Matching Theory and Data: Why Combining Media Content with Survey Data Matters

ANDREAS R.T. SCHUCK, RENS VLIEGENTHART AND CLAES H. DE VREESE*

Understanding media effects is a core challenge for scholars and students of electoral behavior and campaigns. In an era characterized by volatile electorates,1 a rapidly changing media landscape2 and increased mediatization,3 it is of central importance to disentangle the role of the media in these developments. This challenge has obvious relevance for both science and society, and it is one reason that the field of political communication is burgeoning and prominent in journals in both communication science and political science.

A key question is how the media affect political behavior. This is not a new question, but one that has received heightened attention in recent years. In our article ‘Who’s Afraid of Conflict?’4 we test whether individual exposure to conflictual news coverage of European Parliament elections mobilizes citizens, contingent on general polity evaluations. This article is based on an original two-wave panel survey and a media content analysis of television news and newspapers, which enables us to test the dynamics of such questions. The content analysis was conducted within the framework of PIREDEU, where the data are also publicly available.

Our article has prompted a response by Fazekas and Larsen (F&L).5 In ‘Media Content and Political Behavior in Observational Research: A Critical Assessment’ they raise a number of issues with regards to our article and to combining survey and content data in general. We agree that this topic is important, and we are pleased with the opportunity to respond to their Note. Our response is organized as follows. We first address some of their observations about our article. We then discuss the idea of combining survey and content data more generally and the virtue of this approach. Finally, we list a number of issues that may help us looking forward.

WHO’S AFRAID OF CONFLICT?

F&L offer a number of observations about our data and modeling strategy. We respond to these in our supplementary material (see Supplementary Information A–D). First things first: our original article reported being based on twenty-one countries. Due to a file-merging mistake on our side, Bulgaria was not included in the analysis. This is correctly noted by F&L and we regret this error.6 As we will discuss below and in the supplementary material in greater detail, this and

---

* Amsterdam School of Communication Research, University of Amsterdam (emails: a.r.t.schuck@uva.nl; r.vliegenthart@uva.nl; c.h.devreese@uva.nl). Data replication sets are available at http://dataverse.harvard.edu/dataverse/BJPolS, and online appendices are available at http://dx.doi.org/doi:10.1017/S0007123415000228.

1 Dalton 2013.
2 Chadwick 2013.
3 Esser and Stromback 2014.
4 Schuck, Vliegenthart, and de Vreese 2014.
5 Fazekas and Larsen 2015.
6 The journal has informed us that this acknowledgment will serve as the official erratum.
other data-related issues listed by F&L do not affect our interpretation regarding the substantial meaning of our findings. This also pertains to our second hypothesis, which posits that the effect of conflict news depends on polity evaluations (see Supplementary Information A). More generally, F&L object to the way we merge panel survey data and media content data. As will become clear, we believe they raise important points, but we have four observations about their article that we address in turn below.

First, their article largely misses the point when it over-emphasizes the correlation, already noted in our original article, between our weighted news exposure measure and mere exposure. Secondly, F&L disregard the role of theory in this field of research. Thirdly, our empirical results regarding the mobilizing role of conflict news hold up. Fourthly, their empirical illustration, based on a different dataset, is beside the point and does not hold up when applied to the context we are interested in (voter mobilization over the course of a campaign).

WHY COMBINE SURVEY DATA AND CONTENT DATA?

Research on electoral behavior using survey data is often challenged by the ability to draw causal inferences. This applies both conceptually, where the metaphor of the ‘funnel of causality’ has been used, and empirically with the obvious limitations of relying on cross-sectional data. A subset of the literature on electoral behavior is particularly interested in the change or stability of attitudes and behavior during a campaign. Our article focuses on how the contents of political news may change citizens’ likelihood to participate in elections. The goal is to explain voter mobilization during a campaign rather than more static turnout, which most cross-sectional analyses are confined to. Too little is known about the role of media content in such mobilization processes. When trying to fill this gap, we are looking at a change in behavior prompted by the media. In such a situation the bar is high for finding large media effects.

