Public debate in the media matters: evidence from the European refugee crisis

Journal Article

Author(s):
Koch, Caleb M.; Moise, Izabela; Helbing, Dirk; Donnay, Karsten

REGULAR ARTICLE   **Open Access**

Public debate in the media matters: evidence from the European refugee crisis

Caleb M. Koch[1], Izabela Moise[2], Dirk Helbing[1] and Karsten Donnay[3,4*] (iD)

*Correspondence:
donnay@ipz.uzh.ch
[3]Department of Political Science,
University of Zurich, Zurich,
Switzerland
[4]Department of Politics and Public
Administration, University of
Konstanz, Konstanz, Germany
Full list of author information is
available at the end of the article

Abstract

In this paper, we take a novel approach to study the empirical relationship between public debate in the media and asylum acceptance rates in Europe from 2002–2016. In theory, an asylum seeker should experience the same likelihood of being granted refugee status from each of the 20 European countries we study. Yet, in practice, acceptance rates vary widely for nearly every asylum country of origin. We address this inconsistency with a data-driven approach by analyzing refugee-related news articles and data on asylum decisions across 20 Europe countries for more than 100 asylum seekers' countries of origin. We find that: (i) public debate sentiment in the media is strongly associated with European countries' diverging asylum practices, much more so than social, cultural or economic factors, and (ii) by combining different measures of public debate we can make out-of-sample predictions within 3% of true acceptance rates (on average). We conclude by discussing the practical implications of our findings for European asylum practices.

Keywords: Public debate in the media; Asylum acceptance rates; Granger-causality

1 Introduction

The European refugee crisis—a term used by the United Nations High Commissioner for Refugees (UNHCR) and widely adopted by the media to describe the influx of asylum seekers into Europe in 2015—has played a prominent role in EU politics. For example, "Take back control" was a phrase used by Brexiteers to express the campaign goal of helping the UK take back control over immigration policies. Chancellor of Germany Angela Merkel was strongly criticized for easing the entry of asylum seekers into Germany. Sweden was noted for its positive attitude toward asylum seekers, accepting four refugees per one thousand citizens (the highest in Europe); Hungary, on the other hand, built a fence on its eastern boarder to mitigate the entrance of asylum seekers.

In theory, an asylum seeker should experience the same likelihood of being granted refugee status based on international laws. However, as the above examples suggest, the reality is that asylum practices have varied widely across European countries. There has indeed been considerable debate about how Europe has and should respond to the increased inflow of asylum seekers. But it remains unclear whether or not this public debate has indeed impacted national-level asylum practices in Europe, and whether this can account for the stark variation in actual asylum acceptance rates.

Springer

To study this question, we take a novel data-driven approach that incorporates a large database of refugee-related news articles and data on asylum decisions across 20 European countries for more than 100 asylum countries of origin. We draw the news article data from GDELT [1, 2], which is one of the world's largest databases for news reporting on a broad spectrum of events. In particular, it includes coverage from a diverse set of newspapers for each European language. The scope and breadth of this database allow us to define various measures that capture different aspects of public debate in the media. The data on asylum decisions is drawn directly from the official EU database [3]. In addition, we enrich our statistical analysis by incorporating socio-economic and political controls that have been known to influence national-level asylum decisions [4, 5]. The data cover approximately 55,000 observations of asylum acceptance rates, endowed with a rich set of controls and indexed according to time, European country and asylum country of origin. It should be noted that the distinct country-level perspective taken in this paper offers some advantages: (i) we have access to a considerable amount of data, (ii) we can take a Europe-wide perspective in our analysis, which allows us to compare, for example, the UK and Spain, and (iii) we can study the influence of public debate while controlling for key country-specific socio-economic covariates and important explanatory factors such as refugees' countries of origin. These are novel features of our data-driven approach, and we are the first, to the best of our knowledge, to combine a big dataset of refugee-related news articles with official EU asylum statistics.

We approach our research question from three specific perspectives. First, we study the extent to which public debate can predict European asylum practices. Second, we test whether public debate is a better predictor for political change than other mechanisms. Finally, we analyze the causal structure between public debate and European asylum practices.

Our analysis reveals four key empirical insights. First, we show that public debate in the media on refugee-related topics has indeed varied widely across Europe, and this variation can explain the substantial variance observed in asylum acceptance practices. When we look at different measures, we find that the overall *sentiment* of public debate is what influences asylum practices: negative/positive media sentiment associates with lower/higher acceptance rates. Other measures, such as the volume of refugee-related news coverage, i.e., media attention, are less important. Second, by combining different measures of public debate, we can make out-of-sample predictions within 3% of true European acceptance rates (on average). Third, we show that public debate in the media is a better predictor of national-level asylum practices than social, cultural and economic factors captured in our dataset. Finally, by looking at the causal structure of public debate and asylum practices, we find that public debate strongly influences asylum acceptance rates while the reverse effect is statistically negligible. Taken together, our findings thus highlight the prominent role that public debate in the media plays in national-level policy practices.

The remainder of the paper is structured as follows. In Sect. 2, we provide further background on European asylum practices and review the relevant literature. In Sect. 3, we discuss the datasets underlying our analyses. In Sect. 4, we test the relationship between public debate in the media and European asylum acceptance rates by making in-sample and out-of-sample predictions. In Sect. 5, we study the causal structure of public debate in the media and asylum acceptance rates via Granger-causality tests. In Sect. 6, we conclude by discussing avenues for future work.

2 Background and literature review

This section first provides some historical background and information on current asylum practices. We then give an overview of the relevant literature in order to provide context and to situate the contributions of this paper in the broader scholarship on this topic. Readers who are mainly interested in our data analysis may continue to Sect. 3.
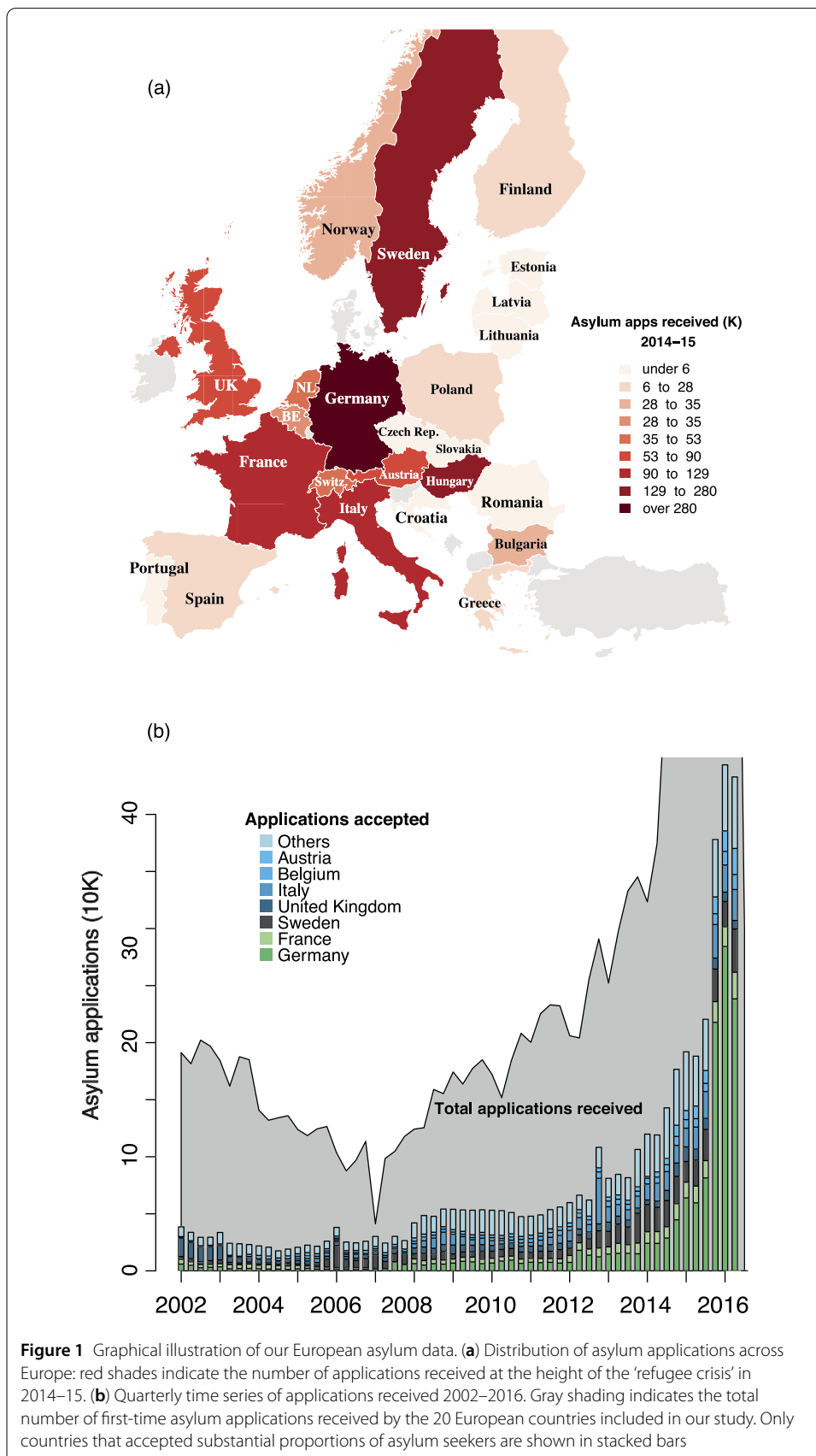
2.1 Historical background

