Humboldt-University Berlin
Faculty of Humanities and Social Sciences
Department of Social Sciences

Qualification thesis for the academic title "Bachelor of Arts"

Government Change, Media Change? Investigating political Bias in German Economic News Reporting:

Are German newspapers politically biased under different governments in their reporting about the German economy?

—

Regierungswechsel, Medienwechsel? Untersuchung von politischem Bias in der deutschen Wirtschaftsberichterstattung:

Sind deutsche Zeitungen unter verschiedenen Regierungen in ihrer Berichterstattung über die deutsche Wirtschaft politisch voreingenommen?

First Advisor: Prof. Dr. Heike Klüver, Humboldt-Universität zu Berlin
Second Advisor: Dr. Bastian Becker, Humboldt-Universität zu Berlin

Submitted by:

Noah Badenhoop
Lübecker Straße, 10559 Berlin
noah.badenhoop@hu-berlin.de

B.A. Social Science
Registration Number: 621442
Date of submission: Berlin, November 15, 2024

Scripts, data, graphs and tables are available here: `https://github.com/noahhoop/BA_thesis`

The scripts for the data analysis were created with the assistance of AI. For the following text I used the tool Grammarly but only for commas and synonyms of words that aren't important concepts.

Contents

List of Figures

List of Tables

1. Introduction

A well-functioning media landscape is vital for democracies. As the so-called Fourth Estate (e.g., McChesney and Pickard, 2014), the media have the role of informing the public so it can make educated political choices. The media also fulfills a "watchdog" function (Schudson, 2014), as they are supposed to investigate those in power, uncover wrongdoings, and critically report about political actors. Over the last years, the usage of social media as a news source has been increasing (Newman et al., 2024), but there are growing concerns about fake news (Aïmeur et al., 2023), especially with developments in AI technology making this more efficient (Bashardoust et al., 2024; Koplin, 2023). The role of traditional media as gatekeeper (Shoemaker and Vos, 2009) has softened due to Web 2.0, but in regard of fake news and AI developments, a gatekeeper for trustworthy information might regain relevance.Some text with a footnote.[1] To be a trustworthy source, the media

[1] Some newspapers already toy with the idea of using AI for their news reporting, while others like *Süddeutsche Zeitung* advertise "Truth cannot be generated but only investi-

should be unbiased, as they can influence people's opinions and political behavior (Foos and Bischof, 2022; Gerber et al., 2009; DellaVigna and Kaplan, 2007)).

While there are many studies on media bias (for a review: Rodrigo-Ginés et al., 2024; Groeling, 2013), there are also a lot of different definitions and types of bias. In one set of definitions media bias is defined as deviation from neutral and balanced reporting (e.g., D'Alessio and Allen, 2000). Here, studies mostly describe what ideological or partisan leaning a newspaper has. This approach, though, implies that the "middle" is the correct point of view, which is a normative question. Others acknowledge that there is a difference between opinion and bias (Groeling, 2013). Another way to define bias is as deviation of truth, which I argue can even be compatible with having different opinions on what truth is. Using this definition, though, it is more difficult to actually measure bias. Another issue in this line of research is the unobserved population problem. This states that we generally only observe news articles, which are a sample of all events happening, and this sample can be subject to selection bias (Groeling, 2013).

In my study, I address these issues by looking at specific economic topics where there's agreement about what is good and bad, and also enables me to control relatively objectively for the "reality". My research question is *whether newspapers are politically biased under different governments in their reporting about the economy.* My contribution to the literature lies in a novel methodological approach where I use the most current AI technology with the Large Language Model (LLM) llama3.1 to conduct the text annotation. This allows me to achieve a sample with specific topic requirements and also use information on the articles' sentiment rather than only the number of articles. While some years ago quantitative computer analysis of text data was concluded to be not good enough for this area (Grimmer and Stewart, 2013), I show that this technology can achieve mostly acceptable results, even using only the smallest model. As it is also relatively simple to use, this opens up many possibilities for researchers as it doesn't necessitate human coders and therefore a lot of time and costs.[2] I also highlight the many possibilities to conceptualize media bias and why researchers should be clear in their research what kind of bias they claim to detect. My study focuses on Germany, which has only been scarcely studied in regards to media bias, during the last two cabinets, one led by center-right parties CDU/CSU and one by center-left party SPD. As the latter cabinet is subject to a lot of criticism regarding the economic performance, it will be interesting to see if this is proportional to the actual economic indicators. The remainder of this article is structured as follows: In section 2, I will lay out the relevant literature and theories regarding media bias. In section 3, I describe the selection of Germany as a case in more detail and present my hypothesis. In section 4, I will describe the data, the procedure of the text analysis, and the methods used. In section 5, I will present descriptive statistics and the results of the analysis. Section 6 will conclude with a summary and discussion of the results, including limitations and prospects for future research.

2. Theory and Literature Review

2.1. Role and Influence of the Media on Political Behavior and Opinions

In theories about democracy, the media play a vital role and are generally considered the Fourth Estate (McChesney and Pickard, 2014; McNair, 2018; Kaid and Holtz-Bacha, 2008;

gated". (https://www.spiegel.de/netzwelt/web/kuenstliche-intelligenz-axel-springer-und-openai-verkuenden-globale-partnerschaft-a-d32dc28b-0382-4ee8-9296-dcab849329a4)

[2]It does, though, require a lot of computational power or the usage of an API, which depending on the amount of text can mean some costs and concerns regarding data protection.

Louw, 2005). In this role, they should inform the public and fulfill a watchdog function over the government and other political actors. In general, journalists in Western democracies hold themselves to standards of objectivity and even neutrality to some degree (Patterson and Donsbagh, 1996; Pfetsch and Adam, 2008, p. 10; Dortmund, 2024). A necessary condition for this function is that the media are free from influence by the government. This is called media capture (Prat, 2015). De jure, in Germany, this is granted via the constitution. In practice, though, a complete independence between the media and the government is not feasible. Gans (2004, p. 116) describes a symbiotic relationship: "The relationship between sources and journalists resembles a dance, for sources seek access to journalists, and journalists seek access to sources." The media are not only transmitters of information and a "forum of the political process" but also political actors themselves (Pfetsch and Adam, 2008, p. 11). Surveys show that many politicians believe the media has strong agenda-setting and even career-controlling power (Lengauer et al., 2014) and even more power than themselves (Maurer, 2011). Expert interviews confirm these findings and highlight the "journalistic power" that publishers like Axel Springer have (Brinkmann, 2018, p. 566). Van Aelst et al. (2008a) argue that "the media may have a greater influence on how politicians communicate to the public than on what they communicate" (p. 495). Politicians can adjust their communication to the media logic to improve their chances of getting media attention (Strömbäck and Nord, 2006; Van Aelst et al., 2008b). Waldherr (2008, p. 18) similarly argue that politicians internalize the rules of media logic. A survey of German journalists also shows that they perceive the influence of journalists on politics to be stronger than the other way around (Dortmund, 2024).

Another point highlighting the importance of studying potential biases in media reporting is the influence of the media on public opinion. Cohen (1964, p. 13) famously wrote the press "may not be successful much of the time in telling people what to think, but it is stunningly successful in telling its readers what to think about." Since then, agenda-setting theory has been widely discussed in the literature (McCombs and Shaw, 1972). The media are gatekeepers (Shoemaker and Vos, 2009) and can decide which topic to publish. They can also choose to frame it in their own way or adopt frames of external actors and choose which experts or other sources to cite (Waldherr, 2008, p. 175). In line with that Eyck and Williment (2003) differentiate between access and coverage. With access, there is direct contact between the journalist and the external actor, giving the latter more opportunity to frame the topic in their own way. At the same time, with coverage, the other party doesn't have the opportunity to tell their side of the story, and the piece is published through the more critical lens of the journalist. While it is intuitive that the media influences people, as it is the primary source of information about political events, it is very difficult to empirically prove this effect due to endogeneity. Newspapers can set the agenda on one side, but they can also react to shifts in public opinion (Foos and Bischof, 2022). Only a few studies have addressed this issue. Foos and Bischof (2022) exploit an exogenous local boycott of a newspaper and show a long-term decrease in euroscepticism in this region. Another example by Gerber et al. (2009) employs a field experiment in a US state, in which they hand out free copies of different newspapers. Here, they find support for an increase in support for Democrats and some support for increased voter turnout. While being able to make causal claims, these studies suffer from external validity as they study specific cases in time and space. Another study by DellaVigna and Kaplan (2007) exploits a natural experiment and finds that the introduction of the Fox News Channel to new cities in the USA increases votes for the republican party.

Economic news reporting even has a term called "vibecession", which describes a disconnect between people's perception of the economy and its actual state. In the US context, news media recently reported on this phenomenon.[3] This doesn't necessarily have to be a bias, as a well-performing economy on the macro level doesn't mean that everyone will benefit from it. The polls showing that republican voters are a lot more likely to believe that the economy is performing worse, though, indicates that this could be due to bias. Using individual data from Turkey, Yagci and Oyvat (2020) indeed find support for the influence of media and partisanship on the perception of the economy. In the German context, Garz (2018) reports worse perceptions of unemployment, with increased reporting on this topic. News coverage about the economy and peoples' opinions about it are also linked to differences in consumption (Doms and Morin, 2004; Ludvigson, 2004).

2.2. What is Media Bias?

The literature about media bias has no dominant theory or definition about what it is. The first crucial question is: "Bias relative to what? Relative to what a 'neutral' or 'fair' or 'balanced' outlet would do? Relative to the average or median preference that citizens have—or would have, if fully informed? Relative to the average preference of voters (sometimes a small and unrepresentative subset of citizens), or the outlet's audience (an even smaller and probably more unrepresentative Subset of citizens)?" (Puglisi and Snyder, 2015, p. 648) In their paper, D'Alessio and Allen (2000) assume that the media should be perfectly balanced towards all sides. "Thus, coverage should be roughly equal for each side, any departure from a '50-50'[4] split could be considered a consequence of some kind of bias" (Ibid. p. 137). They argue that this should be the case for the amount of coverage and opinionated statements about each party. However, this assumption does not account for cases of different parties' performances. If government A receives very positive media coverage while government B receives poor coverage, this could be explained by a poor performance of government B. Or, in the case of climate change, it would be unreasonable to give climate change deniers the same coverage as climate change scientists. Larcinese et al. (2011) try to measure incumbent quality by electoral experience but the validity of such measurements stays unclear. Furthermore, there should be a distinction between opinion and bias: "Even if an individual or a news organization favors one partisan entity over another in their political view, the existence of a political opinion is neither necessary nor sufficient to justify the conclusion that the news they produce would be biased." (Groeling, 2013, p. 133) (Rodrigo-Ginés et al., 2024, p. 4) argue that "the line that separates bias from opinion depends on whether the journalist uses rhetorical artifacts that distort the information to support his opinion or not." Another aspect I would add is leaving out specific parts of a story that paint a (less) favorable picture. There certainly is use in the localization of newspapers in the ideological sphere. Still, researchers should be clear when talking about bias if they're rather describing which party or ideology a newspaper favors or if they claim there's a bias in the stricter sense. This argument entails the normative question of how newspapers should report and how much opinion should be in it, but it also matters empirically as you're measuring different concepts.

One answer to the question "bias relative to what?" is reality. Here, the central problem is the so-called "Problem of the Unobserved Population" (Groeling, 2013, p. 137).

[3] https://www.cnbc.com/2024/08/12/59percent-of-americans-think-the-us-is-in-a-recession-report-finds.html, https://www.theguardian.com/us-news/article/2024/may/22/poll-economy-recession-biden, https://www.nytimes.com/2024/05/23/opinion/biden-trump-vibecession.html

[4] They look at the US case. Therefore the 50-50 split refers to Democrats and Republicans.

Countless events are happening in the world, and journalists have to decide which they will report on as they can't cover every story. The finished news reports are, therefore, a sample of all those events and the problem is that it is very difficult to measure the distribution of all events that happened. Of course, how reality is perceived depends on ideology as well. Table 1 illustrates this problem, albeit in a simplified way. One can infer if a story is biased by comparing the story to reality, but one can only do so when both the story and reality are observed. Events 2 show how the selection of events to report on can result in a bias. Events 3 highlight the problem of opinion versus bias since another problem is knowing if the story is actually the newspaper's opinion or a skewed version of the event. An example could be reporting about a country's debt. German economically right parties hold the position that new debt should be avoided and are proponents of the debt brake. The left parties criticize the debt brake and are in favor of taking on new debt for investments. If an economically right party in government takes on a lot of new debt this would be against their actual position but an economically right newspaper could still try to frame it in a positive way by only highlighting certain aspects.

In the literature, many types of bias are discussed.[5] These range from the selection and omission of certain information to different ways the information is phrased as well as how it is presented in the news. This includes, for example, if a story is well visible on the front page or which pictures are shown. There are many ways to categorize these types of bias. Mullainathan and Shleifer (2002) distinguish between ideology and spin bias. The former means the newspaper wants to influence readers towards a certain ideology. In contrast, the latter means that newspapers want to create a simple and memorable story, which leads to omitting important information. D'Alessio and Allen (2000) distinguish between gatekeeping, coverage, and statement bias. The first refers to the media's role as gatekeepers, as they can decide which stories they want to report. They can paint a distorted picture of reality by selecting only certain stories. Coverage bias signifies how much coverage a newspaper spends on certain entities, such as parties or politicians. Lastly, statement bias means that a news piece is biased when it contains more favorable statements toward one side than another. Groeling (2013) also uses selection bias, the same as gatekeeping bias, as a category. As a second category, he uses presentation bias, which is a broader category than statement bias as it entails all ways the information can be presented in a distorted way. Instead of other categories, he uses an important qualifier for the term media bias and talks about "partisan"[6] media bias to indicate in what regard there is a bias. This can be contrasted with structural types of bias. This is, for example, the tendency of newspapers to report on negative stories (Soroka, 2012; Dewenter et al., 2018) or focus on stories that sell better. While it conceptually makes sense to distinguish types of biases by these qualifiers, in reality, they can overlap. Imagine two parties A and B, where party A has equally positive and negative performances while party B is overall neutral. On average, the performance of these parties is the same. However, a structural bias towards negative stories would lead to overrepresenting party A's negative performances and, therefore, worse news coverage.

[5]For a fine-grained overview of types of bias, view Rodrigo-Ginés et al. (2024)

[6]Depending on the context of research this could of course be replaced with ideological or political in general

Table 1: Conceptual examples

Events 1 perceived by Newspaper A	Events 1 perceived by Newspaper B	Resulting story Newspaper A	Resulting story Newspaper B	Favored by newspaper A	Favored by newspaper B	Biased
Good for party X, bad for party Y	Bad for party X, good for party Y	Good for party X, bad for party Y	Bad for party X, good for party Y	X	Y	No
Events 2 perceived by Newspaper A	Events 2 perceived by Newspaper B	Resulting story Newspaper A	Resulting story Newspaper B	Favored by newspaper A	Favored by newspaper B	Biased
Good for party X, good for party Y	Good for party X, good for party Y	Good for party X	Good for party Y	X	Y	Yes
Events 3 perceived by Newspaper A	Events 3 perceived by Newspaper B	Resulting story Newspaper A	Resulting story Newspaper B	Favored by newspaper A	Favored by newspaper B	Biased
Bad for party X	Bad for party Y	Good for party X	Good for party Y	X	Y	Yes

Another point of contention in the literature is whether bias is intentional. Some studies clearly include in their definition that bias is intentional (Williams, 1975; D'Alessio and Allen, 2000; Gentzkow and Shapiro, 2006; Mullainathan and Shleifer, 2002), while in others, it can be both intentional or unintentional (Groeling, 2013; Spinde et al., 2021). This distinction can help conceptually, as some types of bias can be both. However, it also matters when talking about the nature of bias and the reasons for its existence. Proving that something was intentional, though, is extremely difficult, to say the least. Ultimately, it depends on the specific context and research question, which definition and concepts of bias a researcher should use.

