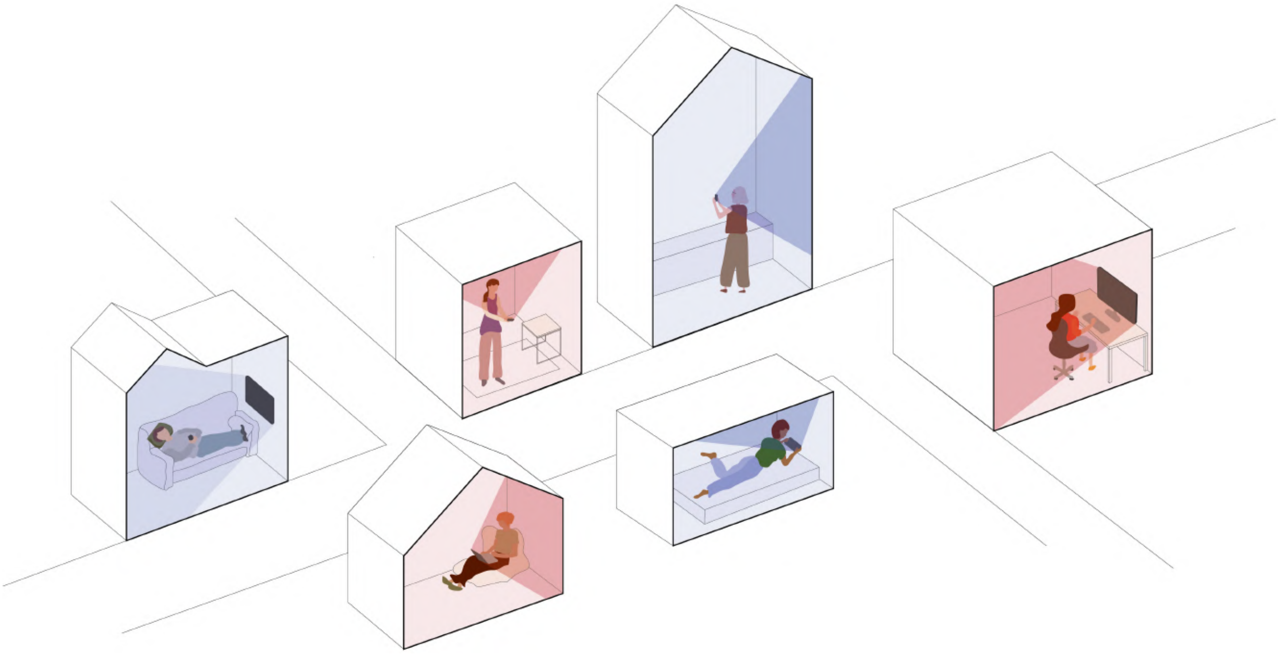
UNIVERSITY OF COPENHAGEN
FACULTY OF SOCIAL SCIENCES



Master Thesis

Knud Anton Højmark Bremholm & Christian Andersen Mølby

Media-Polarisation in Denmark

Beyond the Headlines: Textual and Visual Content from a Decade of News

Type units: 173.937

Advisor: Anne Sofie Beck Knudsen

Submitted: June 1, 2025

Abstract

We study polarisation in article text and lead images from the two Danish online news outlets DR and TV2. By combining computer vision with natural language processing, we extract features from the images. In combination with machine learning techniques and statistical extraction methods, this allows us to distil a wide range of tokens from 168.197 articles published between 2015 and 2024. From a total of 11.954.367 tokens, we obtain unbiased estimates of polarisation using a leave-out estimator and compare model performance to maximum likelihood estimates. We document significant estimates of polarisation in all but 8 out of 260 two-week periods. The average level of polarisation is $0,534$, indicating a 53 pct. probability of a neutral observer with full information correctly identifying the source of an article, given a single, randomly drawn token. This level is comparable to previous findings in American online news. Having established the existence of polarisation in Danish media, we analyse its dynamics over time. We test whether major events - those expected to greatly affect the news stream - influence polarisation. Our findings show that polarisation decreases in the lead-up to general elections. We reject that polarisation increases during the studied period, and find a regime of persistent moderation in polarisation over the past three years. This regime begins around the time of the Russian invasion of Ukraine, which we show has had pronounced effects on the news composition, although we fail to attribute the shift in polarisation directly to the event. Finally, we test the short-term impact of several major events on polarisation, but find no systematic pattern of influence. It remains ambiguous how singular events affect polarisation. Our results indicate that media-polarisation exists in Denmark, though the level of polarisation appears stable.

Distribution key:
Knud Anton Højmark Bremholm (kjh783): 2, 2.2, 3.1, 3.1.2, 3.2, 4.1, 4.1.2, 4.2.1, 4.2.3, 4.2.5, 5.1, 5.3, 5.5, 6, 6.2, 6.4, 6.4.2, 7, 7.1.1, 7.1.3, 7.2.1, 7.2.3, 8.1, 8.3
Christian Andersen Mølby (mjf475): 2.1, 3, 3.1.1, 3.1.3, 4, 4.1.1, 4.2, 4.2.2, 4.2.4, 5, 5.2, 5.4, 5.6, 6.1, 6.3, 6.4.1, 6.5, 7.1, 7.1.2, 7.2, 7.2.2, 8, 8.2
Sections written i collaboration: Abstract, Introduction(1) and Conclusion(9)

Table of content

1 Introduction

High polarisation in a society can have adverse effects, for instance, it poses a potential threat to the effective functioning of democratic institutions (VIVE, 2022). In extension, increases in polarisation "... may have important consequences, including reducing the efficacy of government, increasing the homophily of social groups, and altering economic decisions" Boxell et al. (2020)[1]. These aspects emphasise the importance of assessing polarisation and how it develops over time.

In the United States, a substantial body of research has documented how polarisation has increased over recent decades (Boxell et al., 2020; Gentzkow et al., 2019). As an example of this, Andris et al. (2015) demonstrate that bipartisan cooperation has steadily declined in the US Congress over the past 60 years. Despite the historically consensus-oriented political culture in Denmark, this tendency may translate to a Danish context (Hjorth et al., 2019). Therefore, it raises the central question that motivates our study: is polarisation increasing in Denmark?

There are many types of polarisation, of which some measure to what extent speech, text, content, etc. differs between groups. This paper concerns polarisation in media content, and we use the term to describe differences between media. To understand how these differences arise, we investigate what channels or mechanisms drive polarisation. We narrow the empirical scope to relevant articles and images from two of the most widely used news outlets in Denmark - DR.dk and TV2.dk - in the period 2015-2024. In this paper, we present an approach to estimate polarisation in both text and images from online news. To collect the raw data, we scrape a total of 168.197 articles with accompanying lead images from the two outlets. Recent advances in computer vision, in combination with natural language processing, enable us to achieve state-of-the-art scene tagging from lead images. This is further extended using classical machine learning techniques to extract characteristics of detected faces and to recognise people of interest in the images. Combined with statistical keyword extraction and average sentiment scores from article text, this provides a rich data foundation - unprecedented in the context of Danish media studies.

To obtain unbiased estimates of polarisation, we theoretically evaluate three candidate estimators and identify the leave-out estimator as the most suitable in our context. Using this framework, we estimate both actual polarisation and control polarisation based on random assignment of media to the articles. We estimate significant polarisation in 252 out of 260 two-week periods. Treating polarisation as a time-series, we apply various econometric methods to examine its dynamics across time.

Our approach to studying polarisation in Denmark is inspired by multiple sources. In section 2. we outline the literary context of our contribution, situated at the intersection of the fields of econometrics, social data science, and media studies. We go on to motivate a set of working hypotheses encapsulating the channels to, and temporal dynamics of, polarisation, which we test throughout this paper. From here, we move on to outline the extensive data collection and feature extraction in section 3., which enables our analysis. In section 4. we conduct a descriptive analysis of representation and sentiment across sections and topics. We highlight aspects where the media align and differ to illustrate channels contributing to polarisation. In section 5. we introduce estimators of polarisation of varying complexity, outline our

[1]Boxell et al. ground these consequences in findings from Hetherington and Rudolph (2015), Iyengar et al. (2012), Iyengar et al. (2019), and Gift and Gift (2015)

approach to statistical inference, and present the resulting estimates of polarisation. Having established polarisation in the news coverage, we turn to analysing the dynamics that govern polarisation over time in section 6. Drawing on a range of econometric methods, we examine seasonality, electoral cycles, discontinuities and regime changes, and short-term impacts from major events in the news stream. Further, in section 7., we conduct a robustness check of the estimated polarisation and of the empirical findings from the analysis. Finally, we discuss the results, assumptions, research design, and the broader implications of our findings in section 8., before presenting conclusions and directions for future research in section 9.

2 Literature

> "The political media is biased, but not toward the left or right so much as toward loud, outrageous, colorful, inspirational, confrontational."
> — Ezra Klein, *Why We're Polarized* (2020)

The use of artificial intelligence, machine learning, and other rapidly evolving computational techniques enables the analysis of unconventional data sources, extending the range of empirical contexts suitable for econometric analysis. This paper contributes to the emerging field of empirical polarisation studies, situated at the intersection of econometrics, social data science, and media studies. The latter pertains to the subject of the paper at hand, whereas the former two relate to the methodology applied.

This section has two main components: firstly, it outlines the research that the paper draws on and places the subsequent analysis in a literary context. Secondly, it provides an overview of the theories that we continuously draw on throughout the paper. In doing so, we present and motivate working hypotheses to guide the analysis. The analysis itself is multi-faceted, and these hypotheses serve as a guide to the overall structure, providing reference points that we return to throughout the paper.

2.1 Literary context

Research on polarisation often revolves around affective polarisation, though this is commonly referred to as simply polarisation. Affective polarisation includes, but is not limited to, polarisation in sympathies towards, e.g. political parties, polarisation in stereotypes held against political opponents, or polarisation in social distance to groups or individuals with different political opinions. There are various ways of quantifying affective polarisation, for instance by surveying how respondents answer the question *"How would you feel if your neighbour had a different political affiliation than yourself?"*, as done by Hjorth et al. (2019).

Hjorth et al. (2019) studies affective polarisation in the form of social distance to political opponents based on a survey from 2017. They show signs of social distance to political opponents in Denmark, where it has previously been mostly documented in an American context. Studies of affective polarisation allow for international comparison of the developments in polarisation as well. Boxell et al. (2020) study the trend in affective polarisation across 12 OECD countries, and document increased affective polarisation in five OECD countries over four decades from the 1980's. While polarisation has increased the most in the US, they find an increase in Denmark too. They relate these increases to explanatory factors such as increased ideological distance of political parties and shifts in news consumption patterns. Kingzette et al. (2021) argue that high affective polarisation poses a risk to political norms and has the potential to undermine support for democratic principles, when, e.g., voters in highly polarised environments fail to respect the legitimacy of their political opponents' opinions. VIVE (2022) summarises a number of influential studies that contribute to uncovering polarisation in Denmark, most of which concern affective polarisation. The examples in the study are easily extended to include polarisation from, e.g., hate speech online (Zetland, 2022). VIVE extends the overview of polarisation in Denmark by surveying polarisation in political sympathies and concludes that while there exists some polarisation, it is, to a large degree, a fundamental condition in a democracy.

The paper at hand investigates polarisation in news rather than in affective polarisation. The two are closely related, however, given, how e.g. affective polarisation creates a foundation for echo-chambers in the public debate. If news media attempt to cater to this polarisation, news coverage would likely reinforce affective polarisation. To the best of our knowledge, it remains an open question if that is the case in Denmark. However, Kim et al. (2022) finds that, in an American context, cable news is increasingly polarising, supported by Ash et al. (2023), who finds that news outlets like Fox News evidently have contributed to pushing the US political right rightward, increasing affective polarisation. We relate this to expansions in the Overton window. The Overton window, first presented in the 1990's, describes the window of publicly acceptable speech or policy (Russell, 2006; Mackinac Center for Public Policy, 2023). Singular events, like the election of Donald Trump in 2016, have been linked to an expansion in the Overton window by Bursztyn et al. (2020). They show an increase in the window of publicly acceptable speech following the election, attributable to reduced social consequences of, e.g., expressing xenophobic views. This provides a framework for understanding polarisation in the news coverage The New York Times (2019).

This paper contributes to the existing literature on polarisation in Denmark, most of which examines affective polarisation. A notable contribution to the study of media polarisation in Denmark is Enevoldsen and Hansen (2017), who employ sentiment analysis to assess political bias in Danish newspapers. They analyse 360 articles from Berlingske and Information concerning political parties *Alternativet* and *Liberal Alliance* and find a significant interaction effect between newspapers and parties. The study demonstrates the potential of sentiment analysis, and they suggest future studies construct datasets utilising web-mining methods to gather greater amounts of data. Our paper addresses this and, in doing so, relates the field of polarisation studies to media studies. In combining extensive empirical methods our paper significantly advances the field of media polarisation studies in Denmark. We draw predominantly on three influential methodological contributions in the empirical estimation, which we outline in detail below.

Our primary source of inspiration is the work of Caprini (2024). Caprini explores whether US news media that lean either Democrat or Republican in their political affiliation have a visual bias in the leading images from their respective articles, and documents widespread visual partisanship. The estimated polarisation between Democrat and Republican leaning news outlets in visual language is on average higher than the polarisation from text alone. Constructing the dataset used to estimate polarisation is not trivial. Caprini first collects news articles shared on twitter from a selection of news sources between December 2019 and December 2020. These are filtered to only include unique articles, and the lead images are saved along with the metadata from each article. Employing the computer vision tool Microsoft Azure, an extensive set of features is extracted from each image, creating a substantial set of raw information. The extraction includes people and recognised politicians, coordinates of elements, facial expression, detection of objects and landmarks, and general tags, as examples from a large list of elements. This is passed through a rigorous sorting, filtering, and affiliation process, establishing the subject(s) of each image, the surrounding elements, their relation, and the context of the images, all contributing to decoding the meaning of each image. A central element in the paper is the methodology used to create a comprehensive "Visual vocabulary", forming the basis for which all images in the dataset are mapped to vectors in this vocabulary. These vectors are used to estimate polarisation by a leave-out estimator, following the method developed by (Gentzkow et al., 2019). As the second aspect of the analysis, Caprini presents an experiment where readers are given the same text with different lead images and are asked to give their impression of the news story. This experiment further supports the findings and indicates the importance of not only text focused quality assessment of news and fact checking, but also a control for images. The conclusions of the paper highlight the large degree of partisanship in the American news sphere, with a polarisation in the range between 0,516 and 0,534, with a mean of 0,525. We provide an intuitive interpretation of these measures and relate them to our findings in the subsequent analysis.

This level of polarisation is comparable to the findings of Gentzkow et al. (2019). Gentzkow, Shapiro, and Taddy's paper presents methods of measuring group differences in choices when the choice set is high-dimensional. The paper applies these measures to congressional speech. They consider the problem of measuring differences in a setting where the possible choice set is large compared to the choice outcome observed. They propose methods to combat the severe finite sample bias which traditional approaches suffer from. To illustrate the application, they use text from the United States Congressional Record from the 43$^{\text{rd}}$ to the 114$^{\text{th}}$ congress, split by party and filtered to create a dataset of phrases used by each speaker in each session, including characteristics of each speaker. The structure of the methodology is based on a model of speech constructed with choice probabilities of each phrase at each session, which has proven useful in extracting meaning from text in similar contexts (Groseclose and Milyo, 2005; Taddy, 2013, 2015). Theoretically, they present a maximum likelihood estimation (MLE) and demonstrate the finite sample bias of this traditional method, propose first a leave-out estimator (LO), and then a penalised estimator, which both reduce the bias in different ways. By applying the theoretical estimators, they observe a sharp decrease in polarisation when comparing the results of first the MLE to the LO-estimator, and further when comparing the LO-estimator to penalised estimator. They show that polarisation in congressional speech has been relatively stable, but is far greater in recent years than in the preceding century, and has increased sharply since the 1990's.

While Caprini (2024) establishes that there is visual polarisation in American news, Ash et al. (2022) present ways in which media image content differs. Using online news from New York Times and Fox News over a 20-year time horizon, they apply new methods in computer vision and natural language processing to quantify differences between the media. This allows for an analysis of systematic differences in text and images, as well as investigating how the two domains interact. Of special interest to the paper at hand is their comprehensive dataset, which provides insights into how media coverage might differ in practice. By estimating a vector of identity characteristics, which includes gender and ethnicity as well as the relative dominance in the image frame for people identified in the news images, they obtain frequencies on the representation of different groups. Ash et al. (2022) document negligible differences in the representation of different genders and ethnicities between media and illustrate distinct under-representation of, e.g., females and Hispanics in the images. Conditioning these frequencies on the sections the articles are published under provides an even more granular insight into the composition of news in general and how the media differ specifically. For instance, they find that females are under-represented in all sections apart from *"lifestyle"*, and that the *"sport"*-section over-represents Black people. Differences in how the media represent these groups present an intuitive explanation for some of the polarisation Caprini (2024) estimates. In extension, it points to axes along which Danish media might differ, like their American counterparts.

The three pieces of research outlined above provide an operationally and mathematically appropriate framework for quantifying news-stories and estimating polarisation. However, none of the three studies employ a follow-up study-design of polarisation that aligns with the approach adopted in this study. Instead, we draw on various econometric methods to test hypotheses about polarisation. Our approach integrates models from macroeconomic time-series analysis, political business cycles literature, population-wide policy evaluation, and regression discontinuity designs. We motivate and introduce the relevant econometric techniques throughout the analysis, consistently referencing the assumptions and caveats they impose on our conclusions.

In the remainder of this section and throughout the paper, we include a large number of popular sources. These include, for instance, Danish opinion pieces, news articles, podcasts, and interviews, which we deem relevant when we draw on, e.g., public discourse in the motivation of our working hypotheses. The use of these sources reflects our aim to root the analysis in commonly held beliefs regarding Danish news. Several of the papers we rely on in the estimation of polarisation and from the field of computer vision are working papers. This reflects the recent development of many of the methods we employ, as it has only recently become computationally feasible to, e.g., analyse images at the scale required in this kind of study. Moreover, we include various references to open-source components and Github-repositories, which may be subject to frequent changes and updates. Likewise, we have made all our code pertaining to feature extraction publicly available and reserve the right to continuously update and maintain the repository[2]. The feature extraction is written exclusively in Python (Python Software Foundation, 2025). In addition, we perform all subsequent analysis using R Statistical Software (R Core Team, 2025). Appendix 10.1 contains a list of all R-packages used in the analysis. We use ChatGPT (2025) predominantly in coding and grammar correction. The paper does not include entire pieces of text generated by generative AI.

2.2 Theories and working hypotheses

We define media polarisation as a measure of divergence between how media cover similar material. Theoretically, we expect some polarisation in the news coverage due to market powers: from a micro-economic perspective, media produce news coverage that, if completely objective, would be perfect substitutes of the same product. As economic agents, the media outlets then face incentives to distinguish their coverage from that of competing outlets[3]. They differentiate their products to cater to different consumer tastes by adopting distinct characteristics in their news-products (Thomas J. Nechyba, 2018). Theoretically, these differences are governed by editorial lines that deviate between media. Furthermore, by differentiating the news coverage, the outlets create distinct brands, theoretically leading news-readers to experience some brand-loyalty toward their preferred source of news. Considering brand-loyalty an asset, the media face incentives to reinforce deviating editorial lines to strengthen these effects (Krishnamurthi and Raj, 1991). We refer to this collection of theoretical, micro-economic drivers of polarisation as the competition argument for polarisation. We theorise it results in some natural level of polarisation in Danish media. This leads us to formulate an overarching working hypothesis regarding polarisation in general:

> Main Working Hypothesis:
>
> There is polarisation in Danish media.

The findings in VIVE (2022) and Boxell et al. (2020), as outlined above, support the main working hypothesis. We revisit the hypothesis throughout the paper and investigate the degree to which it holds from different angles using various methods. However, in order to evaluate a

[2] See Mølby and Bremholm (2025).

[3] Assuming that media sequentially decide both product price and differentiation, see Thomas J. Nechyba (2018).

coherent measure of polarisation, we need to investigate the channels that drive polarisation, providing a strengthened ex ante expectation regarding the main working hypothesis. To this end, we formulate three working hypotheses - WH1, WH2, and WH3 - that encapsulate the expected mechanisms. We rely on descriptive analysis to investigate these. Assuming the main working hypothesis holds, it motivates well-defined, verifiable implications, which we summarise in WH4 and WH5. These relate strictly to the dynamics that govern expected polarisation as a time-series. In the following, we present and motivate first WH1, WH2, and WH3, and then go on to present WH4 and WH5.

Ash et al. (2022) presents a natural starting point in investigating theorised polarisation by examining the representation of groups in American media. If Danish media also differ in their visual coverage of specific groups, it is a plausible channel for polarisation. This has the potential to reveal media-wide disparities in the coverage of different groups. That some groups are under-represented in media as a whole is documented in research (Jørndrup, 2021; Pew Research Center, 2022), and suggested by popular sources in the Danish public debate (Berlingske, 2011; Journalisten, 2017). Unless the expected under-representation is equally strong across media, this motivates a specific working hypothesis regarding representation in the media:

Working Hypothesis 1:

The media differ in their visual coverage.

We go on to theorise that the coverage of specific topics in text - somewhat analogue to the coverage of specific groups in images - reveals differences in editorial lines between media. Substantial support for this theory is presented in Caprini (2024). The paper investigates, for instance, the probabilities with which specific topics are covered by different media in different sections, showcasing visible discrepancies across media. We theorise that these differences reflect divergent editorial lines. Regardless, Enevoldsen and Hansen (2017) empirically establishes a link between editorial line and tonality of the news coverage in two Danish news outlets. In conjunction, it allows us to formulate the next working hypothesis:

Working Hypothesis 2:

The media cover similar topics in different tones.

Going back to Gentzkow et al. (2019) we know that some topics can be inherently polarising. In a political context, they document how tax policy is framed as either *"Death-tax"* or *"Estate-tax"* depending on party affiliation. By drawing on the methodological framework presented in Lauridsen et al. (2019b), we are able to investigate an analogue to this in Danish media by illustrating how the topic-specific tonality differs across time and media. Apart from differences in the visual coverage (WH1) and in the editorial line regarding the coverage of specific topics (WH2), we theorise that media might show differences in their political leaning. This theory arises from long-standing public prejudice against the media (Journalisten, 2010; Berlingske, 2010), and from extrapolating differences in Democratic/Republican media coverage in American news outlined above onto a Danish context. To summarise the expectation that media coverage show divergent political bias, we formulate WH3:

Working Hypothesis 3:

The media have distinct political profiles.

In combining the methodological framework of Gentzkow et al. (2019) and Caprini (2024), we are able to quantify a measure of polarisation in Danish news. Immediately, this allows us to test the main working hypothesis: that there is polarisation in Danish media. Intuitively, we expect that if we can find support for the main working hypothesis, then there might be changes to polarisation over time. We ask, what characterises these expected changes? We expect polarisation to increase over time. This expectation is predominately shaped by Boxell et al. (2020), who estimate increased polarisation in five OECD countries - including Denmark. We note that VIVE (2022) contradicts this expectation, as they challenge whether polarisation is actually increasing. As outlined above, however, the findings in both VIVE and Boxell et al. regard affective polarisation, while American literature finds increased polarisation in at least some news forms (Kim et al., 2022), which warrants WH4 regardless of the VIVE study:

> Working Hypothesis 4:
>
> Polarisation increases over time.

Moreover, we can intuitively consider contemporary events as short-term topics in the media coverage. In extension to WH2, we theorise that polarisation might be shaped by specific events at different time-points. Caprini (2023) shows that, in the period following the homicide of George Floyd, the partisanship of the token combination "*police*" and "*weapon*" shifts from Republican to Democratic. This finding empirically supports the expectation that events can shape polarisation. Relating specific events to changes in the Overton window, as outlined above, provides a mechanism through which influential events might contribute to polarisation: when the window expands, it increases the width of the language which news can use to describe events. We go on to formulate WH5, encapsulating these two aspects:

> Working Hypothesis 5:
>
> Contemporaneous events drive polarisation.

Throughout the paper, we return to these working hypotheses to motivate our analytic angles and create a structure for the analysis. We consider testing the main working hypothesis the overarching objective of the paper, while WH1-WH5 guide the analysis. Because we employ various methods to test the working hypotheses, we present preliminary conclusions on the validity of the hypotheses as they arise, and then revisit the robustness of the preliminary conclusions in later sections. We summarise our final conclusions using the structure that the working hypotheses represent.

3 Data

The literature outlined in the previous section provides a theoretical foundation for a quantitative investigation of suspected polarisation. Performing an empirical analysis of polarisation, however, poses a number of challenges and presents various methodological decisions, not least of which relate to constructing a dataset representing the news stories. Understanding how we obtain the data and what it contains is key to the later estimation of polarisation. Therefore, the following section contains a detailed description of the extensive methods applied in collecting the dataset outlined in section 3.1. This subsection details central assumptions and limitations of the data foundation, outlines the advanced computational resources the process utilises, and presents the formatting that allows for subsequent analysis. Section 3.2., describes in detail how the data composition varies over time and its implications.

3.1 Data sources

The analysis is centered around two of the most popular online news outlets in Denmark: DR.dk (Danish Broadcasting Corporation) and TV2.dk (DR, 2025a; TV2, 2025). We aim to collect all relevant news articles and the corresponding leading article image for both news outlets from 01/01/2015 to 31/12/2024. We base our analysis on online news articles due to the combination of both text and image, from which we can extract key characteristics. Accordingly, we discard all news types not following a traditional composition of a main article text with at least one accompanying image.

This traditional newspaper-style news content is created and published by a handful of major news outlets, many of which publish both in-print and online news, and numerous smaller internet-based niche media outlets. By limiting the scope to articles from DR and TV2 we thus discard not only a range of news outlets from the analysis but also limit the data type to online media news stories. Strictly online news might differ from in-print media in both the content covered and the way it is done, with online media having, e.g., a bias towards publishing "breaking" stories or posting live-coverage. We require news articles to be published online, as text and image extraction of archived in-print media is practically infeasible even for online archives[4]. This imposes yet another requirement on the outlets, namely that they have published news online over the entire timeline. For both outlets the period prior to 2015/2016 is characterised by a relatively low daily article count and frequent changes in format and webpage architecture. This imposes a natural cut-off point in the data.

Methodological constraints, such as the dimensionality of our later polarisation estimator, limit the analysis to two news outlets[5]. Given this constraint we choose two media which in many ways operate similarly (they both have public service responsibilities) and typically conform with the traditional news article template. Additionally, polarisation is estimated based on differences in coverage of the same news stories. Therefore, we need to establish that the sections of the outlets chosen overlap, and that they share roughly the same criteria for what constitutes news stories.

Both DR and TV2 are central in daily news coverage in Denmark as large media houses with a broad reader base (Slots- og Kulturstyrelsen, 2017; Journalisten, 2021; Danmarks Radio,

[4] e.g. www.infomedia.dk where the article image often isn't available.

[5] While technically possible a polarisation measure with a higher dimensionality than 2 bears little intuitive meaning.

2025). There exist long-held popular prejudices about the political leaning of the two media houses, stating a more left-leaning political profile for DR than for TV2 (Journalisten, 2010; Berlingske, 2010). This combination hints that, while the outlets appear somewhat similar, each media has a distinct profile and editorial line, which they follow. In the following, we outline how we collect news articles from DR and TV2, and how we extract features from both text and images.

3.1.1 Data collection

Utilising Internet Archive's Wayback Machine (Internet Archive, 2025), we construct a web scraper that parses historical front pages for both news outlets. It outputs links to all relevant articles for each day in the aforementioned time-horizon. We exclude articles from, e.g., live-blogs, articles without a leading main image, and all articles not categorised as news, such as sports. Appendix 10.2 outlines the full set of criteria determining relevancy. The Internet Archive often has multiple archived sites per day, and we target a request for the archived front-page at 12:00 for each date, provided it is available. If it is not, the request is redirected to a more recent archive. For some dates, no archive is available, in which case the entire date request is redirected to the closest archive available in order of recency.

Naturally, some articles are present on the front page across multiple dates and the redirection of pull requests, too, introduces some duplicated links to articles. Both types of duplicates are removed, keeping only the first available published article link.

Next, we build another web scraper, this time aimed directly at DR.dk and TV2.dk. For each historical article link, we issue a pull request to the corresponding outlet web-page. Compiling all these across time and media results in a large dataset, which, for each article, contains metadata on the outlet, section, headline, date of publication, full article text, and a unique image ID. The sections the outlets apply vary over time and overlap partially across media. We align and categorise the sections in accordance with Appendix 10.3. The image ID refers to the leading article image, which we then rescale and reduce in quality before downloading. This procedure returns data on a total of 168.197 unique articles with corresponding images.

3.1.2 Feature extraction

Along with the metadata on each article, we extract multiple features from the main image and the article text, with the aim of creating a set of tokens that represent the text and image content for each article. Here, we distinguish between the list of extracted raw features and the resulting set of tokens. We employ the following three specific text- and image- analysis tools to extract the features:

1. Scene tagging
2. Face recognition
3. Keyword extraction

All features are compiled and formatted to construct a dataset comprising the full set of tokens from all articles. All multi-word features extracted from either the image- or text- analysis, such as keywords from an article or an object in an image, are treated as both one collective token (bi- or trigrams) and as separate tokens (unigrams) for each contributing word.

Scene tagging: In order to analyse how images contribute to polarisation, we need to describe what the images contain. Until recently, the only way to describe image content was to manually encode image content. However, starting with advances in classical machine learning (ML), computers have been trained on hand-coded training data sets of, e.g., cars or humans, and then, out-of-sample, have been able to recognise these objects (Redmon et al., 2016). This is the simplest version of computer vision (CV) at our disposal. This procedure requires large training sets, and the recognition/detection of the objects is limited to objects encoded in the training datasets. The images we have at our disposal can contain virtually anything, and our ex ante expectation of their exact contents is very likely uninformative of their actual contents. In combination this rules out the use of classical machine learning approaches to describing the image contents.

A technologically more demanding approach is the combination of classical object detection, which detects changes in, e.g., depth in an image to determine boundary boxes of objects without being able to label them, and new advances in convolutional neural networks (CNNs). The approach allows for detecting objects in an image and, at some probability determining which object is the closest to some label option. For instance, answering - given the expectation that an image contains a car - where is that car located in the image? (Heidari, 2024). This bypasses the need for manually encoding a labelled training dataset, as the model can leverage pre-trained knowledge. However, in our testing, the results have not been scalable or reliable to a degree that we could use them to describe broader image content. Additionally, it still requires input labels, as the model can only attempt to identify objects it has been explicitly told to look for.

With the development of large language models (LLMs), the field of computer vision is rapidly expanding. At the very vanguard of object detection and scene recognition lies the field of unsupervised scene tagging (Radford et al., 2021; OpenAI, 2021). The most recent advances converge on combining object detection from a classic convolutional neural network with computer vision capabilities with large language models (LMMs), predominantly generative pre-trained transformers (GPTs). This generally encompasses algorithms that, with zero-shot learning, detects objects and identifies them, thus eliminating the need for both training datasets and specifying a label range. One of the most advanced open-source CV models is LLaVa:7B, which we access through the Ollama module (Liu et al., 2024; Ollama, 2024; Liu et al., 2023). We build this zero-shot CV-model into a programme which constitutes the backbone of the image analysis component in the token generation. Running locally, our programme processes all images across media and (1.) identifies the placement within the image of the objects it contains, (2.) recognises internally these objects and the scenes in which they appear, and (3.) parses this information to its generative component (Zheng et al., 2023), which assigns labels, judges their relative importance, and aligns output. Computationally, this process is extremely demanding, with processing time for a single year taking up to three days, thus, for all images across media, approaching 700 hours in total run time[6]. The scalability and applicability of our programme across different image types and the accuracy of the resulting tags, however, outweigh the drawbacks posed by the vast process time. The programme finally returns a string of 10 features per image. This is illustrated by figure 1., in which an example image and the corresponding extracted image features are presented.

[6]We extend our gratitude to Københavns Universitet and KUB-datalab for providing access to state-of-the-art computational resources and hardware, without which the process time surely would have been prohibitively long.

Scene tags post-processing: To ensure compatible tokens, we exclude image features consisting of more than three words per tag. For a limited number of images, we are not able to extract the features, for which we leave the image features blank. We align and format the features and translate them from English to Danish in order to allow for possible overlaps between tokens from images and from text. This is done using software, which is capable of translating simple text at high speed (Finlay and Argos Translate, 2025). Lastly, we divide all multi-word features into separate tokens, as well as keeping the original feature as a tri- or bigram token. This method allows us to keep information contingent on all input words in a feature, such as *"civil unrest"*, while also allowing a mapping to the individual words *"civil"* and *"unrest"*. Finally, we apply a lemmatizer to strip the tokens to their roots, ensuring that we can map tokens when grammar might differ (Lind, 2021).

Scene tags:

Police	Protest	Riot	Emergency	Law enforcement
Crowd control	Public safety	Nighttime	Urban Environment	Civil unrest

Figure 1: Example of object detection and scene tagging

Face recognition and facial attributes: While our approach in unsupervised scene tagging is at the very forefront of the computer vision field, we adopt a more classical machine learning approach in combination with CNNs to quantify characteristics of the faces of the people in the images. Using advanced open-source components, we construct a specific programme to analyse images containing faces. Firstly, we employ the single-shot multilevel face detection model retinaface through the deepface module, which, with very high precision, locates faces in the images (Deng et al., 2019; Serengil, 2024). This is used as an alignment component to direct the actual face recognition model, facenet512, onto the faces (Schroff et al., 2015). The face recognition model is partly zero-shot in the sense that it is pre-trained to identify facial attributes. Using a pre-trained model allows for greater accuracy and again bypasses the need for manually creating training datasets to "learn" to distinguish, e.g., ethnicity. Once directed onto the faces located by the alignment component, the face recognition model extracts probabilities for ethnicity, gender, age, and emotion expressed in the face. We rely on the very high accuracy rates in external tests of the module to support the use of facenet512[7].

We go on to enhance the face recognition component using a more classical machine learning approach. We construct a database of multiple portraits of relevant influential people to identify them in the article images. The training data includes all Danish ministers and party leaders within the time-horizon 01/01/2015-31/12/2024. The database also includes Danish

[7]External validity of facenet512 measured at 98.4 pct. alone, cf. Serengil (2024), and in combination with retinaface surpassing human accuracy out-of-sample, cf. Serengil and Ozpinar (2024).

royalty, the US presidents, the presidential candidates along with vice presidents and vice presidential candidates from US presidential elections within the period. Using the alignment component, we identify the exact location of the faces in the portraits, crop the faces to exclude irrelevant noise, and finally train the face recognition model on the training dataset. This allows us to identify, with some positive distance reflecting the probability of the match, the people in the database out of sample, i.e., in the article images[8]. Having constructed a programme which (1.) locates the exact positioning of the faces, (2.) describes their key characteristics, and (3.) recognises relevant people, we apply it to the entire collection of images.

The method outlined above prioritises accuracy over efficiency by a large margin. By skipping the alignment component, we could drastically improve runtime, however, ensuring proper alignment prior to recognition returns far superior recognition of facial attributes. Additionally, we extract probabilities for all emotions, ethnicities, and both genders, as well as distance on identified people of interest - all of which increase processing time to approx. 200 hours for all years. The face analysis finds and returns attributes for 225.254 faces in images over the ten-year period. Figure 2., provides an example of the face detection by the alignment component along with the corresponding extracted features and identified individuals.

Facial analysis:

Name:	(1) Marie Bjerre (V)	(2) Jakob Engel-Schmidt (M)	(3) Thomas Danielsen (V)
Distance:	0,38	0,13	0,17
Dominant emotion:	Happy	Happy	Neutral
Gender:	Woman	Man	Man
Age:	27	29	38
Race:	White	White	White
Name:	(4) Lars Løkke Rasmussen (M)	(5) Jakob Ellemann-Jensen (V)	(6) Sophie Løhde (V)
Distance:	0,16	0,08	0,33
Dominant emotion:	Sad	Neutral	Happy
Gender:	Man	Man	Woman
Age:	48	41	33
Race:	White	White	White
Name:	(7) Mette Frederiksen (S)	(8) Morten Bødskov (S)	(9) Mathias Tesfaye (S)
Distance:	0,19	0,24	0,23
Dominant emotion:	Happy	Neutral	Neutral
Gender:	Woman	Man	Man
Age:	34	41	43
Race:	White	White	White

Figure 2: Example of face detection, alignment, and facial analysis

We format the resulting features to obtain regularised output. We retain probabilities on all features to discard cases where prediction is uncertain. In the context of race, the accuracy of the prediction depends on several factors and is susceptible to influence by lighting. Empirically, we notice some mislabelling of the ethnicities *"Middle Eastern"*, *"Indian"*, and *"Asian"*. We proceed with caution when specifically commenting on differences between these features.

[8] We enforce a ceiling on match distance at $0,4$ and limit matching to the face with the highest probability for each face located by the alignment component.

We define a white/non-white feature that is robust to this confounding, which we subsequently refer to as the "robust ethnicity". Moreover, there appears to be some inconsistency in the estimation of gender. For instance, some images correctly recognised as Queen Margrethe are labelled as male. This leads us to suspect a degree of bias in estimating faces as male. This does not affect the later estimate of polarisation, as we filter valid tokens, cf. section 5.5. For a list of all possible features, see Appendix 10.4 and Appendix 10.10. for the list of people we recognise.

Text analysis: While the main challenge in the image analysis is to extract features, the challenge in the text analysis is to find out exactly what to extract. Using text as data, we can rely on more established methods from natural language processing than in the image analysis. The key objective is to represent the information stored in the full articles in some tokens that we can use in the further analysis. This implies a decision on what to keep and what to discard without losing relevant information. To extract key information points from the text, we adapt a multilingual Rapid Automatic Keyword Extraction algorithm (mRAKE) (Piskorski et al., 2021; Rose et al., 2010; Grabovets, 2022). mRAKE - being based on the Rapid Automatic Keyword Extraction algorithm (RAKE) - is a purely statistical tool for locating key information in full body text. It scores candidate keywords based on co-occurrences in the text. As RAKE is domain-independent, we can adapt it to the Danish language by adjusting the set of stopwords manually. In doing so, we can optimise it to the specific domain or kind of language that the article text represents. While the domain of the article text potentially differs from that of general Danish written language, our resulting adjusted mRAKE algorithm is now optimised to the specific language the articles contain. See Appendix 10.5. for details on the use of stopwords.

This summarises the article text into 20 keywords, which can at most consist of two words each. Each word is then fed through a lemmatizer, keeping only the roots of each word (Lind, 2021). As for the image features, the lemmatization ensures that the resulting tokens match in cases where the grammar does not. Additionally, we include checks for certain words if they appear in any context in the article. These words include specific topics of interest, e.g., *"Putin"* and *"China"*, and allow us to test hypotheses on more granular changes in coverage over the years. The full list is available in Appendix 10.6. Finally, we perform a standard sentiment analysis on the full content of the article text, providing a general score indicating how positive/negative the phrasing of the article is. Sentiment analysis - especially in low-resource languages - has drawbacks, as the corpora they rely on are typically narrow. Assuming a fairly general use of the Danish language in the article contents, we employ Sentida to estimate sentiment (Lauridsen et al., 2019a,b)). Sentida is not based on the broadest available Danish sentiment corpus, but notably incorporates negations (Nimb et al., 2022). For each article in each media we calculate a sentiment score, and, on a limited subsample of the articles, we empirically examine the resulting sentiment score against the text, and find a sensible overlap between the two. We refer to Schneidermann and Pedersen (2022) for general validity of sentiment analysis on Danish written language, and note that an even smaller sentiment corpus has previously been used to analyse Danish media (Enevoldsen and Hansen, 2017).

3.1.3 Forming tokens

After formatting the features as outlined above, we obtain tokens reflecting the contents of both the article text and the image accompanying it. We combine tokens from the different extraction methods and generate a dataset of 168.197 articles × 140 token columns. The final token dataset comprises all of the information extracted above, apart from ethnicity probabilities, age, sentiment score, and target checks, and we present a descriptive summary of the dataset in table 1.