Following World War II, the international political landscape changed with the creation of intergovernmental organizations such as the United Nations and the International Monetary Fund [6, 7]. These organizations, unprecedented at their time, were the first truly international governing bodies tasked with establishing peaceful coexistence of nations and promoting prosperity. Of all the accomplishments that followed, one of the most important was the *Universal Declaration of Human Rights*, which was adopted by the United Nations in December 1948 and formed the basis of subsequent international treaties, economic agreements, regional human rights, and national constitutions [8–11]. Article 14, in particular, commenced international refugee law as it established the right of any person to seek asylum from persecution in other countries. This article was later elaborated at the more well-known 1951 Geneva Convention, which further spelled out the rights of refugees and the responsibilities of countries that grant asylum [12, 13] (while the original intention was to protect post-WWII European refugees, the 1967 Protocol expanded the Geneva Convention's scope to asylum seekers internationally).
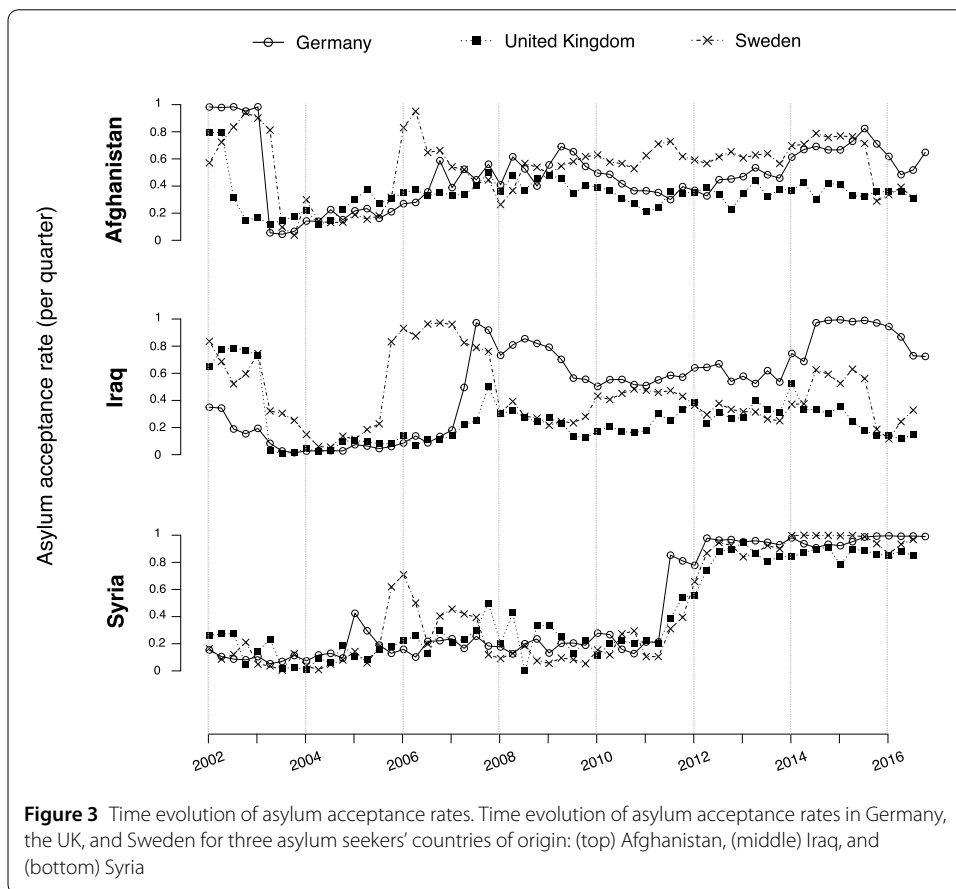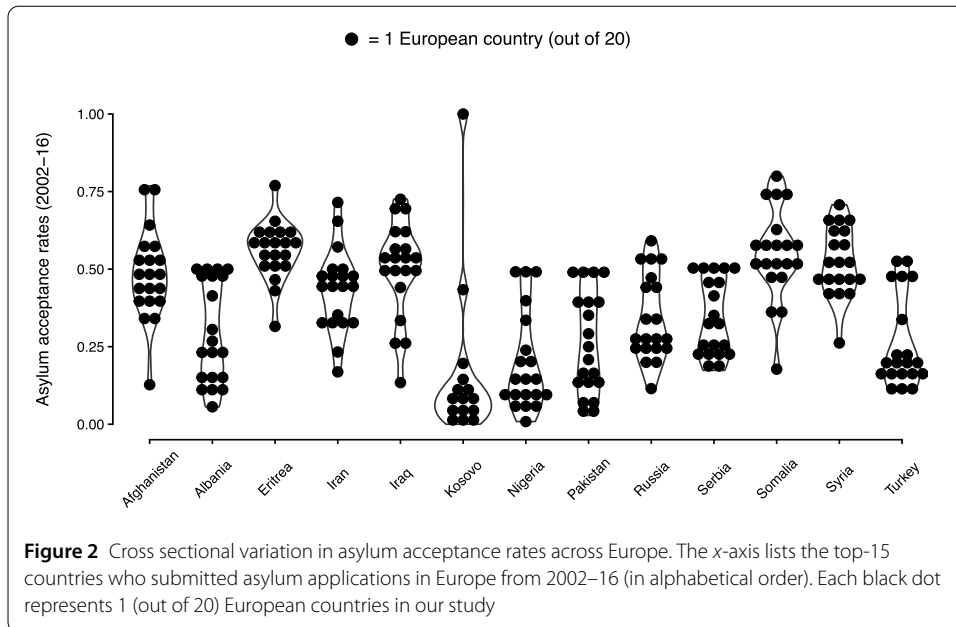
Nearly seventy years later, the 1951 Geneva Convention continues to provide a political mandate for one of the most pressing refugee situations since WWII. Currently, over 20 countries are affected by conflicts with major internal displacement, resulting in a total of 67 million forcibly displaced people—almost one in every 115 humans [14]. Europe has received a particularly large influx of asylum seekers: protracted conflicts in North Africa, the Middle East, and recently Syria have resulted in millions of asylum seekers coming to Europe to seek refugee status (we plot the influx in Fig. 1; see also [15]).[a] The refugee situation has created a divisive political environment, particularly in Europe, where there is large disagreement about how much support can and should be offered to asylum seekers.[b]

2.2 Current asylum practices

According to the 1951 Geneva Convention, the general guideline for granting someone asylum is establishing that he or she *"is unable or unwilling to return to their country of origin owing to a well-founded fear of being persecuted for reasons of race, religion, nationality, membership of a particular social group, or political opinion,"* (see, e.g., [16]). This means that, in theory, asylum seekers from a common country of origin should experience no heterogeneity in the likelihood of being granted 'refugee' status regardless of the country to which they apply.

Yet, in practice, asylum acceptance rates have varied widely across Europe. In Fig. 2, we plot the cross-sectional variation in asylum acceptance rates across the 20 European countries in our study. The starkly different asylum practices are readily apparent.[c] For example, acceptance rates for Afghanistan asylum seekers—who submitted the most asylum applications from 2002–16—ranged from 15 to 75% across Europe. From 2002–2007, the conflict in the Democratic Republic of the Congo prompted more than 60,000 of its citizens to apply for asylum in Europe, and between 10 and 75% of applications were accepted

Figure 1 Graphical illustration of our European asylum data. (**a**) Distribution of asylum applications across Europe: red shades indicate the number of applications received at the height of the 'refugee crisis' in 2014–15. (**b**) Quarterly time series of applications received 2002–2016. Gray shading indicates the total number of first-time asylum applications received by the 20 European countries included in our study. Only countries that accepted substantial proportions of asylum seekers are shown in stacked bars

Figure 2 Cross sectional variation in asylum acceptance rates across Europe. The *x*-axis lists the top-15 countries who submitted asylum applications in Europe from 2002–16 (in alphabetical order). Each black dot represents 1 (out of 20) European countries in our study

Figure 3 Time evolution of asylum acceptance rates. Time evolution of asylum acceptance rates in Germany, the UK, and Sweden for three asylum seekers' countries of origin: (top) Afghanistan, (middle) Iraq, and (bottom) Syria

across Europe. More recently, acceptance rates for the massive influx of Syrian citizens requesting asylum ranged from 27 to 70%. In Fig. 3, we show that this variation is also evident if we look at the time-evolution of asylum acceptance rates: for the same asylum country

of origin, month-to-month acceptance rates can drastically change for different European countries.

How is it possible that, despite the 1951 Geneva Convention, asylum acceptance rates across Europe differ so significantly? One source of this variation is the legal ambiguity of keywords in the definition of a refugee, such as "well-founded" and "persecuted". These words give countries leeway in deciding whether the merit of an asylum seeker's claim matches the country's *interpretation* of refugee [5, 17], which ultimately gives rise to each country's acceptance policies and practices. However, this continues to beg the question: why, then, are countries' legal interpretations of a "refugee" so different across Europe?

These observations are particularly interesting when seen in light of a recent large-scale survey conducted by Bansak et al. [18], who asked 18,000 general citizens across 15 European countries to 'mock' review a total of 180,000 asylum applications. The survey found that if European citizens reviewed asylum applications—rather than civil servants who process asylum applications based on mandates from governmental authorities—the differences in asylum acceptance rates would be insignificant compared to those observed in practice: *"despite the major differences between the countries, there is a considerable consensus in terms of not only the types but also the overall number of asylum seekers that should be admitted,"* ([18], p. 219). The surveys revealed that asylum applicants with high vocation prospects, consistent oppression stories, and who are Christian rather than Muslim received the highest public support.

For all European countries in our study, there thus appears to exist both the same legal mandate and a Europe-wide citizen consensus toward refugees. However, national-level asylum practices have varied widely across Europe. Empirically studying this dichotomy is the main motivation of this paper.

2.3 Prior work on the role of public debate in the media

In this paper, we turn to an explanation that has yet to be explored in the research literature: the extent to which *public debate in the media* shaped national-level asylum practices. The media, sometimes called *"The Fourth Branch of Government"* [19–21], is an institution often viewed as providing a public check on the branches of government. Media is also often viewed as a forum in which public debate on key political issues can take place. Past research has identified different key drivers that shape public debate in the media, such as the media's own political agenda [22], media moguls [23], public opinion [24] and governments engendering public support for policies [25–27]. Beyond the idiosyncratic drivers of public debate in the media, past literature has identified various situations in which public debate in the media has taken an active, and often successful, role in shaping presidential agendas [28] and electoral outcomes [29]. However, we are not interested in how public debate in the media influences electoral outcomes *per se*; we are instead interested in the empirical link with actual policy practices.

It is worth noting that, in the literature, there is an important debate about conditions under which the media takes an active versus passive role in shaping political outcomes, which is sometimes respectively called media power versus media capture [30]. Because we study 20 European countries, which represent a broad spectrum of different political systems, we cannot take a definitive stance about active versus passive role of the media because it is likely to differ across countries and for each news media outlet. For the moment, we focus on the media as a mechanism that could play various roles, which is

without loss of generality for much of the analysis. We take up this discussion again in more detail when we analyze each European country individually.

Why might we expect public debate in the media to have an impact on national-level asylum practices in Europe? Several reasons from past literature point us in this direction. The first is that the European refugee crisis has dominated public debate in the media for the past several years, featuring diverse and often highly polarized opinions, and empirical research has found that politicians are sensitive to this kind of coverage. At one side of the spectrum, pro-refugee news articles have tried to garner support by getting readers to sympathize with humanitarian tragedies of failed treks across the Mediterranean Sea [31, 32]. At the other side, articles opposed to refugees have tended to politicize their presence—particularly that of young men—by depicting them as *"problems"* and *"invaders"* [33]. The second reason is that if public debate in the media reflects citizens' preferences—whose votes might be needed in the next election—or media moguls—whose financial support might be needed in the next election—then political actors might have an incentive to respond to such debate accordingly. Finally, the third reason is that a country's asylum decisions have public and visible consequences, which can increase a political actor's sensitivity to public discontent expressed in the media [34].[d]

It should be noted that our hypothesis—namely that media coverage matters for the asylum seekers' fates—stands in stark contrast to the principles and goals of the 1951 Geneva Convention (see, e.g., [35] and [36]). A refugee claim is made at the *individual*-applicant level: *"... any requirements ... which the particular individual would have to fulfill for the enjoyment of the right in question, if he were not a refugee, must be fulfilled by him,"* (1967 Protocol Relating to the Status of Refugees, Article 6). This claim must also be backed by reasonable evidence that, if this asylum seeker returned to his/her country of origin, then this person would be the target of persecution.[e] However, our hypothesis suggests that asylum decisions are not made vis-à-vis the individual but are influenced by perceptions of refugee-seeking groups as a whole [37]: *"the individual asylum seeker who is escaping persecution is undermined by association with this ostensibly threatening collective,"* ([35], p. 462).[f] In the same vein, [38] found that "threat", "others", "illegality", and "burden" are the four words most closely associated with asylum seekers in newspapers—if our hypothesis is correct, then such words can negatively impact an individual asylum seeker's claim for refugee status.

Perhaps one of the most important means of disseminating public debate in the news is social media, and scholars agree that such platforms have *"... increasing social, economic and political importance"* [39]. A number of studies have shown that individuals experience a sense of self-gratification by sharing news on social media [40–42]. In addition to suppliers of news, social media is the key means by which individuals consume news information, whether or not this is intentional or incidental [43–46]. While the debate on how social media affects news dissemination is still ongoing, one dynamic that has relevance for our study is that social media has been found to polarize those news that do end up in public debate. For example, a recent study in Finland found that social media over-emphasized news articles with crime and threat-oriented themes on refugee issues, while positive articles about refugees where under-emphasized [47].

