Titel/Title: Diverse Effects of Mass Media on Concerns about Immigration: New Evidence from Germany, 2001–2016

Autor*innen/Author(s): Lukas Erhard, Raphael H. Heiberger, Michael Windzio

Veröffentlichungsversion/Published version: Postprint

Publikationsform/Type of publication: Artikel/Aufsatz

Diverse Effects of Mass Media on Concerns about Immigration: New Evidence from Germany, 2001 - 2016

Lukas Erhard[1], Raphael H. Heiberger[1], and Michael Windzio[2]

[1]University of Stuttgart, Seidenstraße 36, 70174 Stuttgart, Germany

[2]University of Bremen, Mary-Somerville-Straße 9, 28359 Bremen, Germany

Corresponding Author:

Lukas Erhard

Seidenstraße 36, 70174 Stuttgart

Germany

Email: lukas.erhard@sowi.uni-stuttgart.de

Abstract

Media discourse is often seen as an important condition of people's attitudes and perceptions. Despite a rich literature, however, it is not well understood how media exposure influences attitudes towards immigrants. In contrast to previous studies, we argue that people rely on "availability heuristics" shaped by mass media. From that point of view, it is the specific content of media discourse on immigration that affects people's concerns. We use "structural topic models" to classify media content of more than 24.000 articles of leading German newspapers from 2001 to 2016. Utilizing "linear fixed effect models" allows us to relate a person's concern towards immigration as reported in the German Socioeconomic Panel to prevalent topics discussed in print media while controlling for several confounding factors (e.g., party preferences, interest in politics, etc.). We find a robust relationship between topic salience and attitudes towards integration. Our results also reveal that specific topics with negative contents (e.g., domestic violence) to increase concerns, while others (e.g., scientific studies, soccer) decrease concerns substantially, underlining the importance of available information provided by media. In addition, people with higher education are generally less affected by media salience of topics.

1 Introduction

Immigration has been a widely and controversially debated topic in German mass media for the past decades. Theories and concepts of mass communication propose different mechanisms of how media influences public opinion and attitudes. Agenda-setting theory, for instance, assumes that frequent reporting on a particular issue widens its salience (Chyi and McCombs, 2004, p. 22). High salience, in turn, increases recipients' exposure to an issue and suggests that it is of particular importance (McCombs, 2004).

In order to be able to map a media discourse about particular issues, however, salience is not the only factor researchers need to consider. It is also important to investigate how these issues are framed[1]. Frames put issues into wider thematic contexts and ask how media present an issue to the audience. A frame is supposed to influence recipients to interpret an issue within the respective context. For instance, media reports on immigration and integration of migrants might apply a frame of the welfare system and, hence, put integration in context with receiving benefits from the state whereas other frames on the integration of migrants might relate to education or popular sports.

It is therefore important to distinguish between the theoretical mechanisms of (first-level) agenda-setting and framing when we try to understand how media reports influence recipients' attitudes towards immigration. Simply attributing media effects to issues and their respective salience, and thereby ignoring content and frames can lead to false generalizations of how the media influences attitudes. As we will show, previous studies conclude that media salience of the immigration issue generally

ll

increases anti-immigrant attitudes but do not reflect on the thematic context of media reports.

Our results confirm that many media articles frame the immigration issue in the context of social problems and thereby increase citizens' concerns. However, we also identify frames that reduce people's concerns about immigration. This bidirectional mechanism is consistent with human cognition in situations of high complexity, where humans are forced to rely on simplifying heuristics. A macro-level societal issue like integration of migrants is such an instance in which people often apply 'availability heuristics' (Tversky and Kahneman, 1973) to assess actually unobservable issues. As a consequence, people do not necessarily increase their stereotyping due to a high salience of the migration issue, but tend to more positive attitudes towards immigration when reports with a positive frame are available.

Our study extends the focus on salience of prior studies (Czymara and Dochow, 2018) with analyses of the content of media reports. In line with recent studies in media effects research, we conceptualize frames as thematic contexts in which media reports on immigration are embedded (Eberl et al., 2018). For that purpose, we utilize "structural topic models" (Roberts et al., 2014) on a comprehensive set of leading newspapers (weekly and daily) in Germany from 2001 to 2016. We combine topic trends on immigration with panel data from the representative German Socio-Economic Panel (GSOEP) (G. G. Wagner, Frick and Schupp, 2007), and apply fixed effects panel regression models. In so doing, we question the assumption that mere prevalence of media reports on migrants generally increases negative attitudes. By distinguishing between different topics, we will show that media exposure can increase and decrease concerns about migration in the German population, conditional on news articles' content.

2 Theoretical Framework

At the core of our theoretical framework lies the concept of an information environment created by agenda setting and media salience. We align this concept with insights from framing theory. The argument can be summarized as follows: people process information in a dynamic 'average' information environment (cf. section 2.1). Not least due to market competition, agenda-setting creates a flow of salient issues therein. In contrast to previous studies (Czymara and Dochow, 2018), however, we do not limit our argument to media salience, but do also consider framing theory (cf. section 2.2). From this, we begin to see that the framing of an issue (i.e., its content) is responsible for potential effects on people's attitudes, although issues certainly need some minimum level of salience to have an effect.

We will argue that media frames depend on the particular content of a news report. Frames influence recipients by combining the immigration issue with other issues, which may represent more positive or problematic contexts. These 'attached' issues, e.g., domestic violence or international students, often suggest an evaluation of the immigration issue and thereby activate a cognitive frame at the recipient's level (audience frames). As we will argue in the following section, it is the combination of media framing and audience framing – and not just the salience – of an issue which corresponds with attitudes and concerns about immigration. Hence, media effect studies should consider the content of news reports in order to capture the framing of the issues under investigation, which is possible by applying structural topic models to the analysis of the dynamic information environment (DiMaggio, Nag and Blei, 2013; Heidenreich et al., 2019; Jacobi, van Atteveldt and Welbers, 2016).

2.1 Information Environment, Agenda-Setting and Media Salience

Media communication creates a general information environment for attitude formation in the public (Jerit, Barabas and Bolsen, 2006). Even persons who are not exposed to media have access to salient issues, for instance, by face-to-face interaction with peers so that "... not everyone who was affected by the news has necessarily been exposed to it" (Boomgaarden and Vliegenthart, 2009, p. 518). Although recipients choose particular media, the German information environment is strongly influenced by news presented in high quality media, particularly in respected newspapers (Pew Research Center, 2018). Despite different political biases, those newspapers relate to a more or less similar stock of information (Pew Research Center, 2018). Taken together, these media create the average information environment. Alternative media often comment on events and issues reported in this environment. They do so in their particular way and from their particular perspective, but do nevertheless often relate to this overall news environment, at least with respect to issues not just of local relevance. Contrariwise, the mainstream media picks up issues from social media and niche media (Gruszczynski and M. W. Wagner, 2017), which also contributes to the emergence of an overall, 'average' information environment. The dynamic flow of the information environment results from the entry of new issues and aspects gaining in salience, so pushing existing issues or certain aspects to the background.

Not all respondents are equally exposed to this average information environment, which imposes measurement error on our core explanatory variables – the topics. In other words, if we find effects of the average information environment in our sample, the "true" effects might be even stronger. Nevertheless, it is essential to control for self-selection into a particular subdomain of the information environment by individual traits such as gender, ethnicity or personality (see also Chapter 3.3).

4

Which topics appear is due to the agenda-setting process. Agenda-setting describes the flow of information through the active selection of topics by the media and their presentation to the audiences (McCombs, 2004; McCombs, 2005; McCombs and Shaw, 1972). Setting an issue on the agenda increases its salience, defined as the prevalence or the intensity of media reports on an issue, that is, "the salience of immigration related topics on the media agenda is often conceptualized as the volume ... or intensity ... of reporting" (Eberl et al., 2018, p. 209). The mere salience of media articles on immigration has been shown to have a mostly negative impact on attitudes towards migration (Boomgaarden and Vliegenthart, 2007; Czymara and Dochow, 2018; Schlueter and Davidov, 2013). If the salience of the immigration issue is high, support for anti-immigrant parties can increase, regardless of the tone or sentiment of the reports (Eberl et al., 2018, p. 211), because "... frequent exposure to out-groups in the media acts as a reminder about people's own identities and their distinct differences from certain out-groups" (van Klingeren et al., 2015, p. 270).

In contrast, a high visibility of immigrant actors in the media can also increase positive attitudes on immigration issues (Eberl et al., 2018, p. 211). Other authors argue, however, that news coverage shapes attitudes only when it is biased in one direction or even one-sided (Boomgaarden and Vliegenthart, 2009, p. 519). Weber (2019) extended the above findings by showing that, dependent on the aggregation level, the share of immigrant people in the respective regions moderate the effects of media salience on people's attitudes. Boomgaarden and Vliegenthart (2007) found similar effects for the Netherlands, connecting their results to a higher tendency to vote for anti-immigrant parties. Walgrave and De Swert (2004) analyzed the uprising of the Vlaams Blok party in Belgium and showed a positive correlation between the growing media attention of newspapers and TV stations to immigration related topics

and the electoral growth of this anti-immigrant party. In order to replicate these findings on media salience and concerns about immigration, we formulate Hypothesis 1:

H1 Higher salience of "migration" in newspapers leads to higher concerns about immigration.

However, current studies based on agenda setting theory are limited to the salience of the migration issue and do not take into account how or in which contexts an issue is presented. In other words, they do not take into account how these media reports are framed. We will explore the framing in greater detail and connect it to people's attitudes towards migrants.

2.2 Media Frames and Audience Frames

Since an issue can be framed in various ways, its mere salience does not necessarily increase concerns about immigration. Mass media communication can be selective not only with respect to the reported issues, but also in the way of how these issues are presented to the audiences, or, in particular, in which thematic context the media reports on an issue. Frames are cognitive dispositions that guide the interpretation of the world (Kahneman, 2011, p. 413), especially if these dispositions correspond with culturally inherited taken-for-granted knowledge. A specific frame can suggest a more or less positive evaluation of a topic by contextualizing it differently. However, frames live a 'double life'. They exist in media communication as well as in the recipients' minds (Lecheler and de Vreese, 2016, p. 5). Media researchers thus distinguish between media frames and audience frames. Media frames result from how journalists

6

and the electoral growth of this anti-immigrant party. In order to replicate these findings on media salience and concerns about immigration, we formulate Hypothesis 1:

H1 Higher salience of "migration" in newspapers leads to higher concerns about immigration.

However, current studies based on agenda setting theory are limited to the salience of the migration issue and do not take into account how or in which contexts an issue is presented. In other words, they do not take into account how these media reports are framed. We will explore the framing in greater detail and connect it to people's attitudes towards migrants.

2.2 Media Frames and Audience Frames

Since an issue can be framed in various ways, its mere salience does not necessarily increase concerns about immigration. Mass media communication can be selective not only with respect to the reported issues, but also in the way of how these issues are presented to the audiences, or, in particular, in which thematic context the media reports on an issue. Frames are cognitive dispositions that guide the interpretation of the world (Kahneman, 2011, p. 413), especially if these dispositions correspond with culturally inherited taken-for-granted knowledge. A specific frame can suggest a more or less positive evaluation of a topic by contextualizing it differently. However, frames live a 'double life'. They exist in media communication as well as in the recipients' minds (Lecheler and de Vreese, 2016, p. 5). Media researchers thus distinguish between media frames and audience frames. Media frames result from how journalists

6

identify, categorize and present information, whereas audience frames describe how

recipients cognitively process this information (Scheufele, 2000, p. 306). Journalists

7

9

contribute to framing by sending frames provided by others, e.g. politicians (frame sending), to the audience, but also by setting their own frames (frame setting). Subsequently, journalists' frames translate into news frames (Brüggemann, 2014, p. 62).

Obviously, immigration and integration of migrants is a broad issue, related to many other issues. When media reports on the migration issue in combination with other issues, some of these implicitly suggest evaluations, for instance 'economic prosperity' or 'domestic violence'. Eberl et al. (2018, p. 212) point to further examples in which different frames are applied to immigrants: While immigrants from North Africa are often associated with 'threat to security', the frames of social benefits and unemployment have been applied to immigrants from Romania. Compared to that, presenting immigrants from Romania as students in British universities, which is an educational frame, is a much more positive context. Alternatively, migration could be framed by 'fragile institutions' and high homicide rates in the Global South, or by environmental collapse due to global warming, which is also a negative thematic context, but might trigger sympathy rather than xenophobia. In sum, journalists and editors can select disproportionately issue-specific frames in which immigration is associated with other topics and thereby bias the public communication (De Vreese, Peter and Semetko, 2010).

Framing is thus an important concept in media effects research, also with respect to audience frames. The sometimes ambiguous application of the framing concept across different disciplines, however, aroused suspicion to stimulate conceptual confusion (Scheufele and Iyengar, 2014, p. 2) because "the question of what exactly constitutes the frame is contested" (Lecheler and de Vreese, 2016, p. 5). Brügge-

mann (2014) defines frames as "patterns of interpretation" (p. 74) and distinguishes between various sub-types.

However, there is no consensus in the literature about defining frames. While we are unable to settle the debate on framing in media effects research, we nevertheless argue that this concept includes useful components for the explanation of media effects. As pointed out by Lecheler and de Vreese (2016, p. 4), framing in cognitive psychology is strongly influenced by Kahneman's and Tversky's studies on the sometimes surprising peculiarities of human cognition (Kahneman, 2011; Tversky and Kahneman, 1981). Different framings (e.g., either gains or losses) of logically equivalent situations can lead to completely different evaluations and decisions at the recipient's side, which is known as equivalence framing in media effects research (Scheufele and Iyengar, 2014). Emphasis framing, in contrast, highlights specific aspects of an issue (issue-specific frame) by putting it into a particular thematic context, so that a frame emphasizes specific aspects of an issue (Matthes, 2007, p. 53).

In our view, extracting topics from a large text corpus of media reports on the immigration issue is close to emphasis framing since topics relate to various other thematic contexts. Two different topics ('immigration and domestic violence' vs. 'immigration and economic prosperity') do not describe logically equivalent situations, but can nevertheless frame the issue in terms of either gains or losses (Tversky and Kahneman, 1981). It is well-established in media effects research to use topics as empirical indicators of framing. To analyze frames in media reports, Jacobi, van Atteveldt and Welbers (2016) propose to use topic models. This is in line with DiMaggio, Nag and Blei (2013), who investigate frames in government arts funding, Heidenreich et al. (2019) analyze the media framing dynamics during the 'European

9

Refugee Crisis' using topic models, as well as Adam et al. (2020), who study the discursive resonance of online climate skepticism.

At the actor's level we follow the psychological view on framing, where frames do in general shape the accessibility of cognitions in the network of humans' memory, so that "…the activation of one unit can spread through the network of interconnected units leading to the activation of related concepts" (Matthes, 2007, p. 54). In its simplest form, framing stimulates the availability heuristics. In accordance, Tversky and Kahneman (1973) explicitly pointed to the role of mass media in making information publicly available:

> Perhaps the most obvious demonstration of availability in real life is the impact of the fortuitous availability of incidents or scenarios. …many must have noticed an increase in the subjective probability that an accident on malfunction will start a thermonuclear war after seeing a movie in which such an occurrence was vividly portrait. (p. 230)

Individuals' thoughts and judgments are therefore considerably influenced by the accessibility of cognitions in the respondent's mind, and this accessibility, in turn, is a result of media frames (Matthes, 2007). Thus, by putting issues in certain thematic contexts (i.e., topics), media shapes people's perceptions, since humans rely on 'availability heuristics' to assess otherwise hard-to-observe issues like migration. The frame provided by media reports is often the only available information and, hence, decisive to raise (or reduce) concerns.

Several studies investigated the potential influence of news articles on people's attitudes towards migration, using the concept of framing. They have come to the conclusion that migrants are framed predominantly negative (Fick, 2009; Merten,

1986; Ruhrmann, 2002). In general, immigration is "predominantly discussed in negative terms" (van Klingeren et al., 2015, p. 279). The same seems to apply to the representation of Muslims and the Islam (Abadi et al., 2016; Namin, 2009; Terman, 2017). Using a theoretical framework on group threats, Schlueter and Davidov (2013) select only negative news reports and come to the conclusion that group threat is positively associated with the prevalence of negative news reports on migration. This effect is stronger in areas with a low number of migrants.

However, not all reports on migrants and migration in the media are negative (see Eberl et al., 2018). As Igartua and Cheng (2009) show in their experimental study, the context in which a news story about immigration is embedded, influences the perception of immigration as a problem. Some studies have found evidence that articles reporting positively on migrants or asylum seekers, it may decrease racial attitudes (Schemer, 2012; Schemer, 2014). We thus assume that the effect of media exposure on concerns about immigration depends on the respective content of the reports. People might evaluate migration issues positively if the media reports on rather positive aspects of migration. Thus, in contrast to previous studies, we account for the heterogeneity of topics in the information environment by utilizing topic models and, hence, consider the various ways the broad issue of 'immigration' is represented in media reports. Using a structural topic model (Roberts et al., 2014, see Section 3.2), we will establish that the issue consists of a variety of frames highlighting multiple aspects of immigration.