More specifically, in our original article we set out to understand what role the media play in mobilizing citizens to participate in European Parliament elections. We know from the previous literature that media use (at least for most news media) tends to correlate positively with mobilization and turnout, though the use of some specific media has no (or a negative) correlation. Our endeavor is to go beyond previously established correlations between media use and behavior. We do so by first of all relying on panel data, so that we can assess the role of the media in mobilizing voters, while controlling for initial turnout intention. Secondly, we use a comprehensive content analysis of television news and newspapers across Europe to assess the content of the news coverage to ask the simple, but fundamental, question whether exposure to media content, rather than just media usage, plays a role. After all, as Lewis-Beck et al. state ‘[f]inding that a particular variable happens to predict voting would be of little theoretical use if we could not understand how it helped “cause” voting’. In communication science, the recommendation has been made to put ‘flesh on the bone’, which in this case would imply combining media content analyses with surveys that include measures for outlet-level media exposure. Others have made that same general observation, including for example

7 Campbell et al. 1960.
8 See, e.g., Berelson, Lazarsfeld, and McPhee 1954.
9 Zaller 1996.
12 See Slater 2004, 2014. We refrain from entering a related and important auxiliary debate on the use of self-reported media exposure measures (see, e.g., Prior 2013), but merely observe that our measures were at the outlet level as recommended by, e.g., Dilliplane, Goldman and Mutz (2013).
Kleinnijenhuis, van Hoof and Oegema, who looked at media content and political trust and stated that in the absence of information about content, ‘media consumption does not offer a very convincing explanation for the level of political trust’.\(^{13}\) However, such combinations of data are not always used, nor are they always possible in the absence of available data, as Graber summarizes: ‘most researchers fail to ascertain, let alone content analyze, the media information that, they assume, their subjects encountered’.\(^{14}\)

The need to combine multiple sources and kinds of data has become widely acknowledged. Publications using the Annenberg National Election Study and, for example, the large-scale German election study (GLES) include multiple sources of data.\(^ {15}\) Obviously, we are proponents of such data integration. In the particular case of media effects on electoral behavior, the absence of content data generally inhibits much of our understanding of media effects and theorizing on the dynamics and processes involved in such effects. The important question, then, is how best to combine content and survey data.

**THE IMPORTANCE OF THEORY**

How should we conceive of the role of the media in political processes, both conceptually and empirically? Different options ranging from time spent, relative reliance on a medium or exposure to certain programs have been considered in past decades.\(^ {16}\) Even in recent research, many observations are based on ‘empty exposure studies’ – that is, using media exposure measures and correlating them with an outcome variable, as data resources that would allow for more rigorous modeling are often not available. Still, these can be very plausible studies. Newton concludes ‘it seems to be the content of the media, rather than its form which is important’.\(^ {17}\) However, he was not able to look at actual content but extrapolated this from medium characteristics such as ‘broadsheet’ versus ‘tabloid’. Kenski and Stroud, for example, used media exposure variables and linked these to political efficacy, knowledge and participation.\(^ {18}\) The scope of their conclusions pertains to the relationship between usage and their dependent variables. One way to further pursue this important agenda would be with media content data so as to say more about what kind of content drives the established relationship.

The combination of survey and content data can take different forms. One might describe media content at the aggregate level during a particular period and link this to the development in public opinion during this period. Such studies can be exploratory in nature\(^ {19}\) or more formal, aggregate-level times series analyses.\(^ {20}\) These studies, albeit valuable for addressing some questions, fall short of drawing inferences at the individual level. When addressing questions at the individual level, a scholar might have individuals’ reported media outlet exposure in a survey and be able to link, at the individual level, these survey measures to the contents of different media outlets. This idea is not novel, but its application is still limited. One prominent example was provided by Erbring, Goldenberg and Miller in their study of media agenda setting.\(^ {21}\) They combined the media exposure measures in the American National Election

---

\(^ {13}\) Kleinnijenhuis, van Hoof, and Oegema 2006, 88.

\(^ {14}\) Graber 2004, 516.

