Chapter 10
Google Trends as a Tool for Public Opinion Research: An Illustration of the Perceived Threats of Immigration

Reilly Lorenz, Jacob Beck, Sophie Horneber, Florian Keusch, and Christopher Antoun

10.1 Introduction

Traditionally, social science researchers have relied on surveys to produce population-level estimates of public opinion and behavior. However, surveys are not always feasible, and they generally require substantial time, effort, and money. Search queries made on Google's search engine, by contrast, can be obtained in aggregate form for free from the website Google Trends (GT). Consequently, researchers are starting to construct population measures based on these data. The research process typically involves selecting the keywords intended to measure a particular construct of interest, and then using GT to extract an estimate of the volume of Google search queries made, containing one or more of these keywords, in a particular time and place (Salganik, 2019). Unfortunately, this process yields measures that are of unknown accuracy.

In this chapter, we describe the fundamental ways that Internet search data differ from surveys, with a focus on the characteristics that could affect population estimates. Then, we describe a case study that empirically evaluates a measure of

R. Lorenz
Life Science Alliance, Heidelberg, Germany
e-mail: r.lorenz@life-science-alliance.org

J. Beck
Ludwig Maximilian University of Munich, Munich, Germany
e-mail: jacob.beck@stat.uni-muenchen.de

S. Horneber · F. Keusch (✉)
University of Mannheim, Mannheim, Germany
e-mail: shornebe@mail.uni-mannheim.de; f.keusch@uni-mannheim.de

C. Antoun
University of Maryland, College Park, MD, USA
e-mail: antoun@umd.edu

the salience of the perceived threats of immigration in Germany based on Google Trends (GT) data. We conclude with a discussion of the issues that our case study raises with respect to the potential advantages and disadvantages of using GT data for social science research.

10.2   Google Trends as a Research Tool

Google Trends (https://trends.google.com/)—first launched on May 11, 2006—is run by Google Analytics. At the time of this research, Google has the highest market share of all search engines (about 90%), and Bing is the second most popular with only 2% of the market share (Statcounter, 2020). GT provides a search volume index (SVI) of a keyword, which is the relative popularity of a search term entered in Google's search engine, and measured as a share of a random sample of Google queries[1] in a specific time unit (e.g., day, week, or month) and location. The values of the SVI range from 0 to 100, with a value of 100 indicating a keyword's maximum share of all Google queries during a chosen time and location. For each other time unit, the SVI is calculated as a fraction of the maximum query share time unit. Thus, GT does not provide the absolute number of searches for a term, but rather an estimate of how the popularity of a keyword changes over time. SVI is available at the global and national level as well as the more fine-grained geographic level of a region or city (given that the number of queries for a term are sufficiently high enough to be in accordance with Google's privacy guidelines). It is possible to search for up to five terms simultaneously and compare their popularity within a chosen time and geographic area.

Recently, researchers have demonstrated that search engine queries can be used to study phenomena that are typically measured using surveys. For example, researchers have utilized GT data for studying consumer trends (Vosen & Schmidt, 2011), tracking of disease outbreaks such as influenza (Ginsberg et al., 2009), tracking of economic crises (Jun et al., 2018), and in migration research (e.g., Wladyka, 2013; Vicéns-Feliberty & Ricketts, 2016; Böhme et al., 2020). Chykina and Crabtree (2018) measured concerns about deportation among immigrants in the United States (US) based on the frequency of the search phrase "will I be deported." Stephens-Davidowitz (2014) measured racial animus in the US based on the volume of searches containing a racial epithet directed towards African Americans, and its association with voter preference in presidential elections. However, GT has only been found to be reliable by some of the methodological studies that have evaluated it using criterion measures. For example, some studies found that search queries for political candidates and parties were able to predict poll and election results (Askitas,

---

[1] See https://newsinitiative.withgoogle.com/training/lessons?tool=Google%20Trends&image=trends

2015; Hauge & Lied, 2017), whereas others found that search queries were not able to predict these outcomes any better than chance (Lui et al., 2011; Harford, 2014).

Given these mixed findings, it is critical to identify the features of Internet search data that distinguish them from survey data and other observational data, which could affect data quality. Search data have at least three potential advantages. First, individuals are presumably less influenced by social desirability pressures when making search queries about sensitive topics (e.g., drug use, racism, sexual practices, income, embarrassing health conditions, etc.) than when answering survey questions about these topics (especially when the surveys are interviewer-administered) due to concerns they may have about protecting their anonymity and privacy (Stephens-Davidowitz, 2014). Second, search queries are recorded at extremely high frequencies rather than at discrete time points, which makes it possible to study events (both expected and unexpected) over time without relying on retrospective survey questions for which forgetting might be a problem. Third, GT data are relatively low cost to obtain and easy to use. Accessing these data does not require advanced levels of programming skills or other data science expertise, and the data are virtually free to everyone with a computer and Internet access. The combination of real-time, low cost data is seen as a solution for the need for timely estimates in many areas, often referred to as *nowcasting* (Zagheni et al., 2017).

