

Tracking the Impact of Media on Voter Choice in Real Time: A Bayesian Dynamic Joint Model

Bhuvanesh Pareek, Pulak Ghosh,
Hugh N. Wilson, Emma K. Macdonald, and Paul Baines

September 9, 2017

Bhuvanesh Pareek is Assistant Professor, Department of Operations Management and Quantitative Techniques at Indian Institute of Management, Indore, India; Pulak Ghosh is Professor, Department of Decision Sciences and Information Systems, Indian Institute of Management, Bangalore, India; Hugh N. Wilson, Emma K. Macdonald and Paul Baines are Professors of Marketing at Cranfield School of Management, Cranfield University, UK; Big thanks to Prof. Rajiv Sinha who has been very instrumental in this work, although he is no more with us, his contribution will be cherished.

Abstract

Commonly used methods of evaluating the impact of marketing communications during political elections struggle to account for respondents' exposures to these communications due to the problems associated with recall bias. In addition, they completely fail to account for the impact of mediated or earned communications such as newspaper articles or television news, that are typically not within the control of the advertising party, nor are they effectively able to monitor consumers' perceptual responses over time. This study based on a new data collection technique using cell-phone text messaging (called real-time experience tracking or RET) offers the potential to address these weaknesses. We propose an RET-based model of the impact of communications and apply it to a unique choice situation: voting behavior during the 2010 UK general election, which was dominated by three political parties. We develop a Bayesian zero-inflated dynamic multinomial choice model that enables the joint modeling of: the inter-play and dynamics associated with the individual voter's choice intentions over time, actual vote, and the heterogeneity in the exposure to marketing communications over time. Results reveal the differential impact over time of paid and earned media, demonstrate a synergy between the two, and show the particular importance of exposure valence and not just frequency, contrary to the predominant practitioner emphasis on share-of-voice metrics. Results also suggest that while earned media have a reducing impact on voting intentions as the final choice approaches, their valence continues to influence the final vote: a difference between drivers of intentions and behavior that implies that exposure valence remains critically important close to the final brand choice.

Key Words: Advertising; Bayesian Methods; Communications - paid and earned; real-time experience tracking; Zero-inflation; Multinomial Logit Model; Voters' Choice;

1 Introduction

Recent advances in information technology, along with the advent of social media and enablement of user-generated content, have resulted in a proliferation of consumer data. Additionally, the widespread use of mobile devices by consumers makes it feasible to collect data about influences on consumer choice in real time. This presents an opportunity for researchers wanting to model the effects of communication on consumer choice, such as organizations working to better understand the voter's choice during an election campaign. However both the volume of consumer data being generated and its real-time nature necessitates the use of new and original analytical tools. This creates fresh opportunities for informed managerial decision-making, particularly for firms working to better understand the voter's choice during an election campaign. This paper addresses two substantive issues concerning the impact of marketing communications on voting choice: the differential effectiveness of alternative marketing communications, and the relative importance of communications frequency on the one hand and the valence of the consumer's perceptual response on the other. Global spend on advertising in 2017 was \$583.91bn(eMarketer 2017). Communications evaluation is a substantial business activity in its own right, providing around 15% of the trade for the \$68bn market research industry(ESOMAR 2016). Many firms also work hard to generate positive communications that are mediated by third parties such as journalists and retailers. Much of this effort aims to influence the consumer's choice of brand, whether it be consumer products and services, business-to-business suppliers, non-profit organizations, or political candidates and parties; the latter being the context for this study. Given the massive expense involved and their potential impact on consumer choice, evaluation and optimization of the effectiveness of communications is critical. While there have been several studies analysing election data (Linzer 2013, Christensen and Florence 2008, Gelman et al. 2008), to our knowledge no study has examined the relative influence of multiple paid and earned media communications on a voter's choice of a political party during an election period.

This study makes use of a new and innovative data collection procedure that permits individuals to report on communications from competing political parties as soon as they encounter them, thereby alleviating the recall bias associated with traditional research techniques. Importantly, we distinguish between exposure to communications that are controlled and paid for by a firm or party ('paid' media such as advertising, newspaper inserts, and billboards) and those that are not controlled by them ('earned' media such as news, editorials, communications about a particular brand by rival brands, etc.). These are referred to as paid and earned communications respectively. Although these are important modes of communication, finding the impact of these communicati-

ons on voters' choice has four key challenges. The first is the typical problem with survey based data. Although surveys can explore voters' attitudes towards a political party and can often, but not always, be quite helpful in forecasting election results (Linzer 2013), most respondents find it difficult to recall their exposure to communications accurately. One can examine the impact of media spend on party choice, but only rarely are data available on the impact of earned communications such as media coverage (which is not controlled by political parties), and it is often difficult to disentangle the effects of the medium from the message. The second problem is in assessing the differential impact of various communications that are controlled by the brands. Models based on media spend are extensively used by media planners to examine the relative impact of different paid media such as TV, radio, and billboards. Although the impact of some earned media such as newspaper editorials has been similarly assessed, this stream of work generally uses entirely separate data sources and models (Goh et al. 2011, Stephen and Galak 2012). However, marketing directors and political strategists need to allocate resources between paid media on the one hand and PR and service improvements focused on influencing earned media, on the other. Third, in assessing communications impact, researchers frequently ignore the crucial role of the consumer's perceptual response to the communication. Theoretical and experimental work confirms that attitude towards communications influences attitude towards the brand (Briñol et al. 2004), but most models of field data, particularly in the case of paid media, focus purely on communications exposure (Verhoef 2003) or the proxy of media spend (Naik and Peters 2009), presumably because perceptual response data are frequently unavailable. Fourth, practitioners have an interest in understanding how the impact of communications on voters' intentions evolve over time and prior to the final brand/party choice. In many consumer choice contexts, particularly in a voting scenario, consumers have extended purchase journeys, within which the timing of communications may be important. Furthermore, communications such as advertising are often intended to have medium-term impacts via attitude change as well as more immediate behavioral impacts; similarly, PR may deliberately or accidentally have lasting effects. Traditionally, a key area of interest has been how to model shifts in voting intention or preference over time and how these shifts impact on subsequent choice decisions. However, what has not been researched, up to now, is why these preference changes have occurred and what role paid and earned communications play in that attitude change process.

In this article, we explore the potential of a real-time experience tracking (RET) technique in addressing many of the aforementioned weaknesses. RET has recently been adopted for communications evaluation by a number of multinationals such as Unilever, BSkyB, PepsiCo, HP, Energizer, InterContinental Hotels and 20th Century Fox, among others (see Macdonald et al. 2012).

The genesis of this technique may be traced to the need for capturing the most essential aspects of consumer encounters with brand communications in real time and in a quick and easy response format. This research method involves asking a panel of trained respondents to send a cell phone text (SMS) message as soon as they encounter any of a set of competitive brands during the study. The respondents, who receive prior training about sending these brief, very structured messages, provide information on three essential aspects of their encounter: the brand name, the communications type (for example, TV advertisement or press editorial), and their perceptual response to the encounter. Respondents also complete a brand attitude survey at the start and end of the study. RET can capture the impact of a wide range of communications on an individual in real time; however, no model of brand choice is currently available to exploit the novelty of these kind of data. We use the RET method to collect data on choice of political parties in the UK's 2010 General Election, with a tracking period - the length of time for which each participant is asked to report their communications exposure-of the complete four-week election period. In addition to reporting all party (brand) encounters throughout the month by text message, we asked participants to provide their initial voting intentions, and to provide weekly updates on their voting intentions, via weekly surveys. For modeling purposes, this context has three benefits: the final party choice is made simultaneously by all participants, thereby alleviating problems of data censoring; the 2010 UK election campaign spanned a four-week period from the announcement of the election to the actual vote, allowing us to track intentions for the entire choice cycle for all participants; and it enables the capture of paid and earned communications on the choice of political party for an individual in a way that many existing research techniques cannot.

Consequently, we propose a RET-based model to capture the impact of these unique, real-time encounters on the choice of political party. Our model incorporates both paid and earned communications, and takes into account both the frequency of exposures and their valence, a measure of the voter's perceptual response to the communication. The ability to track, differentiate and model these two effects is one of the most important contributions of this research. We present a Bayesian Zero-Inflated dynamic multinomial choice model for simultaneously analyzing four critical factors that are associated with the impact of communications on voters' choice: (a) voter's choice intentions and their evolution over time (b) voters' final choice (c) the frequency and valence of paid and earned media, and (d) heterogeneity in the exposures to different communications over time.

In summary, this article contributes to the statistical literature on voter choice theory and, more broadly, on consumer choice in four ways: (1) it unpacks the differential impact of paid and earned communications during a particular purchase window, thereby allowing firms to optimize their re-

source allocation; (2) it demonstrates the complementary roles of communications frequency and communications valence, again with implications for firms' communication strategies; (3) it explores the role of RET in communications evaluation, including recommendations to practitioners on how emerging RET data collection practices need to be redefined; and (4) it presents a dynamic multinomial choice model with zero-inflation that simultaneously models choice intentions, actual choice, and the heterogeneity in the number and valence of communications that respondents are exposed to during the choice period. An important research direction is to explore whether this model applies as equally to commercial brand data as it does to voting data.

The remainder of this article is organized as follows: Section 2 describes the RET method and data from UK 2010 general election; Section 3 describes our proposed zero-inflated multinomial joint model; Section 4 describes the Bayesian inference approach adopted; Section 5 summarizes the empirical findings of our analysis and results from the simulation study followed by a discussion of the theoretical and managerial implications of our findings in Section 6.

2 Motivating Data

Many organizations continue to rely on cross-sectional surveys for collecting data, particularly in politics. One approach used by market research agencies such as Ipsos MORI is self-reported influence, which is based on asking respondents what media influenced them in their voting decisions (Worcester et al. 2011). Consumers are not, however, always reliable at reflecting on the causes of their attitudes and decisions, displaying social desirability bias for example, and biases in emotion recall. A second approach is to correlate recalled exposure or perceptual response to different media with brand attitudes (O'Cass 2002). As well as common method bias, this approach suffers from the inaccuracy of respondents' recall of past events (Belli et al. 1999). The use of a panel responding to repeated surveys, such as the American National Election Studies database, arising from interviews conducted before and after each presidential election (Klein and Ahluwalia 2005), addresses the issue of common method bias but not the problem of recall bias. It is well known that the time of voting decision is important and has been highlighted as a possible segmentation variable, not least because those who decide how to vote during a campaign are affected by media campaigns (Fournier et al. 2004). Our model attempts to address many of these shortcomings by incorporating the impact of real time responses to marketing communications on voting intentions and behavior over the entire campaign period, which we describe below.

2.1 Real-time Experience Tracking (RET)

Figure 1 illustrates how RET works by showing the data collection process for the politics study. RET builds on the observation that the mobile phone is a device for data capture which travels with a consumer wherever they go. A panel of respondents is asked to send a text (SMS) message whenever they come across any of a set of competitive brands during the study period, which typically lasts between one week and one month (Macdonald et al. 2012). The short, very structured messages provide information on three essential aspects of the encounter: (a) the brand name, (b) the communications type or touchpoint (for example, billboard or radio editorial), and (c) the affective response or valence of the encounter. Respondents are briefed with a coding scheme for the message, which has one character for each of these aspects: a letter for each brand, a letter for each communications type, and a number for valence. The valence item was, “How did it make you feel about the party?” on a 5-point scale anchored by ‘Very positive’ to ‘Very negative’. So for example, in a message AF3 sent by the participant, ‘A’ might represent the Conservative Party; ‘F’ represents a billboard; and ‘3’ represents a neutral response. Respondents also complete an online

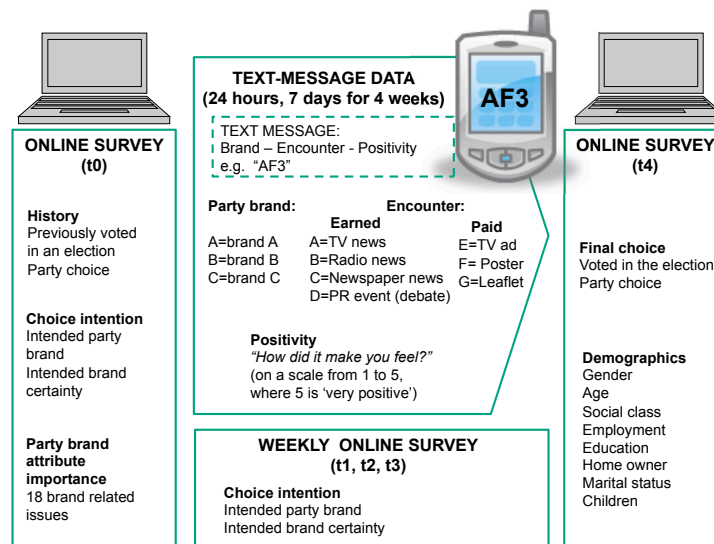


Figure 1: RET Data collection.

survey at the start and end of the study. The pre-study survey (or pre-survey for short) includes a briefing to respondents about the text message coding scheme, which is also summarized by text message to respondents for easy access when texting about exposures. The pre-survey captures brand choice history, while the post-survey captures any final brand choice such as the actual vote.

The pre-survey also captures brand intention at the start of the study. The 4-week study was supplemented with a brief weekly survey that captures the evolution of a voter's intentions, to support a dynamic model.