Emphasizing specific aspects of migration (e.g., violence, welfare system), several studies discussed above show that subsequent frames increase concerns about immigration. Our empirical analysis will corroborate this result. However, it is a neglected aspect in studies of mass communication, and not yet established at all,

lll

that positive frames do also decrease concerns of immigration. Following the concept of 'availability heuristics', we argue that presenting immigration in positive contexts shape people's perceptions towards less concerns. Thus, we assume that, net of an issue's mere salience, the availability of media reports emphasizing positive frames provides information that may lighten people's concerns in respect to this issue:

H2 Media reports on immigration that emphasize positive frames (e.g., economic prosperity, sport, education) decrease concerns about immigration.

We are well aware that it is a matter of interpretation whether a frame is 'positive' or 'negative'. To assess media frames, we use topic models (see below), an exploratory approach. The decomposition of media reports into topics allows us to identify fine-grained media frames, which we then associate with changes in people's concerns towards migration. As we will show, several topics emphasize frames eliciting connotations that are widely assumed to be positive and show a significant pattern of decreasing concerns. Without knowing the descriptive results of the topic modeling, however, the hypotheses on the influence of different topics on survey respondents cannot be more precise and, for instance, state specific frames. This reflects a general problem in the combination of descriptive computational methods of text analysis and causal methods for survey data (cf. 5). Even without specific frames though, we address a crucial research gap with H2, namely, the influence of positive media frames on people's concerns towards migration.

3 Data & Methods

3.1 Newspaper articles

The articles used in our analysis were collected via two common search engines for print media: Factiva and LexisNexis[2]. We selected weekly as well as daily newspapers which represent quality print media in Germany and cover a broad political spectrum. On the basis of seven leading newspapers, we collected a total of 24,099 valid articles ranging from 2001 to 2016.[3]

The original query yields 26,751 results. The cleaning process includes the following steps: First, duplicate analysis was performed and duplicates were removed from the dataset. Articles that consist of more than 10% English words were defined as English articles and dropped. Furthermore, articles that only contained the listing of dates for events, film screenings, panel discussions or similar were excluded. We defined articles as mere date lists if $n * c/N > 0.1$, where n is the number of digits, c the number of the word "Uhr" [4] , and N as total number of words of an article. All sentences containing less than 4 words have been removed. In addition, all URLs, words with less than three letters and 625 stopwords were excluded from further analysis. Part-of-speech tagging (POS) was applied to all sentences before the cleanup to allow lemmatization. The POS tagger was an Average Perceptron trained on the TIGER corpus with a cross-validated accuracy of .972. The GermaLemma Package with the pattern3 extension was used as a lemmatizer with a self-reported accuracy of .994. Following this procedure, all articles consisting of less than 50 lemmas were excluded from further analysis to reduce errors during the topic model (Tang et al., 2014). Finally, bi- and trigrams that occur more than ten times in the whole corpus were added. In total, from initially $26,751$ articles, we excluded

990 duplicates, 1201 short documents, 459 date lists and 2 english articles. Our final dataset therefore comprises $24,099$ valid articles containing a mean of 334 tokens per article. As already mentioned, assuming an average information environment does not necessarily mean that a particular person is exposed to reports presented in our selection of newspapers. Even though these leading newspapers strongly correspond with the general information environment, an imprecise measurement of the actual media exposure of a particular subject results in measurement error. Such an error usually results in attenuation (or 'dilution') of the coefficients, that is, they become biased towards zero (Skrondal and Rabe-Hesketh, 2004, p. 76) and thus lead to conservative estimates. Again, it is the benefit of the FE model to control for stable subject-heterogeneity, e.g., gender, ethnic group, education, personality traits or 'milieu'.

3.2 Structural Topic Model

Even though circa 24.000 articles do not qualify as "Big Data", it still exceeds numbers that could be analysed by qualitative means. Working with such large amounts of texts is a long-standing issue in the field of information retrieval (Deerwester et al., 1990). Therein, the main idea is to summarize a bunch of text documents (the corpus) by reducing their dimensions but to keep most of its relevant information. One popular branch of information retrieval is topic modeling (Jordan and Mitchell, 2015), where a set of documents is assigned to meaningful themes (i.e., topics). These topics are directly derived from the documents by probabilistic algorithms and rely on the notion that words co-occurring in and across documents describe meaningful topics. All words are thereby assigned to all topics, dependent on their context with different association strength (β).

In so-called generative models, each topic is seen as a probability distribution across all words of a given language, describing the likelihood for a chosen word to be part of a certain topic (Griffiths and Steyvers, 2004; Hoffman, Blei and Bach, 2010). Since this likelihood is independent of the position of the word in a text it is sometimes referred to as a "bag-of-words" representation of documents. Although this assumption is clearly not realistic (e.g., grammar is ignored), it has been proven to be very reliable in practical applications and has important applications to the social sciences in particular (DiMaggio, Nag and Blei, 2013; McFarland et al., 2013)

A popular instance of generative models is Latent Dirichlet Allocation (LDA) proposed by Blei, Ng and Jordan (2003). Given a desired number of topics k and a set of D documents containing words from a vocabulary V, LDA models infer k topics which are each a multinomial distribution over words V. Thus the topics are a mixture of words V with probability β, mentioned above, for each word and its association to a topic. The more often words co-occur in documents, the higher the probability that they constitute a topic. At the same time, a document is also considered as a mixture over topics, so that a single document can be assigned to multiple topics. The topic proportions are given by parameter θ. By design, all topics occur within each document, however, the proportion of θ gives us the strength of connection between a topic (itself an ordered vector of words) and a document. Finally, it is important to note that the sampling process of LDA uses the same multinomial distribution (the eponymous Dirichlet distribution) for all documents in a corpus.

In this paper, we use a recently developed advancement of probabilistic topic models called structural topic model (STM) (Roberts et al., 2014). Its key feature is to enable researchers to incorporate document metadata and utilize such information

15

(e.g., year, source, etc.) to improve the estimation of topics. It has been argued that including the date of a document is especially useful for time periods and changing discourses (Farrell, 2016). The covariates of a document d are denoted as X_d. The basic model relies on the same LDA process explained above, each document is still assumed to contain a mixture of k topics[5] and words are aligned to topics with a certain probability β. Opposed to a 'normal' LDA, in a STM the topic proportions θ depend on a logistic-normal generalized regression, such that $\theta \sim LogisticNormal(X_d, \Sigma)$. Thus, for each word a topic is drawn from the specified distribution for one document based on its covariates values X_d, which, in addition, provides a measure of topic prevalence.

In short, conditioning the word and document distribution on additional information about the documents allows the STM to base a word's topic assignment on a document-specific distribution, not only – as in the regular LDA (respectively CTM for correlated topics; see Blei and Lafferty, 2007) – on a general distribution that is the same for all documents. It has been shown that the incorporation of covariates improves the results of the topic quality substantially (Roberts, Stewart and Airoldi, 2016; Roberts et al., 2014). Or in the words of its developers: "These additional covariates provide a way of 'structuring' the prior distributions in the topic model, injecting valuable information into the inference procedure" (Roberts, Stewart and Airoldi, 2016).

Using STM allows us to improve the measurement of media salience and to qualify media reports' content over a comprehensive sample of texts. We also employ 'sentiment analysis', where researchers refer to databases with information on the evaluation of particular terms. The aim of sentiment analysis is to reveal the evaluation of an issue in terms of 'good or bad'. This is done in many text analyses in,

116

for example, political science, since it allows a standardized analysis of how actors evaluate particular issues (Liu, 2012). However, German mass media have a 'code of conduct' which prescribes that reports should not be biased negatively with respect to immigration and should not increase prejudice in the audience. As a consequence, stereotyping reports are criticized and reported to the German Presserat. Thus, we do not expect strong sentiments in news articles from quality mass media.

However, we do assume that migration is framed in many different ways. Reports on migration and integration might focus on areas such as labor, refugees, family, 'welcoming culture' or education. Therefore, considering the frames of media reports on migration captures more variance than an estimate of the issue's raw media salience (Czymara and Dochow, 2018), and allows us to measure whether the exposure to different frames leads to different effects — negative and positive — on concerns about immigration.

3.3   German Socioeconomic Panel

As outlined in Section 2, there exists a rich literature on understanding and tracing media discourse on migration. We seek to extend previous attempts by matching "ups and downs" of media discourse to panel data reporting on attitudes towards migration. For that purpose, we use the German Socio-Economic Panel (GSOEP) (G. G. Wagner, Frick and Schupp, 2007). The GSOEP is a nationwide household survey based on annually repeated interviews conducted by the German Institute for Economic Research (DIW). Starting with the first wave in 1984, it is now one of the leading household panel studies in the world. It currently includes around 30,000 respondents in around 15,000 households. The survey design nowadays consists of several sub-samples, some of which collected explicitly with the aim of compensat-

ing panel attrition and respondents mortality (G. G. Wagner, Frick and Schupp, 2007). In order to allow projections from this data on the overall population in the German Society, the DIW provides survey weights that account for the complex survey design. The representative survey provides yearly information on political attitudes and socioeconomic characteristics[6]. GSOEP participants are asked about their concerns about certain topics. Among those are, for instance, concerns about the economic development. Our dependent variable is the question if the respondent is concerned about immigration with answers given on a three-point scale (3: "not concerned", 2: "somewhat concerned", 1: "very concerned", variable name is `plj0046`). We binarize information by combining 3 and 2, and contrast it with those being "very concerned" (similar procedures are used in Lancee and Pardos-Prado, 2013).[7] Descriptive statistics for all GSOEP-variables can be found in Appendix B.

Unlike Czymara and Dochow (2018), who also examined the connection between concern about immigration and the media salience of immigration issues using the GSOEP, we do not assume that concerns do necessarily indicate negative attitudes about immigrants. It is rather misleading not to clearly distinguish out-group rejection, stereotype and prejudice on the one hand, and "concerns about immigration" on the other hand. Immigration can also be an indicator of global inequality or of severe problems in the sending countries, for that concerns might not be primarily related to stereotypes (Collier, 2013; Windzio, 2018). Instead, we assume more conservatively that such concerns emphasize the sheer importance of an issue in respondent's mental frame (Wlezien, 2005). In so doing, we relax the assumption that issues of importance do necessarily increase negative stereotyping and relate respondent's concern to the salience and framing of an issue.
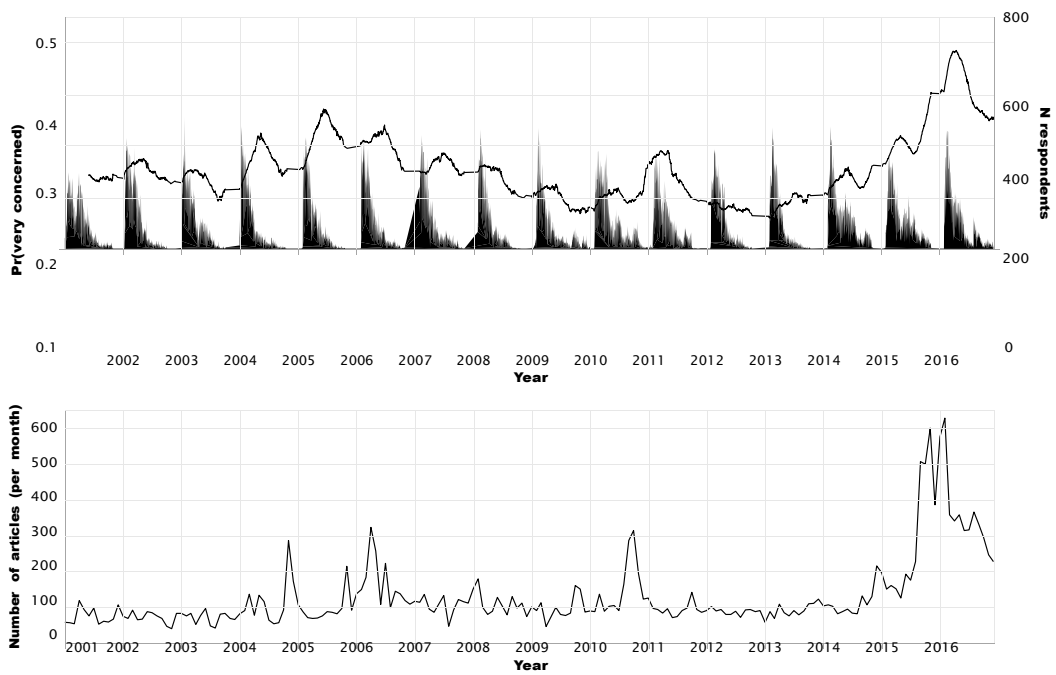
Figure 1: top: the bars depict the number of actual interviews per date; the line shows the percentage of people, who gave "being very concerned" as an answer (rolling mean, lag 150 days); bottom: shows the number of articles in the dataset over time (per month).

119

For our purposes, we need to match media salience and prevalence of topics before and at the time of each interview. Our treatment is the exposure of respondents to the media discourse. It is important for our argument that GSOEP interviews are spread over the entire year, since our central assumption is that interviewees are influenced by current media discourse. Fortunately, they do, as Figure 1 shows. We can therefore make use of changing topics in media discourse. Unlike previous studies, we decompose the German discourse on immigration into topics (an overview of topics is given in Section 4.1). This allows us to trace which topics are prevalent and which are marginal at the time a GSOEP interview takes place and respondents declare how concerned they are about immigration.

To trace the discourse, we normalize each topic's θ (topic load) by z-scaling it. In so doing, we derive a clear metric fitting the fixed effects regression, i.e., whether a topic is more or less prominent in German media discourse (i.e., more "available") than its average between 2001 and 2016. In order to link these ups and downs in media discourse statistically to people's attitudes, we employ linear fixed-effect panel regressions (within-transformation) (Wooldridge, 2002). We discuss this choice in comparison to logit fixed-effect models in Appendix E.

Even though panel attrition and the selectivity of data is always a problem in empirical studies, effects of exposure to the respective information environment can be nevertheless captured by the fixed effects (FE) panel model, which estimates within-effects of changes in the explanatory variable on changes on the outcome. The model controls for any influences of observed and unobserved time-invariant characteristics of the survey respondents. Even if the data would suffer from panel attrition (for which GSOEP is designed to take care of), the model would nevertheless estimate the effect for the remaining sample, which is still much more heterogeneous than,

for example, an experiment based on students only. Moreover, the FE model algebraically removes unobserved heterogeneity at the subject level (Wooldridge, 2002), which includes in our case time-constant traits such as personality characteristics, ethnicity or e.g. particular groups who categorically refuse or are unable to participate in the average information environment for whatever reason. However, the FE model does not account for time-varying unobserved heterogeneity. For instance, respondents could change their media use and switch to another newspaper or media type because its reports on migration better fit to the respondent's worldview.[8] This might surely happen, but our analysis of media communication about immigration describes the dynamic, general information environment generated by German quality newspapers, rather than the exposure to a specific medium or newspaper. Since the selected newspapers are opinion leaders in Germany, we assume that even if not reading those, the agenda setting across all media is largely correlated with prevalent topics in leading print media. In addition, we can show that these newspapers do not considerably differ in sentiments (Appendix I) of their reports nor in topic distributions (Appendix C, Figure 7). Hence, since we analyze effects of this general information environment on respondents' attitudes, self-selection into particular forms of media use due to time varying-processes should not severely bias our estimates.

FE also assume strict exogeneity, which is highly likely in our case since news coverage in leading print media at the day of the interview is not influenced by the same individual characteristics that may affect concerns about migration. As main explanatory variables we use the average topic prevalence over all articles in the 28 days before each interview[9]. This idea expands the work of Czymara and Dochow (2018), who used FE regressions on "concern about immigration", using the number

21

of articles 28 days before each interview as their main explanatory variable. Finally, focusing on within-variations and thereby exploiting the panel structure, we statistically account for all constant person specific attributes which might confound results like social class, race, or sex, i.e., the model eliminates time-constant unobserved heterogeneity.

4 Results

4.1 German Media Discourse on Migration

All topics [10] of the German media discourse on migration, which we interpret as frames, are presented in Appendix F. Each topic consists of an ordered vector of words (β from Section 3.2). The higher a word's rank, the more descriptive it is for the topic. For instance, words like "refugee", "syria", or "mediterranean", make it intuitively clear that Topic50 represents reports on refugees from Syria and other nations surrounding the Mediterranean Sea. The most probable words reveal how we encounter a topic in most newspaper articles (for Topic50, e.g., "refugee", "europe", "syria"), while most exclusive terms (FREX) have a high likelihood to appear in articles devoted exclusively to a certain topic (for Topic50, e.g., "mediterranean", "unhcr", "frontex"). Both metrics are complementary and describe a topic in its most likely and most exclusive configuration. In the remaining paper we will use FREX terms to describe topics unless otherwise specified.