2.3. Empirical Studies on detecting Bias

From an economic background, some papers study news markets and which conditions increase or decrease bias. Generally, these studies examine the role of competition on the amount of bias. One argument is that a newspaper can only slant its news with a monopoly, but readers will choose an alternative when there's competition. This view is supported by the model of Gentzkow and Shapiro (2006). The opposite argument says a more competitive market can cater to more extreme views, leading to more bias (Mullainathan and Shleifer, 2005). In the political sciences, the existing research primarily focuses on partisan bias. But as there are many types of bias their approaches can differ a lot. In the US context, there's a long debate about whether there's a liberal/democrat or conservative/republican bias in the news. In a meta-analysis, D'Alessio and Allen (2000) find no significant bias in either direction when looking at news coverage of presidential campaigns. One widely cited study by Groseclose and Milyo (2005) measures the ideology of think tanks by how much politicians cite them. They then predict the ideology of news based on which think tanks they cite. Overall, they find a clear liberal bias, with only one FOX news program having a republican bias. One issue with this approach is that it could be the case that left-leaning think tanks produce higher quality reports or a higher frequency of reports, leading to more citations. Another point is that this also assumes that the reports by think tanks are equally shared by politicians and the media. Otherwise, the ideological score by a think tank could be produced because more left politicians cite its left-leaning reports, but the media does report more on right-leaning reports of a think tank.

Another study focuses on political scandals (Puglisi and Snyder, 2011). This largely solves the problem of the unobserved population since the population is defined as all scandals that happened. One issue could only be that not all scandals are uncovered and reported on. If the behavior of one party is more heavily scrutinized and therefore leads to more reporting, this could still lead to a biased sample. However, as the study observes, whether a pro-democrat/republican newspaper reports less on democrat/republican scandals, this shouldn't be a problem. The authors indeed find that pro-democrat newspapers are more likely to report on republican scandals and vice versa.

Further studies specifically look at news reporting on the economy. This can also be a solution for the unobserved population problem. The reality, therefore, is how the economy is doing, which can be quite easily measured compared to other concepts, such as government performance. I argue that this works especially well for aspects of the economy where there is no real difference in opinions. Generally, all parties agree that low unemployment, high economic growth, and low inflation are good. One should be careful about which economic indicators to choose as there are different opinions about the level of debt, spending, or tax revenue, for example. Larcinese et al. (2011) examine news reporting about unemployment, trade deficit, and inflation in the US. They find robust bias for unemployment, with pro-democrat newspapers reporting less about high unemployment during a democratic president and vice versa for republican newspapers under Republican presidents. Additionally, they find some support for bias in the news about the trade deficit and no bias in inflation. The study focuses on selection bias as they only look at the number of articles published but not how they are framed. Lott and Hassett (2014) also include the sentiment of the news, even though they only look at the headline. Specifically, they look at the share of positive news about GDP, unemployment, durable goods, and retail sales while controlling for these indicators. Overall, they find that republican presidents receive 20-30% less positive coverage. These results hold when looking at negative coverage. Lowry (2008) find mixed results for a bias in TV news about the stock market. Soroka (2012) has a unique way of thinking about the unobserved population problem. He creates a distribution of real-world events by choosing changes in unemployment and inflation. This distribution is then compared to the distribution of negative and positive stories about these topics to infer the so-called gatekeeping function by which a newspaper selects stories. This way, he identifies a bias against Democrats for unemployment news and a bias against Republicans for news about inflation. In the German context, there are not many empirical studies on media bias. One study by Dewenter et al. (2020) develops an index to "measure (...) the relative positioning of media within the political spectrum." They aim to test if the media are more critical of parties in government by seeing if the media are more right-leaning under left governments and vice versa. Finding support for this hypothesis, they conclude that the media fulfills its role as the fourth estate. This conclusion is questionable as it's unclear whether a simple shift to the right/left means that the media are more critical. And as they admit themselves, it's difficult to control for the government's performance, which they do by using three economic indicators. Garz (2014) specifically examines news reporting about unemployment in Germany and also compares how specific newspapers report under different governments. He mainly finds a bias against the left government by the tabloid BILD.

2.4. Causes of Media Bias

Finally, it is important to discuss the causes of media bias, which are generally divided into supply- and demand-side explanations.

The supply side refers to factors on the side of journalists, editors, and publishers. Those who create the news. Journalists hold different political beliefs and might want to convince readers of their position by framing it in a positive but distorted way. This could also be an unconscious process as they are subject to cognitive biases. Due to confirmation bias (Klayman, 1995), they could be more likely to believe facts that suit their ideology, leading them to ignore other facts. In an experimental setting, Taber and Lodge (2006) show how people are more likely to uncritically believe in facts that support their views while trying to discredit opposing arguments. In the field, it is virtually impossible to measure the intent or unconscious processes of journalists. However, in a study by Patterson and Donsbagh (1996), they employed a quasi-experiment in a survey where they presented journalists with hypothetical situations and tasked them with various news decisions. This resulted in strong correlations between the political leaning of the news decisions and their own political views. The authors also show a strong correlation between German journalists' partisan belief and their perception of the partisan editorial position of their news organization.

Some case studies show the influence of ownership on differences in reporting. Larcinese et al. (2011) show that the previously conservative Los Angeles Times shifted its reporting to the left once a new publisher took over who was known to be more liberal. This shift did not coincide with an increase in voters for the Democrats in California. Durante and Knight (2012) show that the public news television network favored Berlusconi's party once it got into power and through this control over the network. In the case of Israel, Grossman et al. (2022) describe how a billionaire influenced the political sphere by handing out a free newspaper with slanted coverage to the right. Businesses, in general, can also try to influence public opinion. In a study interviewing politicians and lobbyists, left-wing politicians address the problem of discussing topics of taxation, which would be answered by campaigns highlighting alleged negative consequences: "Another communication strategy of the business side is to emphasize adverse effects of taxes on jobs and growth." (Fastenrath et al., 2022, p. 774) If a government doesn't act in the businesses' interest, they could try to push the narrative of a failing economy.[7]

Demand-side factors exist as most newspapers are profit-driven companies and, therefore, cater to what readers want. The same cognitive biases mentioned above also exist for readers. Because of confirmation bias, people prefer to read what confirms their views. Measuring the effect of the demand side is very difficult due to high data requirements and the problem of endogeneity. A few studies try to use proxies to measure the demand side. Gentzkow and Shapiro (2010) use the amount of donations to a party in a zip code to measure the political orientation where the newspaper is sold. They combine that data with newspaper circulation data and find changes in consumer demand predict slant in news coverage. On the other side, they find weaker effects for the ideological position of the owners. Puglisi and Snyder (2011) use the vote share in a county and weight it by how many newspapers are sold there. They find that news coverage responds to demand for local scandals. The same measure is used by Larcinese et al. (2011), but they only find an effect for one of the economic topics they examined. Using aggregated data is prone to the ecological fallacy since it remains unclear if the people voting for one party also actually read the respective newspaper.

3. Case Selection and Hypothesis

In my study, I aim to identify a political bias of newspapers towards different governments in their reporting about economic news. I will base the definition of political media bias on Groeling (2013, p. 133): I define political media bias as a significantly distorted portrayal of reality that systematically and disproportionately favors one political side over the other. As a case, I choose Germany during the current and previous cabinet, the former being run by the center-left party SPD, in a coalition with the Greens and the liberal party FDP, and the latter by the center-right party CDU/CSU, in a coalition with the then smaller SPD.[8] To my knowledge, there are no other studies on current time periods in Germany on this topic. The last one is by Garz (2018), which looks at the period between 2001-2010 and mainly has a different focus.[9]

[7] Among my sample I indeed found articles which are even authored by business owners in which they talk about the state of the economy and what they think needs to be changed.

[8] Data from the Manifesto Project (Lehmann et al., 2024) reports a left-right scale for each party. By taking the mean of the cabinet parties confirming that the SPD cabinet is further to the left than the previous one with a value of -15.1 compared to -9.35. The relative difference between the cabinet stays the roughly same when weighting by the seat share.

[9] One other study by Spinde et al. (2020) examines media bias in Germany but their goal is to find a suitable method for bias detection, and are not very successful.

According to the Reuters Digital News Report 2024, the proportion of German citizens that trusts most "news most of the time" (Newman et al., 2024, p. 83) dropped from 60% in 2015 to 43% in 2024. Simultaneously, the importance of social media as a news source rose from 18% in 2013 to 34% in 2024. Quiring et al. (2024) find a decrease in trust since 2020 even though here the level before 2020 also was as low as currently. Most people, though, at least partly feel like the media doesn't represent reality as they perceive it. Current developments in AI pose the threat of fake news being created more efficiently and frequently shared on social media (Bashardoust et al., 2024; Koplin, 2023). Recent reports suggest that Russia purposefully spreads fake news across Germany to influence the political sphere.[10] These developments highlight the growing importance of having trustworthy and unbiased news.

My focus lies on economic reporting to solve the unobserved population problem and the distinguishment of opinion and bias. In the topics of unemployment, inflation and GDP, there is a general agreement of what is good and bad and they are among the most important indicators for the performance of the economy. Unbiased news should, therefore, solely rely on the changes of these indicators and not be dependent on who is in the government. The specific newspaper I examine are Die Welt and Zeit Online. As there is one more right and one more left leaning government I also choose newspapers with respective ideological leanings. Die Welt is considered a conservative newspaper, while Zeit Online is orientated to the left (Maurer and Reinemann, 2006). The owner and executive of Axel Springer, the publisher of Die Welt, are even known to have personal connections to the CDU (Brinkmann, 2018, p. 576). As I only had resources to examine two newspapers, the choice had to be somewhat arbitrary. Ultimately, I chose these two as they published the most articles on the economy.[11]

My hypothesis is the following:

Under the same economic conditions, a newspaper will report more favorably for the government that is closer to its own ideological leaning.

I will try to identify both selection and presentation bias as I look at the number of articles and the sentiment of them.

4. Data and Method

In this section, I will describe the data sources, the process of text analysis, and lastly, my estimation strategy to answer my hypothesis.

4.1. Data

The relevant time period for my study is the beginning of the previous cabinet, the 14th of March 2018, until June 2024. Data for the economic indicators of unemployment, inflation, and GDP are readily available from the Federal Statistical Office of Germany.[12] As GDP

[10] https://www.tagesschau.de/investigativ/ndr-wdr/russland-propaganda-fakenews-sda-deutschland-100.html

[11] I web scraped articles from the most prominent news sites Bild, Die Welt, FAZ, Spiegel, Süddeutsche Zeitung, taz, and Zeit Online and then filtered them by a keyword search for the relevant topics.

[12] GDP data: https://www.statistischebibliothek.de/mir/receive/DEHeft_mods_00160795, Inflation data: https://www-genesis.destatis.de/genesis/online?sequenz=tabelleErgebnis&selectionname=61111-0002&startjahr=1991#abreadcrumb, Unemployment data: https://www-genesis.destatis.de/genesis//online?operation=table&code=13211-0002&bypass=true&levelindex=1&levelid=1725890763380#abreadcrumb

data is published quarterly, the quarter at the end of June 2024 was the latest available data. Data on the other two indicators are available on a monthly basis. One issue is that there are multiple ways for each of the statistics to calculate changes. First, you can use an absolute number such as the consumer price index, the number of unemployed people, or the total GDP. Next, you can compare the changes in percent to the last month/quarter or the last year. For GDP, there are also raw and price-adjusted numbers. As newspapers generally report about price adjusted numbers, these were also chosen by me. The exact indicators used ultimately are the Consumer Price Index, yearly inflation change, monthly inflation change, unemployment rate, monthly change in unemployment rate, yearly change in unemployment rate, total price adjusted GDP, yearly change in GDP, and quarterly change in GDP. The development over time of the six variables that will be used in the main model can be seen in Figure 1. The most notable changes are the decrease in the COVID-19 pandemic and the increase after, and the increased inflation, especially since the beginning of the war on Ukraine.

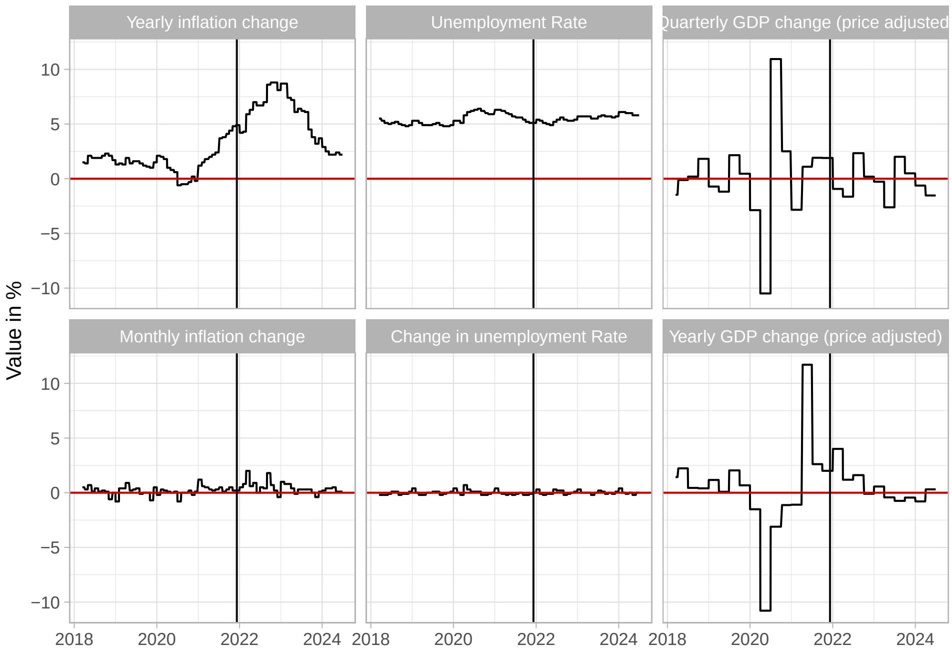

Figure 1: Economic Indicators over time

The data for the newspaper articles was obtained by web scraping the archives.[13] One issue I encountered was the inaccessibility of articles with a paywall that aren't accessible. For Zeit Online, these account for around 6%, and for Die Welt, for around 1.7% of all articles. This could be a problem if these are significantly more or less biased than free articles, in which case the results would only be valid for free articles as the population. However, as these articles don't make up a substantive share of all articles, I don't expect these would have influenced the results. The scraped articles contain information on the date they were published, the title, and the text body. Figure 2 displays the number of monthly articles for each newspaper over time.

Other relevant variables are the newspaper an article belongs to and which government

[13]The script is based on this one: `https://github.com/kssrr/german-media-scrape`
The original ones didn't work anymore, possibly due to being out of date, but were adjusted

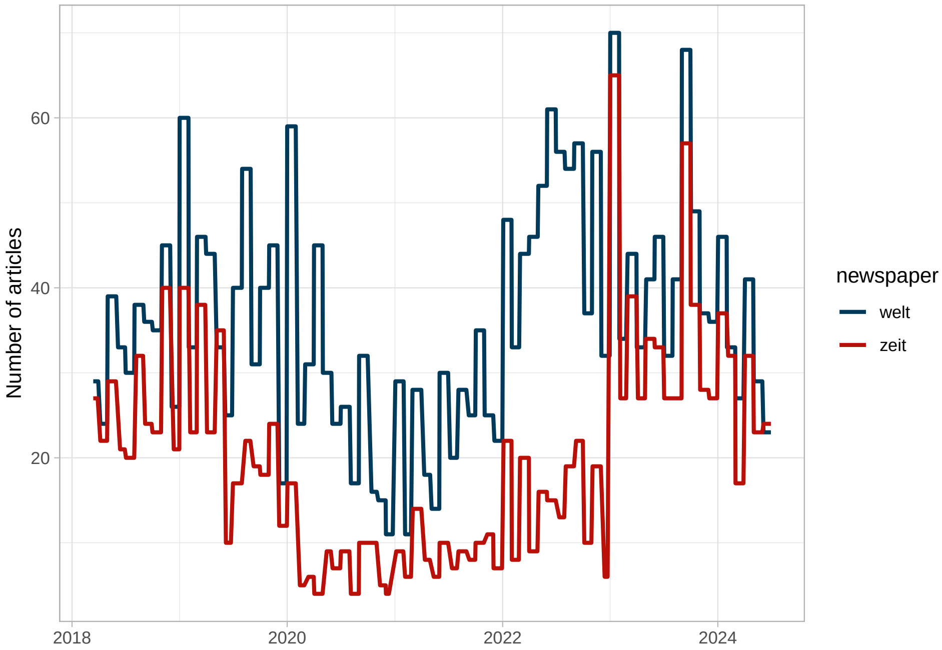

Figure 2: Number of articles over time

is currently in power. These are two time periods, starting with the last Merkel cabinet on the 14th of March and the Scholz cabinet officially starting on the 8th of December 2021.