To our knowledge, there have only been three previous papers using RET in the academic or hybrid academic-practitioner literature. The first two are primarily descriptive. Macdonald et al. (2012), describes the RET method to those practitioners who have not yet come across it. Drawing on only parts of the politics dataset used within this paper, Baines et al. (2011) focus on an interpretation of descriptive data concerning changing communications valence over time, suggesting that these shifting voter responses to communications - and not just their frequency - may have been influential in final voter choice. However, Baines et al. (2011) do not attempt to model this, and their only statistical analysis is a very simple exploratory one: a bivariate chi-squared test to assess the association of dichotomous exposure to different media at any point in the campaign with final vote. Neither does the paper describe or use the weekly survey incorporating intentions data, which plays a central role in our model and is critical in determining model efficacy. The third paper, Baxendale, Macdonald and Wilson (2015) examines the impact of exposures on brand attitude in a consumer goods context but without any dynamic modelling. In order to evaluate the relative impact on the choice of earned and paid/owned communications with more rigor and usefulness, a dynamic model is needed that takes account of: a) communications frequency, and not just dichotomous exposure; b) communications valence, as attitude towards communications influences attitude towards the party c) dynamics over time in both choice intentions and communications exposure and valence; and d) consumers' two-stage decision-making process of whether to vote, and if so which party to choose. We develop such a model in the next section.

First, however, we note that while the RET data collection method appears to have significant advantages for gathering dynamic choice views, it is not without flaw. First, there is likely to be some under-reporting of exposures, whether out of forgetfulness or laziness. For example, if while on the bus listening to radio coverage of an election a participant sees an election poster out of the bus window, the participant would need to be extremely diligent to record both events. We might, though, expect this under-reporting to be less than with surveys which rely on memory of exposures over a preceding week. Surveys are the only method aside from RET currently available for recording the complete range of exposure types that were studied in this paper, as many exposures such as offline word of mouth leave no objectively recorded trace. In support of this argument about under-reporting, our study has similarities with the earlier paper diary method commonly used to track consumer purchases by consumer panels, in that both involve self-reporting of actions by consumers; paper diaries have been experimentally compared with weekly surveys and

found to be closer to objective purchase levels (Wind and Lerner 1979). Nonetheless, an ideal method would capture exposure occurrence objectively and completely. This is currently only feasible for some exposures which leave a behavioural trace, such as receiving word of mouth in an email, or clicking on a digital advertisement. Future research might seek to fuse such objective exposure capture with self-reported RET data for other exposures such as offline WOM or seeing a billboard.

Regarding valence, too, RET has disadvantages as well as advantages. On the positive side, there are well-established problems of asking consumers to recall positivity or other valenced emotions retrospectively, as emotions decay rapidly and are subject to reconstruction for a variety of reasons such as reducing post-purchase cognitive dissonance (Cowley 2007). RET therefore has the advantage that respondents are being asked to report on valence immediately after an exposure. However, a drawback that RET and surveys share is that self-reporting of attitudes is itself an intervention in the relationship, which may encourage respondents to firm up their views more rapidly than non-respondents, reducing their utility as a population sample (Cowley 2007). Capture of exposure valence, like capture of exposure occurrence, would ideally occur objectively and without intervention. Again, this is possible currently only with some exposure types, such as social media (Berger et al. 2010). We return to this matter in the Conclusions, indicating directions for future research to estimate more precisely the validity properties of RET.

2.2 The Campaign: The 2010 UK General Election

The above mentioned RET method was used during the 2010 UK general election to collect our motivating data. The data were collected by market research agency MESH Planning, developers of the formulation of RET described here. Respondents were recruited from an online panel provider, ResearchNow. The sample comprised swing voters who were filtered through asking what party they would vote for if the election was held that day and their certainty of voting for that party on a 10-point scale; only those answering 8 or below were included. Quota sampling was used to recruit a sample of 1,100 respondents, representative of the UK population in terms of region, social status, age group and gender. As the study required a four-week commitment (i.e. the entire election cycle, as UK elections were at that time announced by the incumbent government only shortly before the elections take place), there was a significant drop-out rate resulting in a final sample of 480 respondents (see Table 1). No additional fee was paid per text message, but the cost of the text message was refunded, in order to avoid under-reporting or over-reporting of exposures.

The sampling of swing voters has the benefit for this study of making the best use of the available sample, as undecided voters are more likely to be swayed by communications. Amongst those who voted in this sample, in only 22.5% of cases was the final vote the same as the initial intention in the first survey, so the sample did indeed include a large amount of switching. While the UK political system includes multiple parties, at the time it was dominated by three: Labour, Conservative and Liberal Democrat gaining 88.1% of the total vote share. We focus purely on these in our analysis. The party shares in the sample's final vote were 31.2% Conservative, 22.9% Labour and 34.8% Liberal Democrat. This compares with national actual votes of 36.1%, 29.0% and 23.0% respectively. The Liberal Democrats' high share of switchers is consistent with the party's overall gains over the election period: for example, YouGov's national poll on the first day of this study, showed shares in voting intention of 41% Conservative, 32% Labour and 18% Liberal Democrat, with the Liberal Democrat's final 23% vote being achieved at the expense of both Conservatives and Labour.

As with other practitioner applications of RET, the study began and ended with a survey. The initial survey captured the respondent's voting history: here, whether they had voted in the last general election, and if so, what their vote had been. The initial survey also captured their voting intention and how certain they were of that vote. To capture heterogeneity in customer needs, respondents were profiled on eighteen issue-related items that captured the perceived importance to the respondent of each policy issue. The final survey, conducted immediately after polling, contained information on whether the respondent had voted and the party for which they voted, as well as demographic, socio-economic and political data.

Respondents were briefed with a coding scheme for the message, as described above in section 2.1. The first character represented the party, and codes were given for each party (A=Conservative, etc); the second character represented the type of encounter, with a list of codes such as A for news, etc; and the third character on valence was an answer to the question, "How did it make you feel about the party?", on a scale from 1 to 5. Thus, each respondent's was provided online at the end of the pre-survey and then the coding scheme was summarised by text message to the respondent for easy access when texting about exposures. Thus, each respondent's exposure to different communications could be classified into: (a) the total number of exposures of this communications type (*exposure frequency*) and (b) the average valence, or positivity, score across all exposures of this exposure type (*exposure valence*). While all brand communications were captured equally (including an 'Other' category in the texting framework to capture communications types not anticipated in advance), the analysis focuses on the most commonly reported communications types, as shown in Figure 1; other encounter types such as social media (far less prevalent in this 2010 UK election

than in recent US presidential elections) were present in volumes too small to permit analysis. The conciseness of the message had the aim of minimizing the disruption to the respondent's life. In addition, the date and time of the message were automatically recorded.

Table 1: Summary Statistics ¹

Gender	%	Education	%
Female	53.3	Low (High School and lower)	55.2
Male	46.7	High (College/ University and above)	44.8
Age	%	Socio-Economic Status	%
<25	12.7	A	9.4
25-34	15.4	B	24.8
35-44	22.7	C1	23.5
45-54	18.1	C2	15.2
55-64	16.3	D	9.6
65 +	14.8	E	17.5
Marital Status	%	Occupation	%
Divorced/ Widowed	12.1	Unemployed (retired, student, housewife)	44.4
Married/ Partnership	62.3	Employed (paid, self-employed) full time	40.4
Single	25.6	Employed (paid, self-employed) part time	15.2
Children	%	Home ownership	%
Children of any age	62.3	Mortgagee	36.3
No children	37.7	Owner	29.4
		Tenant: private rental	16.0
		Tenant: social housing	12.9
		Unspecified	5.4

¹For a detailed explanation of the the British socioeconomic status scale, JICNARS (Joint Industry Committee for National Readership Surveys), where A refers to high managerial, or professional class, B refers to the intermediate managerial, or professional class, C1 refers to the supervisory, clerical and junior managerial, administrative or professional class, C2 refers to skilled manual workers, D to semi-skilled and unskilled manual workers, and E to state pensioners, casual or low-grade workers, or the unemployed on state benefits only,
see: https://www.ipsos-mori.com/DownloadPublication/1285_MediaCT_thoughtpiece_Social_Grade_July09_V3_WEB.pdf

After a period of majority Labour rule, the 2010 British general elections returned a LibDem-Conservative coalition government, putting the Liberals back in power for the first time since the 1920s. Interestingly, while the Conservatives maintained a steady lead in voting intentions over Labour from May/June 2009 (Worcester et al. 2011), they failed to win the election, ending 20 seats short of an outright majority. Therefore, in this paper we attempt to understand how the election (marketing) campaign influenced voting intentions and actual voting behavior, as well as the relative role of different communication channels in generating such an outcome.

The essential features of the data, such as the distribution of voters by their choice of party and the frequency of exposures (earned and paid), are presented in Figures 2-3. It is worth pointing out that earned media is considerably more important in the UK than it is in the US. The reason for this is that broadcast political advertising is illegal in the UK and ubiquitous in the US. In the UK, instead a limited quota of short 5-10 minute public service broadcast slots are made available on terrestrial channels.

3 A Zero-Inflated Joint Model of Media Communications and Voter Choice

Our joint model consists of five components that capture the impact of media exposures on choice. These are as follows:

- 1) a flexible multinomial logit model for *choice intentions* and their dynamics over time, including the impact of paid and earned media on party voting intentions.
- 2) a zero-inflated multinomial logit model to capture the *final vote choice* given the decision to vote.
- 3) a zero-inflated count model to capture the exposure *frequency*. Zero-inflated distribution is used in order to account for the excess zeros caused by the fact that respondents are typically not exposed to the majority of marketing communications on a daily or regular basis.
- 4) a zero-inflated log-normal model to capture the *average valence* of exposures.
- 5) finally, we join the models using a correlated random effects approach.

It is important to model the components jointly. Since all the dependent variables, viz., choice intentions, final vote, exposure valence and frequency are correlated for a voter, a joint model would allow us to borrow information and bring estimation efficiency by reducing bias. Another motivation for joint modeling is the fact that even when we are interested in understanding how

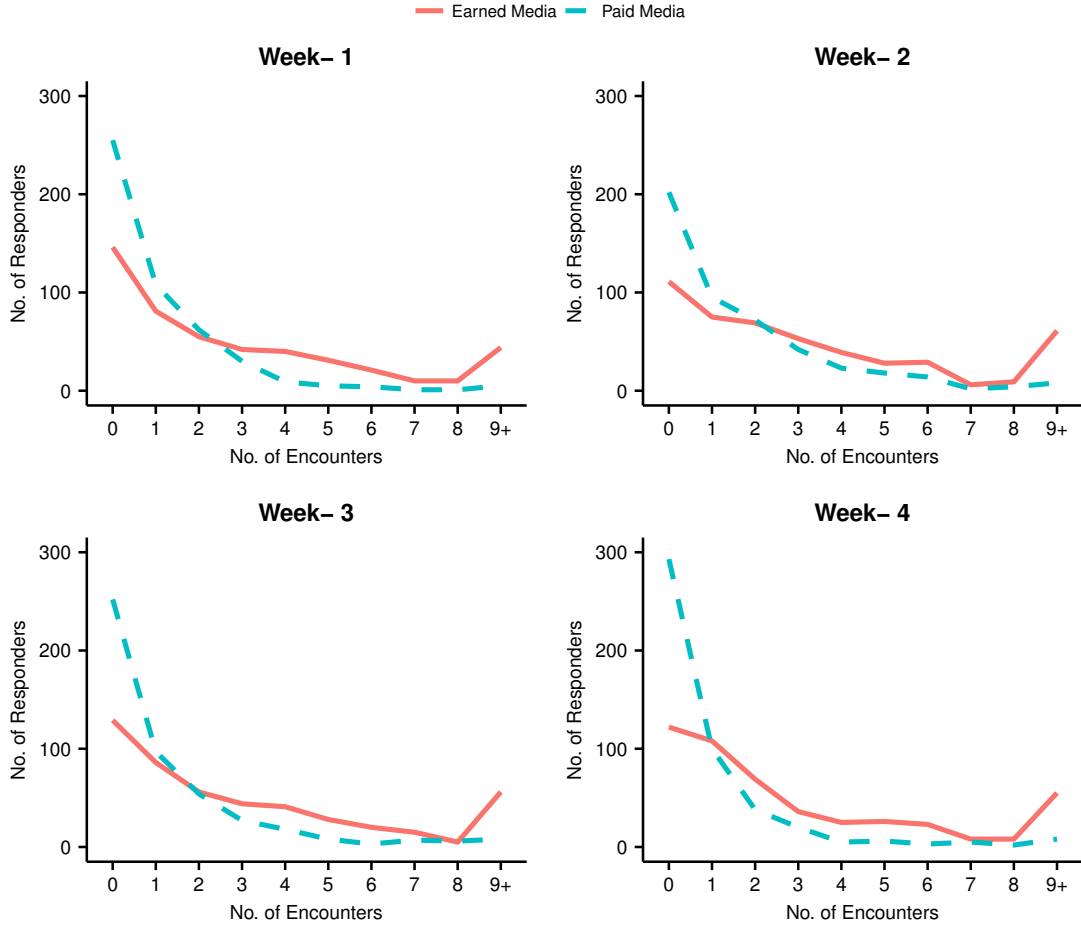


Figure 2: Reporting of paid and earned encounters each week. Earned encounters are encounters with the brand (party) which are not controlled by the political party itself, such as news and editorial comment. More people are exposed to earned communications.

choice intentions depend on exposures, we need to jointly model since these exposures are not exogenously determined.