In addition to β (degree of words describing topics), each topic is assigned to all documents with a certain strength (θ). Summing up this topic load over all documents gives us the prevalence of each topic, i.e., a topic's share in the German print media discourse on migration. Figure 2 depicts its most important topics
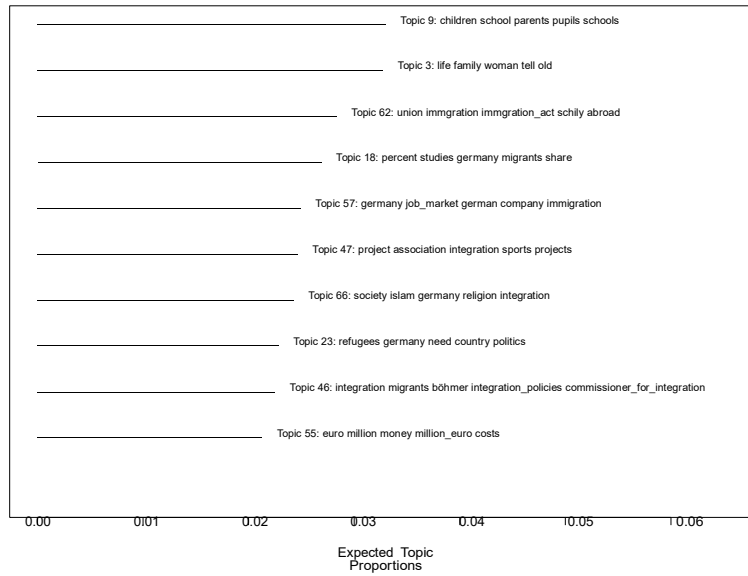
22

Figure 2: Topic Prevalence. Share of 10 topics with largest prevalence in the German discourse on migration, 2001 to 2016. We depict most associated words for each topic (Column "Prob." of the table in Appendix F).

from 2001 to 2016. Topic9 is the topic which appears most often. We labeled it Education, since its content focuses on concepts like "children", "schools", "pisa", or "school system". It is interesting to note that Education occupies the highest share, since the corpus is selected on migration and integration of migrants. Therefore, many articles written on the migration issue embed it into a frame concerning the German educational system which is known to disadvantage children with migration backgrounds (Dollmann and Weißmann, 2019).

Another large portion of the discourse is dedicated to Immigration Law (Topic62). Substantial reforms took place during the early 2000s under chancellor Gerhard Schröder, which resulted in a general revised law on migration. Efforts led by then-to-be Secretary of the Interior Otto Schily, were made possible through negotiations with the CDU led by Peter Müller. Both politicians appear prominently in Topic62.

23

Thus, the words describing Topic62 contain the story of an important legislative reform in Germany.

Topic23 Refugees:General concentrates on more recent events and relates to many other topics. Due to its generality it occurs in many documents and is therefore highly influenced by n.Articles[11]. It is broader than Topic50 Syria and covers the refugee crisis with various aspects ranging from "Willkommenskultur" (welcome culture) to the discussion of a numeric limit of immigrants ("Obergrenze" (upper limit)). A similar prevalent topic is Topic66 Islam. While Topic50 or Topic23 focus on different events of the refugee crisis, Topic66 relates to discussions on societal scale, i.e., how "multiculturalism", "religion", and "fundamentalism" can be coped with in a liberal democracy. Clearly, those questions center around the Muslim religion and how it might be aligned with Germany's mostly Christian tradition (expressed in discussions on "Leitkultur" (dominant host culture)).

Those examples express the variety of the discourse on migration in German media. While it might be a worthwhile task to explore the topics in greater detail, our paper's central goal is to link these topics to changes in attitudes of the German people in order to examine the influence of media salience on people's perceptions of migration.

4.2 Frames Shape People's Attitudes on Migration

In line with previous studies (Czymara and Dochow, 2018; Pardos-Prado, 2011), fixed-effects regressions on SOEP data reveal several attitudes associated with German's concerns about migration. Figure 3 shows the effect sizes and 83% confidence intervals for selected control variables in our model (all results are also reported in Appendix G). Most concerned are people who align themselves to parties of the far-

right. However, the number of people indicating to vote for parties like the NPD ("National Party of Germany") is rather low. The effect of voting for the moderate right-wing CDU ("Christian Democratic Union") is considerably lower, while the effect for the left-leaning, immigration-friendly Green party ("Gruene") points in the opposite direction of decreased concerns.
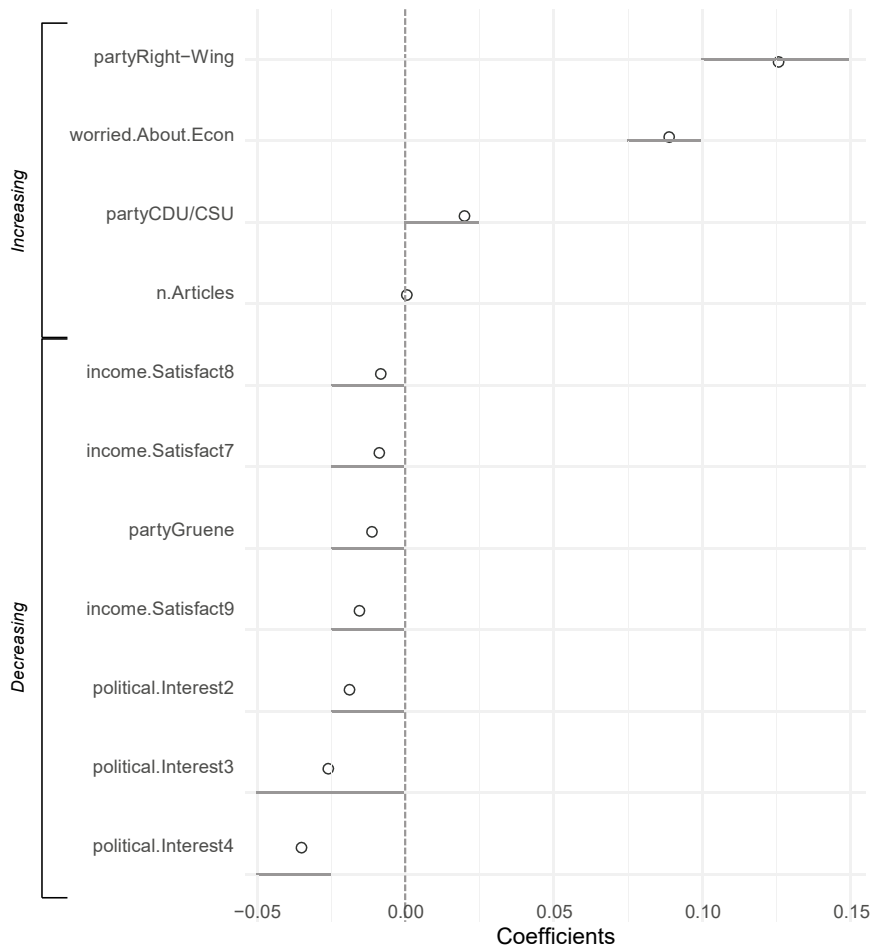
Figure 3: Structural effects in linear fixed effects model. Confidence intervals are at 83 % to allow visual inspection of significant differences (Payton, Greenstone and Schenker, 2003). Only significant effects are reported, all coefficients can be found in Appendix G. Reference values are: party = no party preference; political.Interest = 1 (strong interest in politics); income satisfaction = 6 (low income satisfaction).

While no other political orientation expressed by party preference has a significant effect, the level of political interest (political.Interest) is highly influential. Chances of being "very concerned about migration" clearly decreases with less interest in politics (reference is 1 representing "strong interest in politics"). This means, in turn, that

26

people who think a lot about political issues are more likely to be concerned about migration. It reflects the mechanism proposed by the availability heuristic: only if you are interested in the (larger) subject, you get a "chance" to be worried. The same logic applies when it comes to worries about economic progress (worried.About.Econ). People who are comfortable with the present economic situation are less concerned about migration, vice versa. In that sense, "worry feeds worry".

A small but significant positive effect on the concerns about immigration is illustrated by the number of articles that appeared in the 28 days before questioning. The more articles were published in this time span, the more likely it is for individuals to be very concerned. Those results are in line with theory and mirror previous research by Czymara and Dochow (2018), and Weber (2019). Herein, we see support for our hypothesis H1 which refers to agenda setting and media salience. Yet, our main goal is to trace whether exposure to different topics (i.e., frames) leads to an increase, or decrease respectively, in concerns about immigration. It is important to note that all effects of those frames are net of n.Articles.
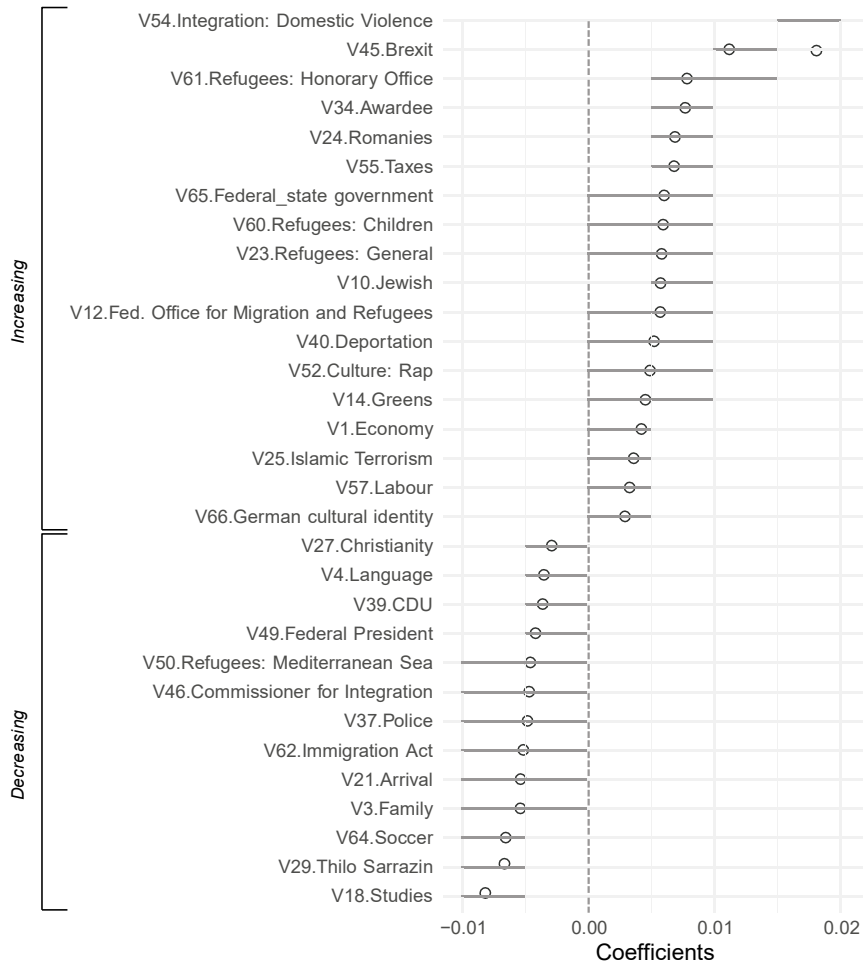
Figure 4: Fixed-effects regression coefficients of selected topics, ordered by effects' strength. Only highly significant ($\alpha < 0.01$) topics are depicted. The regression rests on topics' average prevalence 4 weeks before someone completed the survey. Our findings are robust to varying time windows as reported in Appendix C.

To disentangle media reports' content, we already described the most salient migration frames (cf. Section 4.1). Now, we trace their influence on migration attitudes by including normalized topic loads (z-scale θ) into the fixed-effects panel regression. This means, we measure the ebb and flow of topics before respondents filled out the survey questions. In so doing, we assume to capture a large part of

28

information available for each participant and his or her subsequent heuristics on an issue which is itself directly not observable for individuals. Given the importance of newspapers in German media landscape, migration frames should be captured regardless of individual media preferences. We use a four week window before each participant's interview. Results are robust to the number of weeks (cf. Appendix C).

Most associated with concerns about migration are articles on Domestic Violence (Topic54). Its coefficient is highly significant and with $\beta = 0.018$ comparable to effects of income satisfaction or the strongest effects of party preferences (i.e., Green Party and CDU, respectively). Topic54 reflects articles on marriage, highlighting cultural differences by focusing on domestic violence against women in families with migration background. Hence, it paints troubling images about "women", "(sexual) violence", and "forced marriages" (cf. Appendix F). Consequently, its sentiment is clearly negative (cf. Appendix I). Topic54 also contains a good portion of the discussion around headscarf bans, which is discussed now and then in the media – in particular, when it comes to young women and teachers. Accordingly, it represents problematic contents and its salience drives concerns of survey respondents.

In line with a more cautious interpretation of the dependent variable (cf. 3.3), we observe that concerns are also fueled by other, more general media content. The second strongest influence is due to Topic45 Brexit. Topic45 touches the uncertainty surrounding UK's exit from the European Union. Another example for a rather general media frame increasing concerns is Topic55 Taxes. Topic55 correlates with increased worry connected to state finance's and, hence, potential monetary issues of migration.

Most of the other significant topics driving Germans' concerns about migration are related to reports on specific cultural expressions (topics 10, 23, 24, 52, 60, 66), terrorist threats (topic 25), the economy (topics 1, 57) or state affairs (topics 12, 14, 40, 65). All of those topics reflect subtle, but problematic issues with the potential for economic, political or cultural conflicts[12].

However, opposed to previous research, we also find strong effects of an individual's exposure to specific frames decreasing concerns about migration. Using topic models reveals that certain topics lighten concerns about migration, even when media reports thrive on that issue. The presence of Topic18 has the strongest decreasing effect. It describes articles on scientific studies in Germany's print media and is associated with significant less concerns in the GSOEP. Topic18 Studies refers to a vocabulary that sounds familiar to social scientists, e.g., "respondents", "statistically", "percent german", "survey", or "statistical office". Its strength is comparable to expressing voting preferences for the left-wing, pro-immigration Green party. Thus, frames in information-heavy reports referring to scientific studies in the context of migration decrease concerns about that issue. That might be an interesting side-note in times of worries on simplistic "fake-news" in media outlets (Bennett and Livingston, 2018).

Thilo Sarrazin (Topic29) has been mainly criticized in print media reports for representing controversial positions on immigration and integration, e.g., when the Social Democratic Party debated on his exclusion. According to our results, the overall negative presentation of Sarrazin in high-quality newspapers reduces concerns about immigration.

Soccer (Topic64) has another strong and decreasing effect on concerns. Here, it is important to note that all articles share keywords on migration or the integration

of migrants. Thus, Topic64 does not include (prevalent) reports on soccer results or transfers. Rather, it covers news on football players with migration background or clubs with many migrants. Subsequent articles, hence, focus on "successful" integration of migrants, for instance, about prominent, much adored football players like Mesut Özil.

Other topics (Topic21 Arrival or Topic3 Family) point in the same direction. When media report on contexts of migration or everyday-life situations and struggles (e.g., learning the language, family business), people seem to sympathize with migrant's situation and concerns about migration decrease. In contrast, reports on violence (Topic54), miserable conditions (Topic60) or terrorist threats (Topic25) increase concerns, in line with assumptions of an increased perception of uncertainty when it comes to state business (e.g., Topic45). Considering the different effects for the mentioned topics, we find support for our hypothesis H2.
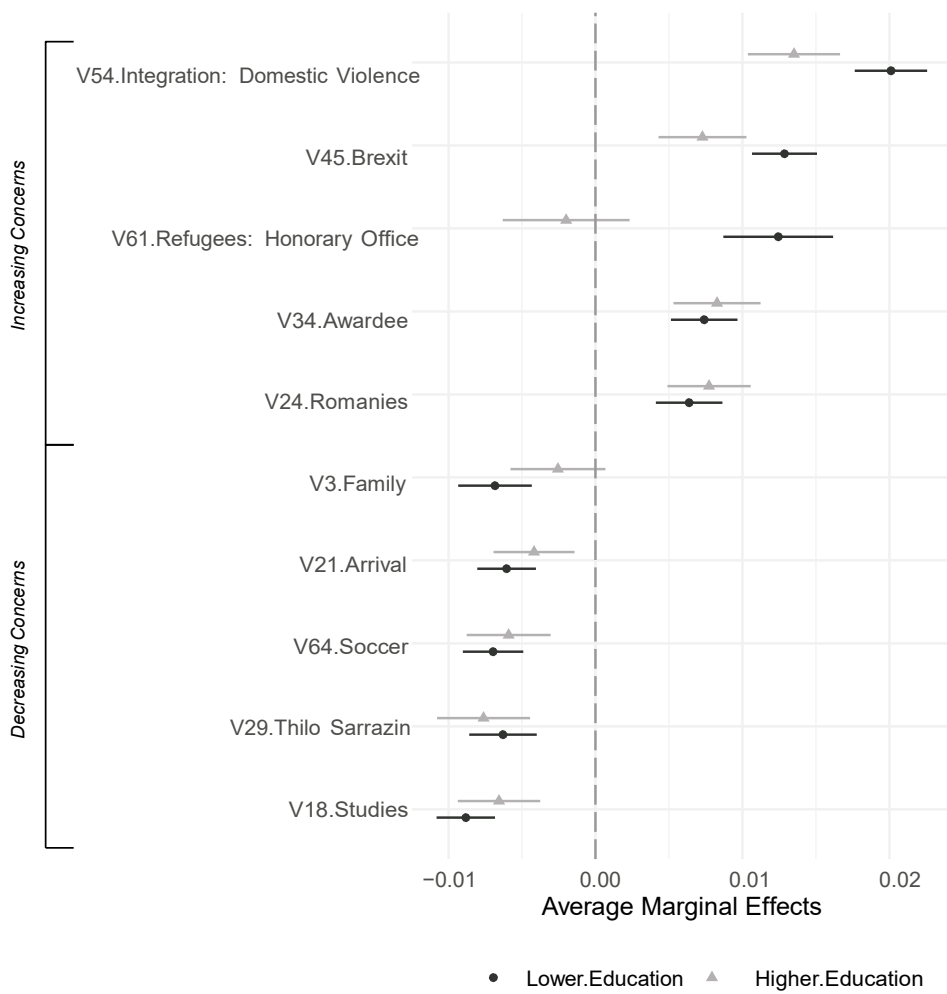
Figure 5: Marginal effects of selected media topics' prevalence on concerns about immigration conditional on education, with 83 % confidence intervals. Topics are selected by strength of effects as reported in Model 3 (cf. Appendix G). We depict only the 5 most decreasing and increasing effects. Marginal effects for all other topics can be provided upon request.