4.2. Text Analysis

The first task is to identify all relevant articles. As a first step, I filtered the articles using a simple keyword search for "inflation OR arbeitslos* OR bruttoinlandsprodukt OR BIP"[14]. This way, all relevant articles should still be in the sample with many false positives. The next step uses the large language model (LLM) llama3.1 8B to identify all relevant articles. LLMs can be especially useful for text analysis as they can detect things like irony, jokes, and have a big pool of knowledge, albeit sometimes faulty. Unfortunately, I could only use the 8B version, which is the smallest model, as only this one can be used in the programming language R. I also used ChatGPT4 for the validation samples.[15] The analysis using llama3.1 was run on a university server since even the smallest model has high hardware requirements for complex tasks with a large input. While it's a very new technology with uncertainty about its usefulness, there are already promising studies testing AI tools for text annotation. Belal et al. (2023) test ChatGPT for sentiment analysis on Amazon reviews and tweets against prominent lexicon-based algorithms. The AI tool reaches an accuracy of 94% for the reviews and 67% for the tweets, outperforming the algorithms by 20-25 percentage points. Reiss (2023) advises caution when using LLMs concerning reliability. Using the temperature setting, one can control the randomness of the output. LLMs are supposed to be mostly deterministic with a temperature of 0 (the maximum is 1) and by setting a seed. I used such a setting for my analysis and achieved

[14]For the first three words I transformed all letters to lower case.

[15]Using it with more data was not possible due to monetary constraints and with ChatGPT one should be cautious for dataprotection reasons as one sends the data via the API. Llama3.1 on the other hand is open source and can be used locally.

100% reliable results with the same prompt. Minor changes in the prompt[16], though, can result in relatively strong variation in the output. Reiss (2023) highlights the importance of validation and proposes to pool the results of different prompts or temperatures to achieve better results. For reproducibility reasons, I only used a 0 temperature setting and a seed.[17] Pooling results did not improve the performance. Krugmann and Hartmann (2024) also compare GPT and llama models against transfer learning models and find them to compare similarly or sometimes even better, reaching high levels of accuracy. Gilardi et al. (2023) even find that ChatGPT outperforms crowd workers at text annotation of tweets and news articles. Even at a temperature of 0.2, ChatGPT's intercoder agreement is close to 100%. Others (Ollion et al., 2023; Pangakis et al., 2023) are more cautious as they find strongly varying performances depending on the task. They highlight the necessity for validation using coded data by humans. This, though, should be a standard for all quantitative text analysis methods. Another issue, also mentioned by Ollion et al. (2023), is the English-centeredness of LLMs, as their performance is worse for other languages. Here, one solution could be to translate texts before the analysis or try to use English prompts. As German is one of the supported languages of llama3.1, the analysis still worked.[18]
To validate the results of the first text analysis to filter out relevant articles I coded 600 randomly selected articles by hand. When developing the prompt, I followed the best practices described by Törnberg (2024).
First, I chose criteria for relevant articles:

- The article reports changes for either inflation, unemployment rate, or GDP. Articles where they are simply mentioned do not count.[19]
- Only articles about statistics on the federal level count and not if it's only about single states or the EU. If both things are mentioned, they are relevant.[20]

Using the hand-coded dataset, I validated the results and searched for patterns of wrongly classified articles. In this process, I tried different wordings for the prompt to improve the results. The LLM also provided a justification for its classification, making it easier to find why there are disagreements. Here is the translated final version of the prompt:

"You are an assistant in a scientific project and help to recognize relevant newspaper articles.
I am looking for articles that explicitly report on changes in inflation, unemployment figures, or economic growth/gross domestic product (GDP) in Germany.

[16] What is a prompt? "AI prompting refers to the process of interacting with an artificial intelligence (AI) system by providing specific instructions or queries to achieve a desired outcome." `https://servicecenter.fsu.edu/s/article/What-is-AI-prompting`

[17] The reproducibility, though, is dependent on the version of the LLM. Currently new versions are updated regularly. For Reproducibility one has to use the same version. I used llama3.1 8B using ollama software version 0.3.3.

[18] The results might have still been better using English texts. Translating them using the Deepl or Goolge API was not possible due to monetary constraints. Using English prompts in my case produced worse results.

[19] The word GDP is often included when talking about debt/spending as a share of GDP and inflation is often mentioned as a side note saying that it's currently high

[20] There are also a lot of articles about the development in the single states but as these have different governments, they might not be suited to identify bias against the federal government.

Only articles that deal with Germany as a whole are relevant. Articles that deal with Europe or the EU/ECB or only individual regions or federal states of Germany should be excluded. Only articles that deal with the change in these statistics are relevant. If they are mentioned in a different context, the articles are irrelevant. It is always only about the nationwide context of Germany. Individual industries or companies are also not relevant.
If the nationwide statistics and individual federal states or the EU/ECB are mentioned at the same time, the article is relevant. If only one federal state or the EU/ECB is reported on, then the article is not relevant.

Enter your answer in JSON format as follows: {'german_economy': 'Yes/No/Uncertain'}

Options:

- Yes
- No
- Uncertain

Ensure accuracy and clarity in your analysis and base your assessment on the context provided and your expertise. If you are unsure about the categorization, select 'Uncertain'. Pay close attention to the above exclusion criteria. Answer 'No' if the criteria do not apply exactly. Article text:

[Here the article text is inserted]

This task is a matter of life and death, so make sure you do everything right!"[21]

Table 2: Confusion Matrix for text classification

	Reference	
Prediction	Ja	Nein
Ja	47	10
Nein	11	532

Table 3: Text classification performance metrics

Statistic	Value
Accuracy	0.96
Precision	0.82
Recall	0.81
F1	0.82

Table 2 contains the confusion matrix with the columns standing for the hand-coded data and the rows for the results of the LLM. The top left cell contains the true positives,

[21]Yes, this last sentence actually increased the accuracy even further. The first version of the prompt achieved an F1 score of 0.71 and was increased by adjusting it.

the top right cell false positives, the bottom left cell false negatives, and the bottom right true negatives. Table 3 displays relevant statistics for the performance. The accuracy is the share of correctly predicted items, which is very high, with 96%, due to the high amount of irrelevant items. Recall indicates how many of the as relevant classified items are actually relevant, therefore accounting for false positives. Precision, on the other hand, considers false negatives and measures how many of the relevant items were classified as relevant. The F1 score then combines the measures of precision and recall. The values of 0.81 and 0.82 are lower than the accuracy but still relatively high overall. Using Krippendorff's Alpha (Krippendorff, 2011) as an indicator to measure agreement between hand-coded and LLM-coded data gives a result of 0.8. Values that are 0.8 or bigger are generally considered as high and an acceptable level of agreement (Marzi et al., 2024). Results for ChatGPT are even higher and reported in Tables A1 and A2 in the appendix.

The second part of the text analysis aims to determine the sentiment of the article. This will serve as the dependent variable of my analysis. "Sentiment analysis measures the polarity or tonality of texts by identifying and assessing expressions people use to evaluate or appraise persons, entities or events" (Haselmayer and Jenny, 2017, p. 2625). In my case, it is supposed to measure if the change in the economic indicator is positive or negative and how this change is framed. For this analysis, I randomly selected a sample of 200 relevant articles and hand-coded them. In this process, I first split each article into natural sentences and coded each sentence into either "1" for positive, "0" for neutral or mixed, "-1" for negative, and "I" for irrelevant if the sentence dealt with another topic. Unfortunately, it is not as objective as one might think. For example, even if the unemployment rate decreased compared to last month, it could have increased compared to the previous year, as monthly changes are often normal due to seasonal effects. Also, if there was a longer period of bad inflation, a following decrease can sound positive even if it is still higher than an increase after a long period of very low inflation, which can sound negative. Often, an unchanging GDP is considered negative rather than neutral, and for inflation, there's no clear dividing line at what percentage it turns negative or positive as the goal of the central bank is 2%. While this might make results less reliable due to subjectivity, it also shows that there are a lot of ways for newspapers to frame these numbers in a different way. Table 4 offers a few examples of hand-coded classifications. A sentiment score was calculated for each article, aggregating each sentence's rating. Here, I adopt the proceeding of Budge (2013), who creates a right-left scale for party manifestos. I chose the ratio scale, which also includes, in my case, the "neutral" category instead of just using negative and positive sentences, as neutral sentences also substantively contribute to the sentiment of the articles. One problem with this scale could be a tendency toward the center in the case of many sentences in the neutral category. This is not the case here. The scale then ranges from -1 to 1. By setting cutoff points at -0.2 and 0.2, which achieved the highest agreement with the LLM results, I created negative, neutral, and positive categories. The procedure for creating the prompt was the same as for the first text analysis. The final translated prompt is the following:

"You are an expert in analyzing news articles. You are given an article that deals with various economic topics. Your task is to analyze only the sections that refer to developments in unemployment, inflation, or economic growth/gross domestic product (GDP) in Germany. Ignore all other topics and also reports on specific regions in Germany, the EU, or the global economy.

First, categorize whether the article has any relevant information at all, as described in the criteria above. Then rate whether the article is positive, neutral, negative or uncertain about developments in Germany. Justify your assessment and give a rating on a scale from 0 to 10, where 0 is completely negative and 10 is completely positive.

Format your answer in JSON format as follows:

{‘relevant’: ‘Yes/No’, ‘sentiment’: ‘Negative/Neutral/Positive/Uncertain’, ‘sentiment_score’: ‘0-10’, ‘justification’: ‘Justify your answer’}

Article text:
[Here the article text is inserted]”

Table 4: Examples for handcoded sentences

Sentence	Sentiment	Justification
Germany’s economy shrank by 0.3 per cent on a price-adjusted basis.	Negative	A decrease in GDP is considered negative
Supply chain problems are decreasing.	Neutral	It’s positive that problems are decreasing but there are still problems
Federal Labour Minister Hubertus Heil (SPD) was pleased, but the Left Party, on the other hand, considers jubilation ‘inappropriate’, as the quality of the work is often not right.	Neutral	The sentence includes two different judgements by different parties, one positive and one negative, resulting in a neutral judgement
Reinhard Houben, economic policy spokesman for the FDP parliamentary group in the Bundestag, described the minimal increase as a “real disappointment”	Negative	Even if there is an increase (in GDP) it is judged as very negative
Record low unemployment, record high number of jobs: according to forecasts, the German labor market will once again be in top form in 2018 at the height of the economic boom.	Positive	Very clearly positive developments on the labor market

The validation results of this analysis can be found in Tables 5 and 6.[22] The overall accuracy lies at 77% which this time is lower but still relatively high. Results for the negative category are the best, with an F1-score of 0.85, while the neutral and positive categories performed worse, with 0.71 and 0.67, respectively. Krippendorff’s alpha is at 0.72, indicating moderate agreement between the hand-coded data and the LLM. This value is still at the “lower bound for tentative conclusions” (Marzi et al. 2024), but the

[22]Results for ChatGPT are in the appendix in Tables A3 and A4

results should be judged with caution. The correlation between the two numeric scales is relatively high (r = 0.67, p <0.001), speaking for acceptably valid results.

Table 5: Confusion Matrix for sentiment analysis

	Reference		
Prediction	Negativ	Neutral	Positiv
Negativ	78	9	3
Neutral	15	46	8
Positiv	1	5	17

Table 6: Sentiment analysis performance metrics

	Precision	Recall	F1
Class: Negativ	0.87	0.83	0.85
Class: Neutral	0.67	0.77	0.71
Class: Positiv	0.74	0.61	0.67

Based on the date variable, the data for the articles is then merged with the economic indicators.

4.3. Estimation Strategy

In order to answer my hypothesis, I will use an OLS regression with the numeric sentiment score as the dependent variable. The unit of analysis is single articles, as this preserves the most information rather than creating an aggregated score or an index for time periods. My main independent variable is the government in power since the goal is to determine if a newspaper reports differently under each one. As my hypothesis states, "under the same economic conditions", I will use control variables for the economy. The six control variables were the most commonly cited statistics in the sample of articles I used to validate the text analysis. My sample contains articles from two different newspapers where one leans to the left and one to the right. By using an interaction term between government and newspaper, I get a distinct coefficient of the government variable for each newspaper. However, the interaction term itself can also indicate a bias if the newspaper changes its reporting more strongly than the other. A non-significant interaction term, though, does not rule out a bias if the reporting changes to the same degree. For further robustness, I will also use logistic regression models on a binary variable indicating if an article is negative and one for positives. This can also give further insight into whether a difference in reporting is due to more negative or positive articles or even both.
I decided against the usage of time fixed effects. First of all, they don't make sense in my case. Time fixed effects are commonly used to control for omitted variables that vary over time. In my case, the government variable is already a time fixed effect as it separates the data into two time periods. Another time fixed effect would further split the data, but only in one of those splits would be variation of the government variable as it only changes at one point in time. Kropko and Kubinec (2020) give an overview of the interpretation and inner workings of fixed effects. The second reason is that time fixed effects would contain time-variant factors like demand for negative/positive news. As this can be a mechanism,

this would reduce my coefficient, and since I can't isolate the effect of demand, I wouldn't know what proportion of the coefficient it affects.
Diagnostics for the linear regression can be found in Figures A1 and A2 in the appendix. Due to heteroscedasticity, all models use robust standard errors. My dependent variable, the sentiment score, does resemble a normal distribution, but due to multiple peaks, it can't be considered normally distributed when tested. This leads to the residuals not being normally distributed, which can result in biased standard errors. The VIF scores do not detect problematic values regarding multicollinearity.

5. Results

5.1. Descriptives

Before delving into the multivariate analysis, I present descriptive statistics. Figure 3 describes the average monthly sentiment score for each newspaper over time. The drop and increase in GDP during the COVID-19 pandemic is clearly visible in the sentiment of the articles. Similarly, the increasing inflation around 2022 leads to lower sentiment scores. Table 7 shows the mean sentiment scores for each newspaper under both governments. For both newspapers, there is a clear decrease in sentiment during the Scholz cabinet. While Die Welt overall reports more negatively, the decrease for Zeit Online during Scholz is stronger. The same comparison but in categories can be seen in Figure 4. Die Welt overall has fewer positive and more negative articles than Zeit Online. Both newspapers have more negative and fewer positive articles than during Merkel. The respective increase and decrease is stronger for Zeit Online. Now, the question is whether these changes can be explained purely by differences in the economic situation.