3 Data

This study leverages two main data sources: (i) a big dataset of refugee-related articles and (ii) official European asylum statistics. We describe both in detail below.

3.1 Media data: GDELT

Data   The first challenge is to represent public debate in the media empirically. To this end, our analysis utilizes a big data repertoire of news articles called the Global Database of Events, Language, and Tone (GDELT; [1, 2]).[g] GDELT is a recent database that includes international and national news coverage from nearly every major online news source (we include a list in Additional file 1). GDELT has been shown to be more comprehensive than other well-known and utilized news article databases, such as ICEWS [48]. Furthermore, GDELT offers updated and sophisticated algorithms, including language translations, which are critical for our analysis. Below, we only highlight the aspects of GDELT that are most pertinent to our analysis; other relevant details are relegated to Additional file 1.

GDELT organizes news articles according to news events, which is crucial for enabling us to build measures capturing the characteristics of refugee-related news coverage. To clarify what this means, it is useful to consider an excerpt from a news article that is included in GDELT:

Example ([49]) "The **United Nations** **will provide** nearly 25,000 tons of emergency **food aid** to **refugees fleeing the civil war in Liberia**, the World Food Program (WFP) said on Monday."

Importantly, GDELT performs what it calls a "principle-actor decomposition analysis". This means that it identifies (1) who is the main actor, (2) what action this actor is taking, and (3) who is the recipient of this action. Concretely, in the above sentence GDELT identified that (1) the United Nations (2) will provide food aid (3) to Liberian refugees. For notation, denote this event as

{United Nations, will provide food aid, Liberian refugees}.

In fact, the principle-actor decomposition analysis is able to abstract beyond specific actors representing a certain country (or its government). This makes it uniquely suitable to construct our country-specific measures of public debate in the media. To showcase this, consider the following excerpt from a 2015 BBC article in the GDELT database:

Example from a 2015 *BBC article included in the GDELT database*:[h] **David Cameron** announced on Monday that the UK **will accept** up to 20,000 **refugees from camps surrounding Syria**.

According to the GDELT algorithm, this event is first identified as

{David Cameron, will accept, Syrian refugees}.

The GDELT algorithm then checks against a dynamically updating database that codifies the country associated with each event. Given that David Cameron is associated with the UK, the example above is codified as

{UK, will accept, refugees}.

The event {UK, will accept, refugees} is then entered into the GDELT database with a date of 08/09/2015 as well as the exact time the article was found.

As mentioned before, the database is fundamentally structured around such news events. Specifically, GDELT codes all other articles that within a fifteen minute time window cover the same event, i.e., {UK, will accept, refugees} still as part of the same news event. From this it then builds two measures for this news event that are of particular interest to us:

(1) the number of articles[i] covering this particular news event and
(2) the average sentiment of these articles.[j]

The average sentiment score is calculated using a standard dictionary-based sentiment algorithm [50, 51]. While perhaps simplistic on the surface, this sentiment measure has been shown to be comparable to human-coded sentiment scores, and it performs comparable to many other state-of-the-art sentiment measures [52].

The data provided by GDELT thus enables us to build country-specific measures for public debate in the media based on their coverage of refugee related events. If we want to find a measure for the *attention* spent on refugee related issues in, e.g. France in March 2015, then we can parse GDELT for all events and count the number of times that articles mention refugee related events. Similar we can use the measures for the average sentiment of each news event recorded by GDELT to determine how *sentiment* of refugee-related media coverage has changed over time. We expand on both below when we describe the measures used for our main analyses.
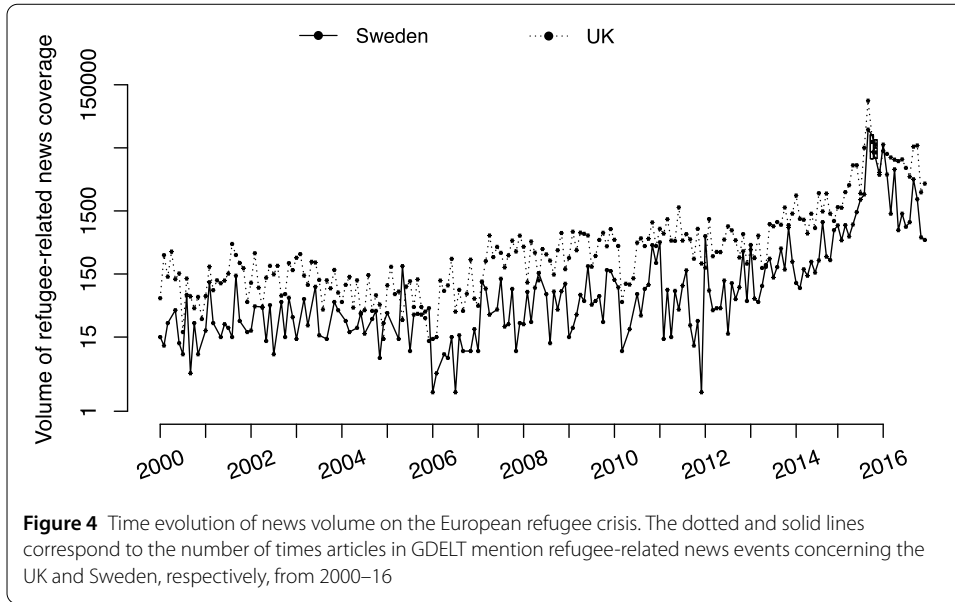
In addition to the article-specific analyses, perhaps the greatest advantage of GDELT is its size and scope. It is known that major news outlets target news audiences [53]. News outlets also have idiosyncratic factors that influence its content, such as geographical reach, political ideology and profit incentives [29, 54]. Our goal is to build empirical variables that capture the debate that results from all of these idiosyncratic features; without the comprehensive nature of GDELT, we could otherwise miss some major news outlets that play an important role. It is also known that both national and international news coverage can influence domestic public debate [55]. Again, GDELT gives us a way to capture both levels of news coverage.

We plot in Fig. 4 the volume of refugee-event related news coverage for Sweden and the United Kingdom resulting from our filter approach described above. As expected, the volume increases during the refugee crisis of 2014–16. However, the absolute volume may be skewed because of changes in the total number of news article published over time, changes in Google News filters, etc. Therefore, in our analyses we normalize our measure of media attention by always considering coverage of refugee-related relative to all events covered in GDELT for each country (see below for details).

Metrics on public debate in the media   We define three general metrics that, on the one hand, leverage the scope of our data and, on the other hand, are subtle enough to detect relevant changes in public debate about refugees. To formally introduce our measures, let

$$\mathcal{E}_{it} = \{\text{all refugee related events about country } i \text{ in time period } t\},$$

where a time period t will typically be a month or quarter (both are used in the analyses below). Note that these are events as captured by GDELT. For each event $E_{it} \in \mathcal{E}_{it}$, let

Figure 4 Time evolution of news volume on the European refugee crisis. The dotted and solid lines correspond to the number of times articles in GDELT mention refugee-related news events concerning the UK and Sweden, respectively, from 2000–16

(1) $N(E_{it})$ denote the number of articles that mention event E_{it} and

(2) $S(E_{it})$ denote the average tone of all articles that mention event E_{it}.

It is also useful to define

$$\mathcal{C}_{it} = \{\text{all events about country } i \text{ in time period } t\}$$

as the set of refugee *and* non-refugee related news events for country i (which means that $\mathcal{E}_{it} \subseteq \mathcal{C}_{it}$). The set $\mathcal{C}_{it}$ allows us to build measures with respect to a baseline.

As our first quantitative measure, we define media *attention* as the number of times that refugee-related events are mentioned in articles divided by the total number of events covered in GDELT:

$$\text{Attention}_{it} = \frac{\sum_{E_{it} \in \mathcal{E}_{it}} N(E_{it})}{\sum_{C_{it} \in \mathcal{C}_{it}} N(C_{it})}. \tag{1}$$
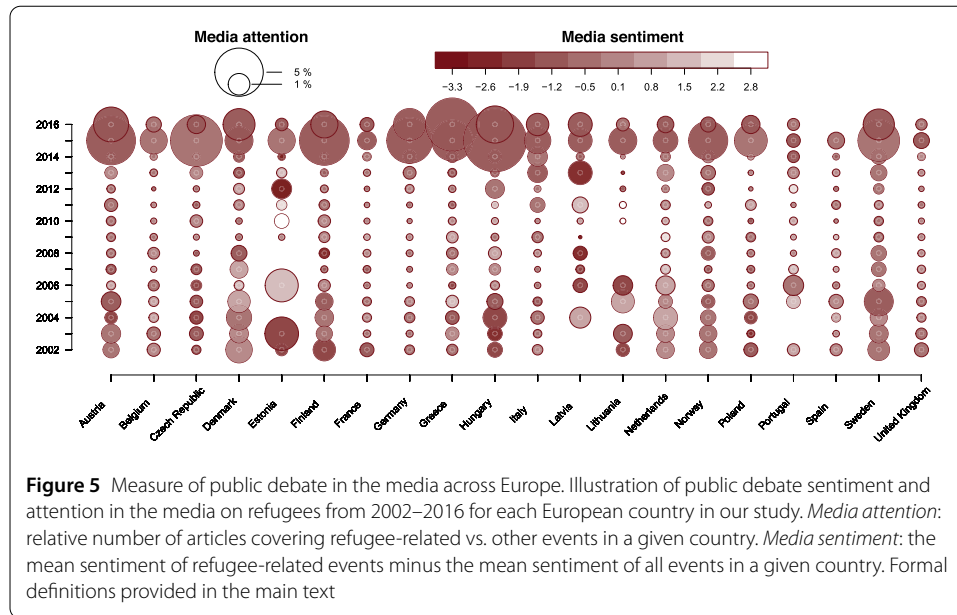
This measure is built for each country i and time period t. A similar metric is used in [28], who studied whether radio attention influences US presidential agendas (see also [56]).

Our second and third quantitative measures focus on the content of the articles covering specific news events. We here leverage the GDELT sentiment scores described above and define two versions of this measure:

$$\text{Sentiment}_{it}^{\text{refugee}} = \frac{1}{|\mathcal{E}_{it}|} \sum_{E_{it} \in \mathcal{E}_{it}} S(E_{it}); \tag{2}$$

$$\text{Sentiment}_{it}^{\text{ref−baseline}} = \text{Sentiment}_{it}^{\text{refugee}} - \frac{1}{|\mathcal{C}_{it}|} \sum_{C_{it} \in \mathcal{C}_{it}} S(C_{it}), \tag{3}$$

where (2) represents the average sentiment of all refugee-related coverage for country i at time period t, and (3) represents the average sentiment of refugee-related coverage relative to the average sentiment of all events (similar metrics can be found in [57].) Note that

Figure 5 Measure of public debate in the media across Europe. Illustration of public debate sentiment and attention in the media on refugees from 2002–2016 for each European country in our study. *Media attention*: relative number of articles covering refugee-related vs. other events in a given country. *Media sentiment*: the mean sentiment of refugee-related events minus the mean sentiment of all events in a given country. Formal definitions provided in the main text

$|\mathcal{C}_{it}|$ denoted the cardinality of $\mathcal{C}_{it}$. Measuring sentiment against a country-level baseline is important because baseline sentiment varies across countries and languages.