While we find considerable and diverse effects of media content on people's attitudes, one potential confounder of our analysis is that the strength of perception is

not uniformly distributed. Education might simultaneously affect the access to specific media content and attitudes towards immigration. We test therefore whether different levels of education mediate the impact of available information on one's attitudes towards minorities (Hainmueller and Hiscox, 2007). For that purpose, we consider the 5 strongest effects in both directions, i.e., decreasing and increasing concerns, and examine those conditioned by participants' education by utilizing average marginal effects.

Figure 5 shows that in most cases topic effects are reduced for people with higher education, i.e., media salience affects people with lower education more strongly. One intuitive explanation might be that people who lived through Germany's higher education system rely on other, maybe more diverse sources of information. Yet, most differences are not very pronounced and statistically indistinguishable. The difference between levels of education is only significant for Topic54 (Domestic Violence), Topic45 (Brexit), and Topic61 (Refugees: Honorary Office). While the concern-increasing effects of Topic54 and Topic45 are still considerably high for people with a higher education, reports on Topic61 has no longer a significant effect for that group. Still, most effects, in particular those with decreasing concerns, are very similar between different levels of education.

Yet, it might be argued that it is not the content but the tone of media reports that drives our effects. Analyzing the tone of newspaper articles about immigration in their study, van Klingeren et al. (2015) found a very limited effect of the positive tone on people's attitudes in the Netherlands but could not show the same for Denmark. To ensure that it is not the tone but the actual content represented by topics that drives our results, we provide a sentiment analysis of our corpus in Appendix I.

It shows that topics are not closely linked to specific sentiments underlining the "neutral" tone used in German quality newspaper.

In addition, our results might rest on different probabilities of topics to appear in specific newspapers. Figure 7 (Appendix C) shows that this is not the case. Instead, topic distributions are very similar across media outlets, indicating that we observe a general information environment and not only thematic preferences of certain media.

5 Discussion

We rely on the framing concept in media research and argue that the thematic context of media reports imposes a particular frame on mass communication. At the level of media recipients, we focus on 'heuristics' of human cognition (Kahneman, 2011; Tversky and Kahneman, 1973). Following this prominent notion, humans depend in complex situations on the availability heuristic: if it is hard to generate a comprehensive representation of a given situation from our immediate environment, we do rely on information at hand. Hence, human's perception of macro-societal issues that are non-observable for a single individual, like immigration, depends on reports from mass media. Mass media provides cognitive access to complex macro-level environments, yet, its communication is highly selective and specific (Luhmann, 2002).

Previous literature stated that mere frequency of media reports on immigration would lead to negative stereotyping and prejudice (Czymara and Dochow, 2018; Weber, 2019). Beyond that notion, our approach combines the prevalence of media frames over time with longitudinal survey data. Using daily and weekly newspapers

generates a time-dependent corpus on the issue of immigration and integration. In line with previous studies, we expected an overall positive effect of media salience

of 'immigration' on concerns (H1). Our results corroborated those general findings: increasing prevalence of immigration-related reports increases concerns.

However, we object the assumption that mere media salience of such a complex issue generally results in concerns about immigration. To gain a more fine-grained perspective on media's content, we apply structural topic models (Roberts et al., 2014). The topics represent thematic contexts embedding the immigration issue, and thereby imposing particular frames on the issue. Underlining the importance of available information provided by media coverage, we find specific topics with negative contents to increase concerns (Igartua and Cheng, 2009; Schlueter and Davidov, 2013), while others substantially decrease concerns. Combining topics with fixed-effects panel regressions reveals that the exposure to topics such as Scientific Studies or Soccer, which have a rather positive connotation, significantly weakens concerns about immigration.

It is an important insight from our study that available information represented in news articles can diminish concerns about a sensitive issue like immigration. It must remain open though whether the effects are influenced by different patterns of media usage. While our corpus covers a broad sample of most important German quality newspapers and we found only small moderating effects of different educational levels, we cannot preclude that, for instance, specialized newsgroups in the internet or communications with peers do not affect people's attitudes differently. Indeed, this might be one of the most important current challenges of analyzing public discourses in the media: is there a general discourse, as we assume here by focusing on popular German quality newspapers, or are we increasingly dealing with a fragmented public where different discourses are held in a variety of 'filter bubbles' and 'echo chambers'? If these alternative media outlets would have a strong impact on the overall

information environment in the population, and the content of communication differed between these media outlets, our focus on quality newspapers would be affected by measurement error. The effect of this error would be, however, to bias the estimate towards zero (Skrondal and Rabe-Hesketh, 2004, p. 76), and therefore drive an underestimation of the 'true' effect. In other words, due to such measurement error our estimates are rather conservative. Therefore, researchers should avoid including strongly biased and socially selective discourses into their data (Ruths and Pfeffer, 2014), unless empirical research can show that these discourses have a high outreach in the society and can influence a considerable share of the population. Aside from that, we applied FE models for panel data, as suggested by Czymara and Dochow (2018), that control for self-selection into particular forms of news consumption due to time-constant characteristics at the subject-level.

In any case, analyzing media effects, particularly with respect to important issues such as attitudes towards immigration, should not neglect the content of media reports. Our approach to include the content via a larger set of topics from comprehensive text data in combination with traditional panel surveys might be a potential way to examine various discourses and their influence on people's attitudes empirically. We believe that the coupling of natural language processing and longitudinal personal data offers great potential to analyze social processes. Yet, bringing together topic models and panel data to track changes in people's concerns towards migration implies a serious pitfall: the outcome of our exploratory analysis is a description of topics, and it is difficult to hypothesize on their effects without knowing the topics in advance. Estimating effects of topics on individual outcomes measured in survey data is thus somewhat incompatible with empirically testing a priori proposed hypotheses. For this reason, we limited our main hypothesis to the effect of 'positive'

37

contexts. Admittedly, the attribute 'positive' is a matter of interpretation, and this underscores again that standardized methods of text analysis strongly depend on the researchers interpretive efforts ((Grimmer and Stewart, 2013)). Thus, how to "translate" topics – which are the result of a massive dimension reduction – statistically, i.e., how to best measure the influence of text-as-data on human behavior is an open question. We answered it in this paper by using multiple ways of measurement to check the robustness of our findings.

More generally, future research should use better indicators of attitudes to immigration and integration of immigrants. The indicator "concerns about immigration" is an imperfect measurement for negative attitudes towards immigration and negative stereotyping because it combines many aspects in which people may express concerns without being generally prejudiced against immigrants. During the "refugee crisis" in 2015 reports on large crowds of refugees marching on streets and highways towards Europe certainly made an impression, but these reports were closely linked to civil wars in Syria, Iraq and Afghanistan. Being "concerned" while watching or reading these touching reports may be the result of a variety of mechanisms associated with the overall topic of immigration and integration of immigrants. With a more fine-grained question, researchers could probably disentangle the complex relationships of concerns and uncertainties. Our study addresses at least some of these aspects and their effects on concerns about immigration by not just analyzing the prevalence of immigration-related reports but also taking their thematic context (i.e., frame) into account. As our results demonstrate, media can therefore influence our attitudes in both directions, towards rejection of an issue as well as towards agreement – conditional on the specific frame that is provided by key media outlets.

38

But why is human cognition highly sensitive towards media framing at all? A famous example of how framing works at the level of human cognition is prospect theory (Kahneman and Tversky, 1979). Humans are more affected by the prospect of losing than gaining something. Prospect theory provides a potential explanation of why media reports on immigration tend overall to increase negative attitudes towards migrants (Czymara and Dochow, 2018). Possibly, media reports frame the issue more often in terms of potential losses (e.g., redistribution of welfare or specific norm-violations) than gains (e.g., qualified human capital or cultural diversity). Future research could test whether reports on immigration do indeed include aspects of gains and losses and, if so, how these aspects influence attitudes towards immigration. If media reports on immigration apply mostly frames that trigger recipients' high sensitivity towards loss aversion, recipients will tend to increase negative attitudes. As our results show, more positive connoted media frames, however, do also shape attitudes towards less concerns about migration (and, potentially, other sensible issues). In this light, media's responsibility for a careful selection of its contents seems therefore more important than ever.

Notes

[1] We are fully aware of the ongoing research on second level agenda setting. In line with Ghanem (1997), who points out that "the principal difference between the research literature on frames and on the second level of agenda setting is that the latter examines the impact of news frames on the public agenda", we stick to the theory of framing as we are interested on effects on the individual rather than the public agenda.

[2] Available at http://www.factiva.com and http://www.lexisnexis.com, respectively.

[3] Query and frequency of articles by newspaper are presented in Appendix A. Python code for cleaning the articles is available upon request.

[4] While in the English language, time specifications are usually ended with either a.m. or p.m., in German, only the word "Uhr" is used.

[5] In this case, we used 69 topics; an explanation for this number can be found in Appendix H.

[6] We dropped all respondents younger than 18 (not entitled to vote), who are still in school or participated less than two times.

[7] A different coding of the dependent variable is part of our robustness checks, cf. Appendix C.

[8] Unfortunately, the GSOEP is not asking what media types respondents consume primarily.

^{9} Our findings are robust to varying time windows, cf. Appendix C.

^{10} All necessary data to reproduce are model will be made available on `https://github.com/luerhard/media_effects_2001_2016` upon publication of this article

^{11} For a sensitivity analysis and multicollinearity checks see Appendix J

^{12} For a closer interpretation of all topics see Appendix F. In addition, please note that the outlined categories do not contain Topic34 Awardee and Topic61 Honorary Office, because both topics' effects are not consistent across different model configurations (cf., Appendix J).

References

Abadi D., d'Haenens L., Roe K., Koeman J. (2016). Leitkultur and Discourse Hegemonies. International Communication Gazette, 78, 557–584.

Adam S., Reber U., Häussler T., Schmid-Petri H. (2020). How Climate Change Skeptics (Try to) Spread Their Ideas: Using Computational Methods to Assess the Resonance among Skeptics' and Legacy Media. PLOS ONE, 15, 1–23.

Allison, P. D. (2009). Fixed Effects Regression Models. Los Angeles: SAGE.

Beck N. (2020). Estimating Grouped Data Models with a Binary-Dependent Variable and Fixed Effects via a Logit versus a Linear Probability Model: The Impact of Dropped Units. Political Analysis, 28, 139–145.

Bennett W. L., Livingston S. (2018). The Disinformation Order: Disruptive Communication and the Decline of Democratic Institutions. European Journal of Communication, 33, 122–139.

Blei D. M., Lafferty J. D. (2007). A Correlated Topic Model of Science. The Annals of Applied Statistics, 1, 17–35.

Blei D. M., Ng A. Y., Jordan M. I. (2003). Latent Dirichlet Allocation. Journal of Machine Learning Research, 3, 993–1022.

Boomgaarden H. G., Vliegenthart R. (2007). Explaining the Rise of Anti-Immigrant Parties: The Role of News Media Content. Electoral Studies, 26, 404–417.

Boomgaarden H. G., Vliegenthart R. (2009). How News Content Influences Anti-Immigration Attitudes: Germany, 1993-2005. European Journal of Political Research, 48, 516–542.

Brüggemann M. (2014). Between Frame Setting and Frame Sending: How Journalists Contribute to News Frames: Between Frame Setting and Frame Sending. Communication Theory, 24, 61–82.

Chyi H. I., McCombs M. E. (2004). Media Salience and the Process of Framing: Coverage of the Columbine School Shootings. Journalism & Mass Communication Quarterly, 81, 22–35.

Collier, P. (2013). Exodus: How Migration Is Changing Our World. Oxford ; New York: Oxford University Press.

Czymara C. S., Dochow S. (2018). Mass Media and Concerns about Immigration in Germany in the 21st Century. European Sociological Review, 34, 381–401.

De Vreese C. H., Peter J., Semetko H. A. (2010). Framing Politics at the Launch of the Euro: A Cross-National Comparative Study of Frames in the News. Political Communication, 18, 107–122.

Deerwester S., Dumais S. T., Furnas G. W., Landauer T. K., Harshman R. (1990). Indexing by Latent Semantic Analysis. Journal of the American Society for Information Science, 41, 391–407.

DiMaggio P., Nag M., Blei D. M. (2013). Exploiting Affinities between Topic Modeling and the Sociological Perspective on Culture. Poetics, 570–606.

Dollmann J., Weißmann M. (2019). The Story after Immigrants' Ambitious Educational Choices: Real Improvement or Back to Square One? European Sociological Review, jcz042.

Eberl J.-M., Meltzer C. E., Heidenreich T., Herrero B., Theorin N., Lind F., Berganza R., Boomgaarden H. G., Schemer C., Strömbäck J. (2018). The European Media Discourse on Immigration and Its Effects: A Literature Review. Annals of the International Communication Association, 42, 207–223.

Farrell J. (2016). Corporate Funding and Ideological Polarization about Climate Change. Proceedings of the National Academy of Sciences, 113, 92–97.

Fick, P. (2009). Der Wandel Der Darstellung von Migranten Am Beispiel Siegener Lokalmedien in Den Jahren 1996 Und 2006. In Geißler R., Pöttker H. (Eds.), Massenmedien Und Die Integration Ethnischer Minderheiten in Deutschland. Forschungsbefunde. Band 2. Bielefeld: Transcript Verl., pp. 235–270.

Ghanem, S. (1997). Filling the Tapestry: The Second Level of Agenda Setting. In McCombs M. E., Shaw D. L., Weaver D. (Eds.), Communication and Democracy: Exploring the Intellectual Frontiers in Agenda-Setting Theory. London: Lawrence Erlbaum Associates, pp. 3–14.

Griffiths T. L., Steyvers M. (2004). Finding Scientific Topics. Proceedings of the National Academy of Sciences of the United States of America, 101 Suppl 1, 5228–5235.

Grimmer J., Stewart B. M. (2013). Text as Data: The Promise and Pitfalls of Automatic Content Analysis Methods for Political Texts. Political Analysis, 21, 267–297.

Gruszczynski M., Wagner M. W. (2017). Information Flow in the 21st Century: The Dynamics of Agenda-Uptake. Mass Communication and Society, 20, 378–402.

Hainmueller J., Hiscox M. J. (2007). Educated Preferences: Explaining Attitudes Toward Immigration in Europe. International Organization, 61.

Heidenreich T., Lind F., Eberl J.-M., Boomgaarden H. G. (2019). Media Framing Dynamics of the 'European Refugee Crisis': A Comparative Topic Modelling Approach. Journal of Refugee Studies, 32, 172–182.

Hoffman M. D., Blei D. M., Bach F. (2010). Online Learning for Latent Dirichlet Allocation. Proceedings of the 23rd International Conference on Neural Information Processing Systems - Volume 1. Red Hook, NY, USA: Curran Associates Inc., 856–864.

Igartua J.-J., Cheng L. (2009). Moderating Effect of Group Cue While Processing News on Immigration: Is the Framing Effect a Heuristic Process? Journal of Communication, 59, 726–749.

Jacobi C., van Atteveldt W., Welbers K. (2016). Quantitative Analysis of Large Amounts of Journalistic Texts Using Topic Modelling. Digital Journalism, 4, 89–106.

Jerit J., Barabas J., Bolsen T. (2006). Citizens, Knowledge, and the Information Environment. American Journal of Political Science, 50, 266–282.

Jordan M. I., Mitchell T. M. (2015). Machine Learning: Trends, Perspectives, and Prospects. Science, 349, 255–260.

Kahneman, D. (2011). Thinking, Fast and Slow. New York: Farrar, Straus and Giroux.

Kahneman D., Tversky A. (1979). Prospect Theory: An Analysis of Decision under Risk. Econometrica, 47, 263–91.

Lancee B., Pardos-Prado S. (2013). Group Conflict Theory in a Longitudinal Perspective: Analyzing the Dynamic Side of Ethnic Competition. International Migration Review, 47, 106–131.