3.1 Dynamic Multinomial Logit Model for Choice Intentions

To build the model for the intention to vote for a particular political party during any week of the campaign, we note that voters were faced with four primary choices, viz., ‘Labour’, ‘Conservatives’, ‘Liberal Democrats’ or ‘Do Not Intend to Vote’. The intention to vote was measured at the end of Weeks 1, 2 and 3 using the following question: “If you had to decide today which party you would vote for, which would it be?” The ‘None’ response in Weeks 1, 2, and 3 is explained by those who did not intend to vote or those who intended to vote for an alternative minor party. For

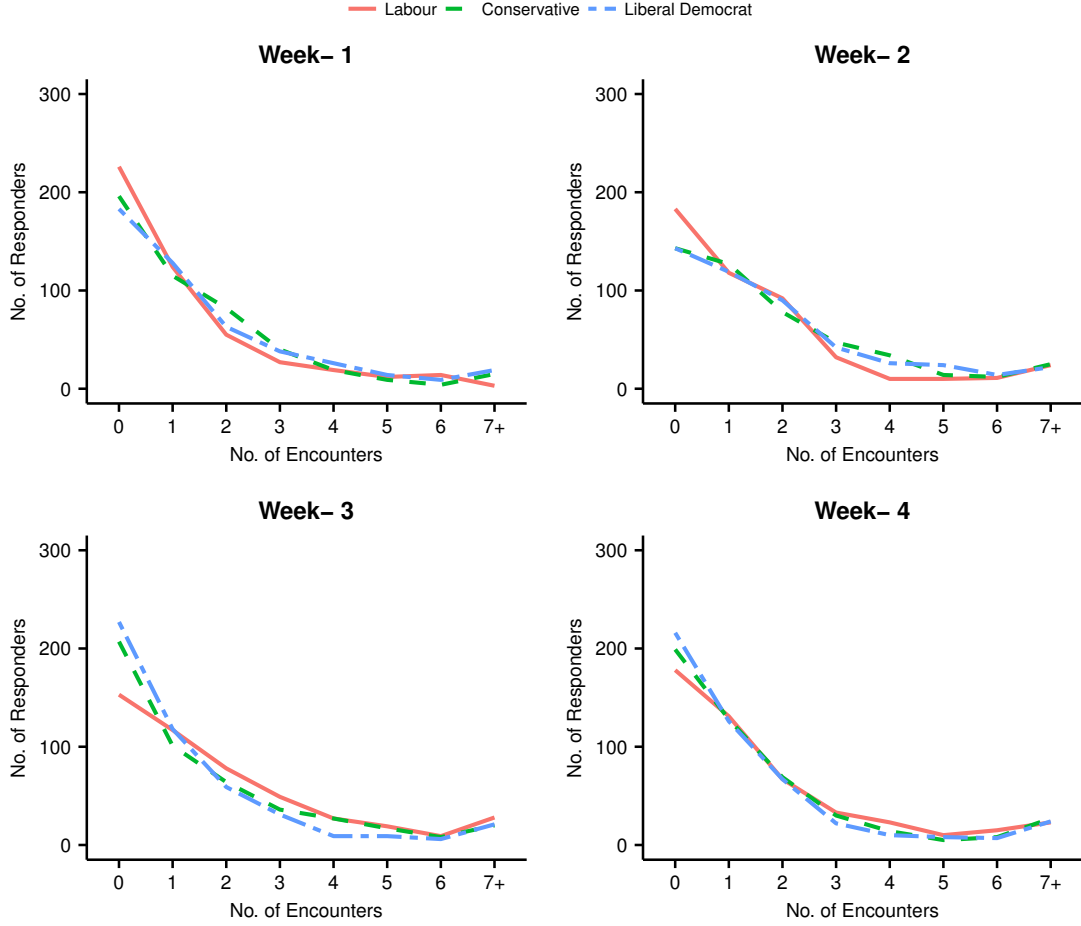


Figure 3: Reporting of encounter from each party each week.

the final vote, this last category had 84 participants, split as follows: 34 did not vote at all, and 50 voted for a minor party. That is, we found a total of 7% of the swing voters in our study did not make a final choice. These choices not only vary over time but they also have a carry over from the choices in the previous week, leading to an autocorrelation among the voters over the campaign time frame.

We model the choices with a flexible multinomial logit model and include a random voter effect to allow for the heterogeneity emanating from the choice behavior of individual voters.

Let k denote the choices available, where $k = 1$ is for No Intention to Vote, $k = 2$ is for Labour party, $k = 3$ is for Conservative party, and $k = 4$ is for Liberal Democrat party. Further, let $(y_{it} = k)$ denote voter i 's intention to choose party $k = (1, 2, 3, 4)$ in week t . We denote the random voter's heterogeneity effect for the i^{th} voter and k^{th} party as b_{ik} . Thus, conditional on b_{ik} , the response variable y_{it} has a multinomial logit model as described below:

$$y_{it}|b_{ik} \sim \text{Multinomial}(\pi_{it1}, \pi_{it2}, \pi_{it3}, \pi_{it4}), \quad (1)$$

where $\pi_{itk} = \text{Prob}(y_{it} = k|b_{ik})$ is the probability of choosing alternative k , conditional on the random effects.

Assuming ‘None’ as the baseline category, we can model the π_{itk} as,

$$\text{Prob}(y_{it} = k|b_{ik}) \propto \exp(U_{itk}) \quad (2)$$

The utility for a voter i , associated with choice k in the t^{th} week may now be described as follows:

$$\begin{aligned} U_{itk} = & \underbrace{x_{itk}^V \beta_t^V + (x_{itk}^V)^2 \mathbf{d}_t^V}_{\text{Avg Valence}} + \underbrace{x_{itk}^F \beta_t^F + (x_{itk}^F)^2 \mathbf{d}_t^F}_{\text{Frequency}} + \underbrace{\ell_1(x_{itk}^{F[\text{earned}]} \times x_{itk}^{F[\text{paid}]})}_{\text{frequency interaction}} + \underbrace{\ell_2(x_{itk}^{V[\text{earned}]} \times x_{itk}^{V[\text{paid}]})}_{\text{valence interaction}} \\ & + \underbrace{\mathbf{w}_i^T \gamma_k + \mathbf{z}_i^T \eta_k}_{\text{predictors}} + \underbrace{\rho_k^1 \mathbb{1}_{\{y_{i,t-1}=k\}} C_{i,t-1} + \rho_k^2 \mathbb{1}_{\{y_{i,t-1}=k\}}}_{\text{dynamic effects}} + \underbrace{\varrho \mathbb{1}_{\{y_{i,2005}=k\}}}_{\text{random effects}} + \underbrace{b_{ik}}_{\text{random effects}} + \epsilon_{itk} \end{aligned} \quad (3)$$

where ϵ_{itk} is an i.i.d extreme value type I distribution, x_{itk}^V and $(x_{itk}^V)^2$ are respectively the linear and quadratic effects of average valence with corresponding parameter being β_t^V and \mathbf{d}_t^V . Likewise, for frequency we have x_{itk}^F and $(x_{itk}^F)^2$ as the linear and quadratic effects with corresponding parameter being β_t^F and \mathbf{d}_t^F . The vector of choice varying predictors pertains to the total frequency (F) and average valence (V) of both paid and earned media exposures and captures the frequency of exposures/average valence to political campaign advertisements. As there is no a priori reason to assume a linear effect of exposure, and it may be more plausible to have higher-order effects of exposure as smaller exposures may be positive, and larger exposures may have a diminished or negative effect, we allow a nonlinear effect of exposure in the model (3). The ℓ_1 and ℓ_2 are coefficients of interaction between earned media and paid media for frequency and valence respectively. These coefficients will help us to determine whether the earned and paid media support each other or cannibalize each other’s effect. The \mathbf{w}_i includes demographic variables, where w_i =(Age, Social Class, Gender, Marital Status, Children, Employment Category, Education Category, Home Ownership), γ_k is the corresponding parameter. In addition, η_k is a vector of parameters corresponding to \mathbf{z}_i which captures the importance of key election issues as perceived by voters. The issues considered in these data are: Health, Economy, Law & Order, Tax & Spend, Climate Change, The Environment, and World Poverty. In general, choice models have no memory. However, the loyalty of a voter to a political party is usually consistent and thus it is important to assess the presence of any stickiness in their intentions to choose from previous periods. We formulate a dynamic model with the lagged term $(y_{i,t-1})$ incorporating the presence of any stickiness in voters’ choices from

previous periods. The ρ_k^2 captures the carry-over effect while the term $(\mathbb{1}_{\{y_{i,t-1}=k\}}C_{i,t-1})$ measures the lagged interaction between voting intentions and a stated measure of the confidence that voters have that their voting intentions will be the same as their actual vote ($C_{i,t-1}$). Here $\mathbb{1}_{\{y_{i,t-1}=k\}}$ is an indicator variable that equals 1 for those who intended to vote for party k . Similarly $\mathbb{1}_{\{y_{i,2005}=k\}}$ is an indicator function which represents if voter i voted for party k in the previous election in year 2005.

Some covariates, including voters' demographics, are included as fixed with respect to time in the models (data on these were collected once at the start of the survey). However, it is likely that only a very small proportion of the respondent's circumstances will have changed with regard to one or more of these variables (i.e. employment, housing type and education category) during the 4 week period of data collection. This is especially likely to be the case in relation to the education category as respondents were aged 18+ and changes in educational status at that age and beyond would be limited to a shift from school education to university undergraduate and postgraduate education, both of which finish in July and September in the UK and the data were collected in April/May. However, voting intention tends to be influenced more by voter psychographics than demographics (Newman and Sheth 1985; Baines et al. 2011), hence there may be a very small accounted-for volatility in these covariates and their effect is likely to be small.

An important feature of our model is the use of time-varying parameters, β_t^V , d_t^V , β_t^F , d_t^F which vary over each week of the campaign. There are several reasons why these coefficients might vary over time, e.g., due to changes in exposures, changes in the policy of the party, or occurrence of unexpected events which may lead the voters to think differently. Hence, it is necessary to develop an empirical model that would be able to absorb these changes, if any, and also be capable of determining the right parameter values in time to help the analyst predict the final outcome. In addition, time-varying coefficients improve the predictive power of a model (see West and Harrison 1997). Therefore, we assume that these parameters have a first-order random walk prior. Thus, for example, β_t^V will have a prior like:

$$\beta_t^V = \beta_{t-1}^V + e_t^V; \quad e_t^V \sim N(0, \tau I) \quad (4)$$

The precision parameter τ in the above distribution of β_t in equation (4) is a hyperparameter and takes an assumed value of 0.25 (Dunson et al. 2003).

3.2 Zero-Inflated Multinomial Logit Model for Final Vote

Since the final choice (vote) took place at the same time for all respondents, the data for the final vote is not time varying. As explained earlier, we model the decision of whether or not to vote with

a binary distribution and the conditional choice decision as a multinomial logit model, resulting in a zero-inflated multinomial model. Letting F_i be the final choice for the i th voter and p_i the probability of voting, we model the outcome of the final vote as:

$$F_i = \begin{cases} 0 & \text{with probability } (1 - p_i), \quad \text{for not voting} \\ G(F_i), & \text{with probability } p_i, \quad \text{for voting} \end{cases} \quad (5)$$

Combining the above two equations, we get the following distribution for the F_i :

$$\pi(F_i|p_i) = (1 - p_i)I_{\{0\}} + p_i G(F_i) \quad (6)$$

We conceptualize the degenerate distribution as representing a “reporting inactive or no vote” state of voter i with probability $(1 - p_i)$, while the probability mass function $G(F_i)$ represents a “reporting active or voting” state. In general when a voter does vote, F_i can take any of the values among $k = 1, 2, 3, 4$, with “1” representing the base category of “Other Party” and “2,3,4” denoting the three parties. Thus, $G(F_i)$ assumes a multinomial choice model. Therefore, the final voting model with predictors is:

Voting Probability:

$$\text{logit}(p_i) = \phi_1 + \phi_2 \times \text{Voted-2005}_i + \underbrace{\varsigma_3 \mathbb{1}_{\{y_{i3} \in (2,3,4)\}} + \varsigma_2 \mathbb{1}_{\{y_{i2} \in (2,3,4)\}} + \varsigma_1 \mathbb{1}_{\{y_{i1} \in (2,3,4)\}}}_{\text{voting intentions for major parties}} + \mathbf{w}_i^T \boldsymbol{\varpi}_k + \mathbf{z}_i^T \boldsymbol{\varkappa}_k \quad (7)$$

Final Vote Conditional on Voting:

$$G(F_i = k) \sim \text{Multinomial}(1; \Pi_{i1}, \Pi_{i2}, \Pi_{i3}, \Pi_{i4})$$

$$\Pi_{ik} \propto \exp \left(\mathbf{c}_{ik}^T \boldsymbol{\theta} + \mathbf{w}_i^T \boldsymbol{\Gamma}_k + \mathbf{z}_i^T \boldsymbol{\kappa}_k + u_i + \underbrace{a_3 \mathbb{1}_{\{y_{i3}=k\}} + a_2 \mathbb{1}_{\{y_{i2}=k\}} + a_1 \mathbb{1}_{\{y_{i1}=k\}}}_{\text{voting intentions}} + \tilde{a} \mathbb{1}_{\{y_{i,2005}=k\}} \right) \quad (8)$$

where \mathbf{c}_{ik}^T is the vector consisting of cumulative frequency and valence for both paid and earned exposures and interaction of paid media and earned media valence and frequency, i.e., $\mathbf{c}_{ik} = (c_{ikl}^F, c_{ikl}^V)$; ($l = 1(\text{paid media}), l = 2(\text{earned media})$), $\boldsymbol{\varpi}_k$ and $\boldsymbol{\varkappa}_k$ are the coefficients associated with the demographics (\mathbf{w}_i^T) and election issues (\mathbf{z}_i^T) respectively in the voting probability model, and $\boldsymbol{\Gamma}_k$ and $\boldsymbol{\kappa}_k$ are the corresponding coefficients in the final voting model. The ϕ_1 is the intercept and ϕ_2 is the coefficient associated with whether or not individuals voted in the year 2005 (i.e., the previous election). We also include the factor of whether the voter intended to vote for any major parties in previous weeks. The coefficients $\varsigma_3, \varsigma_2, \varsigma_1$ in equation (7) are the effects of intending to vote for major parties in week $t = 3, 2, 1$. We assume the effect decays as the week passes and

thus, we write $\varsigma_t = \varsigma \times \phi_3^{3-t}$. Here ς is the overall effect and ϕ_3 captures the decay effect. Note that in equation (8), the final vote is connected with voting intentions to better describe the sequential nature of the data generating process. The u_i represents the heterogeneity corresponding to the i th voter. The, a_t is the effect of the t^{th} week's voting intention on the final vote and since these effects will also have diminishing effects over time, we assume that $a_t = a \times \phi^{3-t}$. Here a is the overall effect and ϕ measures the decay over weeks.