However, Internet search data have several potential disadvantages. First, they are anonymized and aggregated by geography, making it impossible to conduct individual-level analyses. Second, a search query must be interpreted by a researcher who makes inferences about the characteristics of a particular user, which makes it difficult to establish construct validity (e.g., searching for a particular political party/candidate is not a clear indication of the intention to vote for that party/candidate). Carneiro and Mylonakis (2009) have found that search queries on the same topic might even be entered differently, depending on a person's background, such as level of education, culture, and language. Third, Internet search data are collected from users of a particular search engine at a particular point in time, not from representative samples of the population. It has been well-documented that Internet users tend to be younger, higher educated, and wealthier than non-users (e.g., Anderson et al., 2019; Porter & Donthu, 2006). In addition, among Internet users, not everyone uses Google as a search engine (Mellon, 2013). For example, users with high privacy concerns may opt for using alternative search engines that put a strong focus on protecting users' privacy (e.g., DuckDuckGo[2]). Finally, a search engine may change—in how it's designed, who uses it, and how they use it—over time in ways that are out of researchers' control, which may confound real change in longitudinal data analysis (e.g., see Lazer et al., 2014).

We further explore these trade-offs in a case study using GT to measure perceived immigration-related threats in Germany, with a focus on the suitability of GT data for this purpose.

---

[2]However, at the time of the writing of this study, the market share of DuckDuckGo was less than 1% in Germany (Statcounter, 2020).

10.3 Case Study

For this study, we sought to measure the salience of negative opinions towards immigrants in Germany before and after the influx of refugees from the Middle East and North Africa in 2015. GT data were appealing for two reasons. First, our topic was sensitive, and we presumed that GT data would be less susceptible to social desirability bias than a survey-based measure that relied on respondents having to admit their anti-immigrant views. Second, we wanted to study trends, and GT data were available over the time period of interest.

Our main measures of interest pertain to the perceived threats posed by immigrants as an out-group, consistent with Group Threat Theory (Blumer, 1958) in the context of immigration (Zárate et al., 2004; van Klingeren et al., 2015). As shown in Table 10.1, we examined five types of perceived threats: *economic*, *cultural*, *excess*, *security*, and *sexual*. *Economic threat* represents the concerns of natives that immigration will result in a loss of their resources (e.g., lower wages, fewer jobs). *Cultural* concerns focus on whether immigrants will harm society in other ways (e.g., imposing their religious views, needing language accommodations in schools). *Excess threat*, for our purposes, refers to the perception that Germany was unfairly "burdened" by high numbers of migrants in comparison to other European countries. *Security threats* represent concerns about one's physical safety and safety from crime (Larsson, 2017; Fuchs, 2016). Last, *sexual threats* refer to security concerns about sexual violence (Pruitt et al., 2018), which is potentially salient in Germany because of some high-profile cases of sexual violence committed by male migrants (Johnson & Bräuer, 2016; Pruitt et al., 2018).

10.3.1 Keyword Selection

As Table 10.1 shows, we decided on four to five search terms for each threat category. We selected keywords using three steps. First, we acquired a corpus of immigration-related Facebook posts (further described in Lorenz, 2018). Second, we used automated text analysis to determine which terms occurred most often in the posts, which enabled us to discover relevant terms outside our frame of reference. Finally, from this large list of possible search terms, we manually selected terms that we deemed to be conceptually related to the different threat categories. In addition, we used two groups of five search terms each that were not expected to represent perceived threats of immigration, as a sensitivity check for our analysis. One group, labeled as *neutral terms*, contained migration-related terms that we deemed to be about the topic of migration but neutral in tone. The other group, labeled as *randomly selected terms*, contained terms generated by a random word generator.