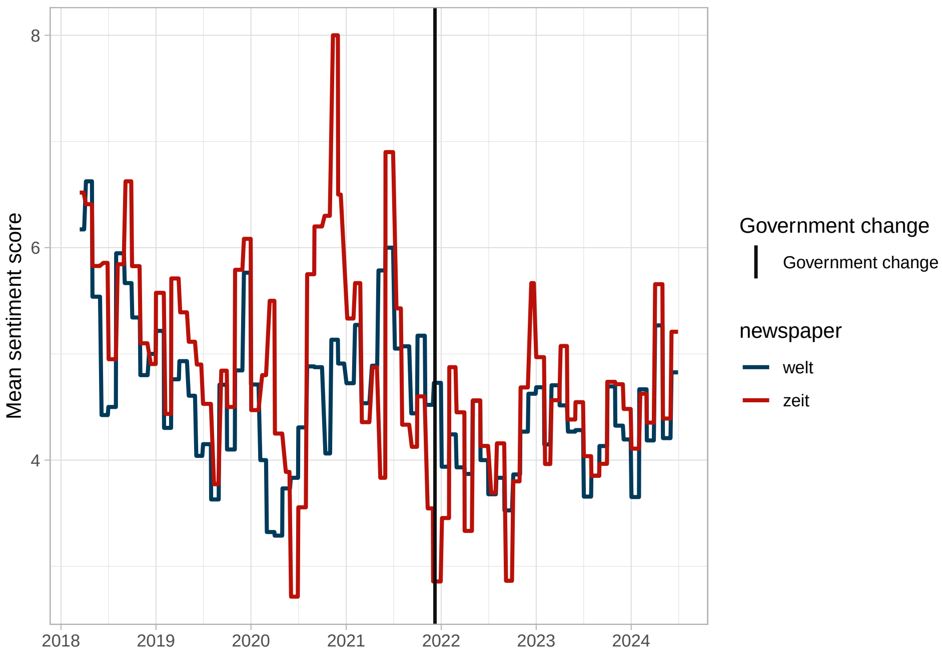

Figure 3: News sentiment over time

Table 7: Mean sentiment by newspaper and government

Newspaper	Government	Mean Sentiment Score
welt	Merkel	4.75
welt	Scholz	4.21
zeit	Merkel	5.29
zeit	Scholz	4.41

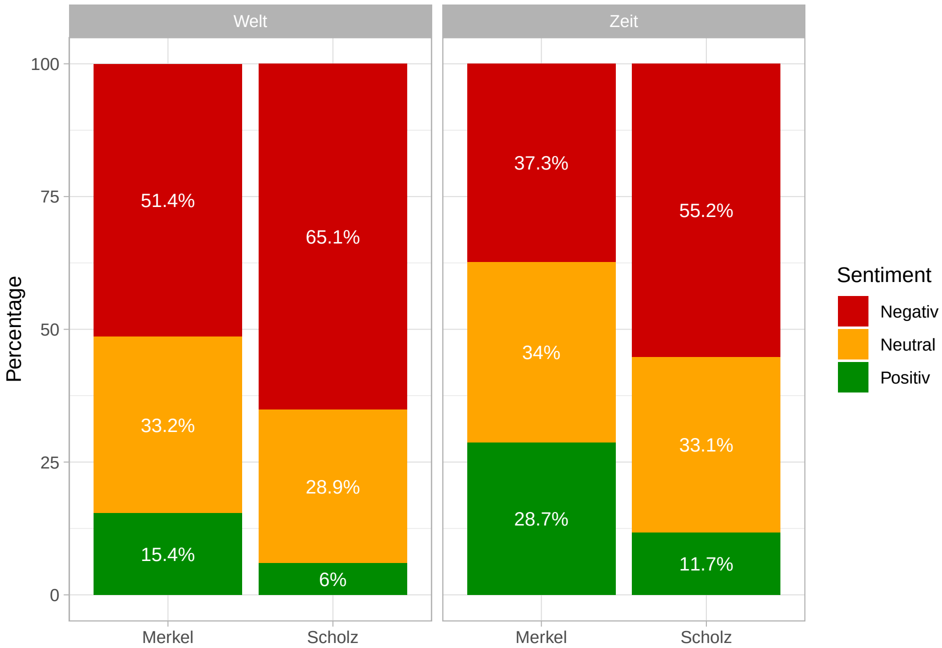

Figure 4: Share of articles by newspaper and government

5.2. Multivariate results

Next, I will discuss the results of the regression models. The two models in Table 8 use the numeric sentiment score as the dependent variable. The first column shows the model with only the interaction of the government and newspaper variables, while the second column contains the model with control variables. With the interaction term in the model, the coefficient of the government variable only accounts for the reference category of the newspaper variable, Die Welt. It, therefore, tells us the difference between the sentiment of articles from Die Welt during the Scholz cabinet compared to Merkel. The interaction term itself signifies if the difference between the cabinets is higher or lower for Zeit Online than Die Welt. By adding the interaction term coefficient to the government coefficient, one has the difference between the cabinets for Zeit Online. As the significance of that difference is not visible in the model output, I calculated the marginal effects for each newspaper separately. These can be found in Figure 5.[23] Both models are largely the same. The reporting of both newspapers got significantly more negative during the Scholz cabinet and even more so for Zeit Online. In the model with the control variables, the coefficients become a bit smaller, indicating that the economic variables explain part of the

[23]The figures are always based on the models with control variables.

Table 8: Regression results for OLS models

Dependent Variable:	Numeric sentiment score (0-10)	
Model:	(1)	(2)
Variables		
Constant	4.746***	3.685***
	(0.0457)	(0.6382)
Newspaper: Zeit (ref. Welt)	0.5459***	0.5174***
	(0.0876)	(0.0884)
Government: Scholz (ref. Merkel)	-0.5410***	-0.4559***
	(0.0602)	(0.0984)
Interaction: Zeit x Scholz	-0.3456**	-0.3206**
	(0.1147)	(0.1176)
Yearly inflation		-0.0295
		(0.0180)
Unemployment rate		0.2071.
		(0.1170)
Quarterly GDP change (price adjusted)		-0.0106
		(0.0105)
Monthly Inflation Change		-0.1036
		(0.0668)
Yearly GDP change (price adjusted)		0.0858***
		(0.0153)
Yearly Change of unemployment rate		-0.0289
		(0.0917)
Fit statistics		
Standard-Errors	Heteroskedasticity-robust	
Observations	4,232	4,232
R^2	0.04820	0.06450
Adjusted R^2	0.04752	0.06251

Heteroskedasticity-robust standard-errors in parentheses
*Signif. Codes: ***: 0.001, **: 0.01, *: 0.05, .: 0.1*

negative articles. Under Scholz, Die Welt has a sentiment score that is 0.46 lower, while Zeit Online has a score that is 0.78 lower. Both of these differences are very significant and the difference between both as well. Figure 6 shows the estimated marginal means for the second model. During the Merkel cabinet, Die Welt has an estimated sentiment score of 4.71 and 4.25 during Scholz. Zeit Online lies at 5.23 during Merkel and at 4.45 during Scholz.

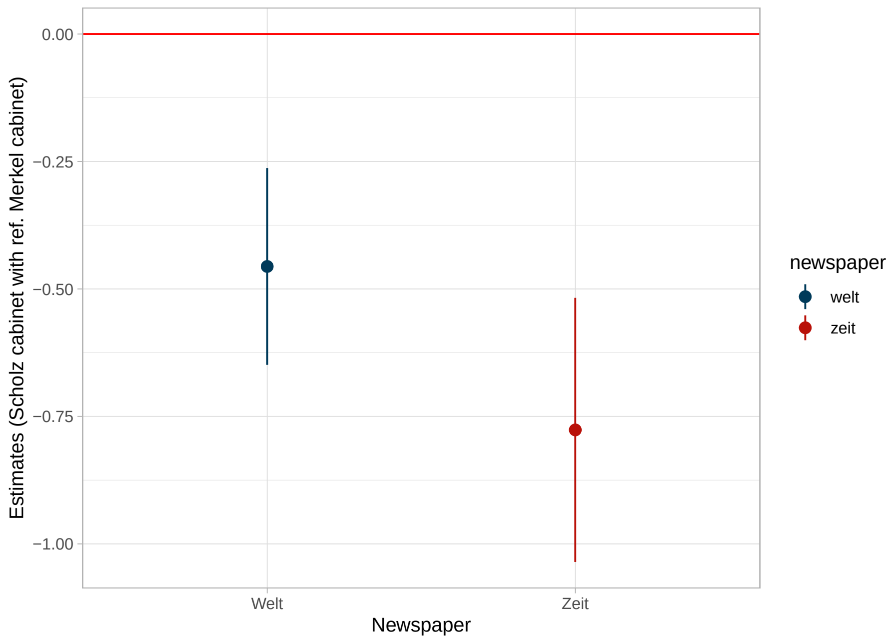

Figure 5: Marginal effects of Scholz government (OLS model)

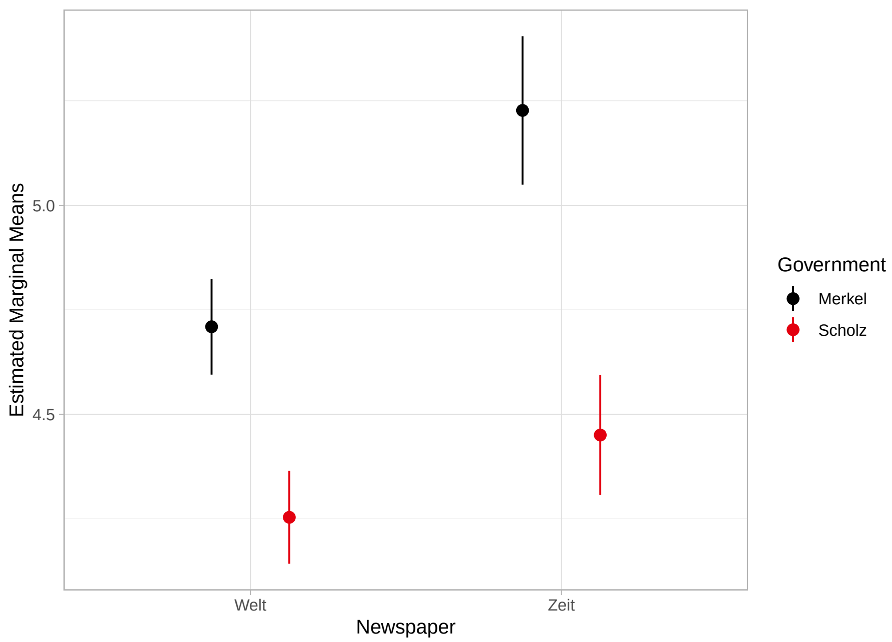

Figure 6: Estimated marginal means by newspaper and government

To determine if these changes are mainly driven by an increase in negative or a decrease in positive articles I also estimate logistic regressions using binary variables indicating if an article is negative/positive or not. The tables for all models can be found in Tables A5 and A6 of the appendix. Figure 7 shows the average marginal effects difference between governments for both newspapers. An article by Die Welt during Scholz has an 11.5% higher probability of being negative, while there is a 16.3% higher probability for Zeit Online. The results for positive articles in Figure 8 show a 10.2% lower probability for positive articles by Die Welt and a 17.9% lower probability by Zeit Online. The probability for an article by Die Welt under Scholz to be negative is 64%, while it is 55% for Zeit Online. In these models, the differences in the changes of the newspapers are not significant. Increased negative reporting under Scholz is therefore both driven by more negative and fewer positive articles.

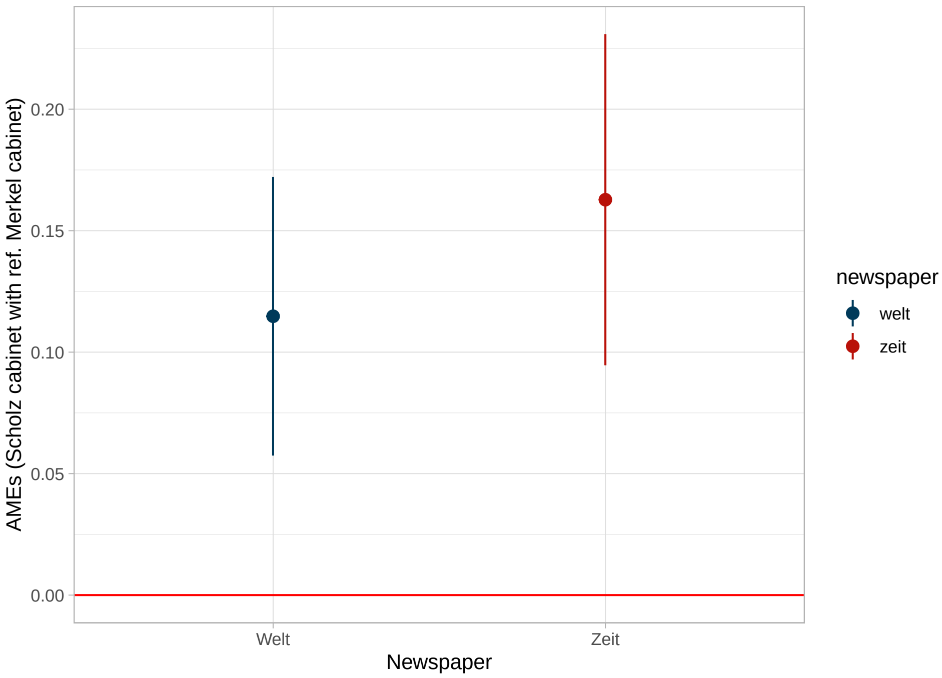

Figure 7: Average marginal effect of Scholz government (negative articles)

These results indicate that the reporting of both newspapers is politically biased since the reporting differs significantly under distinct governments but in the same economic conditions, as achieved by controlling for it. One can't infer from the data if the bias is negative towards the Scholz cabinet or positive towards the Merkel one. It would be a normative claim to choose a reference point that says that, under specific economic conditions, the sentiment of articles should have a specific level. My hypothesis that argues that the bias goes in the direction according to the ideological leaning of the newspaper and government is not supported. Zeit Online is a left-leaning newspaper and, according to the theory, should have favored the Scholz government. The data for Die Welt would make sense in this regard, as the conservative newspaper reports more positively during the cabinet led by the CDU. But both newspapers writing more negative articles during the Scholz cabinet speaks for other reasons explaining this bias.

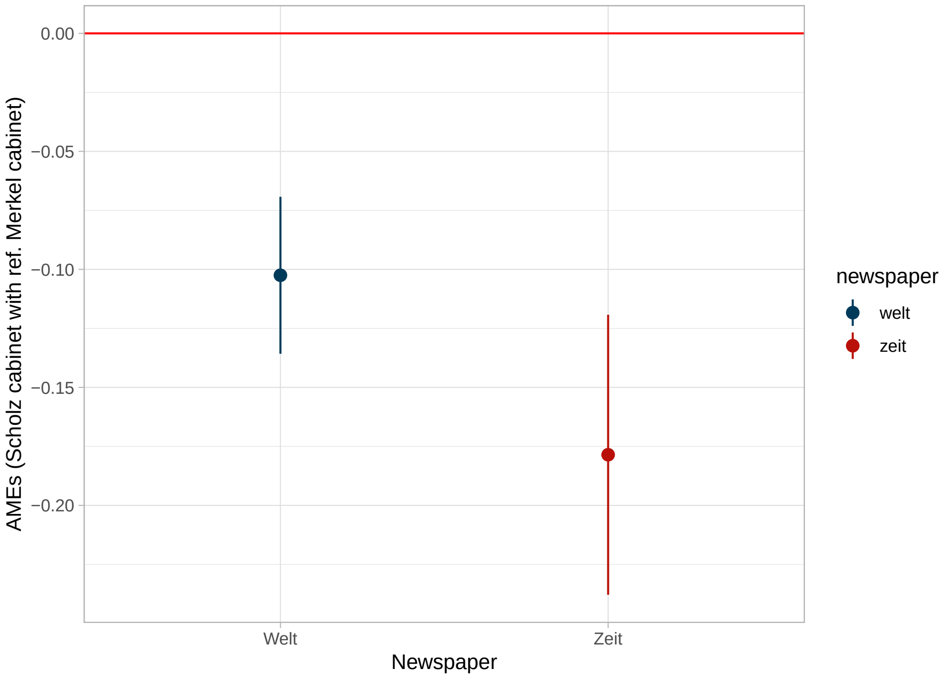

Figure 8: Average marginal effect of Scholz government (positive articles)

5.3. Sensitivity and robustness checks

As there are many different indicators for the economic variables, it could be the case that including or removing certain variables alters the results of my analysis. To make sure the results are not dependent on a specific combination of these variables, I ran additional models. 9 economic variables, 3 for each topic, result in 511 potential combinations of these where 1 to 9 variables are included. I run all these models and report the distribution of estimates and p-values. A narrow distribution of both would indicate that the results are insensitive to small changes in the model. For the estimates, it is important that they do not change the sign, and for p-values, all values should be lower than 0.05. Figure 9 shows the distribution of estimates for both newspapers. The estimates vary to some degree and have a range of 1 point on the 0 to 10 scale, but the peak is near the estimate of the main model, and all values are negative and distinct from 0. The significance level is even clearer, as can be seen in Figure 10. Each model is highly significant.
The same procedure was used for the logistic models as well. All Figures can be found in the appendix (Figure A3 - A6). For the models on negative articles, the Welt estimate does vary relatively strongly from around 6% to 15%, while the Zeit estimate varies regarding the models on positive articles from around -2.5% to -22.5%. The former models are still all very significant. The latter models show most p-values at a very significant level, but a substantial amount crosses the 0.05 line, showing that they are not quite robust.
Next, I also conducted robustness checks regarding the time period used. All Figures for the following analysis can also be found in the appendix (Figure A7 - A33). Firstly, I cut out the first and last year and ran my main model as well as the robustness check from above. The second check uses one year before and after the 2021 election. The last check cuts out the period around the election by 5 months. All of the results are quite similar.