\(^ {15}\) E.g., Kenski, Hardy, and Jamieson 2010.

\(^ {16}\) See McLeod and McDonald 1985.

\(^ {17}\) Newton 1999, 577.

\(^ {18}\) Kenski and Stroud 2006.

\(^ {19}\) E.g., Statham and Tumber 2013.

\(^ {20}\) E.g., Boomgaarden and Vliegenthart 2009; Hester and Gibson 2003.

\(^ {21}\) Erbring, Goldenberg, and Miller 1980.
Study with a content analysis of daily newspapers. All front-page articles were coded for the main issue and merged with the survey data by matching each respondent with information about the content in the particular paper the respondent had read.

Crucially, Erbring, Goldenberg and Miller included a coding of issues in their analysis because their dependent variable was issue salience (this being an agenda-setting study). They therefore made a theoretically informed choice about which content feature to include in their study. This we believe is at the core of the discussion here. Our choice to weigh in conflict framing is because theory predicts that this can be mobilizing, especially in the case of EU politics. Likewise, scholars interested in the effects of news on political cynicism follow theory that suggests that strategy news coverage is a relevant feature to include. Scholars investigating knowledge gaps have linked individual-level exposure measures with news complexity scores. The key matter is that, to the extent that content is considered relevant, the choice of relevant content features is driven by the theoretical puzzle at stake.

FOUR OBSERVATIONS

As mentioned at the outset, we have four observations about F&L’s note. First, F&L raise as a critique of our article that what we do would be ‘merely a rescaling of the news exposure variable’. However, this observation misses the point because we fully agree that our weighted exposure variable relies on the exposure variable. That is exactly what we report in the original article (our first observation). F&L imply that it is largely meaningless because our weighted exposure measure (a) correlates highly with general exposure and (b) one might weigh in other things and get similar empirical results. We see a fundamental difference here, both in theoretical and empirical terms. To start with, it is hardly a ‘discovery’ that our ‘weighted exposure measure’ relies on actual exposure. We explicitly acknowledge that the ‘raw’ exposure measures correlate highly with our ‘weighted exposure measure’ and exposure is a crucial part of that variable, also theoretically: we hypothesize that exposure to conflict framing is a significant factor in explaining mobilization during a campaign. Moreover, we stress that the ability to empirically get similar results with other measures is, in our view, not refuting a theory-driven model.

Secondly, the difference is that we do not ‘merely’ (term used by F&L) rescale. We transform the original variable based on theory. We note that F&L at no point question our theoretical assumptions (in this case that exposure to conflict news can be conducive to mobilization), but we also note that they do not consider the role of theory, and only reason on the basis of statistical considerations in their empirical tests and judgment about model superiority (our second observation; see also the discussion above on the role of theory).

Thirdly, nowhere in our article do we claim that a rescaled variable must per se perform better empirically. We believe that F&L’s elaborate comparison of the effect of raw media exposure and the rescaled news conflict variable is, in essence, beside the point (as we elaborate upon below). Even if we were interested in the formal comparison between weighted and raw exposure – which we are essentially not, for the theoretical reasons outlined above – we note that in all our models (and in F&L’s models) our conflict-weighted exposure variable performs

22 van der Eijk and Franklin 1996.
23 Cappella and Jamieson 1997.
25 See also de Vreese 2014.
26 Fazekas and Larsen 2015, 5.
27 Schuck, Vliegenthart, and de Vreese 2014, 11.
better than alternatives, as is shown consistently by both fit statistics (log likelihood, AIC, BIC) and the size of standardized coefficients (see Table S5 in Supplementary Material D for more information). Also, a comparison of the effects of conflict news, raw exposure and non-conflict news applying bootstrap procedures reveals that conflict news performs statistically better than both raw exposure and non-conflict (see Supplementary Material D) (our third observation).

We finally note that we do attach substantial meaning to the interpretation of the findings pertaining to our second hypothesis of the article, which posited that the effect of conflict news depends on polity evaluations. Here we find relevant and significant country differences regarding the mobilizing effect of conflict news based both on the original country-level measure we reported as well as on the additional measure that we referred to in footnote 47 in the original article (see Supplementary Material A).