3.3 Model for the Frequency and Valence of Communications Exposures

We include both the frequency and the average valence of communications exposures within this model to separately capture the effect of the number of exposures and the encounter valence in the model. Frequency of communications exposure may impact on voter choice in several ways. First, repeated messages of the same type (whether or not the messages are identical) can increase voters' awareness and hence intentions to vote (Yaveroglu and Donthu 2008). Second, any associations that are influenced by the message may grow in attitude strength through familiarity (Erdem and Keane 1996). These party associations may of course be negative as well as positive: for example, inconsistency in the message across communications exposures can negatively impact party loyalty (Erdem and Keane 1996). Third, repetition can increase learning of cognitive message content (Goh et al. 2011).

The valence of communications may also have an impact on the choice of the party. Empirically, this has been observed both in the paid communications context of advertisements and with earned communications such as word-of-mouth (WOM) (Berger et al. 2010), though not to our knowledge in both at the same time. Mechanisms for this may include increased attention when affect is positive (Baumeister et al. 2007), increased source credibility when the source is liked (Kahn and Isen 1993), and direct approach/avoidance effects of some positive and negative emotions (Romani et al. 2012). As these mechanisms are not identical to those by which frequency impacts party choice, it seems plausible that frequency and valence effects may both be present but working in opposite directions. This is supported by Berger et al. (2010), who find that negative book reviews have a negative effect for well-known authors but a positive effect for unfamiliar ones, due to increased awareness.

The exposure frequency variable, i.e., x_{itkl}^F in equation (3), reflects the total frequency of exposures of the i th voter to the campaign of the k th party in the t th week through media l ($= 1$ for paid and $= 2$ for earned exposures). Consistent with the literature, not all voters are exposed to every medium each week, resulting in an excessive number of zero observations for the exposure

frequency. Since the presence of these excessive zeros leads to spurious over-dispersion (Park et al. 2011), we fit a zero-inflated Poisson (ZIP) regression model which is a discrete mixture of true zeros (degenerate distribution) and positive counts for frequency of exposures (Lambert 1992). The zero inflated model overcomes the limiting assumption implicit in basic count models that every voter has some probability of being exposed to communications.

Thus we model the frequency of exposures x_{itkl}^F as:

Exposure Frequency:

$$P(x_{itkl}^F = x) = (1 - \alpha_{il}^F) \mathbb{1}_{\{x_{itkl}^F=0\}} + \alpha_{il}^F \times \text{Poisson}(\lambda_{itkl}) \quad (9)$$

where, $0 < \alpha_{il}^F < 1$ is the probability that the i^{th} person is exposed to the l^{th} category and the total number of exposures is a count process with rate parameter λ_{itkl} . We further assume that:

$$\log(\lambda_{itkl}) = \mathbf{w}_i^T \zeta_l^{F1} + \zeta^{F2} \mathbb{1}_{\{y_{i,t-1}=k\}} C_{i,t-1} + \zeta^{F3} \mathbb{1}_{\{y_{i,t-1}=k\}} + h_{il}^F. \quad (10)$$

Here, the \mathbf{w}_i are the demographic variables described earlier, the h_{il}^F is the unobserved heterogeneity due to voter choice on the media l . We assume a Beta(a, b) prior for the probability α_{il}^F .

Finally, we also need to model average valence due to exposures (x_{itkl}^V). Note that although exposure valence is measured on a five-point scale, in our data we have these scores averaged over each week, which makes x_{itkl}^V a continuous variable. Exposure valence also encounters many zeros and thus we fit a zero-inflated log-normal model for exposure valence with the mean dependent on the demographics.

Exposure Valence:

$$\begin{aligned} x_{itkl}^V &= (1 - \alpha_{il}^V) \mathbb{1}_{\{x_{itkl}^V=0\}} + \alpha_{il}^V \times LN(\mu_{itkl}, \tau_2^2) \mathbb{1}_{\{x_{itkl}^V>0\}} \\ \mu_{itkl} &= \mathbf{w}_i^T \zeta_l^{V1} + \zeta^{V2} \mathbb{1}_{\{y_{i,t-1}=k\}} C_{i,t-1} + \zeta^{V3} \mathbb{1}_{\{y_{i,t-1}=k\}} + h_{il}^V \end{aligned} \quad (11)$$

We assume a Beta(a, b) prior for the probability α_{il}^V .

3.4 Correlation Structure and Heterogeneity: Combining the Models

All the three models described above carry information about the voting pattern of the voters and are therefore inter-related. To get the complete picture and to account for the heterogeneity across voters' outcomes, one must combine these effects by correlating the multiple outcomes. However, since these outcomes are measured on a variety of different scales (viz., multinomial,

ZIP), it is not possible to directly model the joint predictors' effects due to the lack of any natural multivariate distribution for characterizing such a dependency. Moreover, without inter-relating or jointly considering these outcomes, it is hard to answer questions about how the evolution of one response (e.g., intention to vote) is related to the evolution of another (e.g., final vote). In such cases, it is natural to consider models where the dependency among the responses may be incorporated via the presence of one or more latent variables (Dunson 2003; Christel et al. 2008).

A flexible solution is to model the association between different responses by correlating the random heterogeneous effects from each of the responses. In our joint modeling approach, random effects are assumed for each response process and the different processes are associated by imposing a joint multivariate distribution on the random effects. Such a model not only gives us a covariance structure to assess the strength of association between the responses, but also borrows information across the communications types and offers an intuitive way of describing the dependency between the responses. Accordingly, we account for this heterogeneity by assuming correlated random effects, i. e., $\mathbf{V}_i = (b_{i2}, b_{i3}, b_{i4}, u_i, h_{i1}^F, h_{i2}^F, h_{i1}^V, h_{i2}^V)^T$:

$$\mathbf{V}_i \sim N_8(\mathbf{0}, \Sigma) \quad (12)$$

where Σ is the variance-covariance matrix.

There is an additional advantage of having correlated random effects. Due to the IIA property of the multinomial logistic model, it is not possible to find cross effects it is not possible to incorporate cross effects (e.g., how a negative exposure to Labour may affect the voting intention for the Liberal Democrats) directly in the utility equation (3). We do address the issue implicitly in our model following the suggestion of Brownstone et al. (1999), by assuming individual specific correlated random effects across choices (also across exposures).

4 Bayesian Inference

4.1 Likelihood

Let $\Omega = (\beta_t, \mathbf{d}_t, \ell_1, \ell_2, \gamma_k, \eta_k, \rho_k^1, \rho_k^2, \phi_1, \phi_2, \varpi_k, \varkappa_k, \theta, \mathbf{\Gamma}_k, \kappa_k, \delta, \alpha_{il}, \zeta_k^{F,V}, \tau_1^2, \tau_2^2, a_1, a_2, a_3, \mathbf{V}_i, \Sigma)$ be the set of all parameters. To obtain the likelihood function for **voting intentions**, we multiply equation

(2) across the voters, week and choice:

$$f(y|\Omega, \mathbf{v}_i, \mathbf{x}_{it}) \propto \prod_{i,t,k} \left[\frac{\exp(U_{itk})}{\sum_{k=1}^4 \exp(U_{itk})} \right]^{\mathbb{1}_{\{y_{i,t}=k\}}},$$

where $U_{itk} = x_{itk}^V \beta_t^V + (x_{itk}^V)^2 \mathbf{d}_t^V + x_{itk}^F \beta_t^F + (x_{itk}^F)^2 \mathbf{d}_t^F + \ell_1(x_{itk}^{F[earned]} \times x_{itk}^{F[paid]}) + \ell_2(x_{itk}^{V[earned]} \times x_{itk}^{V[paid]})$

$$+ \mathbf{w}_i^T \gamma_k + \mathbf{z}_i^T \eta_k + \rho_k^1 \mathbb{1}_{\{y_{i,t-1}=k\}} C_{i,t-1} + \rho_k^2 \mathbb{1}_{\{y_{i,t-1}=k\}} + \varrho \mathbb{1}_{\{y_{i,2005}=k\}} + b_{ik} \quad (13)$$

and, $\mathbb{1}_{\{y_{i,t}=k\}} = 1$ if the individual i chooses alternative k at week t and zero otherwise.

Similarly, the likelihood function for the **final vote** is given by :

$$f(F|\Omega, \mathbf{v}_i, \mathbf{x}_{it}) \propto \prod_{i,k} \left[\frac{1}{1 + \exp(\phi_1 + \phi_2 \times \text{Voted-2005}_i + \sum_t (\varsigma_t) + \mathbf{w}_i^T \varpi_k + \mathbf{z}_i^T \varkappa_k)} \right]^{\mathbb{1}_{\{F_i=0\}}} \quad (14)$$

$$\times \left[\frac{\exp(\phi_1 + \phi_2 \times \text{Voted-2005}_i + \sum_t (\varsigma_t) + \mathbf{w}_i^T \varpi_k + \mathbf{z}_i^T \varkappa_k)}{1 + \exp(\phi_1 + \phi_2 \times \text{Voted-2005}_i + \sum_t (\varsigma_t) + \mathbf{w}_i^T \varpi_k + \mathbf{z}_i^T \varkappa_k)} \right]^{\mathbb{1}_{\{F_i \neq 0\}}}$$

$$\times \left[\frac{\exp(\mathbf{c}_{ik}^T \theta + \mathbf{w}_i^T \Gamma_k + \mathbf{z}_i^T \kappa_k + u_i + \sum_t (a_t) + \tilde{a} \mathbb{1}_{\{y_{i,2005}=k\}})}{\sum_{k=1}^4 \exp(\mathbf{c}_{ik}^T \theta + \mathbf{w}_i^T \Gamma_k + \mathbf{z}_i^T \kappa_k + u_i + \sum_t (a_t) + \tilde{a} \mathbb{1}_{\{y_{i,2005}=k\}})} \right]^{\mathbb{1}_{\{F_i \neq 0 \text{ \& } F_i=k\}}}$$

where Voted-2005_i represents whether the individual voted in the 2005 UK general election (from equation (7)); $\sum_t (\varsigma_t) = \varsigma_3 \mathbb{1}_{\{y_{i3} \in (2,3,4)\}} + \varsigma_2 \mathbb{1}_{\{y_{i2} \in (2,3,4)\}} + \varsigma_1 \mathbb{1}_{\{y_{i1} \in (2,3,4)\}}$ and $\sum_t (a_t) = [a_3(y_{i3} = k) + a_2(y_{i2} = k) + a_1(y_{i1} = k)]$.

Next, the likelihood function for **exposure frequency** is:

$$f(x_{itkl}^F|\Omega, \mathbf{V}_i) \propto \prod_{i,t,k,l} \left[(1 - \alpha_{il}^F) + \alpha_{il}^F e^{-\lambda_{itkl}} \right]^{1 - \mathbb{1}_{\{x_{itkl}^F=0\}}} \times \left[\frac{\alpha_{il}^F e^{-\lambda_{itkl}} \lambda_{itkl}^{x_{itkl}^F}}{x_{itkl}^F!} \right]^{\mathbb{1}_{\{x_{itkl}^F > 0\}}} \quad (15)$$

where α_{il}^F and λ_{itkl} are given in equation (10).