Table 10.1 Keywords for five perceived threats of immigration, neutral, and random categories

Perceived threat	Google search terms (Original)	Google search terms (Translation)
Economic threat	Flüchtlinge Euro Flüchtlinge Begrüßungsgeld Asylanten Geld Wirtschaftsflüchtlinge Flüchtlinge Kosten	Refugees euro Refugees welcoming money Asylum seekers money Economic refugees Refugees costs
Cultural threat	Islamisierung Scharia Deutschland Islam in Deutschland Salafisten Deutschland Islamisierung Deutschland	Islamization Sharia Germany Islam in Germany Salafists Germany Islamization Germany
Excess threat	Asylflut zu viele Flüchtlinge Flüchtlingswelle in Deutschland Flüchtlingswelle Asylkrise	Asylum flood Too many refugees Refugee wave in Germany Refugee wave Asylum crisis
Security threat	kriminelle Flüchtlinge kriminelle Ausländer Kriminalität Flüchtlinge Deutschland Kriminalität Flüchtlinge Kriminalität Ausländer	Criminal refugees Criminal foreigners Criminality refugees Germany Criminality refugees Criminality foreigners
Sexual threat	Flüchtling vergewaltigt Flüchtlinge Vergewaltigung Vergewaltigung durch Flüchtlinge Flüchtlinge sexuelle Übergriffe	Refugee rapes Refugees rape Rape by refugees Refugee sexual assault
Neutral terms	Asylantrag Immigration Flüchtlinge Migration unbegleitete minderjährige Flüchtlinge	Asylum request Immigration Refugees Migration Unaccompanied minor refugees
Random terms	Alarm Geburtstag gefallen global Wildnis	Alarm Birthday Oblige Global Wilderness

10.3.2   Data and Methods

We extracted the GT data using R version 3.6.0 (R Core Team, 2018) and the R package gtrendsR (Massicotte, 2019).[3] The GT data were collected retrospectively from October 5, 2013, beginning with the first week after the 2013 German Federal

---

[3]Alternatively, one can also download Google Trends data directly from the GT website by specifying the relevant location and time frame. For further explanation, Google provides a "Trends Help" webpage at https://support.google.com/trends/answer/4365538?hl=en. When comparing multiple search terms over a longer period of time, as we do in our case study, using the R package saves time.

Election, until October 5, 2018 (note that Google Trends data on a weekly basis can only be collected for a maximum time period of 5 years). Geographically, we only included searches that came from German IP addresses. We extracted one dataset for each of the 34 search terms individually, which indicated how popular an individual search term was on a given week during the time period of interest. We calculated the weekly summary SVIs for each of the groups of keywords by averaging the individual SVIs across the individual keywords within a group.

To empirically evaluate our GT-based measures, we compared them to polling data for the Alternative für Deutschland (AfD), a German right-wing party that has run on a largely anti-immigrant platform. The AfD was founded in early 2013 with a Euro-critical orientation (Berning, 2017). In 2015, as many refugees from the Middle East and North Africa were coming to Germany, the AfD established a strong anti-immigrant stance and shifted towards a xenophobic right-wing populist orientation (Schmitt-Beck, 2017). Due to this clear positioning regarding the issue of immigration, and several radical anti-immigrant statements made by some AfD politicians, we assumed that the salience of perceived immigration-related threats in Germany should correlate with the polling outcome for the AfD. We also posited that the anti-immigrant searches may precede (e.g., by several weeks) any changes in the polling results for the AfD, consistent with the notion that individuals may gather information before formulating an opinion (e.g., Druckman et al., 2012; Lux, 2009; Chong & Druckman, 2010).

To measure the popularity of the AfD, we used data from weekly telephone surveys conducted by the Forsa Institute for Scientific Research (accessed on the website www.wahlrecht.de, which provides the results of German polls from different research institutes and also aggregates the poll outcomes). Forsa conducts weekly telephone surveys of the German electorate by asking respondents about their voting intentions in the next federal election (Forsa, 2019).[4] We opted to use data from Forsa because of the weekly frequency of the polls; other institutes publish polling results at most biweekly or less frequently. We chose to use polling data rather than election data because the AfD is a fairly new party without extensive data on its election results.[5] Although the polling data itself might have been subject to social desirability bias and other weaknesses, we assumed that the polls could still reliably measure support for the AfD (which on its face does not seem as sensitive as expressing anti-immigrant views). Indeed, polls predictions tend to provide fairly accurate reflections of election outcomes over time (Norpoth & Gschwend, 2003; Wlezien et al., 2013; Wright et al., 2014).

---

[4] Sample sizes range between 1001 and 2510. Around 95% of the surveys interviewed more than 2000 participants.

