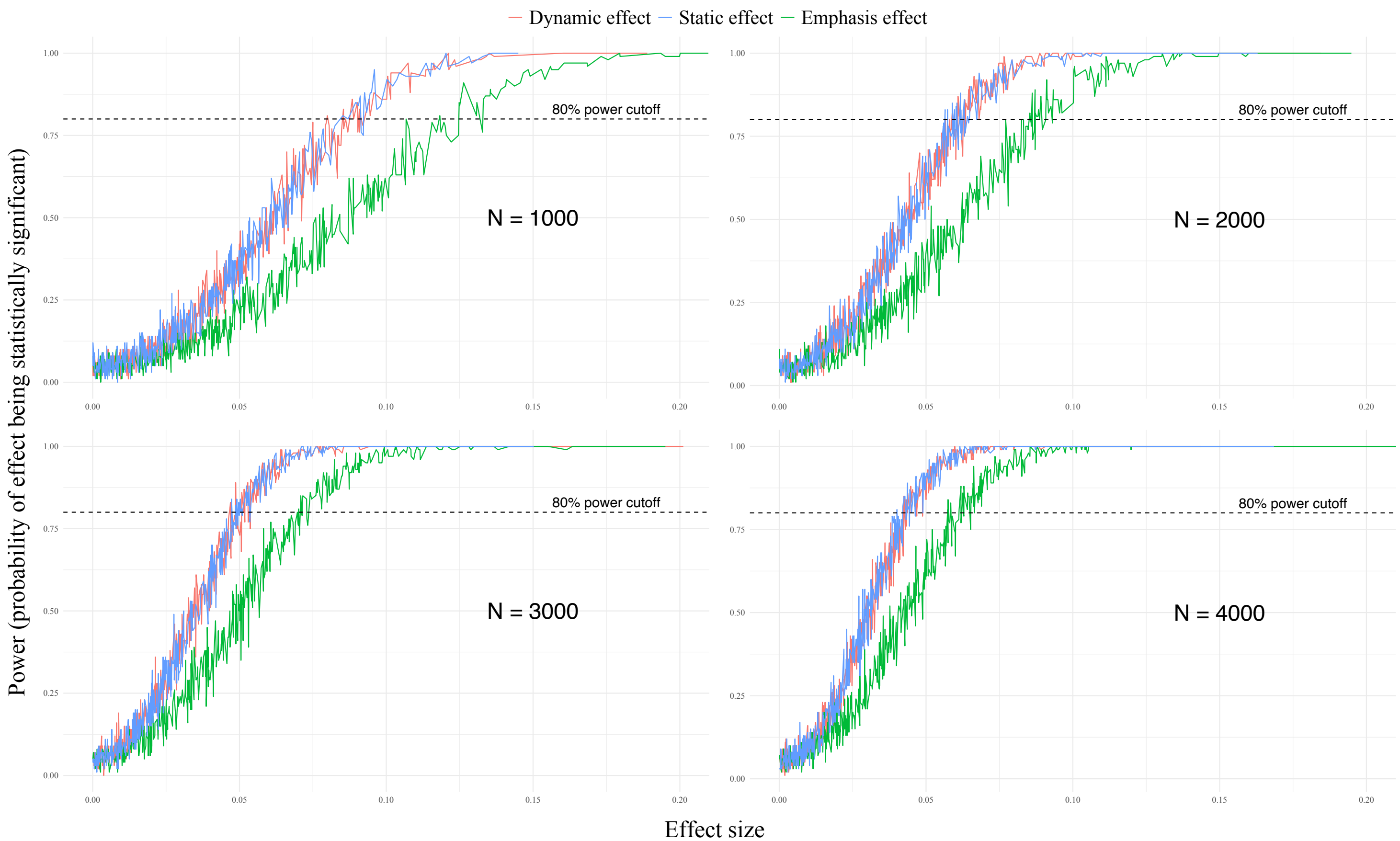


Figure B.1: Probability of detecting statistically significant effect, by absolute effect size and type of effect. Based on 100 simulations at each of 500 different combinations of random effect sizes.

The standard cut-off for sufficient power in experimental design is 80%. The results show that 80% power is only achieved at (static or dynamic) bandwagon effects of approximately 8% (0.08) in a sample size of 1000. With 2000 respondents, an effect of approximately 6% has sufficient power. This reduces to 5% in a sample of 3000, and to 4% in a sample of 4000. Note that it is more difficult to detect emphasis effects because there are four different levels to this feature in the experimental design – but these effects are not of primary interest anyway. Based on past meta-analysis that shows bandwagon effects on average are of only a few percentage points (Hardmeier 2008), I targeted a sample of at least 4000 voters across the four states where I conducted the study. I also decided not to investigate interaction effects between the design features (for example, interacting emphasis with popularity information) because the analysis shows these would be substantially under-powered: Gelman, Hill and Vehtari (2021, 301) explain that 'you need 4 times the sample size to estimate an interaction that is the same size as the main effect.'

Significant rates of non-response meant that I fell short of this number. The final sample used in the analyses in Chapter 3 is 2,661. This implies that the results have sufficient power for effects of approximately 5-6%. So if bandwagon effects were as large as some studies suggest (Cook and Welch 1940), then we would expect my data to detect them.

B.2 Randomisation Check

Figure B.2 shows the predicted probability, and 95% confidence interval, of receiving static or dynamic popularity information that is supportive of Biden in each state, by a series of respondent characteristics: gender, age, education and past vote choice. The dashed horizontal lines in each case denote the probability that would be expected under perfect randomisation. This is nearly always 0.5, except in the case of Texas, where it is 0.66 for dynamic information and 0.33 for static information, because there was no condition in which Joe Biden was in the lead (static) and Trump was gaining ground (dynamic). None of the characteristics has a significant association with whether or not the person received static or dynamic popularity information that was positive for Biden. In other words, randomisation was successful.

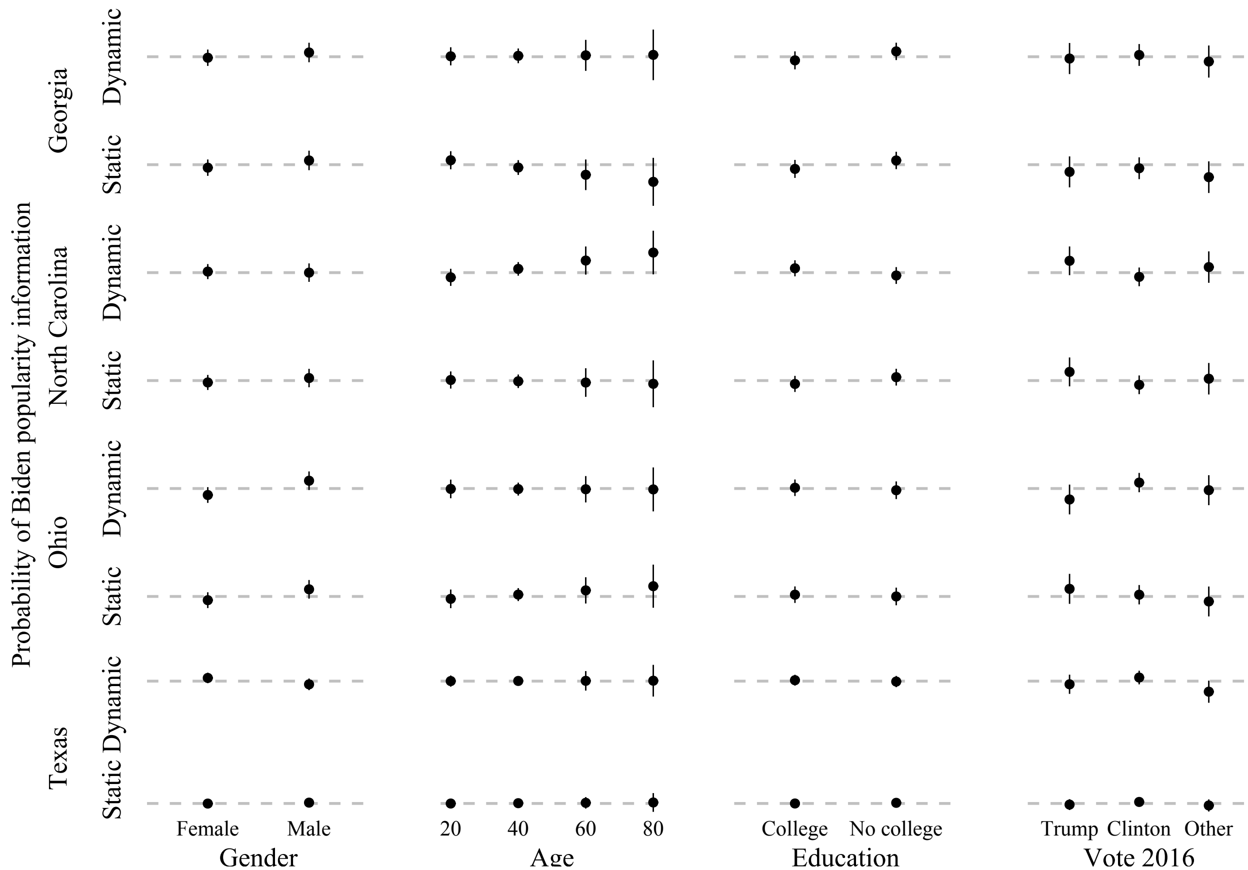


Figure B.2: Randomisation check: Chapter 3. Predicted probabilities of receiving Biden static/dynamic popularity information, rather than Trump in each state, by gender, age, education, and previous election vote choice. Dashed line denotes expected probability under random assignment. None of the characteristics are significantly associated with treatment status, indicating randomisation was successful.

B.3 Poll Sources

Table B.1: Full selection of polls from which results were taken in each state.

Result	Recent poll source	Previous poll source
Georgia		
Biden 47% (-) - 46% (+3) Trump	https://www.aarp.org/ content/dam/aarp/research/ surveys_statistics/politics/2020/2020- election -battleground-states-senate-georgia- annotated- questionnaire.doi.10.26419- 2Fres.00401.016.pdf	https://electjon.com/wp- content/uploads/2020/07/2020.07.22- Ossoff-poll-memo. pdf?__cf_chl_jschl_tk__= b53f1584e6c6fee 058a8570996c235 066d22f7ed-1623959363-0- AWGLwuec5x qqY1mlCHjSY0P HMPZ88rbEwgNEXqy UdVTeIiRu XsPvnuTLXd WsW5J2ojOu DVRWdQZn e0qTPXuTsjv _ZuzyjMaIPy0-NW_1t wnr3IuXW1k _bsIhB40 QaqgJJ3hRa tKbbzHbRsiaAYF QoYS1CrT7ypAb BntX-BaHTDn4OWd BzW44kQX 3L1aQf3nQfpCHY 7C44x6MKcVaPzA AdgqXfr 7iOQeSX96J1t1 CvlvKxOVygMUtz 5UabSSSJvIgiQx RnoeWOWeYk7cdY QvOb2yEv4mE ySs9mbcwE8 GnuJG2x-M- Yhu0wcsZi -4HM4mwSZi 9EXJbUANDxu uRK05qT340 4BDsWxANt 8ZeOl5Ysl2 47qJ9hBp0U fA_1mV8iYDLQ ENj5HxdikkWz 6dHWuLBEe9UJvTb Ztrp1RtZva8 xf9EWKAgdF8 XRnZA6k3tqsbu ys-wS3slFnnGNBfy tigiLTV-7suqvd 7Pa2BfP_PX40vHRt_ j7kwQag9PIsavMJ 1X4TPrlaLqNKKQ 2PcoQYcHnkO0

Table B.1: Full selection of polls from which results were taken in each state. *(continued)*

Result	Recent poll source	Previous poll source
Biden 47% (+3) - 46% (+1) Trump	https://www.aarp.org/content/dam/aarp/research/surveys_statistics/politics/2020/2020-election-battleground-states-senate-georgia-annotated-questionnaire.doi.10.26419-2Fres.00401.016.pdf	https://s3.documentcloud.org/documents/ 6880530/20426-GHRT-GA-Toplines.pdf
Trump 47% (-1) - 46% (+5) Biden	https://www.cbsnews.com/news/trump-biden-opinion-poll-georgia-north-carolina-supreme-court-09-27-2020/	https://s3.documentcloud.org/documents/ 7072527/Landmark-Poll-Georgia-Presidential-Election-8-31.pdf
Trump 47% (+3) - 46% (-) Biden	https://www.cbsnews.com/news/trump-biden-opinion-poll-georgia-north-carolina-supreme-court-09-27-2020/	https://www.surveyusa.com/ client/PollReport.aspx?g=4b9009c4-ef1d-4774-a211-0c4402738ec8
North Carolina		
Biden 48% (+4) - 46% (+1) Trump	https://www.cbsnews.com/news/trump-biden-opinion-poll-georgia-north-carolina-supreme-court-09-27-2020/	https://www.nccivitas.org/ polling/trump-biden-dead-heat/
Biden 48% (+1) - 46% (+3) Trump	https://www.cbsnews.com/news/trump-biden-opinion-poll-georgia-north-carolina-supreme-court-09-27-2020/	https://www.suffolk.edu/-/media/suffolk/documents/academics/research-at-suffolk/suprc/polls/other-states/2020/9_17_2020_final_marginals_pdftxt.pdf
Trump 49% (-) - 48% (+2) Biden	https://www.uml.edu/docs/2020-NC-Sept-Topline_tcm18-330590.pdf	https://morningconsult.com/form/july-presidential-battleground-state-polling/
Trump 49% (+3) - 48% (-) Biden	https://www.uml.edu/docs/2020-NC-Sept-Topline_tcm18-330590.pdf	https://www.monmouth.edu/polling-institute/documents/monmouthpoll_nc_090320.pdf/
Ohio		

Table B.1: Full selection of polls from which results were taken in each state. *(continued)*

Result	**Recent poll source**	**Previous poll source**
Biden 48% (+3) - 46% (-) Trump	https://www.progressivepolicy.org/ wp-content/uploads/2020/09/ ALG.PA_.OH_.energy-poll.09.30.20- .pdf	https://drive.google.com/ file/d/1 Vbmv7EXnwqeubTuvB YNiBvq565sOxi_9/view
Biden 48% (-1) - 46% (+2) Trump	https://www.progressivepolicy.org/ wp-content/uploads/2020/09/ ALG.PA_.OH_.energy-poll.09.30.20- .pdf	https://static.foxnews.com/ foxnews.com/content/ uploads/2020/09/ Fox_September-20-23-2020_Complete_ Ohio_Topline_September-24- Release.pdf
Trump 49% (+3) - 47% (-1) Biden	https://hrc-prod-requests.s3 -us-west-2.amazonaws.com/ ME-13067-HRC-Issues-10-State- 1.pdf?mtime=2020 1002132812&focal=none	https://www.progressivepolicy.org/ wp-content/uploads/2020/09/ ALG.PA_.OH_.energy-poll.09.30.20- .pdf
Trump 47% (-1) - 46% (+2) Biden	https://hrc-prod-requests.s3 -us-west-2.amazonaws.com/ ME-13067-HRC-Issues-10-State- 1.pdf?mtime=2020 1002132812&focal=none	https://morningconsult.com/ 2020/09/09/trump-biden- race-tightens-2020-polling/
Texas		
Biden 48% (+3) - 44% (-) Trump	https://projects.fivethirtyeight.com/ polls/20200921_TX.pdf	https://poll.qu.edu/poll-results/
Trump 48% (-) - 46% (+5) Biden	https://drive.google.com/ file/d/1ZsMPYa9d2L2vg4YYk- DFcwlXosTMmMYP/view	https://www.txhpf.org/ wp-content/uploads/ 2020/08/THPFFinalAug17.pdf
Trump 48% (+3) - 46% (-1) Biden	https://drive.google.com/ file/d/1ZsMPYa9d2L2vg4YYk- DFcwlXosTMmMYP/view	https://www.chrystafortexas.com/ media/press/TX_RR_ Commis- sion_Polling_Memo_F08.14.20.pdf

B.4 Emphasis Sentences

Table B.2: Range of possible emphasis sentences. Ohio used as example.

Level	Sentence
Biden static	
Yes static	With the lead in this poll, it looks like Biden could be set to win in Ohio.
No static	Trailing in this poll, it looks like Biden could be set to lose in Ohio.
Biden dynamic	
Yes dynamic	Biden appears to be winning over an increasing number of voters in Ohio as the campaign goes on.
No dynamic	Biden appears not to be winning over many more voters in Ohio as the campaign goes on.
Trump static	
Yes static	With the lead in this poll, it looks like Trump could be set to win in Ohio.
No static	Trailing in this poll, it looks like Trump could be set to lose in Ohio.
Trump dynamic	
Yes dynamic	Trump appears to be winning over an increasing number of voters in Ohio as the campaign goes on.
No dynamic	Trump appears not to be winning over many more voters in Ohio as the campaign goes on.

B.5 Robustness Check: Texas

As explained in Chapter 3, it was impossible for respondents to receive information suggesting that Biden was in the lead, but Trump had momentum, in Texas. In order to verify that the effect observed in the main text is not an artifact of this,

Figure B.3 conducts the Joe Biden conversion analysis on a reduced dataset in which respondents from Texas are removed. The main finding from the above is reproduced here: people are less likely to switch to vote for Joe Biden when he is in the lead.