Finally, the likelihood function for **average valence** is:

$$f(x_{itkl}^V|\Omega, \mathbf{V}_i) \propto \prod_{i,t,k,l} (1 - \alpha_{il}^V)^{1 - \mathbb{1}_{\{x_{itkl}^V=0\}}} \times \left(\alpha_{il}^V \times LN(\mu_{itkl}, \tau_2^2) \right)^{\mathbb{1}_{\{x_{itkl}^V > 0\}}} \quad (16)$$

Combining these equations, the joint likelihood function of our proposed model is given by:

$$L|\Omega, \mathbf{V}_i = \text{Equation (13)} \times \text{Equation (14)} \times \text{Equation (15)} \times \text{Equation (16)} \quad (17)$$

$$\times f(\mathbf{V}_i|\Sigma) \times \prod_t f(\beta_t^V) \times \prod_t f(\beta_t^F) \times \prod_t f(d_t^V) \times \prod_t f(d_t^F)$$

4.2 Prior Specification and Posterior Inference

We estimate the model parameters using a Bayesian framework. We now describe the prior distributions for the model parameters. We use a zero-mean prior because we are uncertain of the expected value of the parameter. A similar approach has been used in mixed effects model settings (Natarajan and McCulloch 1998). Since our model has a logistic regression structure, the inferences are mainly based on the odds-ratio. Thus, for the regression parameters we assume the prior $N(0, \text{var} = 4)$. This prior implies that the odds ratio (centered at 1) has a 95% interval between $\exp(-4)$ to $\exp(4)$, which has a very wide range and is weakly informative (Dunson et al. 2003; Gelman et al. 2008). For the variance-covariance matrix Σ , we assumed inverse-Wishart($6, 0.1 * I_{4 \times 4}$) as the prior distribution. For the variance parameters, we assume inverse-Gamma (IG) priors of $IG(2.01; 1.01)$, giving rise to a prior mean of 1 and a prior variance of 100 (Ghosh and Tu 2009). This is also vaguely informative. These vaguely informative priors ensure proper posteriors and also allow the data, rather than the assumptions, to dominate the inferences. We assume a Beta(1, 1) prior for α_{it}^F and α_{it}^V .

Posterior inferences were obtained using the MCMC sampler, implemented in the freely available software Bayesian Inference Using Gibbs Sampling, WinBUGS (Spiegelhalter et al. 2005). The sampler from the posterior obtained from the Markov Chain Monte Carlo (MCMC) simulation allowed us to achieve a summary measure of the parameter estimates. For all posterior computations, MCMC simulation was implemented with an initial burn-in of 5000 iterations and a total of 10000 iterations, thinning out in batches of 10. This worked well, with convergence diagnostics using the R package coda showing no evidence of practical convergence problems.

4.3 Model Selection and Model Fit

Before discussing our results, we first compare our proposed model with some alternative models to test the quality of model fit. In order to assess our model, we compare it with a variety of different nested models as follows:

- Model-1: Our proposed dynamic joint estimation model (as outlined in section 3), but with only linear exposure effects (that is, without the quadratic terms in equation 3)
- Model-2: Our proposed dynamic joint estimation model with linear and quadratic exposure effects (that is, the full model described in section 3)
- Model-3: A static model with no dynamic or time-dependent parameter effects in the model

- Model-4: A model which has all the components as independent (no joint model)

We apply deviance information criteria (DIC) proposed by Spiegelhalter et al. (2002), Log-pseudo marginal likelihood (LPML) and Bayesian p-value to determine the best model. In a complicated data likelihood with shared parameters as in our case, DIC4 seems to be a better method (Celeux et al. 2006, Jiang et al. 2015) among the other variants of DICs such as “Complete DICs” and “Conditional DICs”. To compute the DIC4 from MCMC output, we follow the paper by Jiang et al. (2015), which has shown explicitly how to compute DIC4 from the MCMC output when there are latent variables and the posterior does not have a closed form. We also compute Log-Pseudo Marginal Likelihood (LPML) as an additional model selection criterion and Posterior Predictive P-value for model fit.

Let $\mathbf{D} = (y, F, x^F, x^V)$ be the observed data, θ be the set of parameters and \mathbf{b} the set of latent random effects variables.

DIC in its basic form is defined by

$$\text{DIC}(\mathbf{D}) = \overline{\text{D}(\theta)} + \mathbf{p}_{\mathbf{D}} = -4\mathbf{E}_{\theta}[\log \mathbf{p}(\mathbf{D}|\theta)|\mathbf{D}] + 2 \log \mathbf{p}(\mathbf{D}|\mathbf{E}_{\theta}(\theta|\mathbf{D})).$$

However, in our setting, with latent variable \mathbf{b} , $p(\mathbf{D}|\theta)$ is not in a closed form. Hence, we follow the approach in Jiang et al. (2015), and Celeux et al. (2006), and calculate $\text{DIC}(\mathbf{D})$, by first considering the DIC measure with “complete data” with \mathbf{b} and then integrating out the unobserved \mathbf{b} .

$$E_{\mathbf{b}}\{\text{DIC}(\mathbf{D}, \mathbf{b})\} = -4E_{\theta}[\log p(\mathbf{D}, \mathbf{b}|\theta)|\mathbf{D}, \mathbf{b}] + 2 \log p(\mathbf{D}, \mathbf{b}|E_{\theta}(\theta|\mathbf{D}, \mathbf{b})).$$

Integrating out \mathbf{b} gives

$$\text{DIC4} = \text{DIC}(\mathbf{D}) = E_{\mathbf{b}}[-4E_{\theta}[\{\log p(\mathbf{D}, \mathbf{b}|\theta)|\mathbf{D}\} + 2 \log\{(\mathbf{D}, \mathbf{b}|E_{\theta}(\theta|\mathbf{D}, \mathbf{b}))\}] \quad (18)$$

$$= -4E_{\mathbf{b}, \theta}[\log p(\mathbf{D}, \mathbf{b}|\theta)|\mathbf{D}] + 2E_{\mathbf{b}}[\log\{(\mathbf{D}, \mathbf{b}|E_{\theta}(\theta|\mathbf{D}, \mathbf{b}))\}|\mathbf{D}]. \quad (19)$$

where integration over \mathbf{b} is obtained via numerical methods (Jiang et al. 2015). The smaller the DIC4 values the better the model is.

In addition to the DIC4 measure, we also we compute $p(\mathbf{y}_i, \mathbf{F}_i, \mathbf{x}_i^F, \mathbf{x}_i^V | \mathbf{y}_{-i}, \mathbf{F}_{-i}, \mathbf{x}_{-i}^F, \mathbf{x}_{-i}^V)$ (Geisser and Eddy 1979), which is the posterior predictive density of $(\mathbf{y}_i, \mathbf{F}_i, \mathbf{x}_i^F, \mathbf{x}_i^V)$ for subject i conditional on the observed data with a single data point deleted. This value is known as the conditional

predictive ordinate (CPO; Gelfand et al. (1992); Jiang et al. (2015)) and has been widely used for model diagnostics and assessment.

For the i th subject, the CPO statistic under model is defined as:

$$\begin{aligned} \text{CPO}_i &= p(\mathbf{y}_i, \mathbf{F}_i, \mathbf{x}_i^F, \mathbf{x}_i^V | \mathbf{y}_{-i}, \mathbf{F}_{-i}, \mathbf{x}_{-i}^F, \mathbf{x}_{-i}^V) \\ &= E_{\boldsymbol{\theta}}[p(\mathbf{y}_i, \mathbf{F}_i, \mathbf{x}_i^F, \mathbf{x}_i^V | \boldsymbol{\theta}) | \mathbf{y}_{-i}, \mathbf{F}_{-i}, \mathbf{x}_{-i}^F, \mathbf{x}_{-i}^V] \end{aligned}$$

where $-i$ denotes the exclusion of the data from subject i . $p(\mathbf{y}_i, \mathbf{F}_i, \mathbf{x}_i^F, \mathbf{x}_i^V | \boldsymbol{\theta})$ is the sampling density of the model evaluated at the i th observation. The expectation above is taken with respect to the posterior distribution of the model parameters $\boldsymbol{\theta}$ given the cross-validated data $(\mathbf{y}_{-i}, \mathbf{F}_{-i}, \mathbf{x}_{-i}^F, \mathbf{x}_{-i}^V)$. For subject i , the CPO_i can be obtained from the MCMC samples by computing the following weighted average:

$$\widehat{\text{CPO}}_i = \left(\frac{1}{M} \sum_{m=1}^M \frac{1}{f(\mathbf{y}_i, \mathbf{F}_i, \mathbf{x}_i^F, \mathbf{x}_i^V | \boldsymbol{\theta}^{(m)})} \right)^{-1}$$

where M is the number of simulations. $\boldsymbol{\theta}^m$ denotes the parameter samples at the m th iteration. A large CPO value indicates a better fit. A useful summary statistic of the CPO_i is the logarithm of the pseudomarginal likelihood (LPML), defined as $\text{LPML} = \sum_{i=1}^n \log(\widehat{\text{CPO}}_i)$. Models with greater LPML values represent a better fit.

To assess the goodness of fit of the models, we also computed the Bayesian p-value/ posterior predictive p-value (ppp, Gelman et al. (2004), Meng (1994)), which measures the discrepancy between the data and the model by comparing a summary χ^2 statistic of the posterior predictive distribution to the true distribution of the data. The summary statistics from the predicted and observed data are given by $\chi^2(\mathbf{y}, \theta^g)$ and $\chi^2(\mathbf{y}^{rep,g}, \theta^g)$, respectively, where $\mathbf{y}^{rep,g}$ denote the replicated value of \mathbf{y} from the posterior predictive distribution of θ at the g th iteration of the Gibbs sampler. The Bayesian p-value was then calculated as $P(\chi^2(\mathbf{y}^{rep,g}, \theta^g) > \chi^2(\mathbf{y}, \theta^g))$, that is, the proportion of times $\chi^2(\mathbf{y}^{rep,g}, \theta^g)$ exceeds $\chi^2(\mathbf{y}, \theta^g)$ out of $g = 1, 2, \dots, G$ simulated draws from the posterior predictive distribution. A ppp value close to 0.5 is considered to be a good fit, as then the observed pattern would likely be seen in replications of the data under the true model.

Model	DIC4	LPML	ppp
Model-1	3908	-1772	0.47
Model-2	3754	-1563	0.51
Model-3	3978	-1850	0.45
Model-4	4162	-1900	0.58

From the above results, we see that for our proposed model (model-2) the DIC4 value is the lowest and the LPML value is the highest. Also, the ppp value is close to 0.5. Thus, among the models considered, our model is the best in terms of model selection criteria.

5 Empirical Results

5.1 Results for the The Impact of Paid and Earned Media Exposures on Voting Intention

In this section, we describe results pertaining to Equations 1 to 3 (detailed in Section 3.1) including Equation 3, the model for voting intentions.

Table 2 depicts the results of the lagged variables on the voting intentions model. The coefficients for the lagged choice of voters, ρ_k^2 are positive and significant for all parties, implying that the voters for these parties are sticky in the sense they are loyal to their previous choices and intend to repeat their choice. The coefficients ρ_k^1 are also positive and significant, implying that the interaction of voting intentions and a stated measure of confidence is also sticky. The other variable ρ is also positive and significant, implying that if an individual voted for a party in the previous election, he/she is more likely to intend to vote for the same party in this election.

Table 2: Parameter Estimates-Dynamic Coefficients (Intention to Vote Model).

Covariates		Parameter	mean	2.5%	97.5%
Confidence in Voting Intentions (lagged)	Labor	ρ_1^1	0.127	0.023	0.227
	Conservative	ρ_2^1	0.142	0.064	0.216
	Liberal Democrat	ρ_3^1	0.113	0.021	0.196
Voting Intentions (lagged)	Labor	ρ_1^2	1.404	0.731	2.064
	Conservative	ρ_2^2	2.112	1.616	2.647
	Liberal Democrat	ρ_3^2	1.172	0.607	1.820
Party Voted in 2005		ρ	0.717	0.479	0.961

Note: Emboldened variables are significant.

Table 3 depicts the impact of the frequency and valence of both paid and earned media exposures in the intention to vote model. In the case of both frequency and valence, we estimated both linear

and quadratic effects on intention to vote. The coefficients β_{1t}^V capture the linear impact of the voter's affective response to communications with political parties via various earned media, like debates, TV news, newspaper articles and radio news. Likewise d_{1t}^V capture the quadratic impact of earned media valence. Similarly, β_{2t}^V and d_{2t}^V correspond to the linear and quadratic impacts of the exposure valence of paid communications, i.e. exposures with political parties via their own posters, leaflets and party election broadcasts. We find that the valence of both paid and earned media has a positive and significant impact on voting intentions (with the exception of earned media valence for week 2 and paid media valence for week 1, which are non-significant). Furthermore, with the same exceptions, there is an additional positive quadratic effect of valence. This is consistent with psychology research showing that peaks in affective experience have a disproportionate effect on overall retrospective evaluations (Fredrickson 2000).

However, none of the frequency terms in Table 3 is significant, whether for earned media ($\beta_{1,t}^F$ and $d_{1,t}^F$) or for paid media ($\beta_{2,t}^F$ and $d_{2,t}^F$). That is, the frequency of exposures does not significantly affect intention to vote. This is most interesting, as a pervasive assumption in political marketing practice is the need to constantly reinforce a party's messages as an election approaches (O'Cass 2002; Klein and Ahluwalia 2005). Furthermore, this assumption was present among the participants in the 2010 UK election (Worcester et al. 2011). In support of this assumption about the importance of frequency is work on commercial brands suggesting that a higher frequency can help with attitude accessibility (Fazio and Williams 1986) and hence behaviour, via the theory of planned behavior (Herr and Fazio 1993). However, this effect is more prominent with less well known brands, where attitudes are not already accessible (Goh et al. 2011). For better known brands, which would certainly include the three main parties in the UK, exposure frequency of both paid media (Yaveroglu and Donthu 2008) and earned media (Stephen and Galak 2012) can cease to be a positive factor. Indeed, contrary to the assumption of the benefits of repeated exposure, some studies have reported advertisement wearout effects, where the impact of repeated advertisement exposures can tail off or even become negative (Bass et al. 2007). We find that the same can apply to earned media. Table 3 suggests that during the final weeks before an election, voters were thoroughly saturated with the parties' political messages, and that they continued to be influenced not by their quantity but only by their quality, as assessed through exposure valence.