It is worth noting that our sentiment metrics benefit from the massive volume of articles captured in the GDLET dataset. Certainly, the sentiment measure of a single article is subject to noise and error. But by generating a sentiment measure from tens of thousands of articles, we average over such noise and acquire a robust 'averaged sentiment' metric.

Illustration of public debate in the media   With these metrics, the large variation of public debate in the media on refugees across Europe is evident in our dataset. Figure 5 shows the average media *sentiment* and *attention* per year for each country in our study. In 2015—the year with the most asylum applications in Europe in the past 40 years—the refugee crisis dominated media coverage with five times more mentions in news articles than any preceding year; this is in contrast to 2003, when Europe also saw peaking numbers of refugees, but reporting remained relatively small. One can observe a similar trend in sentiment. In 2015, the sentiment of public debate in the media was clearly negative across Europe (see dark colors in Fig. 5), while sentiment varied greatly across Europe in 2003. Another perhaps striking observation is the different trends across Europe: while media sentiment and attention toward the refugee crisis in the UK remained relatively constant, countries such as Denmark, Estonia and Hungary experienced large shifts from year to year.

These metrics intentionally break-down the complexity of what constitutes 'public debate in the media' to provide rather simple indicators. Yet, as we show below, these metrics are sufficient to provide considerable predictive power of European asylum acceptance rates.

3.2 Asylum data

Empirical data on European asylum practices is drawn from the official EU database [3] from which we obtain each country's number of *incoming, accepted,* and *rejected* asylum applications by country of origin; the data is available monthly from 2002–07 and quarterly

from 2008–16.[k] In total, our dataset consists of 20 European countries, more than 100 refugee sending countries, and 60 time period (quarters).

All asylum data in our study concern individuals who have formally submitted an application for international protection (or who have been included in such an application as a family member). It should be noted that our statistics do not include or concern refugees who did not submit asylum applications because (i) such statistics are not relevant for predicting or explaining country-specific asylum acceptance rates and (ii) no such data exist.

Asylum acceptance rate   The dependent variable in our analysis is each European country's *asylum acceptance rate*. Below, we describe how the acceptance rate is defined.

From the EU database, we have information regarding the total number of asylum applications *accepted* each quarter (depending on the year).[l] According to the 1951 Geneva Convention, an asylum application is accepted if and only if the individual is *"someone who is unable or unwilling to return to their country of origin owing to a well-founded fear of being persecuted for reasons of race, religion, nationality, membership of a particular social group, or political opinion"* (see, e.g., [16]).[m]

The EU database also has information regarding the total number of asylum applications *rejected* each quarter. Taken together, we thus know the *fraction* of asylum applications accepted vs. rejected, which we define as the *asylum accepted rate*:

$$R_{ijt} = \frac{Accepted_{ijt}}{Accepted_{ijt} + Rejected_{ijt}},$$

where i represents the European receiving country, j represents the asylum country of origin, and t represents the time period. Whenever $Accepted_{ijt} + Rejected_{ijt} = 0$, we drop this observation from our data.

3.3 Control variables

Refugee inflow   In our analysis, it is important to control for the *"floodgate"* effect that has been pointed out in the literature [37, 58]. This effect amounts to asylum acceptance rates decreasing (or increasing) because of the volume of applicants coming, owing to reasons such as physical and/or financial constraints of a country. We thus include in our analysis the total number of *first-time* asylum applicants, which are asylum seekers who apply for refugee status internationally for the first-time (e.g. an asylum seeker who was denied in Germany and re-applied in France is not considered a first-time applicant). This variable is indexed with respect to asylum country of origin, recipient European country, and time period.

Economic indicators   There is empirical evidence that governments may increase or decrease acceptance rates based on labor demand, economic progress and stability [5]. Therefore, we include national GDP, unemployment rate, consumer price index, and government debt as control variables in our analysis (all from the official EU database; see Additional file 1 for more information). As detailed below, we include (i) a between-country version of these variables to account for structural differences in the economic and institutional capacity to accept applicants and (ii) a within-country version to account for the changing economic environments that governments face.

We include two final measures to control for additional relevant dynamics that could confound our results.

Press freedom index   The first accounts for the freedom with which public debate can take place in the media. In general, newspapers are not entirely free when it comes to what they can and cannot publish, especially pertaining to European refugee policies. Any censorship as such could bias our results, as it could take away from the true public debate that is happening in a country. Therefore, we include the "Press Freedom Index" developed by the organization *Reporters Without Borders* (https://rsf.org/en/). The measure is based on a questionnaire given to reporters around the world that asks about media independence, transparency, legislative infrastructure, *et cetera*—in line with what we want to control, this measure captures journalists' perception of press freedom.

Governmental ideology   The second measure included in our analyses accounts for each European government's ideology. Based on a large dataset collected by the "Manifesto Project" (https://manifesto-project.wzb.eu) we know the number of parliamentary seats held by each political party in each European country in our study from 2002–16. In addition, this dataset includes a measure of left-right ideology for each party based on voting behavior.[n] We combine both pieces of information to build our measure for governmental ideology, which we define as the weighted average of party ideologies in a government with respect to the number of parliamentary seats held by each political party.

4 The predictive power of public debate in the media

The aim of this paper is to investigate the empirical relationship between public debate in the media and national asylum practices in Europe. If public debate in the media indeed influenced these decisions, we would expect at least three patterns to exist in our data. (i) Public debate in the media should be able to explain a considerable amount of variance in asylum acceptance rates across Europe, much more so than other control variables included in our data. (ii) Public debate should not only explain variance, but also accurately *predict* asylum acceptance rates. (iii) Finally, we would expect that public debate in the media at time $(t-1)$ is strongly associated with asylum acceptance rates in time (t), but not the other way around. This third test is fundamentally a causality test and rules out the possibility of confounding factors that would make tests (i) and (ii) true. If this test fails and public debate and acceptance rates were mutually predictive, then we cannot say that public debate influences national asylum practices based on (i) and (ii).

In this section, we focus on the first two hypotheses and then separately test the causal structure between public debate and asylum acceptance rates in the next section. We proceed as follows. First, we describe our model for predicting asylum acceptance rates with our media measures and control variables. Second, we present our findings when we make in-sample predictions, which tests whether public debate can explain Europe-wide asylum acceptance rate variance in our data. Finally, we present our findings when we make out-of-sample predictions, which tests whether public debate in the media has predictive power.

4.1 Empirical model

4.1.1 Model setup

We use a mixed-effect regression design to test whether public debate in the media can explain/predict asylum acceptance rates (see, e.g. [59] for an overview of such models). The

mixed-effect design allows us to (i) disentangle the time-invariant and time-dependent influence of variables and (ii) mix fixed and random effects in order to maximize statistical efficiency. Both design features are discussed below.

Defining time-invariant and time-dependent versions of all variables   Why is it important to disentangle time-invariant and time-dependent aspects of our variables? Consider media attention in the Czech Republic and Denmark as measured in Fig. 5 (i.e., the size of the bubble). These trends are clearly different: the Czech republic has roughly the same degree of media attention except for 2015, while Denmark shifts from high to low to high attention from 2002–15. Yet, the *time-average* is nearly the same, i.e., the average size of the bubbles for the Czech Republic and Denmark are roughly equal. When we study the influence of media attention on asylum acceptance rates, we do not want the Czech Republic and Denmark to be treated equally because they have the same time average. Instead, we somehow want to capture both time-invariant features between countries and time-dependent features that captures country-specific trends, allowing us to differentiate the different dynamics occurring in the Czech Republic and Denmark.

We do so by making a distinction between a time-invariant average vs. a time-dependent trend. If S_{it} is European country-i's public debate sentiment in quarter-t, then $\overline{S}_i = \frac{1}{|\mathcal{T}|} \sum_{t \in \mathcal{T}} S_{it}$ is country i's *mean* sentiment from 2002–16. $\overline{S}_i$ allows us to compare the influence of public debate *across European countries* (in the literature this is typically called a 'between-effect'). Similarly, $(S_{it} - \overline{S}_i)$ represents the quarterly-change in public debate sentiment. The expression $(S_{it} - \overline{S}_i)$ allows us to understand if shifts from positive to negative sentiment in public debate influence national-level asylum acceptance rates (in the literature this is typically called a 'within-effect'). We include similar time-variant and time-dependent versions of *all* variables in our model.

Mixed-effect setup   For the sake of notation, let R_{ijt} represent the acceptance rate of European country i of asylum country j during quarter t. Let $\mathbf{X}_{ijt}$ represent the corresponding vector of control variables (e.g. GDP, unemployment rates, and refugee inflow). Let S_{it} represent public debate *sentiment* in the media in European country i during quarter t. Finally, let A_{it} public debate *attention* in the media in European country i during quarter t.

The full mixed-effect model is given as follows:

$$\text{logit} \, R_{ijt} = \alpha_i + \psi_j + \phi_t + \underbrace{\boldsymbol{\beta}' \cdot (\mathbf{X}_{ijt} - \overline{\mathbf{X}}_{ij}) + \boldsymbol{\beta}'' \cdot \overline{\mathbf{X}}_{ij}}_{\text{control variables}}$$

$$+ \underbrace{\gamma_S' \cdot (\text{logit} \, S_{it} - \text{logit} \, \overline{S}_i) + \gamma_S'' \cdot \text{logit} \, \overline{S}_i}_{\text{media sentiment}}$$

$$+ \underbrace{\gamma_A' \cdot (\text{logit} \, A_{it} - \text{logit} \, \overline{A}_i) + \gamma_A'' \cdot \text{logit} \, \overline{A}_i}_{\text{media attention}} + \varepsilon_{ijt},$$

$$\text{where } \varepsilon_{ijt} \sim \mathcal{N}(0, \sigma^2). \tag{4}$$

We introduced a logit transformation on $(R_{ijt}, S_{it}, A_{it})$ because R_{ijt}, S_{it}, *and* $A_{it} \in [0, 1]$ per definition (see Sect. 3). In (4), $(\alpha_i, \psi_j, \phi_t)$ are dummy variables per European country (i), refugee country (j), and time period (t). $\boldsymbol{\beta}'$ is a vector of parameters for our time-dependent control variables, $(X_{ijt} - \overline{X}_{ij})$. $\boldsymbol{\beta}''$ is a vector of parameters for our time-invariant

control variables, $\overline{X}_{ij}$. $(\gamma_S', \gamma_S'', \gamma_A', \gamma_A'')$ are parameters governing the influence of public debate on acceptance rates. Finally, ε_{ijt} is a normally distributed error term with mean 0 and variance σ^2.

One issue in (4) is statistical efficiency because of the fixed-effects, $(\alpha_i, \psi_j, \phi_t)$. There are 20 European countries, 100 asylum seekers' countries of origin, and 60 time periods, which amounts to 180 separate parameters to estimate. With so many parameters, we lose statistical efficiency in estimating our model parameters and the ability to study our main variables of interest, $(\gamma_S', \gamma_S'', \gamma_A', \gamma_A'')$. We overcome this issue by instead modeling these parameters as random variables. This means that we let

$$\alpha_i \sim \mathcal{N}(\mu_\alpha, \sigma_\alpha^2), \qquad \psi_j \sim \mathcal{N}(\mu_\psi, \sigma_\psi^2), \quad \text{and} \quad \phi_t \sim \mathcal{N}(\mu_\phi, \sigma_\phi^2), \tag{5}$$

where we now only estimate six parameters, $(\mu_\alpha, \mu_\psi, \mu_\phi, \sigma_\alpha^2, \sigma_\psi^2, \sigma_\phi^2)$, rather than 180. The key assumption here is that $(\alpha_i, \psi_j, \phi_t)$ are normally distributed. We revisit the validity of this assumption when we look at out-of-sample predictions in Sect. 4.3. Note that we estimate (4) and (5) with a maximum likelihood estimator.