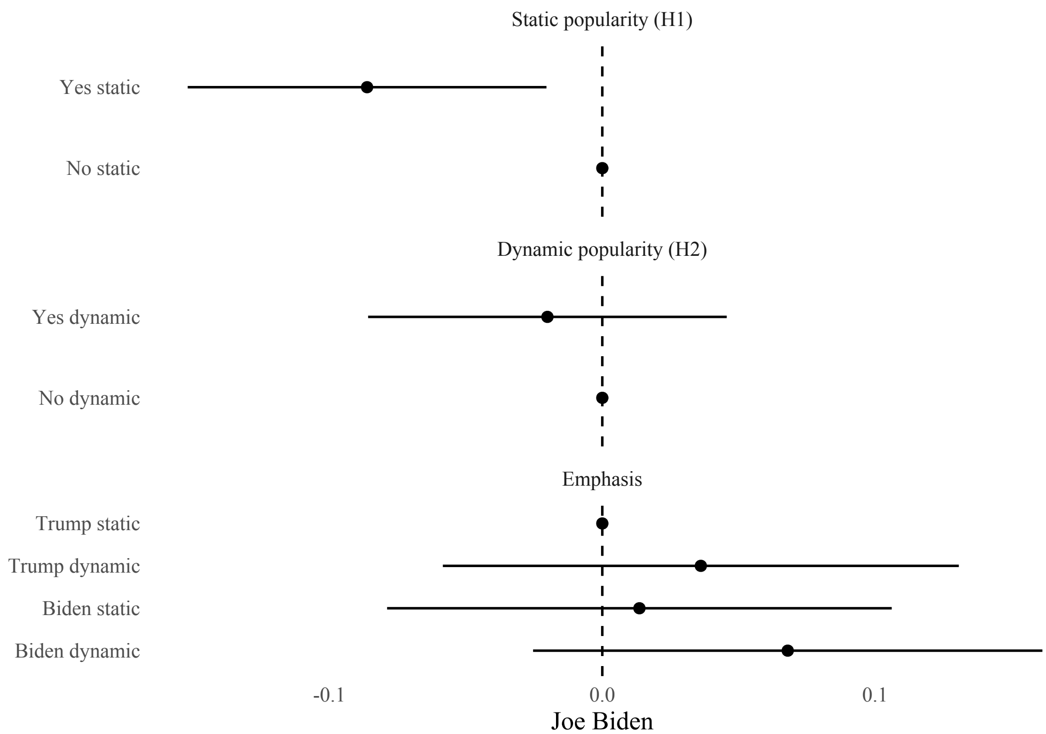


Figure B.3: Effects of popularity information on voting for Biden (left) and Trump (right), for those who do not identify with Democrats (left) or Republicans (right), in all states except Texas.

B.6 Robustness check: Attention

In order to assess whether respondents were paying attention to the polling information they received in the experiment in Chapter 3, I asked them, immediately after they indicated their vote choice, which candidate was ahead in the poll they just read about. The overwhelming majority of participants (2237, or 84%) answered this question correctly, suggesting that on the whole, attention levels were adequate. Nonetheless, in order to validate that the effect on Joe Biden's vote share observed in Chapter 3 is not the result of the behaviour of inattentive respondents, Figure B.4 reports the results of a conversion model that excludes all those who failed the attention check. The main effect discussed in the text is robust to this specification.

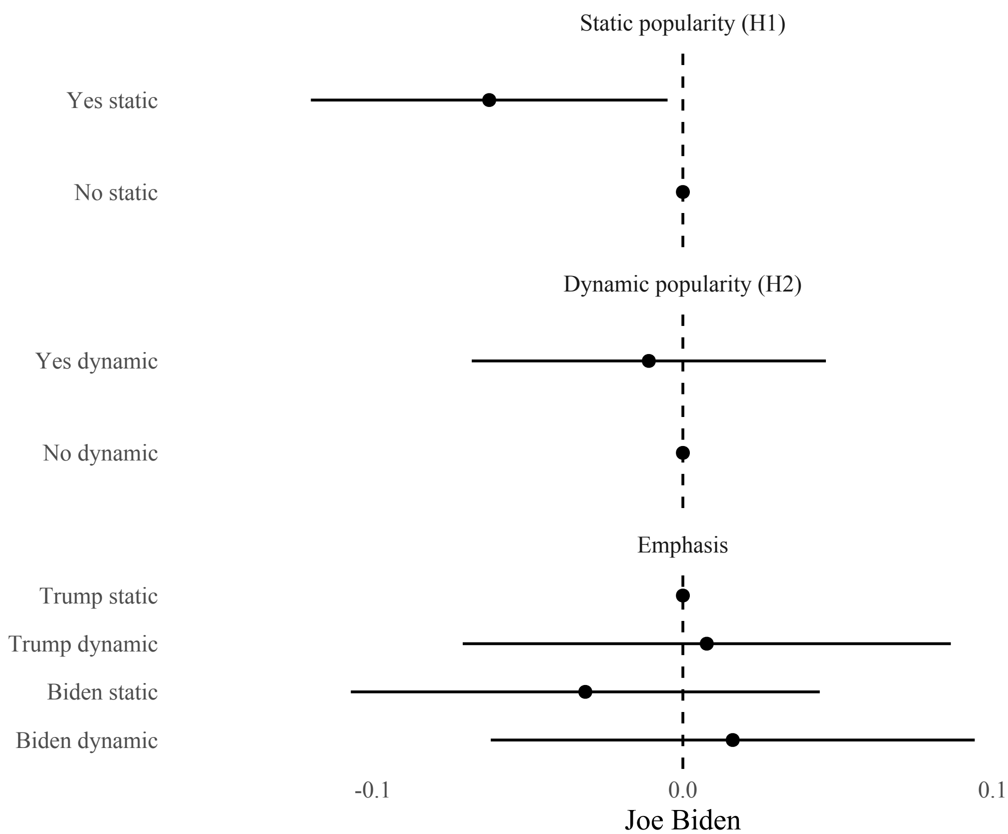


Figure B.4: Effects of popularity information on voting for Biden, among non-Democrats, partially pooled with varying intercepts by state, excluding those who failed the attention check.

B.7 Robustness Check: Already Voted

Due to the timing of the experiment, it was possible that participants had already voted by post in the 2020 election. This could affect the results in multiple ways. First, the real vote choice of those who had already voted was already set in stone, so may be less strongly affected by treatment. Second, in contrast, these people may have behaved in arbitrary ways, taking the experiment less seriously and being overly responsive to treatment. Thirdly, given that postal voters were more likely to vote for Joe Biden, the model might overestimate his support when including these people. However, the overwhelming majority of respondents had not yet voted (2484, or 93%). Figure B.5 shows that the main effect discussed in Chapter 3 is robust to the exclusion of those few participants who voted early.

B.8 Robustness Check: Defining Conversion by 2016 Vote Choice

In Chapter 3, I model conversion using respondents' previously reported party ID from their Prolific profiles. Figure B.6 reports the results of modelling voting for Joe Biden only among those who did not vote for Clinton in 2016, treating these people as potential converters, rather than those who do not identify with the Democrats. The main effect discussed in Chapter 3 is no longer statistically significant at the 5% level in this specification, though it is significant at the 10% level. This likely reflects the fact, as noted in the main text, that this alternative sample includes many people who were too young to vote in 2016, but who

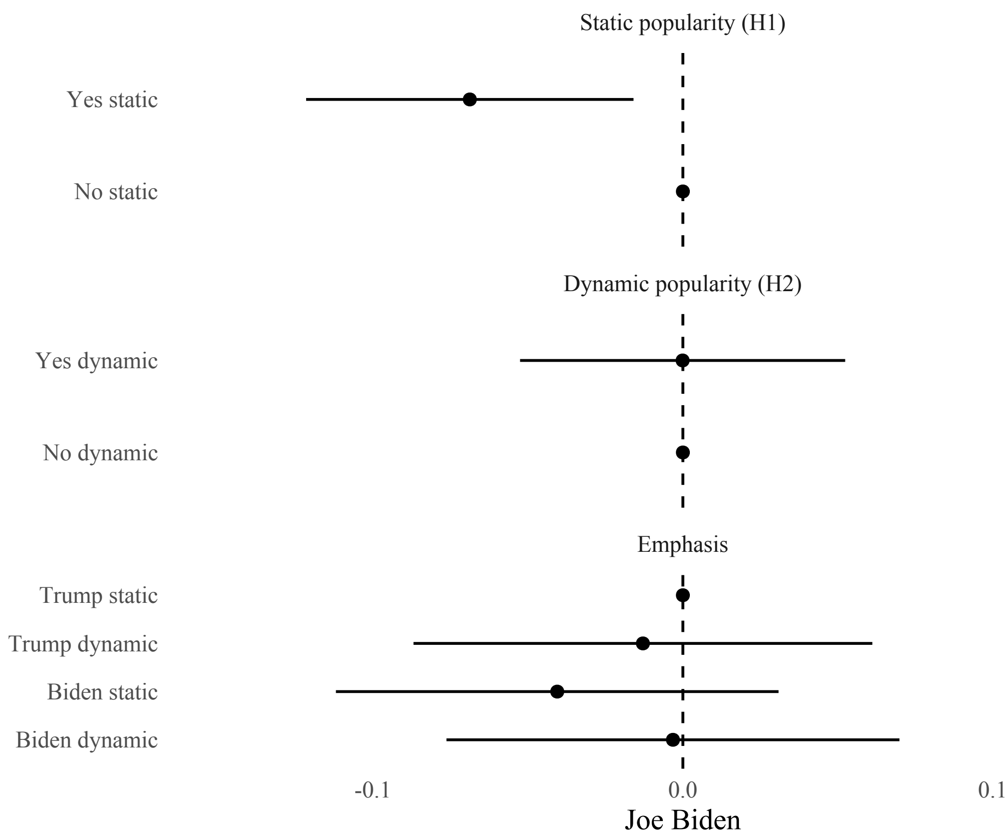


Figure B.5: Effects of popularity information on voting for Biden, among non-Democrats, partially pooled with varying intercepts by state, excluding those who had already voted before taking the experiment.

would have voted Democrat if they could. Many of these people will have voted Biden in the experiment, regardless of the treatment information they received, dampening the observed conversion effect of this information – because these (Democrat) voters were not actually available for conversion to voting Democrat. This complication is, indeed, why I operationalised conversion through the party ID people reported in their Prolific profiles.

Another way to account for this is to only consider people who voted for Donald

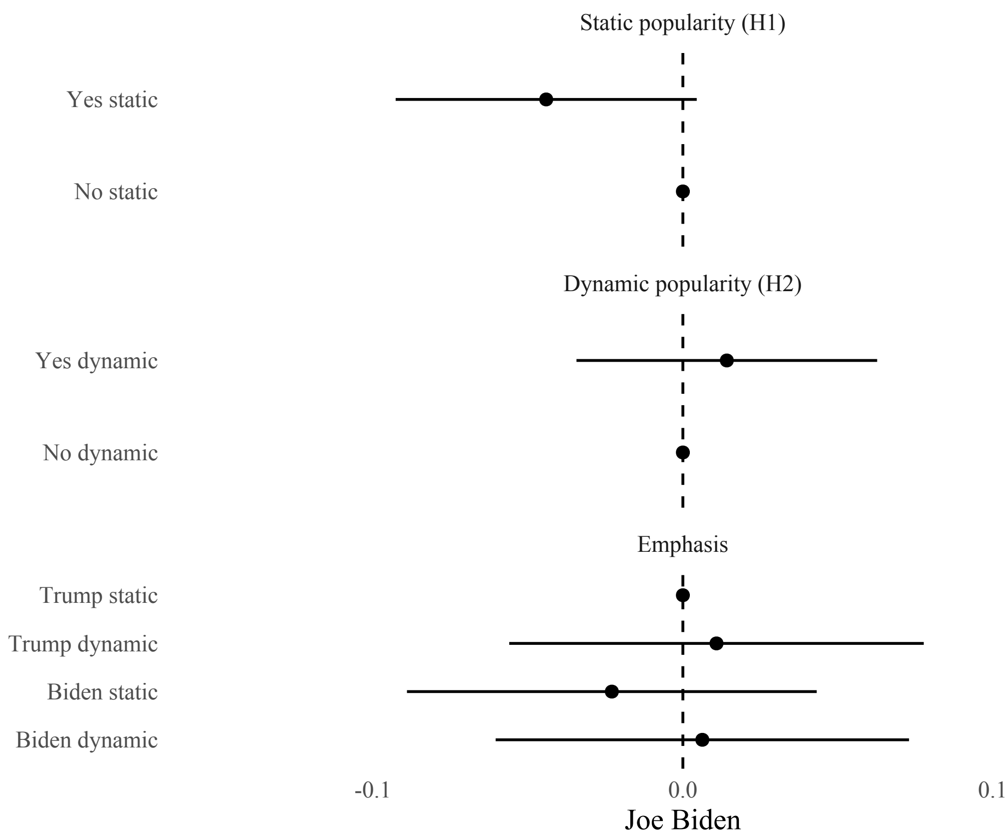


Figure B.6: Effects of popularity information on voting for Biden, among those who did not vote for Hillary Clinton in 2016, partially pooled with varying intercepts by state.

Trump in 2016. Figure B.7 shows that the main conversion effect discussed in Chapter 3 is reproduced at a statistically significant level among these people – even in a very reduced sample of only 438 respondents. People who voted for Trump in 2016 were less likely to switch to vote for Biden when he was in the lead in a poll in their state than if he was behind.

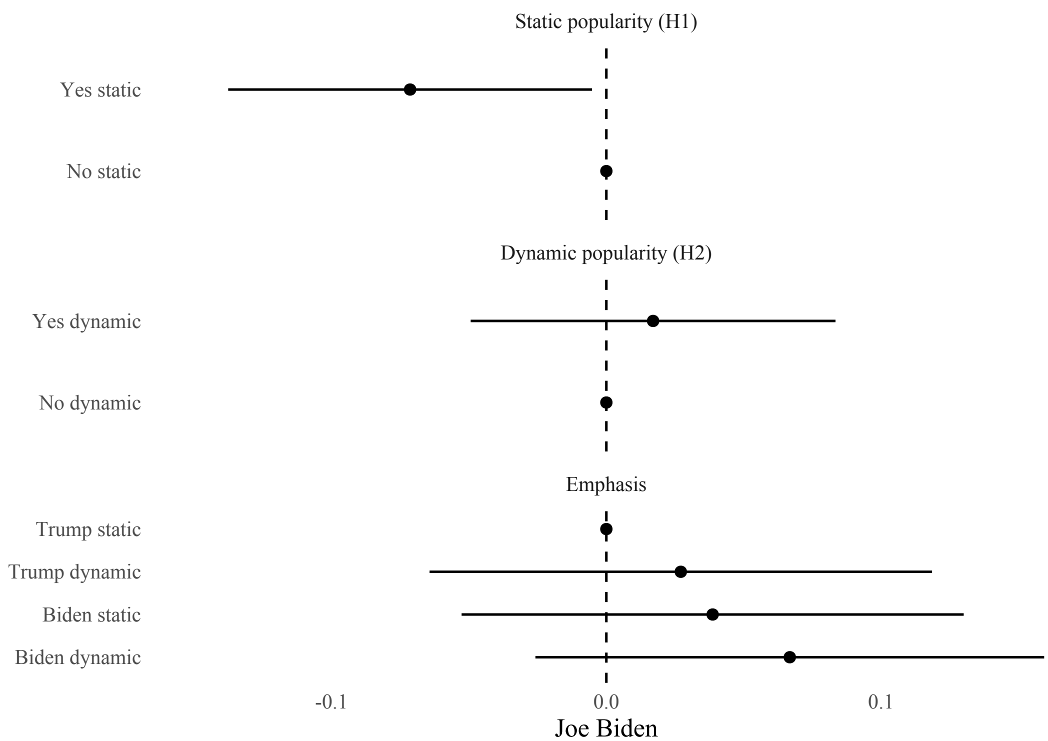


Figure B.7: Effects of popularity information on voting for Biden, among those who voted for Donald Trump in 2016, partially pooled with varying intercepts by state.

B.9 Experiment Screenshots

Please read the short text below and indicate who you would like to vote for, or whether you would like to abstain, by selecting one of the options.

Presidential election 2020: what the polls say in Ohio

A recent poll in Ohio put Biden's support at 47%, a slight decrease of 1 point since a previous poll. Trump's vote share has increased by 3 points, to 49% over the same period. Trailing in this poll, it looks like Biden could be set to lose in Ohio.

Having read this information, who would you like to vote for? Please **select one option** below, and **press the arrow to continue**.

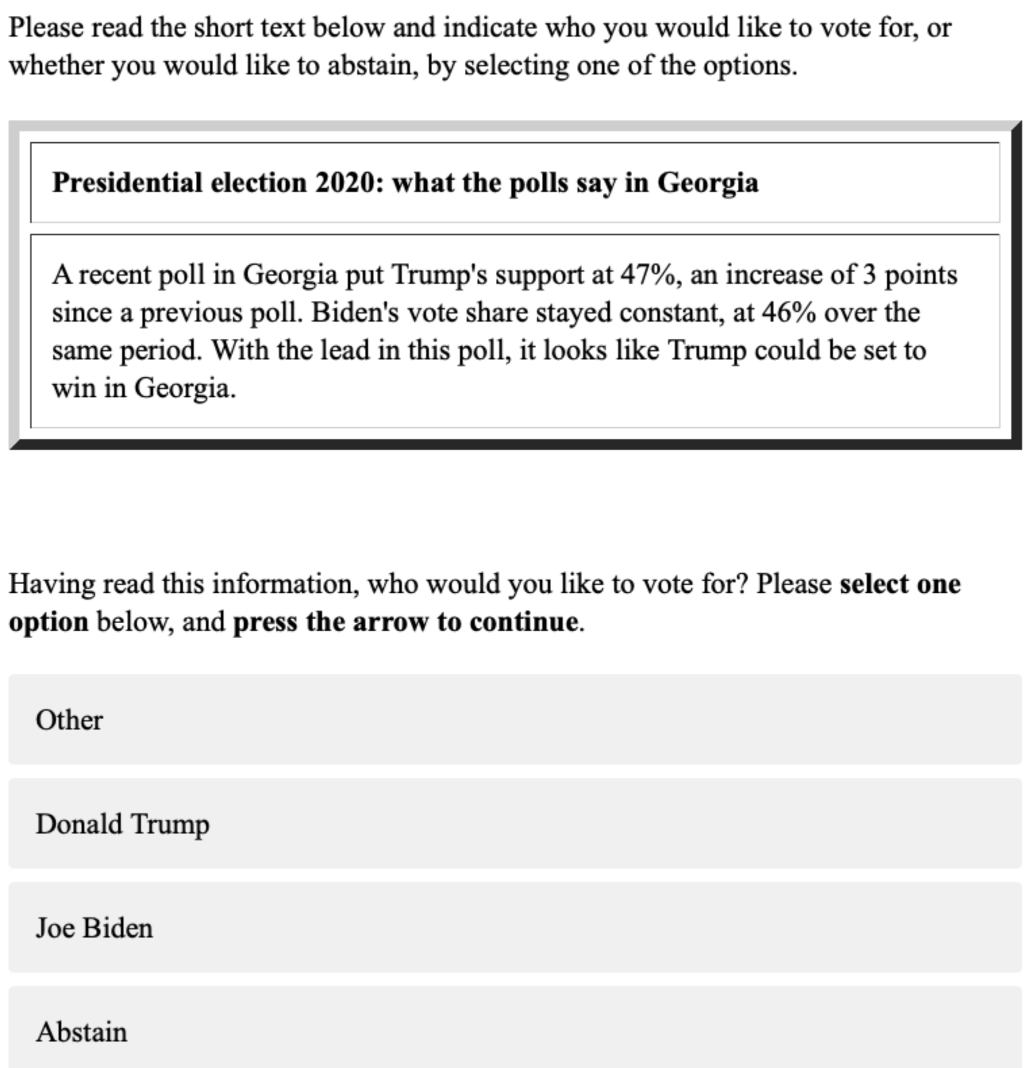
Joe Biden

Abstain

Other

Donald Trump

Figure B.8: Example experimental task from Ohio.