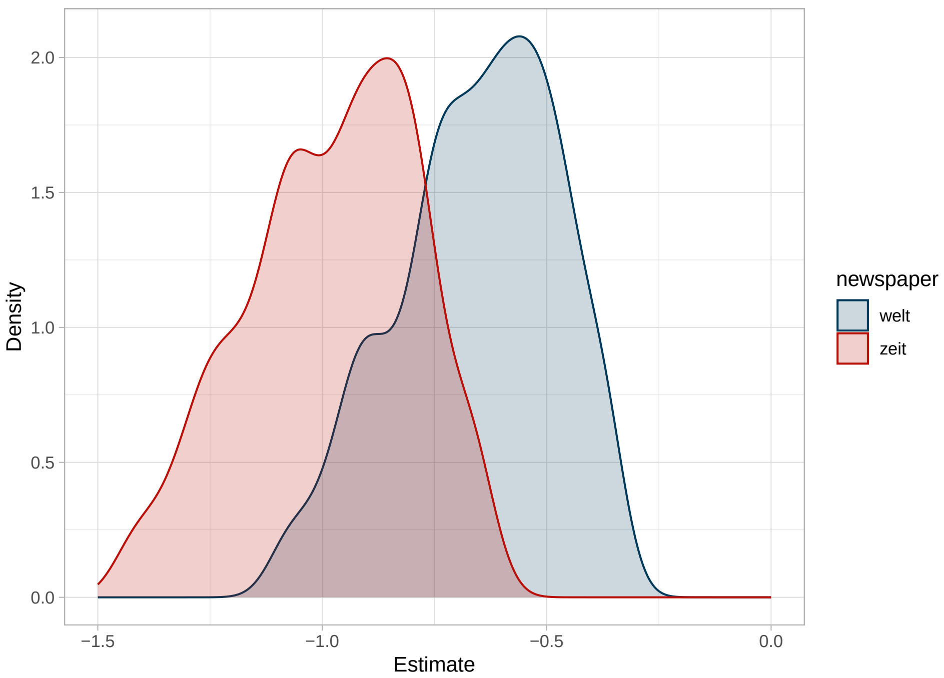

Figure 9: Distribution of estimates with differing control variables (OLS models)

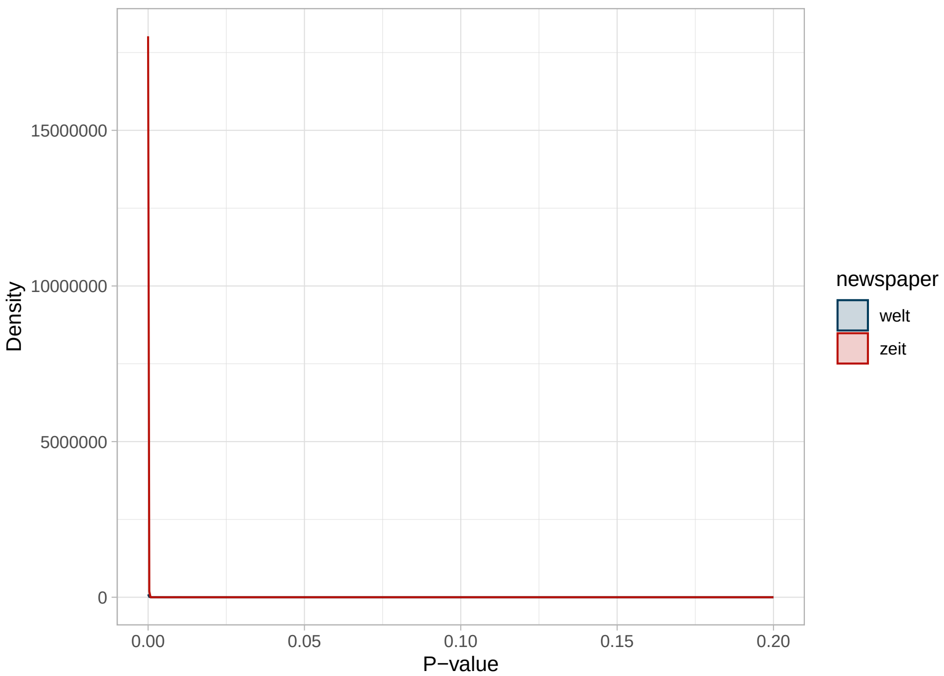

Figure 10: Distribution of p-values with differing control variables (OLS models)

There is some variation in the estimates, but the linear models and the ones on positive articles all stay significant. In the second timeframe the Zeit coefficient for negative articles is not quite significant but the sensitivity analysis shows that almost all other model variations are significant. Only the models on positive articles show a slightly bigger proportion of estimates that are not significant. One reason for this could be that positive articles make up the smallest share of the sample and that the text classification for positive articles performed worse, resulting in more noise. Overall, though, the models prove to be robust.[24]

5.4. Further Investigation on the Nature of the Bias

One potential reason for the bias could be the demand for more negative news during the Scholz cabinet since its popularity declined quickly. As mentioned in the theory section, identifying the effect of demand is difficult due to endogeneity. It is also very difficult to measure because one would need information about the readership of each newspaper and their opinions. Despite these issues, I show the development of government approval and average article sentiment over time in Figure 11. There is no significant correlation between sentiment and the approval rating, even when lagged for 1 or 2 months.

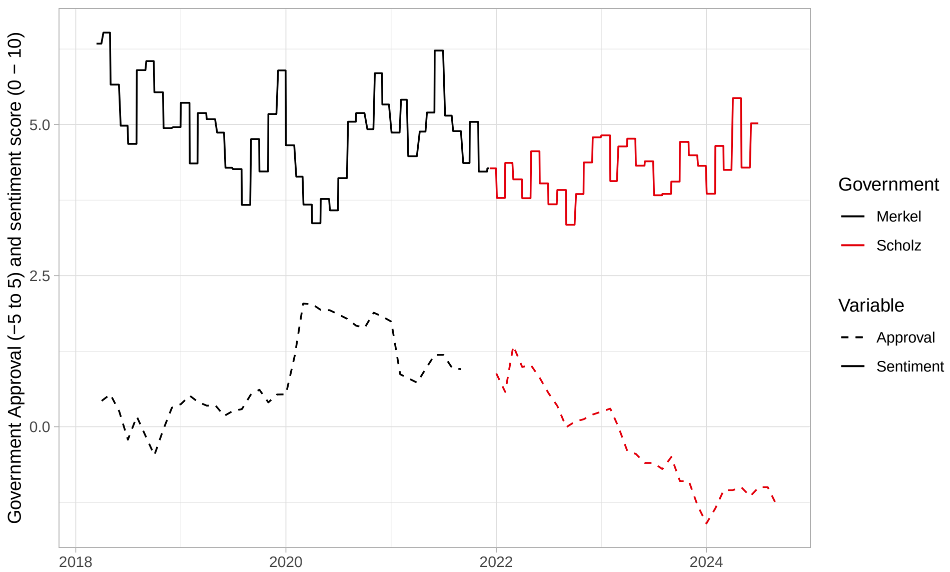

Figure 11: Article sentiment and government approval over time

As mentioned in the method section, adding a normal time fixed effect does not make sense. However, when using a linear fixed effect for the day count in the OLS model, the coefficients of the government variable for both newspapers become insignificant (Figure 12). This hints at a time varying factor that causes the discrepancy of the sentiment in reporting between both governments.

In the Theory section, I discussed the distinction between selection and presentation bias. My main model can account for both types of biases since it takes the sentiment and number of articles into account. To further investigate both types separately, I use a model with the probability of having a negative article and also models with the number of articles per day.[25] In these models, I look at the influence of economic variables on either of these. By interacting them with the government and newspaper variable I can see, for example, if under high inflation under Merkel, there's a higher probability

[24]It has to be noted, though, that no bonferri correction was applied. This can be done when doing a large amount of tests because then some will be significant by pure chance.

[25]All models also use control variables for the respective other economic indicators

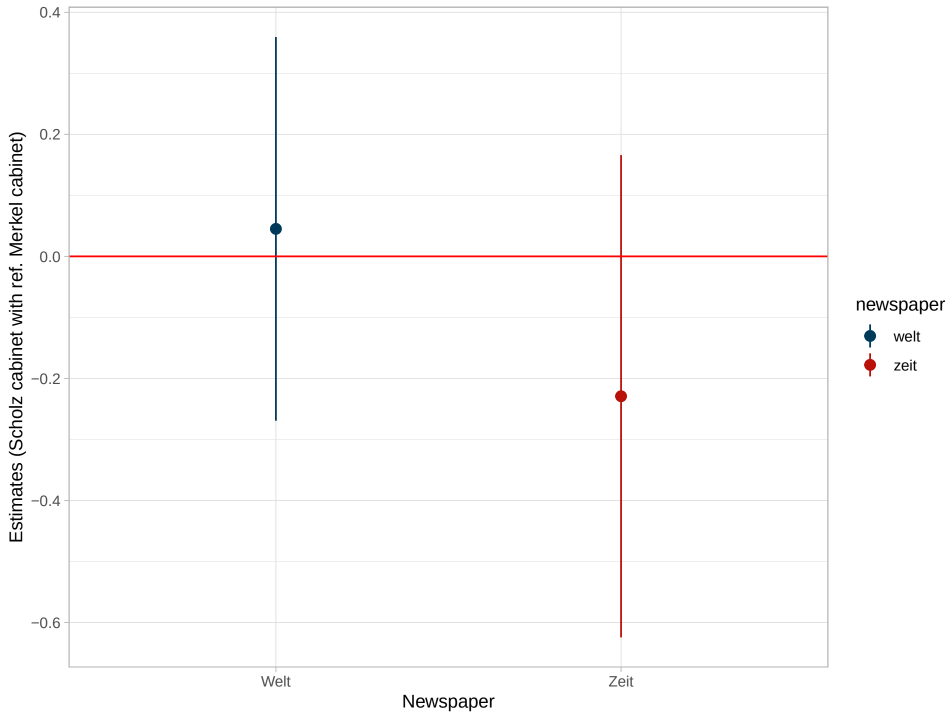

Figure 12: Marginal effects of Scholz government (OLS model with time fixed effect)

of writing a negative article for both newspapers or the high inflation results in a higher number of articles. Figure 13 shows the results for the probability of a negative article using inflation, indicating presentation bias. For Die Welt, an increase in inflation leads to a slightly higher, though insignificant, increase in the probability of publishing a negative article. This slope is almost the same under both governments. Zeit Online shows a different picture. Here, the increase in probability during Merkel is significantly stronger than during Scholz, where there's only a small increase. This indicates a presentation bias of Zeit Online against the Merkel cabinet or a bias in favor of the Scholz cabinet.

In Figure 14 one can see the results for the influence of yearly GDP growth. Here, the results are quite the opposite. The slope for Zeit Online is almost exactly the same under Merkel and Scholz. With growing GDP the probability of a negative article decreases. For Die Welt, this is only the case under Merkel. During the Scholz cabinet, the probability of a negative article increases with higher GDP growth. This speaks for a bias by Die Welt against the Scholz cabinet.
Next, I look at the number of articles published. Figure 15 shows that under Merkel, higher inflation leads to more articles being published. For Welt this increase is slightly higher than under Scholz. For Zeit Online, during Scholz, the number of articles published stays the same irregardless of the change in inflation. None of the differences between the slopes are significant. The results for the yearly change in GDP are shown in Figure 16. During Merkel, there's a slight, though insignificant, increase in articles when GDP is growing. The slopes during Scholz, on the other hand, show that fewer articles are published with growing GDP. Here, the difference between the slopes of Zeit Online is significant. This indicates that there is a selection bias against the Scholz government.

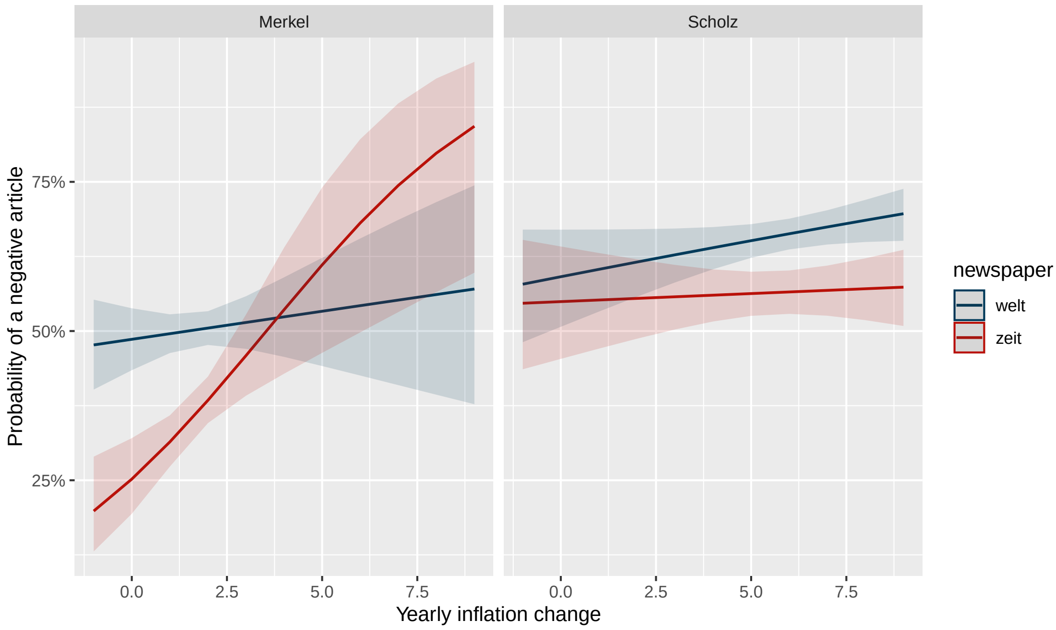

Figure 13: Probability of a negative article by inflation

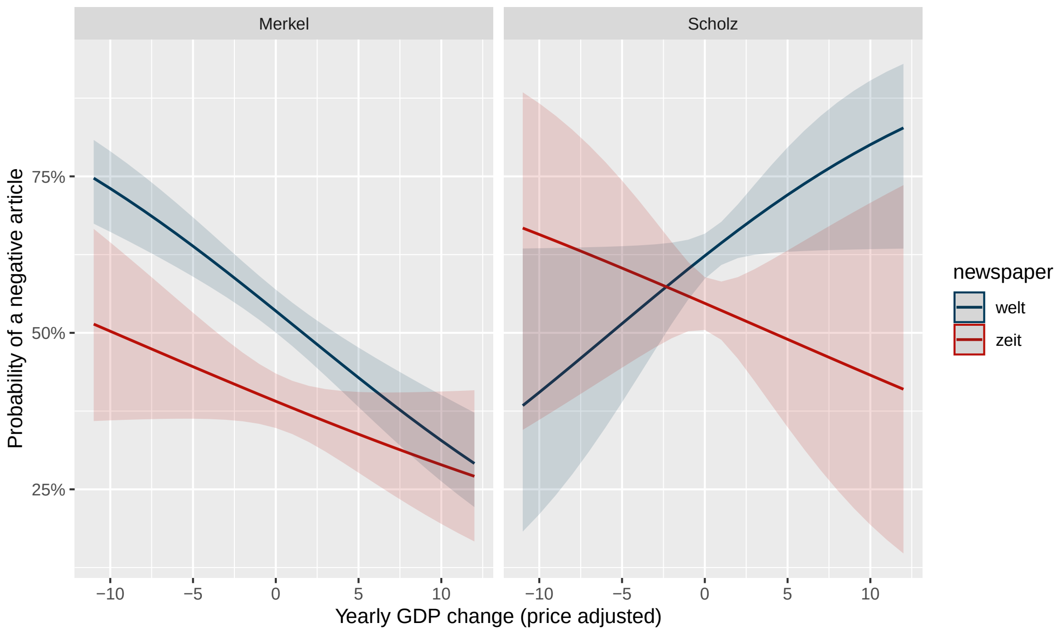

Figure 14: Probability of a negative article by GDP growth

These results offer only little support for a general negativity bias as a reason for the discrepancy between the sentiment during both governments. A worse economy only leads to slightly more articles being published. As during the Scholz cabinet, the economy performed worse this could have been one explanation.