4.1.2 Solving missing data problems: multiple imputations
One problem we must overcome is missing data, which is an issue in our setting. If our data are missing because of non-random reasons, then this 'non-randomness' can be a source of statistical bias [60]. This is an issue in our data because, for example, reporting refugee-related data between 2002–2007 was not obligatory, hence some countries at certain time-points did not report data.

We avoid such bias by utilizing multiple imputations, which has been proven more effective and more robust than the common listwise deletion method [61].[o] The basic idea is to (i) estimate a representative distribution of our data, (ii) generate multiple 'full' versions of the dataset by randomly drawing from this distribution to fill in the missing observations, and then (iii) estimate (4) and (5) utilizing each version of our dataset and combine the results. In doing so, the final results take into account the uncertainty associated with missing data.

Previous studies have suggested that the number of imputations should be approximately equal to the percentage of incomplete observations in the data set (a rule first proposed by [62]). More recently, [61] and [60] proposed that the number of imputations should equal the average percentage of missing data in those columns with any missing data. In our case, the average percentage of missing data is 10.7%; we thus generate 11 imputations.

We combine the estimates of (4) and (5) from each imputation by using the standard 'Rubin combination rule' [63]. This means that: (i) the coefficients are combined using a simple average, (ii) standard errors are combined in a way that accounts for between- and within-variance of the estimated coefficients, and (iii) p-values are computed using the Barnard-Rubin corrected degrees of freedom [64]. We combine p-values by combining test estimates using Rubin's method and then conducting inferences as normal.[p]

4.2 Results: in-sample predictions
With our model at hand, our first goal is to test whether public debate in the media can explain the variance in national-level asylum acceptance rates we observe in Fig. 1(c). We

thus estimate (4) and (5) using all of our data (i.e., via in-sample predictions). We are particularly interested in two aspects of our results: (i) the coefficients of our measures of public debate in the media and (ii) the comparison of R^2 and AICc values between models that use our measures of public debate versus our controls.

Regarding R^2, there exists a large literature on the difficulty of assessing model fitness for mixed-effect models. For our purposes, we utilize two recent metrics proposed in the statistical literature [65, 66]: (i) *Marginal R^2*, describing the proportion of variance explained by the fixed factors alone, and (ii) *Conditional R^2*, describing the proportion of variance explained by both fixed and random factors. In order to combine the Marginal and Conditional R^2 values for each imputation, we use a method proposed by [67], which involves transforming R^2 into a z-score, averaging, and then converting back to an R^2 value.

When we assess model fitness, we also report AICc and BIC in order to understand the trade-off between explaining variance and adding more parameters in our model (this turns out to be important below). We report our empirical estimates in Table 1.

Does public debate in the media explain acceptance rate variance?　　Our first question is perhaps the most important question: do our country-specific measures of public debate in the media explain the stark differences in national-level asylum acceptance rates across Europe?

There are four measures in Table 1 that suggest that they do. The first two are the Marginal and Conditional R^2 values reported toward the bottom of Table 1. Recall that the former represents the variance explained only by fixed-effects while the latter represents the variance explained by fixed-effects plus the random coefficients from (5). As is clear in the table, introducing a normally distributed set of coefficients to represent fixed effects provides an effective means of explaining asylum acceptance rates: Marginal R^2 values range from 0.01 to 0.06, while Conditional R^2 values range from 0.76 to 0.79.

The Conditional R^2 values in particular suggest that a considerable amount of variance can be explained by our media measures. With our measures described above, we can explain up to 76% of the variance of asylum acceptance rates, which is on par with the variance explained by all of our control variables (79%). This means that, rather than having to look at a disparate assortment of economic and political variables to understand national-level asylum trends, we can simply look at trends in public debate in the media to predict asylum acceptance rates.

The final two measures reported in Table 1, namely AICc and BIC, provide further evidence that our public debate measures are effective in explaining asylum acceptance rates. AICc and BIC provide two insights. First, AICc and BIC are defined as measuring the relative amount of information lost/gained about the true underlying process by comparing two models. Lower AICc/BIC values imply that more information is gained, i.e. a better model has been found. Second, AICc and BIC measure the trade-off between explaining data vs. the number of parameters in the model. This means that, if a model has a high R^2 value and a high AICc/BIC, then the data has likely been over-fitted.

With this in mind, the AICc and BIC values reported in Table 1 clearly point to the advantages of our public debate measures. The regression models with only media-related variables—including time-invariant and time-dependent measures of media attention and media sentiment—outperform any model that includes our control variables. Our media models suggests that the explanatory variables offer strong predictive power and justify

Table 1 In-sample results. Regression results for asylum acceptance rates considering public debate in the media and economic indicators as explanatory variables. R^2 values are as in [66]. Coefficients of public debate variables represent logit-logit marginal effects. Coefficients of control variables represent % change in odds with respect to 1 standard deviation change in variable. Small number in parentheses below estimates represents standard error

| | Dep. variable: Logit (Acceptance rates$_{ijt}$) | | | | | |
| | Media models | | Controls models | | Full models | |
	(1)	(2)	(3)	(4)	(5)	(6)
Media measures: $(X_i^t - \overline{X}_i^t)$						
Sentiment, refugee – baseline		0.099				−1.082
		(0.078)				(0.718)
Sentiment, refugee	0.462***				1.186**	
	(0.035)				(0.399)	
Attention	−0.004*	0.006***			0.038*	0.071***
	(0.002)	(0.002)			(0.017)	(0.016)
Media measures: $\overline{X}_i$						
Sentiment, refugee	0.859*	0.97*			0.275	0.983
	(0.38)	(0.38)			(4.515)	(4.592)
Attention	0.029*	0.026*			0.322	0.314
	(0.011)	(0.011)			(0.195)	(0.203)
Time-dependent controls: $(X_i^t - \overline{X}_i^t)$						
Apps. received			0.023	0.01	0.001	0.002
			(0.014)	(0.014)	(0.014)	(0.014)
GDP				−0.066**	−0.074**	−0.072**
				(0.024)	(0.023)	(0.023)
Gov. debt				0.124***	0.111***	0.12***
				(0.027)	(0.026)	(0.027)
Unemployment				0.011	0.011	0.017
				(0.014)	(0.014)	(0.014)
CPI				−0.002	0.003	0.023
				(0.03)	(0.03)	(0.03)
Gov. ideology (left < 0 < right)				−0.019	−0.009	−0.008
				(0.016)	(0.016)	(0.016)
Press freedom index				0.247***	0.211***	0.236***
				(0.021)	(0.023)	(0.022)
Time-invariant controls: $\overline{X}_i$						
Apps. received			0.08	0.038	0.018	0.016
			(0.093)	(0.097)	(0.096)	(0.095)
GDP				−1.297***	−0.812	−0.794
				(0.307)	(0.409)	(0.394)
Gov. debt				1.503**	0.853	0.841
				(0.415)	(0.575)	(0.55)
Unemployment				−0.447**	−0.451**	−0.448**
				(0.149)	(0.15)	(0.149)
CPI				0.309*	0.316*	0.323*
				(0.124)	(0.126)	(0.122)
Gov. ideology (left < 0 < right)				0.089	0.075	0.065
				(0.086)	(0.09)	(0.091)
Press freedom index				0.025	0.351	0.354
				(0.126)	(0.235)	(0.232)
Num. observations	11,593	11,593	11,593	11,593	11,593	11,593
Marginal R^2	0.02	0.01	>0.01	0.06	0.06	0.06
Conditional R^2	0.76	0.76	0.79	0.79	0.79	0.79
AICc	−9662	−9414	42,199	41,724	41,505	41,175
BIC	−9603	−9356	42,243	41,857	41,667	41,337

Note: *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.

additional parameters in the model, while adding economic and political controls adds explanatory power only by adding more parameters in the model and not from the variables themselves. These findings are reinforced when we do out-of-sample predictions in the next subsection, where the under-performance and over-fitting of models (3–6) is even more evident.

Having shown that mass media can explain the variance in our data, we next explore in detail *which* measures of public debate drive these results.

How does public debate in the media influence acceptance rates? To understand how public debate in the media influences acceptance rates, we zoom-in on the media models reported in columns (1) and (2) of Table 1. What is perhaps surprising is that, for both the time-invariant and time-dependent versions of our variables, public debate *attention* does not significantly contribute to explaining national-level asylum acceptance rate variance (in Table 1, note that the media attention coefficients are nearly zero as compared to the media sentiment coefficients). The interpretation is that the volume of public debate is not systematically associated with changes and trends in asylum acceptance rates. This supports previous literature, which has found that media attention is generally not sufficient to shape political change [68].

Instead, we find that the strongest and most significant model parameters are associated with public debate *sentiment*. The emphasis on sentiment versus attention supports previous mass media theory, which suggests that, *"[b]ecause politics is the business of problem solving, negative news automatically turns all heads to politics expecting at least some form of policy reaction,"* [69]. In model (1), we find that a 1-unit increase in the log-odds-ratio in public debate sentiment corresponds to a 0.462 increase in the log-odds-ratio in asylum acceptance rates. Model (5) reports an even higher coefficient of 1.186 (this is explained by the fact that public debate in the media is somewhat correlated with each country's Press Freedom Index, which means that model (1) compensates for this effect with a lower reported coefficient). To gain intuition on the magnitude of this effect, suppose a European country's asylum acceptance rate was 50%, and overall debate sentiment in the media was neutral with $Sentiment_{it}^{\text{refugee}} = 0\%$. Then based on model (5), we estimate that a 5% decrease in quarterly media sentiment—i.e., a somewhat more negative leaning of public debate in the media—correlates with a 3% decrease in asylum acceptance rates. In 2015, this would have translated to a rejection of ~30,000 asylum applications across Europe.

4.3 Results: out-of-sample predictions

Thus far, we have estimated our empirical model using (4) and (5) and all of our data. By doing so, we were able to understand how much variance could be explained with different model setups. In the end, we found strong evidence that our measures of public debate in the media explain considerable variance of asylum acceptance rates across Europe (>75%) and is a better predictor than economic and political variables (based on differences in AICc/BIC values between models 1, 2 and 2–6).

One interesting part of Table 1 is the stark change in AICc and BIC when we include economic and political predictors. The results suggest that this data does not provide additional information for the purposes of explaining asylum acceptance rates. However, with the considerable difference between models (1, 2) vs. models (3–6), we want to further test whether our media models are indeed the best models as suggested by AICc and BIC.