Table 1: Descriptive summary of image and text tokens

	DR	TV2	Total
Total tokens	6.350.867	5.603.500	11.954.367
Unique tokens	901.517	737.336	1.369.352
Image tag tokens	1.208.652	1.067.478	2.276.130
Unique image tag tokens	45.213	40.131	62.319
Image tag token as pct. of total	19,03%	19,05%	19,04%
Avg. tokens per article	70,50	71,74	71,07
Avg. image tokens per article	13,42	13,67	13,53
Total tokens post filter	-	-	670.191
Total articles	90.088	78.109	168.197
Articles with image tag tokens	85.887	75.657	161.544
Faces recognized	131.240	94.014	225.254
People recognized	15.368	13.240	28.608
People recognized pct.	11,71%	14,08%	12,70%
Faces per article image	1,46	1,20	1,34

We combine all the features from the face analysis, the sentiment score, and the target checks in a dedicated dataset for descriptive analysis, as the format, e.g. a number or indicator, is incompatible with the subsequent estimation of polarisation. A sample of tokens as they appear in their final form is presented in table 2.

Table 2: Token source: Face analysis Image analysis Text analysis

Token examples:

sad	maskine	nordjylland politi	blond hår	ulykke ske
angry	ammunitionsrydningstjeneste	hospitalsværelse	dansk pas	neutral
pige	identifikationskort	offentlig	påklædning	briller
møde	ulovlig krysantemumbombe	neutral	kriseberedskab	mark
mand	spirituskørsel	patientseng	voldsom soloulykke	kapitalistisk
medicinsk udstyr	pia kjaersgaard (df)	dialog	sterilt miljø	ung menneske
rød jakke	afslappet påklædning	blå øjne	skaldet	smil
sydkorea	medicinsk facilitet	assens	hvid skjorte	park
tilsyneladende	helle thorning-schmidt (s)	angry	præsident	niveau

As presented in table 1., the token dataset contains 11.954.367 tokens, of which 1.369.352 are unique, implying that on average a token appears ~ 9 times across all articles. The filtering process, as described in section 5.5., reduces our set of tokens to 670.191 valid tokens on which we estimate polarisation. Here, the distinction by source is not applicable, as the filtering is contingent on all tokens from both sources. From table 1., we observe quite uniform dispersion between DR and TV2, but a notable difference in the average number of faces per article image. To determine whether these two averages are significantly different, we perform a Welch's two-sample t-test for difference in means with unequal variances (see Appendix 10.7., table 16.). We reject the null hypothesis of no difference in means at a 1 pct. significance level. Extending this finding to a general conclusion, we observe that DR, on average, uses images with people more frequently than TV2, highlighting a difference in editorial line when deciding the lead images to accompany articles.

3.2 Data composition

Having extracted, filtered, and formatted our features and tokens, as outlined above, we now turn our attention to describing the data. In the following, we provide an overview of the structure and composition of the dataset. We illustrate three aspects of the data composition: the article frequency, the distribution between sections, and the post-filtered valid tokens used in the subsequent estimators. This description is essential to understand the comparability between data input from DR and TV2 over time.

Articles and sections: In figure 3., we show the frequency of articles from DR and TV2, respectively, over two-week periods from 2015-2024. The yearly averages of articles are illustrated by horizontal lines. We refer to the two-week periods as biweeks throughout the paper. This reflects that the biweekly intervals offer both high granularity in data while including enough articles to provide consistent insights. In Appendix 10.7., table 17., we summarise the averages for the entire period, by source, including and excluding 2015. On average, we have 646 articles per biweek, where the DR contribution is marginally larger than TV2's. As described in section 3.1., there is a limited number of archives available in Internet Archive's Wayback Machine before 2015-2016. This results in limited extraction from TV2 and DR in 2015, especially at the start of the year. This is evident when we examine figure 3., as the yearly average number of articles for both media is well below the level in the subsequent years.

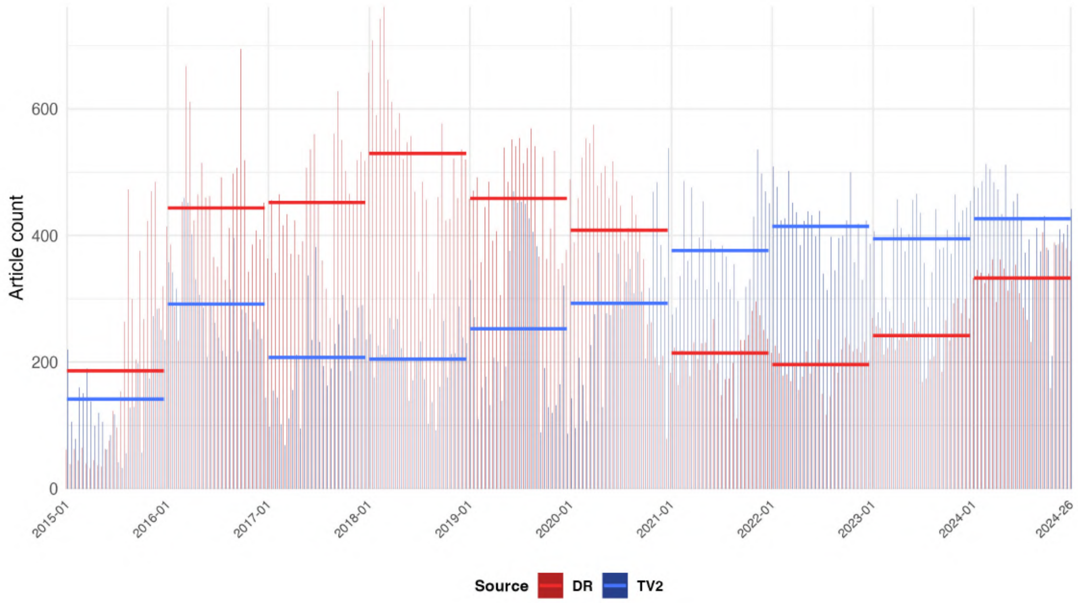
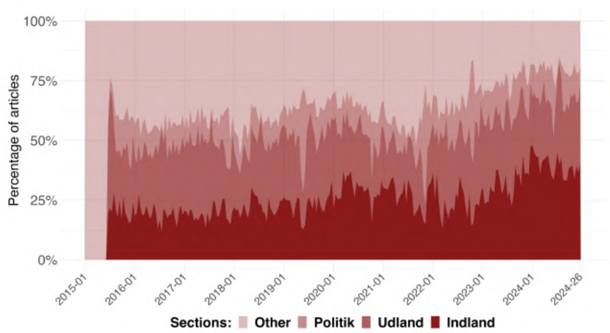


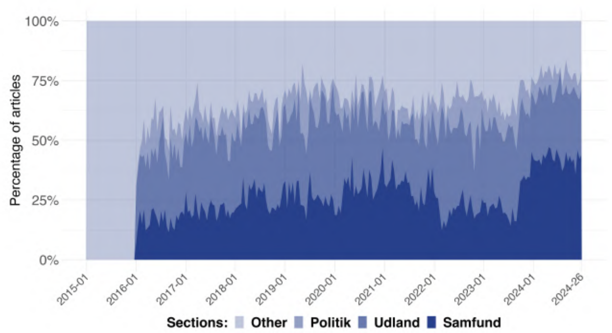
Figure 3: Number of articles per biweek, by source with year average line

To further compare DR and TV2 in terms of compatibility of article input, we examine the section distribution. As outlined in section 3.1.1., the sections vary between media. We have found through manually examining section categorisation across years that the majority of articles fall into the sections *Indland* (DR)/*Samfund* (TV2), *Udland*, or *Politik*. The rest of the sections individually include only a small fraction of the total articles, and some are discontinued or introduced throughout the period. All sections but the three mentioned above are grouped into the category *Other*, which also includes articles with no assigned section. The distribution as a percent of total articles for each given biweek is displayed by media in Figure

4., and the article count by section and source is presented in Appendix 10.7., figure 32. Due to changes in the URL format from DR and TV2 in 2015-2016, we cannot consistently extract sections of articles prior to 2016.



(a) DR



(b) TV2

Figure 4: Comparison of article section distribution

On average, excluding 2015, $26,6$ pct. and $30,2$ pct. of the total number of articles, for DR and TV2 respectively, are published in the section *Udland*. The sections *Indland* (DR)/*Samfund* (TV2) hold $26,2$ pct. and $26,9$ pct. of the number of articles. The two sections are highly comparable, and, as we go forward we refer to *Indland*/*Samfund* as *Indland* collectively. The section *Other* is the largest for both media, containing $35,4$ pct. and $33,6$ pct. This is in contrast to *Politik*, which is the least substantial of the sections excluded from *Other*. It holds $11,8$ pct. and $9,3$ pct. of the articles, respectively. The average section coverage percentages across the period are summarised in Appendix 10.7., table 18. The decomposition into sections lets us examine the comparability between the outlets and the subjects covered. In the later polarisation estimators, it is essential that the material covered by the two media is somewhat comparable for polarisation to signal differences in the bias/editorial lines. Figure 4. indicates a great degree of alignment in the section distribution, both in relative levels and in tendency over time. This suggests that the two media cover similar types of news, providing a solid foundation for the subsequent polarisation analysis.

Tokens: From the articles, we extract tokens according to section 3.1.3. The tokens serve as input to the estimation of polarisation, which is why we investigate their distribution over time. In the ideal case, tokens are evenly distributed across all years. We impose a set of constraints on the token sample, which we outline in 5.5. The process leaves us with a set of valid tokens for each biweek, as presented below in figure 5. As in figure 3., the valid token count in 2015 is well below the average in the other years. It is trivial that when there are few input articles, the valid token set is limited. In all other years, we observe a remarkably stable level of valid tokens, when focusing on the yearly averages. Through the entire period, we have an average of 2.568 valid tokens, increasing to 2.708 when excluding 2015. In the more granular biweekly count of valid tokens, we observe some variance, mimicking that in figure 3. The minimum level of valid tokens is quite high throughout. In the majority of the biweeks, the valid token count is in the interval $(1.750; 3.250)$, approximately ± 750 valid tokens from the average.

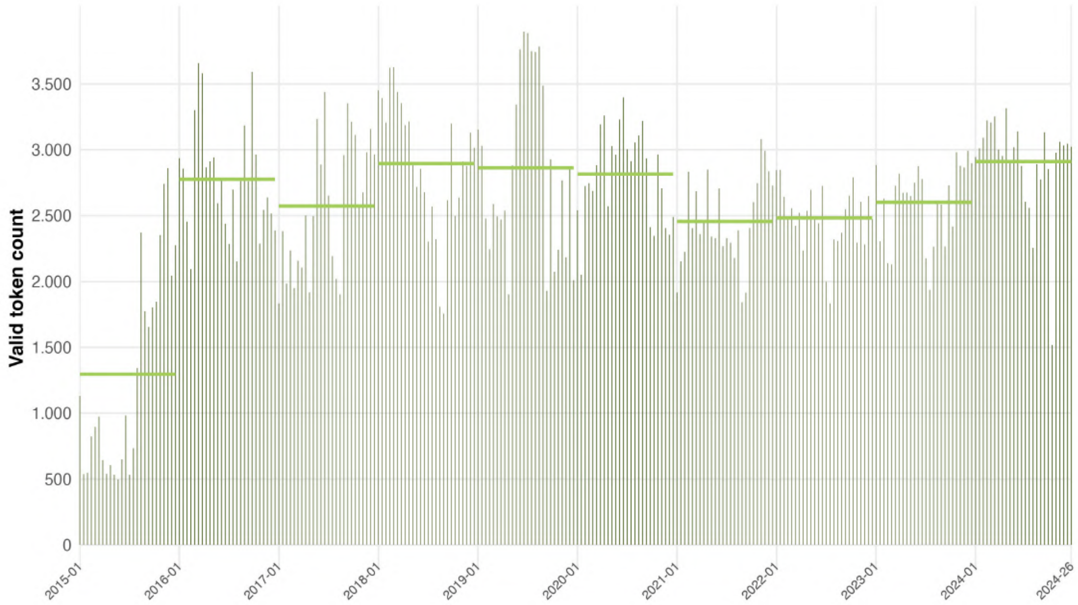


Figure 5: Valid tokens in total

4 Descriptive analysis of media content

In the following section we utilise the rich dataset described in 3. as a basis for an in-depth descriptive analysis of the content in the two media outlets. Throughout this section, we will evaluate WH1, WH2, and WH3. In section 4.1., we investigate how coverage and content differ between media. We then go on to present how tonality in news coverage differs across topics and across the political spectrum in section 4.2.

4.1 Media content and representation

After establishing that the volume and sectional distributions of articles are fairly balanced across media, we explore the news content and representation of groups in our dataset. In section 4.1.1., we examine ethnic, gender, and age representation. These various channels contribute to either support or challenge the hypothesis that the media differ in their visual coverage (WH1). Furthermore, in section 4.1.2., we illustrate the intensity with which the media cover contemporaneous events using target checks for Ukraine, Russia, and Putin as focal points to show how major developments can influence the composition and focus of news coverage. This serves as a foundation for further analysis of the hypothesis that contemporaneous events drive polarisation (WH5).

4.1.1 Representation in media

This section outlines who are visually represented in the news by utilising the facial features detected in lead images across DR and TV2 articles, as described in section 3.1.2. Using the estimated attributes ethnicity, gender, and age, we examine how media representation varies across sections, over time, and between outlets. The descriptive analysis provides insight into whether there exist systematic patterns or disparities in representation in news coverage, which helps to uncover whether the hypothesis that the media differ in their visual coverage (WH1), holds.

By using the estimated ethnicities, we can investigate how the media differ in their coverage of different ethnicities. Immediately, it is clear that the two media outlets are very similar in this regard. This is, e.g., seen in figure 6., where we depict the distribution of the ethnicities in the lead images from the two media, conditioned by section. Both media exhibit very similar patterns in how different ethnicities are represented in lead images. For both media, around 75 pct. of the people detected in the images are white. The largest share of non-white people is found in section *Udland*, where ~ 40 pct. are non-white[9]. The lowest share of non-white is found in the coverage of *Politik*, where less than 20 pct. of faces are non-white. The proportions between the ethnicities that constitute the non-white group are approximately constant across sections, with the largest single ethnicity being Asian. As outlined in section 3.1.2. above, this has a partly methodological explanation. In the coverage of foreign affairs (section *Udland*), the share of Black people for both media is larger than the share of Middle Eastern people, which is the opposite in the domestic coverage (section *Indland*), where Middle Eastern individuals enjoy more frequent representation than Black people. In the domestic coverage, $5-6$ pct. of the faces are Middle Eastern. The actual proportion of the Danish population of Middle Eastern descent is ~ 5 pct., indicating near perfect representation in the domestic

[9] $39,3$ pct. for DR, and $37,4$ pct. for TV2.

coverage of the group[10].

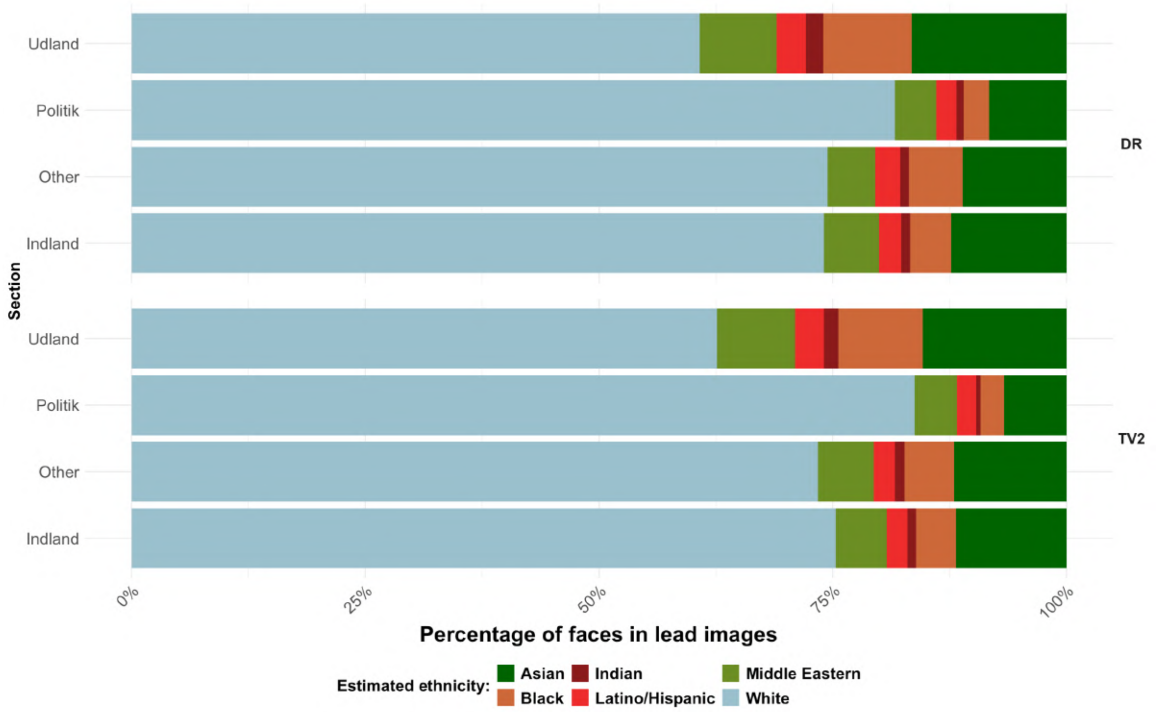


Figure 6: Distribution of ethnicities in news coverage

As an additional insight into our data, we are able to condition TV2 articles on the section *Krimi*, in which representation of Middle Eastern individuals increases to 6.3 pct., indicating a slight over-representation in crime-related coverage. This conditioning is motivated by the fact that immigrants and descendants of MENAPT-origin have the highest representation in crime statistics in Denmark (Udlændinge- og Integrationsministeriet, 2023). This finding aligns with the results of Ash et al. (2022), who document an over-representation of other marginalised groups (Black and Hispanic individuals) in articles concerning crime and poverty. However, the over-representation we document is notably smaller. A corresponding estimate for DR is not available due to differences in section structure between the two media outlets.

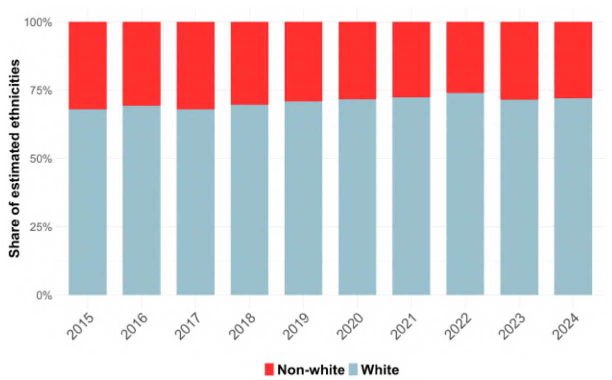


Figure 7: Distribution of robust ethnicities (white/non-white) in news coverage

Having established that the ethnic representation is similar across media, we investigate how it differs over time. In figure 7., there are very limited differences in the robust ethnicity over time. Two similar plots conditioning on the source are shown in Appendix 10.8. There appears to be no trend towards, e.g., greater representation of non-whites over time. However, increasing the granularity and focusing on representation of the Middle Eastern ethnicity uncovers how major news stories have the potential to affect representation. In figure 8., we note that the group constitutes roughly $5-6$ pct. of the total number of people presented in

[10] $5,07$ pct. of immigrants and descendants are of MENAPT-origin, Det Nationale Integrationsbarometer (2025).

the images, and that throughout 2015 the representation level is higher and volatile. This level continues into 2016 and then declines. The years with the largest share of Middle Eastern faces are 2015-2016, which coincide with the migrant crisis that was extensively covered by the two media at the time. Visually, it is clear that while the crisis occupies more and more of the public debate and news coverage in the first half of 2015, the representation follows the same development. Throughout 2016, the representation normalises to the lower level, which is sustained in the following years. This again reflects the developments in the crisis, where the EU signs a deal with Türkiye in 2016, drastically reducing both the inflow of migrants and the coverage of the crisis (Nedergaard, 2025). In the most recent years, the group has experienced a small increase in representation, possibly reflecting the increased share of the group within the total population (Det Nationale Integrationsbarometer, 2025).

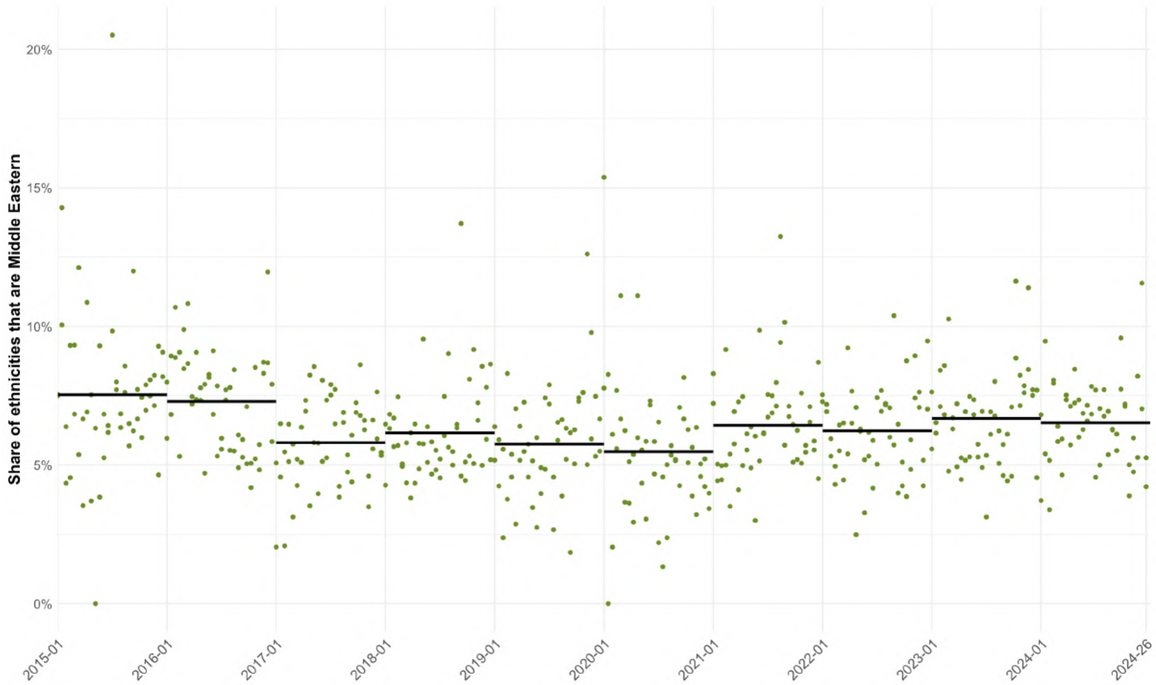


Figure 8: Representation of Middle Eastern people in images across time

We proceed to investigate the representation of genders in the images. In figure 9., we plot the share of female faces across sections for DR and TV2 over the ten-year period, along with the average share across all sections. Firstly, we note a low representation of female faces across all sections - on average $16,0$ and $15,6$ pct. for DR and TV2, respectively. The mean representation of females in DR images is significantly higher than in TV2, despite the limited difference in shares. As outlined in section 3.1.2., the face analysis has a potential bias towards assigning male over female when estimating gender. Quantifying the size of the bias is labour-intensive and not within the scope of our work, so while we consider it highly unlikely, we cannot rule out that the actual representation is 50/50. We base this in part on Ash et al. (2022) that also establishes a pronounced under-representation of female faces in news coverage, with female share of faces estimated at only a slightly higher level than we find.

Secondly, there is a noticeable variation in the female representation across sections. Representation is by far the lowest in the section *Udland*, at a level of 11 pct., while at its highest in *Politik*, nearing 20 pct. for both outlets. If there is a bias in favour of assigning male over female, it should apply equally across sections and thus yield consistent - but systematically skewed - shares, assuming the underlying input is similar. However, this assumption may not hold, as we observe the highest share of white individuals in *Politik* and the lowest in *Udland*. If the estimation of gender is sensitive to ethnicity, the variation in ethnic composition between sections could distort the difference in representation presented in figure 9. Again, we relate the differences in representation to Ash et al. (2022), who also find a greater representation of men in foreign affairs coverage compared to politics.

Thirdly, figure 9. reveals differences in representation within sections for the two media. We cannot ascribe these differences to differences in, e.g., shares of ethnicities, cf. figure 6. There are significant differences between the share of females in the images in DR and TV2 in the sections *Indland* and *Other*, indicating a more equal representation of genders in the domestic coverage in TV2 than in DR. Conversely, the section *Other* has a more equal representation for DR than for TV2.

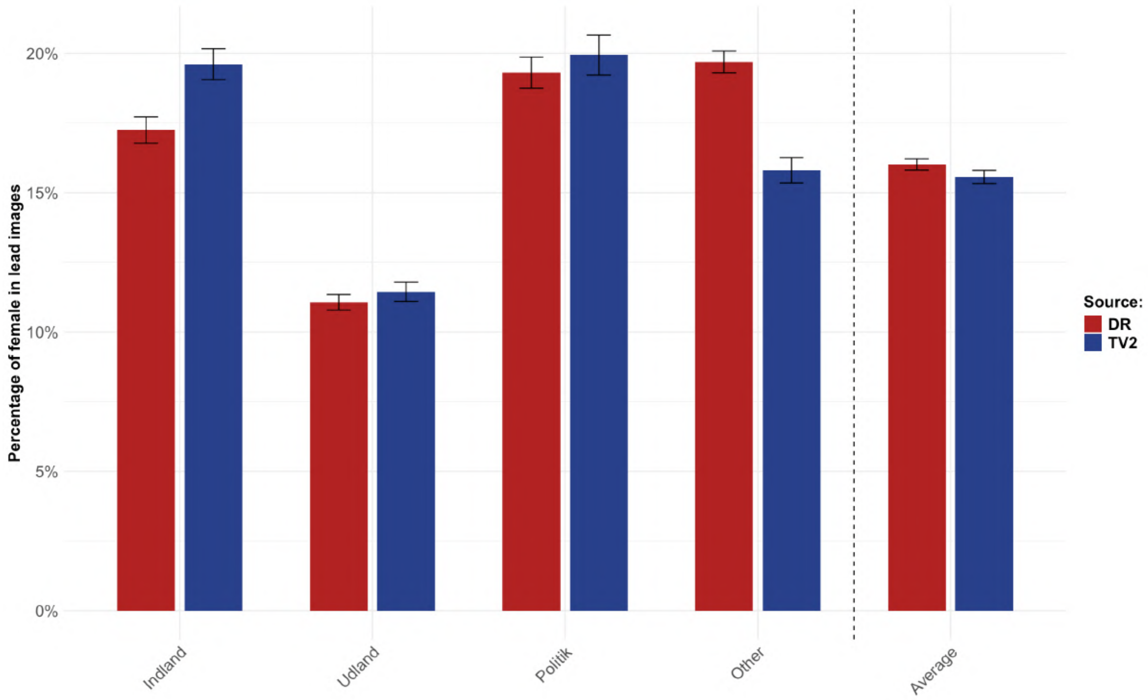


Figure 9: Representation of females in images across time

Over time, we see a trend towards greater equality in representation between the genders, cf. figure 10. From initial low female representation in images of both media at $13 - 14$ pct., the share of female faces in DR images increases each year until 2021. In contrast, TV2 reaches its lowest female representation in 2018, followed by increases of 2 percentage points in both 2018 and 2019. The development stagnates for TV2 in the subsequent years, and for DR declines for two consecutive years, resulting in similar shares of representation at 17 pct. for both media in 2024.

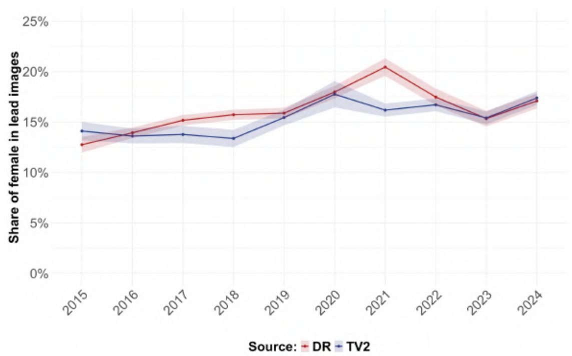


Figure 10: Representation of female faces in images across time

Likewise, we note a positive trend in the average age of the faces in the images. This trend applies to both media. As with the other features of the face analysis, the age estimation appears to be biased, possibly underestimating the true age. As a result, the average estimated ages are relatively low. The average age across sections in 2015 is $33,0$ and $33,9$ for DR and TV2, respectively. This increases over the ten years to $34,4$ and $35,1$, respectively. There are significant differences in the average age of the people that appear in the images of the two media. In all years but 2021, the average age is significantly higher in TV2 images than in DR. Similarly, the average age in TV2 images is significantly higher than the average age in DR images within all sections, cf. table 3. The section *Politik* has the highest average age, while *Indland* has the lowest, for both media.

Table 3: Average age of faces by section and media

Section	**DR**	**TV2**
Average	33,8 (33,7–33,8)	34,5 (34,4–34,5)
Indland	33,1 (33,0–33,2)	33,6 (33,5–33,8)
Udland	33,6 (33,5–33,6)	34,6 (34,5–34,7)
Politik	35,4 (35,3–35,5)	36,6 (36,5–36,8)
Other	33,7 (33,7–33,8)	34,0 (33,9–34,1)

Note: 95 *pct. confidence interval in parentheses*

In conclusion, we have examined different channels through which visual content can differ between DR and TV2. We find no evidence of differences between the two when we focus on the representation of ethnicities. However, when we turn to the representation of gender and age, we find significant differences between DR and TV2, which we view as evidence in support of WH1: that the media differ in their visual coverage. The differences are most pronounced in the sections *Indland* and *Other* in terms of gender representation, while age differences are evident across all sections and most years.

4.1.2 Mentions of Ukraine, etc., in news coverage

To get an indication of how sensitive the general news stream is to specific events, we focus on how extensively Russia's attack on Ukraine has been covered, both around the attack in February of 2022 and in the following years. This is a preliminary examination of how a contemporaneous event can affect the news stream, which motivates a later analysis of the hypothesis that contemporaneous events drive polarisation (WH5).

In figure 11., we illustrate the share of articles from the two media that contain references to Russia, Putin, or Ukraine in relation to the total number of articles within a given biweek. There are more mentions of Russia prior to 2022 than mentions of both Putin and Ukraine; however, around the invasion (biweek 4 of 2022), a record number of articles mention Ukraine, with mentions in around three out of every four articles. This shift clearly showcases the immense focus on the conflict in the start of 2022, and though the mentions quickly drop again, they remain high throughout 2022 and well into 2023, with up to 20 pct. of articles across both media mentioning Ukraine. Similarly, Russia and Putin are mentioned extensively at the beginning of the war, though the levels are lower than for Ukraine. There appears to be no distinct difference in the coverage between the two outlets. TV2 has a higher total number of mentions, reflecting the larger article count of the media in the most recent years, cf. figure 3., but in relation to the total number of articles, the media appear to share the same editorial focus on the conflict. Relating figure 11. to figure 4., the absence of a clear spike in foreign coverage after 2022 indicates that the substantial coverage of the Russia/Ukraine war has displaced other news stories in the section *Udland*, plausibly narrowing the diversity of the foreign coverage in both media.

As the share of articles mentioning Ukraine exceeds the share of articles in the foreign section, the invasion and its repercussions are covered in the other sections too. Thus, rather than affecting the overall distribution of editorial priority between sections, the war appears to have reshaped the content within especially the foreign news section (*Udland*), at least on short term. Following the onset of the war, an increasing number of news stories in the domestic sections - *Indland*, *Politik*, and *Other* - includes mentions of Ukraine/Russia, e.g., relating to military expenditures or refugees. As no media covers Ukraine/Russia more intensively than the other, neither at the start of the war nor in the following years, figure 11., suggests that the subsequent domestic coverage too has held an equally high editorial priority for both media.

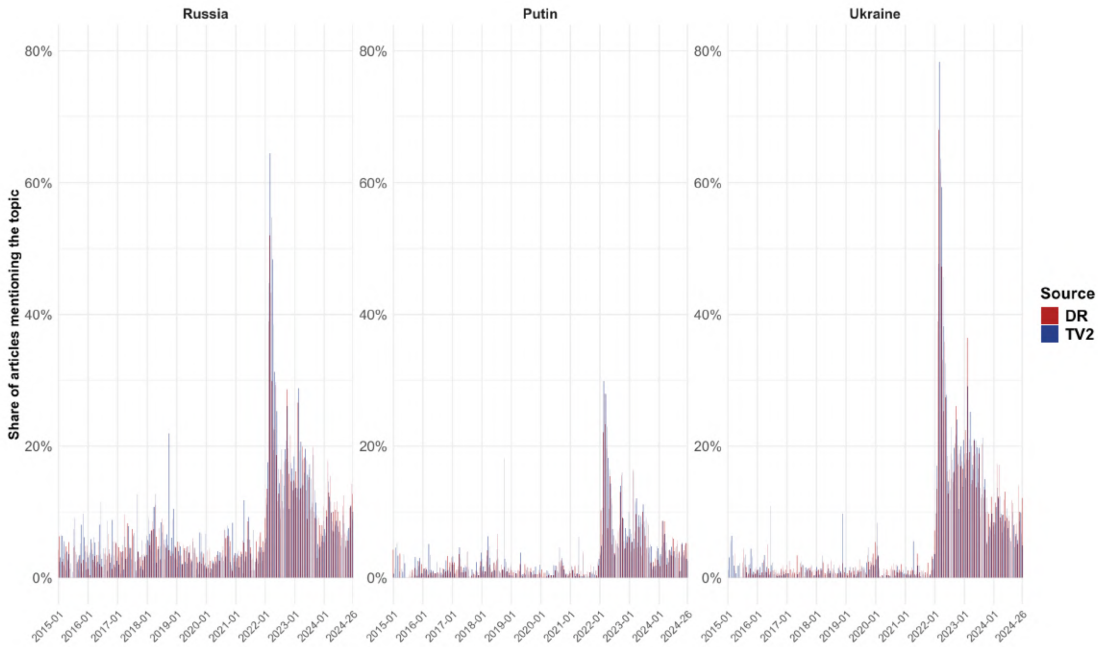


Figure 11: Frequency of the topics Russia, Putin, or Ukraine as share of news

4.2 Sentiment scores of articles

Previously, we have described and illustrated the tendencies, similarities, and differences in article frequency, section composition, facial attributes, and mentions frequency of specific topics, all in an effort to gain an intuitive understanding of what the vast dataset contains. The data largely support the working hypotheses that the media differ in their visual coverage (WH1). Now, we turn our attention to the sentiment score, which quantifies how positive/negative the wording of a text is. We use this measure to examine WH2: that the media cover similar topics in different tones, and WH3: that the media have distinct political profiles. In the following, we use tone, attitude, and sentiment interchangeably, acknowledging that tone and attitude usually have a broader semantic scope than sentiment.

The starting point is a broad outlook, examining overall sentiment for DR and TV2 articles, respectively, from 2015 to 2024. From here, we narrow our focus to specific topics to assess if sentiment differs depending on what type of news material is conveyed. We define a number of subsets of the full dataset, limited to articles with certain news content. This approach allows us to examine whether WH2 holds. Next, we turn our attention to political actors and parties, investigating WH3, which is partly motivated by the perception that DR and TV2 exert some political leaning (Journalisten, 2010). We investigate if sentiment differs in coverage of firstly American politics, and then between red and blue blocs in Danish politics. Extending the focus on Danish politics, we examine if sentiment varies between governing parties, main opposition parties, and fringe parties. Finally, we focus on individual Danish politicians, specifically former and present prime ministers.

4.2.1 General sentiment

The sentiment analysis reveals that DR articles are almost exclusively written in a more positive tone than their TV2 counterparts. In fact, for all but seven biweeks, the average sentiment of DR articles lies above that of the TV2 articles. Three of the seven two-week periods are in 2015, a year with limited data availability. This is illustrated in figure 12., where a 10^{th} order polynomial, matching the number of years, is fitted to the data-series. We refrain from commenting on the absolute scale of the sentiment scores, as the range is uninformative without a relevant anchor-point.

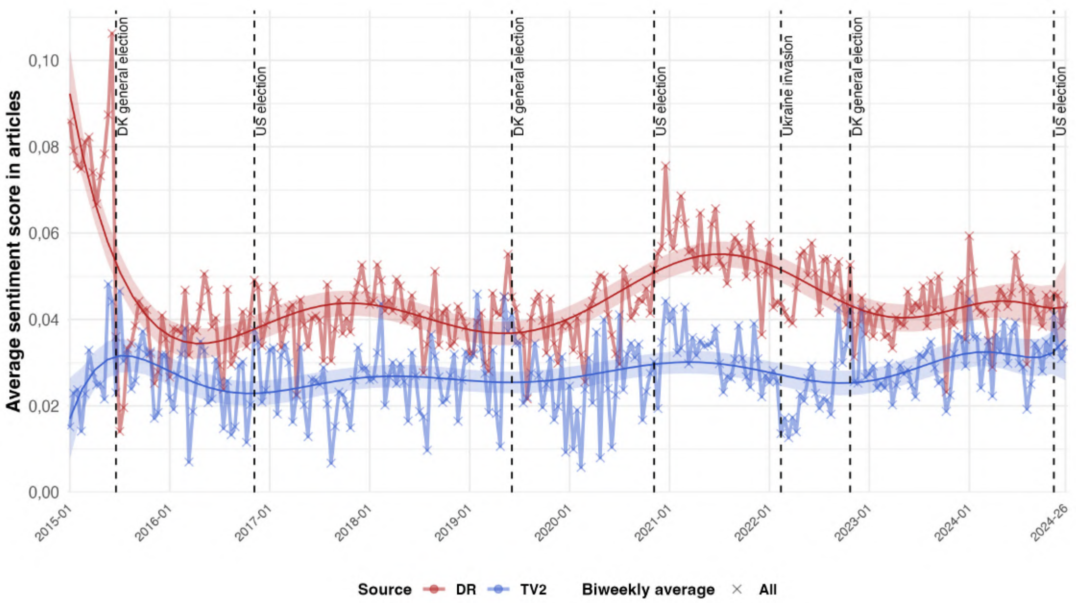


Figure 12: Sentiment score by media, including 10^{th} degree polynomial

There is no clear trend in either of the sentiment time-series shown in figure 12. Further, we observe virtually no overlap of the confidence intervals, except for the last few biweeks, which mechanically have broader intervals, as they are at the end of the time-series. This suggests that average sentiment scores from DR and TV2, when approximated by a 10^{th} degree polynomial, are significantly different from one another. The sentiment in both time-series fluctuates, though it appears more stable for TV2 than for DR. This results in expansions and contractions in the gap between the two polynomials, which exhibits a somewhat cyclical pattern. The timing of the cycles is a subject of special interest, as it could shed light on drivers of differences between DR and TV2. In figure 12., vertical lines mark the time of the general elections in Denmark and the US, and the invasion of Ukraine, which might be associated with the cycles. The contractions in the sentiment gap appear to correlate with the timing of the Danish general elections. This suggests that the sentiment gap possibly follows an electoral cycle. We formally investigate this in section 6.3.

Lastly, we note a feature in the sentiment time-series which is recurrent throughout the analysis for sentiment and polarisation alike: that there exists a period of moderation in the last three years in the time horizon. The volatility both time-series exhibit prior to 2022 is replaced by smaller biweek-to-biweek changes, and the sentiment gap is notably limited. We present a rigorous investigation of this change in section 6.4.

4.2.2 Topic-specific sentiment

Utilising the target checks, as described in 3.1.2., we identify articles containing certain words or phrases/topics. This allows us to compare the within-topic sentiment to the average sentiment for each media. In figures 13., 14., and 15., we isolate the sentiment for articles within six topics. The topics chosen are prevalent in the general media coverage across time and are relevant to the political debate in Denmark. We have chosen topics that are likely polarising or described in a very positive or negative tone. They illustrate if DR and TV2 change attitude in covering specific topics in relation to the average sentiment, cf. figure 12. Moreover, they highlight when there are differences in how the two media cover a topic, e.g. when sentiment changes for a topic in only one media. With this approach, we are able to assess the validity of WH2: that the media cover similar topics in different tones.