Please read the short text below and indicate who you would like to vote for, or whether you would like to abstain, by selecting one of the options.

Presidential election 2020: what the polls say in Georgia

A recent poll in Georgia put Trump's support at 47%, an increase of 3 points since a previous poll. Biden's vote share stayed constant, at 46% over the same period. With the lead in this poll, it looks like Trump could be set to win in Georgia.

Having read this information, who would you like to vote for? Please **select one option** below, and **press the arrow to continue**.

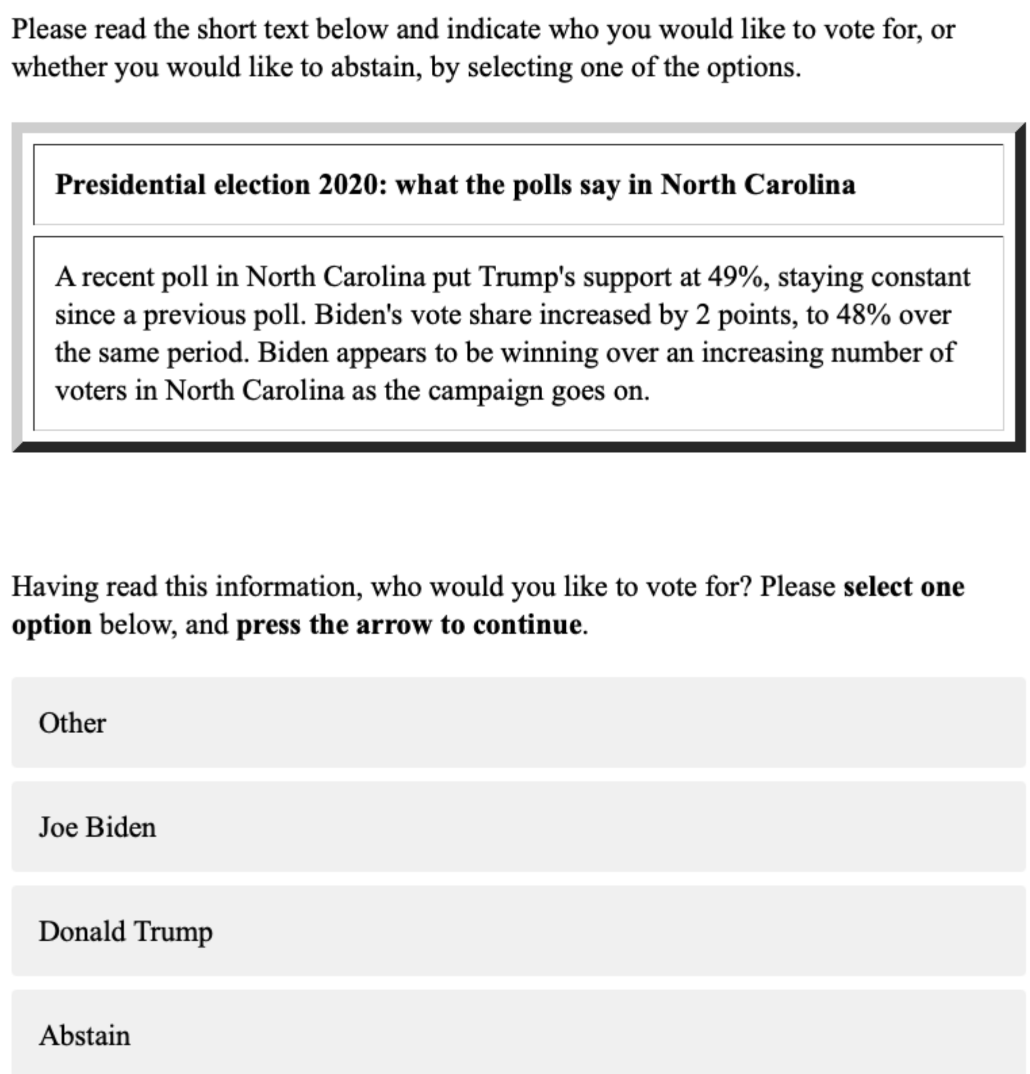
Other

Donald Trump

Joe Biden

Abstain

Figure B.9: Example experimental task from Georgia.



Please read the short text below and indicate who you would like to vote for, or whether you would like to abstain, by selecting one of the options.

Presidential election 2020: what the polls say in North Carolina

A recent poll in North Carolina put Trump's support at 49%, staying constant since a previous poll. Biden's vote share increased by 2 points, to 48% over the same period. Biden appears to be winning over an increasing number of voters in North Carolina as the campaign goes on.

Having read this information, who would you like to vote for? Please **select one option** below, and **press the arrow to continue**.

Other

Joe Biden

Donald Trump

Abstain

Figure B.10: Example experimental task from North Carolina.

Please read the short text below and indicate who you would like to vote for, or whether you would like to abstain, by selecting one of the options.

Presidential election 2020: what the polls say in Texas

A recent poll in Texas put Biden's support at 46%, an increase of 5 points since a previous poll. Trump's vote share has stayed constant, at 48% over the same period. Biden appears to be winning over an increasing number of voters in Texas as the campaign goes on.

Having read this information, who would you like to vote for? Please **select one option** below, and **press the arrow to continue**.

Other

Abstain

Joe Biden

Donald Trump

Figure B.11: Example experimental task from Texas.

Appendix C

C.1 Power Analysis – Study One

Figure C.1 shows the probability of detecting statistically significant effects (power) for a range of effect sizes, at a sample size of 1500, based on the experimental design in Study One, in Chapter 4. This is based on 100 simulations at each of 500 different randomly generated effect sizes. The sample size is based on the fact that the number of respondents for a YouGov omnibus survey, where I knew the experiments would be fielded, is typically just over 1500. I conducted the power analysis in order to assess what size of effect the experiments would have sufficient power to detect, given that they would be included in these surveys. I use a simulation approach because this allows greater flexibility to follow a more realistic data generating process. For example, while a simple calculation could work out, analytically, the sample size needed to detect a given difference in means between treatment and control, this would not factor in the fact that the responses to the experiment are on a 0-10 ordinal scale. The simulation process makes this explicit. The flexibility of this approach also makes it more capable of easily extending to the more complex design of Study Two, as below.

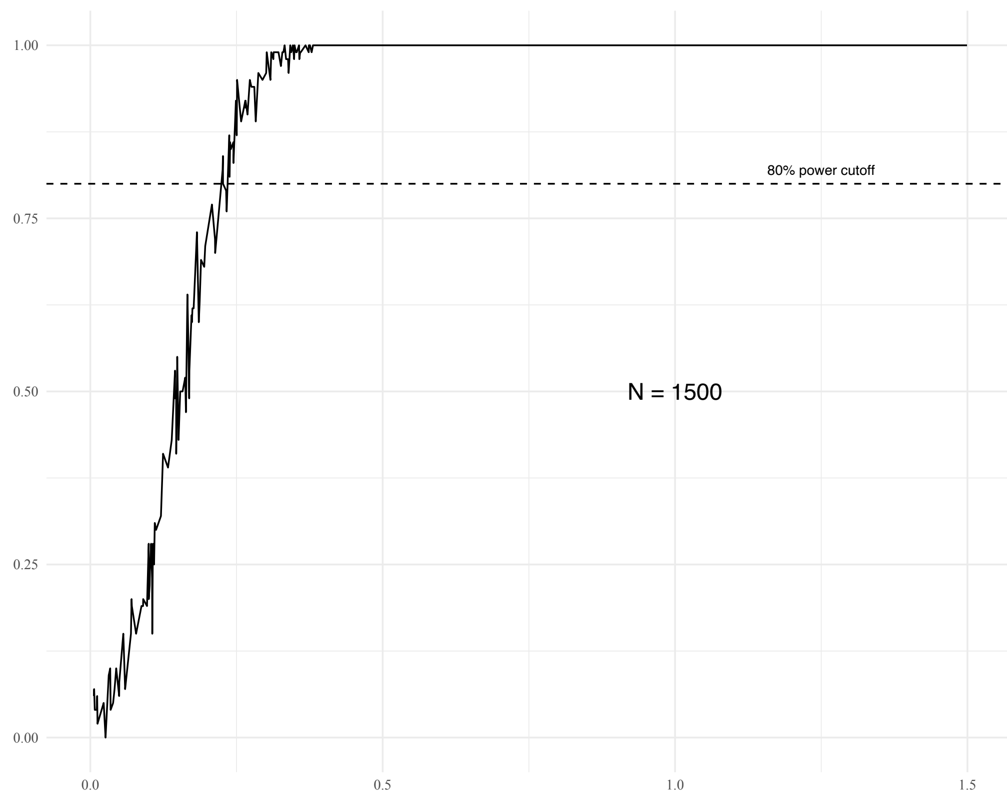


Figure C.1: Probability of detecting statistically significant effect, by absolute effect size, in Study One. Based on 100 simulations at 500 different random effect sizes.

In experimental design, the conventional cut-off for statistical power is 80% (the dashed horizontal line). Figure C.1 shows that, to achieve this level of power with a sample size of 1500, the true underlying effect would need to be of approximately 0.25 or larger – where the line crosses the 80% threshold. This means that the experiments have sufficient power to detect whether the dynamic popularity treatment increases people's expectations by about 1/4 of a point on a 0-10 scale. In fact, the sample sizes I collected were even larger than this, meaning even smaller effects could be detected with sufficient power. The results also showed effects of approximately 0.5 or higher. Figure C.1 shows that, given their sample size, if the underlying effects were of such size, the experiments were almost certain to detect them as statistically significant. In short, the sample sizes in Study One were more than large enough to yield reliable statistical results for effects of any substantively meaningful size.

C.2 Power Analysis – Study Two

Study Two is more complicated both because it also focuses on both two-way and three-way interactions, and because in doing so, it increases the imprecision in the estimation of the main treatment effect. This is necessary in order to detect differences across conditions owing to levels of information, and across parties due to wishful thinking. Figure C.2 repeats the simulated analysis from above but this time, the randomised sample is also split across three different conditions (representing A/B, UK and Canada) and comes from three different parties (representing Conservative, Labour and Other). In order to realistically reflect the numbers of each party's voters that would be present in the experiment, I weighted the randomisation between parties according to the proportions of each in the experiments in Study One, from which I already had data.[1] In this case, I again knew that the experiments would be fielded through YouGov omnibus surveys, but I also knew that the more complex design meant that I would require a larger

[1]This actually made little difference to the results compared to having a uniform split across parties.

sample size in order to replicate the findings of Study One, and certainly in order to detect interaction effects. As such, the sample sizes are based on the assumption that the experiment would feature on two (3000), three (4500) or four (6000) of such surveys.

Firstly, Figure C.2 shows that here a sample of 3000 would be able to reliably replicate the main effect of approximately 0.7 estimated in the +6 experiment of Study One (green line). If average expectations differ across the different conditions (A/B, UK, Canada) by a comparable amount, then this would also be detected (red line). It is plausible that differences across the conditions could be smaller than this though, making a larger sample necessary. Moreover, interaction effects are much harder to detect, as shown by the khaki, blue and pink lines, which all correspond to an effect reported in Chapter 4. Based on this, I decided to go for the largest sample possible, in order to minimise the size of these interaction effects that the experiment would be able to detect. Figure C.2 demonstrates that Study Two is sufficiently powered to detect the following:

- An average treatment effect of approximately 0.4
  - E.g. average expectations in the treatment group (in the A/B condition) are 4.9, versus 4.5 in the control group
- A condition effect of approximately 0.4 (A/B vs UK/Canada)
  - E.g. average expectations (in the control group) in the UK condition are 4.9, versus 4.5 in the A/B condition
- A treatment effect difference, by condition, of approximately 0.6
  - E.g. the difference between the treatment and control group in the UK condition is only 0.2, versus 0.8 in the A/B condition
  - This could happen, e.g., if people incorporate other information into

their expectations in the UK condition, rather than depending as much on the treatment information

- A condition effect difference, by party, of approximately 0.5
  - E.g. the difference between the UK and the A/B condition, among Conservatives, is 0.7, versus 0.2 among Labour supporters
  - This could happen, e.g., if Labour partisans engage in wishful thinking by reporting unduly high expectations for the Labour Party on average
- A treatment effect difference, by condition, by party, of approximately 0.75
  - E.g. the difference between the treatment and control group, in the UK condition, among Conservatives, is 0.1, versus 0.85 in the A/B condition, among Labour supporters
  - This could happen, e.g., if Labour partisans engage in wishful thinking by being more strongly influenced by the treatment information than other groups, specifically in the UK condition

This implies that there could be some small effects that the findings in Chapter 4 rule out as not statistically significant, but which are in fact systematic in the population. However, arguably, to the extent that these effects are of a substantively meaningful size, they are likely to be detected. Indeed, the results reported in Chapter 4 do reveal many significant, complex differences across the experimental conditions, and across partisan groups.

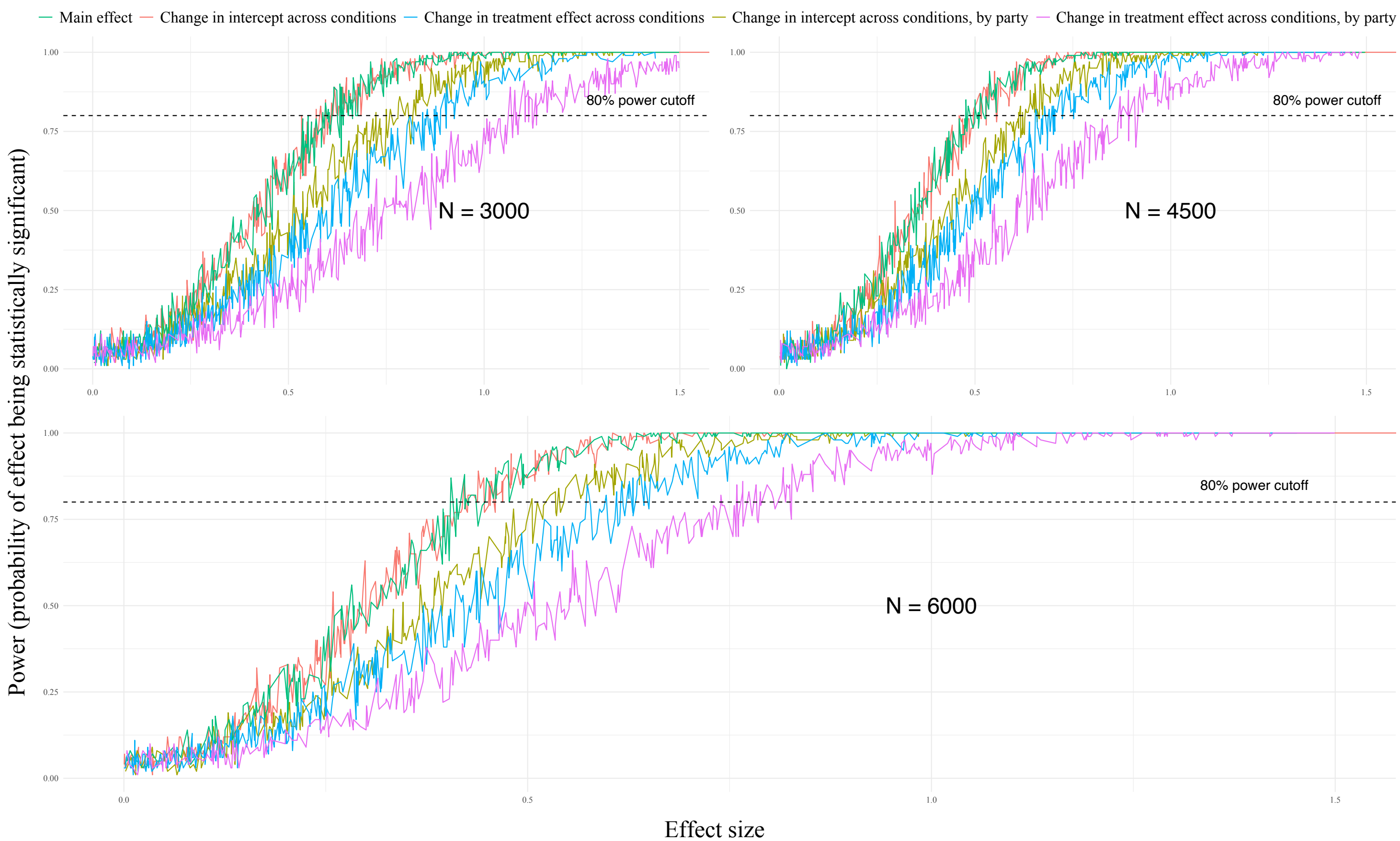


Figure C.2: Probability of detecting statistically significant effect, by absolute effect size and type of effect, in Study Two. Based on 100 simulations at 500 different combinations of random effect sizes.

C.3 Randomisation Check – Study One

Figure C.3 shows that, across all three experiments, there are no systematic differences between the treatment and control groups in terms of gender, age, social grade, past vote choice or Brexit vote choice. These key covariates are balanced. This suggests that the randomisation was effective and should have removed any potential confounding influences on the outcome.