Lecheler S., de Vreese C. H. (2016). How Long Do News Framing Effects Last? A Systematic Review of Longitudinal Studies. Annals of the International Communication Association, 40, 3–30.

Liu B. (2012). Sentiment Analysis and Opinion Mining. Synthesis Lectures on Human Language Technologies, 5, 1–167.

Luhmann, N. (2002). The Cognitive Program of Constructivism and the Reality That Remains Unknown. In Luhmann N., Rasch W. (Eds.), Theories of Distinction. Redescribing the Descriptions of Modernity. Stanford, Calif: Stanford University Press, pp. 128–152.

Matthes J. (2007). Beyond Accessibility? Toward an on-Line and Memory-Based Model of Framing Effects. Communications, 32.

McCombs, M. E. (2004). Setting the Agenda: The Mass Media and Public Opinion. Cambridge, UK : Malden, MA: Polity ; Blackwell Pub.

McCombs M. E. (2005). A Look at Agenda-Setting: Past, Present and Future. Journalism Studies, 6, 543–557.

McCombs M. E., Shaw D. L. (1972). The Agenda-Setting Function of Mass Media. Public Opinion Quarterly, 36, 176.

McFarland D. A., Ramage D., Chuang J., Heer J., Manning C. D., Jurafsky D. (2013). Differentiating Language Usage through Topic Models. Poetics, 41, 607–625.

Meltzer C. E., Eberl J.-M., Theorin N., Heidenreich T., Strömbäck J., Boomgaarden H. G., Schemer C. (2021). Media Effects on Policy Preferences toward Free Movement: Evidence from Five EU Member States. Journal of Ethnic and Migration Studies, 47, 3390–3408.

Merten, K. (1986). Das Bild der Ausländer in der deutschen Presse. Frankfurt a.M.: Dağyeli.

Mimno D., Blei D. M. (2011). Bayesian Checking for Topic Models. Proceedings of the Conference on Empirical Methods in Natural Language Processing. Association for Computational Linguistics, 227–237.

Morgan, S. L., Winship C. (2014). Counterfactuals and Causal Inference: Methods and Principles for Social Research. Cambridge: Cambridge University Press.

Namin, P. J. (2009). Die Darstellung Des Islam in Den Deutschen Printmedien Am Beispiel von Spiegel Und Bild. In Geißler R., Pöttker H. (Eds.), Massenmedien Und Die Integration Ethnischer Minderheiten in Deutschland. Forschungsbefunde. Bielefeld: Transcript Verl., pp. 271–298.

Pardos-Prado S. (2011). Framing Attitudes Towards Immigrants in Europe: When Competition Does Not Matter. Journal of Ethnic and Migration Studies, 37, 999–1015.

Payton M. E., Greenstone M. H., Schenker N. (2003). Overlapping Confidence Intervals or Standard Error Intervals: What Do They Mean in Terms of Statistical Significance? Journal of Insect Science, 3.

Pew Research Center (2018). In Western Europe, Public Attitudes Toward News Media More Divided by Populist Views Than Left-Right Ideology. Washington, D.C. url: https://www.pewresearch.org/journalism/2018/05/14/in-western-europe-public-attitudes-toward-news-media-more-divided-by-populist-views-than-left-right-ideology/ (visited on 10/14/2021).

Remus R., Quasthoff U., Heyer G. (2010). SentiWS - a Publicly Available German-Language Resource for Sentiment Analysis. Proceedings of the 7th International Language Ressources and Evaluation. European Language Resources Association (Ed.), 1168–1171.

Roberts M. E., Stewart B. M., Airoldi E. M. (2016). A Model of Text for Experimentation in the Social Sciences. Journal of the American Statistical Association, 111, 988–1003.

Roberts M. E., Stewart B. M., Tingley D., Lucas C., Leder-Luis J., Gadarian S. K., Albertson B., Rand D. G. (2014). Structural Topic Models for Open-Ended Survey Responses. American Journal of Political Science, 58, 1064–1082.

Ruhrmann, G. (2002). Wie Regionale Tageszeitungen Über Migranten Berichten. In Meier-Braun K.-H., Kilgus A. M. (Eds.), Integration Durch Politik Und Medien.7. Medienforum Migranten Bei Uns. Baden-Baden: Nomos.

Ruths D., Pfeffer J. (2014). Social Media for Large Studies of Behavior. Science, 346, 1063–1064.

Schemer C. (2012). The Influence of News Media on Stereotypic Attitudes Toward Immigrants in a Political Campaign: Influence of News Media on Stereotypes. Journal of Communication, 62, 739–757.

Schemer C. (2014). Media Effects on Racial Attitudes: Evidence from a Three-Wave Panel Survey in a Political Campaign. International Journal of Public Opinion Research, 26, 531–542.

Scheufele D. A. (2000). Agenda-Setting, Priming, and Framing Revisited: Another Look at Cognitive Effects of Political Communication. Mass Communication and Society, 3, 297–316.

Scheufele, D. A., Iyengar S. (2014). The State of Framing Research: A Call for New Directions. In Kenski K., Jamieson K. H. (Eds.), The Oxford Handbook of Political Communication. Vol. 1. Oxford University Press.

Schlueter E., Davidov E. (2013). Contextual Sources of Perceived Group Threat: Negative Immigration-Related News Reports, Immigrant Group Size and Their Interaction, Spain 1996–2007. European Sociological Review, 29, 179–191.

Skrondal, A., Rabe-Hesketh S. (2004). Generalized Latent Variable Modeling: Multilevel, Longitudinal, and Structural Equation Models. Boca Raton: Chapman & Hall/CRC.

Snijders, T. A. B., Bosker R. J. (2012). Multilevel Analysis: An Introduction to Basic and Advanced Multilevel Modeling. Los Angeles: Sage.

Tang J., Meng Z., Nguyen X., Mei Q., Zhang M. (2014). Understanding the Limiting Factors of Topic Modeling via Posterior Contraction Analysis. ICML'14 Proceedings of the 31st International Conference on International Conference on Machine Learning. Beijing, 190–198.

Terman R. (2017). Islamophobia and Media Portrayals of Muslim Women. International Studies Quarterly, 61, 489–502.

Tversky A., Kahneman D. (1981). The Framing of Decisions and the Psychology of Choice. Science, 211, 453–458.

Tversky A., Kahneman D. (1973). Availability: A Heuristic for Judging Frequency and Probability. Cognitive Psychology, 5, 207–232.

Van Klingeren M., Boomgaarden H. G., Vliegenthart R., de Vreese C. H. (2015). Real World Is Not Enough: The Media as an Additional Source of Negative Attitudes To-

ward Immigration, Comparing Denmark and the Netherlands. European Sociological Review, 31, 268–283.

Wagner G. G., Frick J. R., Schupp J. (2007). The German Socio-Economic Panel Study (SOEP) - Evolution, Scope and Enhancements. Schmollers Jahrbuch, 127, 139–169.

Walgrave S., De Swert K. (2004). The Making of the (Issues of the) Vlaams Blok. Political Communication, 21, 479–500.

Weber H. (2019). Attitudes Towards Minorities in Times of High Immigration: A Panel Study among Young Adults in Germany. European Sociological Review, 35, 239–257.

Windzio M. (2018). The Network of Global Migration 1990–2013. Social Networks, 53, 20–29.

Wlezien C. (2005). On the Salience of Political Issues: The Problem with 'Most Important Problem'. Electoral Studies, 24, 555–579.

Wooldridge, J. M. (2002). Econometric Analysis of Cross Section and Panel Data. Cambridge, Massachusetts: MIT Press.

Appendices

A Dataset overview

The following search query was used:

```
integration AND (zuwanderung OR einwanderung* OR migration
    OR migrant* OR ausländer* OR asyl* OR flücht* OR geflü
    chtete*)
```

We tried different variations of this search query. If we chose the word 'migration' instead of 'integration' in the first part of the query, we get around 33% less hits in the database. We also examined some randomly selected articles before we conducted the search. However, the term 'migration' is also related to issues that we want to exclude, such as 'data migration' in information technology or 'bird migration' in biology, which resulted in a lot of false positives. In our view, there are different aspects related to concerns about immigration. Not all of them are related to attitudes towards immigrants (cf. also discussion, Section 5). Media reports on migration between e.g. Ethiopia and Eritrea might not have any effect on attitudes of German residents. German residents interviewed in the GSOEP are particularly affected by immigration due to processes of integration e.g. in schools, workplaces, neighborhoods or in families due to intermarriage. By including the term 'integration' into the search string, we include a component of immigration which is related to respondents who live in Germany and who consider themselves as being affected by immigration – be it in a positive or negative way. After the iterative process of querying the database and manually checking the results, we came to the conclusion

that the presented search query was a good compromise of enough articles and not too many false positives. In the following table, the resulting data set is described:

Newspaper	freq	valid	duplicate	short	onlydates	eng	sum
Spiegel	w	646	57	16	1	0	720
Sueddeutsche	d	7001	258	398	194	0	7851
FrankfurterRundschau	d	6129	534	286	225	0	7174
WELT	d	4766	60	102	10	2	4940
TAZ	d	4968	76	364	26	0	5434
Focus	w	394	5	28	0	0	427
stern	w	195	0	7	3	0	205
sum		24099	990	1201	459	2	26751

Table 1: Number of valid/invalid articles per newspaper.

B GSOEP Variables

vars	n	mean	sd	median	trimmed	mad	min	max	range	skew	kurtosis	se
AV.bin	333934	0.279	0.449	0	0.224	0.000	0	1	1	0.983	-1.033	0.001
party	333934	2.410	2.170	1	2.012	0.000	1	7	6	1.352	0.206	0.004
political.Interest	333934	2.666	0.822	3	2.688	1.483	1	4	3	-0.228	-0.443	0.001
income.Satisfact	333934	2.202	1.278	2	2.042	1.483	1	5	4	0.697	-0.661	0.002
worried.About.Econ	333934	0.306	0.461	0	0.258	0.000	0	1	1	0.842	-1.292	0.001
n.Articles	333934	118.214	96.727	85	95.375	26.687	5	579	574	2.926	8.747	0.167
pmonin	333934	4.932	2.315	4	4.930	2.965	1	10	9	0.017	-1.015	0.004

Table 2: Descriptive statistics for GSOEP variables used in the fixed effects regression.

C   Robustness of regression results
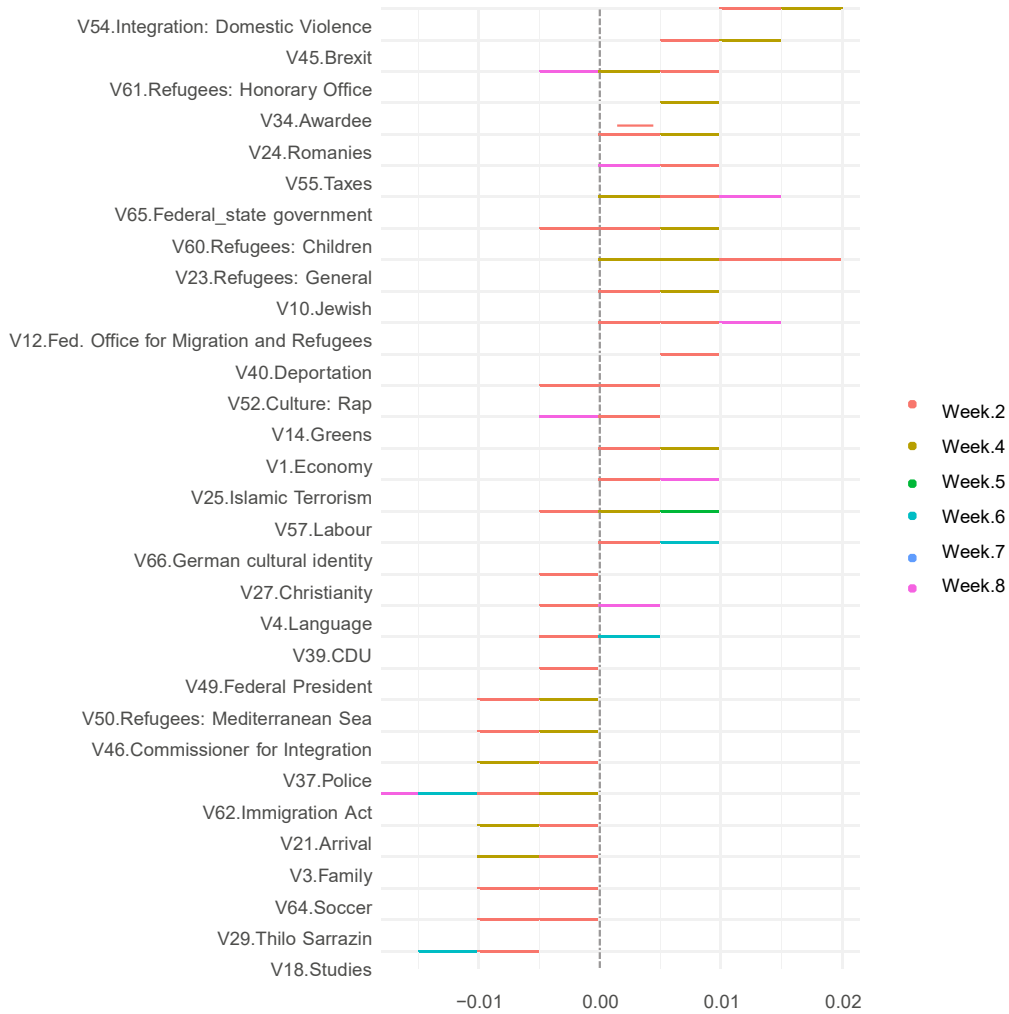
Figure 6: Results of topic effects for different choices of weeks, i.e., the period of time considered before someone filled out the survey. Only effects that are significant in the main model are shown here.

To show that the topic selection does not vary too much across the media outlets, Figure 7 shows the distribution of topic loads across media outlets for the three most positively/negatively associated topics. Other topics show a very similar picture.
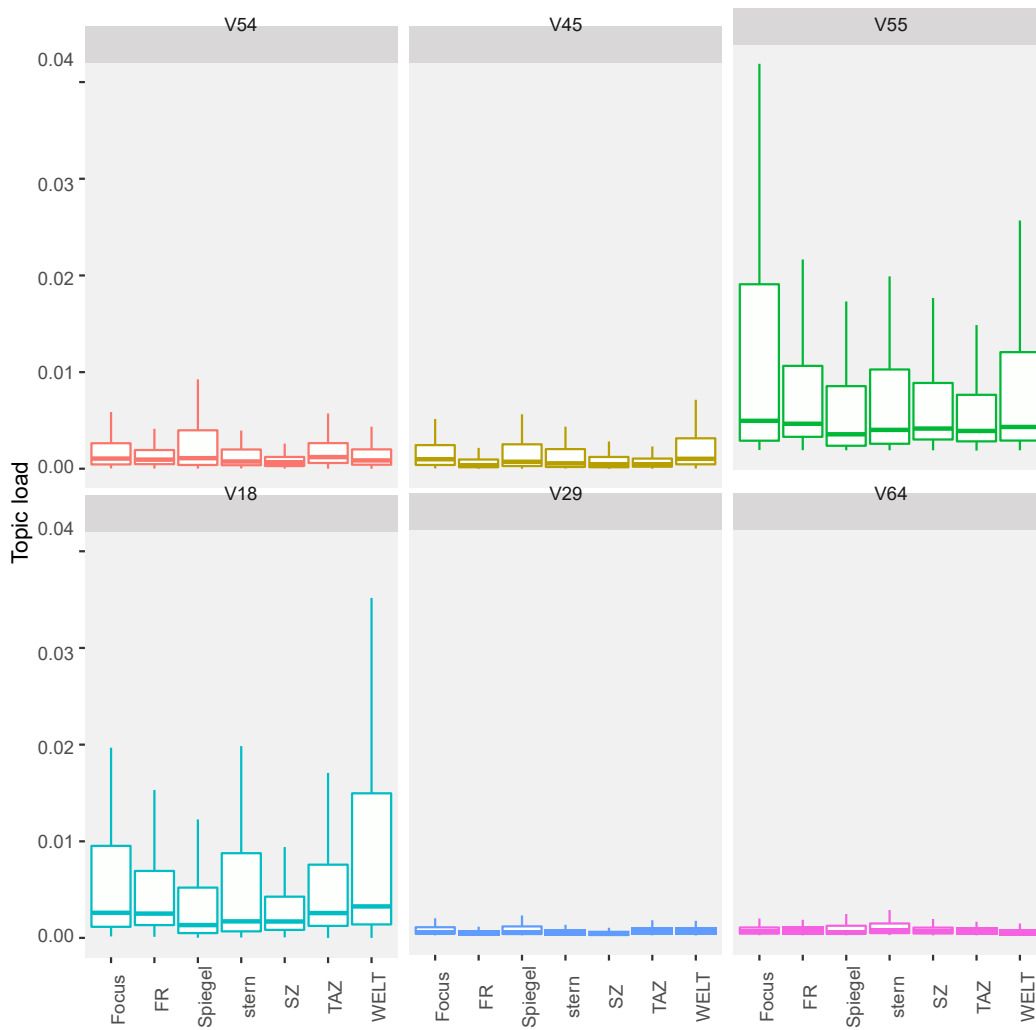
Figure 7: Topic distribution over media outlets.