The first pair of topics we choose to investigate is *"immigration"* and *"islam"*. It is a subject which provokes conflicting viewpoints across the political and public sphere (New York Times, The Daily, 2025)), and immigration in Denmark is broadly covered in academia, politics, and news alike (Bailey-Morley and Kumar, 2022). Cengiz and Eklund Karlsson (2021) compare the narrative on immigration across the Scandinavian countries. They argue that in Denmark, religion and culture are used to explain most issues related to immigration. Further, they state that there is a very limited focus on multiculturalism and racism in Danish news compared to the news coverage in Sweden. Their analysis finds that of the non-Western immigrants, Muslims are covered most extensively in news. We choose to compare articles mentioning *"islam"* or *"muslim"* to articles mentioning *"immigration"*, *"immigrants"*, *"asylum seekers"* or *"foreigners"* on the basis that religious background might affect the language used in the news coverage.

The second topic is climate, where we examine one subset of articles mentioning *"climate"* and another subset mentioning the topic *"climate change"*. The topics receive increasing attention in the news, and like the previous topics they are widely debated. This is exemplified by Journalisten (2024), stating that climate is gaining attention in the press, even though it is a tough topic to cover. This topic has also received increasing political attention, exemplified by the political party *Alternativet* securing representation in the general election in 2015, and the following general election in 2019 being labelled as the "climate election" (Altinget, 2021; Hansen and Stubager, 2021). The expectation as to whether this topic is covered in a more negative or positive tone than the average coverage is ambiguous. Denmark is a self proclaimed "first mover" in relation to tackling climate change and reducing emissions, which would lead to a more positive tone (DR Nyheder, 2023). On the other hand, climate change induces severe natural disasters, which we assume are covered with a more negative tone than the average.

The last topic we investigate is how the two major authoritarian leaders and countries, China with Xi Jinping, and Russia with Putin, are covered. This decision is largely motivated by the increasing geopolitical tensions following Russia's invasion of first Crimea, and later Ukraine as a whole. We compare it against the coverage of China, as it is viewed as the major challenger of the western dominated world order (Pew Research Center, 2024; DR, 2025b). At the same time, it can help put how positive or negative the tone towards Russia and Putin is into perspective.

Using the mentions of certain words to define groups of articles that cover a certain topic has practical limitations, as the target check does not account for the context in which the word appears. An article mentioning *"climate"*, for instance, could describe the political climate of a government coalition. This limits our conclusions on the topics to articles mentioning the specific target words, acknowledging that this is not a perfect selection process. This selection process may result in some noise in the subsets, however, it does not appear to pose a serious concern in this analysis [11].

In the graphs below, the bold lines and corresponding confidence interval represent the topic-specific subset, while the faded lines depict the average sentiment across all articles, cf. figure 12. They serve as anchor-points that we can compare the subset sentiment against. The subset biweekly average sentiment is depicted by triangles, while the average sentiment scores are represented by crosses. Each time-series is approximated by a 10^{th} degree polynomial. We describe and interpret the six topics motivated above pair-wisely in the following.

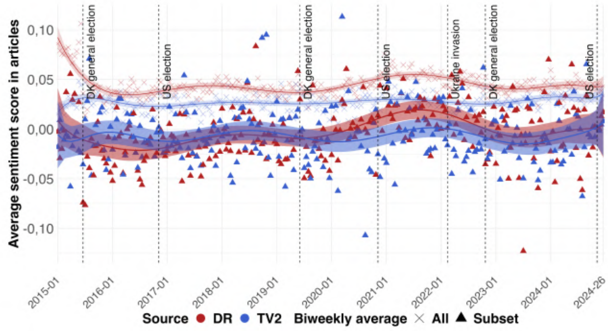
In figure 13., panel (a) shows the sentiment for the subset of articles where *"muslim"* or *"islam"* is mentioned. The polynomials of the subset sentiment are below and significantly different from that of the average sentiment across all articles for most biweeks. Thus, articles where *"muslim"* or *"islam"* is mentioned generally use more negatively worded language, compared to the average article. Further, we observe that the average subset sentiments from DR and TV2 overlap for the majority of the period. This indicates that the two outlets share the same tone when covering material with references to *"muslim"* or *"islam"*. As we observed earlier, the average coverage from DR has a more positive tone than TV2. When the sentiment level for the subset coincides for the two outlets, it indicates that the deviation from an outlet's overall average is greater for DR than for TV2. We compare this to panel (b) of figure 13., where the subset contains articles mentioning either *"asylum seekers"*, *"immigrant"*, *"immigration"*, or *"foreigners"*. For almost all biweeks, we observe no significant difference between the subset sentiment and the full-sample average. If we consider the terms as substitutes that can be used interchangeably, the comparison leads to the partial conclusion that when covering neutral material on e.g., ethnic and minority groups, the media predominantly use terms like *"foreigners"* or *"immigrants"*, while terms like *"islam"* or *"muslim"* are used predominantly in negatively worded articles. We cannot check the degree to which the terms are substitutes, and quite often they likely are not, nevertheless, the distinct difference in sentiment from the two subsets suggests that there exists an editorial or subconscious decision on which tone to use in relation to the terms. Even if the terms are rarely used interchangeably, the difference supports the findings of Cengiz and Eklund Karlsson (2021), stating that Danish news generally describes Muslims and other non-Western immigrants as a threat to Danish cultural values.

Panel (a) in 14. shows the sentiment of the subset of articles mentioning *"climate"*, while

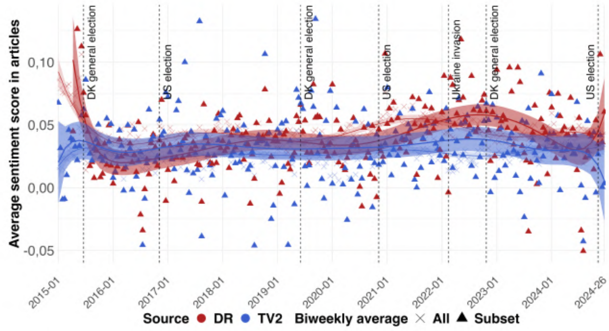
[11] The argument being that noise and sampling error are assumed to be equally distributed between DR and TV2. As the point is to illustrate differences and similarities between media on average, this does not pose a challenge to our analysis.

panel (b) shows the subset with mentions of *"climate change"*. We observe that the attitude in articles mentioning "climate" has only few overlaps with the average sentiments at the start and the end of the time-series, and that it is more positive than the full-sample sentiment for both media. TV2 has a subset sentiment below that of DR in the period 2021-2022, while in all other biweeks, the confidence intervals of the subset polynomials overlap. The deviation between the full-sample and subsample sentiment is greater for TV2 than for DR. As with the previous topic-pair, the subsample sentiments for both media coincide. When examining panel (b) of figure 14., we find that subsample sentiment deviates (positively) from the full-sample average in only a limited number of periods. The reason is somewhat mechanical, given that the subset for *"climate change"* is fully contained within the *"climate"*-subset. The lower number of observations increases variance, resulting in larger confidence intervals of the polynomial. This makes it harder to document statistical differences between the subsample and the full-sample sentiment. Disregarding whether the differences are statistically significant, we take note of the fact that both the coverage of *"climate"* and *"climate change"* are more positively worded than the average article. This is surprising, as we expect articles mentioning *"climate"* and especially *"climate change"* to, at least to some degree, include coverage of natural disasters. Nevertheless, the illustrations in figure 14. indicate that both DR and TV2 choose a more positive line when covering the *"climate"*/*"climate change"* topic-pair, than in their mean coverage.

Moving on to figure 15., where panel (a) shows the subset of articles that mention *"Russia"* or *"Putin"*, and panel (b) depicts the subset with mentions of *"China"* or *"Jinping"*. From panel (a), we observe that the subsample sentiment in DR is much lower than the full-sample sentiment. At the same time, the subsample sentiment in TV2 is indistinguishable from the mean coverage in the period 2015-2022. From 2022 onwards, the tone in the language of TV2 articles mentioning *"Russia"* or *"Putin"* decreases to a level significantly below the full-sample average. This coincides with the Russian invasion of Ukraine in February, 2022. Interestingly, the sentiment of the subsample for DR appears to slightly increase in the period following the invasion. The subsample polynomial even becomes statistically indistinguishable from the full-sample polynomial in the start of 2024. This suggests that DR and TV2 have made distinct editorial choices in what perspective/angle to employ in the coverage of the invasion and the events related to the invasion of Ukraine. One possible explanation for the more positive tone is the increased focus on solidarity among the western countries in support of Ukraine, as opposed to a more negative tone in articles covering the war itself. The development in figure 15., panel (a), supports that editorial differences lead to detectable differences in the coverage. This could explain how the subset sentiments are so different prior to the war even as the intensity of the coverage is similar, cf. figure 11., and could explain the recent divergence in the DR subsample coverage. Contrary to the patterns in figure 15., panel (a), there are no consistent deviations from the full-sample average in panel (b). Here we observe fluctuations in the tone of articles mentioning *"China"* or *"Jinping"* around the full-sample average. There are only a few biweeks where the polynomials are significantly different. This could indicate that the negative deviations in subsample coverage for DR in panel (a) specifically concern *"Russia"*/*"Putin"*, rather than authoritarian regimes in general. Again, this would reflect differences in editorial lines, as we observe no similar difference in the TV2 subsamples.

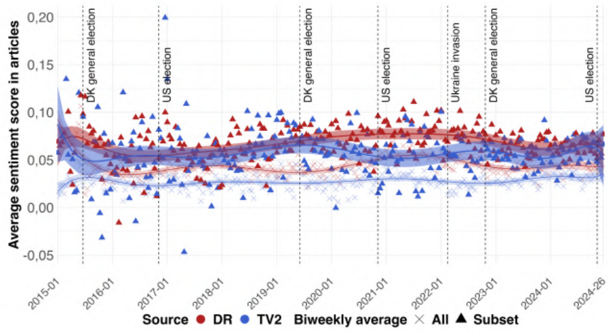


(a) Subset for mentions of *"Muslim"* or *"Islam"*

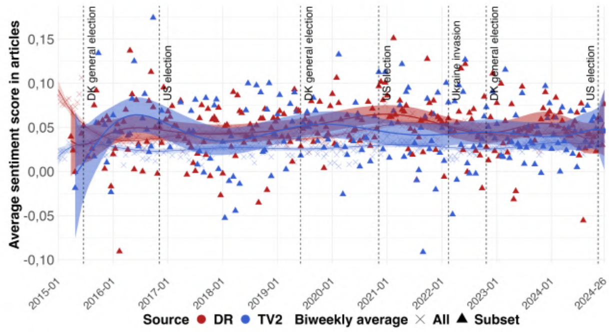


(b) Subset for mentions of *"Asylumn seeker"*, *"Immigrant"*, *"Immigration"* or *"Foreigner"*

Figure 13: Sentiment of topics Islam and Immigrant

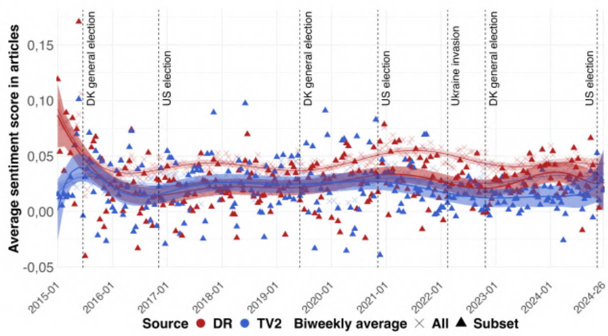


(a) Subset for mentions of *"Climate"*

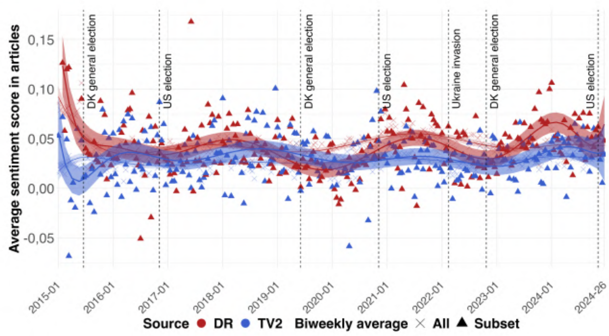


(b) Subset for mentions of *"climate change"*

Figure 14: Comparison of sentiment for topic climate



(a) Subset for mentions of *"Russia"* or *"Putin"*



(b) Subset for mentions of *"China"* or *"Jinping"*

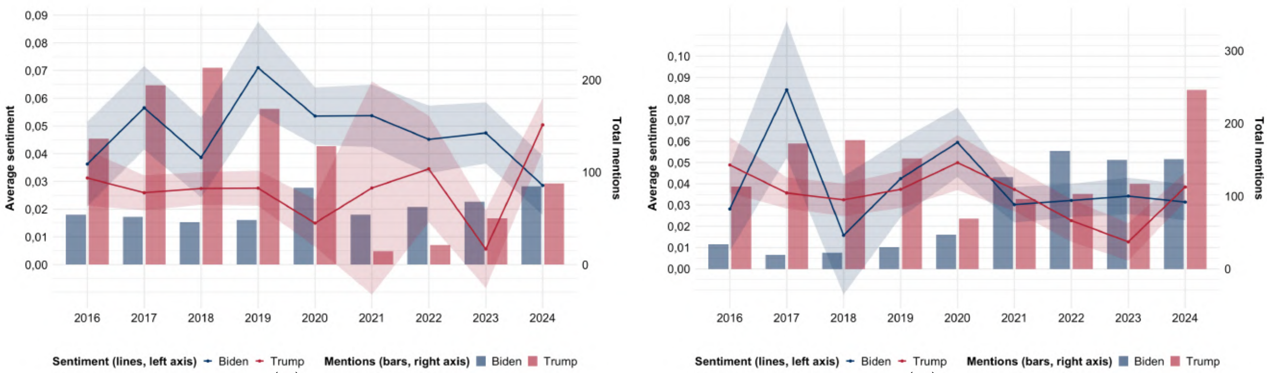
Figure 15: Sentiment of topics Russia/Putin and China/Jinping

To summarise, we find clear differences in the tone between the two media, with sentiment scores varying across topics. The outlets typically deviate in the same direction, and the tone of coverage for a given topic often coincides between them. However, the level differences in sentiment established in section 4.2.1., imply that either DR or TV2 deviate more strongly from the average tone in their articles on certain topics. Further, events like the invasion of Ukraine appear to trigger different approaches in editorial line, which we observe as differences in sentiment. This could contribute to polarisation, which we investigate rigorously in section 6.4. Overall, these findings are consistent with the working hypothesis that the media cover similar topics in different tones (WH2), providing evidence that DR and TV2 deviate differently from their respective average tone in articles mentioning the same topics.

4.2.3 Indications of political bias

We have previously referred to long-held public prejudice against the two media and their political leanings (Berlingske, 2010; Journalisten, 2010). Determining if and how the articles empirically support these prejudices - an aspect of WH3 - is multi-faceted, and we follow several different angles to illustrate the degree to which we can document political bias in the news coverage. Notably, the specific direction of political profile is not our focus, rather, we examine whether the profiles of DR and TV2 are distinct.

The two most dominant political arenas covered in the news are Danish and US politics. Given the animosity with which Danes observe American politics and the public disregard toward especially the Republican presidential candidate (∼ 85 pct. of Danes when asked would vote for the Democratic presidential candidate in the 2024 US presidential election, (Ritzau, 2024)), we start off by investigating how different the media cover the last two American presidents (Donald Trump (Rep) and Joe Biden (Dem)), and if the media differ in their coverage of the two. In figure 16., we plot both mentions and average sentiment for articles in which they appear across time, excluding 2015.



(a) DR (b) TV2

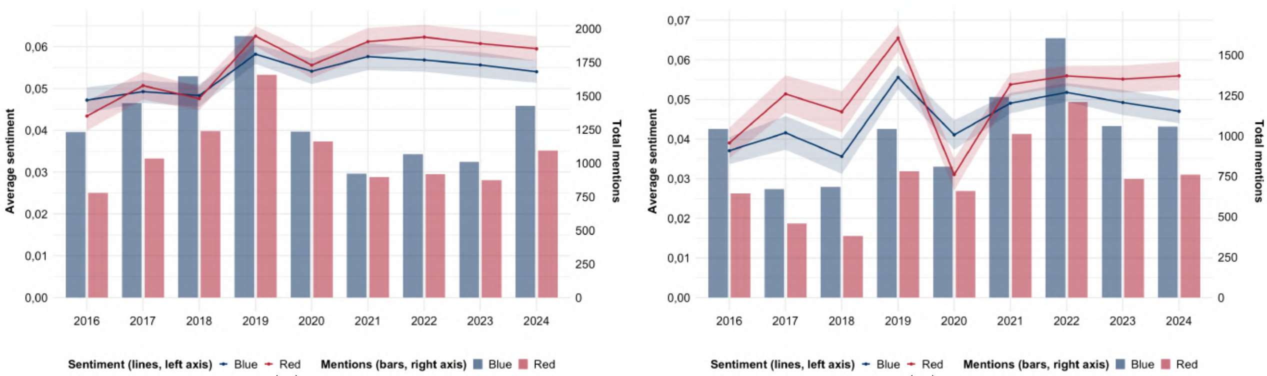
Figure 16: Coverage of US presidents in across media

DR describes Biden more positively than Trump in all but one year, while there is no difference in the TV2 coverage. The effect disappears when we include all presidents, vice presidents, and candidates for both parties over the period. This suggests that, to the extent the coverage of American politics contributes to polarisation, it is driven by differences in how the media portrays the presidents (and to some extent the presidential candidates) rather than by differences in how the two US parties are covered. DR covers the first Trump administration more intensively than the Biden administration, while there is little difference in the mentions between the presidents in TV2, though TV2 covers Trump fairly intensively in the period where he is out of office (2020-2023). Whether the differences reflect conscious differences in the editorial line is unknown, the bottom line remaining that this is a possible channel of polarisation between the media, as we here observe they cover the same material in different ways.

The other political arena, Danish politics, also exhibit differences in the coverage, albeit less pronounced, cf. figure 17. We cannot establish if, and possibly to what degree, this reflects a political leaning of the respective media, but we note that grouping parties by classic blocs on the left/right-spectrum reveals noticeable differences in the coverage between media[12]. The

[12] We group the parties into red/blue blocs as in classical political analysis, disregarding contemporary tendencies to dissolve the classic bloc structure in Danish politics. EL, SF, S and RV form the red bloc and M, V, KF, DF, DD and LA form the blue bloc.

figure reveals that the average sentiments for the two blocs are statistically indistinguishable from each other in most years, though the red bloc is presented significantly more positively than the blue bloc in TV2 in at least 4 years. For TV2, the average sentiment of the red bloc mentions is lower than that of the blue bloc only in 2021. The difference also favours the red bloc in DR, but the difference is much less pronounced, as the confidence bands in average sentiment overlap in all years for DR. For both media in all years, the blue bloc enjoys a more intense coverage, defined by mentions in the article text, which we cannot solely attribute to the higher number of parties in the bloc, given how recent the party *Moderaterne* was formed. The inverse relationship between coverage and sentiment *could* suggest that for the parties, increased visibility can come at the expense of a more critical coverage.



(a) DR (b) TV2

Figure 17: Coverage of Danish party blocs across media

As with the coverage of US politics, we observe differences in how the media covers Danish politics. While determining the precise political leaning of each outlet is complex, our sentiment analysis indicates that coverage varies across the political spectrum, which supports the hypothesis that the media have distinct political profiles (WH3). This presents a viable channel for polarisation, as the media appear to differ in their respective leanings. We address this specifically in section 5.6., by focusing on polarisation in the section *Politik*. In the following, we nuance the aforementioned differences and how they relate to power in general, and the people in power specifically.

4.2.4 Coverage of parties in power

The previous section indicates that there are some differences in the political leanings of the two outlets. To examine how different parties are covered, we distinguish between parties that historically hold the power of government and fringe parties[13]. This is an effort to further assess WH3 and to outline another possible aspect of the political profile of DR and TV2. We define a group of "historically non-governing" parties, which includes parties that have not been in government at any point in the time-period. The group consists of *Enhedslisten* (EL), *Socialistisk Folkeparti* (SF), *Dansk Folkeparti* (DF), and *Danmarksdemokraterne* (DD). The latter is only a part of the group after they are formed in 2022. We go on to define groups of "Currently governing parties" and "Currently non-governing parties", where the group composition varies over time depending on which parties are in government. The parties in government at some point in our time-period are *Socialdemokratiet* (S), *Venstre* (V), *Radikale Venstre* (RV), *Konservative Folkeparti* (KF), *Liberal Alliance* (LA), and *Moderaterne* (M). From

[13]We exclude the party *Alternativet* from the analysis due to ambiguity of the term in the article text.

the start of 2015 to the 28th of June, 2015, the governing parties were (S) and (RV). This government was superseded by a single-party government consisting of only (V) until the 28th of November, 2016. Then (KF) and (LA) formed a government coalition with (V) until the 27th of June, 2019, where the general election was won by the opposing bloc. From here, (S) was the only party in government until the 15th of December, 2022, after which a coalition government was formed, consisting of (S), (V), and (M). This coalition was in government until the end of the time-frame. The parties not in government in a given time-period are part of the group we label "Currently non-governing parties". In figure 18., we have illustrated the yearly average sentiment from each of the three party groups by source. To know whether these averages are significantly different from one another, we have tested if the pairwise differences are significant. This yields three tests per year per source, resulting in 60 tests, all presented in Appendix 10.7., table 19.

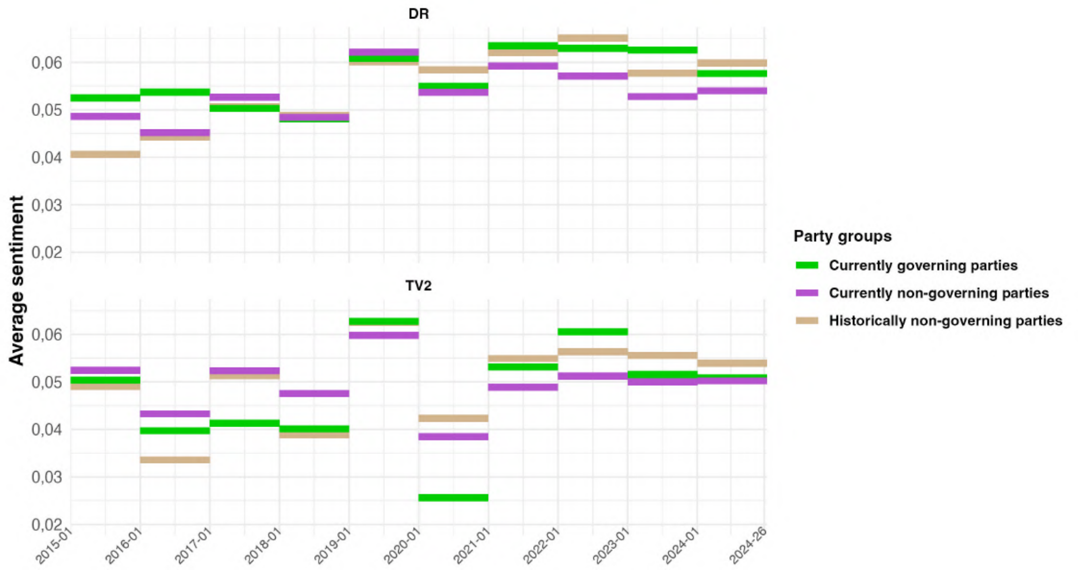


Figure 18: Average yearly sentiment of political party groups in DR and TV2 respectively

From 2021 onwards, the attitude towards the governing parties appears favourable compared to that of the currently non-governing parties. In this time-frame, the average sentiment describing historically non-governing parties is comparable to that of the governing parties. This appears to hold for both TV2 and DR, but the difference is more pronounced for DR. This is in part supported by the t-tests, cf. Appendix 10.7., table 19. Here, we for example establish one year for TV2 and two years for DR, in this period, where the averages for the currently non-governing parties are significantly lower than that of the governing parties.

In the years following COVID-19, the currently non-governing parties are portrayed with a quite low sentiment relative to the other groups. This does not hold for TV2 in 2020, as the currently governing parties are described with an average sentiment at the lowest level of the entire period. This is supported by significant positive differences between currently governing parties and each of the other two groups for TV2, while the differences are insignificant for DR in 2020. We interpret this as a signal of some criticism towards the sitting government from TV2 - which is not observed in DR coverage. This suggests that both political leaning and editorial opposition are possible channels of polarisation.

In 2019, a year with a general election, the party groups are all described with a very positive and similar sentiment in both media, cf. figure 18., supported by no significant differences between the averages. Prior to 2019, TV2 covered the currently non-governing parties with a higher sentiment compared to the two other groups, with significant differences in two out of four years. In the same period, the sentiment in DR articles was approximately the same for the three groups, and there is no clear group that is favoured. Here, we observe some significant differences in 2015 and 2016, but the general direction of differences reverses in the following year, and is insignificant. From 2015 to 2019, DR covers the governing parties relatively more positively, then TV2 does, which further supports that there is differences in the political profiles of the outlets.

We now focus on the average sentiment in articles mentioning the two major parties of power in Denmark, *Socialdemokratiet* (S) and *Venstre* (V), cf. table 4. In all years but 2024 and 2016, DR covers the two parties with similar sentiment. DR has a more favourable coverage of *Socialdemokratiet* (S) in 2024, whereas in 2016, the coverage favoured *Venstre* (V). As the difference between (V)-(S) is negative for TV2 in all years but 2020 and 2021, they generally cover (S) in a more positive tone than that used when covering (V). In fact, the sentiment difference is significantly positive only in 2020 for TV2 - this coincides with the increased editorial opposition during COVID-19 established above. We present an illustration of the yearly averages for all parties in Denmark in Appendix 10.7., 33.

Table 4: Welch's t-tests: *Venstre-Socialdemokratiet* Sentiment Differences

Year	TV2	DR
2015	-0,001	0,007*
	(0,840)	(0,094)
2016	-0,001	0,009***
	(0,771)	(0,004)
2017	-0,010**	0,005*
	(0,038)	(0,079)
2018	-0,002	0,001
	(0,598)	(0,620)
2019	-0,006**	-0,001
	(0,038)	(0,734)
2020	0,013**	0,005
	(0,040)	(0,101)
2021	0,001	0,002
	(0,708)	(0,555)
2022	-0,007***	0,001
	(0,002)	(0,703)
2023	-0,004	-0,002
	(0,231)	(0,423)
2024	-0,013***	-0,007**
	(0,000)	(0,022)

Note: $^{*}p < 0,1$; $^{**}p < 0,05$; $^{***}p < 0,01$

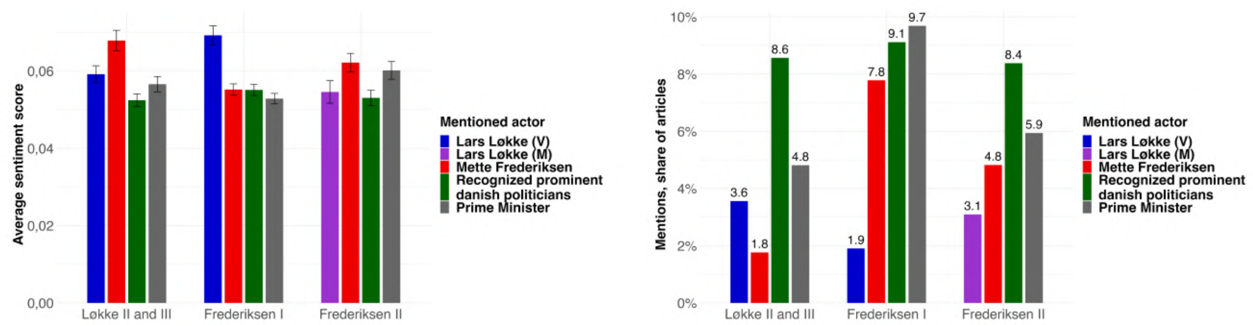
In conclusion, we find that TV2 covers the governing parties significantly more negatively than the other groups in 2020, which coincides with it being the only year where the coverage of (V) is more positive than that of (S). We do not find the same differences for DR, which signals criticism only from TV2 towards the sitting government in that year. Furthermore, DR's coverage is relatively more positive towards the governing parties in the period 2015 to 2019, compared to TV2. Finally, DR has more similar average sentiment when covering (S) and (V) than TV2. These aspects all further support WH3: that the media have distinct political profiles.

4.2.5 Coverage of prime ministers

Recognising people of interest in images allows us to investigate coverage of the prime ministers in the time-frame. This enables an examination of aspects of WH3 beyond political blocs and parties.

In 2015, the observations are too sparse to provide accurate insights, leading us to exclude this year from the following analysis. From 2016 to 2024, Denmark has had two prime ministers: Mette Frederiksen and Lars Løkke Rasmussen. We treat the two as representatives of political power in Denmark over the past 10 years. To illustrate how they are represented in our data, we calculate the average sentiment score used in articles in which they are recognised in the corresponding image, and how frequently they appear as share of total articles, across government periods. We collapse government constellations into single government periods. This results in three government periods: the first with Lars Løkke Rasmussen as prime minister (Løkke II and III, collectively), and the second and third with Mette Frederiksen as prime minister (Mette Frederiksen I and II, separately). We construct two benchmarks to compare against: the first is average sentiment from articles in which other politicians who have held a title of minister or been the leader of their party in our time-frame are recognised in the lead image, and the second sentiment benchmark is based on the articles that mention *"Statsminister"* specifically.

In figure 19, panel (a) and (b) illustrate the sentiments and frequencies described above, respectively. We present corresponding illustrations for each media specifically in Appendix 10.9. Panel (a) reveals that the average sentiment of articles mentioning *"Statsminister"* and in articles for which the lead image includes the sitting prime minister are not significantly different in all three periods. In Løkke II and III and in Frederiksen I, the opposition leader is covered in a more positive tone than the sitting prime minister. The difference in sentiment between the sitting prime minister and the opposition leader during Løkke II and III, and Frederiksen I is mostly driven by TV2, cf. Appendix 10.9. This constitutes the only noticeable difference between the media in the sentiment of their coverage, of the prime ministers.



(a) Average sentiment score (b) Percent of total articles

Figure 19: Representation of politicians and prime ministers

Returning to figure 19., panel (a), we note that in the first and third periods the average sentiment of articles with images in which we recognise other politicians is below the sentiment of articles with lead images depicting the sitting prime minister, while it is on the same level in the second period. Looking at the articles with images in which we recognise Mette Frederiksen during the Frederiksen II government, we note that these articles are more positively worded than articles for which Lars Løkke Rasmussen appears in the image. In the two other government periods, the opposition leader is covered favourably, and the shift indicates a more positive attitude towards the sitting government in the last couple of years. We take note of the fact that this coincides with a period in which there is an absence of a natural opposition leader.

Turning our attention to figure 19., panel (b), showing the coverage of the politicians in relation to the number of articles, we observe that other prominent politicians are mentioned the most in all three periods. This is trivial, as the list of politicians included is long, cf. Appendix 10.10. The sitting prime minister is recognised more than the opposition leader in the first two periods, and in the last, Mette Frederiksen is recognised more than Lars Løkke Rasmussen, who is by then minister of foreign affairs. Across all periods, more articles mention *"Statsminister"* than articles with an image including the prime minister, but the difference is not very pronounced. The prime minister is covered most intensely during the Frederiksen I government, where one in every 13^{th} image includes Mette Frederiksen. To compare, we see that in Løkke II and III and in Frederiksen II, less than every 20^{th} image depicts the prime minister. The Frederiksen I government coincides with COVID-19. The data suggests that this resulted in a lot of media coverage focusing on the top of the sitting government. There are small between-media differences in how intensely the politicians are covered, cf. Appendix 10.9. In combination with the limited difference in sentiment towards the politicians, the data suggests that at least prime ministers do not seem to contribute markedly to polarisation between media. We suspect there might be individual politicians whom we recognise who are more polarising figures, however, investigating this in detail falls outside the time-scope of our work. Hence, this aspect of the media outlets' political profile does not support WH3.

To summarise, in section 4.2.3. we have established differences between the media in how favourable different sides of the political spectrum in both American and Danish politics are covered. When focusing on parties and power dynamics in subsection 4.2.4. we observe indications of editorial opposition to the governing parties for TV2 in 2020, which is not present for DR. Furthermore, TV2 generally covers *Socialdemokratiet* more favourably than *Venstre*, where DR has a more similar coverage of the two parties. Both sections lend empirical support to WH3: that the media have distinct political profiles. We interpret the differences between how the media cover e.g. parties as evidence of these theorised profiles, while taking note of the fact that they vary in strength over time. As section 4.2.5. reveals, we cannot trace these political profiles to the coverage of prime ministers over the time-horizon. Nevertheless, we go forward treating political bias as an empirically viable channel of polarisation.

5 Measures of polarisation

A cornerstone of this paper is to explore quantitatively whether TV2 and DR fundamentally present news differently by employing distinct narrative angles in both text and the accompanying lead image. In doing so, we are able to assess the validity of our main working hypothesis: that there is polarisation in Danish media. In order to uncover systematic differences in the news from the two media outlets, we consider three different estimators of polarisation. We adapt the estimators from the study of party polarisation in congressional speech to polarisation in article content across media (Gentzkow et al., 2019). These estimators can distil high-frequency data into simple measures of difference in tokens between outlets. The resulting measure of either estimator represents the posterior probability that an observer with complete information can accurately determine the source of an article given only a single randomly drawn token.

Notation of the three subsequent estimators aligns to a certain degree. We assign a unique article number, $i \in \{1, 2, ..., 168.197\}$, where the source of a given article is given by $S(i) \in \{DR, TV2\}$. For each article, the observed outcome is the vector $\mathbf{c}_i$, with dimensions $1 \times |\mathbf{J}|$, where $|\mathbf{J}|$ represents the number of unique, valid, tokens present across all articles conditioned on being mentioned at time t. The vector $\mathbf{J}$ contains all valid tokens j at time t. $\mathbf{c}_i$ captures the count of all tokens used in article i. The total count of tokens for article i is given by $m_i = \sum_j c_{i,j}$. Hence, the empirical token frequency from article i is defined as the vector $\hat{\mathbf{q}}_i = \mathbf{c}_i/m_i$. By summing the observed outcome vector $\mathbf{c}_i$ and m_i over source, we obtain $\hat{\mathbf{q}}^S = \sum_{i \in S} \mathbf{c}_i / \sum_{i \in S} m_i$, i.e. the empirical token frequency of source media. From $\hat{\mathbf{q}}^S$ we compute the token-wise probability of appearing in a DR article as opposed to a TV2 article as:

$$\hat{\rho}_j = \frac{\hat{q}_j^{DR}}{\hat{q}_j^{DR} + \hat{q}_j^{TV2}}$$

Stacking $\hat{\rho}_j$ across all tokens, $\mathbf{J}$, yields the vector $\hat{\boldsymbol{\rho}}$. We refer to this measure as the source probability, here of DR, where $(\hat{\boldsymbol{\rho}} - 1)$ would be the source probability of TV2. Intuitively, we interpret this as the relative popularity of tokens between the two media. It can be viewed as an editorial preference, where the choice of image and language signals a distinct media profile.

All the aforementioned terms vary over time, denoted by a subscript t in the following. Intuitively, when the vectors $\hat{\mathbf{q}}_t^{DR}$ and $\hat{\mathbf{q}}_t^{TV2}$ are similar, the articles from the two media outlets are close in terms of linguistic and visual expression. If the vectors are far apart, we interpret it as a higher degree of separation between the media.

5.1 Maximum likelihood estimator

It is straightforward to construct a maximum likelihood estimator (MLE) on the following form:

$$\hat{\pi}_t^{MLE} = \frac{1}{2} \left(\hat{\mathbf{q}}_t^{DR} \cdot \hat{\boldsymbol{\rho}}_t + \hat{\mathbf{q}}_t^{TV2} \cdot (1 - \hat{\boldsymbol{\rho}}_t) \right) \tag{1}$$

In the above, all tokens not used at time t are excluded from $\mathbf{J}$. This reduces the choice set in the maximum likelihood estimator to all tokens used at time t. For each media, the dot product of the token frequency and the source probability yields the media-wise polarisation. Intuitively, if the source probability vector was filled with $0,5$, for all tokens j, the media-wise contribution to the MLE polarisation would be $0,5$, as the sum of $\hat{q}_{t,j}^{DR}$ over $\mathbf{J}$ is 1. It follows that the sum of the two media contributions, divided by two, yields total polarisation, $\hat{\pi}_t^{MLE}$.

When $\hat{\boldsymbol{\rho}}_t$ has a large dispersion, the MLE polarisation will intuitively be large, as some tokens are primarily used by just one of the outlets. As this estimator is based on the empirical tokens derived from a finite sample of text and images from the two sources - and not the entire language - the estimated source probability might not capture the true source probability. In the case where a token is used only once at time t by DR, the source probability $\hat{\rho}_{t,j}$ for that token will be 1, effectively amplifying its role in the resulting polarisation measure. However, the occurrence of this token in DR might not be a signal of strong partisanship towards DR, as it could be down to chance that it is not mentioned in TV2. To clarify the dynamics of this type of estimator, we construct a sandbox example, where each media has two articles and each article has 3 tokens:

$$\mathbf{DR_1} = \begin{bmatrix} b \\ b \\ d \end{bmatrix}, \quad \mathbf{DR_2} = \begin{bmatrix} b \\ c \\ d \end{bmatrix}, \quad \mathbf{TV2_1} = \begin{bmatrix} a \\ a \\ e \end{bmatrix}, \quad \mathbf{TV2_2} = \begin{bmatrix} c \\ c \\ e \end{bmatrix}$$

$$j = \{a, b, c, d, e\}, \quad \hat{\mathbf{q}}^S = \sum_{i \in S} \mathbf{c}_i / \sum_{i \in S} m_i \rightarrow$$

$$\hat{\mathbf{q}}^{DR} = \begin{bmatrix} 0/6 \\ 3/6 \\ 1/6 \\ 2/6 \\ 0/6 \end{bmatrix}, \quad \hat{\mathbf{q}}^{TV2} = \begin{bmatrix} 2/6 \\ 0/6 \\ 2/6 \\ 0/6 \\ 2/6 \end{bmatrix}, \quad \hat{\boldsymbol{\rho}} = \begin{bmatrix} 0 \\ 1 \\ 1/3 \\ 1 \\ 0 \end{bmatrix}, \quad \hat{\mathbf{q}}^{DR} \cdot \hat{\boldsymbol{\rho}} = \frac{32}{36}, \quad \hat{\mathbf{q}}^{TV2} \cdot (1 - \hat{\boldsymbol{\rho}}) = \frac{30}{36}$$

$$\hat{\pi}_t^{MLE} = \frac{1}{2}\left(\frac{32}{36} + \frac{30}{36}\right) = \frac{62}{72} = \frac{31}{36}$$

In the above example, we estimate a polarisation close to 1. The full calculations are in Appendix 10.11. This illustrates that if each media is unique and there is little overlap between them, the polarisation will be large. Further, when one article has the only representation of a specific token, as in case of a in $\mathbf{TV2_1}$, it greatly influences polarisation. This example further illustrates that when i and $\mathbf{J}$ increase, the complexity of vector operations scales linearly, thereby increasing the computational load.

In the MLE, the empirical token frequency of source media $\hat{\mathbf{q}}_t^S$, and the source probability $\hat{\boldsymbol{\rho}}_t$ are both estimated on the same population. This type of estimation introduces a finite sample bias problem, as we outline below.