Finally, the ‘non-conflict’ predictor of turnout in a cross-sectional setting (the ‘illustration’ F&L used to show that similar empirical findings can be obtained using unrelated content features) is theoretically hardly compelling and does not hold up when applied to our context of voter mobilization during a campaign (see Supplementary Information D). To rephrase F&L, getting to the same place empirically while using an irrelevant variable gets us nowhere theoretically. We deem the demonstration of the empirical link between the absence of a content feature and an outcome variable in their illustration irrelevant in this case. Moreover, we note that the example provided by F&L (correlating exposure to non-conflict with self-reported turnout in a cross-sectional setting using the European Election Study, thus relying on different data, different measures and different countries) is not replicated when using our original dataset, looking at mobilization during an election campaign, taking initial turnout intention into account and using panel survey data (see Supplementary Information D). Here, employing bootstrap procedures, we find that the effect size of conflict news is significantly larger than that of non-conflict news. This questions their ‘illustration’ in terms of theoretical relevance and empirical generalizability (our fourth observation).

**IS WEIGHTED BETTER?**

It would be a misunderstanding of our article to suggest that we are trying to do more than improve the link between media and political behavior by infusing theory into an otherwise potentially fairly meaningless variable (raw media exposure). Should a weighted exposure measure *per se* perform better empirically? We do not think so. In fact, one could think of situations in which a rescaled variable might reduce the empirical importance of exposure. Take the fictitious example that a scholar would want to know the impact of newspaper reading on knowledge about EU affairs. If *The Guardian* reported virtually nothing about the EU in a given time period, and we knew that based on content data, would a positive survey data-based coefficient for a respondent reading *The Guardian* daily (raw exposure) be substantively meaningful? Or might it express other underlying relationships? Would a weighted, potentially non-significant, exposure measure that might reduce the empirical importance of reading the newspaper because of the absence of relevant content not be more meaningful? We believe so. In our case, would a significant coefficient of media exposure be theoretically compelling if we knew that the outlet had reported about the elections in a non-mobilizing manner?

High media exposure can mean, de facto, high exposure to celebrity news, sports and weather, so bringing in relevant content features forms a necessary correction for an otherwise potentially inflated or deflated ‘raw exposure’ effect. It makes little sense to think of raw exposure as something entirely distinct and different from weighted exposure. Empirically we
know it is not (because one is partly based on the other) and conceptually high(er) exposure inevitably also implies a high(er) likelihood of exposure to relevant content characteristics, such as conflict in our case, which we know is one of the most prominent features of election coverage. The question thus is not ‘is it exposure or is it conflict?’ that matters because exposure (also) includes conflict, it just does not further specify the amount and does not account for relevant outlet-specific differences, which of course is important in today’s high-choice media landscape. At the same time, conflict without exposure cannot mean anything. But these are not two rival concepts competing with each other (see also Table S6 in the Supplementary Material D, where we disentangle the two and show that indeed both matter). It is essentially the combination of the two we are interested in from a theoretical viewpoint.

It is therefore both possible and illuminating if a weighted exposure measure does not always perform better but sometimes performs worse than general exposure. To take the opposite example, Jebril, Albaek and de Vreese demonstrate that a theoretically grounded weighted exposure can also significantly improve the performance of a raw exposure measure. They showed that while raw exposure was unrelated to their outcome variable, a weighted exposure yielded a significant effect because the relevant content features were infused into the exposure measure. Van Spanje and de Vreese note that mere exposure is insignificant in their model, while weighted exposure helps explain vote choice. In the same vein, Lengauer and Höller provide an example in the context of an Austrian election campaign in which general news exposure had no effect on turnout, but weighing in different content features and combining exposure with content data helped identify one particular content feature as having a demobilizing effect.