Table 2 Out-of-sample results. Comparison of models in Table 1 with respect to out-of-sample-predictions. Mean square error is estimated via randomly sampling 80% of the data to estimate the model, using the remaining 20% to test the model, and repeating 100 times. SD = standard deviation. The best out-of-sample prediction statistics are reported in *bold*

	Out-of-sample prediction measures			
	Mean absolute error		Mean square error	
	| Predicted% – True% |	(SD)	(Predicted% – True%)^2	(SD)
Media models from Table 1				
Model (1)	**3.08%**	(0.05)	**17.2**	(0.6)
Model (2)	**3.07%**	(0.1)	**17.1**	(0.6)
Control models from Table 1				
Model (3)	35.4%	(0.3)	1507.4	(20.0)
Model (4)	35.4%	(0.3)	1504.3	(17.7)
Full models from Table 1				
Model (5)	35.4%	(0.3)	1504.1	(19.7)
Model (6)	35.4%	(0.3)	1504.5	(17.0)

Thus, in this section we test models (1)–(6) by making out-of-sample predictions. Doing so provides a strong indicator of over-fitting vs. under-fitting of models: over- and under-fitted models tend to be poor at making out-of-sample predictions, while models that find a balance between the two perform well.

We follow a standard procedure called *repeated random sub-sample validation*. First, we randomly sample 80% of the data to fit our regression model (i.e., this is our training data). Second, we utilize this model to predict the values of the remaining 20% of the data (i.e., this is our testing data). Finally, we compare the predicted values vs. the true values via the *mean absolute error* (MAE)—which is the average distance between the predicted and true values—and the *mean square error* (MSE)—which is the average *square* distance between the predicted and true value. We repeat this procedure 100 times in order to report a mean and standard deviation of MAE and MSE for each model. We report values in Table 2.

The results in Table 2 support those from Table 1. Using our measures of public debate in the media, *we are able to make out-of-sample predictions within 3% of true asylum acceptance rates (on average).* This is in stark contrast to models based on economic and political predictions, which only manage to make predictions within 35% of true asylum acceptance rates. This supports the AICc and BIC values reported in Table 1 and our hypothesis that public debate in the media can explain *and* predict national-level asylum acceptance rates across Europe.

5 Granger-causality test

Our analysis thus far has focused on explaining and predicting asylum acceptance rates based on our measures of public debate in the media. Our results point to a strong association between the former and latter, so much so that we can predict national-level asylum acceptance rates across Europe with a 3% accuracy (on average). However, these analyses are fundamentally correlation studies that cannot say whether public debate in the media does or does not influence asylum acceptance rates. There are two reasons why. First, it could be that asylum acceptance rates influence public debate—if so, then the strong association is explained by public debate merely reflecting trends in national-level asylum practices. Second, there could be unobserved variables outside of our analysis that influence both asylum acceptance rates and public debate—if so, then the strong association

is explained by a confounding variable problem and there may not exist any causal relation between asylum acceptance rates and public debate in the media. Both reasons do not pose an issue for explaining variance and making predictions, as in Sect. 4. But both reasons do pose an issue if we want to understand *if* and *how* public debate in the media influences national-level asylum acceptance rates.

In this section, we therefore employ Granger-causality tests [70] to study the causal dependencies between public debate in the media and asylum acceptance rates. The test is motivated by [28], who also utilized this test to study the causal relationship between radio media agendas and presidential agendas. Granger-causality tests are powerful insofar as they impose no *a priori* structural conditions on causal relations; instead, *"[w]e merely ask the data to tell us … which, if any, parameter restrictions are appropriate"* [28, p. 176]. In other words, it is the data that clarifies the causal relationship between variables.

In what follows, we first describe the data and the Granger-causality setup before presenting our empirical findings.

5.1 Data
Using the Granger-causality test requires us to utilize data with a short-time scale because, to test the causal relationship between two time series, we must exclude as many confounding variables as possible. We thus limit this analysis to the *monthly* data from GDELT and on asylum acceptance decisions that are available from 2002–07. We proceed assuming that GDP, unemployment rates and other potentially relevant variables affect country policy on somewhat longer time scales than months (this is reasonable assumption considering that countries report such metrics on a quarterly level). The focus on 2002–07 (rather than 2008–16) is due to data availability: asylum acceptance rates from the official EU database are available at a monthly level during this time period, while all asylum statistics after 2007 are available at the quarterly level. GDELT data is available at the monthly level.

5.2 Granger-causality test setup
Months are denoted as $t, \tau = 1, 2, \ldots, T$. The key relationship we are measuring is if public debate in the media during $\tau \in [t - k, t - 1]$ influences asylum acceptance rates at time t, where $[t - k, t - 1]$ represents k-months before t. As in the main text, let R_{ijt} represent European country i's acceptance rate of asylum seekers from country of origin j at time t. We then define *Model 1* as

$$\text{Model 1:} \quad R_{ijt} = \alpha_0 + \underbrace{\sum_{\tau = t-k}^{t-1} \beta_\tau \cdot R_{ij\tau}}_{\text{autoregressive terms}} , \tag{6}$$

where the interpretation is that autoregressive terms in preceding months $[t - k, t - 1]$ are used to explain acceptance rates in month t.

Because we found strong evidence that public debate *sentiment* is a strong predictor of asylum acceptance rates in Sect. 4, we restrict our attention and scope of our Granger-causality tests to studying this case.

As above, S_{it} denotes the relative sentiment of public debate for European country i in month t, as spelled out in (2). The key idea is to test whether the autoregressive terms from

public debate sentiment contributes to *Model 1* significantly. We thus define *Model 2* as

$$\text{Model 2:} \quad R_{ijt} = \alpha_0 + \underbrace{\sum_{\tau=t-k}^{t-1} \beta_\tau \cdot R_{ij\tau}}_{\text{autoregressive terms}} + \underbrace{\sum_{\tau=t-k}^{t-1} \gamma_\tau \cdot (S_{i\tau} - \overline{S}_i)}_{\text{Granger-causing terms}}. \tag{7}$$

In the manner of [70] and [28], we say that public debate sentiment helps explain asylum acceptance rates if we find that *Model 2* indeed significantly improves *Model 1*. Significance is determined by a Wald test comparing *Model 1* and *Model 2*.

As in the robust regressions, we use multiple imputations to avoid any biases due to non-random reasons for missing data [60], which would otherwise be a problem (see Sect. 4.1.2 for details). We combine the Wald test statistics from the multiple imputed datasets as a D_2 statistic as in [71] (p. 239).

In order to establish Granger-causality, we apply the same test as above but in the reverse direction: specifically, we test whether R_{ijt} autoregressive terms significantly contribute to explaining $(S_{it} - \overline{S}_i)$. The major difference in writing down these equations is that N equations—one for each recipient country—are required while only one new variable was introduced to *Model 2*. To parallel this structure, we include the *average asylum acceptance rate* of a country as the additional variable (rather than 10 or 20 terms and their lags).

We run tests for each European country separately. We report results for the 10- and 20-highest sending refugee countries in Tables 3 and 4, respectively. As reported, we run

Table 3 Granger-causality test results. Granger-causality test of interactions between asylum acceptance rates and public debate sentiment in the media. The 10 asylum seekers' countries of origin with the most decisions made (both accepted and rejected) are included in the regressions. We report Granger-causal testing for 1-month *(1-lag)* and 2-month *(2-lag)* lagged VAR models. Locations with *NA* result from monthly data not being available from the EU database. The D_2 statistic is computed as in [71] and is the result of combining Granger tests from multiply imputed datasets

Country	p-value corresponds to Wald test							
	Public debate sentiment → Acc. rate				Acc. rate → Public debate sentiment			
	D_2 statistic		p-value		D_2 statistic		p-value	
	(1-lag)	(2-lag)	(1-lag)	(2-lag)	(1-lag)	(2-lag)	(1-lag)	(2-lag)
Austria	12.10	8.35	<0.001***	<0.001***	9.72	1.66	<0.001***	0.20
Belgium	26.77	10.45	<0.001***	<0.001***	2.49	1.66	0.12	0.20
Bulgaria	12.47	16.12	<0.001***	<0.001***	4.08	0.55	0.04*	0.48
Czech Rep.	12.81	11.68	<0.001***	<0.001***	2.73	0.49	0.10	0.49
Denmark	13.92	11.41	<0.001***	<0.001***	4.71	1.10	0.03*	0.31
Finland	10.03	9.22	<0.001***	<0.001***	5.47	1.25	0.02*	0.26
France	16.44	6.03	<0.001***	0.01*	16.85	5.70	<0.001***	0.02*
Germany	9.04	3.31	<0.001***	0.07	8.75	3.82	<0.001***	0.05
Greece	9.92	7.14	<0.001***	0.01*	0.10	3.61	0.75	0.06
Hungary	5.77	4.44	0.02*	0.05	6.92	6.72	0.01*	0.01*
Italy	0.40	3.83	0.55	0.08	0.04	2.22	0.85	0.14
Netherlands	5.63	3.56	0.02*	0.06	4.76	2.14	0.03*	0.14
Norway	30.80	11.82	<0.001***	<0.001***	3.77	1.73	0.05	0.19
Poland	8.18	3.88	<0.001***	0.06	1.70	1.13	0.19	0.29
Portugal	0.42	1.02	0.52	0.34	*NA*	0.69	*NA*	0.41
Slovakia	8.81	0.45	*0.42*	0.50	1.79	0.60	0.18	0.44
Spain	6.45	8.39	0.01	<0.001***	4.32	3.54	0.04*	0.06
Sweden	9.25	11.44	<0.001***	<0.001***	5.86	2.45	0.02*	0.12
UK	21.53	12.91	<0.001***	<0.001***	29.04	9.72	<0.001***	<0.001***

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

Table 4 Granger-causality test results with alternative specifications. Granger-causality test of interactions between asylum acceptance rates and public debate sentiment in the media. The 20 asylum seekers' countries of origin with the most decisions made (both accepted and rejected) are included in the regressions. We report Granger-causal testing for 1-month *(1-lag)* and 2-month *(2-lag)* lagged VAR models. Locations with *NA* result from monthly data not being available from the EU database. The D_2 statistic is computed as in [71] and is the result of combining Granger tests from multiply imputed datasets

Country	p-value corresponds to Wald test							
	Public debate sentiment → Acc. rate				Acc. rate → Public debate sentiment			
	D_2 statistic		p-value		D_2 statistic		p-value	
	(1-lag)	(2-lag)	(1-lag)	(2-lag)	(1-lag)	(2-lag)	(1-lag)	(2-lag)
Austria	15.55	14.00	<0.001***	<0.001***	7.32	1.45	0.01*	0.23
Belgium	38.38	29.54	<0.001***	<0.001***	3.26	1.09	0.07	0.30
Bulgaria	22.20	33.02	<0.001***	<0.001***	5.21	2.27	0.04*	0.17
Czech Rep.	49.13	27.35	<0.001***	<0.001***	2.53	0.18	0.11	0.67
Denmark	12.60	16.79	<0.001***	<0.001***	2.74	1.57	0.11	0.21
Finland	19.31	15.53	<0.001***	<0.001***	6.76	1.82	0.01*	0.18
France	19.39	9.34	<0.001***	<0.001***	18.59	6.58	<0.001***	0.01*
Germany	18.02	6.61	<0.001***	0.01	8.03	2.35	<0.001***	0.13
Greece	10.74	19.76	<0.001***	<0.001***	0.07	2.44	0.80	0.12
Hungary	16.09	17.67	<0.001***	<0.001***	4.39	4.42	0.05	0.04*
Italy	6.21	11.64	0.02	<0.001***	0.64	1.42	0.42	0.23
Netherlands	11.62	10.11	<0.001***	<0.001***	5.38	2.52	0.02*	0.11
Norway	33.73	29.85	<0.001***	<0.001***	3.88	1.37	0.05	0.24
Poland	10.89	11.46	<0.001***	<0.001***	1.19	0.97	0.27	0.33
Portugal	0.12	0.03	0.75	0.86	0.85	0.27	0.36	0.61
Slovakia	0.50	3.40	0.49	0.07	0.24	0.26	0.63	0.61
Spain	8.75	14.85	0.01*	<0.001***	5.48	2.76	0.02*	0.10
Sweden	16.73	18.16	<0.001***	<0.001***	5.00	2.37	0.03*	0.12
UK	53.66	22.16	<0.001***	<0.001***	32.88	9.75	<0.001***	<0.001***

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

the test for two cases: in the first we only include one lag term, or $R_{ij(t-1)}$ and $(S_{i(t-1)} - \overline{S}_i)$, and in the second we include two lag terms. The latter test is more stringent than the former test, because adding more terms to Model 1 renders it more robust to additional explanatory terms, thereby making it more difficult to identify a significant D_2 test statistic when comparing Models 1 and 2.