Finite sample bias: The estimator does not account for the baseline probability of using a token, defined as a token's frequency in "normal" language. As a consequence, some tokens may be labelled as highly partisan to one media outlet, when in fact they are common and non-polarising tokens. Here, common refers to tokens that are generally characterised as part of "normal" language use and should, in principle, be equally distributed across media sources over time. If such common tokens appear in only one media in a given period t, the estimator assigns them as highly partisan, despite their limited polarising effect. In the following, we showcase the origin of this finite sample bias theoretically, using Jensen's inequality (Jensen, 1906) and drawing on Gentzkow et al. (2019). As $\hat{\pi}_t^{MLE}$ is a convex function of $\hat{\mathbf{q}}_t^{DR}$ and $\hat{\mathbf{q}}_t^{TV2}$, Jensen's inequality implies that this estimator has a positive bias in any finite sample. The estimated source probabilities $\hat{\mathbf{q}}_t^{DR}$ and $\hat{\mathbf{q}}_t^{TV2}$ are used, not the true probabilities $\mathbf{q}_t^{DR}$ and $\mathbf{q}_t^{TV2}$, therefore we have an upward bias by Jensen's inequality: $\mathbb{E}[\hat{\pi}_t^{MLE}] \geq \pi_t^{true}$, which arises from inserting estimates into a convex function. To decompose the bias, we compare the expected difference between the estimated value and the true value:

$$\begin{aligned}&\mathbb{E}\left[\frac{1}{2}\left(\hat{\mathbf{q}}_t^{DR}\cdot\hat{\boldsymbol{\rho}}_t+\hat{\mathbf{q}}_t^{TV2}\cdot(1-\hat{\boldsymbol{\rho}}_t)\right)-\frac{1}{2}\left(\mathbf{q}_t^{DR}\cdot\boldsymbol{\rho}_t+\mathbf{q}_t^{TV2}\cdot(1-\boldsymbol{\rho}_t)\right)\right]\\=&\frac{1}{2}(\mathbf{q}_t^{DR}\cdot\mathbb{E}[\hat{\boldsymbol{\rho}}_t-\boldsymbol{\rho}_t]+\mathrm{Cov}(\hat{\mathbf{q}}_t^{DR}-\mathbf{q}_t^{DR},\hat{\boldsymbol{\rho}}_t-\boldsymbol{\rho}_t)+\\&\mathbf{q}_t^{TV2}\cdot\mathbb{E}[(1-\hat{\boldsymbol{\rho}}_t)-(1-\boldsymbol{\rho}_t)]+\mathrm{Cov}(\hat{\mathbf{q}}_t^{TV2}-\mathbf{q}_t^{TV2},(1-\hat{\boldsymbol{\rho}}_t)-(1-\boldsymbol{\rho}_t)))\end{aligned}$$

Which we reduce to:

$$\mathbb{E}[\hat{\pi}_t^{MLE}]-\pi_t^{true}=\frac{1}{2}\Big((\mathbf{q}_t^{DR}-\mathbf{q}_t^{TV2})\cdot\mathbb{E}[\hat{\boldsymbol{\rho}}_t-\boldsymbol{\rho}_t]+\mathrm{Cov}((\hat{\mathbf{q}}_t^{DR}-\mathbf{q}_t^{DR})-(\hat{\mathbf{q}}_t^{TV2}-\mathbf{q}_t^{TV2}),\hat{\boldsymbol{\rho}}_t-\boldsymbol{\rho}_t)\Big) \tag{2}$$

The first term stems from $\hat{\boldsymbol{\rho}}_t$ being a non-linear transformation of $\hat{\mathbf{q}}_t^{DR}$ and $\hat{\mathbf{q}}_t^{TV2}$, and not the true $\boldsymbol{\rho}_t$ value. Given that the non-linearity in $\hat{\boldsymbol{\rho}}_t$ is modest, this component of the bias is small in practice.

The second term represents the primary source of concern in our context, as it is the main driver of upward bias. Since $\hat{\boldsymbol{\rho}}_t$ is constructed directly from $\hat{\mathbf{q}}_t^{DR}$ and $\hat{\mathbf{q}}_t^{TV2}$, the sampling error in the empirical token frequencies - $\hat{\mathbf{q}}_t^{DR}$ and $\hat{\mathbf{q}}_t^{TV2}$ - transfers mechanically to that of $\hat{\boldsymbol{\rho}}_t$. For example, if $\hat{\mathbf{q}}_t^{DR}$ deviates to some extreme values in a given period, the resulting source probabilities in $\hat{\boldsymbol{\rho}}_t$ will become highly weighted towards DR, thereby inflating the polarisation estimate $\hat{\pi}_t^{MLE}$. Notably, even if the true token frequencies are equal - that is, $\mathbf{q}_t^{DR}=\mathbf{q}_t^{TV2}$ - finite-sample bias alone may lead to polarisation estimates above the true null-polarisation, due to estimation noise.

5.2 Leave-out estimator

We address the severe bias of the MLE by utilising a leave-out (LO) specification of polarisation on the following form:

$$\hat{\pi}_t^{LO}=\frac{1}{2}\left(\frac{1}{|\mathrm{DR}_t|}\sum_{i\in\mathrm{DR}_t}\hat{\mathbf{q}}_{i,t}\cdot\hat{\boldsymbol{\rho}}_{-i,t}+\frac{1}{|\mathrm{TV2}_t|}\sum_{i\in\mathrm{TV2}_t}\hat{\mathbf{q}}_{i,t}\cdot\Big(1-\hat{\boldsymbol{\rho}}_{-i,t}\Big)\right) \tag{3}$$

The key difference compared with the MLE is the token sample on which $\hat{\mathbf{q}}^S$ and $\hat{\boldsymbol{\rho}}$ are constructed. This specification "leaves out" article i when computing the source probability, hence we denote the new source probabilities $\hat{\boldsymbol{\rho}}_{-i}$ in eq. (3). Now the article-wise choice set of possible tokens excludes tokens used in article i. Following the example outlined in section 5.1., a token mentioned only in one article at time t will no longer have an effect on the polarisation, eliminating the amplifying effect it had in the MLE. As a consequence, we construct article-wise token frequency vectors $\hat{\mathbf{q}}_{i,t}$ and source probability vectors $\hat{\boldsymbol{\rho}}_{-i,t}$. We normalise the dot product by the source-specific total number of articles in t to obtain the LO-estimator of polarisation. To show the implied reduction in the bias first presented in section 5.1., we outline the bias term for the LO-estimator, but exclusively for DR articles' contribution, for simplicity:

$$\begin{aligned}&\mathbb{E}\left[\frac{1}{|\mathrm{DR}_t|}\sum_{i\in\mathrm{DR}_t}(\hat{\mathbf{q}}_{i,t}\cdot\hat{\boldsymbol{\rho}}_{-i,t}-\mathbf{q}_{i,t}\cdot\boldsymbol{\rho}_t)\right]\\=&\frac{1}{|\mathrm{DR}_t|}\sum_{i\in\mathrm{DR}_t}\left(\mathbf{q}_{i,t}\cdot\mathbb{E}\left[\hat{\boldsymbol{\rho}}_{-i,t}-\boldsymbol{\rho}_t\right]+\mathrm{Cov}(\hat{\mathbf{q}}_{i,t}-\mathbf{q}_{i,t},\hat{\boldsymbol{\rho}}_{-i,t}-\boldsymbol{\rho}_t)\right)\end{aligned}$$

While this term resembles that of the MLE, the distinction between them is substantial. As $\hat{\boldsymbol{\rho}}_{-i,t}$ is estimated on the population excluding article i, the sampling error becomes independent from that of $\hat{\mathbf{q}}_{i,t}$. This effectively breaks the mechanical relation between the sampling errors of the two, eliminating the second term from equation (2):

$$\mathrm{Cov}(\hat{\mathbf{q}}_{i,t} - \mathbf{q}_{i,t}, \hat{\boldsymbol{\rho}}_{-i,t} - \boldsymbol{\rho}_t) \approx 0$$

Thus, the LO-estimator eliminates the primary source of bias present in the MLE and retains only the bias from the non-linear transformation of $\hat{\boldsymbol{\rho}}_{-i,t}$. By extending the bias term to include TV2 and adopting sample average notation, where $\frac{1}{|\mathrm{S}_t|}\sum_{i\in \mathrm{S}_t} \hat{\mathbf{q}}_{i,t} = \bar{\mathbf{q}}_t^S$, with corresponding true values $\mathbf{q}_t^{DR}$ and $\mathbf{q}_t^{TV2}$, we obtain the total bias term for the LO-estimator:

$$\mathbb{E}[\hat{\pi}_t^{LO}] - \pi_t^{true} = \frac{1}{2}(\mathbf{q}_t^{DR} - \mathbf{q}_t^{TV2}) \cdot \frac{1}{|\mathrm{DR}_t| + |\mathrm{TV2}_t|} \sum_i \mathbb{E}\left[\hat{\boldsymbol{\rho}}_{-i,t} - \boldsymbol{\rho}_t\right] \tag{4}$$

As previously mentioned, this term stems from Jensen's inequality and tends to be small. Above, the notation of estimated source probability $\hat{\boldsymbol{\rho}}_{-i,t}$ differs from that of the true value $\boldsymbol{\rho}_t$. This is because the true (unknown) value of the source probability is fixed at time t, and thus unaffected by dropping one observation.

To illustrate the calculations of the polarisation using the LO-specification, we draw on the sandbox example setup from section 5.1. We present the full calculations in Appendix 10.11. Applying the framework of the LO-estimator to the same setup yields the following results:

$$\hat{\mathbf{q}}_i = \mathbf{c}_i/m_i \rightarrow \quad \hat{\mathbf{q}}_{DR_1} = \begin{bmatrix} 0/3 \\ 2/3 \\ 0/3 \\ 1/3 \\ 0/3 \end{bmatrix}, \quad \hat{\mathbf{q}}_{DR_2} = \begin{bmatrix} 0/3 \\ 1/3 \\ 1/3 \\ 1/3 \\ 0/3 \end{bmatrix}, \quad \hat{\mathbf{q}}_{TV2_1} = \begin{bmatrix} 2/3 \\ 0/3 \\ 0/3 \\ 0/3 \\ 1/3 \end{bmatrix}, \quad \hat{\mathbf{q}}_{TV2_2} = \begin{bmatrix} 0/3 \\ 0/3 \\ 2/3 \\ 0/3 \\ 1/3 \end{bmatrix}$$

$$\hat{\boldsymbol{\rho}}_{-DR_1} = \begin{bmatrix} 0 \\ 1 \\ 1/2 \\ 1 \\ 0 \end{bmatrix}, \quad \hat{\boldsymbol{\rho}}_{-DR_2} = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \quad \hat{\boldsymbol{\rho}}_{-TV2_1} = \begin{bmatrix} 0 \\ 1 \\ 1/5 \\ 1 \\ 0 \end{bmatrix}, \quad \hat{\boldsymbol{\rho}}_{-TV2_2} = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \\ 0 \end{bmatrix}$$

$$\hat{\mathbf{q}}_{DR_1} \cdot \hat{\boldsymbol{\rho}}_{-DR_1} = 1, \quad \hat{\mathbf{q}}_{DR_2} \cdot \hat{\boldsymbol{\rho}}_{-DR_2} = \frac{2}{3}, \quad \hat{\mathbf{q}}_{TV2_1} \cdot (1 - \hat{\boldsymbol{\rho}}_{-TV2_1}) = 1, \quad \hat{\mathbf{q}}_{TV2_2} \cdot (1 - \hat{\boldsymbol{\rho}}_{-TV2_2}) = \frac{1}{3},$$

$$\hat{\pi}^{LO} = \frac{1}{2}\left(\frac{1}{2}\left(1 + \frac{2}{3}\right) + \frac{1}{2}\left(1 + \frac{1}{3}\right)\right) = \frac{3}{4}$$

The example demonstrates how the estimation of polarisation differs between the two methods. In this high-polarisation scenario, where the two outlets only share one of five tokens, the LO-estimate yields a lower polarisation than the MLE: $\hat{\pi}^{LO} = \frac{3}{4} < \hat{\pi}^{MLE} = \frac{31}{36}$. The example also serves to illustrate the added complexity of the LO-estimator. Specifically, for each article i, we compute a separate token frequency vector and a corresponding source probability vector $\hat{\boldsymbol{\rho}}_{t,-i}$. As a result, increasing $\mathbf{J}$ by one token, increases the length of $\hat{\mathbf{q}}_{i,t}^{TV2}$, $\hat{\mathbf{q}}_{i,t}^{DR}$, and $\hat{\boldsymbol{\rho}}_{-i,t}$ vectors. The combined effect of increasing both i and $\mathbf{J}$ leads to an exponential growth in computational complexity. These additional steps impose a substantially greater computational burden in estimating polarisation compared to the MLE.

In consideration of the volume of data outlined in 3.1.1, the computational complexity becomes a limiting factor in estimating polarisation, and an important consideration in selecting the estimator. The LO-estimator is a suitable way to uncover the polarisation in the data, but has certain limitations. It does not allow the use of covariates, and additionally does not yield estimates of the underlying parameters of the model. The latter prevents a direct analysis of the most polarising tokens. In our case, these limitations are of no concern, as the analysis does not aim to investigate the partisanship of specific tokens.

5.3 Penalised estimator

The penalised estimator is proposed by Gentzkow et al. (2019), as the optimal way to reduce finite sample bias, while also generating both token-specific parameter estimates and enabling the use of covariates. In the original specification, a vector of speaker characteristics is a central part of the estimation. This would correspond to article or author characteristics in this paper, but these are left out, as data on authors is limited and complicated to obtain for all journalists across 10 years in the two media. A simplified version of the penalised estimator, excluding characteristics, will be outlined theoretically. However, given the scope and time-constraint of this study, we do not pursue its implementation. The penalised estimator yields coefficient estimates that can be used to plug into the following estimator of token frequency:

$$\begin{aligned} q_{j,t}^{S(i)} &= \frac{e^{u_{i,j,t}}}{\sum_l e^{u_{i,l,t}}} \\ u_{i,j,t} &= \alpha_{j,t} + \varphi_{j,t}\mathbf{1}_{i\in DR} \end{aligned} \tag{5}$$

Where l denotes the length of tokens in i. The parameter estimates are derived by minimising the following objective function:

$$\sum_j \Big\{ \sum_t \sum_i \left[m_{i,t} \exp(\alpha_{j,t} + \varphi_{j,t}\mathbf{1}_{i\in\mathrm{DR}}) - c_{i,j,t}(\alpha_{j,t} + \varphi_{j,t}\mathbf{1}_{i\in\mathrm{DR}}) + \psi\left|\alpha_{j,t}\right| + \lambda_j\left|\varphi_{j,t}\right|\right] \Big\} \tag{6}$$

This yields estimates of $\{\boldsymbol{\alpha}_t, \boldsymbol{\varphi}_t\}$ to be plugged into eq. (5), which returns a vector of token frequencies by source. These vectors can be used in the following to get estimates of the average polarisation, $\bar{\pi}_t$, following the previously defined definition of $\hat{\boldsymbol{\rho}}_t$:

$$\bar{\pi}_t = \frac{1}{|DR_t \cup TV2_t|} \sum_{i\in DR_t\cup TV2_t} \frac{1}{2}\left(\hat{\mathbf{q}}_t^{DR} \cdot \hat{\boldsymbol{\rho}}_t + \hat{\mathbf{q}}_t^{TV2} \cdot (1 - \hat{\boldsymbol{\rho}}_t)\right) \tag{7}$$

The minimand in eq.(6) reflects two key decisions, as outlined by Gentzkow et al. (2019). The first key decision, the likelihood of the multinominal logit model is approximated with the likelihood of a Poisson model, where $c_{i,j,t} \sim \mathrm{Pois}(\exp(\mu_{i,t} + u_{i,j,t})$ and $\hat{\mu}_{i,t} = \log m_{i,t}$. This approach is adopted given that, fixing $\hat{\mu}_{i,t}$, the likelihood is separable across tokens, where Gentzkow et al. argue that it would otherwise be infeasible to compute. The second key decision made by Gentzkow et al., is to use a Lasso penalty (L_1) in formulating the estimator, which imposes sparsity, penalising high values as it shrinks coefficients towards zero. The appeal of this method is strengthened by its ability to limit the effect of sampling error and thus the source of the finite sample bias. Gentzkow et al. determine the penalty size $\boldsymbol{\lambda}$ by regularisation path estimation, where a value large enough to satisfy $\varphi_{j,t} = 0$ is chosen as the starting point. By gradually reducing the size of λ_j, and updating parameter estimates, they choose the optimal λ_j, as the value which minimises a Bayesian Information Criterion (BIC). Gentzkow et al. further implement a minimum penalty ψ, which allows numerical convergence while still treating the covariates flexibly. These steps dramatically increase the computational load, given our volume of data and number of individual tokens in $\mathbf{J}$, as described in section 3.1.3.

In the context of our analysis, this approach is purely a theoretical exploration of how to deal with finite sample bias in the most complete way. As the limitations of the LO-estimator are not of concern in our context, the penalised estimator entails only marginal benefits compared to the LO-specification, that are greatly outweighed by the added complexity and computational load it requires. Based on all of the above, we consider the LO-estimator the theoretically optimal method in our context.

5.4 Inference

To draw inference, we calculate point-wise confidence intervals on the estimates. This is done by performing random subsampling without replacement of articles, drawing $K = 100$ subsamples of size $\tau_{k,t} = \tau_t/10$, with τ_t denoting the total number of articles in biweek t. For the k^{th} subsample, we calculate the subsample estimate of polarisation, $\hat{\pi}_t^k$. The confidence interval around the estimate is then:

$$CI_t \equiv (\hat{\pi}_t - (Q_t^k)_{(90)}/\sqrt{\tau_t}\,;\, \hat{\pi}_t - (Q_t^k)_{(11)}/\sqrt{\tau_t}) \tag{8}$$

Where $(Q_t^k)_{(c)}$ is the c^{th} order statistic of $Q_t^k = \sqrt{\tau_{k,t}}(\hat{\pi}_t^k - \frac{1}{K}\sum_{l=1}^{K}\hat{\pi}_t^l)$. Q_t^k being the normalised deviation of $\hat{\pi}_t^k$ from the subsample average, for each subsample k. Sorting and inserting the 11^{th} and 90^{th} order statistics of Q_t^k into eq. (8) then provides an 80 pct. confidence interval around each estimate of polarisation (Gentzkow et al., 2019; Politis et al., 2001).

We employ the method above generally across estimators. This returns confidence intervals for all estimates across time. Though it presents a simple check for whether the estimate is statistically different from the theoretical null-polarisation, $\hat{\pi}^{\text{null}} = 0,5$, it cannot stand alone. The true null-polarisation, in our context, is the polarisation level when the source of each article is randomised, which follows a distribution centred around the theoretical null-polarisation.

Therefore, we construct a control polarisation that contains the statistical variance we expect in the true null-polarisation. For the full set of articles in t, we scramble the media vector and sample new article vectors without replacement, after which the relationship between article content and media is random. For each time-point, we re-estimate polarisation on the control sample, returning $\hat{\pi}_t^*$, which we denote control polarisation. In the K subsamples with randomly allocated media affiliation, we estimate $\hat{\pi}_t^{k*}$ and Q_t^{k*}. Inserting terms into eq. (8) returns CI_t^*, which provides point-wise confidence intervals for the estimated control polarisation. For each time-point, we draw inference about the presence of polarisation by comparing the confidence interval of the estimate to that of the control polarisation.

Apart from providing a benchmark which we can compare estimated polarisation against, the control polarisation contributes with another important feature. As we estimate $\hat{\pi}_t^*$ using the same estimator as for $\hat{\pi}_t$, control polarisation directly shows how inherently biased the estimator is. Again, exactly because we scramble media affiliation randomly prior to estimating $\hat{\pi}_t^*$, it *should*, for an unbiased estimator, on average equal $\hat{\pi}^{\text{null}} = 0,5$. Any systematic deviation from this unequivocally documents inherent bias in an estimator.

5.5 Token filtering

When constructing the two estimators, the input tokens are of great importance. As visible in the sandbox examples in section 5.1. and 5.2., when a token occurs only once in one article or only in one media, the DR source probability $\hat{\rho}_j$ becomes either 1 or 0, depending on which media used the token. Further, the computational load depends heavily on the length of the J-vector at each time t. In practice, we construct a filtering of all tokens before implementing the estimator. In constructing the filter, we take different aspects into consideration. We do this in two steps: First, we filter tokens based on conditions applicable across the entire time-frame. Second, we filter tokens within the individual periods t.

In determining the limits and parameters by which we filter, we have to factor in several aspects, inspired by the approach of Caprini (2024). The most common tokens are likely not informative of the polarisation. This could be tokens like *"i"*, which is a common token when we split bi- and trigrams into unigrams. In itself, it does not carry any meaning and is likely represented equally between DR and TV2, thus it only contributes to increasing the computational load. Assuming that the most frequent tokens are equally distributed and only contribute to increasing runtime, we choose to filter out the top 0,5 pct. most used tokens. Another aspect is the possible polarisation stemming from very unique articles, e.g. theme weeks concerning very specific topics or other editorial choices that result in highly unique material. To limit the potentially enhancing effect this could have on polarisation, we exclude tokens that are mentioned in less than 10 different periods and tokens that are mentioned less than 50 times across the entire time-frame. These three filters combined constitute the first token filtering step. In the second step we exclude all tokens that are mentioned fewer than three times in the given period. This is done to ensure that when using the LO-estimator for article i containing token j, at least two other articles in t also contain token j. This increases the possibility of getting a meaningful DR probability vector $\hat{\rho}_{-i,j}$. The effect of our filtering is a reduction from 11.954.367 tokens, to 670.191 valid tokens, as presented in table 1. We detail how different criteria-combinations affect estimates of polarisation in section 7.1.1., as well as outline the impact of varying period-lengths in section 7.1.2.

5.6 Estimated polarisation

We have estimated polarisation using both the MLE and the LO estimator in order to test the main working hypothesis: That there is polarisation in Danish media. Comparing MLE and LO-estimates highlights the benefits of our preferred estimator. As argued theoretically, the MLE suffers from a significant finite sample bias, which the LO-estimator drastically reduces by eliminating the term in eq. (2) that contains the most prominent bias. A measure using the penalised estimator is out of scope for this paper, due to the both its complexity, extensive computational load, and the satisfactory performance of our LO-estimator. We predominantly base the following analysis in section 6. on polarisation estimated in biweekly intervals. In section 7., we examine empirically how polarisation depends on these intervals. In figure 20., we illustrate the MLE-estimates of polarisation from 2015-2024 with pointwise confidence intervals and control polarisation. Immediately, we note very high polarisation estimates in a range strictly above $0,575$ and above $0,6$ for long periods, and that the control polarisation is well above $0,55$ in most periods. This supports the theoretical expectation that the MLE-estimates suffers from a large upward bias. We confirm this by comparing the MLE-estimates in figure 20. to the LO-estimate of polarisation illustrated in figure 21. Narrowing the focus to the control polarisation, the level difference between the two estimates is approx. $0,05$ to $0,07$, suggesting a bias of at least this magnitude. The polarisation using the LO-estimator varies, but lies in the range of $0,510$ to $0,575$ in the majority of the periods, along with a control polarisation narrowly fluctuating around $0,5$, as we would expect under random assignment. In fact, the average control polarisation for the LO-estimate differs from 0,5 only at the 4$^{\text{th}}$ decimal place.

The comparison above shows the finite sample upward bias from the MLE and that our LO-estimator reduces the bias very well. This finding supports the theoretical merits of the LO-estimator as a reliable tool to estimate the polarisation between DR and TV2. In the following, we exclusively consider polarisation estimated using the LO-estimator.

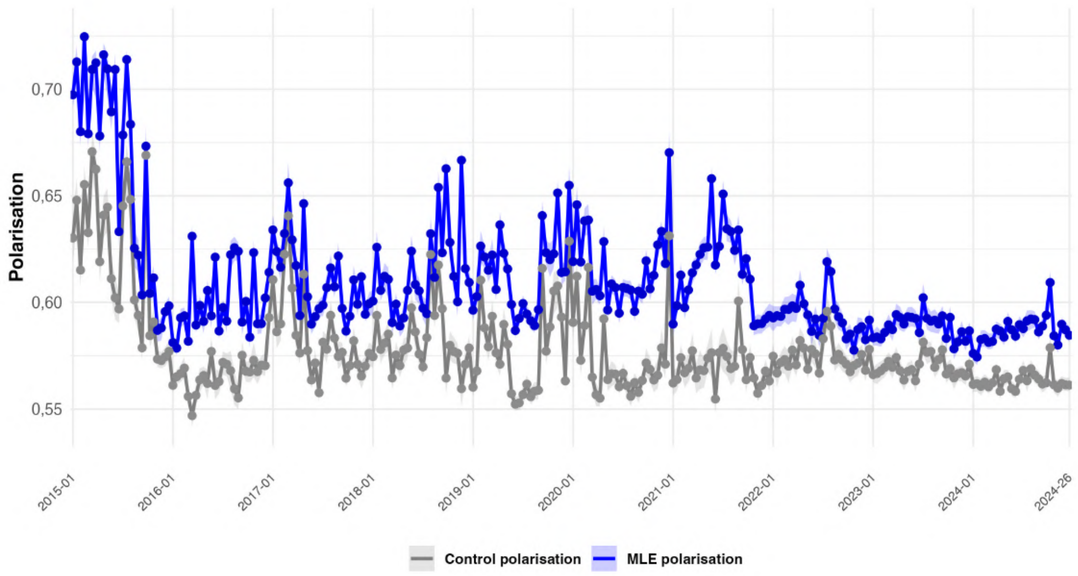


Figure 20: MLE General polarisation 2015-2024

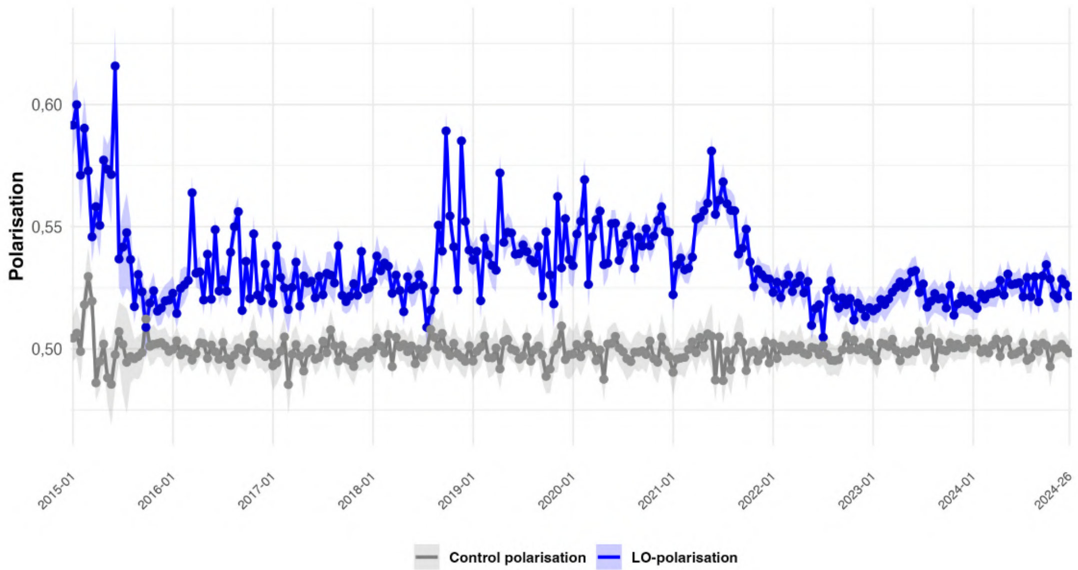


Figure 21: LO General polarisation 2015-2024

In figure 21., we have illustrated polarisation from the LO-estimator. The general measure of polarisation is significantly different from control polarisation for all but 8 biweeks. This suggests that there is polarisation in the token representation of the news coverage. The tokens represent the actual coverage from DR and TV2, and because we have no reason to believe that the feature extraction is biased toward either media, we, by extension, accept our main working hypothesis: that there is polarisation in Danish media. General polarisation between DR and TV2 is on average $0,534$ including and $0,532$ excluding 2015[14]. In line with the result in Caprini (2024) in an American setting, on a much shorter time-horizon, we observe no clear trend in polarisation across time. Polarisation in the first half of 2015 is the highest estimated in the entire ten-year period, with estimates around $0,60$. In the second half of 2015, this is sharply reduced to a low and to some degree steady level in the range $0,52$-$0,55$, which continues until the last quarter of 2018. In these three years, the estimated polarisation exceeds the upper limit of the range in just four biweeks. By the end of 2018, polarisation increases, with estimates nearing $0,59$. This marks a period of relatively high polarisation from the second half of 2018 and three years ahead until the second half of 2021. In this period, polarisation frequently exceeds $0,55$, while just once going below $0,52$. Notably, the polarisation appears to follow a slight upwards trend up until the 13$^{\text{th}}$ biweek in 2021. From the second half of 2021 polarisation declines. At the first biweek of 2022, polarisation has fallen to $0,53$, but the decline continues albeit less sharply throughout 2022. In 2022-2024 polarisation enters a period of moderation, with some of the lowest estimates of the ten-year period, typically around $0,52 - 0,53$. Considering this last period, we find that polarisation does not increase over time, leading us to reject WH4. It is a period characterised by consistency, with some of the smallest biweek-on-biweek changes observed. While polarisation is low and consistency is high, there appears to be a small upward trend in the estimates in this last year of the time period. By the end of 2024, polarisation is approx. $0,53$. The estimates in the last three years are on level with the polarisation Caprini estimates for American news. Here, polarisation in 2022 is estimated at around $0,525$, ranging from $0,518$ to $0,535$, exactly overlapping with the estimated polarisation in Danish news from 2022-2024 as noted above. In this context the estimated polarisation in the years 2019-2021 is relatively high, but we do not have a similar estimate for polarisation in American news in these years to compare against. The estimates of polarisation support the main working hypothesis: that there is polarisation in Danish media.

We assign little validity to the initial high polarisation and subsequent sharp decline in 2015, as both the level and trend in the year are by all accounts explained methodologically by the section composition and data availability. We have on average as little as half of the articles in 2015 as in any other year, cf. figure 5. This affects reliability, as is evident from the large volatility in control polarisation. Importantly, the section categorisation does not apply at all for TV2 in 2015 and is limited to the second half of 2015 for DR, cf. figure 4. This reduces accuracy of the mapping between articles across media, as sections provide a check that the news stories covered by the media are the same. The reduced mapping greatly increases the estimated polarisation. In conjunction, this leads to high and unreliable estimates of polarisation in 2015. We note that while the control polarisation in 2015 is unreliable (it does not include 0,5 in the confidence interval in all biweeks), it is on average centred around 0,5. This signals that we cannot state that the high estimates in 2015 are due to an upward bias from a reduced sample size.

The general polarisation estimates appear to be defined within distinct regimes which govern both level and volatility, cf. figure 21. Visual inspection indicates a period with historically high levels and volatility of polarisation from the second half of 2018 up until and including the first half of 2021. This regime is in sharp contrast with the following period of moderation, but is also visually distinct from the preceding years - excluding 2015 - in which polarisation is lower and more stable. Notably, this development in polarisation somewhat mimics the development in the sentiment score gap between media, see figure 12. We provide a rigorous

[14]We present an intuitive interpretation of these estimates in section 5.

framework for analysing whether there are distinct regimes in polarisation in section 6.4.

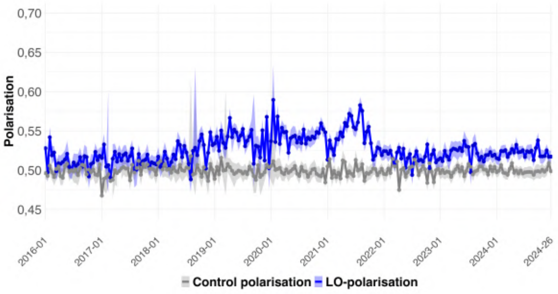
Conditioning the allowed token set to tokens from articles within main topics provides section-specific polarisation. These are illustrated for sections *Indland*, *Udland*, *Politik*, and *Other* in figure 22. in the period 2016-2024. The conditioning by section reduces the potential set of allowed tokens, theoretically upwardly biasing the estimated polarisation by limiting the sample. However, in practice, we do not observe this. We note how polarisation is robust to limitations in the allowed token set, as polarisation in *Politik* is relatively low and represents a smaller section by article count than e.g. *Udland*, cf. figure 4. For all sections, control polarisation is centred around 0,5. We conclude that even when we employ the most restrictive conditioning on the token set, the LO-estimate is not suffering from bias due to sample size. Sample size however does greatly affect reliability. We observed this in the general polarisation estimate in 2015 above, and observe it again in figure 22., panel (c), illustrating polarisation and its wide confidence intervals in the section *Politik*.

For all sections, polarisation follows the same development as the general polarisation, though the level and volatility differ, cf. figure 21. and 22. The sections *Indland* and *Udland* are both of similar size and together constitute roughly half of the article count, but have different contributions to the general polarisation. Polarisation in *Indland* is high with an average of $0,526$, which overlaps with control polarisation in 93 biweeks out of 234. Polarisation in *Udland* is similar, but marginally lower - on average $0,521$ - and overlaps with control polarisation in 128 biweeks. For both *Udland* and *Indland*, the polarisation in the years 2018-2021 is higher than in the rest of the period. Thus, the underlying dynamic is the same, but the level in *Udland* is marginally lower.

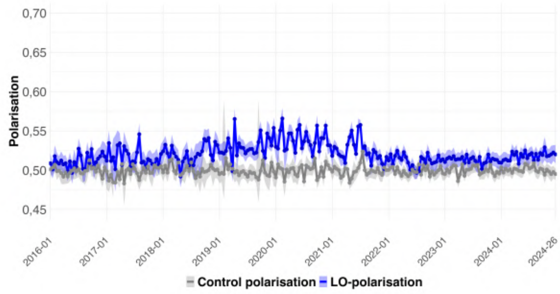
Focusing on *Politik*, the section with the lowest article count, reveals broad confidence intervals around low estimates of polarisation, on average $0,517$. In 180 out of 234 biweeks the confidence interval of the estimate includes the control polarisation, indicating that there is effectively no polarisation within the section.

In the categorisation of sections into four main categories, we have defined a compiled section, *Other*, containing relevant news stories that do not fit into the other three main sections, cf. section 3.2. The average polarisation is $0,546$ in *Other*, compared to $0,526$ for *Indland*. We observe overlaps of the confidence interval of the estimate and control polarisation in 13 biweeks for the section *Other*.

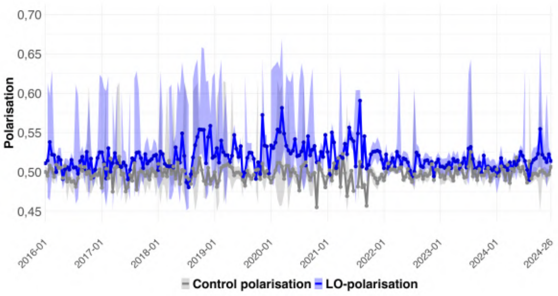
For *Other* there is a somewhat methodological reason behind the high polarisation, as *Other* contains articles with a weaker mapping between media. The articles included constitute a significant and relevant part of the media stream, though, so including them in the general polarisation is warranted. However, we conclude that while the high polarisation within *Other* reflects our study design choices, we have no such explanation for the level of polarisation in *Indland*. The estimated polarisation directly reflects the polarisation we observe in tokens, based on the article and image content, establishing a straight line between the editorial line within the two media and our estimate of polarisation within *Indland*.



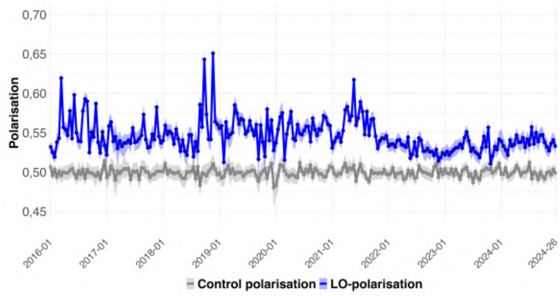
(a) Section-specific polarisation: *Indland*



(b) Section-specific polarisation: *Udland*



(c) Section-specific polarisation: *Politik*



(d) Section-specific polarisation: *Other*

Figure 22: Section specific polarisation, 2016-2024

6 Analysis of polarisation

Having found support for our main working hypothesis in section 5.6.: that there is polarisation in Danish media, we turn our attention to WH4 and WH5. To test our working hypotheses, we start by determining underlying characteristics of the time-series. This includes how polarisation correlates across sections, to determine to what degree general polarisation depends on specific sections and its own past values. Then we propose a model framework to check for seasonality in polarisation. WH4 hypothesises that polarisation increases over time, which we found to be inconsistent with the pattern observed graphically in the general polarisation illustrated in figure 21. We further investigate this in section 6.3. by modelling polarisation with trend components. WH5 hypothesises that contemporaneous events drive polarisation. To explore this, we employ various research designs. In section 6.3., we relate the hypothesis to the development in the sentiment gap and investigate electoral cycles in both time-series. Further, in section 6.4.1., we employ a regression discontinuity design to explore whether a major singular event drives polarisation. This approach is then supported, in section 6.4.2., by detection of change-points in variance. Finally, in section 6.5., we use a segmented regression design to test whether several singular events have short-term impacts on polarisation.

6.1 Correlation across sections

In section 5.6. we have established a difference between polarisation within each section. Given the particularly high levels of polarisation within *Other*, and moderately high levels in *Indland*, we suspect these sections are the main contributors to the general polarisation. To further examine the dynamics of both section-specific and general polarisation, we investigate the extent to which polarisation depends on its own lagged values, as well as the degree of correlation between the different polarisation measures. As outlined in section 2.1., expansions in the domain of publicly accepted speech can have spillover effects (Bursztyn et al., 2020), raising the question of whether for example, polarising events covered in one section might spill over into the coverage in other sections. We address this straightforwardly by estimating correlation coefficients between section-specific polarisation across time.

Plotting the correlations between general polarisation and section-specific polarisation across time reveals a strong correlation between general polarisation and the polarisation within *Indland* and *Other*, cf. figure 23. This correlation supports the preliminary conclusion that these two sections are the main drivers of the general polarisation. Furthermore, figure 23. suggests limited within-period contribution from *Udland* to the general polarisation and effectively zero contribution from *Politik*. These findings align with our expectations based on visual patterns observed in section 5.6. In addition, figure 23. provides insight into how polarisation correlates across sections over time. General polarisation at time t is moderately and symmetrically correlated with its own values in nearby periods, and similarly with polarisation in *Other*. Although *Indland* contributes contemporaneously to general polarisation, these effects are weak and diminish quickly over time. The correlation between general polarisation at time t and polarisation in *Udland* at $t \neq 0$ is minimal. This is indicative of a process in which general polarisation is a result of the contemporaneous contributions from the coverage of *Other*, *Indland*, and *Udland*, and is not, in the short run, imported through leading polarisation in, e.g., foreign news stories in *Udland*.

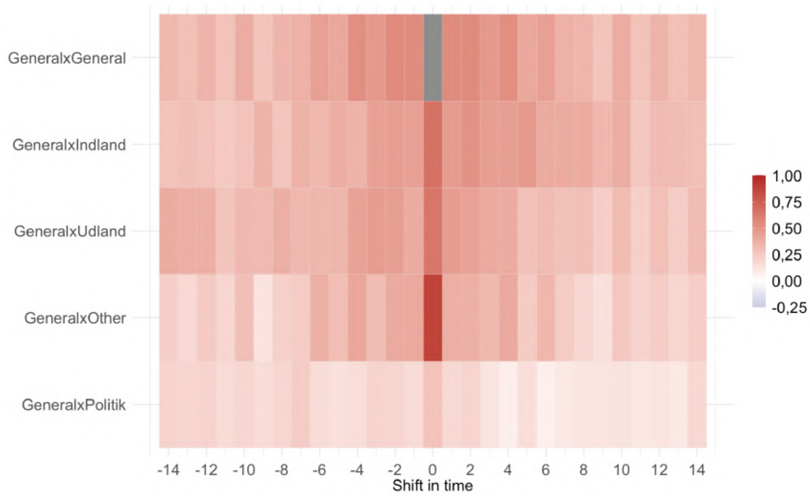


Figure 23: Correlation coefficients for general polarisation across time and sections

We illustrate the same cross-section correlation coefficients for each section in figure 24. These results show no clear evidence of any section leading general polarisation, thereby supporting our previous preliminary conclusion presented above. The distribution of coefficients across *Indland* and *Udland* broadly mirrors that of general polarisation, while we observe no systematic evidence of contributions from *Other* and *Politik*. In fact, for *Politik* the contemporaneous polarisation appears to be almost independent of itself across time too.