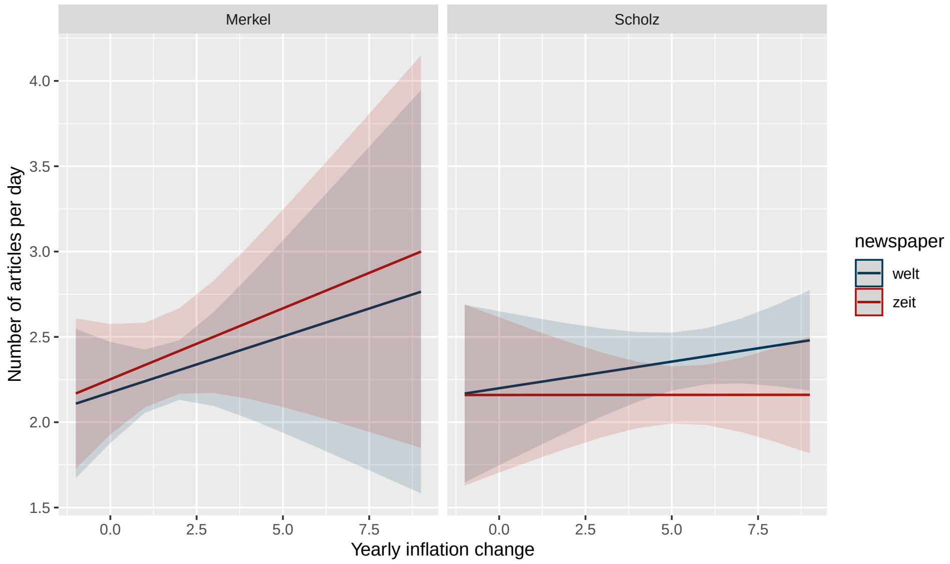

Figure 15: Number of articles a day by inflation

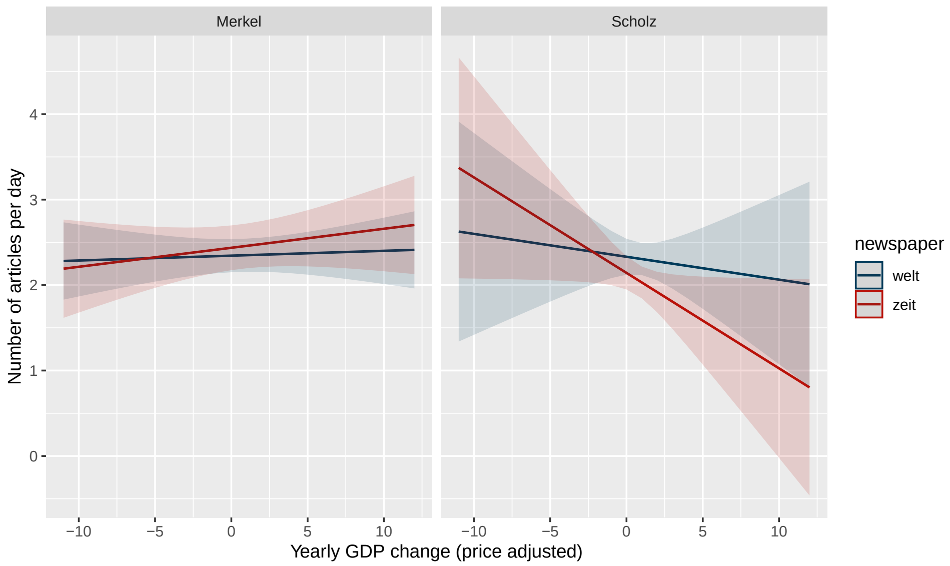

Figure 16: Number of articles a day by GDP growth

Another aspect that I discussed in the theory section is the interdependence between journalists and politicians. It can make a big difference if the media let politicians speak in their own words (Semetko et al., 2011). To shed some light on this relationship, I also examine the quoting patterns of both newspapers. With another LLM analysis, I identified whether a party is quoted in an article. For validation, I used the random sample of 200 articles for hand coding. The F1-scores for all variables range from 0.81 to 1 (Table A7), indicating sufficient validity. Figures 17 and 18 show the difference in the probability of quoting the government and opposition between the cabinets for both newspapers. Die Welt is less likely to quote parties of the Scholz cabinet compared to Merkel, while Zeit Online shows the opposite. Die Welt during Merkel has a 22% probability of quoting the

government and 20% during Scholz. Zeit Online has a 12% probability during Merkel and 14% during Scholz. This aligns with their ideological leanings but both differences are not quite significant.
Regarding quotes of the opposition, only Zeit Online shows a difference between governments with a 1.5 percentage points higher probability, though this is also insignificant. Die Welt has an 8% probability of quoting the opposition, while Zeit Online has a 2% probability during Merkel and 4% during Scholz. Results for the single parties can be found in the appendix (Figure A34 - A39) and also indicate if parties are more likely to be quoted in negative or positive articles. I also included graphs (Figure A40 and A41) for the most commonly cited economic think tanks, the economically liberal ifo institute, and the more left-leaning DIW. Interestingly, the DIW is cited less during the Scholz cabinet, while there are no differences for the ifo institute. All these results should be regarded with caution as they're of descriptive nature. Ultimately, the exact reasons for bias can't be detected, but the previous analyses provide some insight into what the bias looks like.

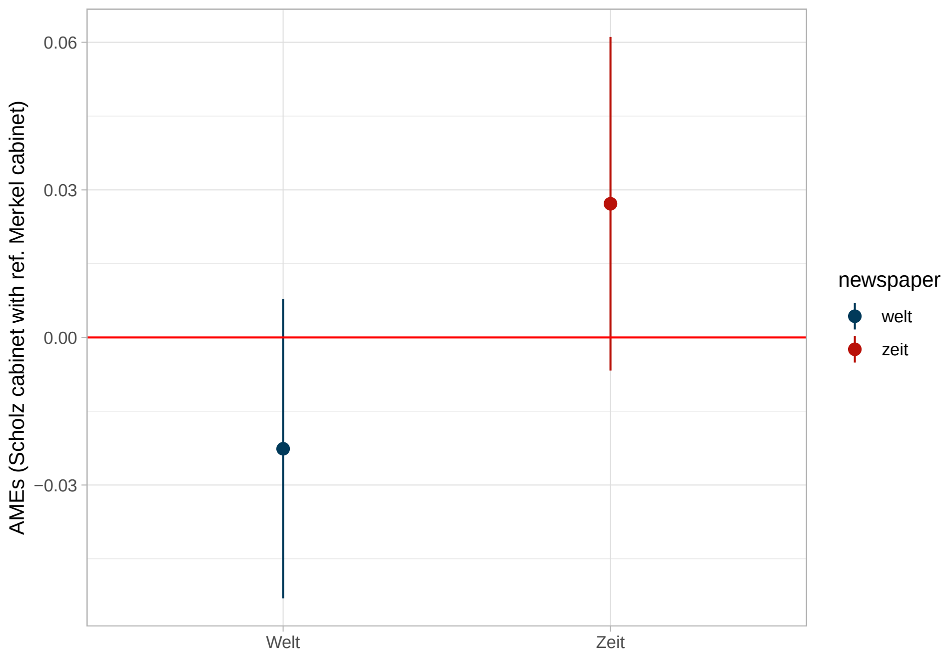

Figure 17: Marginal effect of Scholz government on quoting the government

6. Limitations and Conclusion

6.1. Limitations

This study comes with a few limitations. Using only an observational setting I can't make any causal claims on the influence of a government on news reporting. In my case, I defined bias in such a way that it exists when the reporting doesn't correspond to "reality", which in my case were the actual economic conditions. Here, the problem could be that the news depends on the context. In a case where the economy has been doing good and then worsens news will sound negative. But if the economy has been doing poorly and then improves, the news will sound positive. In absolute terms the second economy

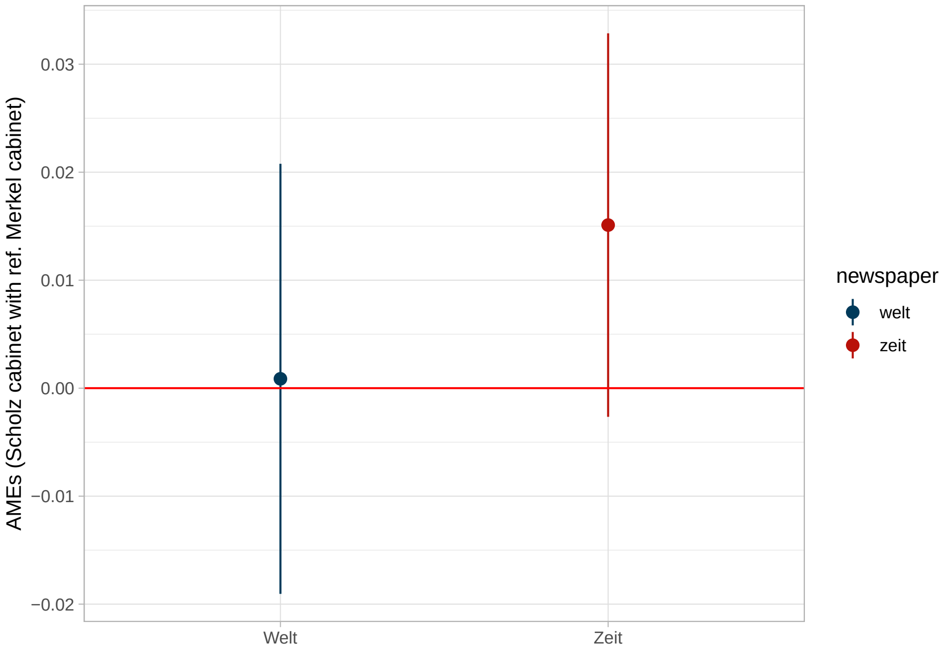

Figure 18: Marginal effect of Scholz government on quoting the opposition

might still perform worse but receives better news coverage.[26] I try to account for this by controlling for different indicators that also measure yearly and monthly changes as well as absolute values.

I also can't make any claims about the mechanisms and if the causes lie on the demand or supply side. Even though simple correlational data does speak against the former. Having access to internal data of newspapers on popularity of articles would be optimal when researching demand to see if newspapers react to shifts in popularity. Though, even then the supply still can have an effect on the demand.

There can also be other explanations that identify the cause of bias in other sources than the newspaper itself. In economic reporting, the newspapers depend a lot on official statistics. These are often shown in monthly/quarterly press releases by respective agencies. It could be the case that these press releases itself are already framed in a biased way. Often, newspapers simply adopt the text and publish it on their site. Future research can examine this relationship to test if news articles on these press releases are identical or if information is added or left out. Comparing these two is also a solution to the unobserved population problem.

Another limitation is caused by the selected case. There were two major events during both governments. First the Covid-pandemic and then the war on Ukraine. Economic reporting also has to do a lot with expectations about the future. It could be the case that the crisis disproportionately affected the outlook on the future, leading to more negative news. Generalizations of these results should be done carefully and future research should use even more comprehensive datasets covering more time periods. Focussing on single federal states can also be fruitful for future research as there is more variance for governments and could allow for difference-in-difference designs where two similar states are examined, but only one has a government change. Furthermore, I combined three different

[26]Reporting in such a way can be considered reasonable. If it is still considered biased depends on the definition of bias.

topics in the analysis. Future research could examine all of these separately to identify potentially heterogeneous effects. In a study on economic voting (Powell and Whitten, 1993), the authors find that left and center governments profit from good performance on unemployment while right-wing parties benefit from lower inflation. It could be the case that the media also judge parties differently based on which topics are more important for them. The study by Soroka (2012), for example, finds different biases for distinct topics. In my case, this could be especially relevant since the ministries for the economic topics are in the hands of three different parties. Again this raises the question if it is considered reasonable that parties are judged differently on topics based on their own priorities and therefore unbiased.

In regards to the methodology, it should be mentioned that while the text analysis using an LLM performed well overall, the identification of positive articles was worse. The results are at the lower bound of acceptability and should, therefore, be regarded with caution. Analyses of the validation samples with a bigger model (ChatGPT4) did consistently outperform the smaller one. The models on positive articles also weren't fully robust as not all model specifications resulted in significant estimates. This could be caused by noise in the text analysis but also by a lower number of positive articles in the sample.

6.2. Conclusion

In this study, my objective was to detect political media bias in economic news reporting. The hypothesis proposed that newspapers will be biased in favor of a government that is closer to their own ideological leaning. I used the case of Germany from March 2018 until June 2024, in which first the center-right party CDU led a cabinet, and since December 2021, the center-left party SPD led a more left-leaning cabinet. Specifically, I looked at articles from the newspapers Die Welt and Zeit Online. In my main models, I examined if a newspaper changed its reporting during the new government under the same economic conditions, which was achieved by controlling for several economic indicators. The results do indicate a negative bias for both newspapers against the Scholz government, at least in relative terms, as it could theoretically also be a positive bias in favor of the Merkel cabinet. Therefore, the hypothesis that both newspapers have a distinct bias is not confirmed.
The reasons for the bias remain unclear, though. Overall, the economic variables had little explanatory power for the news sentiment, which means there are other reasons for the negative reporting. There's no clear picture when looking at the influence of economic conditions on the probability of a negative article and the number of articles. Higher inflation under Merkel even leads to more articles by Zeit Online and them being more likely to be negative. On the other side, there is a decreasing number of articles during Scholz and higher GDP growth.
Even though I couldn't measure it properly, demand-side factors are unlikely the reason as there is a lack of correlation with the sentiment. Quoting patterns also don't serve as an explanation for bias. Both newspapers are a lot more likely to quote the governing parties than the opposition. Notably, though, the more left-leaning think thank DIW is quoted much less during the Scholz cabinet.
Only a linear time fixed effect rendered the coefficient of the bias insignificant. This means that there are important time-varying factors that I couldn't empirically assess in this study. Since both a more left- and right-leaning newspaper reported more negatively during the second government, this suggests that the source of the bias doesn't lie in the newspaper itself but is rather of a structural nature.

My results that speak for an existing bias against the Scholz government have important implications if they can be replicated over a broader time frame and across multiple newspapers. Skewed news reporting can influence people's opinions and ultimatively their voting behavior. Readers should be careful when forming their opinion based on news reporting and be aware that the information might not accurately portray reality even for something relatively objective such as the economy. Newspapers that hold themselves to the standard of objectivity should be aware of their responsibility and carefully assess how they portray the economic conditions across their articles. Finally, it wasn't my aim to examine the impact of the news on government popularity but it is thinkable that the bias I found contributed to it.

References

Aïmeur, E., Amri, S., and Brassard, G. (2023). Fake news, disinformation and misinformation in social media: a review. *Social Network Analysis and Mining*, 13(1):30.

Bashardoust, A., Feuerriegel, S., and Shrestha, Y. R. (2024). Comparing the willingness to share for human-generated vs. AI-generated fake news. Version Number: 1.

Belal, M., She, J., and Wong, S. (2023). Leveraging ChatGPT As Text Annotation Tool For Sentiment Analysis. Version Number: 1.

Brinkmann, J. (2018). *Verlagspolitik in der Zeitungskrise: Theorien, Strukturen, Strategien.* Number Band 12 in Reihe Medienökonomie. Nomos, Baden-Baden, 1. auflage edition.

Budge, I. (2013). The standard left-right scale. *University of Essex.*

Cohen, B. C. (1964). *Press and Foreign Policy.* Princeton University Press.

D'Alessio, D. and Allen, M. (2000). Media Bias in Presidential Elections: A Meta-Analysis. *Journal of Communication*, 50(4):133–156.