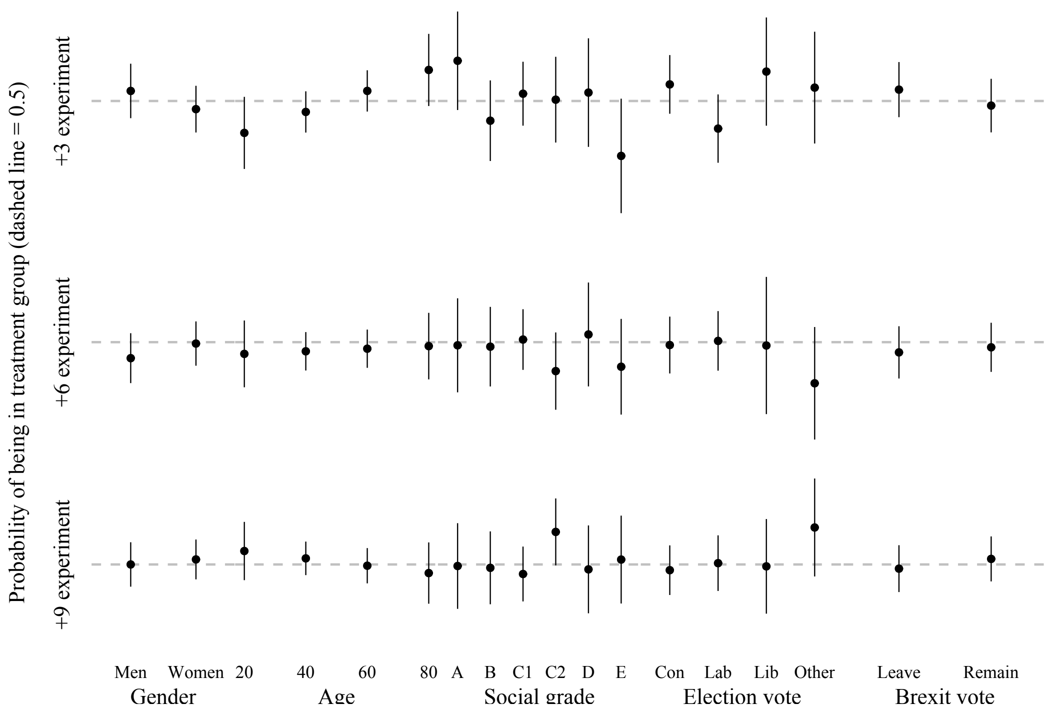


Figure C.3: Randomisation check: Study One, Chapter 4. Predicted probabilities of being in treatment group rather than the control group, by gender, age, social grade, previous election vote choice and Brexit vote choice. Dashed line denotes 0.5 probability. None of the characteristics are significantly associated with treatment status, indicating randomisation was successful.

C.4 Randomisation Check – Study Two

Figure C.4 shows that there are almost no systematic differences between the treatment and control groups in terms of gender, age, social grade, past vote choice or Brexit vote choice. These key covariates are balanced in nearly every case. The exception to this is age in two specific cases. The youngest respondents are slightly more likely to be in the control group in the UK condition, and slightly more likely to be in the treatment group in the A/B condition. The oldest respondents were slightly more likely to be in the treatment group in the UK condition.

In order to establish whether this is likely to have had any substantial impact on the main inferences drawn in Chapter 4, I fit a three-way interaction model equivalent to the one originally presented in Table 4.5 and Figure 4.6, but with an additional control for age. Age is not significantly associated with the expectations reported in Study Two ($\beta =$ -0.002, $p =$ 0.281) when accounting for the interactive effects of treatment, condition, and past vote choice – as I do in the main analyses reported in Study Two. This should demonstrate that the slight deviation from perfect random assignment to treatment by age in two of the conditions is unlikely to have significantly biased the results. Indeed, it could be that this is simply a case of multiple comparisons biasing the statistical tests. That is, in Figure C.4, 54 different coefficients are reported. Three of these are statistically significant at the 95% level. This is approximately equal to 5%, which is precisely the number of tests that would be expected to be significant at this level purely by chance.

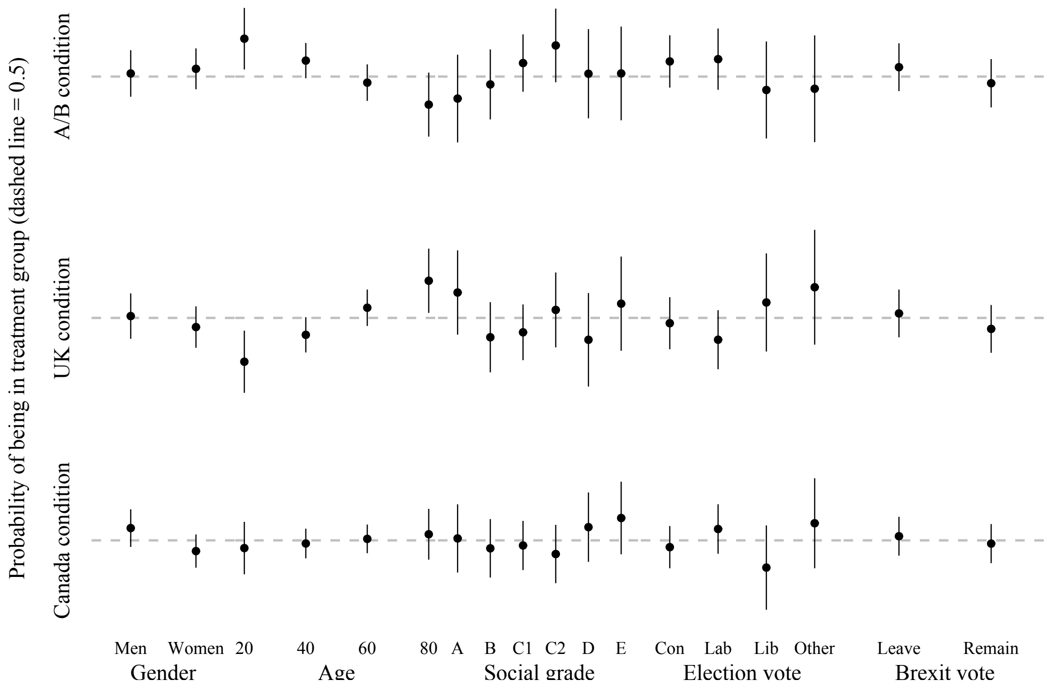


Figure C.4: Randomisation check: Study Two, Chapter 4. Predicted probabilities of being in treatment group rather than the control group, by gender, age, social grade, previous election vote choice and Brexit vote choice. Dashed line denotes 0.5 probability. Randomisation was successful except, in the UK condition, younger respondents were slightly less likely to receive treatment than control, and the opposite for older respondents. In the A/B condition, the youngest respondents were slightly more likely to be in the treatment than in the control group.

C.5 Reduced Sample Model: Study Two

The analyses used to detect wishful thinking in Study Two rely on a reduced sample of only those who provided valid answers about which party they voted for in the last UK general election. In order to verify that the behaviour of this sample does not differ significantly from the full sample used in earlier analyses, Table C.1 reports the results of applying the model used in the first part of Study Two (reported in Table 4.4) to the reduced sample. The results are essentially equivalent.

Table C.1: Linear regression results, reduced sample, Study Two.

Intercept	4.531***
	(0.070)
Treatment	0.849***
	(0.097)
UK condition	−1.301***
	(0.097)
Canada condition	0.255***
	(0.097)
Treatment:UK condition	−0.251*
	(0.137)
Treatment:Canada condition	−0.556***
	(0.136)
Observations	5,207
Adjusted R^2	0.118
Note:	*p<0.1; **p<0.05; ***p<0.01

C.6 Ordinal Regression Models

The effects in both Study One and Study Two in Chapter 4 are measured on a scale that is not truly continuous or 'metric,' but they are modelled using linear regression. According to recent research, this can lead to type I and II errors (saying there is an effect when there is not, and saying there is not an effect when there is), estimates with incorrect signs (e.g. saying an effect is positive when it is negative), and predicted outcomes that lie outside of the measurement scale (e.g. predicting that someone in the treatment group will rate Party B's chances at 12/10) (Liddell and Kruschke 2018). In McElreath's (2020, 380) simpler terms, 'treating ordered categories as continuous measures is not a good idea.'

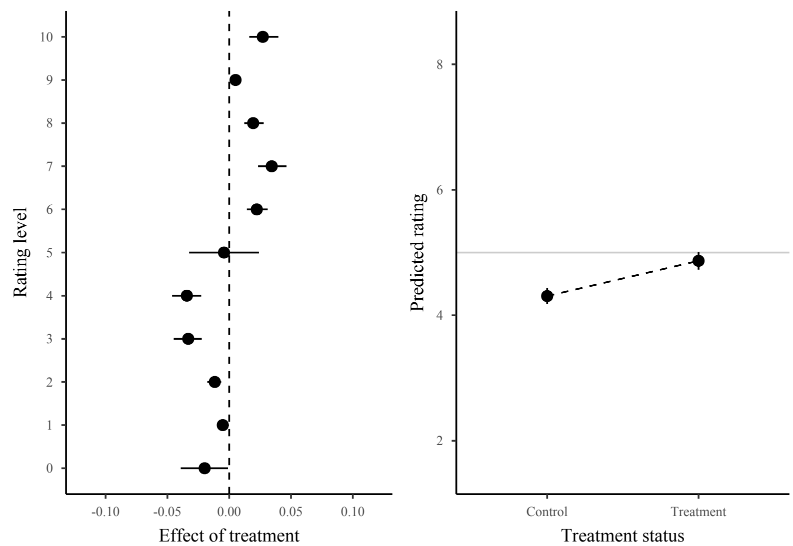


Figure C.5: Summary of ordinal model, Study One, +3 experiment.

These concerns are arguably not entirely applicable here given that my scale contains eleven possible options, which are expressed as numbers. Though not strictly continuous, respondents might essentially treat this outcome as if it were. Some research has indeed shown that the problems above are usually negligible when the outcome scale has seven or more response options (Allen and Seaman 2007; Rhemtulla, Brosseau-Liard, and Savalei 2012). Nonetheless, I also fit ordinal models as a robustness check, following best practices suggested by Bürkner and Vuorre (2019), in order to allay concerns that the linear approach significantly

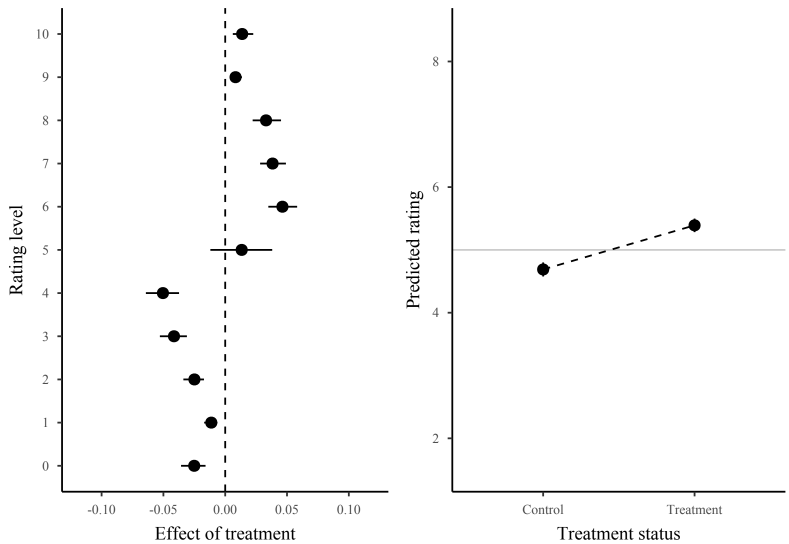


Figure C.6: Summary of ordinal model, Study One, +6 experiment.

misrepresents the estimated effect.

Figures C.5, C.6 and C.7 present the results of these models, for Study One. In each figure, the left-hand plot presents the effect of treatment on each possible level of the outcome. The right-hand plot shows the average predicted response under treatment and control, as in Figure 4.4, derived from these models.[2]

[2] This is derived in the way suggested by Kurz (2020), following an example by McElreath (2020, 386–87).

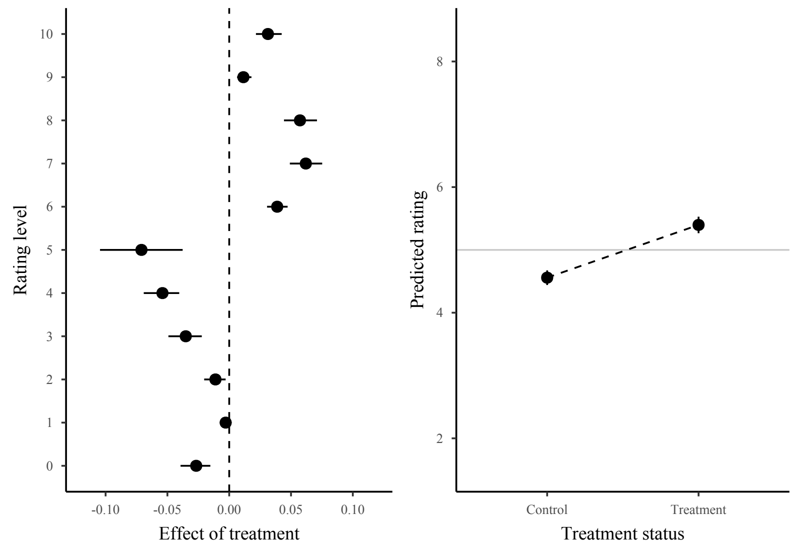


Figure C.7: Summary of ordinal model, Study One, +9 experiment.

Looking first at the plot on the left of Figures C.5, C.6 and C.7, larger values are more likely under treatment, smaller values less likely. There are no rogue values on the scale, at which the effect is in the opposite direction – for example, there is no value above five that is *less* likely under treatment. The effect is larger the stronger the treatment. This can be seen in the way the points and their intervals move further away from zero as the strength of the stimulus increases. In the +9 experiment, the effect is so strong that it even substantially reduces the probability of a 5/10 rating. H_1 and H_2 are again supported. This representation also reveals that this effect varies across the different levels of the outcome, systematically across all three experiments. For example, the values 3, 4, 6, and 7 – that is, those closest to the middle of the scale without being its midpoint – are consistently among the most affected by treatment. Towards the extremes of the scale, the effect tends to be at its smallest. Broadly, this is evidence that Party B's momentum is not convincing people it is definitely going to win, but the fact it is trailing overall is generally not enough to convince them it has no chance either. Rather, knowing Party B has momentum tips the balance slightly more in its favour. The plot on the right of each figure shows that, on average, these effects equal out to predicted responses that are essentially identical to those in the linear regression models above.

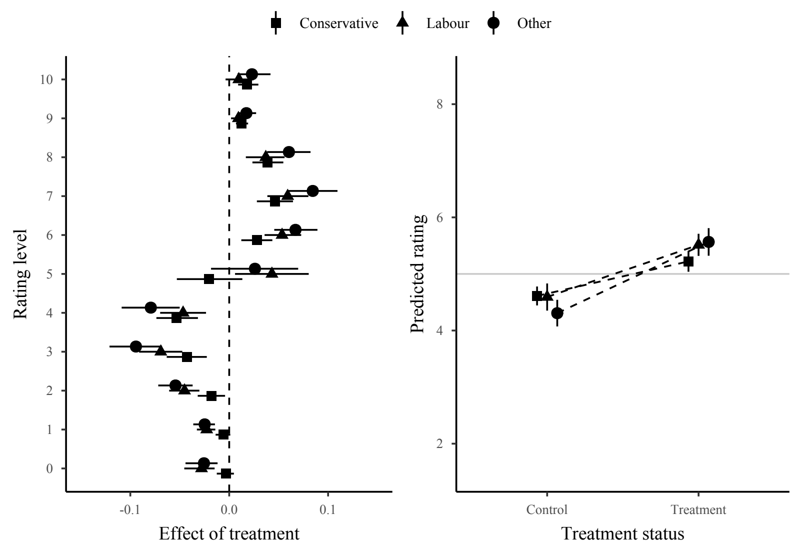


Figure C.8: Summary of ordinal model, Study Two, A/B condition.

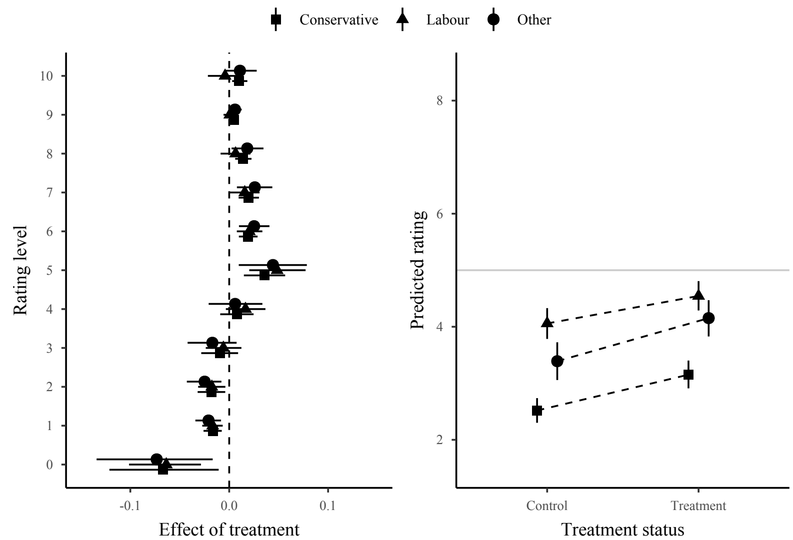


Figure C.9: Summary of ordinal model, Study Two, UK condition.

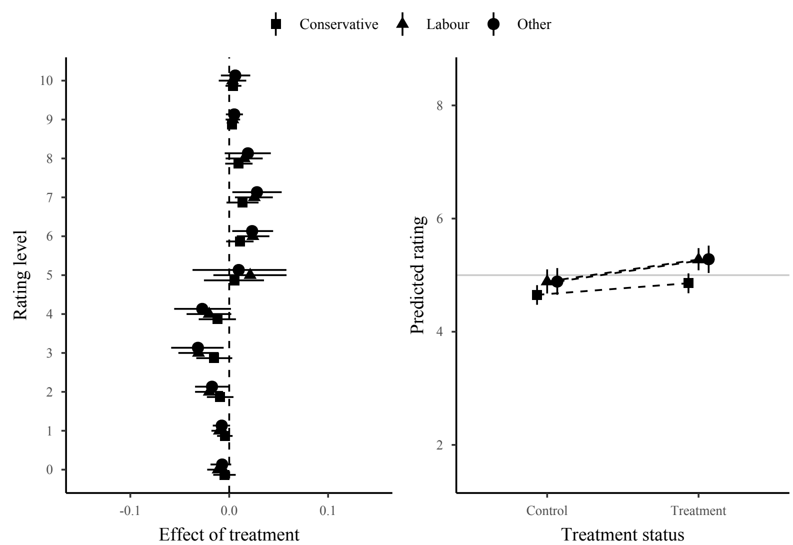


Figure C.10: Summary of ordinal model, Study Two, Canada condition.