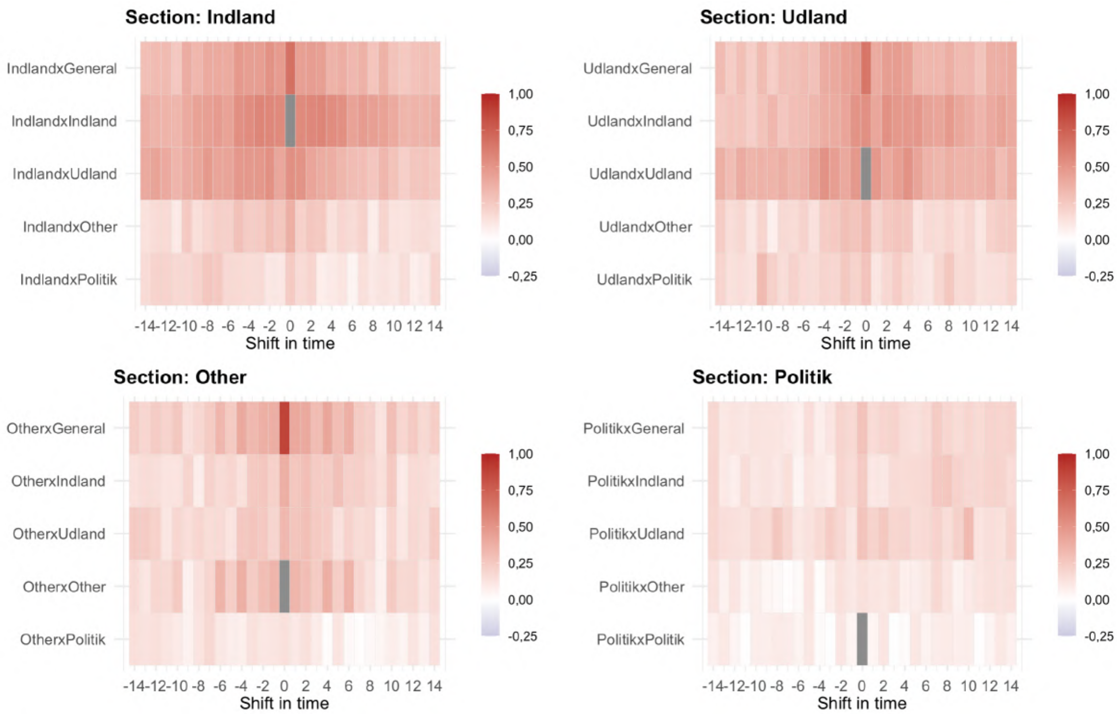


Figure 24: Correlation coefficients for section-specific polarisation across sections and time

6.2 Seasonality and conditional variance

To the best of our knowledge, polarisation in news media has not previously been estimated over a time span like ours - certainly not in a Danish context. Consequently, it remains an open question how polarisation behaves over time. To uncover what dynamics govern polarisation as a time-series, we adopt an exploratory approach, empirically testing a range of factors that *could* explain time-variation in polarisation.

We recognise that there is limited relevant literature to draw on in the formulation of hypotheses and expectations regarding factors that influence polarisation over a time-series as long as the one estimated here. However, classical econometric time-series analysis motivates a check for seasonality and timing effects. Thus, we employ month-, quarter-, and yearly fixed effects and specify dummies for two specific time-periods during a given year: biweek 26 including the Christmas holidays and biweek 15 spanning week 29 and 30, commonly referred to as the height of the *silly season*, *slow-news season*, or simply *"agurketid"*. We test the seasonality effects iteratively by including them in an autoregressive baseline model of polarisation.

In formulating a baseline model of polarisation, we establish that an autoregressive model in levels with a moving average of lag length 1 - ARMA(1,1) - minimises the Bayesian Information Criterion. The ARMA(1,1) model exploits the stationary and autoregressive dynamic in polarisation which we establish using appropriate tests for stationarity and autocorrelation, cf. section 7. and Appendix 10.12. The tests, however, reveal heteroskedasticity in polarisation, warranting a baseline model formulation that accounts for conditional volatility. We cannot accommodate this through robust standard errors due to the MA-component. Therefore, we address this through the use of an ARMA-GARCH model framework. Considering the lack of reliability of polarisation in 2015, we exclude observations from 2015 in the testing and estimation. We revisit how this affects the results in section 7.

To account for methodological drivers of the measure of polarisation, we control for either the total number of tokens available or the total number of valid unique tokens, as is common in qualitative text analysis (Grimmer and Stewart, 2013). The two counts are near-perfectly proportional by a factor of $1:18$ and we include covariates of the valid token count after rescaling and centring the count to achieve numerical stability.

Controlling for exogenous token count expands the baseline model to an ARMAX(1,1) model of conditional mean and a GARCH(1,1) model of conditional variance, allowing for autocorrelation, moving average dynamics, exogenous covariates, and conditional heteroskedasticity. The model is presented in eq. (9)-(10). The mean equation captures covariates of both control and seasonality in $\boldsymbol{\gamma}'\boldsymbol{\Omega}_t$, while accounting for persistence in mean and shock adjustment in β_1 and β_2, respectively. In the variance equation the time-conditional variance, σ_t^2, depends on the ARCH-effect, $\alpha_1\varepsilon_{t-1}^2$, capturing how shocks increase variance, and the GARCH-effect, $\alpha_2\sigma_{t-1}^2$, describing persistence in volatility.

$$\text{Mean equation:} \quad X_t = \beta_0 + \beta_1 X_{t-1} + \beta_2\varepsilon_{t-1} + \boldsymbol{\gamma}'\boldsymbol{\Omega}_t + \varepsilon_t, \tag{9}$$

$$\text{Variance equation:} \quad \sigma_t^2 = \omega + \alpha_1\varepsilon_{t-1}^2 + \alpha_2\sigma_{t-1}^2 \tag{10}$$

Table 5. reports parameter estimates from the baseline ARMAX(1,1)-GARCH(1,1) model, controlling for valid token count in all specifications. M1 includes only the covariates for token count. We expand this in models where seasonality effects are included individually, with month-, quarter-, and yearly fixed effects separately included in models M2, M3, and M4[15]. Finally, M5 includes dummies for biweek 15 and 26.

Table 5: ARMAX(1,1)-GARCH(1,1) Seasonality checks

	Dependent variable: General polarisation, 2015-2024				
	M1	M2	M3	M4	M5
AR(1)	$0,941^{***}$	$0,955^{***}$	$0,951^{***}$	$0,797^{***}$	$0,942^{***}$
	$(0,030)$	$(0,032)$	$(0,029)$	$(0,224)$	$(0,030)$
MA(1)	$-0,604^{***}$	$-0,653^{***}$	$-0,641^{***}$	$-0,555^{***}$	$-0,605^{***}$
	$(0,079)$	$(0,096)$	$(0,087)$	$(0,128)$	$(0,083)$
Intercept	$0,526^{***}$	$0,525^{***}$	$0,525^{***}$	$0,524^{***}$	$0,526^{***}$
	$(0,004)$	$(0,006)$	$(0,005)$	$(0,002)$	$(0,004)$
Valid Token Count	$0,002$	$0,002$	$0,001$	$0,001$	$0,002$
	$(0,001)$	$(0,002)$	$(0,001)$	$(0,001)$	$(0,001)$
Month effects		yes			
Quarter effects			yes		
Year effects				yes	
Specific Timing Effects					yes
Variance Equation:					
ω	$0,000$	$0,000$	$0,000$	$0,000$	$0,000$
	$(0,000)$	$(0,000)$	$(0,000)$	$(0,000)$	$(0,000)$
α_1	$0,140$	$0,146$	$0,146$	$0,190$	$0,139$
	$(0,270)$	$(0,323)$	$(0,300)$	$(0,132)$	$(0,271)$
α_2	$0,859^{***}$	$0,853^{***}$	$0,853^{***}$	$0,809^{***}$	$0,860^{***}$
	$(0,233)$	$(0,287)$	$(0,265)$	$(0,125)$	$(0,233)$
Observations	234	234	234	234	234
Pseudo-R^2	0,439	0,453	0,448	0,452	0,438
p-value (dummies, jointly)	–	0,230	0,044	0,031	0,852

Note: Valid token count is scaled and centred. Robust standard errors in parentheses. The p-value for joint significance of the dummies is estimated by a Likelihood Ratio test against M1. A similar regression table for 2015-2024 is presented in Appendix 10.14.
*p<0,1; **p<0,05; ***p<0,01

In all model specifications the persistence, AR(1), is close to 1 and significant, possibly indicating near-unit root for polarisation. We revisit this in section 7. The shock adjustment, MA(1), is negative and highly significant, showing that shocks to polarisation tend to offset each other. The token count control is insignificant in all specifications M1-M5, supporting the conclusion that remaining bias in the LO-estimate of polarisation is unrelated to token count. This result holds when expanding the time-series to include 2015, which has fewer valid tokens, cf. figure 5. and Appendix 10.14. For the variance equation (eq. 10), both average variance and the ARCH-effect are insignificant while the persistence in volatility, α_2, is pronounced and significant across all specifications. This characterises a time-series in which volatility is clustered, supporting the graphical investigation of polarisation in section 5.6.

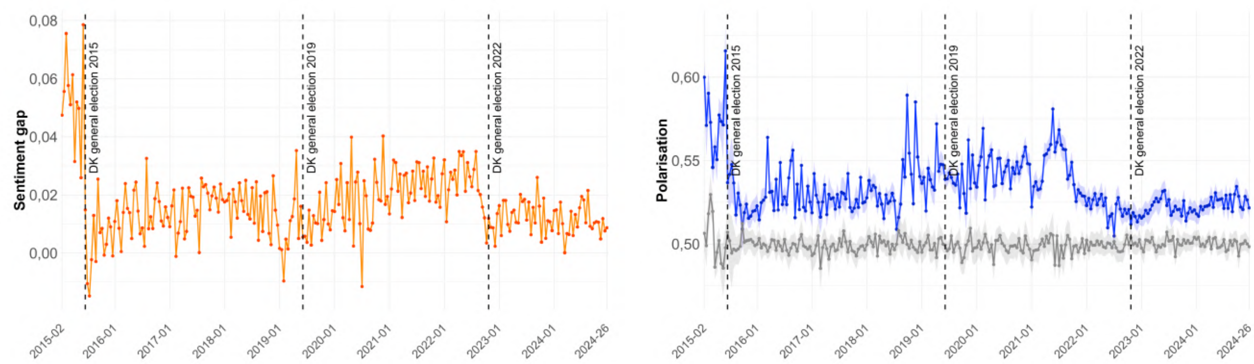
[15] We acknowledge that yearly fixed effects are not technically classified as seasonality effects, but we describe them as such for simplicity.

For each model specification M2-M5, we perform a Likelihood-Ratio test on the inclusion of the model-specific set of seasonal dummies and report the corresponding p-value in table 5. With p-values below $0,05$ for the set of seasonality dummies in M3-M4, the ARMAX(1,1)-GARCH(1,1) estimation shows that polarisation does vary with both quarters and years. There are no significant contributions to the model fit from including monthly effects, nor from including a check for periods with suspected low intensity of the news stream (biweek 15 and 26). This test rejects that polarisation is significantly different in these weeks vis-à-vis the rest of the year, while it remains an open question if and how the general news-stream differs in these weeks, considering that coverage across media can change, while polarisation remains constant.

Going forward, we include quarterly fixed effects when relevant to capture seasonal variation. We conclude that there is yearly variance in polarisation. As we do not go on to specify any models in which we need to control for yearly variance, this finding does not entail any adjustments to later models. We address potential concerns regarding inference from heteroscedasticity as they arise in the subsequent analyses.

6.3 Electoral cycles

The development in sentiment scores of articles published by the two media outlets over the time-horizon shows periodic expansions and contractions in the gap between the outlets, cf. figure 12. The gap itself is depicted in figure 25., panel (a), along with polarisation, panel (b). The timing of the general elections in the plots serves as suggestive evidence that the developments in potentially both time-series follow an electoral cycle. This observation motivates the following analysis, which could support the hypothesis that contemporaneous events drive polarisation (WH5). Though general elections are typically covered extensively in mainstream media, it is not immediately obvious why or how this appears to be directly traceable in sentiment scores of articles.



(a) Sentiment gap (b) Polarisation

Figure 25: Electoral cycles in the time-series

Hopmann et al. (2012) study political bias in news and find that the period leading into an election is characterised by a large degree of political coverage, with varying intensity and tone of coverage between political parties. This partly supports that media content changes in the pre-election period. Hopmann et al. relate to our study in the use of e.g. tone and empirical subject, but notably they do not investigate polarisation nor cyclical patterns. Overall, we find a limited theoretical link from political business cycles, to the field of media studies and polarisation.

As outlined above we observe periodic contractions in the sentiment gap immediately prior to an election, cf. figure 25. The cyclical pattern mimics what Schultz (1995) refers to as "opportunistic" political business cycles, where policy makers manipulate economic conditions in the run-up to elections to influence voter behaviour.

In order to test the hypothesis that article content reflects political cycles we borrow simple model specifications from Davidson et al. (1992). The link from electoral cycles to sentiment and polarisation may take the form of a level effect in the unconditional mean of the dependent variable and in the shape, form, and strength of the resulting dynamic describing the relationship between the cycle and the outcome. We adapt the notation of outcome/dependent variable here to highlight that we first establish the presence of the hypothesised electoral cycle component in the sentiment gap, after which we investigate to what degree it is present in the estimated polarisation measure. The sentiment gap and estimated polarisation follow stationary autoregressive processes with heteroskedasticity, cf. Appendix 10.12. We observe both a high parameter estimate of the AR term in table 5. and non-stationarity in polarisation for some, non-critical, lag-lengths and provide more detail on the implications this has for inference in section 7.

We employ two simple autoregressive models from the field of political business cycles to which we add a trend component (Davidson et al., 1992). The models are outlined in equation (11) and (12) below:

$$X_t = \beta_0 + \beta_1 X_{t-1} + \beta_2 Z_{\text{Election},\,t} + \left[\beta_3 + \beta_4 Z_{\text{Election},\,t}\right] t + \boldsymbol{\gamma}'\boldsymbol{\Omega}_t + \varepsilon_t \tag{11}$$

$$X_t = \beta_0 + \left[\beta_1 + \beta_2 Z_{\text{Election},\,t}\right] X_{t-1} + \left[\beta_3 + \beta_4 Z_{\text{Election},\,t}\right] t + \boldsymbol{\gamma}'\boldsymbol{\Omega}_t + \varepsilon_t \tag{12}$$

Both models take as input an indicator variable, $Z_{Election,\,t}$ with value 1 in some time-window prior to general elections. By evaluating β_2 in eq. (11) we establish whether an electoral cycle component explains differences in the unconditional mean of the outcome variable in an election window. This is refined in eq. (12) where the coefficient of the interaction term of the indicator variable and autoregressive component, β_2, allows for a different persistence in the AR dynamic in the election window. In both eq. (11) and (12), the coefficient on the interaction of trend and election window, β_4, captures deviations from the global trend within the election window.

In adopting the simple autoregressive models from political business cycles literature, we draw on an established econometric method to analyse electoral cycles. This methodological choice departs from the ARMAX(1,1)-GARCH(1,1) baseline specification previously used. This modelling choice deliberately omits the moving average and conditional heteroskedasticity components, thereby avoiding the risk that short-term shock adjustments absorb variation that may be attributable to electoral dynamics. The caveat being that this specification is likely mis-specified in terms of capturing the true error structure: as shown in Table 5., the MA(1) term contributes significantly to model fit, suggesting that some of the dynamics in the data may be better captured through past errors. Thus, while we retain the AR(1) component to model persistence, we assign limited interpretation to its coefficient, β_1, recognising that it may be partially capturing dynamics that an MA(1) term would otherwise model. As motivated above, we control for quarterly effects included in $\boldsymbol{\Omega}_t$. Year fixed effects are excluded in the models, as they risk over-controlling for time variation that is plausibly introduced by the electoral cycle. Considering the heteroskedasticity in both time-series, we report robust standard errors for all estimates. Reflecting that the time-series includes just three elections, one of which is in 2015, and that token count has no significant effect on estimates in section 6.2., we fit the models to the full time-series including observations from 2015.

Parameter estimates describing the level- and persistence effects of election cycles on the sentiment gap and polarisation are reported in table 6. In model M6-M9, we define an indicator variable for the window leading up to the election from the time of election and half a year (13 biweeks) into the past, $Z_{election}$. Model M6-M7 reports parameter estimates for eq. (11) showing a higher unconditional mean in the sentiment gap in the lead-up to an election compared to the rest of the electoral cycle. The parameter estimate of $Z_{election}$ in M6 is significant at a 5 pct. significance level and corresponds to a more than doubling of the unconditional mean of the sentiment gap prior to general elections[16]. We establish a similarly directed, albeit much smaller, effect on polarisation, where the conditional mean of polarisation is around $3,6$ pct. higher in the 13 biweeks leading up to a general election compared to the unconditional mean[17]. From figure 25., we note the initial high level of both polarisation and the sentiment gap in 2015, which likely fully explains the estimates of unconditional mean in the period leading up to the election. As mentioned, data from 2015 is more unreliable and we place limited interpretative value on the changes to the unconditional mean.

Table 6: Electoral cycles with a leading election window of 13 biweeks, 2015-2024

	Specification			
	M6: Eq (11)	M7: Eq (11)	M8: Eq (12)	M9: Eq (12)
Lag(1)	0,338***	0,453***	0,186**	0,455***
	(0,079)	(0,070)	(0,078)	(0,069)
$Z_{election}$	0,016**	0,019***		
	(0,006)	(0,007)		
Lag(1) $\times Z_{election}$			0,398***	0,031***
			(0,124)	(0,012)
Trend	0,00001	−0,00001	0,00001	−0,00001
	(0,00001)	(0,00001)	(0,00001)	(0,00001)
Trend $\times Z_{election}$	−0,0001**	−0,0001***	−0,00004**	−0,0001***
	(0,00004)	(0,00004)	(0,00002)	(0,00004)
Constant	0,009***	0,291***	0,012***	0,290***
	(0,002)	(0,037)	(0,002)	(0,037)
Quarterly effects	yes	yes	yes	yes
Dependent variable:	Sentiment	Polarisation	Sentiment	Polarisation
Observations	259	259	259	259
R^2	0,333	0,457	0,348	0,454

Note: Robust standard errors in parentheses. *p<0,1; **p<0,05; ***p<0,01

Model M8 and M9 report estimates for the eq. (12) specification. In the lead-up to a general election both the sentiment gap and polarisation appear to follow a more persistent

[16] With predictive margins for the sentiment gap at $0,015$ and $0,031$ when $Z_{Election}$ is at $(0,1)$, respectively, the sentiment gap is about 106 pct. higher during the defined election window than during non-election periods.

[17] Again, with predictive margins for polarisation at $0,533$ and $0,552$ for $Z_{Election}$ being $(0,1)$ the change corresponds to an increase of $3,6$ pct.

autoregressive process, as the parameter estimates on Lag(1)$\times Z_{election}$ are positive and significant at a 1 pct. significance level. No model specification in table 6. exhibits a global trend (the time-series are both stationary processes), which constitutes empirical evidence against WH4: that polarisation increases over time. However, in the defined electoral window both polarisation and the sentiment gap exhibit significant but numerically small negative trends. This finding holds across all model specifications M6-M9. In the lead-up to general elections, we have thus documented that both the sentiment gap and polarisation in the articles are higher than normal and that in the period prior to the election, both measures fall as the election approaches. The finding suggest that general elections in Denmark generally have a de-polarising effect in the news coverage which relates to research showing that polarisation remains high during an election period in America (Fasching et al., 2024).

In summary, we find that both polarisation and the sentiment gap follow electoral cycles. This reveals a clear common pattern in the lead-up to an election, and that is exists in both time-series suggest that sentiment and polarisation may have some shared dependency. Relating this to our findings in section 4.2.2., it strengthens the preliminary conclusion that variations in sentiment across topics may contribute to polarisation. We observe both significant level and trend effects for the sentiment gap and polarisation. However, we interpret the level effect with caution, as it appears to be driven by unusually high levels in 2015. The negative trend effect leading up to elections suggests an increase in the overlap of content and tonality of DR and TV2 news coverage. This finding supports WH5: that contemporaneous events drive polarisation. However, as elections are recurring and anticipated events, they differ from sudden external events or shocks to the system, which we focus on in the subsequent sections 6.4. and 6.5.

6.4 Regime change

We have in section 5.6. suggested that, in recent years, a regime change in polarisation may have occurred, leading to a period of both low polarisation and low volatility. In the following section, we analyse this period of moderation more systematically and examine whether it could be caused by a single event. This would further support WH5: that contemporaneous events drive polarisation. To do so, we employ a regression discontinuity in time approach and expand the findings using a change-point analysis of volatility.

6.4.1 Regression discontinuity in time

We note that the onset of moderation coincides, to some extent, with the invasion of Ukraine in biweek 4 of 2022. Using a regression discontinuity in time framework, we formally assess this relationship. This allows us to test whether the observed changes in polarisation can be attributed to the timing of the invasion.

Grouping polarisation estimates for biweeks in bins and fitting trends conditional on the timing of the invasion graphically reveals the suspected level differences in the unconditional mean of polarisation. This is depicted for the general LO-polarisation estimates from 2016-2024 in figure 26. It can be traced back to the sections *Indland* and *Other* with varying differences in the level around the event timing. For the sections *Politik* and *Udland*, the event appears to explain changes in the squared trend only, cf. Appendix 10.13.

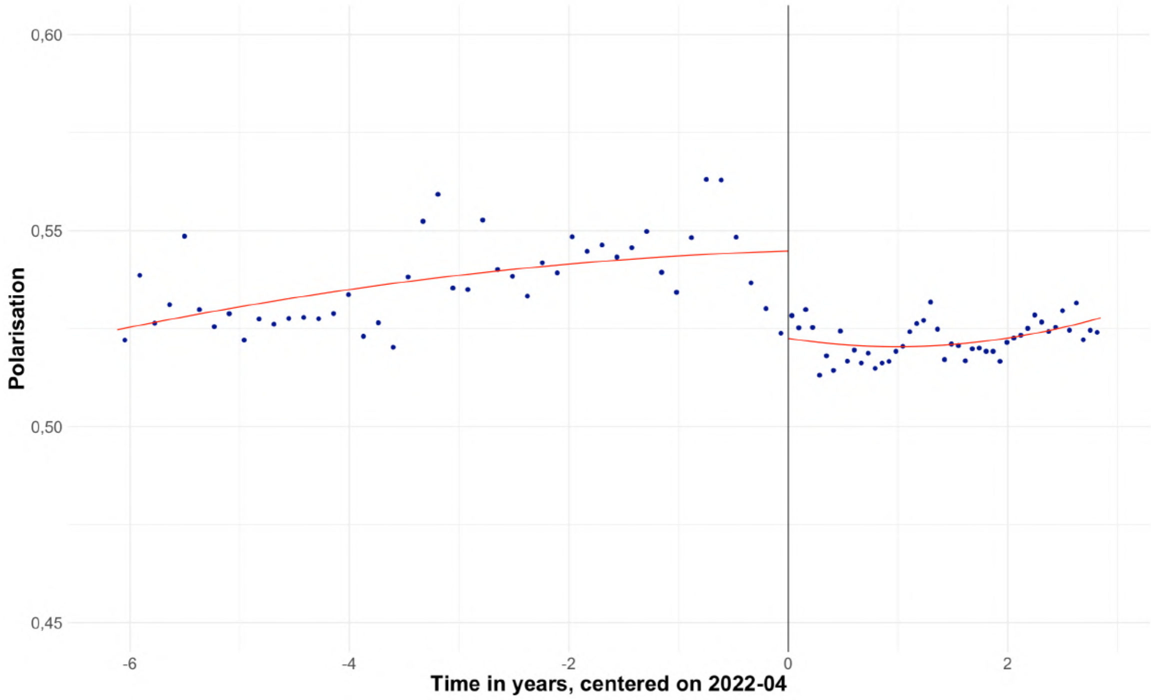


Figure 26: Discontinuity around 2022, biweek 4. Time-span 2016-2024.

We define a window of approximately six years around the invasion in biweek 4 of 2022. This window consists of the three years leading up to the invasion, and the rest of our time-frame after (close to three years). This fully contains the period of moderation and importantly captures the higher volatility characteristic of the period prior to the event. We adopt a classic regression discontinuity (RD) approach to a time-series specification (RDiT) (Angrist and Pischke, 2008; Arutyunov, 2025). The application of RDiT reflects two key features of the data: that for each point in time we have a single observation of polarisation and that we can define a sharp cut-off after which we consider all observations as treated.

The use of RDiT assumes smoothness of potential time-varying confounders around the threshold that determines treatment status. This assumption is critical, as it requires the conditional expectation of both the outcome and potential confounders to vary continuously with the running variable time. We argue this assumption is likely satisfied for two reasons. First, we have no *a priori* reason to believe that any unobserved confounders change discontinuously at the event threshold. Second, the frequency of the polarisation data is relatively high. With a high frequency of observations, the risk that the expectation function is not continuous and differentiable around the threshold decreases. Additionally, employing a regression discontinuity in time model on an autocorrelated time-series naturally risks introducing autocorrelation in the residuals. To account for this we apply robust standard errors.

Given the uniform distribution of time, we cannot implement any standard tests for e.g. manipulation. This implies that we address potential concerns of anticipation, avoidance, etc., informally: as we document later, the data appears to show some anticipation of the event. We can justify this immediately prior to the event, given how coverage of related topics increases prior to the event, cf. figure 11. However, we can graphically trace this increased coverage back only four biweeks, to the start of 2022. This does not fully account for the suspected anticipation we observe, the implications of which we discuss later in the section. Considering the nature of the data, other forms of potential manipulation, for instance sorting or selection, are non-intuitive. If the media were to engage in some form of sorting - by e.g. withholding articles with a distinct token composition that influences polarisation - this would be a feature of the polarisation measure itself rather than a confounding factor.

We can infer causality of the regression discontinuity in time when the assumptions above hold. In a classic RDiT-design, it is standard to perform a donut-regression, omitting observations around the threshold to combat potential bunching/selection. This is used as a robustness check and if the parameter estimates hold when omitting these potentially critical observations, it strengthens the inference of the results. It is, however, not straightforward to draw the same conclusions in our analysis of polarisation. If the results of a donut-regression strengthen the later results, it would imply that the discontinuity is stronger as we omit the observations around the event. Considering that we have no suspected manipulation, this result would reflect that the regime change might predate the event. In that case, a strong result of a donut-regression presents a challenge to, rather than a support for, inference.

We estimate the treatment effects using the RDiT model proposed in eq. (13). In this model we include the unconditional mean, β_0, effects from both linear and quadratic trends β_1 and β_2, as well as quarterly effects, $\boldsymbol{\gamma}'$. We allow the event to affect the level of polarisation, β_3 (average treatment effect on impact), and both the linear and quadratic trend of polarisation, β_4 and β_5, respectively. We define a running time variable, t, and define the event dummy, $Z_{Event,t} = 1$ for all t after the Russian invasion of Ukraine in biweek 4 of 2022.

$$X_t = \beta_0 + \beta_1 t + \beta_2 t^2 + \beta_3 Z_{Event,t} + \beta_4 Z_{Event,t} t + \beta_5 Z_{Event,t} t^2 + \boldsymbol{\gamma}' \boldsymbol{\Omega}_t + \varepsilon_t \tag{13}$$

In table 7., we present parameter estimates for eq. (13) when estimated on general polarisation and for each of the four sections. The average treatment effect on impact in general polarisation at the cutoff is negative and significant at a 5 pct. significance level. When comparing predicted values exactly at the threshold with $Z_{Event,t} \in \{0, 1\}$, the average treatment effect on impact corresponds to a drop in general polarisation from $0,534$ to $0,518$, i.e. a difference of $-2,9$ pct.[18]. There is a significant, negative average treatment effect on impact in the sections *Indland* and *Other*, with the effect in *Indland* being slightly more pronounced. The conditional mean of polarisation in both *Politik* and *Udland* is unchanged at the event time, however, for *Udland* exclusively, there is a significant increase in the linear trend. In the general polarisation there is a significant increase in the squared trend following the event. This comes from the sections *Udland* and *Politik*. In fact, there is no significant change in any trend for the sections *Indland* and *Other*. Taken together, this suggests that the event causes a drop in mean general polarisation, driven by level reductions in mean polarisation of *Indland* and *Other*. The increase in the squared trend appears to reflect changes in the curvature of the trend within *Politik* and *Udland*. To the degree that we can ascribe the effect to the event, we conclude that it leads to varying shifts in polarisation across sections.

[18]Using $\frac{\hat{X}_{0+} - \hat{X}_{0-}}{\hat{X}_{0-}} \times 100$ with $\hat{X}_{0+} = 0,518$ and $\hat{X}_{0-} = 0,534$ returns a drop in the conditional mean of $2,9$ pct.

Table 7: Regression discontinuity at biweek 2022-04, by section

	Section:				
	General	Indland	Udland	Politik	Other
t	$-0{,}013^{*}$	$-0{,}010$	$-0{,}032^{***}$	$-0{,}024^{**}$	$-0{,}003$
	(0,007)	(0,009)	(0,005)	(0,011)	(0,009)
t^2	$-0{,}005^{**}$	$-0{,}003$	$-0{,}009^{***}$	$-0{,}007^{**}$	$-0{,}001$
	(0,002)	(0,003)	(0,002)	(0,003)	(0,003)
$Z_{Event,t}$	$-0{,}016^{**}$	$-0{,}024^{***}$	0,004	0,008	$-0{,}023^{***}$
	(0,007)	(0,008)	(0,005)	(0,010)	(0,008)
$Z_{Event,t} \times t$	0,010	$0{,}017^{*}$	$0{,}028^{***}$	0,014	$-0{,}002$
	(0,008)	(0,010)	(0,006)	(0,012)	(0,011)
$Z_{Event,t} \times t^2$	$0{,}007^{**}$	0,002	$0{,}012^{***}$	$0{,}011^{***}$	0,005
	(0,003)	(0,003)	(0,002)	(0,004)	(0,004)
Constant	$0{,}534^{***}$	$0{,}535^{***}$	$0{,}509^{***}$	$0{,}509^{***}$	$0{,}549^{***}$
	(0,005)	(0,007)	(0,004)	(0,008)	(0,007)
Quarterly effects	yes	yes	yes	yes	yes
Observations	153	153	153	153	153
R^2	0,635	0,477	0,503	0,102	0,485

Note: Robust standard errors in parentheses. $^{*}p < 0{,}1$; $^{**}p < 0{,}05$; $^{***}p < 0{,}01$

A concern is whether there is some anticipation in polarisation. Intuitively, we cannot justify any anticipation earlier than a few biweeks prior to the event. Therefore, if the period of moderation we have noticed graphically starts earlier than the event takes place, we fail to adequately justify that the assumptions above hold and that the event causes the changes we observe. We see this challenge to inference in figure 26., where the event coincides with the start of the period of moderation, but where some observations prior to the event appear to belong to the new regime. The transition from one regime to the next indeed appears to begin in the last quarter of 2021 rather than in the beginning of 2022.

We perform a donut-RD, which is analogous to the specification previously used, but omits eight biweeks before and after the event. This effectively places the event after biweek 21 in 2021 - far earlier than any spike in mentions of related topics, cf. figure 11. The results are presented in table 8., column RD-donut, and show a more pronounced average treatment effect on impact in the general polarisation, with higher significance than in the standard specification in table 7. The fact that we have no reason to suspect anticipation so far from the event and that the treatment effects increase in conjunction implies two preliminary conclusions: firstly, it is increasingly unlikely that the true event causing the effect is located in biweek 4 of 2022 and, secondly, that the changes - regardless of the true cause - are persistent across time. If the changes in polarisation due to the invasion were transitory, the treatment effect would likely decrease as we omit the observations right around the event, however, we note the opposite.

Table 8: Donut and Placebo RD

	Specification:	
	RD-Donut	Placebo RD
t	0,004	0,021
	(0,0095)	(0,0132)
t^2	0,000	0,005
	(0,0026)	(0,0041)
$Z_{Event,t}$	$-0{,}034^{***}$	$-0{,}048^{***}$
	(0,0085)	(0,0104)
$Z_{Event,t} \times t$	$-0{,}001$	$-0{,}043^{***}$
	(0,0108)	(0,0141)
$Z_{Event,t} \times t^2$	0,000	0,003
	(0,0030)	(0,0044)
Constant	$0{,}547^{***}$	$-0{,}597^{***}$
	(0,0080)	(0,0099)
Quarterly effects	yes	yes
Observations	138	157
R^2	0,683	0,646

Note: Robust standard errors in parentheses. $^{*}p < 0{,}1$; $^{**}p < 0{,}05$; $^{***}p < 0{,}01$

The former preliminary conclusion - that the true event is unlikely to coincide with the timing of the invasion - indicates that the true threshold lies somewhere prior to the previous definition of $Z_{Event,t}$. To test this, we adopt a placebo test. The placebo test, where a fake threshold is placed at some time-point not related to the actual event, is a common robustness check in regression discontinuity (Arutyunov, 2025). Specifically, we place a placebo-treatment dummy in biweek 22 of 2021 and again define the estimation window as six years centred around the event. The results are reported in table 8., column Placebo-RD. As with the donut-RD, we note a stronger, and more significant, negative treatment effect compared to the original specification, cf. table 7., effectively proving that the true event is not located in biweek 4, 2022.

Naturally, this raises the question of whether the period of persistent moderation, which the regression discontinuity supports, is actually attributable to any one specific event. On the contrary, considering that both the placebo and donut specification of the RDiT increase predictive power, it appears that polarisation is quite indifferent to influence from specific events.

6.4.2 Change-points in volatility

The observed period of persistent moderation implies both a lower mean polarisation, as the RDiT results support, and a lower volatility in polarisation in recent years, which we, too, can investigate empirically. If the data shows not only a lower mean polarisation but also a distinctly lower volatility-regime starting at the same time, it supports that there is some foundational and persistent shift in how distinctly media present news in recent years. That the new dynamic is sustained across time suggests that it is likely not attributable to a singular event, however major, but may stem from a cumulative effect of recurring events that continuously exert downward pressure on polarisation. To validate this, we check for the existence of a volatility regime distinct to the last years of the polarisation data[19].

[19]With slight abuse of language, we use volatility and variance interchangeably in this section.

Starting off, we remove observations from 2015, take 1st differences, and compute a rolling average (RA) for polarisation using a window of 3 biweeks, centred. From this we obtain point-wise 95 pct. confidence intervals for the time-series. Based on the 1st differenced general polarisation, we perform a change-point detection by minimising eq. (14).

$$\min_{\{\tau_1,\dots,\tau_m\}} \sum_{i=1}^{m+1} [\mathcal{C}(y_{(\tau_{i-1}+1):\tau_i})] + \beta f(m) \tag{14}$$

$$\mathcal{C}(y_{a:b}) = (b-a+1)\log(\hat{\sigma}^2_{a:b}) \tag{15}$$

Equation (15) defines the cost of a regime $(a:b)$, which is scaled by the natural logarithm of the regime-specific variance, $\hat{\sigma}^2_{a:b}$ (Killick et al., 2012). This is solved using the Pruned Exact Linear Time algorithm (PELT) (Killick and Eckley, 2014; Killick et al., 2024). We employ PELT, reflecting that it is relatively straightforward to implement and that it allows for change-point detection in variance specifically. The application is trivial, but reflecting features of the polarisation data, we enforce a manual penalty parameter of $\beta = 7$ and a minimum regime length of $\mid b-a \mid \geq 20$ biweeks. A relatively long minimum regime length lowers the risk of identifying false positives, while the moderately conservative penalty parameter reduces the number of identified change points.

The change-point detection in variance returns six distinct regimes in volatility demarcated by four change-points. These are located in 2017-4, 2018-18, 2020-7, and 2022-16, where the last digits represent the biweek. Out of the four change-points, only biweek 7 of 2020 immediately coincides with a major event that we suspect could influence the composition of the news stream fundamentally: the first lockdown in Denmark was announced in biweek 6 in 2020. We separately analyse the short-term effects of this event on polarisation in section 6.5. We take note of the last detected change-point, where conditional variance drops, as we will go on to show below. It indicates a shift 12 biweeks after the invasion in biweek 4 of 2022, and lies more than six months after the timing used in the placebo-test. While there is a partly mechanical explanation for why the level and variance cannot drop simultaneously, this indicates that the shift in volatility trails the shift in level by some margin.

Within each regime we calculate upper and lower variance bands as mean ± standard deviation (SD). In figure 27., we illustrate the estimated change-points with bands on variance within each regime. The figure includes the 1st differenced time-series, the RA and its pointwise confidence intervals, along with the timing of the Russian invasion of Ukraine. The beginning of the series is characterised by high volatility, which from 2017-4 to 2018-18 is replaced by a lower volatility regime. Comparing figure 27. to the general polarisation, cf. figure 21., this corresponds to a period in which the level of polarisation remains approximately the same while stability improves. This regime is then superseded by a two year period, up until 2020-7, with historically high volatility. In this period, we observe the largest biweek-on-biweek shocks to polarisation in the time-series. At the change-point in 2020-7, volatility drops to the same low level as immediately prior to 2018-18. Volatility is further reduced to a historically low level at the change-point in 2022-16. This regime continues until the end of the time-series. The standard deviation in the last regime is ~ 50 pct. lower than the standard deviation in the previous regime and almost 80 pct. lower than the average standard deviation in the regime from 2018-18 to 2020-7[20].

[20]With standard deviations in the regimes $m \in \{3,4,5\}$ being $SD_{m=3} = 0,023$, $SD_{m=4} = 0,010$, and $SD_{m=5} = 0,005$, the $SD_{m=5}$ is $49,1$ pct. lower than $SD_{m=4}$ and $77,6$ pct. lower than $SD_{m=3}$.

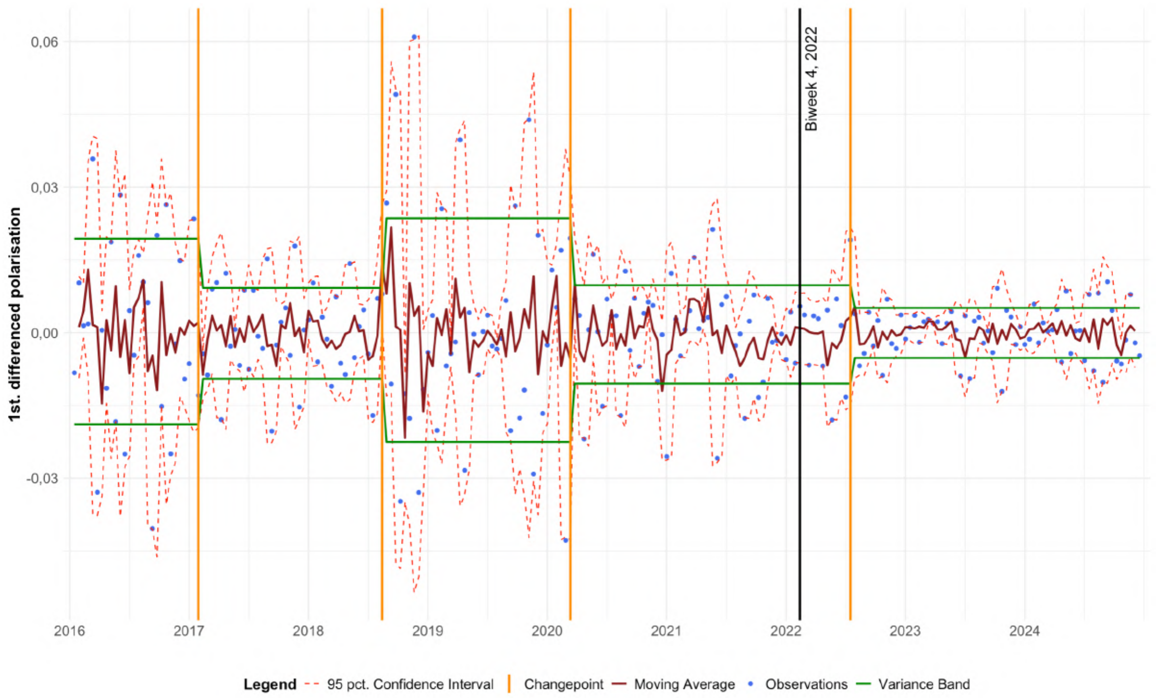


Figure 27: Change-points in variance, 2016-2024

The change-point detection establishes that in recent years there exists a historically anomalous volatility regime with very low fluctuations in observed polarisation compared to previous years. Placing the results of the change-point detection in conjunction with the results of the regression discontinuity analysis above, we document that polarisation is atypical in both level and volatility in the last years of the time-series. While polarisation changes along both axes, we fail to relate any of the changes to specific events. We cannot with certainty reject that this is due to investigating an event that has no true effect on polarisation, but considering the profound changes the event has on the news composition, as outlined when motivating the RDiT, it appears increasingly likely that polarisation is not explained by singular events in general. By extending the analysis of changes in the level/trend using RDiT to changes in volatility using PELT, we have further strengthened the case that at least the invasion of Ukraine, and possibly singular events in general, have little direct effect on polarisation.