54

D Regression results with reversed AV

	Dependent variable:	
	Not at all concerned	
	(1)	(2)
partyCDU/CSU	–0.024***	–0.023***
	(0.003)	(0.003)
partyFDP	–0.006	–0.010
	(0.007)	(0.007)
partyGruene	0.028***	0.021***
	(0.006)	(0.006)
partyLinke	0.001	–0.003
	(0.006)	(0.006)
partyRight-Wing	–0.025***	–0.014*
	(0.006)	(0.006)
partySPD	–0.008*	–0.004
	(0.003)	(0.003)
political.Interest2	0.001	0.0004
	(0.003)	(0.003)
political.Interest3	0.005	0.004
	(0.004)	(0.004)
political.Interest4	0.029***	0.028***
	(0.005)	(0.005)
income.Satisfact7	0.006**	0.006**
	(0.002)	(0.002)
income.Satisfact8	0.006**	0.006**
	(0.002)	(0.002)
income.Satisfact9	0.015***	0.015***
	(0.003)	(0.003)
income.Satisfact10	0.015***	0.017***
	(0.004)	(0.004)
worried.About.Econ	–0.052***	–0.044***
	(0.002)	(0.002)
n.Articles	–0.001***	–0.0004***
	(0.00001)	(0.00002)
pmoninApril	0.027***	–0.004

55

	(0.003)	(0.005)
pmoninAugust	0.006	−0.024***

56

	Dependent variable:	
	Not at all concerned	
	(1)	(2)
	(0.006)	(0.007)
pmoninFebruary	0.032***	0.018***
	(0.003)	(0.004)
pmoninJuly	0.030***	0.001
	(0.005)	(0.006)
pmoninJune	0.032***	0.018***
	(0.004)	(0.005)
pmoninMarch	0.028***	0.010*
	(0.003)	(0.004)
pmoninMay	0.021***	0.004
	(0.004)	(0.005)
pmoninOct/Nov/Dec	0.023**	−0.005
	(0.008)	(0.009)
pmoninSeptember	0.016*	−0.013
	(0.007)	(0.007)
V1		−0.005***
		(0.001)
V3		0.006***
		(0.001)
V4		0.004***
		(0.001)
V5		−0.004***
		(0.001)
V8		−0.0005
		(0.001)
V9		−0.0003
		(0.001)
V10		−0.007***
		(0.001)
V12		0.0003
		(0.002)
V14		−0.004***
		(0.001)
V16		0.001

57

(0.001)

58

	Not at all concerned	
	(1)	(2)
V18		0.008***
		(0.001)
V21		0.005***
		(0.001)
V23		0.005*
		(0.002)
V24		−0.002
		(0.001)
V25		−0.004***
		(0.001)
V26		−0.001
		(0.001)
V27		0.003**
		(0.001)
V28		−0.001
		(0.001)
V29		0.008***
		(0.001)
V31		0.004**
		(0.001)
V32		0.003**
		(0.001)
V33		−0.001
		(0.001)
V34		−0.005***
		(0.001)
V35		−0.002*
		(0.001)
V37		0.006***
		(0.001)
V39		0.003**
		(0.001)
V40		−0.007***

59

V42

$$\begin{matrix} (0.001) \\ -0.001 \end{matrix}$$

60

| | Dependent variable: | |
| | Not at all concerned | |
	(1)	(2)
		(0.001)
V45		−0.011***
		(0.001)
V46		0.007***
		(0.001)
V47		−0.001
		(0.001)
V49		0.009***
		(0.001)
V50		0.001
		(0.001)
V52		−0.003***
		(0.001)
V53		0.003***
		(0.001)
V54		−0.014***
		(0.001)
V55		−0.008***
		(0.001)
V56		0.001
		(0.001)
V57		0.001
		(0.001)
V60		−0.007***
		(0.002)
V61		−0.009***
		(0.002)
V62		−0.003*
		(0.002)
V64		0.007***
		(0.001)
V65		−0.011***
		(0.001)
V66		−0.004***

61

(0.001)

60

	Dependent variable:	
	Not at all concerned	
	(1)	(2)
V67		0.0004
		(0.001)
V68		0.001
		(0.001)
Observations	333,934	333,934
R^2	0.523	0.530
Adjusted R^2	0.425	0.433
Residual Std. Error	0.342 (df = 276856)	0.339 (df = 276809)

Note: $^*p{<}0.05$; $^{**}p{<}0.01$; $^{***}p{<}0.001$

E   Details of the fixed effects model

In a fixed-effects panel regression researchers are usually interested in a "within" effect: a change in the explanatory variable x within the respective subject has an impact on a change in the dependent variable y within the same subject. From a causal inference perspective, this within estimator is often regarded as a proxy for a counterfactual because all states of interest (Y and D) are observed in the panel. The fixed effects approach excludes all time-constant variation in x, which brings the estimator close to the average treatment effect on the treated (ATT). In other words, what is observed in the data is the expected difference σ in the outcome of the treated subjects Y^1 compared with the counterfactual outcome Y^0 if these subjects had not received a treatment D (Morgan and Winship, 2014, p. 55).

$$E[\sigma|D=1] = E[Y^1 - Y^0|D=1] = E[Y^1|D=1] - [Y^0|D=1] \qquad (1)$$

59

$$L_c(\beta, y, X) = \prod_{i=1}^{Y} P_i \left(\frac{y_{i1}, \dots, y_{iT}}{\sum_{i=1}^{T} y_{it}} \right) \tag{2}$$

In our research design there are different ways of predicting fixed effects estimators. The survey item in the GSOEP on concerns about immigration is an ordinal scale with three categories (1. N=93297; 2. N=146022; 3. N=94615). Why not using all three categories in an ordinal logistic hybrid model, which yields the appropriate within estimators for the time varying covariates (Allison, 2009, p. 49)? The hybrid approach gives approximately the same within-estimates as the fixed effects model, so the procedure is the same for hybrid ordinal logistic regression (Allison, 2009). However, a recent overview highlighted the problem of fixed effects logit models (Beck, 2020): the estimator drops all information where the dependent variable y is constant within subjects. In the hybrid (binary and ordinal) logit approach, the contribution of subjects with time-invariant y on the within-estimator is ignored. The linear probability FE model, in contrast, includes time invariant-information on y and thereby shrinks the coefficients toward zero (Beck, 2020). In other words, this shrinkage toward the zero is removed from the FE logit model, which thus implies less conservative estimates. Beck (2020) leaves it to the researchers to decide between these two models. In our view, the shrinkage toward zero provides important counter-evidence against the researchers' hypotheses. It would be 'bad news' for our hypotheses if x changed but y usually remained constant within a subject. Since we are interested in highlighting particular effects and postulate that these effects exist, opting for the FE logit model would be a decision in favor of our hypotheses – simply by removing this 'bad news'. We thus prefer to include the shrinkage-toward-zero

60

implied in the FE linear probability model and opt in favor of a more conservative, strict and rigorous test of our hypotheses.

The prevalence of topics is standardized in the fixed effects models. Hence, if the coefficient of a topic would be one, the likelihood of a person becoming very concerned about immigration would increase by one standard deviation of the corresponding topic (cf. topics' deviations Table ??).
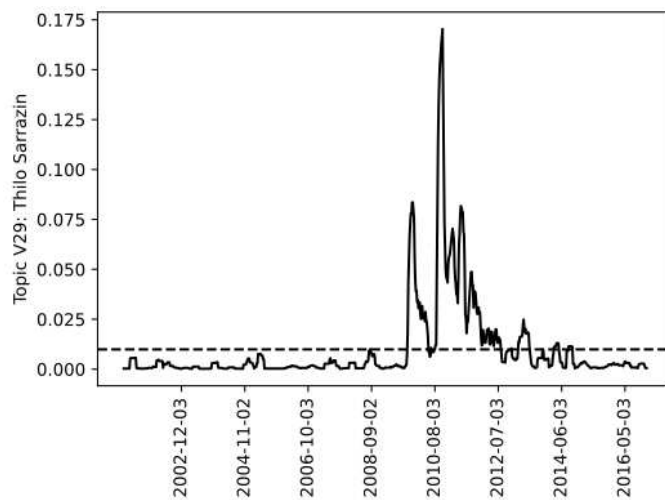
Figure 8: Topic prevalence of Topic29 over time (in weeks). The dashed line indicates the overall mean prevalence.

We will present a concrete example to illustrate how we estimate the influence of topics. Figure 8 shows the topic prevalence of topic 29. Its thematic content revolves around Thilo Sarrazin, a former social democratic politician. He published a very controversial book about Muslim immigrants in Germany in 2010, which sparked a discussion about multiculturalism across all political camps. The coefficient of Topic29 in model 2 (Appendix G) denotes -0.007 ($p < .001$). Given a rise in prominence (compared to its average prevalence, because its normalized), the negative coefficient indicates that individuals, who were very concerned about immigration beforehand, tend to reduce their concern to "somewhat concerned" or "not at all con-

cerned". To facilitate the interpretation of the coefficients, the following table shows

all means and standard deviations for the respective topic variables.

Topic	Mean	Std
V1	0.0109	0.0692
V3	0.0326	0.0778
V4	0.0176	0.0756
V5	0.0092	0.0669
V6	0.0094	0.0606
V8	0.0144	0.0731
V9	0.0329	0.1016
V10	0.0079	0.0643
V12	0.0194	0.0757
V14	0.0065	0.0523
V16	0.0165	0.0836
V18	0.0268	0.0866
V21	0.0083	0.0668
V23	0.0228	0.0737
V24	0.0070	0.0591
V25	0.0054	0.0479
V26	0.0189	0.0713
V27	0.0073	0.0546
V28	0.0174	0.0813
V29	0.0116	0.0779
V31	0.0159	0.0717
V32	0.0190	0.0806
V33	0.0113	0.0735
V34	0.0127	0.0754
V35	0.0184	0.0706
V37	0.0145	0.0681
V39	0.0130	0.0745
V40	0.0158	0.0809
V42	0.0166	0.0736
V45	0.0183	0.0899
V46	0.0225	0.0822
V47	0.0246	0.0870
V49	0.0085	0.0630
V50	0.0131	0.0663
V52	0.0140	0.0784

62

Topic	Mean	Std
V53	0.0017	0.0325
V54	0.0134	0.0664
V55	0.0212	0.0644
V56	0.0114	0.0751
V57	0.0249	0.0935
V60	0.0154	0.0781
V61	0.0203	0.0886
V62	0.0283	0.1223
V64	0.0136	0.0855
V65	0.0145	0.0643
V66	0.0242	0.0893
V67	0.0118	0.0725
V68	0.0164	0.0833

F   Topics of STM model

This table shows the topic number, our given label, and the top four words for each topic by Prob and FREX.

Nr	Label	Prob	Frex
1	Economy	german organisation germany economy	deutsch-bank stockholders-meeting go ernment-bond company-share
2	Region: Bavaria	csu bavaria bavarian seehofer	söder joachim-herrmann home-secretary-joachim minister-presiden-horst
3	Family	life family woman tell	mother daughter father drink
4	Language	german language learn course	adult-education-center courses dzif participants
5	Media	media german integration radio	radio-multicultural wdr station shows
6	Region: Berlin	berlin john dresden npd	stadtkewitz barbara-john tillich npd
7	Region: Switzerland	political switzerland germany politics	switzerland swiss svp blocher

Continued on next page

63

Nr	Label	Prob	Frex
8	Islam	muslims muslim islamic islam	görüs milli-görüs milli ditib
9	Pisa	children school parents pupils	pisa type-of-school pisa-study school-system
10	Jewish	jewish jews israel germany	jewish-communities israel's scharon israel
11	Region: Frankfurt	frankfurt city integration church	amka feldmann multicultural-matter eskandarigrünberg
12	Fed. Office for Migration and Refugees	refugees asylum-seeker employment employment-market	bamf employment-agency jobcentre federal-agency
13	Region: Hanau	city integration offenbach hanau	weissthiel hanau großgerau intercultural-weeks
14	Greens	greens fischer özdemir german	özdemir joschka westerwelle cem-özdemir
15	Region: NRW	laschet spd nrw armin	laschet nrw armin-laschet öney
16	Integration: Intercultural	migrants integration intercultural german	germering leupold plot gardens
17	Region: Serbia	kosovo serbian serbia government	serbs serbian mladic belgrade
18	Studies	percent studies germany migrants	respondents study percent-respondents share-percent
19	Region: Netherlands	dutch van netherlands wilders	theo-van theo-van-gogh pim-fortuyn balkenende
20	Region: Berlin_SPD	berlin senate spd wowereit	dilek-kolat senator-for-integration günter-piening piening
21	Arrival	germany life german homeland	vietnamese elvis wolfsburg eke
22	Region: Munich	county asylum-seeker refugees municipality	karmasin göbel christoph-göbel loderer
23	Refugees: General	refugees germany need country	refugee-crisis obergrenze(upper-limit) influx-of-refugees welcoming-culture

64

| 24 | Romanies | romanies romania bulgaria live | romanies bulgarians sinti sinti-romanies | romanians- |

65

Nr	Label	Prob	Frex
25	Islamic Terrorism	germany young live german	kurnaz morsal jihad alqaida
26	Party donations	green spd cdu party	kretschmann blackgreen greens winfried-kretschmann
27	Christianity	church life world catholic	pope luther roman chr
28	Nationalization	german germany abroad citizenship	naturalization naturalizations nationality german-citizenship
29	Thilo Sarrazin	sarrazin thilo thilo-sarrazin germany	sarrazin's sarrazin german-bundesbank-manager thilo-sarrazin
30	Region: Hesse	hesse hessian koch wiesbaden	jörguwe jörguwe-hahn bebenburg pitt-bebenburg
31	Merkel	merkel spd gabriel schäuble	chancellor-angela-merkel chancellor-angela merkel-cdu angela-merkel-cdu
32	Rent	city housing refugees reside	tenant skotnik vacancy housing-association
33	Integration: Work	integration germany german society	fue must can ueb
34	Awardee	integration prize euro munich	ismair kjr tutzing honor
35	Erdogan	turkish turkey german turks	erdogan turks-germany tayyip-erdogan turkish-minster-president
36	Region: USA	usa russia russian world	bush georgia putin moscow
37	Police	police perpetrator young victim	perpetrator sanel investigations police
38	Region: Dachau	county dachau asylum-seeker ebersberg	löwl kirchseeon ebersberg herbertshausen
39	CDU	cdu union merkel party	özkan aygül aygül-özkan giousouf
40	Deportation	deportation family germany right-of-residence	hardship-commission residence-right regulation-of-residence sürücü
41	Region: Dietzenbach	foreigners-advisory-council city offenbach integration	dietzenbach dietzenbachs giesler butterweck
42	Globalization	social political society societal	economy globalization research transnational

66

Nr	Label	Prob	Frex
43	Region: Homburg	bad refugees district city	maintaunuskreis vilbel bad-vilbel xyriax
44	Region: Berlin Neukölln	neukölln berlin buschkowsky kreuzberg	neukölln rütlischool buschkowsky neuköllns
45	Brexit	european europe union brussels	member-state brexit juncker single-market
46	Commissioner for Integration	integration migrants böhmer integration-policies	maria-böhmer bömer böhmer-cdu integration-summit
47	Associatve life	project association integration sports	sports-club sports-youth hertiefoundation sports-portfolio
48	Region: Europe	france sarkozy government sweden	danish sarkozy övp stockholm
49	Federal President	federal-president wulff gauck germany	gauck joachim-gauck federal-president rau
50	Refugees: Mediterranean Sea	refugees germany europe syria	mediterranean-sea syrian-refugee traffickers unchr
51	Region: Hamburg	hamburg bremen senate integration	goetsch ahlhaus schnieberjastram hanseatic-city
52	Culture: Rap	german germany language write	bushido german-russians roman rapper
53	Culture: Movies	movie usa germany tip	cinema bhv movie-service bhv-del
54	Integration: Domestic Violence	women men girls woman	kelek necla necla-kelek seyran
55	Taxes	euro million money million-euro	million-euro billion-euro billions revenue
56	BDV	german poland germany history	bdv flight-displacement displacement displaced-persons
57	Labour	germany job-market german company	skilled-worker lack-of-skilled-labour qualified highly-qualified
58	Region:	city city-council mayor	puchheim geretsried seidl

67

District Munich	geretsried	reiter

68

Nr	Label	Prob	Frex
59	Region: Africa	france french paris country	rwanda congo marseille militias
60	Refugees: Children	young refugees teenager refugee-child	unaccompanied-minor minor-refugee unaccompanied vocational-school
61	Refugees: Honorary Office	refugees asylum-seeker helper voluntary	helper helperss voluntary-helper volunteer-helpers
62	Immigration Act	union immgration immgration-act schily	schily home-secretary-otto otto-schily home-secretary-otto-schily
63	Region: Karlsfeld	integration munich dachau municipality	naz karlsfeld hasenbergl dachauost
64	Soccer	soccer play german player	türkiyemspor player dfb bundesliga
65	Federal_state government	municipality federal state refugees	federal-state maly association-of-german-cities states-municipalities
66	German cultural identity	society islam germany religion	leitkultur multiculturalism secular religion
67	Party leadership	merkel afd chancellor germany	wagenknecht gauland petry afd
68	Culture: Theatre	theatre art music culture	theatre fesitval ballhaus artistic
69	Region: Anzing	asylum-seeker anzing give live	anzing anzingian stranglmeier oellerer

69

G   Full regression results

The following table show the main results of this article. Model 1 only contains our covariates (no topic variables), model 2 presents our key findings and contains all non-interaction variables. Model 3 only contains all covariates, a subset of topic variables and their interaction effects with education. The overall increase in R^2 is not large but comparable to effects found in previous studies (Meltzer et al., 2021; Schemer, 2012; Schemer, 2014).