DellaVigna, S. and Kaplan, E. (2007). The Fox News Effect: Media Bias and Voting. *The Quarterly Journal of Economics*, 122(3):1187–1234.

Dewenter, R., Dulleck, U., and Thomas, T. (2018). The political coverage index and its application to government capture. *EcoAustria - Institute for Economic Research, Vienna.*

Dewenter, R., Dulleck, U., and Thomas, T. (2020). Does the 4th estate deliver? The Political Coverage Index and its application to media capture. *Constitutional Political Economy*, 31(3):292–328.

Doms, M. and Morin, N. J. (2004). Consumer sentiment, the economy, and the news media. Working Paper Series 2004-09, Federal Reserve Bank of San Francisco.

Dortmund, T. (2024). Journalismus & Demokratie.

Durante, R. and Knight, B. (2012). PARTISAN CONTROL, MEDIA BIAS, AND VIEWER RESPONSES: EVIDENCE FROM BERLUSCONI'S ITALY. *Journal of the European Economic Association*, 10(3):451–481.

Eyck, T. A. T. and Williment, M. (2003). The National Media and Things Genetic: Coverage in the New York Times (1971–2001) and the Washington Post (1977-2001). *Science Communication*, 25(2):129–152.

Fastenrath, F., Marx, P., Truger, A., and Vitt, H. (2022). Why is it so difficult to tax the rich? Evidence from German policy-makers. *Journal of European Public Policy*, 29(5):767–786.

Foos, F. and Bischof, D. (2022). Tabloid Media Campaigns and Public Opinion: Quasi-Experimental Evidence on Euroscepticism in England. *American Political Science Review*, 116(1):19–37.

Gans, H. J. (2004). *Deciding what's news: a study of "CBS Evening News, "NBC Nightly News, "Newsweek", and "Time"*. Visions of the American press. Northwestern university press, Evanston (Ill.).

Garz, M. (2014). Good news and bad news: evidence of media bias in unemployment reports. *Public Choice*, 161(3-4):499–515.

Garz, M. (2018). Effects of unemployment news on economic perceptions – Evidence from German Federal States. *Regional Science and Urban Economics*, 68:172–190.

Gentzkow, M. and Shapiro, J. (2006). Media Bias and Reputation. *Journal of Political Economy*, 114(2):280–316.

Gentzkow, M. and Shapiro, J. M. (2010). What Drives Media Slant? Evidence From U.S. Daily Newspapers. *Econometrica*, 78(1):35–71.

Gerber, A. S., Karlan, D., and Bergan, D. (2009). Does the Media Matter? A Field Experiment Measuring the Effect of Newspapers on Voting Behavior and Political Opinions. *American Economic Journal: Applied Economics*, 1(2):35–52. Publisher: American Economic Association.

Gilardi, F., Alizadeh, M., and Kubli, M. (2023). ChatGPT outperforms crowd workers for text-annotation tasks. *Proceedings of the National Academy of Sciences*, 120(30):e2305016120.

Grimmer, J. and Stewart, B. M. (2013). Text as Data: The Promise and Pitfalls of Automatic Content Analysis Methods for Political Texts. *Political Analysis*, 21(3):267–297.

Groeling, T. (2013). Media Bias by the Numbers: Challenges and Opportunities in the Empirical Study of Partisan News. *Annual Review of Political Science*, 16(1):129–151.

Groseclose, T. and Milyo, J. (2005). A Measure of Media Bias. *The Quarterly Journal of Economics*, 120(4):1191–1237.

Grossman, G., Margalit, Y., and Mitts, T. (2022). How the Ultrarich Use Media Ownership as a Political Investment. *The Journal of Politics*, 84(4):1913–1931.

Haselmayer, M. and Jenny, M. (2017). Sentiment analysis of political communication: combining a dictionary approach with crowdcoding. *Quality & Quantity*, 51(6):2623–2646.

Kaid, L. L. and Holtz-Bacha, C. (2008). *Encyclopedia of political communication*. Sage Publications, Los Angeles. OCLC: 192020698.

Klayman, J. (1995). Varieties of Confirmation Bias. In *Psychology of Learning and Motivation*, volume 32, pages 385–418. Elsevier.

Koplin, J. J. (2023). Dual-use implications of AI text generation. *Ethics and Information Technology*, 25(2):32.

Krippendorff, K. (2011). Computing Krippendorff's Alpha-Reliability.

Kropko, J. and Kubinec, R. (2020). Interpretation and identification of within-unit and cross-sectional variation in panel data models. *PLOS ONE*, 15(4):e0231349.

Krugmann, J. O. and Hartmann, J. (2024). Sentiment Analysis in the Age of Generative AI. *Customer Needs and Solutions*, 11(1):3.

Larcinese, V., Puglisi, R., and Snyder, J. M. (2011). Partisan bias in economic news: Evidence on the agenda-setting behavior of U.S. newspapers. *Journal of Public Economics*, 95(9-10):1178–1189.

Lehmann, P., Franzmann, S., Al-Gaddooa, D., Burst, T., Ivanusch, C., Regel, S., Riethmüller, F., Volkens, A., Weßels, B., Zehnter, L., Wissenschaftszentrum Berlin Für Sozialforschung (WZB), and Institut Für Demokratieforschung Göttingen (IfDem) (2024). Manifesto Project Dataset.

Lengauer, G., Donges, P., and Plasser, F. (2014). Media Power in Politics. In Pfetsch, B., editor, *Political Communication Cultures in Europe: Attitudes of Political Actors and Journalists in Nine Countries*, pages 171–195. Palgrave Macmillan UK, London.

Lott, J. R. and Hassett, K. A. (2014). Is newspaper coverage of economic events politically biased? *Public Choice*, 160(1-2):65–108.

Louw, P. E. (2005). *The media and political process*. SAGE, London ; Thousand Oaks, Calif. OCLC: ocm57428807.

Lowry, D. T. (2008). Network Tv News Framing of Good Vs. Bad Economic News under Democrat and Republican Presidents: A Lexical Analysis of Political Bias. *Journalism & Mass Communication Quarterly*, 85(3):483–498.

Ludvigson, S. C. (2004). Consumer Confidence and Consumer Spending. *Journal of Economic Perspectives*, 18(2):29–50.

Marzi, G., Balzano, M., and Marchiori, D. (2024). K-Alpha Calculator–Krippendorff's Alpha Calculator: A user-friendly tool for computing Krippendorff's Alpha inter-rater reliability coefficient. *MethodsX*, 12:102545.

Maurer, M. and Reinemann, C. (2006). *Medieninhalte*. VS Verlag für Sozialwissenschaften, Wiesbaden.

Maurer, P. (2011). Explaining perceived media influence in politics: An analysis of the interplay of context and attitudes in four European democracies. *Publizistik*, 56(1):27–50.

McChesney, R. W. and Pickard, V. (2014). *News Media as Political Institutions*, volume 1. Oxford University Press.

McCombs, M. E. and Shaw, D. L. (1972). The Agenda-Setting Function of Mass Media. *The Public Opinion Quarterly*, 36(2):176–187. Publisher: [Oxford University Press, American Association for Public Opinion Research].

McNair, B. (2018). *An introduction to political communication*. Communication and society. Routledge, Taylor & Francis Group, London New York, sixth edition edition.

Mullainathan, S. and Shleifer, A. (2002). Media Bias. NBER Working Papers 9295, National Bureau of Economic Research, Inc.

Mullainathan, S. and Shleifer, A. (2005). The Market for News. *The American Economic Review*, 95(4):1031–1053. Publisher: American Economic Association.

Newman, N., Fletcher, R., Robertson, C. T., Arguedas, A. R., and Nielsen, R. K. (2024). Digital News Report 2024.

Ollion, E., Shen, R., Macanovic, A., and Chatelain, A. (2023). ChatGPT for Text Annotation? Mind the Hype!

Pangakis, N., Wolken, S., and Fasching, N. (2023). Automated Annotation with Generative AI Requires Validation. Version Number: 1.

Patterson, T. E. and Donsbagh, W. (1996). News decisions: Journalists as partisan actors. *Political Communication*, 13(4):455–468.

Pfetsch, B. and Adam, S., editors (2008). *Massenmedien als politische Akteure*. VS Verlag für Sozialwissenschaften, Wiesbaden.

Powell, G. B. and Whitten, G. D. (1993). A Cross-National Analysis of Economic Voting: Taking Account of the Political Context. *American Journal of Political Science*, 37(2):391.

Prat, A. (2015). Media Capture and Media Power. In *Handbook of Media Economics*, volume 1, pages 669–686. Elsevier.

Puglisi, R. and Snyder, J. M. (2011). Newspaper Coverage of Political Scandals. *The Journal of Politics*, 73(3):931–950.

Puglisi, R. and Snyder, J. M. (2015). Empirical Studies of Media Bias. In *Handbook of Media Economics*, volume 1, pages 647–667. Elsevier.

Quiring, O., Ziegele, M., Schultz, T., Fawzi, N., Jackob, N., Jakobs, I., Schemer, C., Stegmann, D., and Viehmann, C. (2024). Zurück zum Niveau vor der Pandemie – Konsolidierung von Vertrauen und Misstrauen. Mainzer Langzeitstudie Medienvertrauen 2023. *Media Perspektiven*, 9:1–14.

Reiss, M. V. (2023). Testing the Reliability of ChatGPT for Text Annotation and Classification: A Cautionary Remark. Version Number: 1.

Rodrigo-Ginés, F.-J., Carrillo-de Albornoz, J., and Plaza, L. (2024). A systematic review on media bias detection: What is media bias, how it is expressed, and how to detect it. *Expert Systems with Applications*, 237:121641.

Schudson, M. (2014). *How to Think Normatively About News and Democracy*, volume 1. Oxford University Press.

Semetko, H. A., Blumler, J. G., Gurevitch, M., Weaver, D. H., Barkin, S. M., and Wilhoit, G. C. (2011). *The formation of campaign agendas: a comparative analysis of party and media roles in recent American and British elections*. Routledge, London. OCLC: 773432068.

Shoemaker, P. J. and Vos, T. P. (2009). *Gatekeeping theory*. Routledge, New York. OCLC: ocn258331230.

Soroka, S. N. (2012). The Gatekeeping Function: Distributions of Information in Media and the Real World. *The Journal of Politics*, 74(2):514–528.

Spinde, T., Hamborg, F., and Gipp, B. (2020). Media Bias in German News Articles: A Combined Approach. In Koprinska, I., Kamp, M., Appice, A., Loglisci, C., Antonie, L., Zimmermann, A., Guidotti, R., Özgöbek, , Ribeiro, R. P., Gavaldà, R., Gama, J., Adilova, L., Krishnamurthy, Y., Ferreira, P. M., Malerba, D., Medeiros, I., Ceci, M., Manco, G., Masciari, E., Ras, Z. W., Christen, P., Ntoutsi, E., Schubert, E., Zimek, A., Monreale, A., Biecek, P., Rinzivillo, S., Kille, B., Lommatzsch, A., and Gulla, J. A., editors, *ECML PKDD 2020 Workshops*, volume 1323, pages 581–590. Springer International Publishing, Cham. Series Title: Communications in Computer and Information Science.

Spinde, T., Rudnitckaia, L., Mitrović, J., Hamborg, F., Granitzer, M., Gipp, B., and Donnay, K. (2021). Automated identification of bias inducing words in news articles using linguistic and context-oriented features. *Information Processing & Management*, 58(3):102505.

Strömbäck, J. and Nord, L. W. (2006). Do Politicians Lead the Tango?: A Study of the Relationship between Swedish Journalists and their Political Sources in the Context of Election Campaigns. *European Journal of Communication*, 21(2):147–164.

Taber, C. S. and Lodge, M. (2006). Motivated Skepticism in the Evaluation of Political Beliefs. *American Journal of Political Science*, 50(3):755–769.

Törnberg, P. (2024). Best Practices for Text Annotation with Large Language Models. Version Number: 1.

Van Aelst, P., Brants, K., Van Praag, P., De Vreese, C., Nuytemans, M., and Van Dalen, A. (2008a). THE FOURTH ESTATE AS SUPERPOWER?: An empirical study of perceptions of media power in Belgium and the Netherlands. *Journalism Studies*, 9(4):494–511.

Van Aelst, P., Maddens, B., Noppe, J., and Fiers, S. (2008b). Politicians in the News: Media or Party Logic?: Media Attention and Electoral Success in the Belgian Election Campaign of 2003. *European Journal of Communication*, 23(2):193–210.

Waldherr, A. (2008). Gatekeeper, Diskursproduzenten und Agenda-Setter — Akteursrollen von Massenmedien in Innovationsprozessen. In Pfetsch, B. and Adam, S., editors, *Massenmedien als politische Akteure*, pages 171–195. VS Verlag für Sozialwissenschaften, Wiesbaden.

Williams, A. (1975). Unbiased Study of Television News Bias. *Journal of Communication*, 25(4):190–199.

Yagci, A. H. and Oyvat, C. (2020). Partisanship, media and the objective economy: Sources of individual-level economic assessments. *Electoral Studies*, 66:102135.

A. Appendix

A.1. Validation results

Table A1: Confusion Matrix for text classification (ChatGPT)

	Reference	
Prediction	Ja	Nein
Ja	50	6
Nein	8	536

Table A2: Text classification performance metrics (ChatGPT)

Statistic	Value
Accuracy	0.98
Precision	0.89
Recall	0.86
F1	0.88

Table A3: Confusion Matrix for sentiment analysis (ChatGPT)

	Reference		
Prediction	Negativ	Neutral	Positiv
Negativ	85	5	1
Neutral	9	8	2
Positiv	1	6	24

Table A4: Sentiment analysis performance metrics (ChatGPT)

	Precision	Recall	F1
Class: Negativ	0.93	0.89	0.91
Class: Neutral	0.42	0.42	0.42
Class: Positiv	0.77	0.89	0.83