5.3 Main results

Our main findings are surprisingly unambiguous. We find that public debate sentiment significantly contributes to explaining asylum acceptance rates in nearly every European country in our study. Conversely, we find essentially no statistical evidence that the reverse effect exists: in nearly every country in our study (with the exception of the UK), asylum acceptance rates do not help explain public debate sentiment. This trend is systematic in both Tables 3 and 4 (with the exception of the UK). The fact that nearly all countries exhibit a unilateral direction of causality is somewhat noteworthy from a Granger-causality standpoint. The standard interpretation of such a finding is that public debate sentiment Granger-causes asylum acceptance rates.[q]

These results rule out the concerns that were spelled out at the beginning of this section. If there was an unobserved confounding variable influencing both public debate and asylum acceptance rates, then we would not have observed any significant results in Tables 3 and 4. This is because adding new terms to *Model 2* would have added no new informa-

tion to *Model 1*, as the autoregressive terms would have already contained the information from the driving confounding variable. On the other hand, if public debate was influenced by asylum acceptance rates but not the other way around, then we would have observed the opposite results from Tables 3 and 4. Instead, our results provide clear evidence that public debate in the media influences—and is not influenced by—national-level asylum acceptance rates. Taken together with Sect. 4, we have strong evidence that public debate is indeed being heard at the policy level.

6 Discussion and conclusion

Our study was motivated by the observation that—despite the 1951 Geneva Convention, several follow-up international conventions, and a Europe-wide consensus on citizen preferences towards migration (vis-à-vis measured in [18])—we observe significant national-level variation in asylum practices across Europe. One might expect that European countries would have converged on asylum practices, giving equal and judicious treatment of asylum applications in a unified way. However, real-world data suggests the opposite. For nearly every asylum-sending country, national-level asylum acceptance rates differ widely across Europe.

In this paper, we test the hypothesis that public debate in the media can explain much of the variation in asylum practices across Europe. Our data-driven analysis draws on a comprehensive collection of refugee-related news articles (GDELT) and data on 20 European countries, 100 asylum countries of origin, all spanning 2002–16. By taking a Europe-wide perspective, we incorporate many different types of media outlets, thereby many different possible mechanisms that could drive public debate in the media such as the media's own agenda, media moguls, interest groups, citizens' opinions, and governmental efforts to engender public support for policies. While it is beyond the scope of this study to adjudicate between different explanations for the divergence in public debate in the media we observe, our quantitative findings nevertheless suggest that refugee-related public debate in the media is highly indicative of variation in asylum practices.

We find three clear patterns in the data. First, we find that public debate in the media is strongly predictive of European asylum acceptance rate, accounting for nearly 80% of the variation. It turns out that changes in public debate *sentiment* explains most of this variation, much more so than social, economic and political variables captured in our dataset. Second, by combining different measures of public debate in the media, we can make out-of-sample predictions within 3% of true asylum acceptance rates (on average). Third, going beyond correlation analyses, we study the causal relationship between public debate sentiment and asylum acceptance rates via Granger-causality tests with monthly-level data. This offers two advantages: (i) we can rule out many possible confounding factors, especially those that might induce spurious correlations with our quarterly-level analyses, and (ii) it gives us further evidence that public debate is not only associated with national-level asylum acceptance rates, but also contains distinct signals that seem to set a precedence for European asylum practices (we want to emphasize that this is not causal evidence *per se*, rather, it is statistical evidence that points in this direction).

How can we understand the strong predictive power of our media-based measurements, especially in light of the weak predictive power of structural socio-economic and political indicators? Past literature offers a few insights. For example, a study analyzing the US found that hostile policy-level debate toward immigrants is most likely to take place when

communities undergo sudden demographic changes at the same time that news rhetoric politicizes immigration [56]. In other words, the study suggests that the news media offer some type of trigger that builds on sudden demographic stress in a community. Building on this idea, [72] found that German civil servants who are making the asylum applications decisions are more influenced by regional sentiment toward refugees than management directives. Taken together, these studies suggest that public debate in the media—influenced by sudden demographic stress and regional sentiment—can influence asylum acceptance rates directly via civil servants. This is one among several mechanisms that could theoretically explain our results; testing each is left for future work.

We pursued a comparative cross-national research design in order to gain a Europe-wide perspective on the empirical relationship between public debate in the media and asylum practices. This advantage, however, comes at the cost of not being able to adjudicate between country-specific mechanisms that drive public debate in the media. This opens up several interesting avenues for future work. Given data on party deliberations and country-level policy processes, for example, it should be possible to test whether the media indeed serves as a conduit for governments to promote specific asylum policies. In addition, it would be particularly interesting to study the role of social media platforms, for example, whether social media polarizes public debate about issues such as asylum practices and, therefore, contributes to shaping the asylum policy-making process. Another avenue is studying the effect of international versus national public debate on refugees, as well as the effect of different media outlets (such as the New York Times versus CNN) [73]. Finally, another interesting avenue would be examining more detailed data on refugee inflow in order to study the extent to which heterogeneity of refugee profiles across Europe has shaped media coverage indirectly. It is likely that wealthy, educated refugees have better prospects of integrating into society more easily, and perhaps less noticeably, than refugees coming from other backgrounds. If public debate in the media is spurred by social tensions (as suggested by past literature), then European countries that received the latter profile of refugees might have exhibited stronger (and more negative) debates during the European refugee crisis than countries that received more refugees of the former profile. As such, heterogeneity in refugee profiles across Europe could also contribute to the patterns we observe in our study.

Supplementary information

Supplementary information accompanies this paper at https://doi.org/10.1140/epjds/s13688-020-00229-8.

Additional file 1. In the additional file, we (i) describe the GDELT dataset in further detail, and (ii) provide additional information on the EU asylum data and control variables used in this study. (PDF 145 kB)

Abbreviations
AIC, Akaike information criterion; AICc, Akaike information criterion with a correction for small sample sizes; BBC, British Broadcasting Corporation; BIC, Bayesian Information Criterion; CNN, Cable News Network; EU, European Union; GDELT, Global Database of Events, Language, and Tone; GDP, Gross Domestic Product; ICEWS, Integrated Crisis Early Warning System; MAE, Mean Absolute Error; MSE, Mean Square Error; NA, Not Available; UK, United Kingdom; UN, United Nations; UNHCR, United Nations High Commissioner for Refugees; US, United States (of America); VAR, Vector Autoregressive; WWII, Second World War.

Author details
[1]Department of Humanities, Social and Political Sciences, ETH Zurich, Zurich, Switzerland. [2]Swiss Data Science Center, ETH Zurich, Zurich, Switzerland. [3]Department of Political Science, University of Zurich, Zurich, Switzerland. [4]Department of Politics and Public Administration, University of Konstanz, Konstanz, Germany.

Endnotes
[a] Note that the numbers in Fig. 1 only pertain to asylum applications submitted and not the total number of asylum seekers hosted by a country. For example, Germany reported over 1 million asylum seekers having arrived in 2015, yet less than 500,000 submitted applications [15].

[b] It is perhaps worth noting that the current refugee situation has spurred considerable financial commitments from international organizations and countries alike. For example, the UNHCR alone invests more than $1.8 billion in cash-based interventions and assistance (see, e.g. https://www.unhcr.org/5c5c5acc4).

[c] We restrict our analysis to those 20 European countries for which comprehensive asylum statistics are available. For a full list of countries see Fig. 5 or Table 3.

[d] A poll in the UK in 2004, for example, found that 71% of voters believed that asylum seekers were receiving priority in provision of public services over citizens [35].

[e] See [74] for a detailed discussion of individual-level requirements in the case of Sri Lanka Tamil applicants in the late 1980s.

[f] [75, 76], credited as the first to study the link between the media and refugees, took a close look at how the media evolved as Tamil asylum seekers arrived in Western European countries in 1985. van Dijk found two distinct trends that have been confirmed by subsequent case studies [35, 38, 77–80]: (i) the media tended to refer to asylum seekers as a collective rather than individually, and (ii) initial positive discourse eventually declined into language that instantiated threat of 'self' and 'our territory.'

[g] See, e.g., [81] for an analysis of protests using GDELT.

[h] https://www.bbc.co.uk/news/uk-politics-34183150

[i] In the GDELT database, this measure is equivalent to column 'NumArticles'.

[j] In the GDELT database, this measure is equivalent to column 'AvgTone'.

[k] We gathered the data in January 2017.

[l] According to the EU meta-documentation, the number of asylum application acceptances is: "… *the sum of decisions granting refugee status, subsidiary protection status, authorization to stay for humanitarian reasons (for countries where applicable) and temporary protection.*"

[m] Formally, the number of asylum applications accepted is the *'total positive decisions'* made regarding asylum seeking applications. According to the EU meta-documentation, a positive decision refers to, "… *the sum of decisions granting refugee status, subsidiary protection status, authorization to stay for humanitarian reasons (for countries where applicable) and temporary protection.*" Accordingly, a rejected asylum application refers to those that received a 'negative decision.'

[n] This ideology score is updated annual based on past voting behavior; see the information at the Manifesto Project website.

[o] We implement multiple imputations by using the *Amelia II* algorithm for *R* written by [82], which implements a multivariate normal multiple imputation and employs a bootstrapping EM-maximizer to take draws from the posterior distribution.

[p] It is important to note that Rubin's original methods assumes the degrees of freedom for the complete data is infinite. In our case study, the sufficiently high number of observations makes this assumption valid.

[q] A reader has perhaps noted that no coefficients from the regressions are reported. This is because individual coefficients are not readily interpretable as terms autoregressive terms are highly correlated, which yields biased coefficients. However, R^2 terms remain robust with maximum likelihood or OLS estimators. Therefore, in our analyses, we rely solely on model significance and not on β_τ or γ_τ coefficient interpretations—the former is sufficient for establishing Granger-causality.