An increasingly likely interpretation is that polarisation is the observed outcome of a high-dimensional news stream in which various events each have a relatively low influence on the overall polarisation at a given time. In this light, we can interpret the results of the regression discontinuity and change-point analysis as the invasion marking a shift in the geopolitical reality, which has had pronounced effects on the composition of the news stream in the two media, together with an unobserved number of other events, possibly reinforcing each other. Placing the invasion as a news story in a context of other news stories that were also prevalent at the time, and even predated the event, such as stories about inflation, general geopolitical instability, and a lingering trade war with China, suggests that the period of moderation that the RDiT supports is attributable to a new geopolitical reality. Theories of "rally around the flag" (Steiner et al., 2023; Baker and Oneal, 2001) lend credibility to this interpretation, though we cannot with certainty establish a causality between geopolitics and news content.

From section 6.4. we draw two main conclusions. Firstly, we find both a regime of historically low volatility and a significant, negative level effect on polarisation in the last three years of our time-frame, which leads us to definitively reject WH4: that polarisation increases over time. Secondly, when examining specifically the coverage of the invasion of Ukraine, we have found evidence against WH5: that contemporaneous events drive polarisation. Therefore, we propose an alternative interpretation: that polarisation is shaped not by singular events, but rather by the sum of continuous events. In the following section, we focus more broadly on short-term effects of singular events.

6.5 Short-term impact of events

Motivated by WH5, we are specifically interested in investigating whether specific events in our time-frame have an explicit impact on polarisation. In the analyses above, we have established the effect on polarisation from electoral cycles, the moderation in polarisation levels and volatility in recent years, and shown distinct regimes in volatility. To understand in detail if and how DR and TV2 cover major global and domestic events differently, the gold standard would be to use randomised controlled trials (RCTs) on polarisation. An RCT uses randomised assignment of treatment to a subset of a group, to allow comparison of the behaviour of the treated to a control group that has not been exposed to treatment. In our context, an RCT is infeasible due to data constraints and we propose a weaker event study design to capture the short-term effects of singular events.

The objective is to unveil event-specific channels contributing to polarisation. Our aim is to test whether specific events can impact polarisation, why we include a broad set of candidate events, where some, naturally, will be more intensely covered than others. We choose events to test for short-term impact on section-specific polarisation in contrast to section 6.4., where we examined sustained effects of one specific event. Further reinforcing this short-term approach is the non-controversial assumption that news media and their editorial lines are influenced by a large number of factors, from different competition strategies to global events, domestic events, public service obligations, and others, which all potentially contribute to shaping polarisation. Therefore, we assume that the effect of a singular event on polarisation is limited to a short-term impact. We analyse the effects on section-specific polarisation based on which section the singular events are most likely to be covered in.

In the following we employ segmented regression (SR) on interrupted time series (ITS) data (Bernal et al., 2016). To obtain interrupted time series data, we limit our time-frame to 26 periods before and 12 periods after the event, where each event is classified as an intervention at a clear-cut date. This leaves a well-defined pre- and post- period. Prior to inspecting the data around each event, we propose an impact model. Proposing model design prior to data inspection ensures that the model design truly reflects prior theoretical expectations of how the event shapes polarisation (Bernal et al., 2016). In our case, the impact model allows for changes in level and trend in the full period following the event, as opposed to, e.g., a model capturing temporary level change, etc. We choose this formulation precisely because we expect the events to shape polarisation in both trend and level within the full post-treatment window. We utilise a one-week specification of our LO-estimator to generate weekly polarisation. This doubles the data-point frequency compared to using biweekly data. A sufficient number of data-points is important in a well-defined segmented regression design of short-term effects (Bottomley et al., 2019), but using weekly estimates entails a trade-off between frequency and reliability. As we have shown in section 5.6., when the token set is small, uncertainty is large.

Here, we place high emphasis on short-term effects, why we accept larger uncertainty in order to capture week-on-week changes in polarisation. It does not increase bias of the estimates, however, cf. section 6.2., where we documented that the valid token count has no significant effects on polarisation.

The SR-design on ITS-data entails a set of assumptions that mimic those of the regression discontinuity design in section 6.4. Most importantly, we assume smoothness of the unobserved confounders around the event. We cannot test this directly in our setup, but by including an unsystematic, diverse set of candidate events, we reduce the likelihood that unobserved confounders change sharply in all events. Looking at the broad picture, not focusing on the singular event effects but on effects across all or most events, provides a guardrail against this. Moreover, we assume the model captures all underlying, relevant trends in the pre-event window. This differs slightly from the assumptions in the RDiT, where we allowed for fundamental changes to the underlying trends. The distinction being that in the RDiT we expect changes to the underlying trend, whereas in the SR-design we restrict regression window to a period shorter than a year, and we expect minimal changes in underlying trend.

We propose eq. (16) to estimate the intervention effect of the listed events. In our proposed impact model β_0 and β_1 captures the level prior to the event the pre-event trend, respectively. The level effect after the event is captured by β_2 while the change in slope after the event is captured by β_3. We define a running time variable t, the event dummy, $Z_{Event,t} = 1$ for the event and the following weeks, and a second running time variable with event week $t_{Post} = 1$.

$$X_t = \beta_0 + \beta_1 t + \beta_2 Z_{Event,t} + \beta_3 Z_{Event,t} t_{Post} + \varepsilon_t \tag{16}$$

The segmented regression (SR) approach is applied to eight singular events within our timeframe, each of which can be considered to have had a major impact on specific news sections. Each event is labelled SR(X) and the results from the segmented regressions are shown in table 9. and plotted in figure 28. We employ robust standard errors in line with our previous findings from section 6.2. The figure includes the estimated polarisation for an additional 26 weeks before and after each regression window to illustrate the context of the short-term effect.

We define a set of eight candidate events distributed over the time-horizon, which we motivate individually in the following. For each candidate event, we connect it to the section we expect it to affect the most. The first event, SR(1), is the Brexit referendum, held on 24$^{\text{th}}$ of June 2016. As this event includes 2015 in the pre-event period, we estimate its effect on general polarisation, since section-specific polarisation estimates are only available from 2016, cf. section 5.6. The second event, SR(2), is the election of Donald Trump as President of the United States on 6$^{\text{th}}$ of November 2016. In this case, we focus on changes in section-specific polarisation within *Udland*. Third, we examine the culmination of the Fridays for Future movement, exemplified by Greta Thunberg's speech at the United Nations on 23$^{\text{th}}$ of September 2019, SR(3). This event is expected to affect coverage within *Indland*, as the speech provoked polarising reactions in the public debate (Broberg, 2024). The fourth event, SR(4), is the announcement of the first COVID-19 lockdown in Denmark, delivered at a press conference on 11$^{\text{th}}$ of March 2020. Here, we analyse its impact on polarisation in the section *Indland*, and note that the timing coincides with that of a change-point in volatility, as mentioned in section 6.4.2. The fifth event, SR(5), also linked to *Indland* coverage, concerns "the mink case": the nationwide order to cull all mink issued on 4$^{\text{th}}$ of November 2020, which was subsequently revealed - two days later - to lack legal authority, triggering a political scandal (Den Store Danske, 2024). The sixth event, SR(6), is the storming of the U.S. Capitol on 6$^{\text{th}}$ of January 2021, which we examine in relation to polarisation in *Udland*. The seventh event, SR(7), is

Russia's invasion of Ukraine on 24 February 2022, for which we again assess changes in the *Udland* section. Finally, the eighth event, SR(8), is the Hamas terrorist attack on Israel on the 7$^{\text{th}}$ of October 2023 which reignited the longstanding Israeli-Palestinian conflict[21]. We anticipate this event to influence section-specific polarisation within *Udland*.

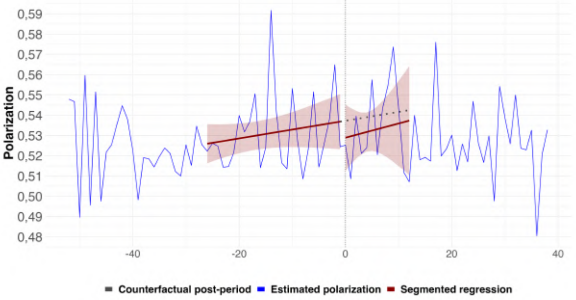
Table 9: Segmented Regression Results

	SR(1)	SR(2)	SR(3)	SR(4)	SR(5)	SR(6)	SR(7)	SR(8)
Intercept	0,525***	0,513***	0,557***	0,508***	0,537***	0,536***	0,509***	0,518***
	(0,005)	(0,006)	(0,012)	(0,014)	(0,006)	(0,007)	(0,010)	(0,003)
t	0,000	–0,000	–0,001	0,002*	0,000	0,000	0,000	–0,000*
	(0,000)	(0,000)	(0,001)	(0,001)	(0,000)	(0,000)	(0,001)	(0,000)
Z_{Event}	–0,009	0,007	–0,006	0,001	0,019*	–0,030**	0,012	0,009
	(0,012)	(0,013)	(0,023)	(0,020)	(0,010)	(0,009)	(0,007)	(0,009)
$Z_{Event} \times t_{Post}$	0,000	0,000	0,001	–0,004**	–0,002*	0,001	–0,002*	–0,000
	(0,002)	(0,002)	(0,003)	(0,001)	(0,001)	(0,001)	(0,001)	(0,001)
Polarisation	*General*	*Udland*	*Indland*	*Indland*	*Indland*	*Udland*	*Udland*	*Udland*
Observations	39	38	39	39	39	39	39	39
R^2	0,028	0,029	0,130	0,142	0,106	0,405	0,062	0,126

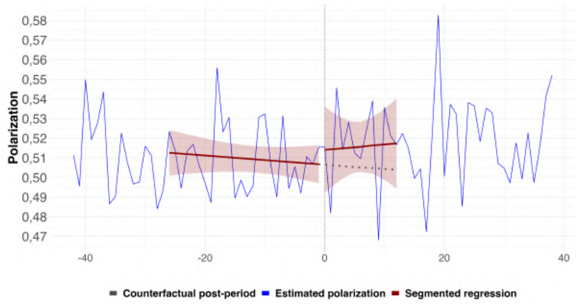
Note: Robust standard errors in parentheses. $^{*}p < 0{,}1$; $^{**}p < 0{,}05$; $^{***}p < 0{,}01$

In table 9. we present the parameter estimates from all eight segmented regressions, SR(1)-SR(8). The intercept in all regressions is significant at a 1 pct. significance level with coefficients ranging from $0,508$ to $0,557$, well in line with the levels presented in section 5.6. In SR(1)-SR(8) the general trend is insignificant at a 5 pct. significance level. In all regressions but SR(6) there are no significant level effects following the event. In SR(6) the coefficient on the level effect is relatively large and negative at $-0,03$. It is significant at a 5 pct. significance level, indicating that at the event, the conditional mean drops notably. The drop is visually distinct in figure 28., panel (f). In SR(1)-SR(3) and SR(5)-SR(8) none of the coefficients on the change in trend after the event are significant at a 5 pct. significance level. In SR(4) the parameter estimate on the change in trend is $-0,004$ and significant. Again, in figure 28., panel (d), the change is clearly visible. We take note of the fact that for SR(5) the level effect is significant at a 10 pct. significance level and the same holds for the change in slope for SR(5) and SR(7). We consider these results too weak to conclude on them. In all regressions but SR(6), where there is a significant and negative level effect the models explain relatively little of the variance in the ITS, especially considering how short the time-series is. For the two models where we note significant and large changes after the event (SR(4) and SR(6)), figure 28. reveals that if the short-term changes are indeed attributable to the events, then they are relatively short-lived, as the series quickly converge to the pre-event mean. We conclude that there appears to be no systematic evidence that the set of candidate events shape short-term polarisation in the relevant sections. We find some effects for two of the events, but fail to establish an overall pattern across SR(1)-SR(8).

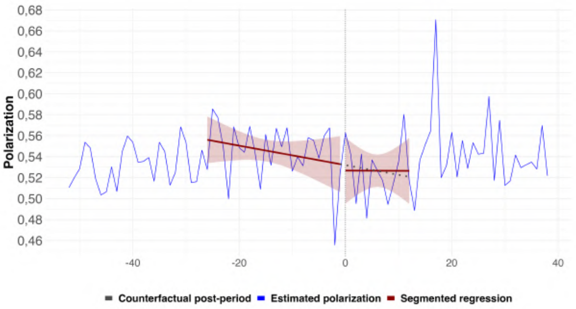
[21] We recognise that this wording does not even attempt to capture the scale of the atrocities in the aftermath of the attack.



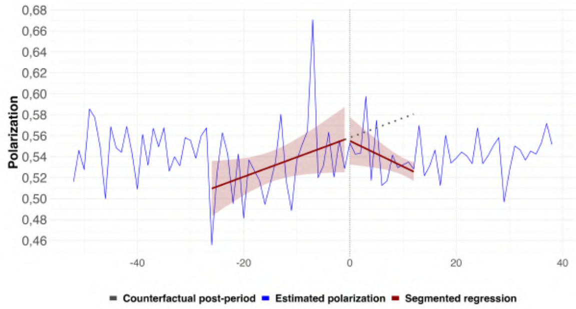
(a) SR(1), General polerisation



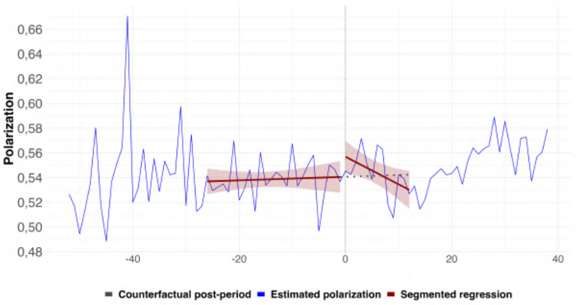
(b) SR(2), *Udland* polarisation



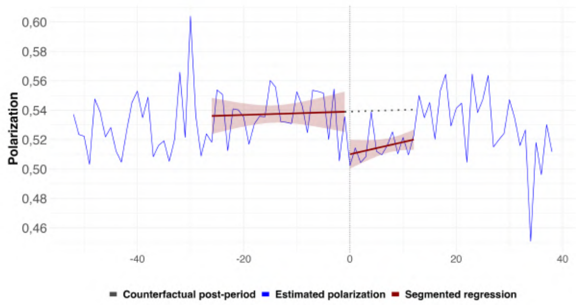
(c) SR(3), *Indland* polarisation



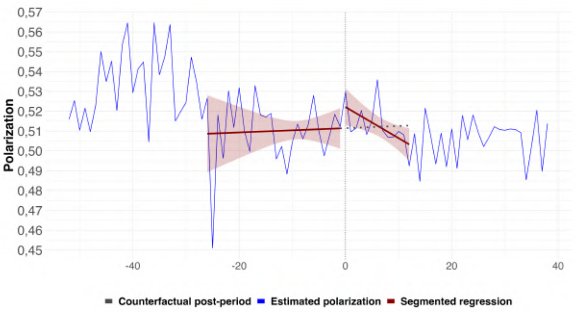
(d) SR(4), *Indland* polarisation



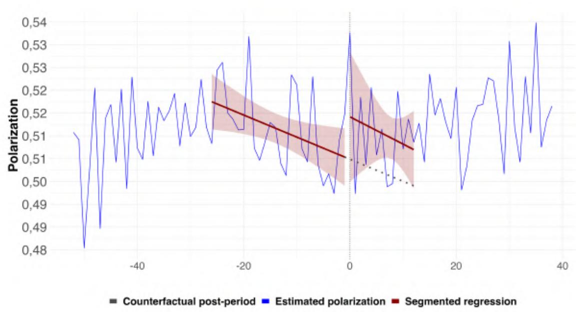
(e) SR(5), *Indland* polarisation



(f) SR(6), *Udland* polarisation



(g) SR(7), *Udland* polarisation



(h) SR(8), *Udland* polarisation

Figure 28: Short-term event effects on polarisation

In the robustness section 7.2.3., we conduct a placebo test, by shifting the events six weeks earlier to assess whether the observed effects are truly event-driven or if they reflect spurious variations at a convenient point in time. We have no reason to believe that there would be any other relevant events systematically predating the candidate events by 6 weeks. This check does not challenge the results presented above. The results presented here do not, in general, support the presence of systematic short-term effects from singular events on polarisation. While some of the candidate events appear to have short-term effects, these appear to be isolated cases rather than evidence for a consistent pattern. The most influential of these appears to be the announcement of the first lockdown in Denmark, modelled by SR(4), which might be an example of an event that has an explicit impact, as we observe impacts to both variance and level/trend in the biweek following the event. Out of the eight candidate events, however, only two show short-term effects that are significant at a 5 pct. significance level. These results offer only limited support for WH5, which posits that contemporaneous events drive polarisation.

7 Robustness

In the following, we revisit firstly three aspects of the robustness of the estimates of polarisation presented in section 5.6., and then secondly perform robustness checks on the results of the analysis in section 6.

7.1 Robustness of estimated polarisation

For the estimates of polarisation, we revisit separately the assumptions and criteria regarding the LO-estimate of polarisation. We do this by estimating polarisation using different sets of criteria for the J-vector in section 7.1.1. In section 7.1.2. and 7.1.3., we examine the implications of the frequency of polarisation estimates and the width of the confidence intervals, respectively.

7.1.1 Determining the criteria for the J-vector

In section 5.5., we outline the token filtering process. The second filtering step, excluding all tokens that are mentioned less than three times for period t, remains unadjusted in the following, as it is necessary for the LO-estimator to yield a valid output. The criteria of the first filtering step are not theoretically fixed, therefore the determination of the criteria reflects a "weighted" decision addressing both computational load from an excessive amount of tokens and validity from either too few tokens or selective tokens being filtered out. We test different specifications of the following three criteria in order to find an optimal criteria-combination:

C.1. Minimum number of times a token is used across all periods.

C.2. Minimum number of unique periods where a token has to be used at least once.

C.3. Share of top-frequency tokens excluded based on the total number of times tokens are mentioned across all periods.

To be able to compare the effects of the three criteria, we estimate polarisation using a subset of the possible combinations of the following criteria C.1.$\in \{25, 50\}$, C.2.$\in \{2, 5, 10, 25\}$, and C.3.$\in \{0,1 \text{ pct.}, 0,5 \text{ pct.}, 1 \text{ pct.}, 5 \text{ pct.}\}$.

In figure 29., we plot estimates in three years - 2016, 2019, and 2022 - for nine different combinations of the criteria above. We observe that polarisation exits regardless of the criteria combination. We conclude that polarisation is a feature inherent in the tokens themselves and not a reflection of design choices when determining the J-vector.

Figure 29: Specifications of the J-vector and the resulting LO-estimated polarisation for the subset 2016, 2019 and 2022

The specific combination of criteria used to construct our polarisation, as mentioned in section 5.5., is 50 x 10 x 0,5 pct. In figure 29., this iteration is marked as slightly larger pink dots. Compared to the other combinations, we have chosen a conservative level, as to not overstate polarisation. Furthermore, we observe that the underlying fluctuations between the combinations are the same, but the choice affects the level and scale of the estimated polarisation. Another important aspect is that the control polarisation has to be fairly stable around 0,5. This is true for most of the combinations, but the fluctuations vary in scale. We observe that control polarisation for the preferred set of combinations fluctuates less than most of the other sets. Amongst these, a subset of combinations exhibit fairly similar output at the conservative end, and we decide one of these as the preferred combination. Choosing a criteria set that yields conservative estimates is non-controversial and we do not conduct further tests in this regard.

7.1.2 Interval length impact on polarisation

There is no theoretical upper limit to the frequency of estimated polarisation. By defining shorter and shorter time intervals, t, we increase granularity of the estimates. The caveat being that when the intervals shorten, the token input reduces, inferring less reliable estimates. Analysing polarisation on intervals spanning two weeks offers a sensible trade-off between reliability and frequency. We demonstrate this in figure 30., where we plot polarisation in 2024 estimated using weekly, biweekly, and monthly intervals. Mechanically, as intervals increase in length, we observe less volatile estimates with narrower confidence intervals. Confidence intervals are narrow for both biweekly and monthly intervals. When using weekly intervals the confidence intervals are much wider, and for some weeks overlap with control polarisation, though biweekly and monthly estimates suggest that polarisation is significantly positive. We observe no level differences in the estimated polarisation, again supporting that the LO-estimator yields unbiased estimates even for very limited token input. These aspects motivate that we rely on biweekly estimates of polarisation in the analysis, apart from section 6.5., where frequency is particularly critical.

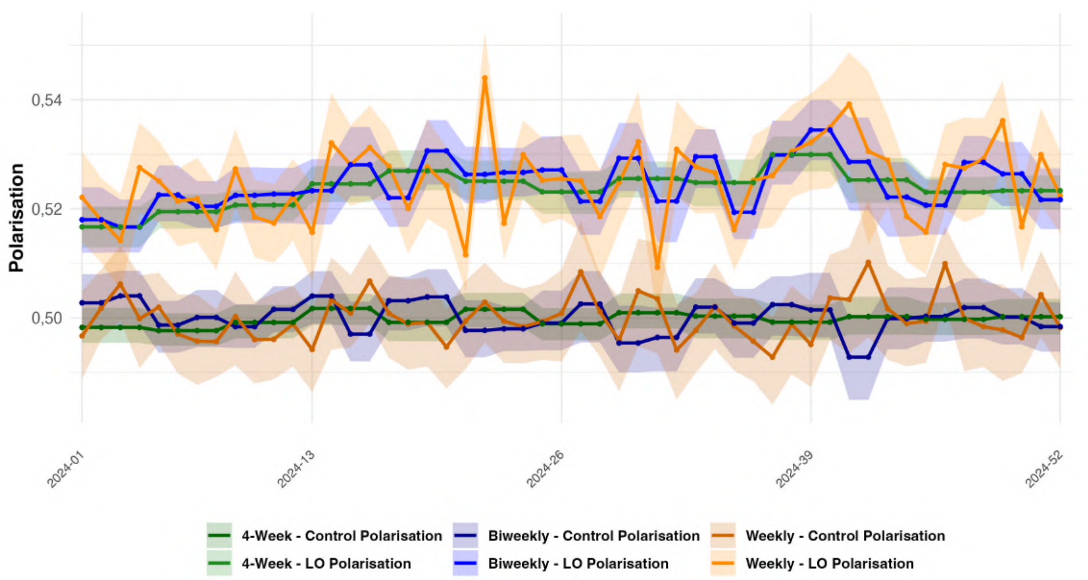


Figure 30: LO-polarisation, variable time interval, 2024

7.1.3 Statistical significance of estimated polarisation

As we explicitly outline in the theoretical foundation for the inference of the polarisation estimate in section 5.4., we employ 80 pct. confidence intervals on all estimates of polarisation throughout the paper. This reflects the practice in the proposal of the method behind the interval (Gentzkow et al., 2019). However, the interval is relatively narrow, and we require stricter confidence on all other estimates in the analysis. Therefore, we examine if polarisation estimates remain significantly different from control polarisation after replacing eq. (8) with eq. (17) below:

$$CI_t \equiv (\hat{\pi}_t - (Q_t^k)_{(98)}/\sqrt{\tau_t}\,;\, \hat{\pi}_t - (Q_t^k)_{(3)}/\sqrt{\tau_t}) \tag{17}$$

The equation defines a 95 pct. point-wise confidence interval on polarisation. We illustrate general polarisation using this interval in figure 31. When using an 80 pct. confidence interval, polarisation in 8 biweeks in the period 2015-2024 overlaps with control polarisation, cf. figure 21. This corresponds to 3 pct. of biweeks. In figure 31., the number has increased to 61 biweeks, corresponding to a 20 percentage points increase to 23 pct. of observations. We conclude that polarisation in some biweeks that were previously significantly different from control polarisation is now insignificant, but notably, for a clear majority of periods, polarisation remains significantly positive. This corroborates the main working hypothesis using a stricter threshold for statistical significance.

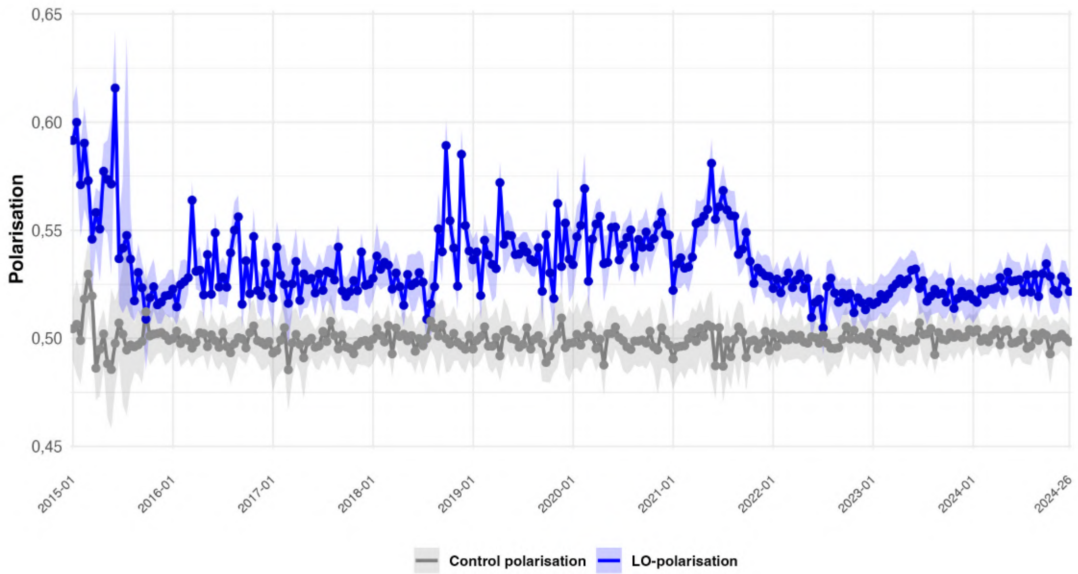


Figure 31: General LO-polarisation, 95 pct. confidence interval

7.2 Robustness of analysis results

The subsequent three sections contain formal and informal checks to investigate the robustness of the results presented in section 6. The results rely on entirely new data, where these methods have not previously been implemented. Therefore, the robustness checks are key in determining the validity of the conclusions from the analysis. Specifically, we revisit the tests for stationarity and the results of the electoral cycles- and short-term impact analysis.

7.2.1 Stationarity of polarisation and sentiment gap.

For both time-series we test for stationarity using an Augmented Dickey Fuller (ADF) test with drift, the results of which are reported in Appendix 10.12. For the sentiment gap, we reject the null hypothesis of non-stationarity at a 1 pct. significance level at all relevant lags $k \in \{1, 10\}$. However, for polarisation, the ADF test statistics are contingent on the lag length. Specifically, we fail to reject non-stationarity at lag length $k = 6$. For all other relevant lags lengths, we reject non-stationarity at a 5 pct. significance level. This indicates that stationarity is potentially weak or borderline in polarisation and could pose a concern for inference, as the conclusions of all previous econometric analysis on the levels of polarisation rest on the assumption of stationarity.

To assess whether non-stationarity poses a credible concern for inference, we determine optimal lag length in an autoregressive specification to determine if the optimal lag length is close to the lag lengths with weak stationarity. Automatic model selection of AR-model specifications minimising BIC returns an optimal lag length of $k = 1$. This supports the use of models with a single lag in the regressions in section 6., keeping the lag length well below the lag lengths where stationarity is weak. Thus, for both the sentiment gap and polarisation, the tests adequately support stationarity at relevant lag lengths, and we do not expect it to compromise the robustness of our model-based inference. Appendix 10.12. reports tests for autocorrelation and heteroskedasticity too. None of the related test statistics indicate issues challenging the robustness of the inference we draw from the models employed in section 6.

7.2.2 Robustness of the electoral cycles

The design of the electoral cycles analysis does not represent a fixed methodology. For instance, the specification of the leading electoral window reflects the assumption that any effect on either polarisation or sentiment gap is predominant in the lead-up to an election, cf. section 6.3. Given the scarce literature on the field, the length of the electoral window reflects our assumptions on news composition. To validate these, we therefore extend the electoral cycles analysis to see how the results depend on the length of window. In section 6.3., the electoral window is defined in the time-span from the election and some time into the past. In this section the specification is extended using an election window centred around the election. While we do not report the parameter estimates here, we note that the electoral cycles analysis is contingent on the model specification and that when extending the baseline ARMAX(1,1)-GARCH(1,1) model from section 6.2. to include the terms capturing electoral cycles in eq. (11) and (12), we fail to document any electoral cycles effects, as these are absorbed as shocks to the conditional variance. In the following, we examine how the results in section 6.3. depend on different specifications of the electoral window.

Firstly, we centre the election window. Table 10. presents estimates analogous to those in table 6., with an unchanged length of the electoral window (13 biweeks). Comparing the two, we note that the level effect in the conditional mean and the distinct negative trend within the electoral window are insignificant for both sentiment gap and polarisation. This supports the assumption that the predominant electoral effect on the news stream is in the lead-up to an election. The dynamic effect, seen as a distinct increase in persistence in both time-series is no longer present for either variable.

Table 10: Robust Regression Results: Centred election window, 13 biweeks

	Specification:			
	M10: Eq (11)	M11: Eq (11)	M12: Eq (12)	M13: Eq (12)
Lag(1)	0,457***	0,557***	0,461***	0,556***
	(0,078)	(0,072)	(0,082)	(0,072)
$D_{election}$	0,002	0,007		
	(0,006)	(0,006)		
Lag(1) $\times D_{election}$			−0,003	0,012
			(0,170)	(0,012)
Trend	−0,00000	−0,00002	−0,00001	−0,00002
	(0,00001)	(0,00001)	(0,00001)	(0,00001)
Trend $\times D_{election}$	−0,00003	−0,0001	−0,00001	−0,0001
	(0,00004)	(0,00003)	(0,00002)	(0,00003)
Constant	0,010***	0,238***	0,011***	0,239***
	(0,002)	(0,039)	(0,002)	(0,038)
Quarterly effects	yes	yes	yes	yes
Dependent variable:	Sentiment	Polarisation	Sentiment	Polarisation
Observations	259	259	259	259
R^2	0,290	0,427	0,289	0,426

Note: Robust standard errors in parentheses. $^{*}p < 0{,}1$; $^{**}p < 0{,}05$; $^{***}p < 0{,}01$

Secondly, we estimate the level and dynamic effect models with election windows spanning 6, 13 and 26 biweeks, representing 1, 2, and 4 quarters, cf. Appendix 10.17. When focusing on the level effect, cf. Appendix 10.17., table 24., we cannot establish an electoral cycle when the electoral window is defined just a quarter prior to the election. However, for both the window of 13 and 26 biweeks, there are significant level effects when using a leading electoral window. Similarly, there are significant changes to the trend in both these window lengths. Generally, few of the coefficients when using the centred window are significant, lending support to the working assumption that the leading window specification reflects the true dynamic best. When turning to the dynamic effect, cf. Appendix 10.17., table 25., again, for the shortest window length the model does not capture any electoral cycle. Specifically, both the change in persistence and in trend remain insignificant at a 5 pct. significance level. And as above, the leading election window specification appears to capture some electoral cycle effects that the centred specification is not capturing.

In conclusion, the robustness check above support that both sentiment gap and polarisation appear to follow an electoral cycle. This supports the conclusion that in the window prior to an election both polarisation and sentiment gap are higher than normal and that both decline as the election approaches, in line with the results presented in section 6.3. As previously argued, we assign little interpretation to the level difference in the electoral window, as it appears to be exclusively driven by observations in 2015. In fact, when we estimate the cycles on data excluding 2015 the level effect disappears - the trend effect does not. We do not present these results, as we consider using just two elections to be inadequate when establishing an electoral cycle. The robustness check supports the main finding in section 6.3.: that elections appear to function as de-polarising events in the news coverage. While the cycle is present in both series under some specifications, we note that we cannot establish it unequivocally across models or across different election window lengths. This indicates that the changes are limited and take place specifically prior to an election.

7.2.3 Robustness of short-term impact of event

In line with our approach to validate the RDiT in section 6.4., we conduct a placebo test on each segmented regression. As described in section 6.5., segmented regression is used to assess whether singular events lead to short-term changes in polarisation. A key concern with this approach - particularly when applied to high volatility data like ours - is that observed changes might not be caused by the event itself, but rather reflect random variation, underlying trends not fully captured in the pre-period, or coincide with sharp changes in the unobserved confounders (Huntington-Klein, 2021). To address this concern, we shift each event six weeks earlier and re-estimate the segmented regressions. This method is not without limitations, as we cannot rule out the possibility that other events, unrelated to the ones we study, may have influenced polarisation six weeks prior to the actual eight events. However, this uncertainty reflects the nature of the news stream, where coverage occurs continuously, regardless of the significance of the material covered. We consider this placebo test a reasonable robustness check of the limited short-term event effects identified in section 6.5.

Table 11. presents the results from the placebo segmented regressions, PSR(X), corresponding to the events analysed in section 6.5. We find no effects that are significant at a 5 pct. significance level for any model PSR(1)-PSR(8). There are negative level effects significant only at a 10 pct. significance level for PSR(2) and PSR(8), and similarly for change in slope in PSR(6). This suggests that the effects that were weakly significant only at a 10 pct.

significance level in section 6.5. reflect random variation in the data rather than being actual event-driven effects. Regarding the effects that were significant at 5 pct. in SR(4) and SR(6) they disappear or lose significance in PSR(4) and PSR(6), indicating that they are robust to this test.

Table 11: Segmented Regression Results from placebo event dates

	PSR(1)	PSR(2)	PSR(3)	PSR(4)	PSR(5)	PSR(6)	PSR(7)	PSR(8)
Intercept	0,520***	0,502***	0,543***	0,524***	0,541***	0,541***	0,522***	0,515***
	(0,004)	(0,005)	(0,010)	(0,016)	(0,012)	(0,012)	(0,008)	(0,003)
t	0,001	0,000	0,000	0,001	–0,000	–0,000	–0,001*	–0,000
	(0,000)	(0,000)	(0,001)	(0,001)	(0,001)	(0,001)	(0,000)	(0,000)
Z_{Event}	0,001	–0,017*	0,001	–0,013	–0,013	0,007	0,006	–0,010*
	(0,013)	(0,009)	(0,018)	(0,025)	(0,016)	(0,011)	(0,008)	(0,005)
$Z_{Event} \times t_{Post}$	–0,001	0,001	–0,003	0,000	0,002	–0,003*	0,001	0,001*
	(0,001)	(0,001)	(0,002)	(0,002)	(0,002)	(0,001)	(0,001)	(0,000)
Polarisation	*General*	*Udland*	*Indland*	*Indland*	*Indland*	*Udland*	*Udland*	*Udland*
Observations	39	39	39	39	39	39	39	39
R^2	0,061	0,082	0,119	0,037	0,086	0,218	0,081	0,124

Note: Robust standard errors in parentheses. $^{*}p < 0,1$; $^{**}p < 0,05$; $^{***}p < 0,01$

In conclusion, the main finding in section 6.5 was that there was only limited empirical support for WH5. The findings above strengthen this preliminary conclusion in the sense that they weaken the previous results, further limiting support for WH5. We see this as there are events in the placebo test that are weakly significant for events that should not have predictive power. In section 6.5., however, we attributed some effect to SR(4) and SR(6), both of which have insignificant coefficients for the relevant parameters in the placebo test, indicating that possibly these two events do indeed explain polarisation on short-term. Nevertheless, we fail to accept WH5: that contemporaneous events drive polarisation, in general.

8 Discussion

This section contains discussions of the methods, assumptions, and design choices reflected in our results. We discuss the extent to which the estimates of polarisation between DR and TV2 generalise to media-polarisation more broadly, what the estimates imply, and propose alternative designs for future research.

8.1 Broad validity of polarisation estimates

The paper sets out to examine polarisation in Danish media. For various methodological reasons outlined in section 3.1., the empirical analysis is based on news content from two major mainstream news outlets. In light of this limitation, how do the estimates in section 5.6. mirror polarisation in Danish news broadly speaking? We present two opposing arguments for whether the estimated polarisation is indicative of polarisation in general. The first challenges any broad validity of the results, whereas the second supports the interpretation that our findings are indicative of a broader media polarisation. The claim against any broad implications of our results is based partly on VIVE (2022), and argues that we should expect true polarisation in news to stem at least in some part from reinforced echo-chambers, which are likely niche online media outlets. As social media becomes an increasingly integral platform for news consumption, these echo-chamber outlets accrue a larger reader-base (Kulturministeriet, 2021). Therefore, the estimates of polarisation in mainstream media does not capture whether true polarisation increases when or if more readers substitute mainstream media consumption with that of fringe media. For this argument to hold, the reader-base of the mainstream media should shrink due to substitution (which it does not (Danmarks Radio, 2025)) or the fringe media should become increasingly polarising (which we cannot estimate in our framework).

The counter-argument, supporting broad validity of the estimated polarisation, is based on the competition argument, presented in section 2.2. In this line of interpretation, the low polarisation estimated in recent years actually reflects low polarisation across Danish media broadly. The argument states that the media, to some degree, adjust editorial lines to cater to reader-base preferences. Consequently, if some polarisation is optimal in Danish media, from a competition point of view, the mainstream media outlets would not be immune to its influence. For Fox News in an American context, Bursztyn et al. (2020) observe how a mainstream media outlet can act more or less polarising. Through the lens of the competition argument, we interpret this as possibly being a reaction to a saturated centre coverage, where increased distinction has proved profitable. For DR and TV2 their public service commitments somewhat limit the ways in which they might react to these incentives, nevertheless, this argument presents a mechanism by which general polarisation in media would be mirrored by polarisation in the mainstream media.

While the latter suggests that we can extrapolate general tendencies in estimated polarisation to hold for media as a whole, the coverage within the section *Udland* specifically challenges this. In *Udland*, there is less polarisation than in *Indland* - a difference that could be explained by the large number of articles based on AP/Reuters news telegrams in *Udland* in both media. That the media rely so heavily on these telegrams in their coverage of foreign affairs somewhat sidetracks the competition argument. At the very least it remains an open research question why - following the competition argument - it appears optimal to have different levels of polarisation in the coverage of news in sections *Indland* and *Udland*.

8.2 Implications of the main results

In section 5.6., we definitively accept our main working hypothesis: that there is polarisation in Danish media. But documenting a strictly positive polarisation does not imply very much. While the estimate of polarisation is absolute, its interpretation becomes relative because the straightforward interpretation of the estimate - the probability with which a neutral observer with full information would correctly guess which media has published an article after observing only a single randomly drawn token - does not have a natural anchor-point. For instance, if an observer can guess the media for an article with 53 pct. chance of being right, does that characterise a media landscape in which polarisation is high? Offering a relevant reference point to compare estimates against then becomes key. The set of external reference points is in our case limited to the estimates of Caprini (2024) and possibly, taking a leap of intuition, Gentzkow et al. (2019). Compared to Caprini, by estimating polarisation over a ten-year period, we extend the set of possible anchor-points to historical levels of polarisation. In a historical context, we find that polarisation in recent years is low, and we unequivocally reject WH4: that polarisation increases over time. We also find that our estimates in recent years are in line with Caprini's estimates for American news. Considering the contemporary sources we present in section 2. on polarisation in America vis-à-vis the ones we present on polarisation in Denmark, this is unexpected. The estimated polarisation in the years 2019-2021, when polarisation is high, appears quite dramatic in this context. It raises the natural question of whether the LO-estimate of polarisation does, in fact, overestimate polarisation? In section 5.6., we compare estimated polarisation against the control polarisation, which appears unbiased for all LO-estimates. While this implies that polarisation is not upwardly biased, it does not exclude that we overestimate polarisation compared to Caprini due to differences in token composition. Caprini uses a more exhaustive feature extraction process for image content, while we use tokens from a combination of text and images. These input differences could explain level differences in estimated polarisation that are not informative of the true polarisation. Gentzkow et al. do - like us - estimate large shifts in polarisation over time, and so we cannot reasonably rule out that the polarisation, and notably changes in polarisation, we estimate reflect the true polarisation. Nor can we rule out that the estimated polarisation in the paper at hand is, in fact, a better signal of the true polarisation than the one presented by Caprini.

8.3 Research design choices

In our extraction of data and the subsequent estimation of polarisation, we have made several consequential decisions. In the following, we discuss what these decisions imply in terms of data foundation and limitations on the specifications of the polarisation estimator.