	Media topics and concerns about immigration		
Variable	Model1	Model2	Model3
partyCDU/CSU	0.021***	0.020***	0.019***
	(0.003)	(0.003)	(0.003)
partyFDP	0.002	0.008	0.007
	(0.007)	(0.007)	(0.007)
partyGruene	–0.014***	–0.011**	–0.012**
	(0.004)	(0.004)	(0.004)
partyLinke	–0.012	–0.009	–0.009
	(0.006)	(0.006)	(0.006)
partyRight-Wing	0.137***	0.126***	0.129***
	(0.010)	(0.010)	(0.010)
partySPD	–0.004	–0.006	–0.005
	(0.003)	(0.003)	(0.003)
political.Interest2	–0.019***	–0.019***	–0.019***
	(0.003)	(0.003)	(0.003)
political.Interest3	–0.026***	–0.026***	–0.025***
	(0.004)	(0.004)	(0.004)
political.Interest4	–0.034***	–0.035***	–0.033***
	(0.005)	(0.005)	(0.005)
income.Satisfact7	–0.008***	–0.009***	–0.008***
	(0.002)	(0.002)	(0.002)
income.Satisfact8	–0.007**	–0.008***	–0.007**
	(0.002)	(0.002)	(0.002)
income.Satisfact9	–0.014***	–0.016***	–0.014***
	(0.003)	(0.003)	(0.003)

Continued on next page

70

Media topics and concerns about immigration			
Variable	Model1	Model2	Model3
income.Satisfact10	–0.008*	–0.010*	–0.008*
	(0.004)	(0.004)	(0.004)
worried.About.Econ	0.092***	0.089***	0.090***
	(0.002)	(0.002)	(0.002)
n.Articles	0.001***	0.0005***	0.001***
	(0.00001)	(0.00002)	(0.00001)
pmoninApril	–0.005	0.006	0.006
	(0.004)	(0.005)	(0.004)
pmoninAugust	0.014**	0.025***	0.028***
	(0.005)	(0.006)	(0.006)
pmoninFebruary	–0.013***	–0.011*	–0.017***
	(0.003)	(0.004)	(0.003)
pmoninJuly	–0.014**	–0.002	–0.018***
	(0.005)	(0.006)	(0.005)
pmoninJune	–0.003	0.001	–0.005
	(0.004)	(0.005)	(0.005)
pmoninMarch	–0.008*	–0.006	–0.010**
	(0.003)	(0.004)	(0.003)
pmoninMay	–0.001	0.004	0.004
	(0.004)	(0.005)	(0.004)
pmoninOct/Nov/Dec	–0.020*	–0.007	–0.004
	(0.008)	(0.009)	(0.009)
pmoninSeptember	0.003	0.011	0.020**
	(0.007)	(0.007)	(0.007)
V1		0.004***	
		(0.001)	
V3		–0.005***	–0.005***
		(0.001)	(0.001)
V4		–0.004***	
		(0.001)	
V5		0.002*	
		(0.001)	
V8		–0.001	
		(0.001)	
V9		–0.002*	
		(0.001)	

71

Variable	Model1	Model2	Model3
		Media topics and concerns about immigration	
V10		0.006***	
		(0.001)	
V12		0.006***	
		(0.002)	
V14		0.004***	
		(0.001)	
V16		−0.0004	
		(0.001)	
V18		−0.008***	−0.011***
		(0.001)	(0.001)
V21		−0.005***	−0.006***
		(0.001)	(0.001)
V23		0.006**	
		(0.002)	
V54:educ.bin			−0.006***
			(0.002)
educ.bin:V45			−0.006***
			(0.002)
educ.bin:V61			−0.014***
			(0.002)
educ.bin:V34			0.001
			(0.001)
educ.bin:V24			0.001
			(0.001)
educ.bin:V18			0.002
			(0.001)
educ.bin:V29			−0.001
			(0.002)
educ.bin:V64			0.001
			(0.001)
educ.bin:V3			0.004**
			(0.002)
educ.bin:V21			0.002
			(0.001)
V24		0.007***	0.002*
		(0.001)	(0.001)

72

Media topics and concerns about immigration			
Variable	Model1	Model2	Model3
V25		0.004***	
		(0.001)	
V26		0.002	
		(0.001)	
V27		−0.003**	
		(0.001)	
V28		−0.0003	
		(0.001)	
V29		−0.007***	−0.006***
		(0.001)	(0.001)
V31		0.00002	
		(0.001)	
V32		−0.002	
		(0.001)	
V33		−0.00004	
		(0.001)	
V34		0.008***	0.007***
		(0.001)	(0.001)
V35		0.002	
		(0.001)	
V37		−0.005***	
		(0.001)	
V39		−0.004***	
		(0.001)	
V40		0.005***	
		(0.001)	
V42		0.001	
		(0.001)	
V45		0.011***	0.018***
		(0.001)	(0.001)
V46		−0.005***	
		(0.001)	
V47		0.0001	
		(0.001)	
V49		−0.004***	
		(0.001)	

73

		Media topics and concerns about immigration	
Variable	Model1	Model2	Model3
V50		$-0.005{***}$	
		(0.001)	
V52		$0.005{***}$	
		(0.001)	
V53		0.001	
		(0.001)	
V54		$0.018{***}$	$0.022{***}$
		(0.001)	(0.001)
V55		$0.007{***}$	
		(0.001)	
V56		$0.002*$	
		(0.001)	
V57		$0.003{**}$	
		(0.001)	
V60		$0.006{***}$	
		(0.002)	
educ.bin			$-0.021{***}$
			(0.006)
V61		$0.008{***}$	$0.021{***}$
		(0.002)	(0.001)
V62		$-0.005{***}$	
		(0.001)	
V64		$-0.007{***}$	$-0.004{***}$
		(0.001)	(0.001)
V65		$0.006{***}$	
		(0.001)	
V66		$0.003{**}$	
		(0.001)	
V67		-0.002	
		(0.001)	
V68		$-0.002*$	
		(0.001)	
Observations	333,934	333,934	333,934
R^2	0.502	0.508	0.507
Adjusted R^2	0.400	0.406	0.405

74

Media topics and concerns about immigration			
Variable	Model1	Model2	Model3
Residual Std. Error	0.348 (df = 276856)	0.346 (df = 276809)	0.346 (df = 276835)

H   Finding k

The idea of validating K is to identify a model with topics that best reflect weighted bags of words which are used by newspaper articles. For that purpose, semantic coherence and exclusivity are widely used measures (Mimno and Blei, 2011; Roberts, Stewart and Airoldi, 2016).

The coherence of a semantic space addresses whether a topic is internally consistent by calculating the frequency with which high probability topic words tend to co-occur in documents. Coherence grows with the likelihood of a topic's most probable words co-occurring together. The authors who introduced this measure validated it for academic writing (Mimno and Blei, 2011). They showed that it had high correspondence with the judgments of NIH officials on a set of NIH grants.
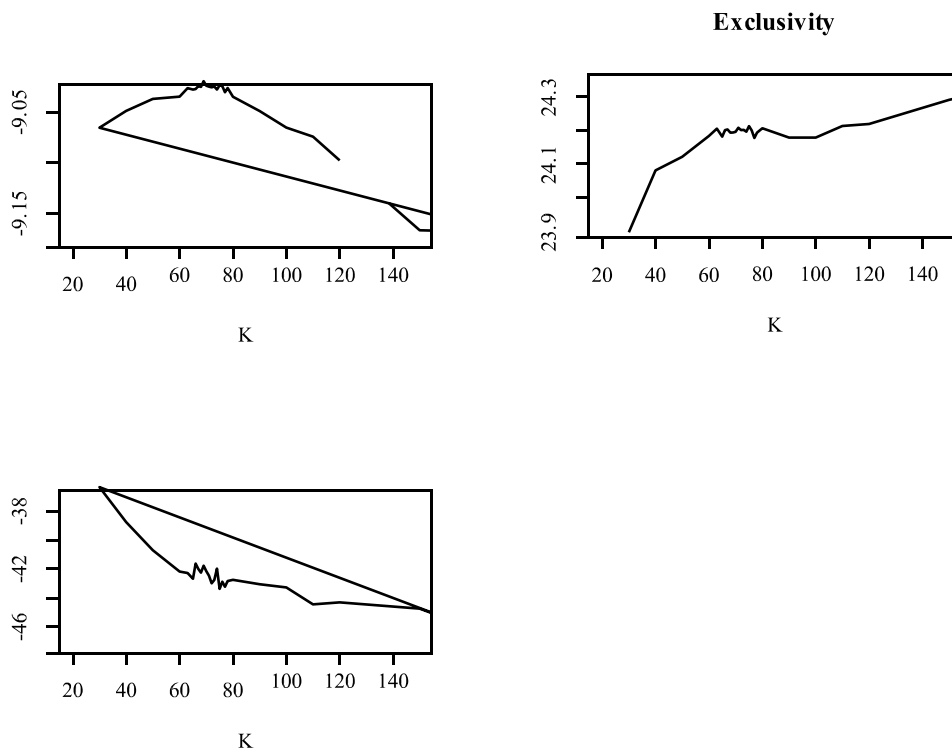
However, semantic coherence alone can be misleading since high values can simply be obtained by very common words of a topic that occur together in most documents. To account for the desired statistical discrimination between topics we may consider a second metric proposed by Roberts et al. (2014) and measure the exclusivity of a topic. Exclusivity provides us with the extent to which the words of a topic are distinct to it.

The developers of STM recommend that researchers look for the "semantic coherence-exclusivity frontier" — namely the specification after which allowing for more topics fails to produce models that dominate others in terms of semantic coherence and exclusivity (Roberts et al., 2014, p. 1070). Thus, what we are looking for is when both indicators build a plateau.

In addition, we use held-out likelihood. This is "a measure of predictive power to evaluate comparative performance", in this case among models that allow for different
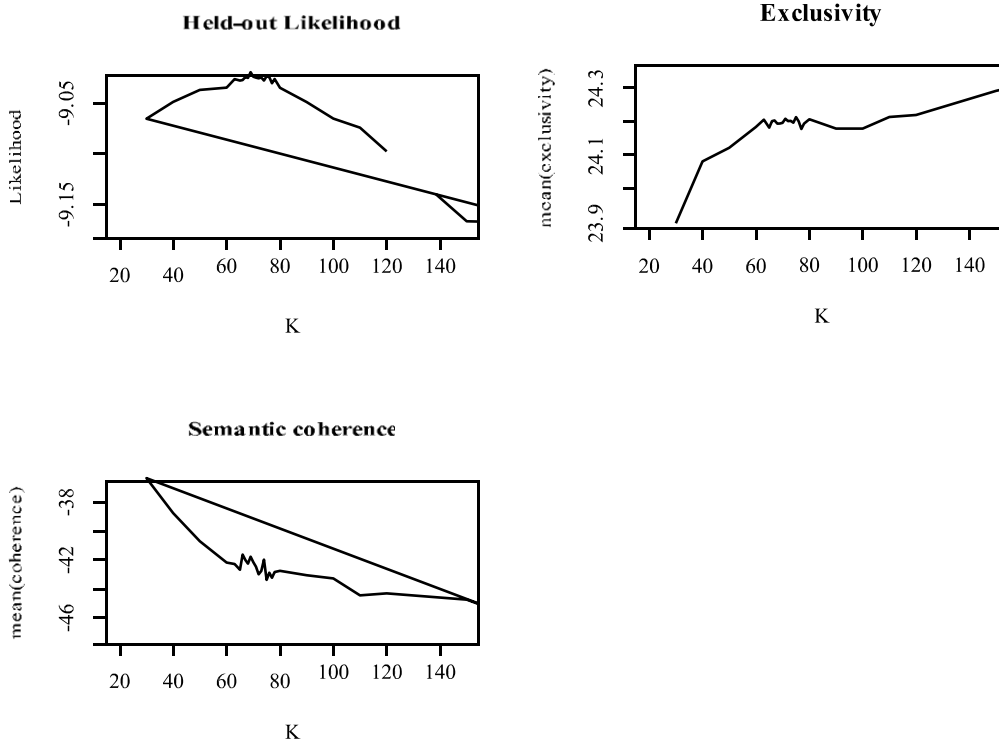
76

numbers of topics (Roberts, Stewart and Airoldi, 2016, p. 22). To obtain the held-out likelihood of an STM we first subset 10% of the documents in the corpus and hold out half of the words in them. We then evaluate the likelihood of the held-out words. Higher likelihoods indicate a more predictive model.

We choose a topic model with 69 topics. Having run a number of models with different parameters, $K = 69$ appears to possess two desired properties derived from the metrics explained above: (a) largest heldout-likelihood with -9.03; (b) $K = 69$ lies right on top of the plateaus of exclusivity and semantic coherence, indicating a good mixture between both measures.

Exclusivity

numbers of topics (Roberts, Stewart and Airoldi, 2016, p. 22). To obtain the held-out likelihood of an STM we first subset 10% of the documents in the corpus and hold out half of the words in them. We then evaluate the likelihood of the held-out words. Higher likelihoods indicate a more predictive model.

We choose a topic model with 69 topics. Having run a number of models with different parameters, $K = 69$ appears to possess two desired properties derived from the metrics explained above: (a) largest heldout-likelihood with -9.03; (b) $K = 69$ lies right on top of the plateaus of exclusivity and semantic coherence, indicating a good mixture between both measures.

77

I Sentiment analysis

Sentiment analysis is often used to identify the "tone" of a newspaper article, or any kind of media output, quantitatively (Liu, 2012). In our study, the presence of sentiments might confound the association with topics, our primary research interest.

We utilize the established "SentiWS" dictionary which contains around 1,650 positive and 1,800 negative words (Remus, Quasthoff and Heyer, 2010). To apply sentiments on our corpus and find associations with topics, we have to assign representative documents to topics. For that purpose, we can simply use θ (topic load) and find the, by design, most associated documents of a topic. Because every text contains positive and negative words, a "sentiment score" $\mathbf{P}_{positive} - \mathbf{P}_{negative}$ is calculated.