[5] A note on the data: Even if GT data were available for our chosen keywords in all regions, we, unfortunately, lacked the appropriate polling data to run an analysis on a regional basis. While the keyword data were available on a weekly basis for some of the German states (many German states did not surpass the privacy threshold, and therefore GT data were not provided), we did not have enough observations for poll or election data because the poll data were not collected often enough or had fewer than 30 observations. This issue highlights one of the drawbacks of using GT in combination with other data sources with respect to data availability and comparability.
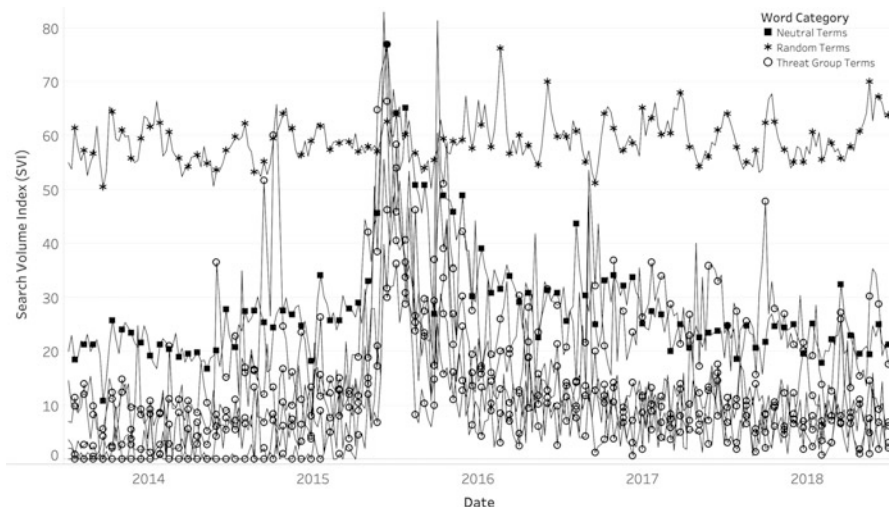
Fig. 10.1 Search Volume Index (SVI) from Google Trends for immigration-related threat group terms, neutral terms, and random terms in Germany from October 2013 to October 2018

We recoded the polling data to the calendar week in which the publication date occurred to be on the same timescale as the GT data. In a case where two polling estimates were published in the same week, we used the arithmetic mean. If no data were available for the week, the mean of the previous week and the following week was imputed. Thus, our analytical data set has no missing values.

We investigated the relationship between the GT data and polling results by using simple bivariate correlation coefficients. We also computed *real-time* correlations between the searches and poll results from the same period. In addition, we computed *temporally-lagged* correlations (Keane & Adrian, 1992; Podobnik & Stanley, 2008) between the keyword searches and poll results that occurred at a later time period than the searches.

10.3.3   Results

Figure 10.1 shows the SVIs for the five *threat* groups, as well as the *neutral* terms and *randomly generated* terms. While variation occurs across the *threat* groups with respect to search volume, the trends across the different *threat* groups is largely consistent: search volume increases in early 2015 and then again later that year, and then decreases to about the original volume in 2016 and later. The increases correspond to when the peak of the influx of refugees from the Middle East and North Africa to Germany occurred. While the SVI of *neutral* terms related to immigration seems to move in a similar pattern, it begins at a higher level before peaking in 2015, and also returns to this higher level compared to the search volume

Table 10.2 Pearson's r for groups of search terms and AfD polling results

| Keyword Group	Correlation coefficient with AfD polling results
Economic threat	−.189**
Cultural threat	−.217***
Excess threat	−.129*
Security threat	.286***
Sexual threat	.057
Neutral terms	−.117
Random terms	.162**

*p < .05; **p < .01; ***p < .001

for the thread terms. This indicates that the volume of searches for the neutral terms was affected by the events in Germany in 2015, but not as much as the search volume for the threat terms. As expected, the SVI for the *randomly selected* terms stayed more or less consistent across our 5 year reference period. If the SVIs of the threat groups reveal actual changes in anti-immigrant opinions over time, we would expect them to correlate with the AfD polling results either in real-time or after a lag period.

As shown in Table 10.2, the real-time correlations between the aggregated Google searches and the AfD polling results varied substantially across the different threat groups. The search terms related to *security threat* were significantly and positively correlated with AfD polling results (r = .29; p < .001). The search terms related to *sexual threat* were also positively correlated with AfD polling results, although the effect did not reach statistical significance (r = .06; p = .361). Surprisingly, each of the other three *threat* groups were significantly and negatively correlated with AfD polling results. As expected, the *neutral* terms showed no significant correlation with AfD polling results (r = −.12; p = .059). However, the SVI of the group of *randomly selected* terms showed a significant positive correlation with AfD polling results (r = .162; p = .009).