We begin by discussing the implications and alternatives to our data restriction and selection strategy. A straightforward approach to estimating polarisation between media based on token data is to rely solely on keyword extraction from article text. Scraping, processing, and extracting meaning from large bodies of text using natural language processing is more widely established and requires fewer computational resources than image analysis. For instance, Gentzkow et al. (2019) employs a database of scripted speech to estimate polarisation exclusively on textual data. Our objective is to examine polarisation between DR and TV2, where article text is not the only channel through which editorial stance is conveyed. Findings by Ash et al. (2022) and Caprini (2024) illustrate that image content, when analysed using computer vision and NLP-tools, captures distinct differences between media outlets in an American context. This approach has been enabled by recent advances in CV and neural networks, but it omits the textual dimension of news articles. We therefore combine image and text to construct a more comprehensive dataset for each article, trading off the efficiency of a text-only method for the added insight afforded by visual data. Although beyond the scope of our paper, comparing polarisation estimates based on text-only and image-only tokens could further substantiate our findings and help identify the source of polarisation. Caprini does this and finds that image content contributes disproportionally to polarisation. Whether the same holds in Danish news remains a promising direction for future research on media polarisation.

Restricting the data foundation to DR and TV2 implies the absence of comparable polarisation estimates for other outlets, and, consequently, we lack of a viable control group in the analysis. Without such a group, we cannot apply a difference-in-difference (DiD) approach to study causal effects of events - an extension that would otherwise enrich our analysis. To construct a panel dataset with a viable control group, we suggest two alternative strategies regarding data foundation. The first approach is to estimate section-specific polarisation in a wide range of national media outlets. The resulting dataset would include pairwise polarisation estimates within sections between, e.g., DR.dk and TV2.dk, and DR.dk and Politiken.dk, and so on[22]. We would then be able to estimate causal effects from an event to a group of section specific pairwise estimates of polarisation by using an unrelated section as a control group in a DiD-framework. For example, we could identify whether specific foreign events give rise to changes in polarisation in section *Udland* across media. Here, polarisation in *Indland* across media would serve as a control group, identifying underlying trends in the counterfactual no-treatment case, assuming the parallel pre-trend holds. Moreover, extending the dataset on nation-wide media polarisation with estimates on political leaning of the media provides an empirical foundation to examine whether polarisation correlates with political bias. A second possibility is to compare polarisation in domestic media with that of a set of prominent media outlets from comparable countries. Using these as a synthetic control, it would facilitate comparisons of Danish polarisation dynamics with those of similar media outlets abroad. Such an approach could, for instance, isolate the effects of political scandals on polarisation within Danish news by accounting for a common trend in the synthetic control group. This approach would provide evidence on whether, e.g., the ”mink case” has had a polarising effect in the news coverage. Furthermore, it would enable cross-country comparisons of media-polarisation, in line with to the analysis of cross-country affective polarisation of Boxell et al. (2020).

Turning to the implications of using the LO-estimator as our measure of polarisation, we limit the discussion to two core aspects: the definition of $\boldsymbol{\rho}_t$ and the decision not to estimate polarisation using the penalised estimator. The definition of time-varying source probability $\boldsymbol{\rho}_t$ is the principal contributor to the bias in the MLE-estimator, addressed in the LO-specification by excluding the current article from $\boldsymbol{\rho}_{t,-i}$. However, we theorise that a more comprehensive measure of source probability is achievable within this framework. Rather than limiting $\boldsymbol{\rho}_{t,-i}$ to contemporaneous articles (within the period), it could be extended to include tokens from all articles in a quarter or even a year. This would increase the set of valid tokens markedly and provide a more stable distribution of source probabilities for tokens uninformative of the true polarisation. These are the tokens that we label ”common” in section 5.1. Extending the token input to $\boldsymbol{\rho}_t$ would lower the risk that tokens are considered partisan by chance. Such an approach would substantially increase computational demands but could reduce volatility in polarisation by maintaining source token frequencies, $\mathbf{q}_t^S$, that are defined biweekly or even weekly, while computing source probability $\boldsymbol{\rho}_{year,-i}$ over a longer horizon. This represents a

[22] Hosseinmardi et al. (2025), published on the 21st of May 2025, implement this approach on TV news from a sample of the largest cable and broadcasting stations in the US. The recent publication date reflects how rapidly the field is developing and also explains why we do not relate our results to their study.

potential methodological refinement of the framework we have applied. The second core aspect concerns our choice of using the LO-estimator, which sacrifices the ability to model token-level covariates, only obtainable with the penalised estimator, as mentioned in section 5.3. As a result, we are unable to identify which tokens contribute most strongly to polarisation. An analysis using the penalised estimator could reveal whether the differences in sentiment associated with particular keywords or topics, as outlined in section 4.2.2., contribute significantly to polarisation. We consider most of the elements discussed above potential future research subjects. Many of these elements have been hypothetical options for our study design, but we discarded them due to time constraints, data availability, or because they entail an excessive computational load.

9 Conclusion

The conclusion follows the structure presented by the working hypotheses. As the paper reveals several aspects of media coverage and polarisation, we provide a concise overview of our main findings before outlining in greater detail the specific evidence for or against each working hypothesis iteratively. Lastly, we conclude on the methodology applied throughout the paper.

Overview

The overarching objective of the paper has been to test whether there is polarisation in Danish media. This is summarised by the main working hypothesis, which we accept in section 5., based on polarisation between DR and TV2.

We present viable, theoretical channels of polarisation, and we find support for these in section 4. Using articles published in the period 2015-2024, we document significant differences between the two media in their representation of different groups in images and in their political profiles. Sentiment analysis of articles uncovers varying editorial lines between the two media in general, and specifically for topics prevalent in the public debate. Based on 11.954.367 tokens from 168.197 articles and images, we estimate measures of polarisation and observe that polarisation in recent years is low and stable, cf. section 5. In section 6., we reject the hypothesis that polarisation increases over time, and test whether contemporaneous events drive polarisation. We find limited support for the latter and conclude that the evidence on the empirical link between events and polarisation is ambiguous.

Main Working Hypothesis: There is polarisation in Danish media

To test the main working hypothesis, we employ a rigorous theoretical framework adapted from a comparable study of polarisation in speech. We collect articles and images over a ten-year period from DR.dk and TV2.dk, and extract features from both using advanced computational resources. This includes combining recent advances in computer vision with natural language processing to achieve state-of-the-art scene tagging in article images. This is extended using facial analysis to estimate characteristics of people in images. Lastly, we apply classical machine learning techniques to train a model to recognise people of interest in the visual news coverage. In conjunction with statistical keyword extraction, this provides an unprecedentedly rich data foundation on which we estimate polarisation. We present estimators of varying complexity, and motivate our preferred method, the leave-out approach, by theoretically demonstrating its low bias. We empirically confirm the suspected bias when using a less advanced estimation technique.

In section 5., we present LO-estimates of polarisation. We find significant polarisation in all but 8 out of 260 two-week periods in the time-frame. Polarisation is on average $0,534$, which corresponds to a $53,4$ pct. chance of being correct when guessing which media has published an article, after being shown only a single token from that article, cf. section 5. Robustness checks in section 7. reveals that polarisation remains significant in a clear majority of periods when requiring stricter thresholds for statistical significance and when adjusting the interval length. We notice that most of the polarisation comes from the coverage under the compiled section *Other* and the section *Indland*, while polarisation in *Udland* is lower and polarisation in *Politik* is insignificant. We discuss why polarisation in *Udland* is lower than in *Indland* in section 8., and how it challenges what we throughout the paper refer to as the competition argument for polarisation, cf. section 2. It remains an avenue for future research to examine optimal polarisation and what drives differences in polarisation between sections.

Estimating polarisation empirically, as we have done in the paper at hand, constitutes a novel contribution to the field of polarisation studies in Denmark. Using an extensive empirical framework, we have proven that there is significant polarisation in the token representation of the news coverage and - in extension - we accept the main working hypothesis and conclude that there is polarisation in Danish news.

WH1: The media differ in their visual coverage

With facial analysis of the lead images we assign gender, ethnicity, emotion, and age to the faces that appear in the images. This presents a number of insights into how the visual coverage differs between the media, and potentially reveals how differences in editorial line extend into the visual coverage. In section 3., we observe that there are on average 1,5 faces per image in DR images and that the corresponding number in TV2 images is 1,2. The difference is significant at a 1 pct. significance level. We find a relatively low share of female faces in the images, around 16 pct. for both media, cf. section 4. There is a small but significant difference in the shares across the general coverage, but notably pronounced discrepancies in the female share in sections, highlighting a more equal representation of genders in the domestic coverage in TV2. This supports WH1, while the fact that we see no pronounced differences in the ethnic representation between media challenges it. Observing significant differences in the average age of the people depicted, with individuals in TV2 images being older on average, we conclude that the media do differ in their visual coverage in three out of the four aspects we have investigated.

WH2: The media cover similar topics in different tones

As a relatively simple extension of the feature extraction of text and images in section 3., we estimate average sentiment scores for the full article text. This shows that DR articles are on average significantly more positively worded than their TV2 counterparts. By conditioning the article set on the presence of specific topics, we notice clear discrepancies in the tonality of the articles within the topics. The media generally align in whether a topic is covered in a more positive or negative tone than the average coverage, but the difference from their respective average tone varies significantly over time and between media. The latter leads us to accept WH2.

WH3: The media have distinct political profiles

Using the sentiment scores, we investigate different aspects of WH3. In section 4., we document that the media differ both in the intensity and tone with which they cover US politics and Danish party blocs. This indicates distinct political profiles of the outlets, which supports WH3. We go on to document how TV2 appears to embrace a harder editorial opposition in 2020, which we do not observe for DR, again lending support to the hypothesis that the media have distinct profiles. In conjunction with significant differences in the favourability the media show toward different parties, we conclude that for certain aspects of WH3, we accept the hypothesis. Notably, though, we fail to document different profiles in their coverage of prime ministers in both text and images.

WH4: Polarisation increases over time

In addition to the main working hypothesis, WH4 provides a testable hypothesis on the dynamic that governs estimated polarisation. In order to test it, we analyse the characteristics of polarisation as a time-series in section 6. We observe heteroskedasticity in polarisation and model it using an ARMAX-GARCH framework, revealing that the time-series is characterised by both seasonality and pronounced persistence in volatility. These insights form the foundation for an analysis of electoral cycles, which include a trend term, cf. section 6. In all model specifications, the global trend remains insignificant, and we reject WH4. This is in line with the graphical analysis of polarisation in section 5.

WH5: Contemporaneous events drive polarisation

Testing WH5 requires candidate events. The descriptive analysis in section 4., illustrates how Ukraine was mentioned in about three of every four articles in the two-week period around the start of the Russian invasion of Ukraine in 2022. The profound displacement of "normal" coverage this entails implies that the invasion is such a candidate event. In section 6., we use a regression discontinuity in time framework to test if we can ascribe the change in level and trend to the event. We conclude that, if the changes are attributable to any one event, then the invasion is likely not the true event. As we analyse first level/trend differences and later change-points in variance, we conclude that it is increasingly unlikely that the fundamental shift and long-term changes we observe arise from any singular event, which challenges WH5.

Observing expansions and contractions in polarisation implies that we consider general elections as candidate events too. We propose a model framework to test for electoral cycles and document a significant negative trend in a recurring pre-election window. This suggests that general elections have a de-polarising effect, and that WH5 likely holds for elections.

Lastly, we propose eight candidate events to broadly test short-term impacts on polarisation. We test the effects of the events using an interrupted time-series approach with segmented regression. For two of the eight candidate events in this analysis, we document robust short-term effects. These two events are the first COVID-19 lockdown in Denmark and the 6$^{\text{th}}$ of January attack on the US Capitol, which suggests that, possibly, for some singular events, WH5 might hold. However, we fail to establish any effects for the other six candidate events and fail to establish any systematic pattern in what determines if singular events drive polarisation. In conjunction with the aspects tested above, this implies that support for WH5 is ambiguous.

Final remarks and directions for future research

The conclusions to the main working hypothesis and WH1-WH5 present novel insights into media-polarisation and its drivers in Denmark. Our principal finding - that media-polarisation exists but is not intensifying - corroborates the findings of VIVE (2022) regarding affective polarisation. To the degree that increases in media-polarisation undermine the effective functioning of democratic institutions, our results provide a measure of reassurance: we find no evidence of a rising trend in media-polarisation that would signal growing cause for concern in a Danish context. The data foundation on which the conclusions rest is exceptionally exhaustive, allowing us to draw conclusions on polarisation in both images and text. Finally, the paper advances the study of media-polarisation in Denmark by integrating computational complexity with a theoretical scope that is unprecedented in a Danish context. This approach demonstrates how the combination of econometrics and recent technological advances makes it possible to quantify measures previously inaccessible to empirical analysis. We suggest that future studies of polarisation build on the framework we have outlined, as it presents numerous opportunities for further analysis.

10 Appendix

10.1 List of R-packages

Package	Link	Package	Link
lubridate	CRAN link	data.table	CRAN link
dplyr	CRAN link	ggplot2	CRAN link
ISOweek	CRAN link	scales	CRAN link
tidyr	CRAN link	sandwich	CRAN link
lmtest	CRAN link	zoo	CRAN link
purrr	CRAN link	grid	CRAN link
kableExtra	CRAN link	reshape2	CRAN link
gridExtra	CRAN link	stringr	CRAN link
glmnet	CRAN link	xtable	CRAN link
margins	CRAN link	rugarch	CRAN link
forecast	CRAN link	stargazer	CRAN link
tseries	CRAN link	estimatr	CRAN link
forcats	CRAN link	rdrobust	CRAN link

Table 12: R packages used in the analysis with CRAN links

10.2 Criteria for relevancy of article links

For both news outlets we aim to discard live updates and news stories in other formats than traditional articles which are accompanied by at least one article image. The web page design is not 1:1 for the two outlets however we discard the same type of unwanted article types by enforcing the following criteria on the URL-string of the article links.

For DR.dk we require that ”/nyheder/” or ”/Nyheder/” is present in the URL and that neither ”/om-dr/”, ”/reel/”, ”/seneste/”, ”/ultra/”, ”/p3/”, ”/tv-guide/”, ”/etik-og-rettelser/”, or ”/det-bedste-fra-dr/” is present.

Likewise, for TV2.dk we discard unrelevant article types by requiring that the URL-string contains either ”nyheder.”, ”nyhederne.”, ”politik.”, ”finans.”, or ”vejret’. The check is case-insensitive. Further, the presence of ”/reel/” or ”/live/” disqualifies the article link.

We post-process by removing duplicates based on articles title, and further filter out any form of ”live” newsfeed, as these articles have distinct titles (e.g. ”livecenter 12:35”) but the exact same contents, and appear several times a day. This post-processing includes removal of COVID-19 update blogs etc.

10.3 Article categories for both media

DR sections	TV2 sections
Indland	*Krimi*, *Trafik*, *Samfund*
Politik	*Politik*
Udland	*Udland*
Regionale	*Lokalt*
Viden, *Vejret*	*Klima*, *Tech*
Penge	*Finans*, *Business*, *Penge*, *Erhverv*
Detektor	
Kultur	
Webfeature	

We determine the section in which an article is published based on the article category in the URL-string of the article link. The list of sections for both outlets and how they correspond is listed below:

Some sections overlap and e.g. DR.dk articles matching the TV2.dk section *Klima* is found in both *Viden* and *Vejret*. However, the DR.dk section *Viden* overlaps with the TV2.dk section *Tech*. This justifies that we retain a high mapping accuracy in the compiled section *Other*.

10.4 Features from facial analysis

The extraction of facial attributes assign the following features. Age is estimated as the most likely year-number out of all ages.

Category:	**Probability of:**					
Ethnicity	White	Black	Indian	Asian	Middle Eastern	
Robust ethnicity	White	Non-white				
Emotions	Happy	Sad	Angry	Neutral	Fear	Surprise
Gender	Male	Female				
Age	Age in years					

Table 13: Facial attributes.

10.5 mRAKE and stopwords in Danish

We use the stop-words presented in table 14. in the mRAKE process. We rely on Grabovets (2022) in our implementation of the mRAKE algorithm.

Table 14: List of Danish stopwords

af	alle	andet	andre	at	begge	da
de	den	denne	der	deres	det	dette
dig	din	dog	du	ej	eller	en
end	ene	eneste	enhver	et	fem	fire
flere	fleste	for	fordi	forrige	fra	få
før	god	han	hans	har	hendes	her
hun	hvad	hvem	hver	hvilken	hvis	hvor
hvordan	hvorfor	hvornår	i	ikke	ind	ingen
intet	jeg	jeres	kan	kom	kommer	lav
lidt	lille	man	mand	mange	med	meget
men	mens	mere	mig	ned	ni	nogen
noget	ny	nyt	nær	næste	næsten	og
op	otte	over	på	se	seks	ses
som	stor	store	syv	ti	til	to
tre	ud	var				

10.6 List of target checks

. In table 15., we present the list of target checks for mentions in articles.

Table 15: Overview of target keyword checks and detection logic

Target	Detection Logic
`rusland`	Substring match
`putin`	Substring match
`kina`	Substring match
`jinping`	Substring match
`klimaforandring`	Substring match
`klima`	Substring match
`indvandrer`	Substring match
`indvandring`	Substring match
`udlænding`	Substring match
`asylansøger`	Substring match
`muslim`	Substring match
`islam`	Substring match
`mette frederiksen`	Substring match
`lars løkke`	Substring match
`statsminister`	Substring match
`enhedslisten`	Substring match
`sf`	Regex: exact whole word match (`\bsf\b`)
`radikale venstre`	Substring match
`socialdemo`	Substring match
`venstre`	Match if `venstre` appears not directly after `radikale`
`konservativ`	Substring match
`liberal alliance`	Substring match
`dansk folkeparti`	Substring match
`danmarksdemokrat`	Substring match
`moderaterne`	Substring match
`ukrain`	Substring match

10.7 Further descriptive insights

Table 16: Welch Two-Sample t-test: Faces per article image (DR vs TV2)

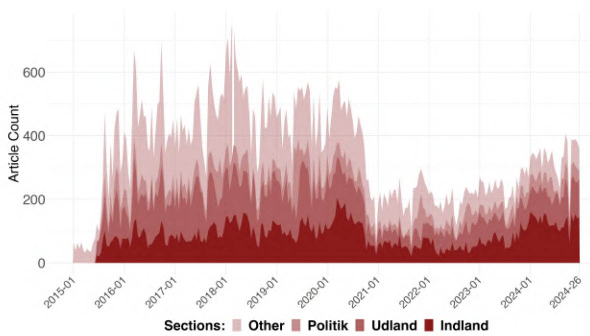
	DR	TV2
Mean number of faces	1,457	1,204
	Test summary	
Test statistic (t)	20,84	
Degrees of freedom	168.125	
95% Confidence interval	[0,229 ; 0,277]	
p-value	0,000	
Alternative hypothesis	True difference in means $\neq 0$	

Table 17: Summary of Article Counts and Averages

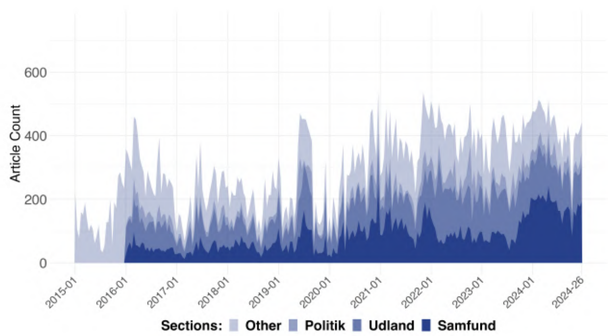
Metric	DR	TV2	Total
Average	345	301	646
Average Excluding 2015	361	315	676
Total Articles	90,088	78,109	168,197
Total Articles Excluding 2015	85,138	74,342	159,480

Table 18: Average share of coverage per section in DR and TV2, with and without 2015 included.

Theme	DR	TV2	DR excl. 2015	TV2 excl. 2015
Udland	25,6%	27,1%	26,6%	30,2%
Indland/Samfund	24,7%	24,1%	26,2%	26,9%
Politik	11,2%	8,4%	11,8%	9,3%
Other	38,5%	40,4%	35,4%	33,6%



(a) Stacked article count by theme (DR)



(b) Stacked article count by theme (TV2)

Figure 32: Comparison of article theme distribution between DR and TV2

Table 19: Welch's t-tests of pairwise difference in sentiment score, by source

	Test of difference in sentiment					
	TV2			DR		
Year	A	B	C	A	B	C
2015	0,004	-0,000	-0,004	-0,004	-0,012***	-0,008**
	(0,329)	(0,895)	(0,196)	(0,319)	(0,004)	(0,029)
2016	0,003	-0,006*	-0,009***	-0,009***	-0,010***	-0,001
	(0,299)	(0,069)	(0,001)	(0,003)	(0,001)	(0,704)
2017	0,011***	0,010***	-0,001	0,002	0,000	-0,002
	(0,004)	(0,008)	(0,739)	(0,329)	(0,873)	(0,418)
2018	0,008*	-0,001	-0,009**	0,000	0,001	0,001
	(0,052)	(0,769)	(0,034)	(0,890)	(0,710)	(0,811)
2019	-0,003	-0,000	0,003	0,001	-0,001	-0,002
	(0,198)	(0,886)	(0,261)	(0,508)	(0,660)	(0,271)
2020	0,012**	0,016**	0,004	-0,001	0,003	0,005*
	(0,044)	(0,015)	(0,459)	(0,628)	(0,201)	(0,051)
2021	-0,004	0,002	0,006***	-0,004	-0,001	0,003
	(0,101)	(0,369)	(0,002)	(0,125)	(0,599)	(0,268)
2022	-0,009***	-0,005**	0,005**	-0,006**	0,002	0,008***
	(0,000)	(0,030)	(0,011)	(0,018)	(0,320)	(0,000)
2023	-0,002	0,004	0,006*	-0,010***	-0,005**	0,005*
	(0,549)	(0,139)	(0,051)	(0,000)	(0,045)	(0,087)
2024	-0,001	0,003	0,004	-0,004	0,002	0,006**
	(0,785)	(0,266)	(0,181)	(0,113)	(0,326)	(0,012)

Note: $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

H_0 = No difference in the average sentiment
Following groups are tested:
A = "current non-governing" - "current governing"
B = "historically non-governing" -"current governing"
C = "historically non-governing" - "current non-governing"

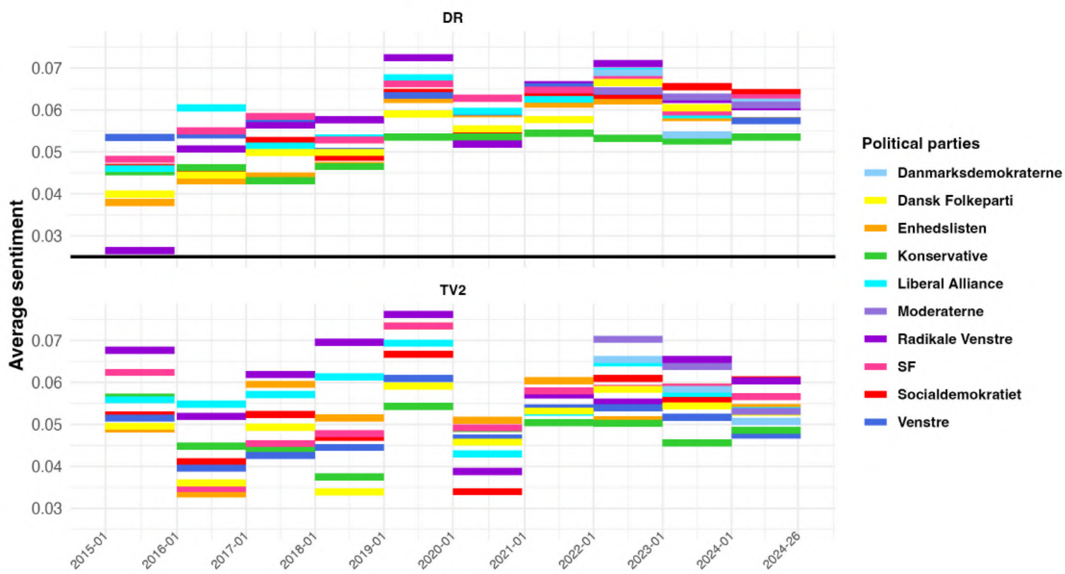
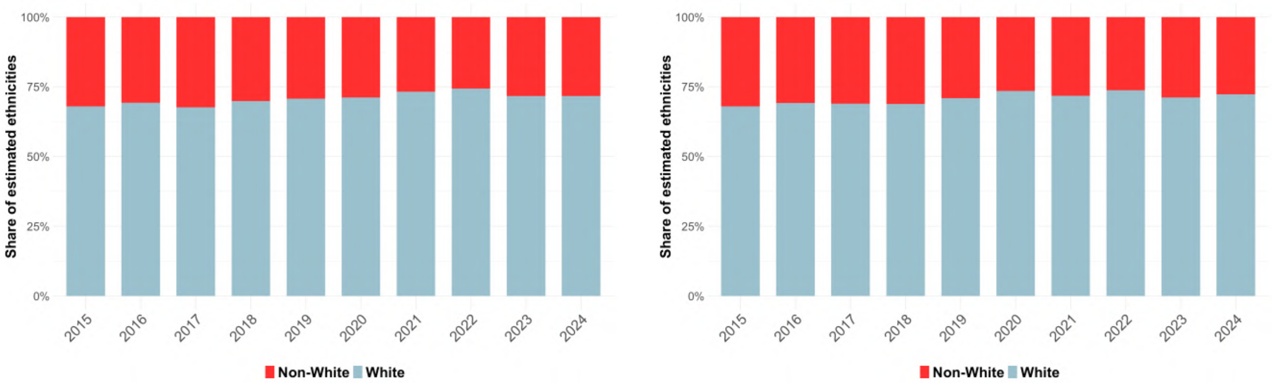


Figure 33: Average yearly sentiment of all political parties by source

Figure 34: Average yearly sentiment of political party groups

10.8 Distribution of white/non-white faces across time and media

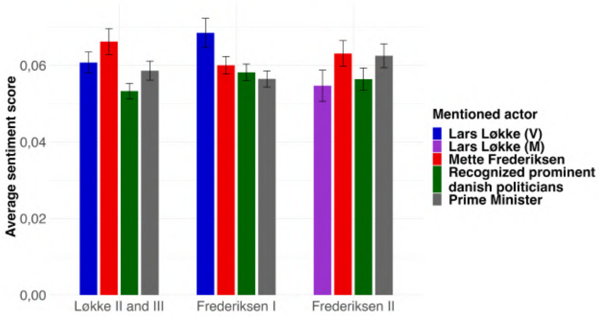


(a) DR

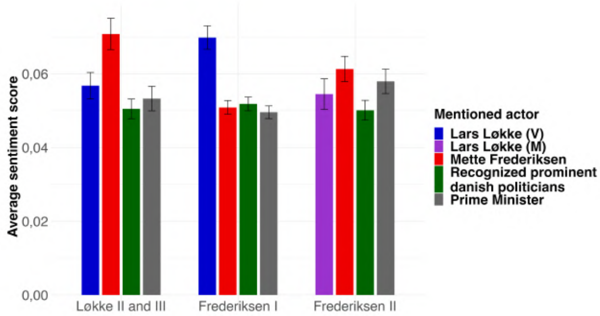
(b) TV2

Figure 35: Distribution of ethnicities (white/non-white) in news coverage by source

10.9 Sentiment for prime ministers and politicians by media

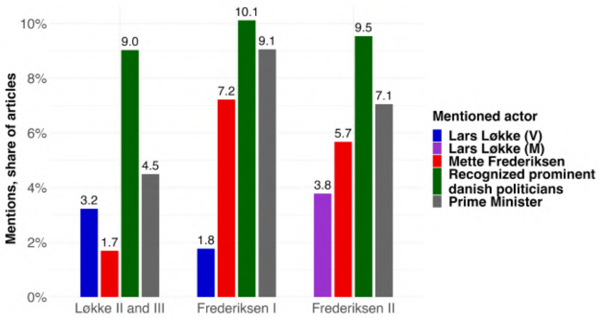


(a) Average sentiment score DR

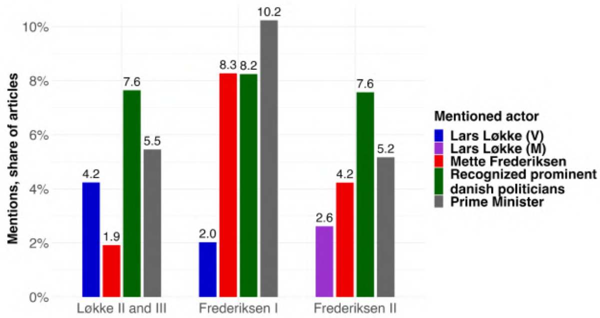


(b) Average sentiment score TV2

Figure 36: Representation of politicians and prime ministers



(a) Mentions as percent of total articals DR



(b) Mentions as percent of total articles TV2

Figure 37: Representation of politicians and prime ministers

10.10 List of people we recognise in images

Name	Party	Name	Party	Name	Party
Alex Vanopslagh	LA	Joy Mogensen	S	Marianne Jelved	RV
Ane Halsboe-Joergensen	S	Kaare Dybvad	S	Margrethe Vestager	RV
Anette Vilhelmsen	SF	Kaare Dybvad Bek	S	Martin Lidegaard	RV
Astrid Krag	S	Kamala Harris	Dem	Mattias Tesfaye	S
Barack Obama	Dem	Karen Ellemann	V	Mette Bock	LA
Benedikte Kiaer	KF	Karen Haekkerup	S	Mette Frederiksen	S
Benny Engelbrecht	S	Karen Jespersen	V	Mette Kierkgaard	M
Bertel Haarder	V	Karsten Lauritzen	V	Mia Wagner	V
Birthe Roenn Hornbech	V	Kirsten Brosboel	S	Mogens Jensen	S
Bjarne Corydon	S	Kristian Jensen	V	Mona Juul	KF
Brian Mikkelsen	KF	Kristian Thulesen Dahl	DF	Morten Boedskov	S
Carina Christensen	KF	Lars Aagaard	M	Morten Dahlin	V
Caroline Stage Olsen	M	Lars Barfoed	KF	Morten Messerschmidt	DF
Christian Friis Bach	RV	Lars Christian Lilleholt	V	Morten Oestergaard	RV
Christian Rabjerg Madsen	S	Lars Loekke Rasmussen	V	Nick Haekkerup	S
Christina Egelund	M	Lea Wermelin	S	Nicolai Wammen	S
Christine Antorini	S	Lene Espersen	KF	Ole Birk Olesen	LA
Claus Hjort Frederiksen	V	Louise Schack Elholm	V	Ole Sohn	SF
Connie Hedegaard	KF	Lykke Friis	V	Pelle Dragsted	Oe
Dan Joergensen	S	Mai Mercado	KF	Per Stig Moeller	KF
Donald Trump	Rep	Mai Villadsen	Oe	Peter Christensen	V
Ellen Trane Noerby	V	Magnus Heunicke	S	Peter Hummelgaard	S
Esben Lunde Larsen	V	Manu Sareen	RV	Pernille Rosenkrantz-Theil	S
Eva Kjer Hansen	V	Marie Bjerre	V	Pernille Skipper	Oe
Finn Poulsen	KF	Marianne Jelved	RV	Pernille Vermund	NB
Franciska Rosenkilde	Aa	Margrethe Vestager	RV	Pia Kjaersgaard	DF
Gitte Lillelund Bech	V	Martin Lidegaard	RV	Pia Olsen Dyhr	SF
Helle Thorning-Schmidt	S	Mattias Tesfaye	S	Rasmus Helveg Petersen	RV
Henrik Dam Kristensen	S	Mette Bock	LA	Rasmus Jarlov	KF
Henrik Hoeegh	V	Mette Frederiksen	S	Rasmus Prehn	S
Henrik Sass Larsen	S	Mette Kierkgaard	M	Rasmus Stoklund	S
Hillary Clinton	Dem	Mia Wagner	V	Simon Emil Ammitzboell	LA
Ida Auken	SF	Mogens Jensen	S	Simon Kollerup	S
Inger Stoejberg	V	Mona Juul	KF	Soeren Gade	V
Jacob Jensen	V	Morten Boedskov	S	Soeren Pape Poulsen	KF
Jakob Ellemann-Jensen	V	Morten Dahlin	V	Soeren Pind	V
Jakob Engel-Schmidt	M	Morten Messerschmidt	DF	Sofie Carsten Nielsen	RV
Jakob Axel Nielsen	KF	Morten Oestergaard	RV	Sophie Haestorp Andersen	S
J.D. Vance	Rep	Nick Haekkerup	S	Sophie Loehde	V
Jeppe Bruus	S	Nicolai Wammen	S	Stephanie Lose	V
Jeppe Kofod	S	Ole Birk Olesen	LA	Thor Möger Pedersen	SF
Joergen Neergaard Larsen	V	Ole Sohn	SF	Thomas Danielsen	V
Joe Biden	Dem	Pelle Dragsted	Oe	Thyra Frank	LA
Johanne Schmidt-Nielsen	Oe	Per Stig Moeller	KF	Tim Kaine	Dem
Jonas Dahl	SF	Peter Christensen	V	Tim Walz	Dem
Josephine Fock	Aa	Peter Hummelgaard	S	Tina Nedergaard	V
Pernille Rosenkrantz-Theil	S	Pernille Skipper	Oe	Torsten Schack Pedersen	V
Pernille Vermund	NB	Pia Kjaersgaard	DF	Trine Bramsen	S
Pia Olsen Dyhr	SF	Rasmus Helveg Petersen	RV	Troels Lund Poulsen	V
Rasmus Jarlov	KF	Rasmus Prehn	S	Uffe Elbaek	Aa
Rasmus Stoklund	S	Simon Emil Ammitzboell	LA	Ulla Toernaes	V
Simon Kollerup	S	Soeren Gade	V	Villy Soevndal	SF
Queen Margrethe II	-	King Frederik	-	Queen Mary	-

10.11 Full calculations of MLE and LO examples

$$\mathbf{DR_1} = \begin{bmatrix} b \\ b \\ d \end{bmatrix}, \quad \mathbf{DR_2} = \begin{bmatrix} b \\ c \\ d \end{bmatrix}, \quad \mathbf{TV2_1} = \begin{bmatrix} a \\ a \\ e \end{bmatrix}, \quad \mathbf{TV2_2} = \begin{bmatrix} c \\ c \\ e \end{bmatrix}$$

$$j = \{a, b, c, d, e\}, \quad \hat{\mathbf{q}}^S = \sum_{i \in S} \mathbf{c}_i / \sum_{i \in S} m_i \rightarrow$$

$$\hat{\mathbf{q}}^{DR} = \begin{bmatrix} 0/6 \\ 3/6 \\ 1/6 \\ 2/6 \\ 0/6 \end{bmatrix}, \quad \hat{\mathbf{q}}^{TV2} = \begin{bmatrix} 2/6 \\ 0/6 \\ 2/6 \\ 0/6 \\ 2/6 \end{bmatrix}, \quad \hat{\boldsymbol{\rho}} = \begin{bmatrix} 0/(0+2/6)=0 \\ 3/6/(3/6+0)=1 \\ 1/6/(1/6+2/6)=1/3 \\ 2/6/(2/6+0)=1 \\ 0/(0+2/6)=0 \end{bmatrix}$$

$$\hat{\mathbf{q}}^{DR} \cdot \hat{\boldsymbol{\rho}} = \begin{bmatrix} (0/6) \cdot 0 = 0 \\ (3/6) \cdot 1 = 3/6 \\ (1/6) \cdot 1/3 = 1/18 \\ (2/6) \cdot 1 = 2/6 \\ (0/6) \cdot 0 = 0 \end{bmatrix} = \left(0 + \frac{3}{6} + \frac{1}{18} + \frac{2}{6} + 0\right) = \frac{32}{36}$$

$$\hat{\mathbf{q}}^{TV2} \cdot (1 - \hat{\boldsymbol{\rho}}) = \begin{bmatrix} (2/6) \cdot 1 = 2/6 \\ (0/6) \cdot 0 = 0 \\ (2/6) \cdot 2/3 = 4/18 \\ (0/6) \cdot 0 = 0 \\ (2/6) \cdot 1 = 2/6 \end{bmatrix} = \left(\frac{2}{6} + 0 + \frac{4}{18} + 0 + \frac{2}{6}\right) = \frac{30}{36}$$

$$\pi_t^{MLE} = \frac{1}{2}\left(\frac{32}{36} + \frac{30}{36}\right) = \frac{62}{72} = \frac{31}{36}$$

Leave out sandbox example full calculations:

$$\hat{\mathbf{q}}_i = \mathbf{c}_i/m_i \rightarrow$$

$$\hat{\mathbf{q}}_{DR_1} = \begin{bmatrix} 0/3 \\ 2/3 \\ 0/3 \\ 1/3 \\ 0/3 \end{bmatrix}, \quad \hat{\mathbf{q}}_{DR_2} = \begin{bmatrix} 0/3 \\ 1/3 \\ 1/3 \\ 1/3 \\ 0/3 \end{bmatrix}, \quad \hat{\mathbf{q}}_{TV2_1} = \begin{bmatrix} 2/3 \\ 0/3 \\ 0/3 \\ 0/3 \\ 1/3 \end{bmatrix}, \quad \hat{\mathbf{q}}_{TV2_2} = \begin{bmatrix} 0/3 \\ 0/3 \\ 2/3 \\ 0/3 \\ 1/3 \end{bmatrix}$$

$$\hat{\boldsymbol{\rho}}_{-DR_1} = \begin{bmatrix} 0/(0+2/6)=0 \\ 1/3/(1/3+0)=1 \\ 1/3/(1/3+2/6)=1/2 \\ 1/3/(1/3+0)=1 \\ 0/(0+2/6)=0 \end{bmatrix}, \quad \hat{\boldsymbol{\rho}}_{-DR_2} = \begin{bmatrix} 0/(0+2/6=0 \\ 2/3/(2/3+0)=1 \\ 0/(0+2/6)=0 \\ 1/3/(1/3+0)=1 \\ 0/(0+2/6)=0 \end{bmatrix}$$

$$\hat{\boldsymbol{\rho}}_{-TV2_1} = \begin{bmatrix} 0/(0+0)=0 \\ 3/6/(3/6+0)=1 \\ 1/6/(1/6+2/3)=1/5 \\ 2/6/(2/6+0)=1 \\ 0/(0+1/3)=0 \end{bmatrix}, \quad \hat{\boldsymbol{\rho}}_{-TV2_2} = \begin{bmatrix} 0/(0+2/3)=0 \\ 3/6/(3/6+0)=1 \\ 1/6/(1/6+0)=1 \\ 2/6/(2/6+0)=1 \\ 0/(0+1/3)=0 \end{bmatrix}$$

$$\hat{\mathbf{q}}_{DR_1} \cdot \hat{\boldsymbol{\rho}}_{-DR_1} = \begin{bmatrix} 0\cdot 0=0 \\ (2/3)\cdot 1=2/3 \\ 0\cdot 1/2=0 \\ (1/3)\cdot 1=1/3 \\ 0\cdot 0=0 \end{bmatrix} = 1, \quad \hat{\mathbf{q}}_{DR_2} \cdot \hat{\boldsymbol{\rho}}_{-DR_2} = \begin{bmatrix} 0\cdot 0=0 \\ (1/3)\cdot 1=1/3 \\ (1/3)\cdot 0=0 \\ (1/3)\cdot 1=1/3 \\ 0\cdot 0=0 \end{bmatrix} = \frac{2}{3},$$

$$\hat{\mathbf{q}}_{TV2_1} \cdot (1-\hat{\boldsymbol{\rho}}_{-TV2_1}) = \begin{bmatrix} (2/3)\cdot 1=2/3 \\ 0\cdot 0=0 \\ 0\cdot 4/5=0 \\ 0\cdot 0=0 \\ (1/3)\cdot 1=1/3 \end{bmatrix} = 1, \quad \hat{\mathbf{q}}_{TV2_2} \cdot (1-\hat{\boldsymbol{\rho}}_{-TV2_2}) = \begin{bmatrix} 0\cdot 1=0 \\ 0\cdot 0=0 \\ (2/3)\cdot 0=0 \\ 0\cdot 0=0 \\ (1/3)\cdot 1=1/3 \end{bmatrix} = \frac{1}{3},$$

$$\pi^{LO} = \frac{1}{2}\left(\frac{1}{|\text{DR}|}\sum_{i\in\text{DR}} \hat{\mathbf{q}}_i \cdot \hat{\boldsymbol{\rho}}_{-i} + \frac{1}{|\text{TV2}|}\sum_{i\in\text{TV2}} \hat{\mathbf{q}}_i \cdot \left(1-\hat{\boldsymbol{\rho}}_{-i}\right)\right) \rightarrow$$

$$\pi^{LO} = \frac{1}{2}\left(\frac{1}{2}\left(1+\frac{2}{3}\right) + \frac{1}{2}\left(1+\frac{1}{3}\right)\right) = \frac{3}{4}$$

10.12 Characteristics of the polarisation and sentiment gap processes

This note establishes that both polarisation defined by the LO-estimator, $\hat{\pi}_t^{LO}$ and the sentiment gap, $\tilde{X}_t^S = \bar{X}_{DR,t}^S - \bar{X}_{TV2,t}^S$, follow stationary autoregressive processes with heteroskedasticity. If not specified the note draws on test statistics for the period 2015-2024. When the test statistics for the period 2016-2024 deviate from those above we highlight it explicitly. The autoregressive nature of both time-series is presented by the corresponding autocorrelation function plots in figure 38. Further, we report the test statistics for an ADF-test with lag length $k = 1$ in table 20., for both time-series rejecting the null-hypothesis of non-stationarity. We go on to test the range of models used in investigating electoral cycles here exemplified by model M1-M4. For all models we reject the null-hypothesis of homoskedasticity and use HC1 robust std. errors to draw inference from all subsequent estimations, cf. the Breusch-Pagan (BP) statistic in table 20. We do not reject the null hypothesis for homoskedasticity for sentiment gap when looking af 2016-2024, which have no real implications on the results in section 6, cf. table 21. Lastly, based on the Durbin Watson (DW) test statistic we fail to reject the null-hypothesis of no positive autocorrelation in the error term strengthening the case that HC1 robust std. errors are valid for drawing inference based on the selected model specifications.