Figures 9 and 10 report sentiments in the 500 most associated documents for each of the significant positive and negative topics reported in Figure 4. The dotted line gives the "average sentiment score" across all documents. We see that neither the positive nor the negative topics are significantly deviating from that mean, or, for that matter, any topic at all. That provides evidence for the general "neutral" tone used in German quality newspapers. We also tried various numbers of documents besides 500, results do not change though (available upon request).
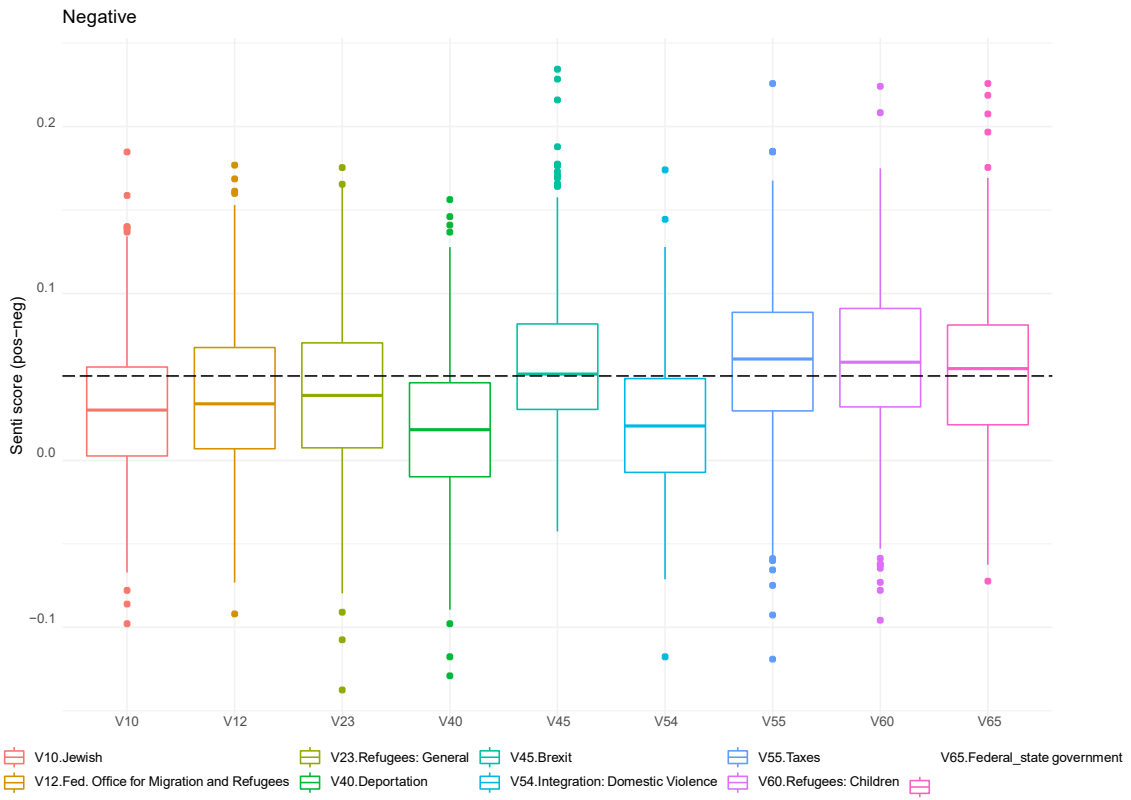
78

Figure 9: Sentiment score of topics having a significant, negative association with concerns on migration.
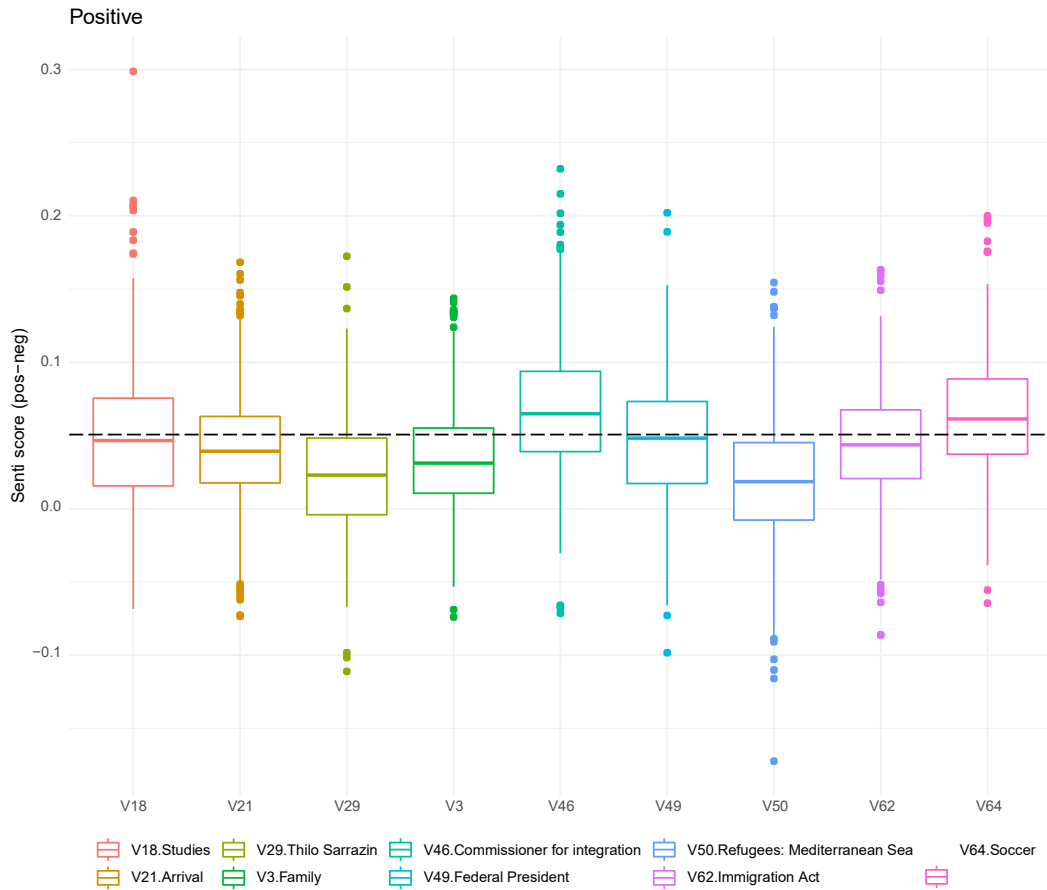
Figure 10: Sentiment score of topics having a significant, positive association with concerns on migration.

J Multicollinearity in Topics

One concern regarding our model is potential multicollinearity between topics. To account for this issue, we computed the model with only a subset of the topics. These 23 topics are the highly significant topics from model 2 (that is, the effects depicted in Figure 4 of the main text). In addition, the number of articles n.Articles could be strongly associated with the topic load of prevalent topics. We therefore run the subset-model twice, with and without n.Articles. Figure 11 shows the diverging effects for the full model (model 2 in Appendix G) and the two subset models 4.
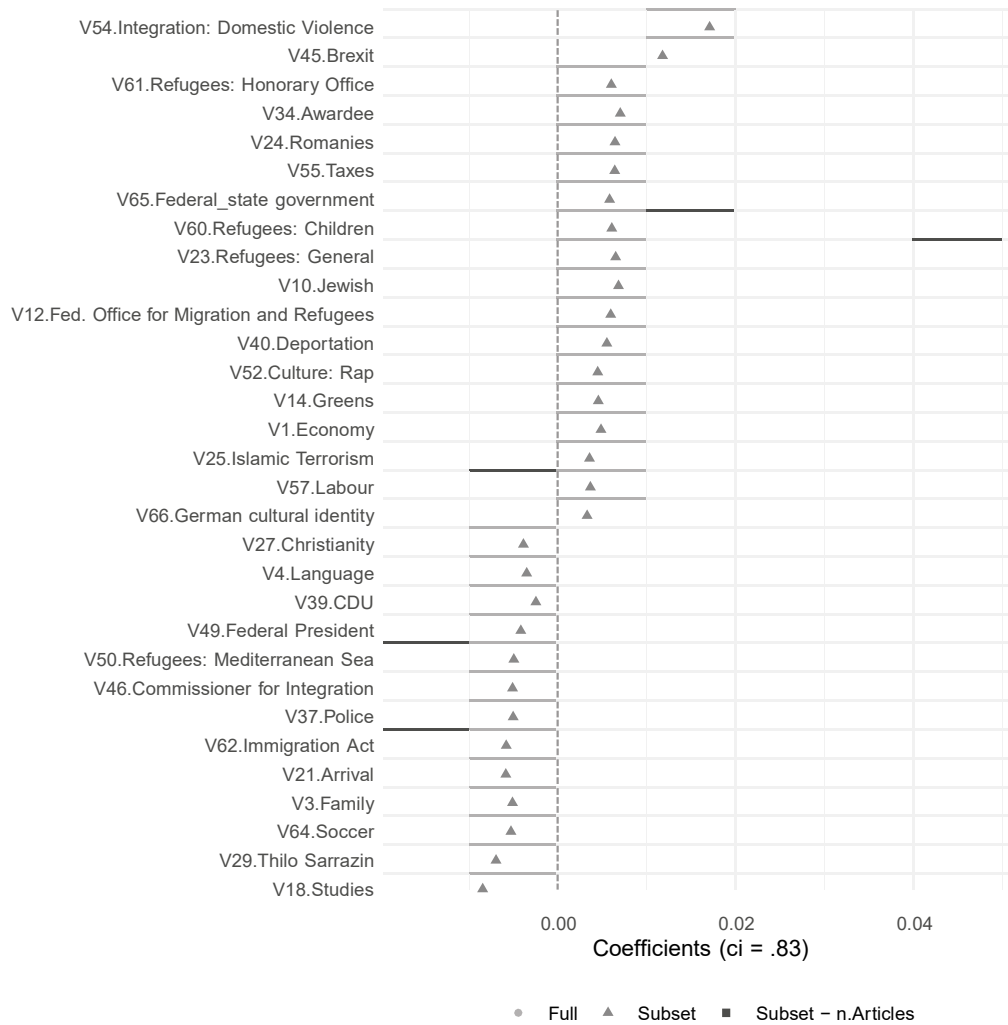
80

Figure 11: Diverging Effects for selected Topics

The effect of topic V23 "Refugees: General" shows the largest difference between the
full model specification and the model with a subset of topics (see Figure 11). This
seems reasonable, because during summer 2015 many refugees came to Germany, and
all media reported intensely on this topic. The topic dominated mass communication
in this period, but at the same time it was highly polarizing. On the one hand, the
people adhering to the "welcoming culture" mobilized support to refugees e.g. by
donating furniture or helping refugees in daily transactions. On the other hand,
right-wing populists strongly opposed the immigration of mostly Muslim refugees.

81

This polarization might be reflected in the effect of topic V23 "Refugees: General". Multicollinearity in the full model leads to a reduced effect, so that the effect is even stronger in the subset model. Accordingly, the estimation of the full model gives a conservative result.

Multicollinearity can be tested by the variance inflation factor (VIF), which is 1 divided by tolerance. Here, each explanatory variable x is regarded as an outcome y in a regression on all other explanatory variables. For each x, tolerance is 1-R2 of the respective regression. Since the constant term in the fixed effects regression is not the same as in OLS regression, however, VIF for fixed effects models is not defined in the same way as it is for OLS models. Fortunately, it is possible to approximate the within-estimator by integrating mean values of time-varying covariates and estimating the linear model by using OLS (Snijders and Bosker, 2012, pp. 26–31). The benefit of this approach is to get approximate within-effects, and at the same time being able to give interpretable values of this in order to test the multicollinearity in the model. Table 7 shows those approximated VIF values for the full model in descending order.

Variable	VIF	1/VIF
n.Articles	13.06	0.076578
V23	12.43	0.080448
political.Interest3	10.37	0.096394
pmoninMarch	9.69	0.103235
pmoninFebruary	9.20	0.108696
V61	9.15	0.109242
V12	7.85	0.127415
political.Interest4	7.12	0.140538
pmoninApril	7.11	0.140609
political.Interest2	6.75	0.148161
V60	6.29	0.159040
V62	6.03	0.165898

<div align="center">Continued on next page</div>

Variable	VIF	1/VIF
V31	5.68	0.176192
V33	5.47	0.182775
V67	5.45	0.183362
V65	5.07	0.197339
pmoninMay	4.98	0.200982
pmoninOct/Nov/Dec	4.27	0.234428
pmoninJune	4.04	0.247387
V50	3.91	0.255783
V3	3.78	0.264868
V54	3.75	0.266432
V9	3.67	0.272315
V32	3.65	0.273619
partyCDU/CSU	3.61	0.276954
pmoninJune	3.52	0.284277
partyGruene	3.31	0.301724
V46	3.31	0.302425
V56	3.25	0.307778
partySPD	3.16	0.316897
V24	3.13	0.319492
V34	3.12	0.320214
V42	3.12	0.320490
V47	3.09	0.323479
V29	3.04	0.328553
V14	3.01	0.332002
V16	2.97	0.336873
V28	2.97	0.337012
pmoninAugust	2.96	0.337948
V26	2.89	0.346502
V45	2.87	0.348811
V57	2.87	0.348996
V66	2.85	0.350599
partyLinke	2.82	0.354502
V35	2.82	0.354649
V55	2.78	0.359798
pmoninSeptember	2.72	0.367051
V39	2.65	0.377649
V68	2.63	0.379984
V10	2.56	0.390336

83

Variable	VIF	1/VIF
income.Satisfact10	2.53	0.395218
V53	2.46	0.405987
V27	2.46	0.407142
V64	2.45	0.407938
V18	2.35	0.425140
income.Satisfact8	2.34	0.426618
partyFDP	2.33	0.428491
V49	2.28	0.438679
income.Satisfact9	2.24	0.445498
V1	2.21	0.453364
V21	2.20	0.455274
V8	2.19	0.456207
V4	2.14	0.467824
V40	2.13	0.469196
V5	2.08	0.480028
V37	2.06	0.485393
V25	1.92	0.520934
V52	1.88	0.530639
partyRight-Wing	1.83	0.544996
worried.About.Econ	1.81	0.550995
income.Satisfact7	1.80	0.555764

Table 7: VIF calculated by context variable regression

To ensure the robustness of our results, we excluded all Variables that show a VIF
> 7, namely n.Articles, pmonin, political.Interest, and the topics V23, V61, V12. To
investigate the influence of n.Articles further, we calculated this 'Low VIF' Model
twice and added only the variable n.Articles in the second run. The results are
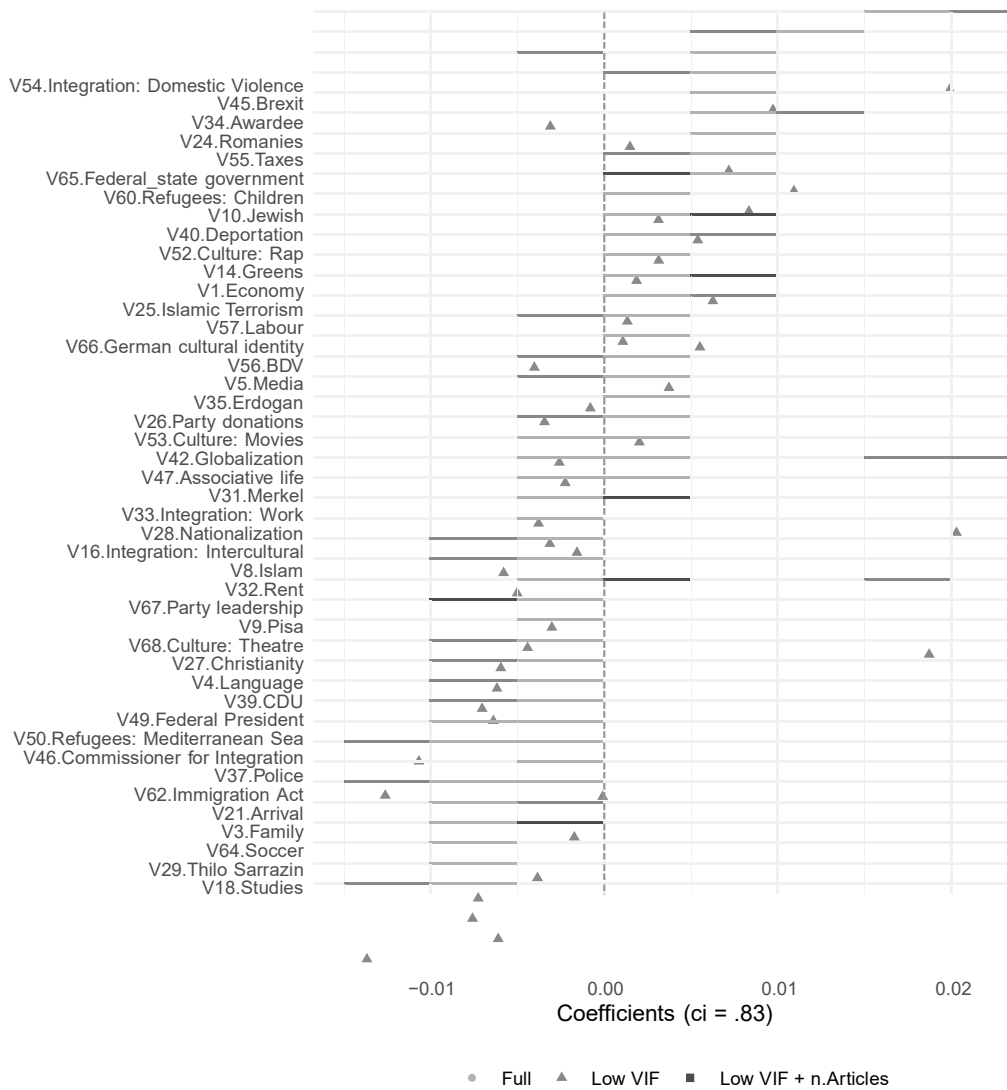shown below in Figure 12.

84

Figure 12: Effects of multicollinearity

Figure 12 reveals multicollinearity in the data, which is in some cases considerable. A major part is due to the number of articles (n.articles, VIF=13.06, Table 6, appendix). This variable is strongly correlated with V23.Refugees General (Figure 4) and Topic V31.Merkel. Consequently, there is a horizontal shift of the effect of Topic V31.Merkel in Figure 12. Moreover, we excluded Topic V23.Refugees General from the models shown in Figure 12. Strong collinearity with n.Articles is due to the refugee crisis in two 2015 (in which chancellor Merkel played an important role) when these topics totally dominated most reports in almost all media. Hence,

85

excludingn.Articles (Figure 12) comes along with a strong change in the effect of these two topics. Overall, however, the results of our main model 2 in Appendix G (including n.Articles) are in line with hypothesis 2, despite of the variance inflation. The robustness check in Figure 12 shows that aside from these two topics, the basic pattern remains stable. Accordingly, there is a large and meaningful set of topics and these topics actually decrease concerns about immigration. In addition, the result for H1 (significant positive effect of n.Articles) is robust and even larger, when no topic variables are included in the model (see Model 1 in Appendix G).