In our example in the original article, the difference between the raw exposure and the weighted exposure is not large. But in all models (also the ones offered by F&L) the conflict-weighted exposure is important and significant and, again, as we show in the supplementary material it also performs better than other operationalizations (see Table S5 in Supplementary Information D). Our argument is that the theoretical relevance of that relationship is more meaningful than the relationship with raw exposure. In their Note, F&L create the impression that not including ‘raw exposure’ was conducive to our argument. However, including this variable in our analyses augments the effect of the theory-driven weighted exposure (compare model in Table S2 with the first model in Table S3 in Supplementary Information D). But – as acknowledged in our article – we present a model with the weighted exposure based on the theoretical grounds outlined here and because of multicollinearity issues with a model including both variables. Either way, the observation that weighted exposure is largely a rescaling of reported exposure is correct. We assign great importance, however, to this transformation because the precision of the indicator increases, based on theoretical grounds.

THE WAY AHEAD

At the core of our endeavor to understand the role of media content in electoral behavior is a question of the ability to draw causal inferences. We believe that a combination of designs can be helpful. In our research we often use experimentation, survey data and media content data. While experimentation is obviously superior in making inferences about causality, we do also believe that the combination of panel survey data and media content analyses is a strong complementary ‘real-world’ design that enhances our ability to make claims about causality.

29 van Spanje and de Vreese 2014, 340.
30 Lengauer and Höller 2012.
Scholarly exchanges about theory and modeling should serve the purpose of retrospective learning and prospective improvements. We welcome this exchange, and with this response we hope to have contributed to a general discussion of why combining media content with survey data not only matters, but should (also) be driven by theoretical ambitions. We propose that we do not need less, but more explicit theorizing and links between survey and content data and about the consequences of different modeling strategies.\textsuperscript{31}

Combining data sources such as survey data and media data is a research endeavor that is ‘state of the art’ (term used by F&L). This is not to say that there is no need of further improvement, disentanglement, and refinement of measures and modeling techniques. The question of how to assess the role of content in media effects has not been settled, and we agree that future research should address the nature of the combination in more detail and more explicitly. More attention is needed for model specification, alternatives and robustness. The weighting of content is not carved in stone – much akin to modeling choices in (pooled) time series studies in political science in which time lags or other dynamic properties are open for discussion.\textsuperscript{32} Currently there are no established standards, and future research should investigate alternative ways of combining data as well as alternative ways of conceptualizing the content features that matter and the weight they should be assigned, relative to another, when combining with survey data.

Likewise, future research should spend more time isolating the effects of different content features. In our case, the response by F&L made us consider the following: while we fundamentally disagree that a better empirical performance of a weighted variable vis-à-vis a raw exposure measure is an important validity criterion for the reasons outlined above, we propose a first way to isolate the effects of exposure and conflict separately. In the example provided in Supplementary Information D (especially in Table S6) we demonstrate that both exposure and conflict matter, substantively and empirically independently. This, again, is in line with our reasoning put forward in the original article and yields a different conclusion than the one offered by F&L.

The above example merely serves to show that the ‘burden of evidence’ in future studies has both an empirical and a theoretical component. In that process there will (and should) be trial and error. But theory should be our guiding light when making these choices. For media effects research in general we see a progression: scholars have gone from asking questions such as whether time spent on media use is good or bad or whether television or newspapers are good for democracy to investigating more appropriate questions such as the effects of specific programs or outlet types. We go a step further with the observation that it is not program viewing or reading a specific paper that creates an effect, but rather the content of the program or paper. In the same vein, we currently see a new generation of studies looking at the role of the internet and social media, where a similar progression can be observed. The questions have developed from a focus on the implications of ‘time spent on the internet’ to what kinds of activities are undertaken, to what kinds of political messages are being created, shared and posted.\textsuperscript{33} This is a logical and needed step in using content in combination with (social) media use data to predict behavior. These developments in the media landscape, we believe, underscore that the study of media effects and political behavior is well served by more, not fewer, combinations of media content and survey data.

\textsuperscript{31} See Wolling 2002.
\textsuperscript{32} See, e.g., Wilson and Butler 2007.
\textsuperscript{33} See, e.g., Gil de Zuniga, Jung, and Valenzuela 2012.
REFERENCES