Turning to Study Two, Figures C.8, C.9 and C.10 show the corresponding effects for the A/B, UK and Canada conditions, split by 2019 vote. Figure C.8 tells a very similar story to that seen in the +6 experiment in Study One, as expected. The effect is strongest on levels close to the midpoint of the outcome scale, and the average treatment effect is the same as estimated by the linear model. These effects are slightly stronger for other party supporters than Labour supporters, on whom the effect is in turn stronger than on Conservative supporters. Figure C.9 shows that the effect in the UK condition is instead concentrated lower down the scale. Treatment here makes people substantially less likely to give the Labour Party 0/10 chance of winning, and slightly less likely to give it 1/10 or 2/10. Ratings of 5/10 and 6/10 are correspondingly slightly more likely to be chosen. This reflects an average treatment effect that is substantial, but again concentrated among low outcome levels. These effects appear to be slightly stronger for other party supporters than for Labour or Conservative voters. Finally, Figure C.10 confirms that effects in the Canada condition are essentially negligible across the full outcome scale. Though the pattern of effects on the outcome scale is similar to that seen in Study One, only for Labour and other party supporters does treatment appear at any point to have a discernible effect on the outcome, reducing their probability of giving the Liberal Party a 3/10 chance and making them more likely to give it a 6/10 or 7/10 chance. This is reflected in small average effects, consistent with the linear model.

These findings serve to demonstrate that the main discussion in Chapter 4 does not misrepresent the results, which are robust to an alternative, ordinal specification. They also provide additional information, however, about how the effects are expressed more directly in terms of the outcome scale.

Appendix D

D.1 Complex Causal Models

The models presented to illustrate my key points in Chapter 5 are kept deliberately simple and parsimonious, in order to avoid unnecessary complexity that would confuse the argument. Here, I demonstrate what these models would look like when reintroducing this complexity. Figure D.1 shows how distinguishing between predispositions and preferences changes the picture slightly. This model more accurately asserts that it is predispositions, along with popularity information, that drive expectations, rather than preferences. Recall that preferences are convenient measures of relevant aspects of predispositions. Because there is still an indirect effect of popularity information on the vote through preferences here, the conclusions of Chapter 5 are unchanged.

Figure D.2 makes a further distinction between static and dynamic popularity information. This model is equivalent to the main model presented in Chapter 2, with the addition of predispositions. Again, this does not change the conclusions drawn in Chapter 5 and only further demonstrates the problem of distinguishing between static and dynamic effects using expectations data.

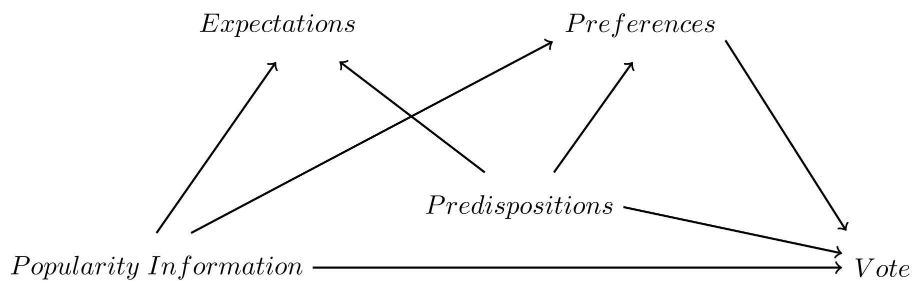


Figure D.1: Causal model of the relationship between expectations, predispositions, preferences, popularity information, and vote choice.

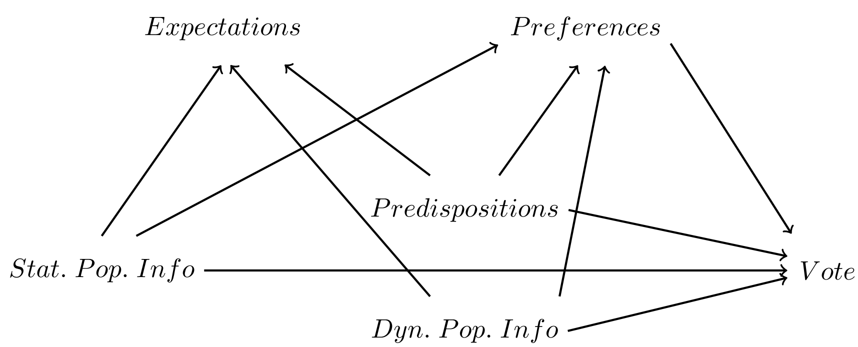


Figure D.2: Causal model of the relationship between expectations, predispositions, preferences, static and dynamic popularity information, and vote choice.

D.2 Full GCLM Fit Statistics

Table D.1 reports the full range of fit statistics for the GCLM in Chapter 5. In the main text, I only focused on the three most common, and easiest to interpret, measures. This full range of statistics consistently demonstrates a good level of fit.

Table D.1: Fit statistics for reduced bandwagon GCLM.

Fit Measure	Estimate	Fit Measure	Estimate
npar	45.000	bic2	290114.600
fmin	0.012	rmsea	0.051
chisq	284.276	rmsea.ci.lower	0.046
df	9.000	rmsea.ci.upper	0.057
pvalue	0.000	rmsea.pvalue	0.313
baseline.chisq	93423.914	rmr	0.038
baseline.df	36.000	rmr_nomean	0.042
baseline.pvalue	0.000	srmr	0.010
cfi	0.997	srmr_bentler	0.010
tli	0.988	srmr_bentler_nomean	0.011
nnfi	0.988	crmr	0.008
rfi	0.988	crmr_nomean	0.009
nfi	0.997	srmr_mplus	0.008
pnfi	0.249	srmr_mplus_nomean	0.009
ifi	0.997	cn_05	689.303
rni	0.997	cn_01	882.423
logl	-144918.298	gfi	0.996
unrestricted.logl	-144776.160	agfi	0.979
aic	289926.597	pgfi	0.166
bic	290257.605	mfi	0.988
ntotal	11565.000	ecvi	0.032

References

Abramowitz, Alan I. 1987. “Candidate Choice Before the Convention: The Democrats in 1984.” *Political Behavior* 9 (1): 49–61. https://doi.org/10.1007/BF00987278.

———. 1989. “Viability, Electability, and Candidate Choice in a Presidential Primary Election: A Test of Competing Models.” *The Journal of Politics* 51 (4): 977–92. https://doi.org/10.2307/2131544.

Abramson, Paul R., John H. Aldrich, Phil Paolino, and David W. Rohde. 1992. “"Sophisticated" Voting in the 1988 Presidential Primaries.” *The American Political Science Review* 86 (1): 55–69. https://doi.org/10.2307/1964015.

Acharya, Avidit, Matthew Blackwell, and Maya Sen. 2018a. “Analyzing Causal Mechanisms in Survey Experiments.” *Political Analysis* Forthcoming.

———. 2018b. “Explaining Preferences from Behavior: A Cognitive Dissonance Approach.” *The Journal of Politics* 80 (2): 400–411. https://doi.org/10.1086/694541.

Achen, Christopher, and Larry Bartels. 2015. *Democracy for Realists - Why Elections Do Not Produce Responsive Government*. Oxford: Princeton University

Press.

Agranov, Marina, Jacob K Goeree, Julian Romero, and Leeat Yariv. 2018. "What Makes Voters Turn Out: The Effects of Polls and Beliefs." *Journal of the European Economic Association* 16 (3): 825–56. https://doi.org/10.1093/jeea/jvx023.

Albrechsten, Kathrine. 2013. "Enigt Folketing Vil Afskaffe Meningsma° Linger Pa° Valgdagen." *Berlingske Nyhedsbureau* 20th November. https://www.berlingske.dk/politik/enigt-folketing-vil-afskaffe-meningsmaalinger-paa-valgdagen.

Allard, Winston. 1941. "A Test of Propaganda Values in Public Opinion Surveys." *Social Forces* 20 (2): 206–13. https://doi.org/10.2307/2571339.

Allegretti, Aubrey. 2021. "MPs Criticise Academics for Sending Them Fictitious Emails for Research." *The Guardian* 24th February. https://www.theguardian.com/politics/2021/feb/24/mps-criticise-academics-for-sending-them-fictitious-emails-for-research.

Allen, I Elaine, and Christopher A Seaman. 2007. "Likert Scales and Data Analyses." *Quality Progress - Statistics Roundtable*.

Allen, Peter. 2020. "Political Science, Punditry, and the Corbyn Problem." *British Politics* 15 (1): 69–87. https://doi.org/10.1057/s41293-019-00115-6.

Alvarez, R. M., F. J. Boehmke, and J. Nagler. 2006. "Strategic Voting in British Elections." *Electoral Studies* 25 (1): 1–19.

Alvarez, R. Michael, and Jonathan Nagler. 2000. "A New Approach for Modelling Strategic Voting in Multiparty Elections." *British Journal of Political Science*

30 (1): 57–75. www.jstor.org/stable/194113.

Anderson, Christopher J., Silvia M. Mendes, and Yuliya V. Tverdova. 2004. "Endogenous Economic Voting: Evidence from the 1997 British Election." *Electoral Studies* 23 (4): 683–708. https://doi.org/10.1016/j.electstud.2003.10.001.

Ansolabehere, Stephen, and Shanto Iyengar. 1994. "Of Horseshoes and Horse Races: Experimental Studies of the Impact of Poll Results on Electoral Behavior." *Political Communication* 11 (4): 413–30. https://doi.org/10.1080/10584609.1994.9963048.

Araújo, Victor, and Malu Gatto. 2021. "Casting Ballots When Knowing Results." *APSA Preprints*. https://preprints.apsanet.org/engage/apsa/article-details/606c0f1a25de03434435c523.

Arendt, Hannah. 1972. *Crises of the Republic*. New York: Harcourt.

Arkes, Jeremy. 2011. "Do Gamblers Correctly Price Momentum in NBA Betting Markets?" *Journal of Prediction Markets* 5 (1): 31–50.

Arnesen, Sveinung, Mikael P. Johannesson, Jonas Linde, and Stefan Dahlberg. 2017. "Do Polls Influence Opinions? Investigating Poll Feedback Loops Using the Novel Dynamic Response Feedback Experimental Procedure." *Social Science Computer Review* 36 (6): 735–43. https://doi.org/10.1177/0894439317731721.

Atkin, Charles K. 1969. "The Impact of Political Poll Reports on Candidate and Issue Preferences." *Journalism Quarterly* 46 (3): 515–21. https://doi.org/10.1177/107769906904600308.

Ayton, Peter, and Ilan Fischer. 2004. "The Hot Hand Fallacy and the Gambler's Fallacy: Two Faces of Subjective Randomness?" *Memory & Cognition* 32 (8): 1369–78.

Babad, Elisha. 1995. "Can Accurate Knowledge Reduce Wishful Thinking in Voters' Predictions of Election Outcomes?" *The Journal of Psychology* 129 (3): 285–300. https://doi.org/10.1080/00223980.1995.9914966.

———. 1997. "Wishful Thinking Among Voters: Motivational and Cognitive Influences." *International Journal of Public Opinion Research* 9 (2): 105–25. https://doi.org/10.1093/ijpor/9.2.105.

Babad, Elisha, and Eitan Yacobos. 1993. "Wish and Reality in Voters' Predictions of Election Outcomes." *Political Psychology* 14 (1): 37–54. https://doi.org/10.2307/3791392.

Bailey, Jack. 2019a. "Accounting for Endogeneity in Voters' Economic Perceptions: A Psychometric Approach." *Paper Presented at the 2019 Montréal-Leuven Winter School on Elections, KU Leuven*.

———. 2019b. "The Fact Remains: Party ID Moderates How Voters Respond to Economic Change." *Electoral Studies* 61 (October): 102071. https://doi.org/10.1016/j.electstud.2019.102071.

Bailey, Jack, and Matthew Barnfield. 2021. "Vaccine Bounce: Why We Should Take It with a Pinch of Salt." *UK in a Changing Europe - Commentary* 9th April. https://ukandeu.ac.uk/vaccine-bounce-why-we-should-take-it-with-a-pinch-of-salt/.

Bailey, Jack, Mark Pack, and Luke Mansillo. 2021. "PollBasePro: Daily Estimates of Aggregate Voting Intention in Britain from 1955 to the Present Day." *APSA Preprints*. https://doi.org/10.33774/apsa-2021-2hbgw.

Bale, Tim. 2002. "Restricting the Broadcast and Publication of Pre-Election and Exit Polls: Some Selected Examples." *Representation* 39 (1): 15–22. https://doi.org/10.1080/00344890208523210.

Barnfield, Matthew. 2019. "Think Twice Before Jumping on the Bandwagon: Clarifying Concepts in Research on the Bandwagon Effect." *Political Studies Review* 18 (4): 553–74. https://doi.org/10.1177/1478929919870691.

Bartels, Larry. 1985. "Expectations and Preferences in Presidential Nominating Campaigns." *American Political Science Review* 79: 804–15.

———. 1987. "Candidate Choice and the Dynamics of the Presidential Nominating Process." *American Journal of Political Science* 31: 1–30.

Bauer, Wilhelm. 1930. "Public Opinion." In *Encyclopaedia of the Social Sciences*, edited by E. Seligman. New York: Macmillian.

BBC. 2015a. "Jeremy Corbyn Elected Labour Leader: How Did He Win?" *BBC News* 12th September. https://www.bbc.co.uk/news/uk-politics-34126758.

———. 2015b. "Labour Leadership: Jeremy Corbyn Enters Race." *BBC News* 3rd June. https://www.bbc.co.uk/news/uk-politics-33000155.

Beniger, James R. 1976. "Winning the Presidential Nomination: National Polls and State Primary Elections, 1936-1972." *Public Opinion Quarterly* 40 (1): 22–38. https://doi.org/10.1086/268265.

Berelson, Bernard, Paul F. Lazarsfeld, and William McPhee. 1954. *Voting*. Chicago: University of Chicago Press.

Berghmans, Ron L. P. 2007. "Misleading Research Participants: Moral Aspects of Deception in Psychological Research." *Netherlands Journal of Psychology* 63 (1): 12–17. https://doi.org/10.1007/BF03061057.

Blais, André. 2000. *To Vote or Not to Vote: The Merits and Limits of Rational Choice Theory*. Pittsburgh: University of Pittsburgh Press.

Blais, André, and Marc André Bodet. 2006. "How Do Voters Form Expectations about the Parties' Chances of Winning the Election?" *Social Science Quarterly* 87 (3): 477–93. https://doi.org/10.1111/j.1540-6237.2006.00392.x.

Blais, André, Elisabeth Gidengil, and Neil Nevitte. 2006. "Do Polls Influence the Vote?" In *Capturing Campaign Effects*, edited by Henry E. Brady and Richard Johnston, 263–79. Michigan: The University of Michigan Press.

Blais, André, and Richard Nadeau. 1996. "Measuring Strategic Voting: A Two-Step Procedure." *Electoral Studies* 15 (1): 39–52. https://doi.org/10.1016/0261-3794(94)00014-X.

Bok, Sissela. 1995. "Shading the Truth in Seeking Informed Consent for Research Purposes." *Kennedy Institute of Ethics Journal* 5 (1): 1–17. https://doi.org/10.1353/ken.0.0116.

Bolsen, Toby, James N. Druckman, and Fay Lomax Cook. 2014. "The Influence of Partisan Motivated Reasoning on Public Opinion." *Political Behavior* 36 (2): 235–62. https://doi.org/10.1007/s11109-013-9238-0.

Bonneau, Chris W, and Damon M Cann. 2015. "Party Identification and Vote Choice in Partisan and Nonpartisan Elections." *Political Behavior* 37 (1): 43–66.

Bowler, Shaun, and Todd Donovan. 1998. *Demanding Choices: Opinion, Voting, and Direct Democracy*. Ann Arbor: The University of Michigan Press.

Bowler, Shaun, David J. Lanoue, and Paul Savoie. 1994. "Electoral Systems, Party Competition, and Strength of Partisan Attachment: Evidence from Three Countries." *The Journal of Politics* 56 (4): 991–1007. https://doi.org/10.2307/2132070.

Bowler, Shaun, Gail McElroy, and Stefan Müller. 2021. "Voter Expectations of Government Formation in Coalition Systems: The Importance of the Information Context." *European Journal of Political Research* Early View (March). https://doi.org/10.1111/1475-6765.12441.

Brady, Henry E., and Richard Johnston. 1987. "What's the Primary Message: Horse Race or Issue Journalism?" In *Media and Momentum: The New Hampshire Primary and Nomination Politics*, edited by Gary R Orren and Nelson W Polsby. Chatham, New Jersey: Chatham House.

Brown, Katy, Aurelien Mondon, and Aaron Winter. 2021. "The Far Right, the Mainstream and Mainstreaming: Towards a Heuristic Framework." *Journal of Political Ideologies*, July, 1–18. https://doi.org/10.1080/13569317.2021.1949829.

Bullock, John G., and Shang E. Ha. 2011. "Mediation Analysis Is Harder Than It Looks." In *Cambridge Handbook of Experimental Political Science*, edited by

James N. Druckman, Donald P. Green, James H. Kuklinski, and Arthur Lupia, 508–21. New York: Cambridge University Press.

Bullock, John G, Donald P Green, and Shang E Ha. 2010. "Yes, but What's the Mechanism?(don't Expect an Easy Answer)." *Journal of Personality and Social Psychology* 98 (4): 550.

Burgis, Ben. 2020. "Bernie Sanders Has the Momentum." *Jacobin*, March. https://jacobinmag.com/2020/03/bernie-sanders-has-the-momentum.

Burn-Murdoch, John, Martin Stabe, and Anna Leach. 2017. "UK General Election 2017 Poll Tracker." *Financial Times*, August. https://doi.org/https://ig.ft.com/elections/uk/2017/polls/.

Bürkner, Paul-Christian, and Matti Vuorre. 2019. "Ordinal Regression Models in Psychology: A Tutorial." *Advances in Methods and Practices in Psychological Science* 2 (1): 77–101. https://doi.org/10.1177/2515245918823199.

Bølstad, Jørgen, Elias Dinas, and Pedro Riera. 2013. "Tactical Voting and Party Preferences: A Test of Cognitive Dissonance Theory." *Political Behavior* 35 (3): 429–52. https://doi.org/10.1007/s11109-012-9205-1.

Callander, Steven. 2007. "Bandwagons and Momentum in Sequential Voting." *The Review of Economic Studies* 74 (3): 653–84.

Callard, Agnes. 2018. *Aspiration: The Agency of Becoming*. Oxford: Oxford University Press.