This raises the question whether exposure wearout, a phenomenon reported in empirical work which only measures exposure frequency (Bass et al. 2007), also applies to valence. That is, does a high (or low) valence of exposures have a lesser impact on voters as the voting choice approaches? In the case of earned media, there is some indication that this is so, with parameter values somewhat higher in week 1 than weeks 2 and 3 (compare $\beta_{1,1}^V$ with $\beta_{1,2}^V$ and $\beta_{1,3}^V$; similarly,

Table 3: Parameter Estimates for Exposures (Intention to Vote Model)

Covariates		Parameter	Mean	2.50%	97.50%
Earned Media Exposure Valence Linear	Week 1	$\beta_{1,1}^V$	1.022	0.300	2.8
	Week 2	$\beta_{1,2}^V$	-0.125	-0.569	0.497
	Week 3	$\beta_{1,3}^V$	0.395	0.0131	1.251
Earned Media Exposure Valence Quadratic	Week 1	$d_{1,1}^V$	0.322	0.238	0.914
	Week 2	$d_{1,2}^V$	0.005	-0.061	0.167
	Week 3	$d_{1,3}^V$	0.154	0.046	0.3
Paid Media Exposure Valence Linear	Week 1	$\beta_{2,1}^V$	0.81	-0.0218	1.507
	Week 2	$\beta_{2,2}^V$	1.12	0.466	3.191
	Week 3	$\beta_{2,3}^V$	1.355	0.288	3.755
Paid Media Exposure Valence Quadratic	Week 1	$d_{2,1}^V$	0.186	-0.012	0.355
	Week 2	$d_{2,2}^V$	0.251	0.102	0.751
	Week 3	$d_{2,3}^V$	0.074	-0.095	0.271
Earned Media Exposure Frequency Linear	Week 1	$\beta_{1,1}^F$	-0.967	-3.851	1.33
	Week 2	$\beta_{1,2}^F$	-1.312	-3.819	0.922
	Week 3	$\beta_{1,3}^F$	-1.49	-4.02	0.784
Earned Media Exposure Frequency Quadratic	Week 1	$d_{1,1}^F$	-0.530	-1.676	0.418
	Week 2	$d_{1,2}^F$	-0.454	-1.478	0.472
	Week 3	$d_{1,3}^F$	-0.786	-1.719	0.181
Paid Media Exposure Frequency Linear	Week 1	$\beta_{2,1}^F$	-1.222	-4.041	1.583
	Week 2	$\beta_{2,2}^F$	-2.180	-5.067	0.724
	Week 3	$\beta_{2,3}^F$	-0.089	-2.444	2.059
Paid Media Exposure Frequency Quadratic	Week 1	$d_{2,1}^F$	-0.486	-1.669	0.706
	Week 2	$d_{2,2}^F$	-0.992	-2.298	0.314
	Week 3	$d_{2,3}^F$	-0.402	-1.361	0.386
Interaction (Paid *Earned Media)	Valence	ℓ_1	0.146	0.0852	0.3445
	Frequency	ℓ_2	-0.4822	-2.031	0.452

Note: Emboldened variables are significant.

compare $d_{1,1}^V$ with $d_{1,2}^V$ and $d_{1,3}^V$). No such pattern is evident with paid media, however. Decision-making research suggests a possible mechanism for such a difference between earned and paid media in wearout of exposure valence. Object-evaluation associations such as an attitude towards a party are more likely to guide intentions and behaviour when these associations are strong and accessible (Fazio and Williams 1986), hence the strong role of valence overall in Table 3. Once an attitude is formed, individuals are motivated to deliberate as to its validity when approaching a high-involvement decision (Herr and Fazio 1993). This deliberation may involve a greater willingness to pay attention to direct communications from the brand- in this case, the party (Baxendale, Macdonald and Wilson 2015). Further work is required to check these observations about valence wearout, which ideally would be studied over longer periods and in alternative decision contexts.

The significance of the interaction between paid and earned media valence shows that positively received (and indeed negatively received) media exposures from these different sources are mutually reinforcing in relation to voting intention. Where both paid and earned media give a similarly positive (or a similarly negative) impression, this presumably adds to the voter’s confidence in their assessment. Such synergies have been observed in online versus offline paid media (Naik and Peters 2009); we find a similar phenomenon in paid versus earned media.

Parameter estimates for the impact of demographics and socio-political issues on voting intentions also provide some interesting insight. For example, while social class and gender do not impact voting, we find that age has a negative and significant impact on voting intentions for Labour and the Liberal Democrat parties. Those concerned with climate change were less likely to vote for any of the main parties, perhaps due to mistrust of the establishment among these voters due to the UK’s “Climategate” controversy. For these socio-demographic results and their interpretation, see Appendix A1.

5.1.1 Impact of Increased Exposure Frequency and Valence

Having established the importance of these media effects, we assess the impact of marginal increases in exposure frequency or valence on choice behavior. Let i , k , and t denote individuals, brand choices and time respectively; $k = 1, 2, 3$ and $t = 1, 2, 3$ (as in the voting data) and $Y_{i,t}$ be the choice of i^{th} individual at time t . Individual probabilities of choice are represented as:

$$Y_{i,t} \sim \text{Multinomial}(p_{i,t}^k, k = 1, 2, 3 \text{ brands}) \quad (20)$$

where

$$p_{i,t}^k = \frac{q_{i,t}^k}{q_{i,t}^1 + q_{i,t}^2 + q_{i,t}^3} \quad (21)$$

and

$$q_{i,t}^k = \exp \left(a_0 + \beta_{1,t}^V \cdot V_{i,t,k}^1 + \beta_{2,t}^V \cdot V_{i,t,k}^2 + d_{1,t}^V \cdot (V_{i,t,k}^1)^2 + d_{2,t}^V \cdot (V_{i,t,k}^2)^2 + \beta_{1,t}^F \cdot F_{i,t,k}^1 + \beta_{2,t}^F \cdot F_{i,t,k}^2 + d_{1,t}^F \cdot (F_{i,t,k}^1)^2 + d_{2,t}^F \cdot (F_{i,t,k}^2)^2 + \ell_1 (V_{i,t,k}^1 \cdot V_{i,t,k}^2) + \ell_2 (F_{i,t,k}^1 \cdot F_{i,t,k}^2) + \gamma_1 Z + b_{i,1} \right) \quad (22)$$

$V_{i,t,k}^l$ is average valence of individual i after encountering $F_{i,t,k}^l$ number of encounters via media l at time t of brand k . Z are other control variables and b are random effects. Political parties are particularly interested in knowing the impact of improvements in their communications on their share of vote. Since our data pertain to exposure frequencies and their average valence, we can compute the change in likelihood of choosing party k if they encounter $x\%$ more exposures from paid and earned media for party k at time t :

$$\frac{\partial p_{i,t}^k}{\partial F_{i,t,k}^1} = p_{i,t}^k (1 - p_{i,t}^k) (\beta_{1,t}^F + 2 \cdot d_{1,t}^F \cdot F_{i,t,k}^1 + \ell_2 \cdot F_{i,t,k}^2) \quad (23)$$

$$\frac{\partial p_{i,t}^k}{\partial F_{i,t,k}^2} = p_{i,t}^k (1 - p_{i,t}^k) (\beta_{2,t}^F + 2 \cdot d_{2,t}^F \cdot F_{i,t,k}^2 + \ell_2 \cdot F_{i,t,k}^1) \quad (24)$$

Table 4: Average Frequency and Average Valence of Encounters

For Avg Encounters	Earned Media		Paid Media	
	Frequency	Avg Valence	Frequency	Avg Valence
Week 1	2.859	3.000	0.875	2.986
Week 2	3.032	3.000	1.192	2.998
Week 3	3.046	3.143	1.071	3.043

The impact for changes in valence can be assessed in a similar manner. We can compare effects of changes in exposure frequencies using $\beta_{1,t}^F$ and $\beta_{2,t}^F$ and exposure frequencies for other media as well as compare the values of $\beta_{i,t}^F$ for different values of t . The sensitivity depends on the probability of choosing a brand and the number of exposure across different media. The results of this exercise are shown below for the average level of exposure in our sample (Table 4, 5). For illustrative purposes, we show the impact of a small increase in earned and paid media frequency and valence for the average voter with 30%, 50% and 80% probability of voting for party 3 (Liberal Democrats) at one, two and three weeks prior to the final vote. Similar calculations can be made to

assess smaller or larger increases in earned or paid exposures using equations (23) and (24). Our model shows that a small increase in the frequency of paid and earned media exposures had the greatest impact in week 3. The marginal impact of paid media is initially negative; however, it has a positive effect on probability of party selection in week 3. We find that the marginal effect of improved earned media valence is positive and, by some margin, highest in week 1. The marginal effect of paid media valence, while also positive, shows no discernible differences through the weeks. Interestingly, as the original probability of intending to vote for a party varies, an improved valence continues to have a comparable effect, though perhaps with the largest effect among those who are seriously on the fence with a probability around .5.

Table 5: Sensitivity Analysis for Average Level of Encounters

Rate of change in Probability	Week 1	Week 2	Week 3	Week 1	Week 2	Week 3	Week 1	Week 2	Week 3
	0.3	0.3	0.3	0.5	0.5	0.5	0.8	0.8	0.8
Earned Media Freq	0.429	0.296	0.649	0.510	0.353	0.773	0.327	0.226	0.495
Earned Media Val	0.686	0.107	0.174	0.817	0.127	0.207	0.523	0.081	0.132
Paid Media Freq	-0.094	0.022	0.146	-0.111	0.027	0.173	-0.071	0.017	0.111
Paid Media Val	0.191	0.220	0.142	0.227	0.262	0.169	0.146	0.168	0.108

5.2 Results from the Final Vote model

In this section, we describe results pertaining to Equations 7 and 8 (detailed in Section 3.2).

Table 6 contains the results from the logistic regression analysis of whether or not a person voted. The positive and significant coefficient ϕ_2 suggests that those who voted in 2005 were also more likely to vote in the current election, as were those who intended to vote for one of the three major parties (ϕ_3). In addition, the results suggest that those voters for whom international terrorism was a salient issue had higher voting likelihood while those who were concerned with climate change had lower voting likelihood, probably because of the skepticism induced by the “Climategate” scandal.

The results of the conditional final voting model based on the covariates described in Table 7 show that the previous voting intentions have a significant, positive impact on the final vote and are a precursor to the final voting decision. As expected, the decay parameter tells us that the most recent decisions have the greatest impact. As exposure valence continues to have an effect on intentions in each week, this emphasizes the need to prioritize the quality of media exposures right up until the final choice decision. With regard to the effect of different exposures on the final

vote, it is interesting to note from Table 8 that the cumulative valence of earned media had a very significant and positive impact on the final vote, consistent with its impact on voting intentions. Also, a being positive and significant reaffirms our assumption that the effectiveness of these exposures depreciates over time. In fact, it decays slowly and has a positive effect on the final voting decision. This is consistent with the relatively stable nature of brand attitudes which may be influenced by strongly positive or negative encounters (MacKenzie et al. 1986). Once again, then, our model demonstrates the managerial importance of valence over frequency and, in particular, of maintaining positive valence across both earned and paid media.

Parameter estimates and their interpretations for the impact of demographics and socio-political issues on actual voting are explained in detail in the Appendix A1 (see Tables 3 and 4 in that Appendix).

Table 6: Decision to Vote: Final Vote Model

Description	Parameter	mean	2.5%	97.5%
Intercept	ϕ_1	0.207	-1.88	3.11
Voted in 2005	ϕ_2	3.757	1.672	6.829
Intended to Vote Major Party	ϕ_3	1.548	0.276	3.523
Decay Factor	ϕ_4	0.490	-0.671	1.905
The economy	κ_1	0.368	-0.196	1.209
Education	κ_2	0.394	-0.627	1.513
Tax & spend	κ_3	-0.216	-1.289	0.582
Family life and childcare	κ_4	-0.284	-1.360	0.596
Climate change	κ_5	-0.493	-1.386	-0.066
International terrorism	κ_6	0.827	0.169	1.754
Housing	κ_7	-0.285	-1.217	0.292
Social Category	ϖ_1	-0.842	-2.304	0.710
Gender	ϖ_2	-0.719	-3.138	0.997
Age	ϖ_3	0.567	-0.122	1.512
Marital Status	ϖ_4	-0.845	-2.737	0.779
No. of Children	ϖ_5	-0.912	-3.883	1.379
Employment Category	ϖ_6	-0.008	-1.869	1.356
Education Category	ϖ_7	0.165	-0.418	1.039
Home Ownership	ϖ_8	-0.016	-0.059	0.014

Note: Emboldened variables are significant.

Table 7: Previous Intentions (Final Vote Model)

Description	Parameter	mean	2.5%	97.5%
Effects of Previous Intentions	a	2.598	2.125	3.089
Decay factor	ϕ	0.470	0.331	0.617
Party Voted in 2005	\tilde{a}	0.075	-0.414	0.550

Note: Emboldened variables are significant.