The correlation results for individual search terms are provided in the Appendix (Table 10.3). Four individual search terms had a small (r = .2) to medium (r = .5) correlation with the criterion measure in the expected direction, including three terms in the security threat group—*kriminelle Flüchtlinge* (*criminal refugees*; r = .31; p < .001), *Kriminalität Flüchtlinge* (*criminality refugees*; r = .32; p < .001), and *Kriminalität Flüchtlinge Deutschland* (*criminality refugees Germany*; r = .27; p < .001)—as well as one term in the sexual threat group—*Flüchtlinge vergewaltigt* (*refugees rapes*; r = .26; p < .001). At the same time, four individual search terms had a small to medium negative correlation with the AfD polling results: *Flüchtlinge Begrüßungsgeld* (*refugees welcoming money*; r = −.27; p < .001), *Asylanten Geld* (*asylum seekers money*; r = -.34; p < .001), *Wirtschaftsflüchtlinge* (*economic refugees*; r = −.28; p < .001), and *Asylflut* (*asylum flood*; r = −.25; p < .001). Similarly, we found that two of the randomly selected terms had a medium correlation with the criterion—*Geburtstag* (*birthday*; r = .51; p < .001) and *gefallen* (*oblige*; r = .36; p < .001)—and one had a small negative correlation: *Wildnis* (*wilderness*; r = −.22; p < .001).
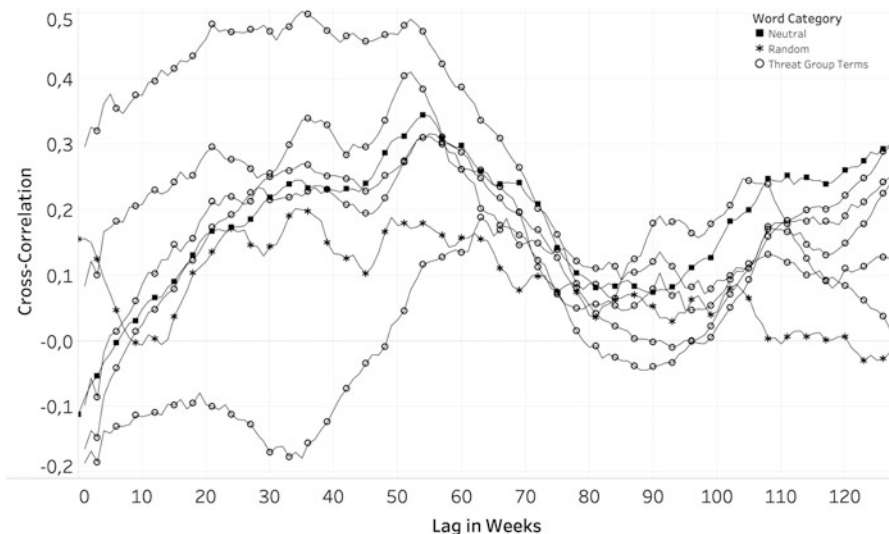
Fig. 10.2 Cross-correlations of Search Volume Index (SVI) from Google Trends for immigration-related threat group terms, neutral terms, and random terms with AfD polling results in Germany from October 2013 to October 2018

Next, we examined the temporally-lagged correlations with a time lag ranging from 1 to 127 weeks. Figure 10.2 shows these correlations for each of the search term groups. This figure reveals that all of the threat categories, even those that showed negative correlation values without a lag, reached medium positive correlation values once a lag was implemented. The peak correlation coefficients for the five threats ranged between .25 (*cultural threat* after 104 weeks) and .50 (*security threat* after 35 weeks). With the exception of the cultural threat group, the curves for the other threat groups showed similar patterns: the correlations increased with a lag up until 35–55 weeks of lag, and decreased rapidly until approximately 85 weeks of lag, and then stayed relatively stable.

The curve for the migration-related neutral terms showed a similar lag pattern and also peaked at a lag of 54 weeks. As expected, the group of randomly selected terms yielded the lowest maximum correlation with the AfD polls and stayed within a range of −.05 and .2, depending on the lag. This finding suggests that the connection between migration-related searches (regardless of their connotation) and AfD poll results was more than just random noise.