Campanile, Carl. 2020. "Trump Pollster Accuses Media of Voter Suppression Through Bad Polling." *New York Post*, November. https://nypost.com/2020/11/05/trump-pollster-accuses-media-of-voter-suppression-through-bad-polling/.

Campbell, Angus, Philip E. Converse, Warren E. Miller, and Donald E. Stokes. 1960. *The American Voter*. New York: John Wiley.

Campbell, Donald T. 1963. “From Description to Experimentation: Interpreting Trends as Quasi-Experiments.” In *Problems in Measuring Change*. Madison: University of Wisconsin Press.

Campbell, Rosie, and Philip Cowley. 2014. “What Voters Want: Reactions to Candidate Characteristics in a Survey Experiment.” *Political Studies* 62 (4): 745–65. https://doi.org/10.1111/1467-9248.12048.

Campbell, Rosie, Philip Cowley, Nick Vivyan, and Markus Wagner. 2016. “Legislator Dissent as a Valence Signal.” *British Journal of Political Science*, 1–24. https://doi.org/10.1017/S0007123416000223.

Carson, Thomas L. 2006. “The Definition of Lying.” *Noûs* 40 (2): 284–306. https://doi.org/10.1111/j.0029-4624.2006.00610.x.

Ceci, Stephen J., and Edward L. Kain. 1982. “Jumping on the Bandwagon With the Underdog: The Impact of Attitude Polls on Polling Behavior.” *The Public Opinion Quarterly* 46 (2): 228–42.

Chaiken, Shelly. 1980. “Heuristic Versus Systematic Information Processing and the Use of Source Versus Message Cues in Persuasion.” *Journal of Personality and Social Psychology* 39 (5): 752–66. https://doi.org/10.1037/0022-3514.39.5.752.

Chung, Sungeun, Yu-Jin Heo, and Jung-Hyun Moon. 2018. “Perceived Versus Actual Polling Effects: Biases in Perceptions of Election Poll Effects on Candidate Evaluations." *International Journal of Public Opinion Research* 30 (3): 420–42. https://doi.org/10.1093/ijpor/edx004.

Clifford, Scott, Geoffrey Sheagley, and Spencer Piston. 2021. "Increasing Precision Without Altering Treatment Effects: Repeated Measures Designs in Survey Experiments." *American Political Science Review*, 1–18. https://doi.org/10.1017/S0003055421000241.

Cloutier, Edouard, Richard Nadeau, and Jean Guay. 1989. "Bandwagoning and Underdoging on North-American Free Trade: A Quasi-Experimental Panel Study of Opinion Movement." *International Journal of Public Opinion Research* 1 (3): 206–20. https://doi.org/10.1093/ijpor/1.3.206.

Cohen, Ben. 2020. *The Hot Hand: The Mystery and Science of Streaks*. New York: Custom House.

Cook, Stuart A., and Alfred C. Welch. 1940. "Methods of Measuring the Practical Effects of Polls of Public Opinion." *Journal of Applied Psychology* 24 (4): 441–54. https://doi.org/10.1037/h0061199.

Coppock, Alexander, and Donald P. Green. 2016. "Is Voting Habit Forming? New Evidence from Experiments and Regression Discontinuities." *American Journal of Political Science* 60 (4): 1044–62. https://doi.org/10.1111/ajps.12210.

Cowden, Jonathan A., and Rose M. McDermott. 2000. "Short-Term Forces and Partisanship." *Political Behavior* 22 (3): 197–222. https://doi.org/10.1023/A:1026610113325.

Cowley, Philip. 2001. "The Observer: Good at Observing, Less Good at Influencing?" *Political Studies* 49 (5): 957–68. https://doi.org/10.1111/1467-9248.00349.

———. 2021. "Fake Letters to Politicians: The Timing May Be Bad but It's OK to Test Whether MPs Reply to Constituents." *The Conversation* 26th February. https://theconversation.com/fake-letters-to-politicians-the-timing-may-be-bad-but-its-ok-to-test-whether-mps-reply-to-constituents-156137.

Cowley, Philip, and Dennis Kavanagh. 2018. *The British General Election of 2017*. London: Palgrave Macmillan.

Crewe, Ivor. 1992. "A Nation of Liars? Opinion Polls and the 1992 Election." *Parliamentary Affairs* 45 (4): 475–95.

Cunow, Saul, Scott Desposato, Andrew Janusz, and Cameron Sells. 2021. "Less Is More: The Paradox of Choice in Voting Behavior." *Electoral Studies* 69 (February). https://doi.org/10.1016/j.electstud.2020.102230.

Cutts, David, Edward Fieldhouse, and Peter John. 2009. "Is Voting Habit Forming? The Longitudinal Impact of a GOTV Campaign in the UK." *Journal of Elections, Public Opinion and Parties* 19 (3): 251–63. https://doi.org/10.1080/17457280903073914.

Dahlgaard, Jens Olav, Jonas Hedegaard Hansen, Kasper M. Hansen, and Martin V. Larsen. 2017. "How Election Polls Shape Voting Behaviour." *Scandinavian Political Studies* 40 (3): 330–43. https://doi.org/10.1111/1467-9477.12094.

Dahlgreen, Will. 2015. "With One Month to Go, Corbyn's Lead Increases." *YouGov* 10th August. https://yougov.co.uk/topics/politics/articles-reports/2015

/08/10/corbyn-pull-ahead.

Daks, Alon, Nishant Desai, and Lisa R. Goldberg. 2018. “Do the Golden State Warriors Have Hot Hands?” *The Mathematical Intelligencer* 40 (4): 1–5. https://doi.org/10.1007/s00283-018-9825-3.

Dalton, Russell J. 2013. *The Apartisan American: Dealignment and Changing Electoral Politics*. Washington: CQ Press.

Dassonneville, Ruth. 2016. “Volatile Voters, Short-Term Choices? An Analysis of the Vote Choice Determinants of Stable and Volatile Voters in Great Britain.” *Journal of Elections, Public Opinion and Parties* 26 (3): 273–92. https://doi.org/10.1080/17457289.2016.1158181.

de Bock, Harold. 1976. “Influence of In-State Election Poll Reports on Candidate Preference in 1972.” *Journalism Quarterly* 53 (3): 457–62. https://doi.org/10.1177/107769907605300309.

de la Cuesta, Brandon, Naoki Egami, and Kosuke Imai. 2021. “Improving the External Validity of Conjoint Analysis: The Essential Role of Profile Distribution.” *Political Analysis*, 1–27. https://doi.org/10.1017/pan.2020.40.

———. 2021. “Improving the External Validity of Conjoint Analysis: The Essential Role of Profile Distribution.” *Political Analysis*, 1–27. https://doi.org/10.1017/pan.2020.40.

Delli Carpini, Michael X. 1984. “Scooping the Voters? The Consequences of the Networks’ Early Call of the 1980 Presidential Race.” *The Journal of Politics* 46 (3): 866–85.

Deutsch, Morton, and Harold B. Gerard. 1955. “A Study of Normative and Informational Social Influences Upon Individual Judgment.” *Journal of Abnormal and Social Psychology* 51 (3): 629–36.

Dickson, Eric S. 2011. “Economics Versus Psychology Experiments: Stylization, Incentives, and Deception.” In *Cambridge Handbook of Experimental Political Science*, edited by James N. Druckman, Donald P. Green, James H. Kuklinski, and Arthur Lupia, 58–69. New York: Cambridge University Press.

Dilliplane, Susanna. 2011. “All the News You Want to Hear: The Impact of Partisan News Exposure on Political Participation.” *Public Opinion Quarterly* 75 (2): 287–316. https://doi.org/10.1093/poq/nfr006.

———. 2013. “Activation, Conversion, or Reinforcement? The Impact of Partisan News Exposure on Vote Choice.” *American Journal of Political Science* 58 (1): 79–94. https://doi.org/10.1111/ajps.12046.

Dionne Jr, E. J. 1987. “Gary Hart, the Elusive Front-Runner.” *The New York Times Magazine* 3rd May. https://www.nytimes.com/2014/09/21/magazine/gary-hart-the-elusive-front-runner.html.

Dizney, Henry F., and Ronald W. Roskens. 1962. “An Investigation of the "Bandwagon Effect" in a College Straw Election.” *The Journal of Educational Sociology* 36 (3): 108–14. https://doi.org/10.2307/2264188.

Donsbach, Wolfgang, and Uwe Hartung. 2008. “The Legal Status of Public Opinion Research in the World.” In *The SAGE Handbook of Public Opinion Research*, edited by Wolfgang Donsbach and Michael W. Traugott, 431–40. London: SAGE Publications Ltd. https://doi.org/10.4135/9781848607910.

Dowding, Keith. 2017. *Power, Luck and Freedom: Collected Essays*. Manchester: Manchester University Press.

Downs, Anthony. 1957. *An Economic Theory of Democracy*. New York: Harper and Row.

Druckman, James N. 2014. "Pathologies of Studying Public Opinion, Political Communication, and Democratic Responsiveness." *Political Communication* 31 (3): 467–92. https://doi.org/10.1080/10584609.2013.852643.

Efferson, Charles, Rafael Lalive, Peter J. Richerson, Richard McElreath, and Mark Lubell. 2008. "Conformists and Mavericks: The Empirics of Frequency-Dependent Cultural Transmission." *Evolution and Human Behavior* 29 (1): 56–64. https://doi.org/10.1016/j.evolhumbehav.2007.08.003.

Eggers, Andrew C, and Nick Vivyan. 2020. "Who Votes More Strategically?" *American Political Science Review* 114 (2): 470–85. https://doi.org/10.1017/S0003055419000820.

Ehsan, Rakib, James Sloam, and Matt Henn. 2018. "Youthquake Was Real – Here's How We Know It Was More Than a Myth." *The Conversation*, January.

Eisenstein, Elizabeth. 1979. *The Printing Press as an Agent of Change: Communications and Cultural Transformations in Early-Modern Europe*. Cambridge: Cambridge University Press.

Elving, Ron. 2018. "The Florida Recount Of 2000: A Nightmare That Goes On Haunting." *NPR* 12th November. https://www.npr.org/2018/11/12/666812854/the-florida-recount-of-2000-a-nightmare-that-goes-on-haunting?t=1629639836324.

Evrenk, Haldun, and Chien-Yuan Sher. 2015. "Social Interactions in Voting Behavior: Distinguishing Between Strategic Voting and the Bandwagon Effect." *Public Choice* 162 (3): 405–23. https://doi.org/10.1007/s11127-015-0241-3.

Faas, Thorsten, Christian Mackenrodt, and Rüdiger Schmitt-Beck. 2008. "Polls That Mattered: Effects of Media Polls on Voters' Coalition Expectations and Party Preferences in the 2005 German Parliamentary Election." *International Journal of Public Opinion Research* 20 (3): 299–325. https://doi.org/10.1093/ijpor/edn034.

Farjam, Mike. 2020. "The Bandwagon Effect in an Online Voting Experiment With Real Political Organizations." *International Journal of Public Opinion Research*, June. https://doi.org/10.1093/ijpor/edaa008.

Ferguson, Niall. 2020. "Three Words That Haunt Joe Biden: "Dewey Defeats Truman"." *Bloomberg Opinion* Politics and Policy (November). https://www.bloomberg.com/opinion/articles/2020-11-01/three-words-that-haunt-joe-biden-dewey-defeats-truman.

Fieldhouse, Edward, Jane Green, Geoffrey Evans, John Mellon, and Chris Prosser. 2020. *British Election Study Internet Panel, 2014-2023*.

Fiorina, Morris. 1981. *Retrospective Voting in American National Elections*. New Haven, Connecticut: Yale University Press.

Fisher, Stephen D. 2004. "Definition and Measurement of Tactical Voting: The Role of Rational Choice." *British Journal of Political Science* 34 (1): 152–66.

———. 2014. "One Party Can Be the Biggest Winner and the Biggest Loser: Tactical Voting." In *Sex, Lies & the Ballot Box*, 73–77. London: Biteback

Publishing.

Fleitas, Daniel W. 1971. “Bandwagon and Underdog Effects in Minimal-Information Elections.” *The American Political Science Review* 65 (2): 434–38. https://doi.org/10.2307/1954459.

Ford, Robert. 2017. “Corbyn and Labour Surge but Tories Are Favourites Thanks to ‘Grey Vote’.” *The Observer* 4th June. https://www.theguardian.com/politics/2017/jun/03/labour-conservatives-poll-grey-vote.

Fowler, Anthony. 2019. “But Shouldn’t That Work Against Me?” The Political Methodologist. May 24, 2019. https://polmeth.org/blog/shouldn\T1\textquoterightt-work-against-me.

franceinfo. 2017. “Les 4 Vérités - Marine Le Pen : "Je Vais Gagner Cette Présidentielle".” *Franceinfo:élections* 7th April. https://www.francetvinfo.fr/elections/presidentielle/les-4-verites-marine-le-pen-je-vais-gagner-cette-presidentielle_2134501.html.

Frankfurt, Harry. 1998. *The Importance of What We Care about*. Cambridge: Cambridge University Press.

Frankovic, Kathy, Timothy Johnson, and Marina Stavrakantonaki. 2018. “The Freedom to Publish Opinion Poll Results: A Worldwide Update of 2012.” ESOMAR and WAPOR.

Freud, Sigmund. 1922. *Group Psychology and the Analysis of the Ego*. London: International Psychoanalytical Press.

Gallup, George, and Saul Forbes Rae. 1940. “Is There a Bandwagon Vote?” *Public Opinion Quarterly* 4 (2): 244–49. https://doi.org/10.1086/265397.

Ganser, Christian, and Patrick Riordan. 2015. "Vote Expectations at the Next Level. Trying to Predict Vote Shares in the 2013 German Federal Election by Polling Expectations." *Electoral Studies* 40 (December): 115–26. https://doi.org/10.1016/j.electstud.2015.08.001.

Gartner, Manfred. 1976. "Endogenous Bandwagon and Underdog Effects in a Rational Choice Model." *Public Choice* 25: 83–89.

Gauriot, Romain, and Lionel Page. 2018. "Psychological Momentum in Contests: The Case of Scoring Before Half-Time in Football." *Journal of Economic Behavior & Organization* 149 (May): 137–68. https://doi.org/10.1016/j.jebo.2018.02.015.

Geddes, Barbara. 1990. "How the Cases You Choose Affect the Answers You Get: Selection Bias in Comparative Politics." *Political Analysis* 2: 131–50. http://www.jstor.org/stable/23317768.

Gelman, Andrew. 2013. "A Poll Is a Snapshot, Not a Forecast." *Monkey Cage - Washington Post* December 10. https://www.washingtonpost.com/news/monkey-cage/wp/2013/12/10/a-poll-is-a-snapshot-not-a-forecast/.

Gelman, Andrew, and John Carlin. 2014. "Beyond Power Calculations: Assessing Type S (Sign) and Type M (Magnitude) Errors." *Perspectives on Psychological Science : A Journal of the Association for Psychological Science* 9 (6): 641–51. https://doi.org/10.1177/1745691614551642.

Gelman, Andrew, Jennifer Hill, and Aki Vehtari. 2021. *Regression and Other Stories*. Cambridge: Cambridge University Press.

Gelman, Andrew, and Gary King. 1993. "Why Are American Presidential Election Campaign Polls so Variable When Votes Are so Predictable?" *British Journal of Political Science* 23 (4): 409–51. http://www.jstor.org/stable/194212.

Gelman, Andrew, G Elliott Morris, Liberty Vittert, and Xiao-Li Meng. 2020. "Post-Election Interview with Andrew Gelman and G. Elliott Morris." *Harvard Data Science Review* 2 (4). https://hdsr.mitpress.mit.edu/pub/qqms394f/release/1.

Gilovich, Thomas, Robert Vallone, and Amos Tversky. 1985. "The Hot Hand in Basketball: On the Misperception of Random Sequences." *Cognitive Psychology* 17 (3): 295–314. https://doi.org/10.1016/0010-0285(85)90010-6.

Gimpel, James G., and Diane Hollern Harvey. 1997. "Forecasts and Preferences in the 1992 General Election." *Political Behavior* 19 (2): 157–75. https://doi.org/10.1023/A:1024810225641.

Glynn, Carroll J., Susan Herbst, Garrett J. O'Keefe, Robert Y. Shapiro, and Mark Lindeman. 2004. *Public Opinion*. 2nd Edition. Oxford: Westview Press.

Glynn, Carroll J., and Jack M. Mcleod. 1985. "Implications of the Spiral of Silence Theory for Communication and Public Opinion Research." In *Political Communication Yearbook 1984*, edited by K R Sanders, L L Kaid, and D Nimmo. Carbondale: Southern Illinois University Press.

Goeree, Jacob K., and Jens Großer. 2007. "Welfare Reducing Polls." *Economic Theory* 31 (1): 51–68. https://doi.org/10.1007/s00199-006-0082-x.

Goidel, Robert K., and Todd G. Shields. 1994. "The Vanishing Marginals, the Bandwagon, and the Mass Media." *The Journal of Politics* 56 (3): 802–10. https://doi.org/10.2307/2132194.

Goodwin, Matthew J. 2015. "Ukip, the 2015 General Election and Britain's EU Referendum." *Political Insight* 6 (3): 12–15. https://doi.org/10.1111/2041-9066.12107.

Graefe, Andreas. 2014. "Accuracy of Vote Expectation Surveys in Forecasting Elections." *Public Opinion Quarterly* 78 (S1): 204–32. https://doi.org/10.1093/poq/nfu008.

Granberg, Donald, and Edward Brent. 1983. "When Prophecy Bends: The Preference-Expectation Link in U.S. Presidential Elections, 1952-1980." *Journal of Personality and Social Psychology* 45 (3): 477–91. https://doi.org/10.1037/0022-3514.45.3.477.

Green, Donald P., and Alan S. Gerber. 2003. "The Underprovision of Experiments in Political Science." *The Annals of the American Academy of Political and Social Science* 589: 94–112.

Green, Donald P., Bradley Palmquist, and Eric Schickler. 1998. "Macropartisanship: A Replication and Critique." *American Political Science Review* 92: 883–99.

———. 2002. *Partisan Hearts and Minds: Political Parties and the Social Identities of Voters*. New Haven, Connecticut: Yale University Press.

Grenier, Éric. 2021. "Poll Tracker." *CBC News* May 31. https://newsinteractives.cbc.ca/elections/poll-tracker/canada/.

Grillo, Alberto. 2017. "Risk Aversion and Bandwagon Effect in the Pivotal Voter Model." *Public Choice* 172 (3): 465–82. https://doi.org/10.1007/s11127-017-0457-5.