Table 8: Parameter Estimates of Party Exposures (Final Vote Model)

Descriptions		Parameter	Mean	2.50%	97.50%
Cumulative Valence	Earned Media	θ_1	1.934	0.907	3.128
	Paid Media	θ_2	0.713	-0.272	1.64
Cumulative Frequency	Earned Media	θ_3	-1.393	-2.464	-0.596
	Paid Media	θ_4	0.018	-0.951	0.871

Note: Emboldened variables are significant.

5.3 Results from the Exposure Frequency and Average Valence Model

In this section, we describe results pertaining to Equations 10 and 11 (detailed in Section 3.3). The results from estimating exposure frequency and valence (Tables 9 and 10) provide several interesting results. Those from lower social categories (manual, semi-manual and lowest grade workers) have higher paid media frequencies ($\zeta_{1,2}^{F1}$); this may result in part from their greater television consumption. Older age voters report higher frequencies particularly of earned media ($\zeta_{3,1}^{F1}$), and they report a lower valence for these earned media exposures ($\zeta_{3,1}^{V1}$). This is consistent with a higher engagement in election processes with age, illustrated by higher voting rates, and a decreased trust in politicians with age (Hetherington 1998). Those with children report lower earned media frequencies ($\zeta_{5,1}^{F1}$), perhaps due to having less time for media consumption; the valence of these exposures is, though, slightly higher ($\zeta_{5,1}^{V1}$). People in employment report higher exposure frequencies ($\zeta_{6,1}^{F1}, \zeta_{6,2}^{F1}$). This corresponds with their higher likelihood of voting, particularly for those in the public sector (Corey and Garand, 2002). Finally, higher education is associated with a higher earned media frequency, which again may reflect a higher engagement in political processes.

5.4 Correlation Matrix

In this section, we describe results pertaining to Equation 12 (detailed in Section 3.4). The estimated value of the correlation matrix is given in Table 11. This is the estimate of the variance-covariance matrix from section 3.4 and equation 12. In order to interpret the results, we label the strength of the associations using 0-0.19 as very weak, 0.2-0.39 weak, 0.40-0.59 moderate, 0.6-0.79 strong and 0.8-1 very strong. However, these cut-offs reflect somewhat arbitrary limits, and results should be considered in the context from which they are derived. We find a moderate

Table 9: Model Exposure Frequency : Demographics

Description		Parameter	mean	2.5%	97.5%
Social Category	Earned Media	$\zeta_{1,1}^{F1}$	0.035	-0.195	0.259
Social Category	Paid Media	$\zeta_{1,2}^{F1}$	0.243	0.017	0.502
Gender	Earned Media	$\zeta_{2,1}^{F1}$	0.051	-0.151	0.259
Gender	Paid Media	$\zeta_{2,2}^{F1}$	0.167	-0.066	0.400
Age	Earned Media	$\zeta_{3,1}^{F1}$	0.286	0.215	0.355
Age	Paid Media	$\zeta_{3,2}^{F1}$	0.073	0.009	0.150
Marital Status	Earned Media	$\zeta_{4,1}^{F1}$	-0.043	-0.253	0.160
Marital Status	Paid Media	$\zeta_{4,2}^{F1}$	-0.104	-0.348	0.137
No. of Children	Earned Media	$\zeta_{5,1}^{F1}$	-0.255	-0.507	-0.003
No. of Children	Paid Media	$\zeta_{5,2}^{F1}$	-0.023	-0.320	0.284
Employment Category	Earned Media	$\zeta_{6,1}^{F1}$	0.175	0.039	0.387
Employment Category	Paid Media	$\zeta_{6,2}^{F1}$	0.379	0.142	0.619
Education Category	Earned Media	$\zeta_{7,1}^{F1}$	0.059	0.010	0.129
Education Category	Paid Media	$\zeta_{7,2}^{F1}$	0.033	-0.044	0.111
Home Ownership	Earned Media	$\zeta_{8,1}^{F1}$	0.001	-0.004	0.006
Home Ownership	Paid Media	$\zeta_{8,2}^{F1}$	0.003	-0.003	0.008
Lagged Confidence on Intention		ζ^{F2}	0.002	-0.004	0.0038
Lagged Intention to Vote		ζ^{F3}	0.082	-0.132	0.148

Note: Emboldened variables are significant.

Table 10: Average Valence Model : Demographics

Description		Parameter	mean	2.5%	97.5%
Social Category	Earned Media	$\zeta_{1,1}^{V1}$	0.014	-0.040	0.068
Social Category	Paid Media	$\zeta_{1,2}^{V1}$	-0.002	-0.072	0.061
Gender	Earned Media	$\zeta_{2,1}^{V1}$	-0.002	-0.045	0.040
Gender	Paid Media	$\zeta_{2,2}^{V1}$	0.037	-0.020	0.099
Age	Earned Media	$\zeta_{3,1}^{V1}$	-0.029	-0.052	-0.012
Age	Paid Media	$\zeta_{3,2}^{V1}$	-0.013	-0.030	0.010
Marital Status	Earned Media	$\zeta_{4,1}^{V1}$	-0.031	-0.079	0.017
Marital Status	Paid Media	$\zeta_{4,2}^{V1}$	-0.026	-0.099	0.039
No. of Children	Earned Media	$\zeta_{5,1}^{V1}$	0.050	0.008	0.120
No. of Children	Paid Media	$\zeta_{5,2}^{V1}$	0.039	-0.045	0.122
Employment Category	Earned Media	$\zeta_{6,1}^{V1}$	0.003	-0.045	0.053
Employment Category	Paid Media	$\zeta_{6,2}^{V1}$	0.010	-0.055	0.063
Education Category	Earned Media	$\zeta_{7,1}^{V1}$	0.007	-0.009	0.024
Education Category	Paid Media	$\zeta_{7,2}^{V1}$	0.006	-0.016	0.029
Home Ownership	Earned Media	$\zeta_{8,1}^{V1}$	0.000	-0.001	0.002
Home Ownership	Paid Media	$\zeta_{8,2}^{V1}$	0.000	-0.002	0.001
Lagged Confidence on Intention		ζ^{V2}	0.021	-0.006	1.1
Lagged Intention to Vote		ζ^{V3}	0.392	0.212	0.673

Note: Emboldened variables are significant.

positive linear association between the intention to vote Liberal Democrat and exposure frequency of earned media ($\text{corr}(b_{i4}, h_{i2}^F) = 0.521$, significant). This indicates that earned media was differentially more persuasive for those intending to vote Liberal Democrat than Labour or Conservative, perhaps because LibDem voters tend to hold less loyalty to their party and conviction in their party identification (Rose and McAllister 1990) and so are more persuadable. We find a weak positive linear association between exposure valence of paid media and earned media ($\text{corr}(h_{i1}^V, h_{i2}^V) = 0.368$, significant), which suggests that these two forms of communication reinforce each others' effects (i.e. if you feel positive about encounters via earned media, you are likely to feel the same about encounters via paid media). We find a very weak positive linear association between exposure valence of paid media and exposure frequency of paid media ($\text{corr}(h_{i1}^V, h_{i1}^F) = 0.175$, significant), which would indicate that how positive or negative an encounter makes someone feel is weakly reinforced by how often they experience such paid communications. We find a very weak positive linear association between exposure valence of paid media and exposure frequency of earned media ($\text{corr}(h_{i1}^V, h_{i2}^F) = 0.168$, significant) indicating that how positive or negative a paid communication is perceived to be is only weakly related with how often earned media is experienced. We find a weak positive linear association between earned exposure frequency and earned exposure valence ($\text{corr}(h_{i2}^V, h_{i2}^F) = 0.321$, significant), which indicates that how often an earned communication is received is associated, but not necessarily caused by, how often voters encountered earned communications. Again, this indicates that how positive or negative an earned media encounter makes someone feel is weakly reinforced by how often they experience such communications. We find a moderate positive linear association between paid exposure frequency and earned exposure frequency ($\text{corr}(h_{i1}^F, h_{i2}^F) = 0.399$, significant), which would indicate that how often a paid communication is received is associated, but not necessarily caused by, how often voters encounter earned communications. This may be explained by two possible mechanisms: 1) people's awareness of politics is raised and so they notice other forms of political communications and 2) paid political advertising campaigns in the UK are often discussed in depth by journalists in the broadcast and press news in British politics (particularly the poster campaigns).

5.5 Simulation Study

We conduct a simulation study to evaluate the performance of our proposed model. The simulation is focused on evaluating the finite sample performance of the proposed Bayesian estimation when the data are generated, mimicking the real data we have. The scheme of generating the data and detailed results are in Appendix B.1. Based on the results, we find that the estimates under the

Table 11: Posterior estimates of correlation matrix $\widehat{\Sigma}$.

	b_{i2}	b_{i3}	b_{i4}	u_i	h_{i1}^V	h_{i2}^V	h_{i1}^F	h_{i2}^F
b_{i2}	1.000	0.226	0.378	0.094	0.066	0.161	0.149	0.520
b_{i3}		1.000	0.218	0.032	0.083	0.103	0.092	0.190
b_{i4}			1.000	0.113	0.123	0.220	-0.042	0.521
u_i				1.000	0.048	0.061	0.188	0.194
h_{i1}^V					1.000	0.368	0.175	0.168
h_{i2}^V						1.000	0.137	0.321
h_{i1}^F							1.000	0.399
h_{i2}^F								1.000

Note: Emboldened correlation values are significant

proposed model reliably recovered the true estimates with reasonable coverage probability.

6 Conclusions

In this paper, we have reviewed the relative strengths and limitations of existing methods of evaluating voters' response to communications, proposed the use of RET for this purpose, presented a dynamic model of the impact of communications on voters' choice, and applied the model to a three-way political party choice. Results show that the valence of customers' perceptual response to communications, and not just communications frequency, is important in predicting voters' choice. Indeed, communications wearout leads to exposure frequency becoming insignificant in determining voting intentions, and only earned media frequency being significant in final voting behavior. It's worth pointing out that the whole advertising industry model is currently structured around weight/dissemination ('impressions') irrespective of how these are received - what might be termed a measure of production evaluation rather than measure of effect. So we have an approach here to evaluate the relative influence and not just the reach of communication types. This confirms the importance of tracking both valence and frequency in communications evaluation, as otherwise, entirely wrong conclusions might be drawn. Another important focus in our analysis is the role of paid versus earned communications. In the politics context, while both are important, earned media have a double role, influencing both weekly voting intentions and additionally influencing the final vote, as voters seek to check their emerging evaluations against trusted sources immediately prior to their high-involvement decision (it is important to note that this may be the

case in the UK only where paid broadcast advertisements are illegal). Hence, our results extend the findings in Stephen and Galak (2012) regarding the importance of social earned media relative to paid media, demonstrating that this applies to earned media as a whole, and adding that the valence of both earned and paid media are important and not just their frequency. This once again emphasizes the managerial importance of understanding communications valence and not just exposure, contrary to the accepted gold standard in commercial practice of market mix modeling based on exposure or proxies for it, and contrary to the common managerial emphasis on ‘share of voice’ metrics. Complementing these decay effects, we find that earned media are particularly influential early in the decision-making cycle, and that the danger of wearout is less pronounced closer to the final brand choice, particularly in the case of paid media. Parties may therefore wish to phase their communications efforts relative to the individual’s voting cycle where this is possible.

Results from our proposed model are congruous with the previous descriptive work of Baines et al. (2011). Notably, Baines et al. (2011) speculate that the lower actual Liberal Democrat vote compared with the final polls was due to their less strongly held voting intentions. Our findings support that interpretation: the final brand choice is determined not only by the final intention but also (albeit with a decay) by previous intentions. So a longer-held view is more likely to be followed through. This is an important point for the opinion polling industry, which tends to take final voting intention as the sole predictor of the final vote. Nonetheless, the Liberal Democrat’s campaign overall was successful and resulted in them sharing power. Baines et al. (2011) provide some qualitative discussion as to why their leader performed particularly well in earned media early in the campaign. Our findings suggest that this might have had a residual effect on the final vote, even after party touchpoint valences evened out towards the election day, as the decay for the impact of touchpoint valence is lower than that for touchpoint frequency; furthermore, positively-valenced exposures continue to have an impact beyond their influence on voting intentions, also appearing in the final vote model.

Our study also illustrates the benefits of an RET-based approach to communications evaluation, which distinguishes it from other current methods. It allows modeling at the individual level. This provides a granularity of insight that is not available through aggregated market mix models (Gatignon 1990). In our sample of swing voters, for example, overall party market shares demonstrated only modest shifts between the start and end of the election campaign, but aggregate analysis would have masked a considerable level of switching in all directions, with only 39% of the sample maintaining a consistent preference throughout the study period. Our results show that communications played a significant role in this switching behavior. Importantly, we found that both exposure frequency and exposure valence are dependent on individual consumer character-

ristics and previous exposures. This demonstrates a validity challenge to the plethora of previous literature that treats communications exposure as exogenous. Our results also show that at least in the context of these elections, earned media are both pervasive and persuasive. While PR is a \$14.2bn industry in its own right (Holmes Report 2016), much of the current PR effort focuses on overall corporate reputation rather than direct brand sales; and when compared with the \$584bn spent on paid-for media, it seems highly likely that many firms are under-investing in influencing earned media commentary about their products. RET studies would allow practitioners to check for this possibility. RET captures exposures in real time. This allows the capture of the consumer's perceptual responses with greater validity than retrospective surveys. The valence of this perceptual response, in the form of a valence measure, is significant in our data: for example, the loss of the Liberal Democrats' early lead was not due to any fall in share of voice, but rather to a loss of their initial advantage in encounter valence which followed its leader's superior performance in the first televised debate. While market mix models based on media spend or reach are useful for a summative evaluation such as relative spend, for formative purposes such as improving communications content this perceptual response data is invaluable, while its real-time capture may allow amendments to campaigns in near-real time, too.