10.4   Discussion

Social science researchers are increasingly using alternative non-survey data sources to answer substantive research questions. The aim of this chapter was to explore the advantages and disadvantages of using one of these new data sources, Google

Trends (GT), based on a case study on the perceived threats of immigration. We found that GT data did not consistently correlate with the polling data for the right-wing German AfD in the expected direction in real-time, but rather was consistently predictive of future polling trends (35–104 weeks later) at a moderate level (r = .25 to .50), although the size of the correlations varied across time periods and groups of keywords. By contrast, although a group of randomly selected search terms had a small but significant positive correlation with current AfD polling results, it had the lowest correlations with future polling results. We take this finding as an indication that the correlation between the salience of specific threat-perceptions and AfD polling results is more than just random noise that could be expected from such an amount of data, and, moreover, it seems plausible that a sizable share of these searches was associated with virulent anti-immigrant attitudes.

Our case study highlighted several of the advantages of using GT data. First, we were able to gather the data quickly and at no cost. The data, covering a time period from 2013 to 2018, provided information on how the immigration debate changed in Germany over a long period. Conducting a longitudinal survey on this topic would have been expensive and not possible in retrospect. Second, we measured a topic that triggers social desirability concerns without relying on survey self-reports for which bias is a concern. We assumed that Google users had less concern when typing certain search terms as compared to openly admitting to anti-immigration sentiments in a survey. Third, we conducted a longitudinal analysis with measures recorded at discrete time points (weekly) over a relatively long period of time (5 years). As the AfD began to own the issue of immigration, Internet searches on immigration also increased. We were able to follow the timeline of when the public began searching for crime statistics with a connection to immigration. For example, a large spike in searches about immigration and crime and sexual threat perceptions occurred directly after the New Year's Eve events in Cologne at the start of 2016.[6]

Our case study also highlighted several of the disadvantages of using GT. First, with access to aggregate-level data only, we were unable to explore individual-level correlates (e.g., gender, education) of the perceived threats of immigration. If information on individual voters, such as partisanship, was available, we could have considered whether certain threat cues actually affected voters differently (Lahav & Courtemanche, 2012). Second, the search data were collected from Google users, who are most likely not a representative sample of the population of Germany, and probably not even representative of Internet users in Germany. Without access to sociodemographic and other auxiliary information about the searchers, we could not adjust for any potential bias due to the selectivity of the users who produced these data. Third, the search terms we selected were not equally valid measures of our construct of interest: some keywords were positively corre-lated with a variable that they ought to be related to (AfD support), whereas others

---

[6]The growing opposition to immigration in Germany often is attributed to the 2015/2016 New Year's Eve events in Cologne where a large number of sexual assaults were attributed to male immigrants (e.g., Ingulfsen, 2016).

were negatively correlated with the same variable. The choice of keywords cannot be validated directly, and so this is a significant issue for any research using GT data. Even though we found that the security threats correlated with the AfD polling data, another explanation for our results may be that the topic of crime and immigration is easier to operationalize than other topics using GT. For example, the idea that the Islamic culture will take over Germany was rather difficult to define using only keywords for Internet searches, and GT may not be able to capture these complicated nuances, especially since *culture* is a highly debated concept subject to personal opinion. Fourth, Google's algorithm automatically suggests search terms once a user begins typing. We checked whether the first words of our search terms returned suggestions that were particularly negative towards immigrants, and we found that, at the time of our data collection, this was not the case for our keywords. However, the algorithm may have returned different suggestions over time when users typed immigration-related words, and Google may change its algorithm in the future, which could potentially jeopardize the measurement of long-term trends.

Despite these issues, our case study demonstrates that GT data can be predictive of public opinion, which supports the notion that GT has value for social science researchers as a real-time monitoring tool or leading indicator of public opinion, and it may be especially well suited for measuring socially undesirable views. Future research should investigate which events or phenomena can be reliably measured using GT. Our methodology provides an approach for doing this through the validation of GT-based measures with benchmark survey data. Future research must also address the important questions regarding keyword selection. For example, in the absence of a validation measure, how should keywords be selected and how should they be aggregated into summary measures? These efforts will expand the ways in which social science researchers can leverage Internet search data to produce population-level estimates of public opinion and behavior.