Großer, Jens, and Arthur Schram. 2010. "Public Opinion Polls, Voter Turnout, and Welfare: An Experimental Study." *American Journal of Political Science* 54 (3): 700–717.

Gschwend, Thomas. 2007. "Ticket-Splitting and Strategic Voting Under Mixed Electoral Rules: Evidence from Germany." *European Journal of Political Research* 46 (1): 1–23. https://doi.org/10.1111/j.1475-6765.2006.00641.x.

Gschwend, Thomas, Michael F Meffert, and Lukas F Stoetzer. 2017. "Weighting Parties and Coalitions: How Coalition Signals Influence Voting Behavior." *The Journal of Politics* 79 (2): 642–55.

Gschwend, Thomas, Lukas Stoetzer, and Steffen Zittlau. 2016. "What Drives Rental Votes? How Coalitions Signals Facilitate Strategic Coalition Voting." *Electoral Studies* 44 (December): 293–306. https://doi.org/10.1016/j.electstud.2016.06.011.

Guay, Brian, and Christopher D. Johnston. 2021. "Ideological Asymmetries and the Determinants of Politically Motivated Reasoning." *American Journal of Political Science*, June. https://doi.org/10.1111/ajps.12624.

Hamaker, Ellen L., Rebecca M. Kuiper, and Raoul P. P. P. Grasman. 2015. "A Critique of the Cross-Lagged Panel Model." *Psychological Methods* 20 (1): 102–16. https://doi.org/10.1037/a0038889.

Hardmeier, Sibylle. 2008. "The Effects of Published Polls on Citizens." In *The SAGE Handbook of Public Opinion Research*, edited by Wolfgang Donsbach and Michael W. Traugott, 504–13. London: SAGE Publications Ltd. https://doi.org/10.4135/9781848607910.

Hayes, S P Jr. 1932. “The Predictive Ability of Voters.” *Journal of Social Psychology* 7: 183–91.

Henshel, Richard L. 1982. “The Boundary of the Self-Fulfilling Prophecy and the Dilemma of Social Prediction.” *British Journal of Sociology* 33 (4): 511–28. https://doi.org/10.2307/589359.

Henshel, Richard L., and William Johnston. 1987. “The Emergence of Bandwagon Effects: A Theory.” *The Sociological Quarterly* 28 (4): 493–511.

Herbst, Susan. 1993. *Numbered Voices: How Opinion Polling Has Shaped American Politics*. Chicago: University of Chicago Press.

———. 2021. *A Troubled Birth: The 1930s and American Public Opinion*. Chicago: University of Chicago Press.

Hernadi, Alexandra. 2010. “Statsvetare Skeptisk till Opinionsförbud.” *Svenska Dagbladet* 3rd September. https://www.svd.se/statsvetare-skeptisk-till-opinionsforbud.

Hertwig, Ralph, and Andreas Ortmann. 2008. “Deception in Experiments: Revisiting the Arguments in Its Defense.” *Ethics & Behavior* 18 (1): 59–92. https://doi.org/10.1080/10508420701712990.

Hill, Seth J. 2017. “Learning Together Slowly: Bayesian Learning about Political Facts.” *The Journal of Politics* 79 (4): 1403–18. https://doi.org/10.1086/692739.

Ho, Daniel E, and Kosuke Imai. 2006. “Randomization Inference With Natural Experiments.” *Journal of the American Statistical Association* 101 (475): 888–900. https://doi.org/10.1198/016214505000001258.

Hobolt, Sara B. 2018. "Brexit and the 2017 UK General Election." *Journal of Common Market Studies* Early View (July): 1–12. https://doi.org/10.1111/jcms.12751.

Hodgson, Robert, and John Maloney. 2013. "Bandwagon Effects in British Elections, 18851910." *Public Choice* 157 (1): 73–90. https://doi.org/10.1007/s11127-012-0027-9.

Howell, Susan E. 1981. "Short Term Forces and Changing Partisanship." *Political Behavior* 3 (2): 163–80. https://doi.org/10.1007/BF00989844.

Huddy, Leonie, Lilliana Mason, and Lene Aarøe. 2015. "Expressive Partisanship: Campaign Involvement, Political Emotion, and Partisan Identity." *American Political Science Review* 109 (1): 1–17. https://doi.org/10.1017/S0003055414000604.

Irwin, G. A., and J. J. M. Van Holsteyn. 2000. "Bandwagons, Underdogs, the Titanic, and the Red Cross. The Influence of Public Opinion Polls on Voters." In *Communication Presented at the 18th Congress of the International Political Science Association, Quebec*.

Irwin, Galen A. 2002. "According to the Polls: The Influence of Opinion Polls on Expectations." *The Public Opinion Quarterly* 66 (1): 92–104.

Iyengar, Shanto, Helmut Norpoth, and Kyu S Hahn. 2004. "Consumer Demand for Election News: The Horserace Sells." *The Journal of Politics* 66 (1): 157–75.

Jackson, John E. 1983. "Election Night Reporting and Voter Turnout." *American Journal of Political Science* 27 (4): 615–35. https://doi.org/10.2307/2110886.

Jennings, Will, and Christopher Wlezien. 2018. "Election Polling Errors Across Time and Space." *Nature Human Behaviour* 2 (4): 276–83. https://doi.org/10.1038/s41562-018-0315-6.

Jones, Owen. 2015. "Jeremy Corbyn's Supporters Aren't Mad – They're Fleeing a Bankrupt New Labour." *The Guardian* 3rd August. https://www.theguardian.com/commentisfree/2015/aug/03/jeremy-corbyn-new-labour-centre-left.

Jost, John T. 2017. "Ideological Asymmetries and the Essence of Political Psychology." *Political Psychology* 38 (2): 167–208. https://doi.org/10.1111/pops.12407.

Kant, Immanuel. 1978 (1798). *Anthropology from a Pragmatic Point of View*. Translated by Victor Lyle Dowdell. London: Feffer and Simons, Inc.

Kaplan, Noah, David K. Park, and Andrew Gelman. 2012. "Polls and Elections Understanding Persuasion and Activation in Presidential Campaigns: The Random Walk and Mean Reversion Models." *Presidential Studies Quarterly* 42 (4): 843–66. https://doi.org/10.1111/j.1741-5705.2012.04021.x.

Karp, Jeffrey A., and Maarja Lühiste. 2016. "Explaining Political Engagement with Online Panels: Comparing the British and American Election Studies." *Public Opinion Quarterly* 80 (3): 666–93. https://doi.org/10.1093/poq/nfw014.

Keeter, Scott, Nick Hatley, Arnold Lau, and Courtney Kennedy. 2021. "What 2020's Election Poll Errors Tell Us About the Accuracy of Issue Polling." *Pew Research Center* March 2. https://www.pewresearch.org/methods/2021/03/02/what-2020s-election-poll-errors-tell-us-about-the-accuracy-of-issue-polling/.

Kellner, Peter. 2015. "Comment: Corbyn Takes Early Lead in Labour Leadership Race." *YouGov* 22nd July. https://yougov.co.uk/topics/politics/articles-reports/2015/07/22/comment-corbyn-ahead-labours-leadership-contest.

Kelman, Herbert C. 1967. "Human Use of Human Subjects: The Problem of Deception in Social Psychological Experiments." *Psychological Bulletin* 67 (1): 1–11. https://doi.org/10.1037/h0024072.

Kenney, Patrick J, and Tom W. Rice. 1994. "The Psychology of Political Momentum." *Political Research Quarterly* 47 (4): 923–38.

Key, V. O. 1966. *The Responsible Electorate*. Cambridge. Harvard University Press.

Kirschbaum, Erik. 2009. "German Pollsters, TV Raise Guard on Twitter Leaks." *Reuters* 20th September. https://www.reuters.com/article/us-germany-election-twitter-idUSTRE58J0JB20090920.

Kiss, 'Aron, and G'abor Simonovits. 2014. "Identifying the Bandwagon Effect in Two-Round Elections." *Public Choice* 160 (3): 327–44. https://doi.org/10.1007/s11127-013-0146-y.

Konner, Joan, James Risser, and Ben Wattenburg. 2001. "Television's Performance on Election Night 2000." *CNN* 29th January. http://i.cnn.net/cnn/2001/ALLPOLITICS/stories/02/02/cnn.report/cnn.pdf.

Krizan, Zlatan, and Kate Sweeny. 2013. "Causes and Consequences of Expectation Trajectories: 'High' on Optimism in a Public Ballot Initiative." *Psychological Science* 24 (5): 706–14. https://doi.org/10.1177/0956797612460690.

Kropko, Jonathan, and Robert Kubinec. 2020. "Interpretation and Identification of Within-Unit and Cross-Sectional Variation in Panel Data Models." *PLOS ONE* 15 (4): e0231349. https://doi.org/10.1371/journal.pone.0231349.

Kurtzleben, Danielle. 2020. "Why Were The Polls Off? Pollsters Have Some Early Theories." *NPR* Elections (November). https://www.npr.org/2020/11/19/936317341/why-were-the-polls-off-pollsters-have-some-early-theories?t=1620209103358.

Kurz, A Solomon. 2020. "Statistical Rethinking with Brms, Ggplot2, and the Tidyverse." Bookdown.org. 2020. https://bookdown.org/ajkurz/Statistical_Rethinking_recoded/.

Lai, Keke, and Samuel B Green. 2016. "The Problem with Having Two Watches: Assessment of Fit When RMSEA and CFI Disagree." *Multivariate Behavioral Research* 51 (2-3): 220–39.

Landgrave, Michelangelo. 2020. "Can We Reduce Deception in Elite Field Experiments? Evidence from a Field Experiment with State Legislative Offices." *State Politics & Policy Quarterly* 20 (4): 489–507. https://doi.org/10.1177/1532440020925723.

Lang, Kurt, and Gladys Engel Lang. 1984. "The Impact of Polls on Public Opinion." *The ANNALS of the American Academy of Political and Social Science* 472 (1): 129–42. https://doi.org/10.1177/0002716284472001012.

Lanoue, David J., and Shaun Bowler. 1998. "Picking the Winners: Perceptions of Party Viability and Their Impact on Voting Behavior." *Social Science Quarterly* 79 (2): 361–77.

Larsen, Erik Gahner, and Zoltán Fazekas. 2021. "Reactions and Implications: How Do the Elite and the Public Respond to Polls?" In *Reporting Public Opinion: How the Media Turns Boring Polls into Biased News*, edited by Erik Gahner Larsen and Zoltán Fazekas, 83–108. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-75350-4_5.

Lasswell, Harold D. 1936. *Politics: Who Gets What, When, How*. New York: Whittlesey House.

Lau, Richard R, and David P. Redlawsk. 2006. *How Voters Decide: Information Processing During Election Campaigns*. Cambridge: Cambridge University Press.

Lavrakas, Paul J, Jack K Holley, and Peter V Miller. 1991. "Public Reactions to Polling News During the 1988 Presidential Election Campaign." In *Polling and Presidential Election Coverage*, edited by Paul J Lavrakas and Jack K Holley, 151–83. Newbury Park: SAGE Publications Ltd.

Lazarsfeld, Paul F., Bernard Berelson, and Hazel Gaudet. 1948. *The People's Choice - How the Voter Makes up His Mind in a Presidential Campaign*. 2nd ed. New York: Columbia University Press.

Le Bon, Gustave. 1948. *The Crowd: A Study of the Popular Mind*. London: Unwin.

Le Guin, Ursula K. 2017. "Introduction." In *The Left Hand of Darkness*, by Ursula K Le Guin, xiii–xvii. London: Gollancz.

———. 2019. *Words Are My Matter: Writings on Life and Books*. New York: Mariner Books.

Leemann, Lucas, Lukas F. Stoetzer, and Richard Traunmüller. 2021. "Eliciting Beliefs as Distributions in Online Surveys." *Political Analysis*, 1–13. https://doi.org/10.1017/pan.2020.42.

Leeper, Thomas J. 2018. "Interpreting Regression Results Using Average Marginal Effects with R's Margins." *CRAN Working Paper*, May.

Leeper, Thomas J. 2019. "Where Have the Respondents Gone? Perhaps We Ate Them All." *Public Opinion Quarterly* 83 (S1): 280–88. https://doi.org/10.1093/poq/nfz010.

Leeper, Thomas J., Sara B. Hobolt, and James Tilley. 2019. "Measuring Subgroup Preferences in Conjoint Experiments." *Political Analysis*, 1–15. https://doi.org/10.1017/pan.2019.30.

Leeper, Thomas J., and Rune Slothuus. 2014. "Political Parties, Motivated Reasoning, and Public Opinion Formation." *Political Psychology* 35 (S1): 129–56. https://doi.org/10.1111/pops.12164.

Leszczensky, Lars, and Tobias Wolbring. 2019. "How to Deal With Reverse Causality Using Panel Data? Recommendations for Researchers Based on a Simulation Study." *Sociological Methods & Research*, November, OnlineFirst. https://doi.org/10.1177/0049124119882473.

Lewis, David. 1978. "Truth in Fiction." *American Philosophical Quarterly* 15 (1): 37–46. http://www.jstor.org/stable/20009693.

Lewis-Beck, Michael S., and Mary Stegmaier. 2011. "Citizen Forecasting: Can UK Voters See the Future?" *Special Symposium: Electoral Forecasting Symposium* 30 (2): 264–68. https://doi.org/10.1016/j.electstud.2010.09.012.

Lewis-Beck, Michael S, W G Jacoby, H Norpoth, and Herbert F. Weisberg. 2008. *The American Voter Revisited.* Ann Arbor: University of Michigan.

Liddell, Torrin M., and John K. Kruschke. 2018. "Analyzing Ordinal Data with Metric Models: What Could Possibly Go Wrong?" *Journal of Experimental Social Psychology* 79: 328–48. https://doi.org/10.1016/j.jesp.2018.08.009.

Little, Andrew T. 2021. "Directional Motives and Different Priors Are Observationally Equivalent." *Working Paper*.

Lodge, Milton, and Charles S. Taber. 2013. *The Rationalizing Voter*. Cambridge: Cambridge University Press.

Madson, Gabriel J., and D. Sunshine Hillygus. 2019. "All the Best Polls Agree with Me: Bias in Evaluations of Political Polling." *Political Behavior*, February. https://doi.org/10.1007/s11109-019-09532-1.

Marcus, G, W Neuman, and M MacKuen. 2000. *Affective Intelligence and Political Judgement*. Chicago: University of Chicago Press.

Marsh, Catherine. 1985. "Back on the Bandwagon: The Effect of Opinion Polls on Public Opinion." *British Journal of Political Science* 15 (1): 51–74.

Mason, Lilliana. 2018. *Uncivil Agreement: How Politics Became Our Identity*. Chicago: The University of Chicago Press.

McAllister, Ian, and Donley T. Studlar. 1991. "Bandwagon, Underdog, or Projection? Opinion Polls and Electoral Choice in Britain, 1979-1987." *The Journal of Politics* 53 (3): 720–41. https://doi.org/10.2307/2131577.

McDermott, Rose. 2011. "Internal and External Validity." In *Cambridge Handbook of Experimental Political Science*, edited by James N. Druckman, Donald P. Green, James H. Kuklinski, and Arthur Lupia, 27–40. New York: Cambridge University Press.

McElreath, Richard. 2020. *Statistical Rethinking: A Bayesian Course with Examples in R and Stan*. Second. London: CRC Press.

Medeiros, João. 2017. "Final Election Poll Gives Jeremy Corbyn the Lead over Theresa May for the First Time." *Wired*, June. https://www.wired.co.uk/article/election-polls-labour-conservative-winner.

Meffert, Michael F., and Thomas Gschwend. 2011. "Polls, Coalition Signals and Strategic Voting: An Experimental Investigation of Perceptions and Effects." *European Journal of Political Research* 50 (5): 636–67. https://doi.org/10.1111/j.1475-6765.2010.01986.x.

Meffert, Michael F., Sascha Huber, Thomas Gschwend, and Franz Urban Pappi. 2011. "More Than Wishful Thinking: Causes and Consequences of Voters' Electoral Expectations about Parties and Coalitions." *Electoral Studies* 30 (4): 804–15. https://doi.org/10.1016/j.electstud.2011.08.001.

Mehrabian, Lbert. 1998. "Effects of Poll Reports on Voter Preferences." *Journal of Applied Social Psychology* 28 (23): 2119–30. https://doi.org/10.1111/j.1559-1816.1998.tb01363.x.

Meier, Philippe, Raphael Flepp, Maximilian Ruedisser, and Egon Franck. 2020. "Separating Psychological Momentum from Strategic Momentum: Evidence from Men's Professional Tennis." *Journal of Economic Psychology* 78 (June):

102269. https://doi.org/10.1016/j.joep.2020.102269.

Mendelsohn, Harold, and Irving Crespi. 1970. *Polls, Television, and the New Politics*. Scranton: Chandler Publishing Company.

Mercer, Andrew. 2016. "5 Key Things to Know about the Margin of Error in Election Polls." *Pew Research Center*, September. https://doi.org/https://www.pewresearch.org/fact-tank/2016/09/08/understanding-the-margin-of-error-in-election-polls/.

Miller, Joshua B., and Adam Sanjurjo. 2018. "Surprised by the Hot Hand Fallacy? A Truth in the Law of Small Numbers." *Econometrica* 86 (6): 2019–47. https://doi.org/10.3982/ECTA14943.

Mongrain, Philippe. 2021. "Did You See It Coming? Explaining the Accuracy of Voter Expectations for District and (Sub)national Election Outcomes in Multi-Party Systems." *Electoral Studies* 71 (June): 102317. https://doi.org/10.1016/j.electstud.2021.102317.

Mortensen, Chad R., Rebecca Neel, Robert B. Cialdini, Christine M. Jaeger, Ryan P. Jacobson, and Megan M. Ringel. 2019. "Trending Norms: A Lever for Encouraging Behaviors Performed by the Minority." *Social Psychological and Personality Science* 10 (2): 201–10. https://doi.org/10.1177/1948550617734615.

Morton, Rebecca B., Daniel Muller, Lionel Page, and Benno Torgler. 2015. "Exit Polls, Turnout, and Bandwagon Voting: Evidence from a Natural Experiment." *European Economic Review* 77 (Supplement C): 65–81. https://doi.org/10.1016/j.euroecorev.2015.03.012.

Morton, Rebecca B., and Kai Ou. 2015. "What Motivates Bandwagon Voting Behavior: Altruism or a Desire to Win?" *Behavioral Political Economy* 40 (Part B): 224–41. https://doi.org/10.1016/j.ejpoleco.2015.04.009.