Publisher's Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References
1. Leetaru KH (2012) Fulltext geocoding versus spatial metadata for large text archives: towards a geographically enriched Wikipedia. D-Lib Mag 18(9):5
2. Leetaru K, Schrodt PA (2013) GDELT: global data on events, location, and tone, 1979–2012. In: ISA annual convention, vol 2, p 4
3. EUROSTAT (2019) Database of the Statistical Office of the European Union. https://ec.europa.eu/eurostat/data/database (last accessed: 10/10/2019)
4. Neumayer E (2004) Asylum destination choice what makes some west European countries more attractive than others? Eur Union Polit 5(2):155–180
5. Neumayer E (2005) Asylum recognition rates in Western Europe: their determinants, variation, and lack of convergence. J Confl Resolut 49(1):43–66
6. Paul JC (1995) The United Nations and the creation of an international law of development. Harvard Int Law J 36:307
7. Schlesinger SC (2003) Act of creation: the founding of the United Nations: a story of superpowers, secret agents, wartime allies and enemies, and their quest for a peaceful world. Westview Press, Boulder
8. Nickel JW (1987) Making sense of human rights: philosophical reflections on the Universal Declaration of Human Rights. University of California Press, Oakland
9. Hannum H (1995) The status of the Universal Declaration of Human Rights in national and international law. Ga J Int Comp Law 25:287–397
10. Morsink J (1999) The Universal Declaration of Human Rights: origins, drafting, and intent. University of Pennsylvania Press, Philadelphia
11. Risse-Kappen T, Risse T, Ropp SC, Sikkink K (1999) The power of human rights: international norms and domestic change. Cambridge University Press, Cambridge
12. Goodwin-Gill GS, McAdam J (2007) The refugee in international law. Oxford University Press, Oxford
13. Zimmermann A, Doerschner J, Machts F (2011) The 1951 Convention relating to the status of refugees and its 1967 Protocol: a commentary. Oxford University Press, Oxford
14. UNHCR (2016) Global report. Technical report, The UN Refugee Agence, Geneva
15. Aiyar S, Barkbu B, Batini N, Berger H, Detragiache E, Dizioli A, Ebeke C, Lin HH, Kaltani L, Sosa S et al (2016) The refugee surge in Europe: economic challenges. Technical report, International Monetary Fund
16. UNHCR (2010) Convention and protocol relating to the status of refugees. Technical report, The UN Refugee Agency, Geneva
17. Noll G (2000) Negotiating asylum: the EU acquis, extraterritorial protection and the common market of deflection. Nijhoff, Boston
18. Bansak K, Hainmueller J, Hangartner D (2016) How economic, humanitarian, and religious concerns shape European attitudes toward asylum seekers. Science 354(6309):217–222
19. Carter D (1959) The fourth branch of government. Houghton Mifflin, Boston
20. Strauss PL (1984) The place of agencies in government: separation of powers and the fourth branch. Columbia Law Rev 84:573
21. Yackee SW (2005) Sweet-talking the fourth branch: the influence of interest group comments on federal agency rulemaking. J Public Adm Res Theory 16(1):103–124
22. Entman RM (2007) Framing bias: media in the distribution of power. J Commun 57(1):163–173
23. Tunstall J, Palmer M (1991) Media moguls. Routledge, New York
24. Jennings W (2009) The public thermostat, political responsiveness and error-correction: border control and asylum in Britain, 1994–2007. Br J Polit Sci 39(4):847–870
25. Negrine R (2003) Politics and the mass media in Britain. Routledge, London
26. Gehlbach S, Sonin K (2014) Government control of the media. J Public Econ 118:163–171
27. Enikolopov R, Petrova M (2015) Media capture: empirical evidence. In: Anderson SP, Waldfogel J, Stroemberg D (eds) Handbook of media economics, vol 1B. Elsevier, London, pp 687–700
28. Wood BD, Peake JS (1998) The dynamics of foreign policy agenda setting. Am Polit Sci Rev 92(1):173–184
29. Prat A, Stroemberg D (2011) The political economy of mass media. In: Acemoglu D, Arellano M, Dekel E (eds) Advances in economics and econometrics. Cambridge University Press, Cambridge, pp 135–187
30. Prat A (2015) Media capture and media power. In: Anderson SP, Waldfogel J, Stroemberg D (eds) Handbook of media economics, vol 1B. Elsevier, London, pp 669–686
31. Johann D, Thomas K (2018) Need for support or economic competition? Implicit associations with immigrants during the 2015 migrant crisis. Res Polit 5(2):1–8
32. Lenette C, Miskovic N (2018) "Some viewers may find the following images disturbing": visual representations of refugee deaths at border crossings. Crime, Media, Culture 14(1):111–120
33. Ward DG (2019) Public attitudes toward young immigrant men. Am Polit Sci Rev 13(1):264–269
34. Soroka SN (2002) Agenda-setting dynamics in Canada. UBC Press, Vancouver
35. Innes AJ (2010) When the threatened become the threat: the construction of asylum seekers in British media narratives. Int Relat 24(4):456–477
36. Tazreiter C (2017) Asylum seekers and the state: the politics of protection in a security-conscious world. Routledge, New York
37. Hathaway JC (2005) The rights of refugees under international law. Cambridge University Press, Cambridge
38. Klocker N, Dunn KM (2003) Who's driving the asylum debate? Newspaper and government representations of asylum seekers. Media Int Aust Inc Cult Policy 109(1):71–92
39. Lee CS, Ma L (2012) News sharing in social media: the effect of gratifications and prior experience. Comput Hum Behav 28(2):331–339

40. Khuntia J, Sun H, Yim D (2016) Sharing news through social networks. Int J Media Manag 18(1):59–74
41. Karnowski V, Leonhard L, Kümpel AS (2018) Why users share the news: a theory of reasoned action-based study on the antecedents of news-sharing behavior. Commun Res Rep 35(2):91–100
42. Thompson N, Wang X, Daya P (2019) Determinants of news sharing behavior on social media. J Comput Inf Syst. https://doi.org/10.1080/08874417.2019.1566803
43. Gil de Zúñiga H, Weeks B, Ardèvol-Abreu A (2017) Effects of the news-finds-me perception in communication: social media use implications for news seeking and learning about politics. J Comput-Mediat Commun 22(3):105–123
44. Boczkowski PJ, Mitchelstein E, Matassi M (2018) "News comes across when I'm in a moment of leisure": understanding the practices of incidental news consumption on social media. New Media Soc 20(10):3523–3539
45. Fletcher R, Nielsen RK (2018) Are people incidentally exposed to news on social media? A comparative analysis. New Media Soc 20(7):2450–2468
46. Masip P, Suau-Martínez J, Ruiz-Caballero C (2018) Questioning the selective exposure to news: understanding the impact of social networks on political news consumption. Am Behav Sci 62(3):300–319
47. Pöyhtäri R, Nelimarkka M, Nikunen K, Ojala M, Pantti M, Pääkkönen J (2019) Refugee debate and networked framing in the hybrid media environment. Int Commun Gaz. https://doi.org/10.1177/1748048519883520
48. Wang W, Kennedy R, Lazer D, Ramakrishnan N (2016) Growing pains for global monitoring of societal events. Science 353(6307):1502–1503
49. Schrodt PA (2012) CAMEO: conflict and mediation event observations event, event and actor codebook. Technical report, Pennsylvania State University, University Park, PA. http://data.gdeltproject.org/documentation/CAMEO.Manual.1.1b3.pdf
50. Esuli A, Sebastiani F (2006) Sentiwordnet: a publicly available lexical resource for opinion mining. In: LREC, vol 6, pp 417–422
51. Baccianella S, Esuli A, Sebastiani F (2010) Sentiwordnet 3.0: an enhanced lexical resource for sentiment analysis and opinion mining. In: LREC, vol 10, pp 2200–2204
52. Ribeiro FN, Araujo M, Gonçalves P, Gonçalves MA, Benevenuto F (2016) Sentibench—a benchmark comparison of state-of-the-practice sentiment analysis methods. EPJ Data Sci 5(23):1
53. Elejalde E, Ferres L, Schifanella R (2019) Understanding news outlets' audience-targeting patterns. EPJ Data Sci 8(1):1
54. Herman ES, Chomsky N (2002) Manufacturing consent: the political economy of the mass media. Pantheon, New York
55. Shaw M (2000) Media and public sphere without borders: news coverage and power from Kurdistan to Kosovo. In: Decision making in a glass house: mass media, public opinion and American and European foreign policy in the 21st century, pp 27–40
56. Hopkins DJ (2010) Politicized places: explaining where and when immigrants provoke local opposition. Am Polit Sci Rev 104(1):40–60
57. Boudemagh E, Moise I (2017) News media coverage of refugees in 2016: a GDELT case study. In: International AAAI conference on web and social media
58. Najem S, Faour G (2018) Debye–Hueckel theory for refugees' migration. EPJ Data Sci 7(1):1
59. McLean RA, Sanders WL, Stroup WW (1991) A unified approach to mixed linear models. Am Stat 45(1):54–64
60. Lall R (2016) How multiple imputation makes a difference. Polit Anal 24(4):414–433
61. van Buuren S (2012) Flexible imputation of missing data. Taylor & Francis, Boca Raton
62. von Hippel PT (2009) How to impute interactions, squares, and other transformed variables. Sociol Method 39(1):265–291
63. Rubin DB (1987) Multiple imputation for nonresponse in surveys. Wiley, New York
64. Barnard J, Rubin DB (1999) Small-sample degrees of freedom with multiple imputation. Biometrika 86(4):948–955
65. Nakagawa S, Schielzeth H (2013) A general and simple method for obtaining R^2 from generalized linear mixed-effects models. Methods Ecol Evol 4(2):133–142
66. Johnson PC (2014) Extension of Nakagawa & Schielzeth's R2GLMM to random slopes models. Methods Ecol Evol 5(9):944–946
67. Harel O (2009) The estimation of R^2 and adjusted R^2 in incomplete data sets using multiple imputation. J Appl Stat 36(10):1109–1118
68. Baumgartner FR, Jones BD, Leech BL (1997) Media attention and congressional agendas. In: Iyengar S, Reeves R (eds) Do the media govern? Politicians, voters, and reporters in America. Sage, Thousand Oaks, pp 349–363
69. Walgrave S, Van Aelst P (2006) The contingency of the mass media's political agenda setting power: toward a preliminary theory. J Commun 56(1):88–109
70. Granger CW (1969) Investigating causal relations by econometric models and cross-spectral methods. Econometrica 37(3):424–438
71. Enders CK (2010) Applied missing data analysis. Guilford, New York
72. Riedel L, Schneider G (2017) Dezentraler Asylvollzug diskriminiert: Anerkennungsquoten von Fluechtlingen im bundesdeutschen Vergleich, 2010–2015. Polit Vierteljahresschr 58(1):23–50
73. Roberts M, Wanta W, Dzwo T-H (2002) Agenda setting and issue salience online. Commun Res 29(4):452–465
74. Hyndman P (1987) The 1951 Convention definition of refugee: an appraisal with particular reference to the case of Sri Lankan Tamil applicants. Hum Rights Q 9:49
75. van Dijk TA (1988) Semantics of a press panic: the Tamil 'invasion'. Eur J Commun 3(2):167–187
76. van Dijk TA (1991) Racism in the press: critical studies in racism and migration. Routledge, New York
77. Brosius H-B, Eps P (1995) Prototyping through key events: news selection in the case of violence against aliens and asylum seekers in Germany. Eur J Commun 10(3):391–412
78. Lubbers M, Scheepers P, Wester F (1998) Ethnic minorities in Dutch newspapers 1990–5: patterns of criminalization and problematization. Gazette 60(5):415–431
79. Khosravinik M (2010) The representation of refugees, asylum seekers and immigrants in British newspapers: a critical discourse analysis. J Lang Polit 9(1):1–28
80. Esses VM, Medianu S, Lawson AS (2013) Uncertainty, threat, and the role of the media in promoting the dehumanization of immigrants and refugees. J Soc Issues 69(3):518–536
81. Steinert-Threlkeld Z, Mocanu D, Vespignani A, Flowler J (2015) Online social networks and offline protest. EPJ Data Sci 4(19):1
82. Honaker J, King G, Blackwell M et al (2011) Amelia II: a program for missing data. J Stat Softw 45(7):1–47