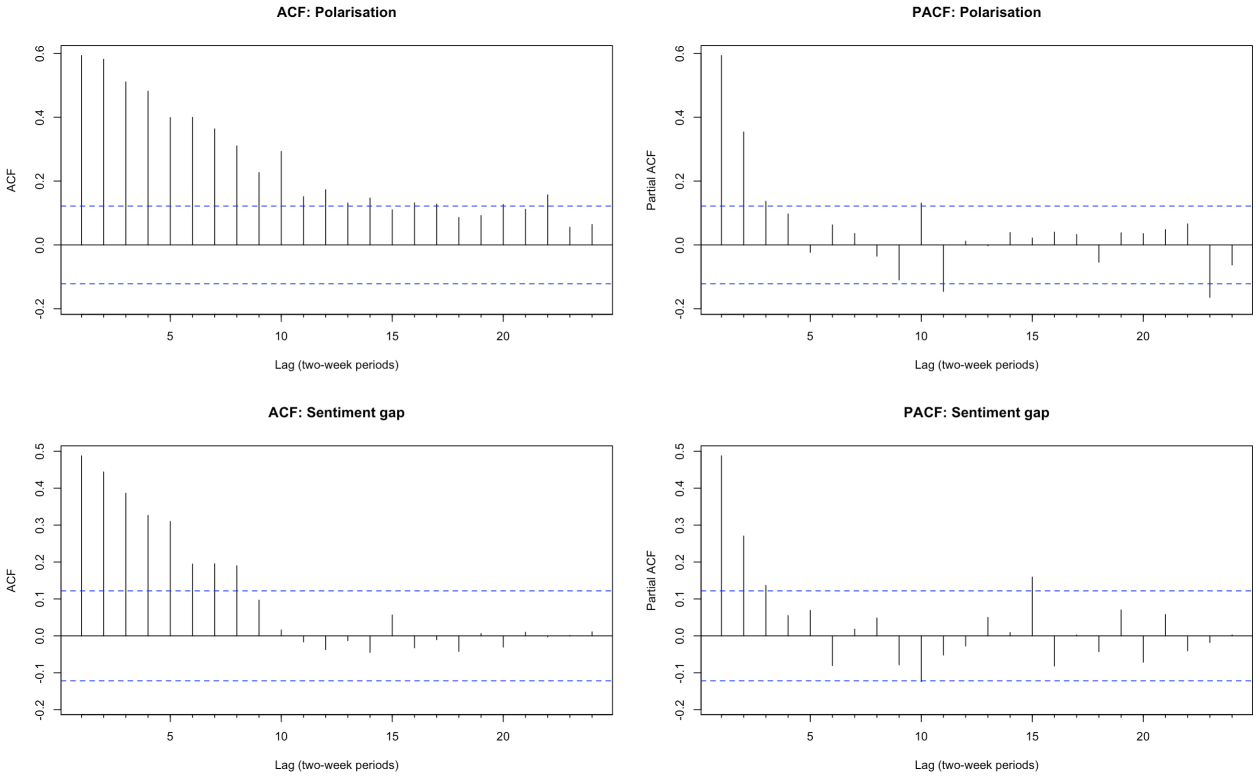


Figure 38: ACF and PACF for polarisation and sentiment gap, 2015-2024.

Table 20: Stationarity, Heteroskedasticity, and Serial Correlation Tests (2015–2024)

Stationarity Tests (Augmented Dickey–Fuller)			
Series	Dickey–Fuller	Lag order	p-value
Polarisation	-5,43	1	0,01
Sentiment gap	-6,49	1	0,01

Breusch–Pagan and Durbin–Watson Tests					
Model	BP stat	df	BP p-value	DW stat	DW p-value
M1: Eq. (11)	52,42	4	$1{,}13 \times 10^{-10}$	2,14	0,83
M2: Eq. (11)	24,42	4	$6{,}57 \times 10^{-5}$	2,32	0,99
M3: Eq. (12)	44,37	4	$5{,}38 \times 10^{-9}$	2,18	0,90
M4: Eq. (12)	25,40	4	$4{,}18 \times 10^{-5}$	2,33	0,99

Table 21: Stationarity, Heteroskedasticity, and Serial Correlation Tests (2016–2024)

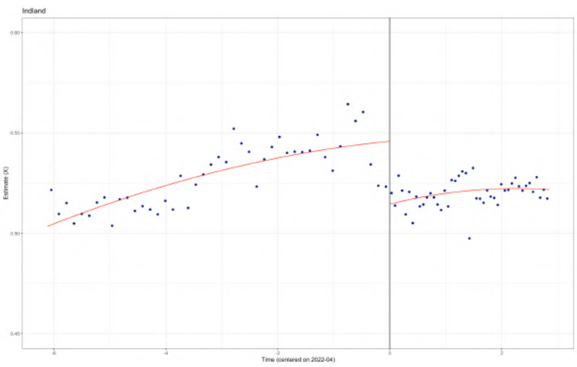
Stationarity Tests (Augmented Dickey–Fuller)

Series	Dickey–Fuller	Lag order	p-value
Polarisation	-5,19	1	0,01
Sentiment gap	-7,22	1	0,01

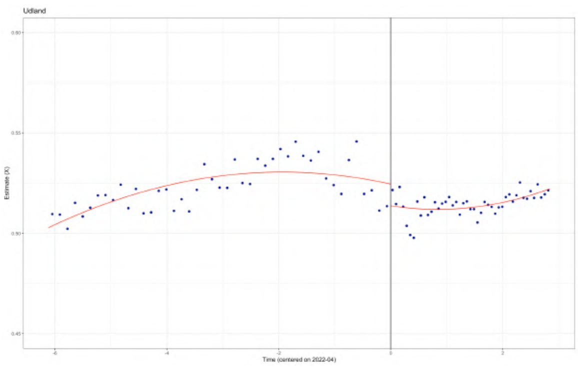
Breusch–Pagan and Durbin–Watson Tests

Model	BP stat	df	BP p-value	DW stat	DW p-value
M1: Eq. (11)	0,96	4	0,92	2,08	0,65
M2: Eq. (11)	10,51	4	0,03	2,35	0,99
M3: Eq. (12)	6,74	4	0,15	2,12	0,78
M4: Eq. (12)	10,52	4	0,03	2,35	0,99

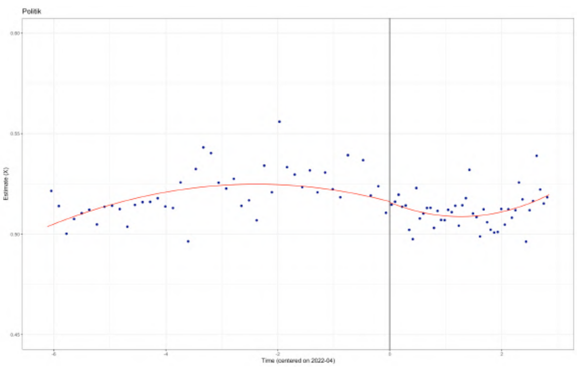
10.13 Discontinuities across sections, centred around biweek 4, 2022.



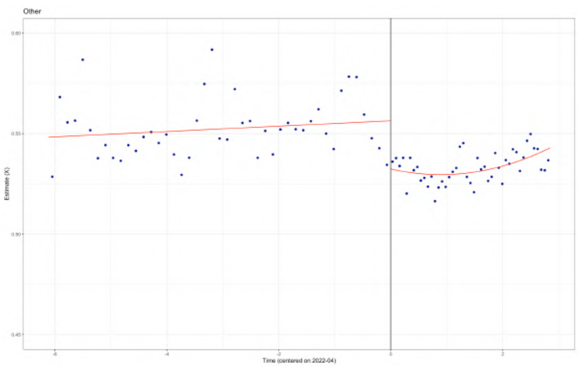
(a) Indland



(b) Udland



(c) Politik



(d) Other

Figure 39: Section specific discontinuities

10.14 Regression results: Seasonality checks

Table 22: ARMAX(1,1)-GARCH(1,1), Polarisation, 2015-2024

	M14	M15	M16	M17	M18
AR(1)	0,999***	0,999***	0,999***	0,845***	0,999***
	(0,004)	(0,003)	(0,003)	(0,123)	(0,004)
MA(1)	−0,619***	−0,663***	−0,650***	−0,565***	−0,619***
	(0,072)	(0,079)	(0,076)	(0,175)	(0,074)
Intercept	0,587***	0,587***	0,587***	0,525***	0,588***
	(0,015)	(0,016)	(0,017)	(0,003)	(0,016)
Valid Token Count	0,002	0,002	0,001	0,000	0,002
	(0,002)	(0,002)	(0,002)	(0,001)	(0,002)
Month Dummies		yes			
Quarter Dummies			yes		
Year Dummies				yes	
Specific Timing Effects					yes
Variance Equation:					
ω	0,000	0,000	0,000	0,000	0,000
	(0,000)	(0,000)	(0,000)	(0,000)	(0,000)
α_1	0,135	0,143	0,137	0,181	0,135
	(0,143)	(0,137)	(0,140)	(0,453)	(0,146)
β_1	0,865***	0,856***	0,862***	0,816**	0,864***
	(0,123)	(0,124)	(0,123)	(0,397)	(0,126)
Observations	260	260	260	260	260
Pseudo-R^2	0,528	0,538	0,537	0,434	0,528
p-value (dummies, jointly)	–	0,127	0,015	1,000	0,898

Note: Valid token count is scaled and centred. Robust standard errors in parentheses. The p-value for joint significance of the dummies is estimated by a Likelihood-Ratio test against M14. *p<0,1; **p<0,05; ***p<0,01

10.15 Front page illustration

We extend our gratitude to Kathrine Andersen Mølby for illustrating the influence of polarisation from online news, cf. figure 40.

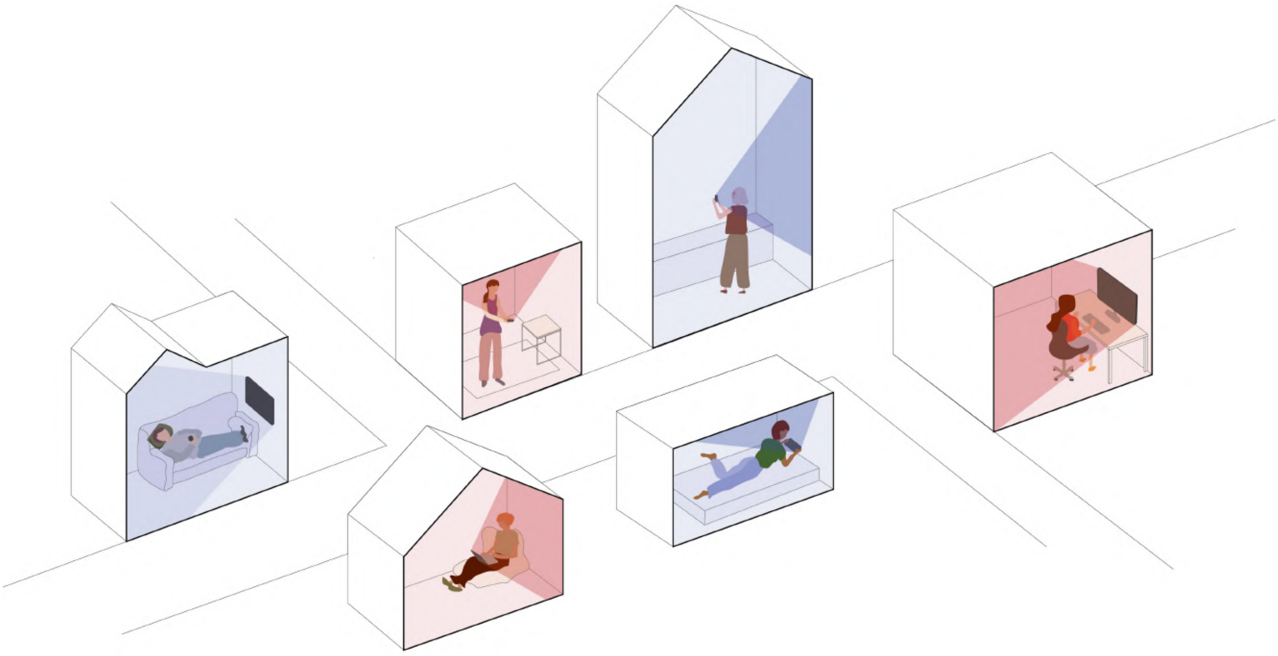


Figure 40: Front page illustration of polarisation in news

10.16 Character count in tex.

This text contains 0+26+0 (1/0/0/0) File: main.tex 1801+8+0 (1/0/1/0) Included file: ./0Abstract.tex 3500+12+50 (1/0/0/0) Included file: ./1Introduction.tex 16986+53+75 (3/0/2/0) Included file: ./2Litterature.tex 55611+434+2404 (22/23/56/0) Included file: ./3Data.tex 23092+143+286 (8/3/173/14) Included file: ./4MeasuringPolarisation.tex 37511+183+1322 (8/11/93/7) Included file: ./5Analysis.tex 11961+290+273 (9/5/17/1) Included file: ./6Robustness.tex 10326+95+227 (4/0/9/0) Included file: ./7Discussion.tex 8768+309+0 (9/0/5/0) Included file: ./8Conclusion.tex 2818+577+1229 (18/23/3/2) Included file: ./9Appendix.tex Sum of files: main.tex 172374+2130+5866 (84/65/359/24) File(s) total: main.tex characters.

Section	Body	Header	Footnote	Total
`main.tex`	0	26	0	26
`1Introduction.tex`	3.500	12	50	3.562
`2Litterature.tex`	16.986	53	75	17.114
`3Data.tex`	55.611	434	2.404	58.449
`4MeasuringPolarisation.tex`	23.092	143	286	23.521
`5Analysis.tex`	37.511	183	1.322	39.016
`6Robustness.tex`	11.961	290	273	12.524
`7Discussion.tex`	10.326	95	227	10.648
`8Conclusion.tex`	8.768	309	0	9.077
Total (excl. Abstract & Appendix)	**167.755**	**1.545**	**4.637**	**173.937**

Table 23: Character counts by section (excluding Abstract and Appendix).

10.17 Robust Regression Results: Election window specifications

Table 24: Robust Regression Results: Eq. (11) - Level effects

	Specification:											
	S, l:6	P, l:6	S, l:13	P, l:13	S, l:26	P, l:26	S, c:3	P, c:3	S, c:6	P, c:6	S, c:13	P, c:13
Lag(1)	0,419***	0,538***	0,338***	0,453***	0,365***	0,450***	0,461***	0,558***	0,457***	0,557***	0,449***	0,540***
	(0,084)	(0,070)	(0,079)	(0,070)	(0,075)	(0,077)	(0,073)	(0,069)	(0,078)	(0,072)	(0,073)	(0,072)
D_{election}	0,011	0,014	0,016**	0,019***	0,012**	0,019***	-0,004	0,008	0,002	0,007	0,003	0,006
	(0,010)	(0,010)	(0,006)	(0,007)	(0,005)	(0,006)	(0,010)	(0,010)	(0,006)	(0,006)	(0,004)	(0,004)
Trend (biweeks)	-0,000	-0,000*	0,000	-0,000	0,000	-0,000	-0,000	-0,000*	-0,000	-0,000	0,000	-0,000
	(0,000)	(0,000)	(0,000)	(0,000)	(0,000)	(0,000)	(0,000)	(0,000)	(0,000)	(0,000)	(0,000)	(0,000)
Trend $\times D_{\text{election}}$	-0,000	-0,000	-0,000**	-0,000***	-0,000*	-0,000***	-0,000	-0,000	-0,000	-0,000	-0,000	-0,000**
	(0,000)	(0,000)	(0,000)	(0,000)	(0,000)	(0,000)	(0,000)	(0,000)	(0,000)	(0,000)	(0,000)	(0,000)
Constant	0,011***	0,248***	0,009***	0,291***	0,009***	0,292***	0,011***	0,238***	0,010***	0,238***	0,010***	0,246***
	(0,002)	(0,037)	(0,002)	(0,037)	(0,002)	(0,041)	(0,002)	(0,036)	(0,002)	(0,039)	(0,002)	(0,038)
Quarterly effects	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Observations	259	259	259	259	259	259	259	259	259	259	259	259
R^2	0,302	0,434	0,333	0,457	0,325	0,462	0,299	0,424	0,290	0,427	0,292	0,431

Note: P: Polarisation, S: Sentiment gap. l: Leading window, c: Centred window. The number indicates the number of biweeks prior to the election included. E.g. "P, c:13" refers to estimates for polarisation, using a centred election window with 13 biweeks before and after the election. Standard errors in parentheses are robust. $^*p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$.

Table 25: Robust Regression Results: Eq. (12) - Dynamic effects

	Dependent variable:											
	S, l:6	P, l:6	S, l:13	P, l:13	S, l:26	P, l:26	S, c:3	P, c:3	S, c:6	P, c:6	S, c:13	P, c:13
Lag(1)	0,454***	0,540***	0,186**	0,455***	0,182**	0,446***	0,489***	0,559***	0,461***	0,556***	0,204**	0,535***
	(0,079)	(0,069)	(0,078)	(0,069)	(0,084)	(0,077)	(0,076)	(0,067)	(0,082)	(0,072)	(0,082)	(0,071)
Lag(1) $\times D_{election}$	0,028	0,023	0,398***	0,031***	0,369***	0,032***	-0,142	0,012	-0,003	0,012	0,341***	0,011
	(0,180)	(0,019)	(0,124)	(0,012)	(0,123)	(0,011)	(0,195)	(0,018)	(0,170)	(0,012)	(0,119)	(0,008)
Trend (biweeks)	-0,000	-0,000*	0,000	-0,000	0,000	-0,000	-0,000	-0,000*	-0,000	-0,000	0,000	-0,000
	(0,000)	(0,000)	(0,000)	(0,000)	(0,000)	(0,000)	(0,000)	(0,000)	(0,000)	(0,000)	(0,000)	(0,000)
Trend $\times D_{election}$	-0,000	-0,000	-0,000**	-0,000***	-0,000	-0,000***	-0,000	-0,000	-0,000	-0,000	-0,000***	-0,000**
	(0,000)	(0,000)	(0,000)	(0,000)	(0,000)	(0,000)	(0,000)	(0,000)	(0,000)	(0,000)	(0,000)	(0,000)
Constant	0,011***	0,247***	0,012***	0,290***	0,012***	0,295***	0,010***	0,238***	0,011***	0,239***	0,012***	0,249***
	(0,002)	(0,036)	(0,002)	(0,037)	(0,002)	(0,041)	(0,002)	(0,036)	(0,002)	(0,038)	(0,002)	(0,038)
Quarterly effects	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Observations	259	259	259	259	259	259	259	259	259	259	259	259
R^2	0,287	0,432	0,348	0,454	0,340	0,458	0,302	0,424	0,289	0,426	0,332	0,431

Note: P: Polarisation, S: Sentiment gap. l: Leading window, c: Centred window. The number indicates the number of biweeks prior to the election included. E.g. "P, c:13" refers to estimates for polarisation, using a centred election window with 13 biweeks before and after the election. Standard errors in parentheses are robust. *p<0.1; **p<0.05; ***p<0.01.

References

Altinget (2021). Valgforskere afblæser generationskampen: Folketingsvalget var et klimavalg for både unge og ældre. Accessed: 2025-05-22. www.altinget.dk/artikel/valgforskere-afblaeser-generationskampen-folketingsvalget-var-et-klimavalg-for-baade-unge-og-aeldre.

Andris, C., D. Lee, M. J. Hamilton, M. Martino, C. E. Gunning, and J. A. Selden (2015, 04). The rise of partisanship and super-cooperators in the u.s. house of representatives. PLOS ONE 10(4), 1–14.

Angrist, J. D. and J.-S. Pischke (2008). Mostly Harmless Econometrics: An Empiricist's Companion. Princeton, NJ: Princeton University Press.

Arutyunov, V. (2025). Regression discontinuity in time (rdit). Accessed: 2025-05-25. https://tilburgsciencehub.com/topics/analyze/causal-inference/rdd/regression-discontinuity-in-time-rdit/.

Ash, E., R. Durante, M. Grebenshchikova, and C. Schwarz (2022, April). Visual Representation and Stereotypes in News Media. CESifo Working Paper Series (9686).

Ash, E., S. Galletta, M. Pinna, and C. S. Warshaw (2023). From viewers to voters: Tracing fox news' impact on american democracy. Journal of Public Economics 222, 104882. Short communication.

Bailey-Morley, A. and C. Kumar (2022). Offentlige fortællinger og holdninger til flygtninge og andre indvandrere: Danmarks landeprofil. Country study, ODI, London. Accessed: 2025-05-20. https://www.odi.org/publications/public-narratives-and-attitudes-towards-refugees-and-other-migrants-denmark-country-profile.

Baker, W. D. and J. R. Oneal (2001). Patriotism or opinion leadership?: The nature and origins of the "rally 'round the flag" effect. The Journal of Conflict Resolution 45(5), 661–687.

Berlingske (2010). Er dr-journalister røde? er tv2's blå. Accessed: 2025-05-22. https://www.berlingske.dk/kronikker/er-dr-journalister-roede-er-tv2s-blaa.

Berlingske (2011, September). Hvor er kvinderne i den offentlige debat. Opinion piece by Anne Sophia Hermansen. Accessed: 2025-05-24. https://www.berlingske.dk/kommentarer/hvor-er-kvinderne-i-den-offentlige-debat.

Bernal, J. L., S. Cummins, and A. Gasparrini (2016, 06). Interrupted time series regression for the evaluation of public health interventions: a tutorial. International Journal of Epidemiology 46(1), 348–355. Accessed: 2025-05-15. https://doi.org/10.1093/ije/dyw098.

Bottomley, C., J. A. G. Scott, and V. Isham (2019). Analysing interrupted time series with a control. Epidemiologic Methods 8(1), 20180010. Accessed: 2025-05-15. https://doi.org/10.1515/em-2018-0010.

Boxell, L., M. Gentzkow, and J. M. Shapiro (2020, January). Cross-country trends in affective polarization. NBER Working Paper 26669, National Bureau of Economic Research, Cambridge, MA. Revised November 2021.

Broberg, F. H. (2024). Hvor vover hun! : Vrede og autisme i mediedækningen af greta thunberg. Kvinder, Køn & Forskning 37(2), 15. Accessed: 2025-05-12. https://tidsskrift.dk/KKF/article/view/141343.

Bursztyn, L., G. Egorov, and S. Fiorin (2020). From extreme to mainstream: The erosion of social norms. The American Economic Review 110(11), pp. 3522–3548.

Caprini, G. (2023, May). Visual Bias. Economics Series Working Papers 1016, University of Oxford, Department of Economics.

Caprini, G. (2024, November). Visual bias. Economics series working papers, University of Oxford, Department of Economics.

Cengiz, P.-M. and L. Eklund Karlsson (2021, 05). Portrayal of immigrants in danish media—a qualitative content analysis. Societies 11, 45.

ChatGPT (2025, May). Chatgpt (may 2025 version) from openai. https://chat.openai.com.

Danmarks Radio (2025). Medieudviklingen 2024. Årsrapport, Danmarks Radio. Accessed: 2025-04-12. https://www.dr.dk/om-dr/fakta-om-dr/dr-analyse/hent-medieudviklingen-2024-som-pdf.

Davidson, L. S., M. Fratianni, and J. von Hagen (1992). Testing the satisficing version of the political business cycle: 1905–1984. Public Choice 73(1), 21–35.

Den Store Danske (2024). Minksagen. Accessed: 2025-05-25. https://lex.dk/minksagen.

Deng, J., J. Guo, Y. Zhou, J. Yu, I. Kotsia, and S. Zafeiriou (2019). Retinaface: Single-stage dense face localisation in the wild. CoRR abs/1905.00641.

Det Nationale Integrationsbarometer (2025). Det nationale integrationsbarometer. Accessed: 2025-05-25. https://integrationsbarometer.dk/.

DR (2025a). Danmarks radio. Accessed 2025-03-22. https://www.dr.dk.

DR (2025b). Vestens storhed og fald 2:3 – vesten set med Østens øjne. https://www.dr.dk/lyd/p1/kampen-om-historien/kampen-om-historien-2025/vestens-storhed-og-fald-2-3-vesten-set-med-oestens-oejne-11032515132. Podcast episode, Kampen om historien, DR P1, 25. marts 2025, Accessed: 2025-05-25.

DR Nyheder (2023). Dan jørgensen om global varmerekord: Alle alarmklokker bimler og bamler. Accessed: 2025-05-22. https://www.dr.dk/nyheder/politik/dan-joergensen-om-global-varmerekord-alle-alarmklokker-bimler-og-bamler.

Enevoldsen, K. C. and L. Hansen (2017). Analysing political biases in danish newspapers using sentiment analysis. Language Works 2(2), 88–98.

Fasching, N., S. Iyengar, Y. Lelkes, and S. J. Westwood (2024). Persistent polarization: The unexpected durability of political animosity around us elections. Science Advances 10(36), eadm9198.

Finlay, P. and C. Argos Translate (2025). Argos translate. GitHub repository: https://github.com/argosopentech/argos-translate. Accessed: 2025-02-26.

Gentzkow, M., J. M. Shapiro, and M. Taddy (2019). Measuring group differences in high-dimensional choices: Method and application to congressional speech. Econometrica 87(4), 1307–1340.

Gift, K. and T. Gift (2015). Does politics influence hiring? evidence from a randomized experiment. Political Behavior 37(3), 653–675.

Grabovets, V. (2022). multi_rake: Multilingual implementation of the rake keyword extraction algorithm. https://github.com/vgrabovets/multi_rake. Accessed: 2025-05-30.

Grimmer, J. and B. M. Stewart (2013). Text as data: The promise and pitfalls of automatic content analysis methods for political texts. Political Analysis 21(3), 267–297.

Groseclose, T. and J. Milyo (2005). A measure of media bias. The Quarterly Journal of Economics 120(4), 1191–1237.

Hansen, K. M. and R. Stubager (Eds.) (2021). Klimavalget: Folketingsvalget 2019, Volume 8 of Studier i dansk politik. København: Djøf Forlag.

Heidari, A. (2024). Clip zero-shot object detection. GitHub repository: https://github.com/deepmancer/clip-object-detection. Accessed: 2025-05-23.

Hetherington, M. J. and T. J. Rudolph (2015). Why Washington Won't Work: Polarization, Political Trust, and the Governing Crisis. Chicago: University of Chicago Press.

Hjorth, F., K. M. Sønderskov, and P. T. Dinesen (2019). Affektiv polarisering i danmark: Et listeeksperiment om social distance til politiske modstandere. Økonomi & Politik 92(3). Udgives af Djøf Forlag.

Hopmann, D. N., P. V. Aelst, and G. Legnante (2012). Political balance in the news: A review of concepts, operationalizations and key findings. Journalism 13(2), 240–257.

Hosseinmardi, H., S. Wolken, D. Rothschild, and D. Watts (2025, 05). Unpacking media bias in the growing divide between cable and network news. Scientific Reports 15.

Huntington-Klein, N. (2021). Event studies. In The Effect: An Introduction to Research Design and Causality, Chapter 17. New York: CRC Press. Accessed: 2025-05-25. https://theeffectbook.net/ch-EventStudies.html.

Internet Archive (2025). Wayback machine. Accessed: 2025-05-22. https://archive.org.

Iyengar, S., Y. Lelkes, M. Levendusky, N. Malhotra, and S. J. Westwood (2019). The origins and consequences of affective polarization in the united states. Annual Review of Political Science 22, 129–146.

Iyengar, S., G. Sood, and Y. Lelkes (2012). Affect, not ideology: A social identity perspective on polarization. Public Opinion Quarterly 76(3), 405–431.

Jensen, J. L. W. V. (1906). Sur les fonctions convexes et les inégalités entre les valeurs moyennes. Acta Mathematica 30, 175–193.

Journalisten (2010). Danskerne: DR er rød, TV 2 er blå. Accessed: 2025-05-22. https://journalisten.dk/danskerne-dr-er-rod-tv-2-er-bla/.

Journalisten (2017, June). Vi er nået dertil, hvor etniske minoriteter godt kan udtale sig som genkendelige medborgere. Accessed: 2025-05-24. https://journalisten.dk/vi-er-naet-dertil-hvor-etniske-minoriteter-godt-kan-udtale-sig-som-genkendelige-medborgere/.

Journalisten (2021). Trafikmålingen lever igen: Her er de 10 største online medier. Accessed: 2025-05-22. https://journalisten.dk/trafikmaalingen-lever-igen-her-er-de-10-stoerste-online-medier/.

Journalisten (2024). Boring but important: Klima fylder mere på mediernes dagsorden. Accessed: 2025-05-22. https://journalisten.dk/boring-but-important-klima-fylder-mere-paa-mediernes-dagsorden/.

Jørndrup, H. (2021, sep). Ligestilling i medierne 2020: Den danske del af undersøgelsen who makes the news? Global Media Monitoring Project.

Killick, R. and I. A. Eckley (2014). changepoint: An r package for changepoint analysis. Journal of Statistical Software 58(3), 1–19. Published: May 6, 2013 (preprint), Journal version: 2014.

Killick, R., P. Fearnhead, and I. A. Eckley (2012). Optimal detection of changepoints with a linear computational cost. Journal of the American Statistical Association 107(500), 1590–1598.

Killick, R., K. Haynes, H. Hullait, I. Eckley, P. Fearnhead, and R. Long (2024). changepoint: Methods for Changepoint Detection. https://CRAN.R-project.org/package=changepoint: CRAN. R package version 2.3. Accesed: 2025-05-29.

Kim, E., Y. Lelkes, and J. McCrain (2022). Measuring dynamic media bias. Proceedings of the National Academy of Sciences 119(32), e2202197119.

Kingzette, J., J. N. Druckman, S. Klar, Y. Krupnikov, M. Levendusky, and J. B. Ryan (2021). How affective polarization undermines support for democratic norms. Public Opinion Quarterly 85(2), 663–677.

Krishnamurthi, L. and S. P. Raj (1991). An empirical analysis of the relationship between brand loyalty and consumer price elasticity. Marketing Science 10(2), 172–183.

Kulturministeriet (2021). Internetbrug og sociale medier 2021. Analyse, basisrapport, Kulturministeriet. Accessed: 2025-05-22. https://kum.dk/fileadmin/_mediernesudvikling/2021/Internetbrug_og_sociale_medier_2021.pdf.

Lauridsen, G. A., J. A. Dalsgaard, and L. K. B. Svendsen (2019a). Sentida, github repository. https://github.com/Guscode/Sentida. Accessed: 2025-05-23.

Lauridsen, G. A., J. A. Dalsgaard, and L. K. B. Svendsen (2019b, Sep.). Sentida paper. Journal of Language Works - Sprogvidenskabeligt Studentertidsskrift 4(1), 38–53.

Lind, S. (2021). Lemmy: A lemmatizer for danish and swedish. https://github.com/sorenlind/lemmy. Accessed: 2025-05-23.

Liu, H. et al. (2024). Llava: Large language and vision assistant. GitHub repository: https://github.com/haotian-liu/LLaVA. Accessed: 2025-02-24.

Liu, H., C. Li, Q. Wu, and Y. J. Lee (2023). Visual instruction tuning. In NeurIPS.

Mackinac Center for Public Policy (2023). The overton window. Accessed: 2025-05-26. https://www.mackinac.org/OvertonWindow.

Mølby, C. A. and K. A. H. Bremholm (2025). Feature extraction Github repository. GitHub repository: https://github.com/Christianmoelby/Feature-extraction-Danish-news. Contains essential code to scrape and collect Danish news and extract features from text, images and faces. Formats features to a token-dataset. We reserve the right to continuously update and maintain the repository.

Nedergaard, P. (2025). EU's flygtningekrise 2015. Lex.dk. Accessed: 2025-05-22. https://lex.dk/EU%27s_flygtningekrise_2015.

New York Times, The Daily (2025, may). A conversation with vice president vance. Podcast audio, 6:00, approx. https://www.nytimes.com/2025/05/22/podcasts/the-daily/vice-president-vance-pope-politics.html.

Nimb, S., S. Olsen, and T. Troelsgård (2022). Dds: The danish sentiment lexicon. https://github.com/dsldk/danish-sentiment-lexicon. Compiled by Det Danske Sprog- og Litteraturselskab (DSL) and Center for Sprogteknologi (CST), University of Copenhagen. Licensed under CC-BY-SA 4.0 International. Accessed: 2025-05-23.

Ollama (2024). Llava model on ollama. https://ollama.com/library/llava. Accessed: 2025-05-23.

OpenAI (2021). Clip: Connecting text and images. https://openai.com/index/clip/. Accessed: 2025-05-23.

Pew Research Center (2022, June). Journalists sense turmoil in their industry amid continued passion for their work. Accessed: 2025-05-22. https://www.pewresearch.org/journalism/2022/06/14/journalists-express-high-job-satisfaction-but-are-concerned-about-the-state-of-the-news-industry/,.

Pew Research Center (2024, May). Americans remain critical of china. Accessed: 2025-05-22. https://www.pewresearch.org/global/2024/05/01/americans-remain-critical-of-china/.

Piskorski, J., N. Stefanovitch, G. Jacquet, and A. Podavini (2021). Exploring linguistically-lightweight keyword extraction techniques for indexing news articles in a multilingual set-up. In Proceedings of the EACL Hackashop on News Media Content Analysis and Automated Report Generation, Online, pp. 35–44. Association for Computational Linguistics.

Politis, D. N., J. P. Romano, and M. Wolf (2001). On the asymptotic theory of subsampling. Statistica Sinica 11(4), 1105–1124.

Python Software Foundation (2025). Python language reference, version 3.12. https://www.python.org/. Accessed: 2025-05-24.

R Core Team (2025). R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing. Accessed: 2025-05-22. https://www.R-project.org/.

Radford, A., J. W. Kim, C. Hallacy, A. Ramesh, G. Goh, S. Agarwal, G. Sastry, A. Askell, P. Mishkin, J. Clark, G. Krueger, and I. Sutskever (2021). Learning transferable visual models from natural language supervision. arXiv preprint arXiv:2103.00020.

Redmon, J., S. Divvala, R. Girshick, and A. Farhadi (2016). You only look once: Unified, real-time object detection. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 779–788.

Ritzau (2024). Danskerne: Dr er rød, tv 2 er blå. Accessed: 2025-05-22. https://danishnews.ritzau.com/article/583886e7-8186-4e52-a61a-6eff9f6994041.

Rose, S., D. Engel, N. Cramer, and W. Cowley (2010). Automatic keyword extraction from individual documents. In M. W. Berry and J. Kogan (Eds.), Text Mining: Applications and Theory, pp. 1–20. John Wiley & Sons.

Russell, N. J. (2006, January 4). An introduction to the overton window of political possibilities. Mackinac Center for Public Policy. Accessed: 2025-05-26. https://www.mackinac.org/7504.

Schneidermann, N. S. and B. S. Pedersen (2022). Evaluating a new danish sentiment resource: the danish sentiment lexicon, dsl. In Proceedings of the SALLD-2 Workshop @ LREC 2022, Marseille, France, pp. 19–24. European Language Resources Association (ELRA).

Schroff, F., D. Kalenichenko, and J. Philbin (2015). Facenet: A unified embedding for face recognition and clustering. CoRR abs/1503.03832.

Schultz, K. A. (1995). The politics of the political business cycle. British Journal of Political Science 25(1), 79–99.

Serengil, S. I. (2024). Deepface: A lightweight face recognition and facial attribute analysis framework. GitHub repository :https://github.com/serengil/deepface. Accessed: 2025-02-24.

Serengil, S. I. and A. Ozpinar (2024). A benchmark of facial recognition pipelines and co-usability performances of modules. Bilisim Teknolojileri Dergisi 17(2), 95–107.

Slots- og Kulturstyrelsen (2017). Internettrafik – Mediernes Udvikling i Danmark 2017. Technical report, Slots- og Kulturstyrelsen, København. Accessed: 2025-05-30. https://kum.dk/fileadmin/_mediernesudvikling/2017/Internettrafik_2017.pdf.

Steiner, N. D., R. Berlinschi, E. Farvaque, J. Fidrmuc, P. Harms, A. Mihailov, M. Neugart, and P. Stanek (2023). Rallying around the eu flag: Russia's invasion of ukraine and attitudes toward european integration. JCMS: Journal of Common Market Studies 61(2), 283–301.

Taddy, M. (2013). Measuring political sentiment on twitter: Factor optimal design for multinomial inverse regression. Technometrics 55(4), 415–425.

Taddy, M. (2015, September). Distributed multinomial regression. The Annals of Applied Statistics 9(3).

The New York Times (2019, February 26). How the politically unthinkable can become mainstream. The New York Times. Accessed: 2025-05-26. https://www.nytimes.com/2019/02/26/us/politics/overton-window-democrats.html.

Thomas J. Nechyba (2018). Intermediate Microeconomics - An intuitive approach with calculus. Cengage Learning EMEA. Chapter 26. Product differentiation, subsection A.

TV2 (2025). Tv2 nyheder. Accessed: 2025-05-22 https://tv2.dk.

Udlændinge- og Integrationsministeriet (2023, February). Kriminalitet blandt mandlige indvandrere og efterkommere fra menapt-landene. Technical report, Det Nationale Integrationsbarometer. Accessed: 2025-05-22. https://integrationsbarometer.dk/tal-og-analyser/filer-tal-og-analyser/arkiv/NotatvedrrendekriminalitetenblandtMENAPT.pdf.

VIVE (2022). Polarisering og tilslutning til demokratiske normer – Danskernes syn på politiske modstandere og demokratiet. København: VIVE – Det Nationale Forsknings- og Analysecenter for Velfærd. Written by Rasmus T. Pedersen, Julian Christensen, and Niels Bjørn Grund Petersen.

Zetland (2022, June). Polariseringen vokser i vesten, men ikke danmark. et nyt studie forklarer hvorfor. Accessed: 2025-05-24. https://www.zetland.dk/historie/seg6Q4a2-aekdm0J7-94432.

Zheng, L., W.-L. Chiang, Y. Sheng, S. Zhuang, Z. Wu, Y. Zhuang, Z. Lin, Z. Li, D. Li, E. P. Xing, H. Zhang, J. E. Gonzalez, and I. Stoica (2023). Judging llm-as-a-judge with mt-bench and chatbot arena.