Morwitz, Vicki G., and Carol Pluzinski. 1996. "Do Polls Reflect Opinions or Do Opinions Reflect Polls? The Impact of Political Polling on Voter's Expectations, Preferences, and Behavior." *Journal of Consumer Research* 23 (1): 53–67.

Moy, Patricia, and Eike Mark Rinke. 2012. "Attitudinal and Behavioral Consequences of Published Opinion Polls." In *Opinion Polls and the Media: Reflecting and Shaping Public Opinion*, edited by Christina Holtz-Bacha and Jesper Strömbäck, 225–45. London: Palgrave Macmillan UK. https://doi.org/10.1057/9780230374959_11.

Murdoch, Iris. 1992. *Metaphysics as a Guide to Morals*. New York: Allen Lane.

Murr, Andreas E. 2016. "The Wisdom of Crowds: What Do Citizens Forecast for the 2015 British General Election?" *Electoral Studies* 41 (March): 283–88. https://doi.org/10.1016/j.electstud.2015.11.018.

Murr, Andreas E., Mary Stegmaier, and Michael S. Lewis-Beck. 2019a. "Vote Expectations Versus Vote Intentions: Rival Forecasting Strategies." *British Journal of Political Science*, 1–8. https://doi.org/10.1017/S0007123419000061.

Murr, Andreas E., Mary Stegmaier, and Michael S Lewis-Beck. 2019b. "Citizen Forecasting 2019: A Big Win for the Conservatives." *LSE BPP*, December. https://doi.org/https://blogs.lse.ac.uk/politicsandpolicy/citizen-forecasting-2019-a-big-win-for-the-conservatives/.

Murr, Andreas Erwin. 2011. "'Wisdom of Crowds?' A Decentralised Election Forecasting Model That Uses Citizens' Local Expectations." *Electoral Studies* 30 (4): 771–83. https://doi.org/10.1016/j.electstud.2011.07.005.

Muthukrishna, Michael, and Joseph Henrich. 2019. "A Problem in Theory." *Nature Human Behaviour* 3 (3): 221–29. https://doi.org/10.1038/s41562-018-0522-1.

Mutz, Diana C. 1997. "Mechanisms of Momentum: Does Thinking Make It So?" *The Journal of Politics* 59 (1): 104–25. https://doi.org/10.2307/2998217.

———. 1998. *Impersonal Influence: How Perceptions of Mass Collectives Affect Political Attitudes*. Cambridge: Cambridge University Press.

Nadeau, Richard, Edouard Cloutier, and J.-H. Guay. 1993. "New Evidence About the Existence of a Bandwagon Effect in the Opinion Formation Process." *International Political Science Review* 14 (2): 203–13. https://doi.org/10.1177/019251219301400204.

Nadeau, Richard, Richard G. Niemi, and Timothy Amato. 1994. "Expectations and Preferences in British General Elections." *The American Political Science Review* 88 (2): 371–83. https://doi.org/10.2307/2944710.

Navazio, Robert. 1977. "An Experimental Approach to Bandwagon Research." *The Public Opinion Quarterly* 41 (2): 217–25.

Nickell, Stephen. 1981. "Biases in Dynamic Models with Fixed Effects." *Econometrica* 49 (6): 1417–26. https://doi.org/10.2307/1911408.

Noelle-Neumann, Elisabeth. 1977. "Turbulences in the Climate of Opinion: Methodological Applications of the Spiral of Silence Theory." *Public Opinion Quarterly* 41 (2): 143–58.

Orne, Martin T. 1962. "On the Social Psychology of the Psychological Experiment: With Particular Reference to Demand Characteristics and Their Implications." *American Psychologist* 17 (11): 776–83. https://doi.org/10.1037/h0043424.

Page, Benjamin, and Robert Y Shapiro. 1992. *The Rational Public: Fifty Years of Trend in America's Policy Preferences*. Chicago: The University of Chicago Press.

Palan, Stefan, and Christian Schitter. 2018. "Prolific.ac—A Subject Pool for Online Experiments." *Journal of Behavioral and Experimental Finance* 17 (March): 22–27. https://doi.org/10.1016/j.jbef.2017.12.004.

Panova, Elena. 2015. "A Passion for Voting." *Games and Economic Behavior* 90 (March): 44–65. https://doi.org/10.1016/j.geb.2015.01.003.

Parry, Douglas A., Brittany I. Davidson, Craig J. R. Sewall, Jacob T. Fisher, Hannah Mieczkowski, and Daniel S. Quintana. 2021. "A Systematic Review and Meta-Analysis of Discrepancies Between Logged and Self-Reported Digital Media Use." *Nature Human Behaviour*, May. https://doi.org/10.1038/s41562-021-01117-5.

Paul, L. A., and Kieran Healy. 2018. "Transformative Treatments." *Noûs* 52 (2): 320–35. https://doi.org/10.1111/nous.12180.

Pearl, Judea. 2000. *Causality: Models, Reasoning, and Inference*. Cambridge: Cambridge University Press.

———. 2010. "An Introduction to Causal Inference." *The International Journal of Biostatistics* 6 (2): 7–7. https://doi.org/10.2202/1557-4679.1203.

Pearl, Judea, and Dana MacKenzie. 2018. *The Book of Why: The New Science of Cause and Effect*. London: Allen Lane.

Petty, Richard E, and John T Cacioppo. 1986. "The Elaboration Likelihood Model of Persuasion." *Advances in Experimental Social Psychology* 19: 123–205. https://doi.org/https://doi.org/10.1016/S0065-2601(08)60214-2.

———. 1996. *Attitudes and Persuasion: Classic and Contemporary Approaches*. Boulder, Colorado: Westview Press.

Pickup, Mark, and Geoffrey Evans. 2013. "Addressing the Endogeneity of Economic Evaluations in Models of Political Choice." *Public Opinion Quarterly* 77 (3): 735–54.

Pierce, Walter M. 1940. "Climbing on the Bandwagon." *Public Opinion Quarterly* 4 (2): 241–43. https://doi.org/10.1086/265396.

Platt, Robert W. 1998. "ANOVA, t Tests, and Linear Regression." *Injury Prevention* 4 (1): 52–53.

Plescia, Carolina. 2018. "On the Subjectivity of the Experience of Victory: Who Are the Election Winners?" *Political Psychology* Early View: 1–18. https://doi.org/10.1111/pops.12562.

Popkin, Samuel L. 1991. *The Reasoning Voter: Communication and Persuasion in Presidential Campaigns*. Chicago: University of Chicago Press.

Price, Vincent. 1992. *Public Opinion*. Newbury Park: SAGE Publications Ltd.

Prosser, Christopher. 2018. "The Strange Death of Multi-Party Britain: The UK General Election of 2017." *West European Politics* 41 (5): 1226–36.

https://doi.org/10.1080/01402382.2018.1424838.

Prosser, Christopher, Edward Fieldhouse, Jane Green, Jonathan Mellon, and Geoffrey Evans. 2020. "Tremors but No Youthquake: Measuring Changes in the Age and Turnout Gradients at the 2015 and 2017 British General Elections." *Electoral Studies* 64 (April): 102129. https://doi.org/10.1016/j.electstud.2020.102129.

Prosser, Christopher, and Jonathan Mellon. 2018. "The Twilight of the Polls? A Review of Trends in Polling Accuracy and the Causes of Polling Misses." *Government and Opposition* 53 (4): 757–90. https://doi.org/10.1017/gov.2018.7.

Rhemtulla, Mijke, Patricia É Brosseau-Liard, and Victoria Savalei. 2012. "When Can Categorical Variables Be Treated as Continuous? A Comparison of Robust Continuous and Categorical SEM Estimation Methods Under Suboptimal Conditions." *Psychological Methods* 17 (3): 354–73.

Riambau, Guillem. 2015. "Bandwagon or Strategic Voting in Israel? Note on Bargsted and Kedar 2009." *Electoral Studies* 37 (March): 63–72. https://doi.org/10.1016/j.electstud.2014.10.007.

———. 2016. "Do Citizens Vote for Parties, Policies or the Expected Winner in Proportional Representation Systems? Evidence from Four Different Countries Using a Multiple-Type Model." *Party Politics* 24 (5): 549–62. https://doi.org/10.1177/1354068816668669.

Rohrer, Julia M. 2018. "Thinking Clearly About Correlations and Causation: Graphical Causal Models for Observational Data." *Advances in Methods and*

Practices in Psychological Science 1 (1): 27–42. https://doi.org/10.1177/2515245917745629.

Rosseel, Yves. 2012. “Lavaan: An R Package for Structural Equation Modeling.” *Journal of Statistical Software; Vol 1, Issue 2 (2012)*. https://doi.org/10.18637/jss.v048.i02.

Roy, Jason, Shane P. Singh, Patrick Fournier, and Blake Andrew. 2015. “An Experimental Analysis of the Impact of Campaign Polls on Electoral Information Seeking.” *Electoral Studies* 40 (Supplement C): 146–57. https://doi.org/10.1016/j.electstud.2015.08.005.

Sayare, Scott. 2012. “French Media Question Election Reporting Rules.” *New York Times* 10th April. https://www.nytimes.com/2012/04/21/world/europe/french-media-question-election-reporting-rules.html.

Schmitt-Beck, Rüdiger. 1996. “Mass Media, the Electorate, and the Bandwagon. A Study of Communication Effects on Vote Choice in Germany.” *International Journal of Public Opinion Research* 8 (3): 266–91. https://doi.org/10.1093/ijpor/8.3.266.

Schwartz, Ian. 2020. “Trump: I Should Get A "Redo Of 4 Years" Because My Campaign Was Spied On.” *RealClear Politics* 17th August. https://www.realclearpolitics.com/video/2020/08/17/trump_we_should_get_a_redo_of_four_years_because_my_campaign_was_spied_on.html.

Select Committee on Political Polling and Digital Media. 2018. “The Politics of Polling.” *Parliament.uk* HL Paper 106. https://publications.parliament.uk/pa/ld201719/ldselect/ldppdm/106/10602.htm.

Shi, Dexin, Taehun Lee, and Alberto Maydeu-Olivares. 2018. "Understanding the Model Size Effect on SEM Fit Indices." *Educational and Psychological Measurement* 79 (2): 310–34. https://doi.org/10.1177/0013164418783530.

Simon, Herbert A. 1954. "Bandwagon and Underdog Effects and the Possibility of Election Predictions." *The Public Opinion Quarterly* 18 (3): 245–53.

Skalaban, Andrew. 1988. "Do the Polls Affect Elections? Some 1980 Evidence." *Political Behavior* 10 (2): 136–50.

Sneddon, Andrew. 2021. "Alternative Motivation and Lies." *Analysis* 81 (1): 46–52. https://doi.org/10.1093/analys/anaa027.

Sparkman, Gregg, and Gregory M. Walton. 2019. "Witnessing Change: Dynamic Norms Help Resolve Diverse Barriers to Personal Change." *Journal of Experimental Social Psychology* 82 (May): 238–52. https://doi.org/10.1016/j.jesp.2019.01.007.

Stacey, Kiran. 2015. "Labour MPs Call on Gordon Brown to Halt Jeremy Corbyn Bandwagon." *Financial Times* 11th August. https://www.ft.com/content/4f9bc81a-401a-11e5-9abe-5b335da3a90e.

Stewart, Marianne, Harold D Clarke, Matthew Goodwin, and Paul Whiteley. 2018. "Yes, There Was a 'Youthquake' in the 2017 Snap Election – and It Mattered." *New Statesman*, February.

Stiers, Dieter, Jean-François Daoust, and André Blais. 2018. "What Makes People Believe That Their Party Won the Election?" *Electoral Studies* 55: 21–29.

Stiers, Dieter, and Ruth Dassonneville. 2018. "Affect Versus Cognition: Wishful Thinking on Election Day: An Analysis Using Exit Poll Data from Belgium."

International Journal of Forecasting 34 (2): 199–215. https://doi.org/10.1016/j.ijforecast.2017.12.001.

Stock, James H., and Mark W. Watson. 2001. "Vector Autoregressions." *Journal of Economic Perspectives* 15 (4): 101–15. https://doi.org/10.1257/jep.15.4.101.

Stoetzer, Lukas F., and Matthias Orlowski. 2020. "Estimating Coalition Majorities During Political Campaigns Based on Pre-Election Polls." *Journal of Elections, Public Opinion and Parties* 30 (1): 126–37. https://doi.org/10.1080/17457289.2019.1582533.

Stolwijk, Sjoerd B, Andreas R T Schuck, and Claes H de Vreese. 2017. "How Anxiety and Enthusiasm Help Explain the Bandwagon Effect." *International Journal of Public Opinion Research* 29 (4): 554–74. https://doi.org/10.1093/ijpor/edw018.

Stolwijk, Sjoerd B, and Andreas RT Schuck. 2019. "More Interest in Interest: Does Poll Coverage Help or Hurt Efforts to Make More Young Voters Show up at the Ballot Box?" *European Union Politics* 20 (3): 341–60. https://doi.org/10.1177/1465116519837351.

Strong, James. 2018. "Confidence and Caretakers: Some Less-Obvious Implications of the Fixed-Term Parliaments Act." *The Political Quarterly* 89 (3): 466–73. https://doi.org/10.1111/1467-923X.12485.

Sturgis, Patrick. 2016. "Structural Equation Modeling: What Is It and What Can We Use It for? (Part 1 of 6)." YouTube Video.

Sturgis, Patrick, and Juni Kuha. 2018. "'New Poll Suggests…': How to Tell When Public Opinion Has Really Changed." *LSE Blogs* 3rd July. https://blogs.lse.ac

.uk/politicsandpolicy/how-to-tell-when-public-opinion-has-really-changed/.

Sudman, Seymour. 1986. "Do Exit Polls Influence Voting Behavior?" *The Public Opinion Quarterly* 50 (3): 331–39. http://www.jstor.org/stable/2748722.

Thorson, Kjerstin, and Chris Wells. 2015. "Curated Flows: A Framework for Mapping Media Exposure in the Digital Age." *Communication Theory* 26 (3): 309–28. https://doi.org/10.1111/comt.12087.

Tudor, Jack, and Matthew Wall. 2021. "A Moving Target? An Analysis of the Impact of Electoral Context on Polling Error Variation in Both British and International General Elections." *Journal of Elections, Public Opinion and Parties*, August, 1–23. https://doi.org/10.1080/17457289.2021.1972305.

Utych, Stephen M., and Cindy D. Kam. 2014. "Viability, Information Seeking, and Vote Choice." *The Journal of Politics* 76 (1): 152–66. https://doi.org/10.1017/S0022381613001126.

van der Meer, Tom W. G., Armen Hakhverdian, and Loes Aaldering. 2016. "Off the Fence, Onto the Bandwagon? A Large-Scale Survey Experiment on Effect of Real-Life Poll Outcomes on Subsequent Vote Intentions." *International Journal of Public Opinion Research* 28 (1): 46–72. https://doi.org/10.1093/ijpor/edu041.

van der Meer, Tom W G, Rozemarijn Lubbe, Erika van Elsas, Martin Elff, and Wouter van der Brug. 2012. "Bounded Volatility in the Dutch Electoral Battlefield: A Panel Study on the Structure of Changing Vote Intentions in the Netherlands During 20062010." *Acta Politica* 47 (4): 333–55. https://doi.org/10.1057/ap.2012.5.

VanderWeele, Tyler J. 2015. *Explanation in Causal Inference: Methods for Mediation and Interaction*. New York: Oxford University Press.

Veit, Walter. 2020. "Why Joe Biden Is Winning." *Psychology Today*, March. https://www.psychologytoday.com/gb/blog/science-and-philosophy/202003/why-joe-biden-is-winning.

Vergin, Roger. 2000. "Winning Streaks in Sports and the Mispreception of Momentum." *Journal of Sport Behavior* 23 (2): 181. https://www.proquest.com/scholarly-journals/winning-streaks-sports-mispreception-momentum/docview/1311960695/se-2?accountid=13375.

Volckart, Oliver. 2021. "Voting Like Your Betters: The Bandwagon Effect in the Diet of the Holy Roman Empire." *LSE Economic History Working Papers*. http://eprints.lse.ac.uk/111613/.

Westwood, Sean Jeremy, Solomon Messing, and Yphtach Lelkes. 2020. "Projecting Confidence: How the Probabilistic Horse Race Confuses and Demobilizes the Public." *The Journal of Politics* 82 (4): 1530–44.

Whiteley, Paul, and Harold D Clarke. 2017. "Understanding Labour's 'Youthquake'." *The Conversation*, July.

Zaller, John R. 1992. *The Nature and Origins of Mass Opinion*. Cambridge: Cambridge University Press.

Zaller, John R., and Stanley Feldman. 1992. "A Simple Theory of the Survey Response: Answering Questions Versus Revealing Preferences." *American Journal of Political Science* 36 (3): 579–616. https://doi.org/10.2307/2111583.

Zerback, Thomas, Carsten Reinemann, and Matthew Barnfield. 2021. "Total Recall? Examining the Accuracy of Poll Recall During an Election Campaign." *Mass Communication and Society*, June, 1–23. https://doi.org/10.1080/15205436.2021.1936559.

Zerback, Thomas, Carsten Reinemann, and Angela Nienierza. 2015. "Who's Hot and Who's Not? Factors Influencing Public Perceptions of Current Party Popularity and Electoral Expectations." *The International Journal of Press/Politics* 20 (4): 458–77. https://doi.org/10.1177/1940161215596986.

Zizzo, Daniel John. 2010. "Experimenter Demand Effects in Economic Experiments." *Experimental Economics* 13 (1): 75–98. https://doi.org/10.1007/s10683-009-9230-z.

Zyphur, Michael J., Paul D. Allison, Louis Tay, Manuel C. Voelkle, Kristopher J. Preacher, Zhen Zhang, Ellen L. Hamaker, et al. 2019. "From Data to Causes I: Building A General Cross-Lagged Panel Model (GCLM)." *Organizational Research Methods* 23 (4): 651–87. https://doi.org/10.1177/1094428119847278.

Zyphur, Michael J., Manuel C. Voelkle, Louis Tay, Paul D. Allison, Kristopher J. Preacher, Zhen Zhang, Ellen L. Hamaker, et al. 2019. "From Data to Causes II: Comparing Approaches to Panel Data Analysis." *Organizational Research Methods* 23 (4): 688–716. https://doi.org/10.1177/1094428119847280.