Appendix

Table 10.3 Correlation overview of search terms with AfD poll outcomes

Keyword group	Google search terms (Translation)	Individual correlation coefficients	Combined correlation for perceived threat
Economic threat	Refugees euro	.003	−.189**
	Refugees welcoming money	−.272***	
	Asylum seekers money	−.344***	
	Economic refugees	−.281***	
	Refugees costs	.077	

Table 10.3 (continued)

Keyword group	Google search terms (Translation)	Individual correlation coefficients	Combined correlation for perceived threat
Cultural threat	Islamization	−.195**	−.217***
	Sharia Germany	−.060	
	Islam in Germany	−.199**	
	Salafists Germany	−.143*	
	Islamization Germany	−.160**	
Excess threat	Asylum flood	−.247***	−.129
	Too many refugees	−.178**	
	Refugee wave in Germany	−.115	
	Refugee wave	−.086	
	Asylum crisis	.091	
Security threat	Criminal refugees	.305***	.286***
	Criminal foreigners	−.001	
	Criminality refugees Germany	.269***	
	Criminality refugees	.321***	
	Criminality foreigners	.043	
Sexual threat	Refugee rapes	.255***	.057
	Refugees rape	−.021	
	Rape by refugees	−.072	
	Refugee sexual assault	.061	
Neutral terms	Asylum request	−.004	-.117
	Immigration	−.072	
	Migration	−.151*	
	Refugees	−.162**	
	Unaccompanied minor refugees	−.078	
Random terms	Alarm	−.119	.162**
	Birthday	.511***	
	Oblige	.362***	
	Global	−.002	
	Wilderness	−.222***	

*p < .05; **p < .01; ***p < .001

References

Anderson, M., Perrin, A., Jiang, J., & Kumar, M. (2019). *10% of Americans don't use the internet. Who are they?* Pew Research Center. Retrieved from https://www.pewresearch.org/fact-tank/2019/04/22/some-americans-dont-use-the-internet-who-are-they/

Askitas, N. (2015). *Calling the Greek Referendum on the Nose with Google Trends* (IZA discussion paper). Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2708382

Berning, C. C. (2017). Alternative für Deutschland (AfD) – Germany's New Radical Right-wing Populist Party. *ifo DICE Report*, ifo Institut – Leibniz Institut für Wirtschaftsforschung an der Universität München, München, *15*(4), 16–19.

Blumer, H. (1958). Race prejudice as a sense of group position. *Pacific Sociological Review, 1*(1), 3–7.

Böhme, M. H., Gröger, A., & Stöhr, T. (2020). Searching for a better life: Predicting international migration with online search keywords. *Journal of Development Economics, 142*, 1–14.

Carneiro, H. A., & Mylonakis, E. (2009). Google Trends: A web-based tool for real-time surveillance of disease outbreaks. *Clinical Infectious Diseases, 49*, 1557–1564.

Chong, D., & Druckman, J. N. (2010). Dynamic public opinion. *American Political Science Review, 104*(4), 663–680.

Chykina, V., & Crabtree, C. (2018). Using Google Trends to measure issue salience for hard-to-survey populations. *Socius*. https://doi.org/10.1177/2378023118760414

Druckman, J. N., Fein, J., & Leeper, T. J. (2012). A source of bias in public opinion stability. *American Political Science Review, 106*(2), 430–454. https://doi.org/10.1017/S0003055412000123

Forsa. (2019). Methoden. (Version: May 22 2019) [Website Article]. Retrieved from https://www.forsa.de/methoden/

Fuchs, C. (2016). Racism, nationalism and right-wing extremism online: The Austrian Presidential Election 2016 on Facebook. *Momentum Quarterly, 5*(3), 172–196.

Ginsberg, J., Mohebbi, M. H., Patel, R. S., Brammer, L., Smolinski, M. S., & Brilliant, L. (2009). Detecting influenza epidemics using search engine query data. *Nature, 457*(7232), 1012–1014.

Harford, T. (2014). Big data: Are we making a big mistake? *Financial Times*.

Hauge, H. S., & Lied, T. B. (2017). *Explaining Election Outcomes Using Web Search Data: Evidence from the U.S. Presidential Elections 2008–2016*. Unpublished master's thesis). Norwegian School of Economics.

Ingulfsen, I. (2016, February 18). Why aren't European feminists arguing against the anti-immigrant right? *Open Democracy*. Retrieved from https://www.opendemocracy.net/en/5050/why-are-european-feminists-failing-to-strike-back-against-anti-immigrant-right/

Johnson, H., & Bräuer, T. (2016). Migrant crisis: Changing attitudes of a German city. *BBC News*. Retrieved from https://www.bbc.com/news/world-europe-36148418

Jun, S.-P., Hyoung, S. Y., & Choi, S. (2018). Ten years of research change using Google Trends: From the perspective of big data utilizations and applications. *Technological Forecasting & Social Change, 130*, 69–87.

Keane, R. D., & Adrian, R. J. (1992). Theory of cross-correlation analysis of PIV images. *Applied Scientific Research, 49*, 191–215.