For commercial brands adopting RET, our findings would suggest that ideally, data collection should follow the approach we have described, tracking individuals over time prior to the purchase decision, as it is only in this way that the relative effects of different communication types on final brand choice, and their relative importance at different distances from the brand choice, can be estimated. It is not uncommon for practitioners in sectors such as automotives, consumer technology or financial services to conduct 'purchase window' surveys, recruiting individuals who plan to change their car, mobile handset or mortgage within the next month or two. To extend such a survey to a multiple-week RET study up until the purchase would enable the model we have described to be applied with very few modifications. Indeed, if the purchase window occurs over longer than the four weeks of the UK's election cycle, as is the case with US presidential elections for example, then individuals can be tracked for longer than four weeks: we are aware of several commercial examples of this among lead users of RET. The main difficulty this raises is that of motivating continued participation in the study, which necessitates allowing for a proportion of participants to drop out in the initial recruitment; however, we have seen successful examples of consumers staying in the field for 8-10 weeks with a manageable drop-out level of 40-50%.

By contrast, many early adopters of RET such as LG, Delta, Coca-Cola and Unilever currently track individuals for just one week, primarily using the data to examine communications impact on changes in brand attitudes. While this has many benefits as compared with the cross-sectional

survey, clearly a brand manager would ideally want insight into drivers of sales, not just attitude shift: after all, our final choice model shows that it is not just the final week's attitudes which shape the actual decision. For impulse and portfolio categories such as snacks and soft drinks, one-week tracking may suffice for determining communications impact on purchase, by applying an adapted version of our model to those individuals who have made a brand purchase by the end of the week. Here, indeed, time effects may be profitably examined at the level of days rather than weeks. For more considered purchases, though, such as televisions or cars, we would suggest that managers extend the tracking of individuals until they make their purchase, if at all possible.

Overall, we believe that RET could have a profound impact on the measurement of communications effectiveness, given its ability to model individual consumer response to communication encounters and, importantly, evaluate the relative influence of those encounters. To support practitioners in the appropriate application of RET, further research is needed to better understand its validity properties, and to locate its best role in the market researcher's armoury. Notably, field experiments could usefully compare its performance in predicting consumer choice with the survey, along the lines of the comparison of surveys and purchase diaries by Wind and Lerner (1979) comparison of surveys and purchase diaries. The inclusion in such a study of an objective data source such as clickstream data could enable the extent of under-reporting in RET to be checked, and compared with that of the retrospective survey. Further work is also needed to fuse RET with other data sources such as objective purchase recording, using retailer loyalty card data for example.

References

- Bass, F.M., Bruce, N., Majumdar, S. and Murthi, B.P.S., 2007. Wearout effects of different advertising themes: A dynamic Bayesian model of the advertising-sales relationship. *Marketing Science*, 26, pp.179-195.
- Baines, P., Macdonald, E. K., Wilson, H. N., and Blades, F. 2011. Measuring communication channel experiences and their influence on voting in the 2010 British General Election. *Journal of Marketing Management* **27** 694–720.
- Baumeister, R. F., Vohs, K. D., DeWall, C. N., and Zhang, L. 2007. How emotion shapes behavior: Feedback, anticipation, and reflection, rather than direct causation. *Personality and Social Psychology Review* **11** 167–203.
- Baxendale, S., Macdonald, E. K., and Wilson, H. N. 2015. The impact of different touchpoints on brand consideration. *Journal of Retailing* **91**(2) 235-253.
- Belli, R.F., Traugott, M. W. , Young, M., and McGonagle, K. A. 1999. Reducing vote overreporting in surveys: Social desirability, memory failure, and source monitoring. *The Public Opinion Quarterly* **63**(1) 90–108.
- Berger, J., Sorensen, A.T., and Rasmussen, S. J. 2010. Positive effects of negative publicity: When negative reviews increase sales. *Marketing Sci.* **29**(5) 815–827.
- Briñol P., Petty, R. E., and Tormagla, Z. L. 2004. Self-validation of cognitive responses to advertisements. *J. Consumer Res.* **30**(4) 559–572.
- Brownstone, D., and Train, K. 1999. Forecasting new product penetration with flexible substitution patterns. *Journal of econometrics* **89**(1) 109–129.
- Celeux, G., Forbes, F., Robert, C. P., and Titterton, D. M. 2006. Deviance information criteria for missing data models. *Bayesian Analysis* **1**(4) 651-673.
- Christel F., Marc, A., Molenberghs, G., Helena, G., Greet, T., and Luc, B. 2008. A high-dimensional joint model for longitudinal outcomes of different nature. *Statistics in Medicine* **27** 4408–4427.
- Christensen, W. F., and Florence, L. W. 2008. Predicting presidential and other multistage election outcomes using state-level pre-election polls. *The American Statistician* **62**, 1–10
- Corey, E.C., and Garand, J.C. 2002. Are government employees more likely to vote? An analysis of turnout in the 1996 U.S. national election. *Public Choice*, **111**, 259-283
- Cowley, E. 2007. How enjoyable was it? Remembering an affective reaction to a previous consumption experience. *Journal of Consumer Research* **34**(4) 494-505.
- Dunson, D. B. 2003. Dynamic latent trait models for multidimensional longitudinal data. *Journal of the American Statistical Association* **98** 555–563.

- Dunson, D.B., Chen, Z., and Harry, J. 2003. A Bayesian approach for joint modeling of cluster size and subunit-specific outcomes. *Biometrics* **59** 521–530.
- eMarketer. 2017. Worldwide ad spending: The eMarketer forecast for 2017. eMarketer, 12 April. Retrieve from: <https://www.emarketer.com/Report/Worldwide-Ad-Spending-eMarketer-Forecast-2017/2002019> [accessed 12 August 2017].
- ESOMAR. 2016. Global Market Research Report 2016. Amsterdam: ESOMAR. Retrieve from: <https://www.esomar.org/uploads/public/publications-store/reports/global-market-research-2016/ESOMAR-GMR2016> [accessed 12 August 2017].
- Erdem, T., and Keane, M. P. 1996. Decision-making under uncertainty: capturing dynamic brand choice processes in turbulent consumer goods markets. *Marketing Sci.* **15**(1) 1–20.
- Fazio, R.H., and Williams, C. J. 1986. Attitude accessibility as a moderator of the attitude-perception and attitude-behavior relations: an investigation of the 1984 presidential election. *Journal of Personality and Social Psychology* **51** 505–514.
- Fredrickson, B.L. 2000. Extracting meaning from past experiences: The importance of peaks, ends, and specific emotions. *Cognition and Emotion* **14**, 577–606.
- Fournier, P., Nadeau, R., Blaise, A., Gidengil, E., and Nevitte, N. 2004. Time-of-voting decision and susceptibility to campaign effects. *Electoral Studies* **23**(4) 661–681.
- Gatignon, H. 1993. Marketing-mix models. *Handbooks in operations research and management science* **5** 697–732.
- Geisser, S., and Eddy, W. F. 1979. A predictive approach to model selection. *Journal of the American Statistical Association* **74**(365) 153–160.
- Gelfand, A. E., Dey, D. K., and Chang, H. 1992. Model determination using predictive distributions with implementation via sampling-based methods (No. TR-462). Stanford University Department of Statistics.
- Gelman, A., Carlin, J. B., Stern, H., and Rubin, D. B. 2004. *Bayesian Data Analysis*. London Chapman & Hall/CRC.
- Gelman, A., Jakulin, A., Pittau, M. G., and Su, Y. S. 2008. A weakly informative default prior distribution for logistic and other regression models. *The Annals of Applied Statistics* **2**(4) 1360–1383.
- Ghosh, P., and Tu, W. 2009. Assessing Sexual Attitudes and Behaviors of Young Women: A Joint Model with Nonlinear Time Effects, Time Varying Covariates, and Dropouts. *Journal of the American Statistical Association* **104** (486) 474–485.
- Goh, K.Y., Hui, K.I., and Png, I.P.L. 2011. Newspaper report and consumer choice: Evidence from the do not call registry. *Management Sci.* **57**(9) 1640–1654.
- Hetherington, M.J. 1998. The political relevance of political trust. *American Political Sci. Review.* **92**, 791–808.

- Herr, P.M., and Fazio, R. H. 1993. The attitude-to-behavior process: implications for consumer behavior. In A. Mitchell (Ed.), *Advertising exposure, memory, and choice*. Hillsdale, NJ: Erlbaum.
- Holmes Report. 2016. Infographic: 2016 Global PR industry at a glance. New York & London: Holmes Report. Retrieve from: <https://www.holmesreport.com/research/article/infographic-2016-global-pr-industry-at-a-glance> [accessed 12 August 2017].
- Jiang, B., Wang, N., Sammel, M. D., and Elliott, M. R. 2015. Modelling short and longterm characteristics of follicle stimulating hormone as predictors of severe hot flashes in the Penn Ovarian Aging Study. *Journal of the Royal Statistical Society: Series C (Applied Statistics)* **64**(5) 731-753.
- Kahn, B.E., and Isen, A.M. 1993. The influence of positive affect on variety among safe, enjoyable products. *Journal of Consumer Research* **20** 257-270.
- Klein, J.G., and Ahluwalia, R. 2005. Negativity in the evaluation of political candidates. *Journal of Marketing* **69** 131-142.
- Lambert, D. 1992. Zero-inflated Poisson regression with an application to defects in manufacturing. *Technometrics* **34** 1-14.
- Linzer, D.A. 2013. Dynamic Bayesian forecasting of presidential elections in the states. *Journal of the American Statistical Association* 108(501): 124-134.
- Macdonald, E.K., H.N. Wilson, and U. Konus. 2012. Better customer insight in real time. *Harvard Business Review* **90**(9) 102-108.
- MacKenzie S. B., Lutz, R. J., and Belch, G. E. 1986. The role of attitude toward the ad as a mediator of advertising effectiveness: A test of competing explanations. *Journal of Marketing Res.* **23** 130-43.
- Meng, X. L. 1994. Posterior predictive p-values. *The Annals of Statistics* 1142-1160.
- Naik, P. A., and Peters, K. 2009. A hierarchical marketing communications model of online and offline media synergies. *Journal of Interactive Marketing* **23** 288-99.
- Natarajan, R., and McCulloch, C. E. 1998. Gibbs sampling with diffuse priors: A valid approach to data-driven inference. *Journal of Computational and Graphical Statistics* **7** 267-277.
- Newman, B. I., and Sheth, J. N. 1985. A model of primary voter behavior. *Journal of Consumer Research* **12**(2) 178-187.
- O'Cass A. 2002. Political advertising believability and information source value during elections. *J. Advertising* **31**(1) 63-74.
- Park, Y. H., Chang, H. P., and Ghosh, P. 2011, Modelling member behaviour in online user-generated content sites: A semiparametric Bayesian approach. *Journal of the Royal Statistical Society: Series A* **174**(4) 1051-1069.
- Romani, S., Grappi, S., and Dalli, D. 2012. Emotions that drive consumers away from brands: Measuring negative emotions towards brands and their behavioral effects. *International Journal of Research in Marketing* **29**(1) 55-67.

- Rose, R., and McAllister, I. 1990. *The loyalties of voters: A lifetime learning model*. London: SAGE Publications.
- Spiegelhalter, D. J., Best, N.G., Carlin, B. P., and Van Der Linde, A. 2002. Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **64**(4) 583–639.
- Spiegelhalter, D., Thomas, A., Best, N., and Lunn, D. 2005. WinBUGS User Manual, Version 1.4.2, <http://www.mrc-bsu.cam.ac.uk/bugs.html>.
- Stephen, A.T., and Galak, J. 2012. The effects of traditional and social earned media on sales: A study of a microlending marketplace. *Journal of Marketing Research* **49**(5) 624-639.
- Verhoef, P. C. 2003. Understanding the effect of customer relationship management efforts on customer retention and customer share development. *Journal of Marketing* **67** 30–45.
- West, M., and Harrison, J. 1997. *Bayesian Forecasting and Dynamic Models (2nd edition)*. Berlin: Springer.
- Wind, Y., and Lerner, D. 1979. On the measurement of purchase data: Surveys versus purchase diaries. *Journal of Marketing Research* **16** 39–47.
- Worcester, R., Mortimore, R., Baines, P., and Gill, M. 2011. *Explaining Cameron's Coalition: The Outcome of the 2010 British General Election*, London: Biteback Publishing.
- Yaveroglu, I., and Donthu, N. 2008. Advertising repetition and placement issues in on-line environments. *Journal of Advertising* **37**(2) 31–44.

Tracking the impact of media on voter choice in real time: A Bayesian dynamic joint model

Pareek, Bhuvanesh

2018-01-15

Attribution-NonCommercial 4.0 International

Pareek B, Ghosh P, Wilson HN, Macdonald EK, Baines P, Tracking the impact of media on voter choice in real time: A Bayesian dynamic joint model, Journal of the American Statistical Association, 2018, Volume 113, Issue 524, pp. 1457-1475

<http://dx.doi.org/10.1080/01621459.2017.1419134>

Downloaded from CERES Research Repository, Cranfield University