Lahav, G., & Courtemanche, M. (2012). The ideological effects of framing threat on immigration and civil liberties. *Political Behavior, 34*(3), 477–505.

Larsson, A. O. (2017). Going viral? Comparing parties on social media during the 2014 Swedish election. *Convergence: The International Journal of Research into New Media Technologies, 23*(2), 117–131.

Lazer, D., Kennedy, R., King, G., & Vespignani, A. (2014). The parable of Google flu: Traps in Big Data analysis. *Science, 343*(6176), 1203–1205.

Lorenz, R. E. (2018). *Right-wing extremism online: How the AfD frames immigration on Facebook*. Unpublished master's thesis. Universität Mannheim.

Lui, C., Metaxas, P. T., & Mustafaraj, E. (2011). *On the predictability of the U.S. elections through search volume activity*. Proceedings of the IADIS International Conference on e-Society, Avila, Spain.

Lux, T. (2009). Rational forecasts or social opinion dynamics? Identification of interaction effects in a business climate survey. *Journal of Economic Behavior & Organization, 72*(2), 638–655. https://doi.org/10.1016/j.jebo.2009.07.003

Massicotte, P. (2019). *gtrendsR: Perform and display Google Trends queries*. R package version 1.4.3.

Mellon, J. (2013). Internet search data and issue salience: The properties of Google Trends as a measure of issue salience. *Journal of Elections, Public Opinion and Parties, 24*(1), 45–72.

Norpoth, H., & Gschwend, T. (2003). Politbarometer und Wahlprognosen: Die Kanzlerfrage. In A. M. Wüst (Ed.), *Politbarometer*. VS Verlag für Sozialwissenschaften.

Podobnik, B., & Stanley, H. (2008). Detrended cross-correlation analysis: A new method for analyzing two nonstationary time series. *Physical Review Letters, 100*(8).

Porter, C. E., & Donthu, N. (2006). Using the technology acceptance model to explain how attitudes determine Internet usage: The role of perceived access barriers and demographics. *Journal of Business Research, 59*(9), 999–1007.

Pruitt, L., Berents, H., & Munro, G. (2018). Gender and the age in the construction of male youth in the European migration "crisis". *Journal of Women in Culture and Society, 43*(3), 687–709.

R Core Team. (2018). *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/

Salganik, M. (2019). *Bit by bit: Social research in the digital age*. Princeton University Press.

Schmitt-Beck, R. (2017). The 'Alternative für Deutschland in the Electorate': Between single-issue and right-wing populist party. *German Politics, 26*(1), 124–148. https://doi.org/10.1080/09644008.2016.1184650

Statcounter. (2020). *Search engine market share Germany*. Retrieved from https://gs.statcounter.com/search-engine-market-share/all/germany

Stephens-Davidowitz, S. (2014). The cost of racial animus on a black candidate: Evidence using Google search data. *Journal of Public Economics, 118*, 26–40.

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 toward immigration, comparing Denmark and the Netherlands. *European Sociological Review, 31*(3), 268–283.

Vicéns-Feliberty, M. A., & Ricketts, C. F. (2016). An analysis of Puerto Rican interest to migrate to the United States using Google trends. *The Journal of Developing Areas, 50*(2), 411–430. https://doi.org/10.1353/jda.2016.0090

Vosen, S., & Schmidt, T. (2011). *Forecasting private consumption: Survey-based indicators vs. Google trends* (Ruhr Economic Papers 155). RWI – Leibniz-Institut für Wirtschaftsforschung, Ruhr-University Bochum, TU Dortmund University, University of Duisburg-Essen.

Wladyka, D. (2013). The queries to Google Search as predictors of migration flows from Latin America to Spain. *UTB/UTPA electronic theses and dissertations*, 10.

Wlezien, C., Jennings, W., Fisher, S., Ford, R., & Pickup, M. (2013). Polls and the vote in Britain. *Political Studies, 61*(1), 66–91. https://doi.org/10.1111/1467-9248.12008

Wright, M. J., Farrar, D. P., & Russell, D. F. (2014). Polling accuracy in a multiparty election. *International Journal of Public Opinion Research, 26*(1), 113–124. https://doi.org/10.1093/ijpor/edt009

Zagheni, E., Weber, I., & Gummadi, K. (2017). Leveraging Facebook's advertising platform to monitor stocks of migrants. *Population and Development Review, 43*, 721–734.

Zárate, M. A., Garcia, B., Garza, A. A., & Hitlan, R. T. (2004). Cultural threat and perceived realistic group conflict as dual predictors of prejudice. *Journal of Experimental Social Psychology, 40*, 99–105.