A.2. Regression Diagnostics

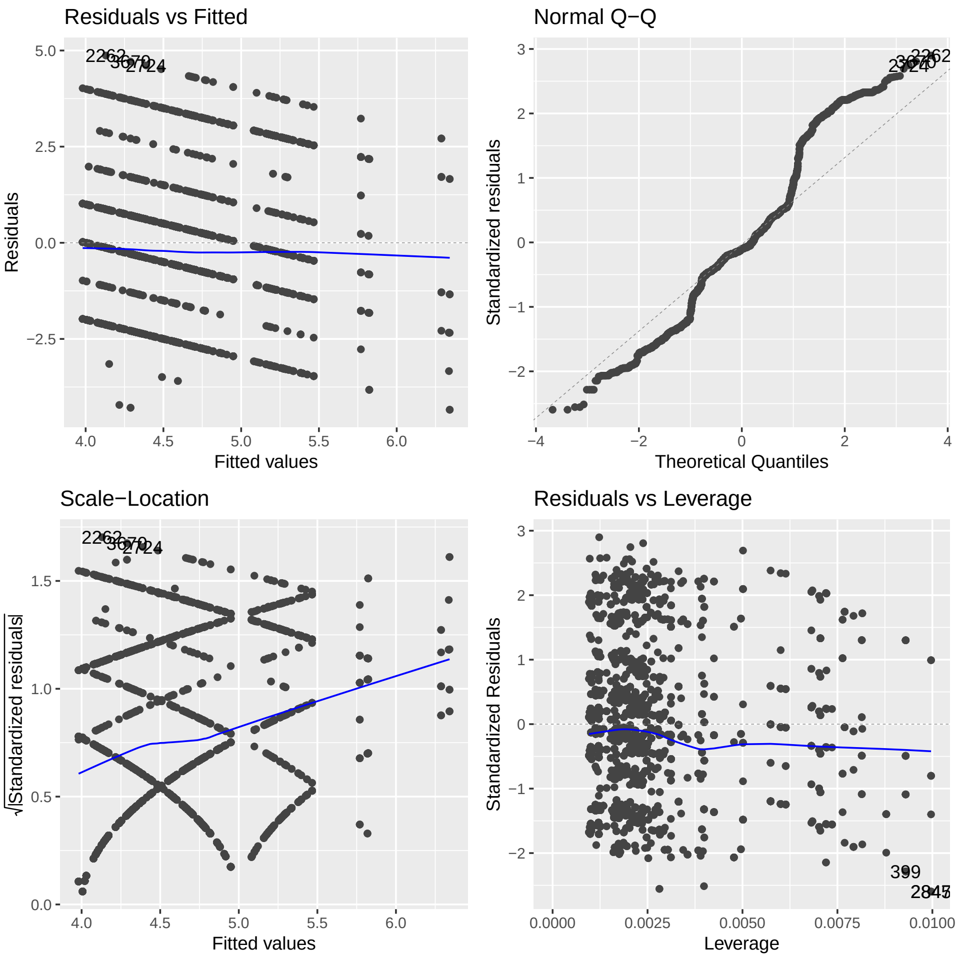

Figure A1: Regression diagnostic plots for the OLS model

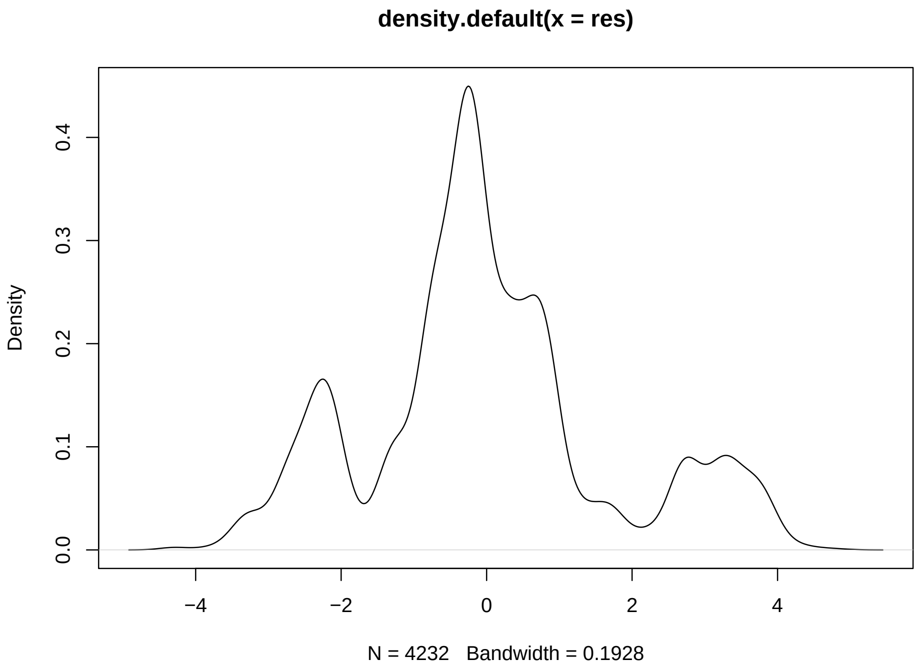

Figure A2: Regression diagnostic plot: Error distribution

A.3. Logistic regression results

Table A5: Regression results for logistic regression (Negative articles)

Dependent Variable:	Article is negative	
Model:	(1)	(2)
Variables		
Constant	0.0558	1.546*
	(0.0535)	(0.7359)
Newspaper: Zeit (ref. Welt	-0.5746***	-0.5750***
	(0.0935)	(0.0946)
Government: Scholz (ref. Merkel)	0.5691***	0.4741***
	(0.0787)	(0.1220)
Interaction: Zeit x Scholz	0.1585	0.1863
	(0.1313)	(0.1333)
Yearly inflation		0.0420·
		(0.0218)
Unemployment rate		-0.2914*
		(0.1350)
Quarterly GDP change (price adjusted)		0.0123
		(0.0126)
Monthly Inflation Change		0.0489
		(0.0765)
Yearly GDP change (price adjusted)		-0.0770***
		(0.0174)
Yearly Change of unemployment rate		-0.0408
		(0.1074)
Fit statistics		
Standard-Errors	IID	
Observations	4,232	4,232
Squared Correlation	0.03594	0.04588
Pseudo R^2	0.02630	0.03327
BIC	5,719.9	5,729.3

IID standard-errors in parentheses
*Signif. Codes: ***: 0.001, **: 0.01, *: 0.05, .: 0.1*

Table A6: Regression results for logistic regression (Positive articles)

Dependent Variable:	Article is positive	
Model:	(1)	(2)
Variables		
Constant	-1.706***	-4.507***
	(0.0742)	(0.9710)
Newspaper: Zeit (ref. Welt)	0.7945***	0.7880***
	(0.1105)	(0.1124)
Government: Scholz (ref. Merkel)	-1.052***	-1.177***
	(0.1377)	(0.1949)
Interaction: Zeit x Scholz	-0.0592	-0.0035
	(0.1953)	(0.1982)
Yearly inflation		-0.0086
		(0.0346)
Unemployment rate		0.5083**
		(0.1782)
Quarterly GDP change (price adjusted)		0.0004
		(0.0184)
Monthly Inflation Change		0.2201*
		(0.1084)
Yearly GDP change (price adjusted)		0.0472*
		(0.0203)
Yearly Change of unemployment rate		-0.4326**
		(0.1463)
Fit statistics		
Standard-Errors	IID	
Observations	4,232	4,232
Squared Correlation	0.04882	0.06268
Pseudo R^2	0.05787	0.07191
BIC	3,267.9	3,269.8

IID standard-errors in parentheses
*Signif. Codes: ***: 0.001, **: 0.01, *: 0.05, .: 0.1*

A.4. Sensitivity results for logistic regressions

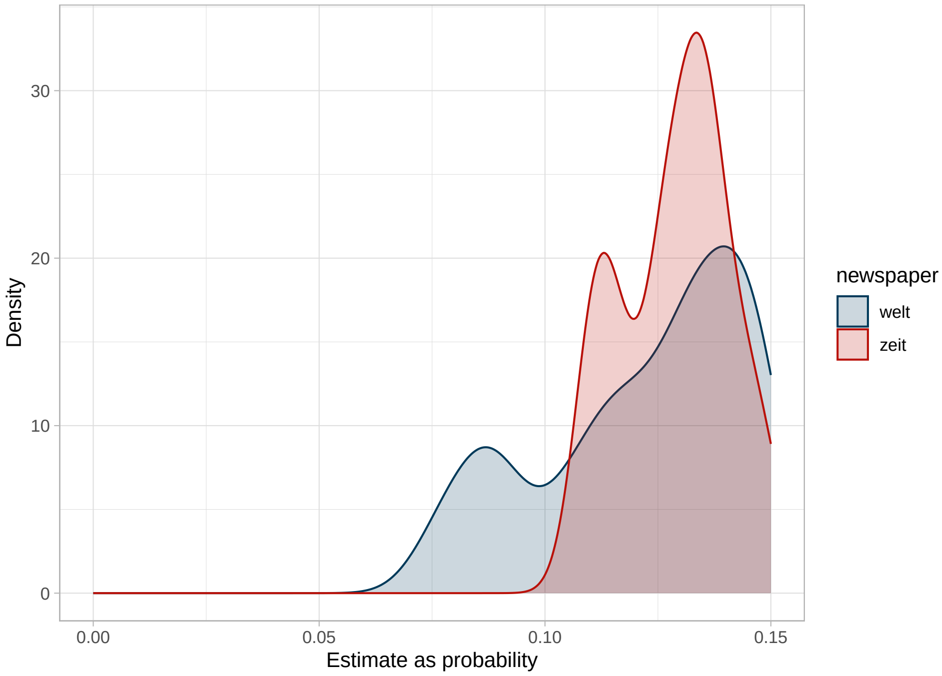

Figure A3: Distribution of estimates with differing control variables (logistic regression on negative articles)

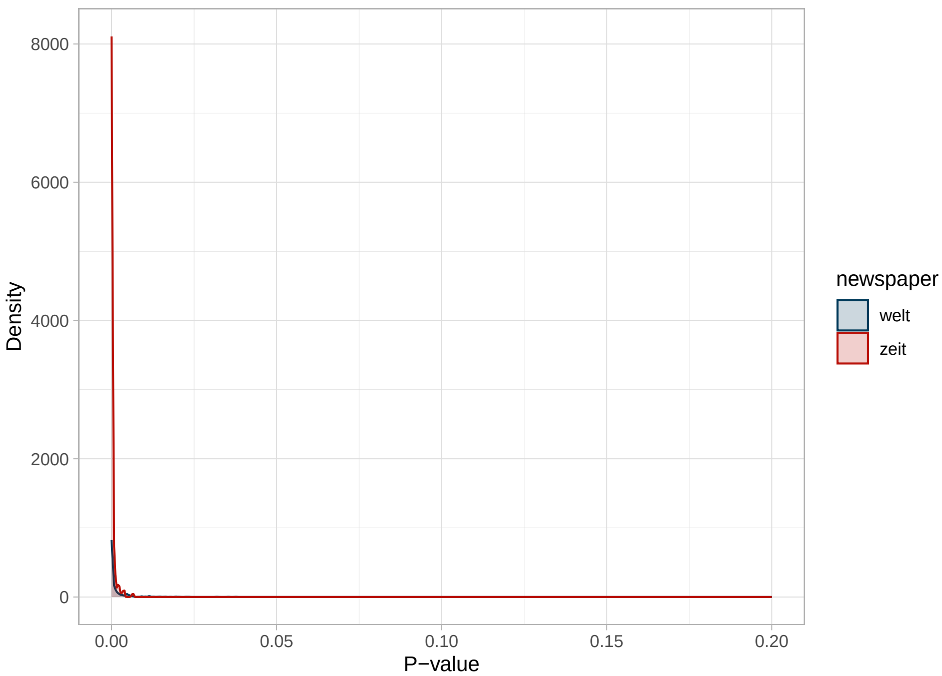

Figure A4: Distribution of p-values with differing control variables (logistic regression on negative articles)

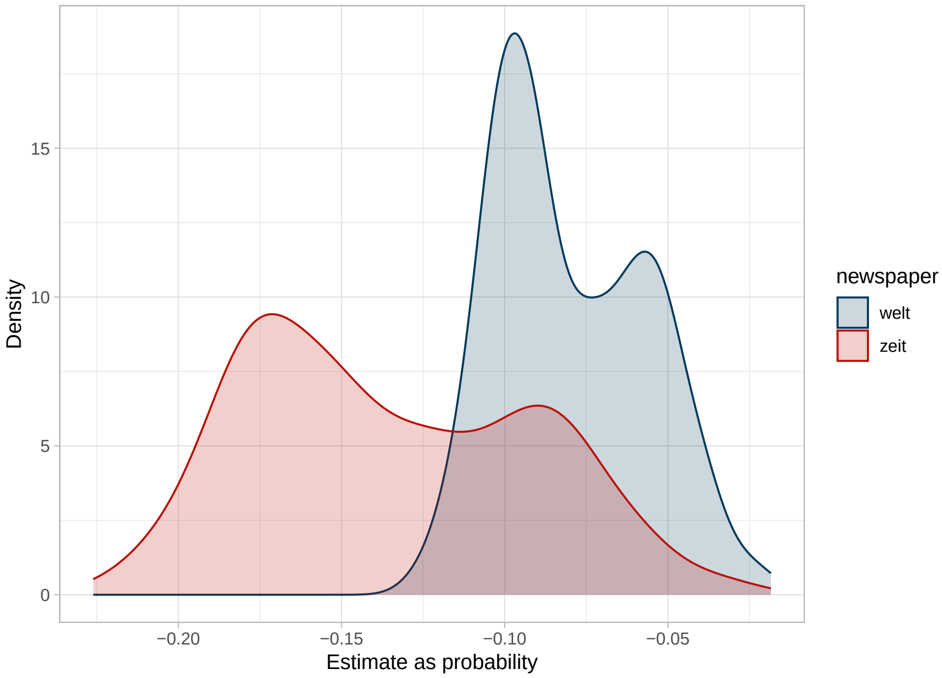

Figure A5: Distribution of estimates with differing control variables (logistic regression on positive articles)

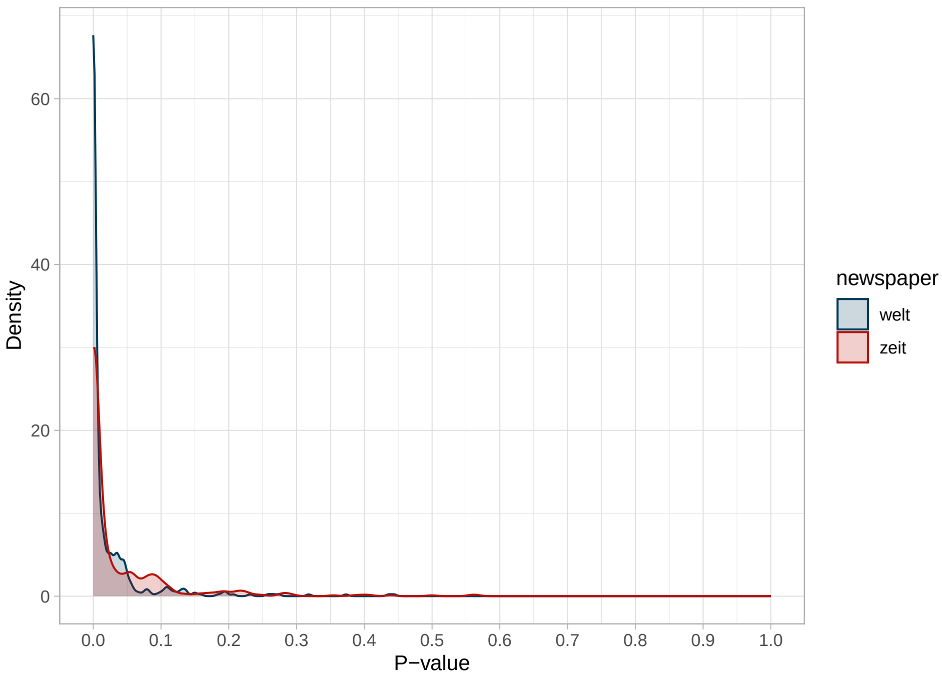

Figure A6: Distribution of p-values with differing control variables (logistic regression on positive articles)

A.5. First robustness analysis

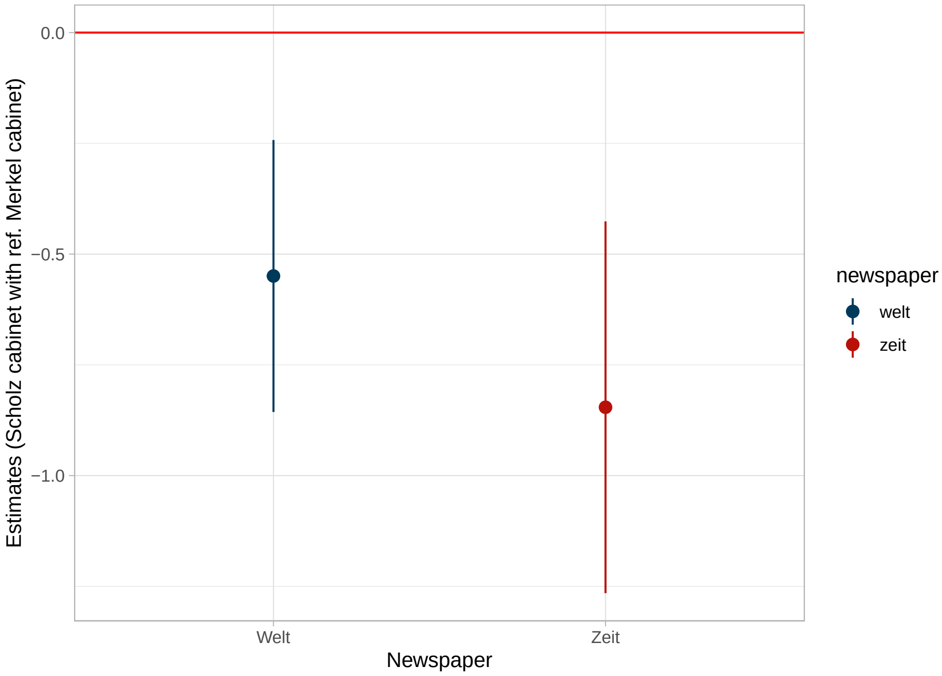

Figure A7: Marginal effects of Scholz government (OLS model)

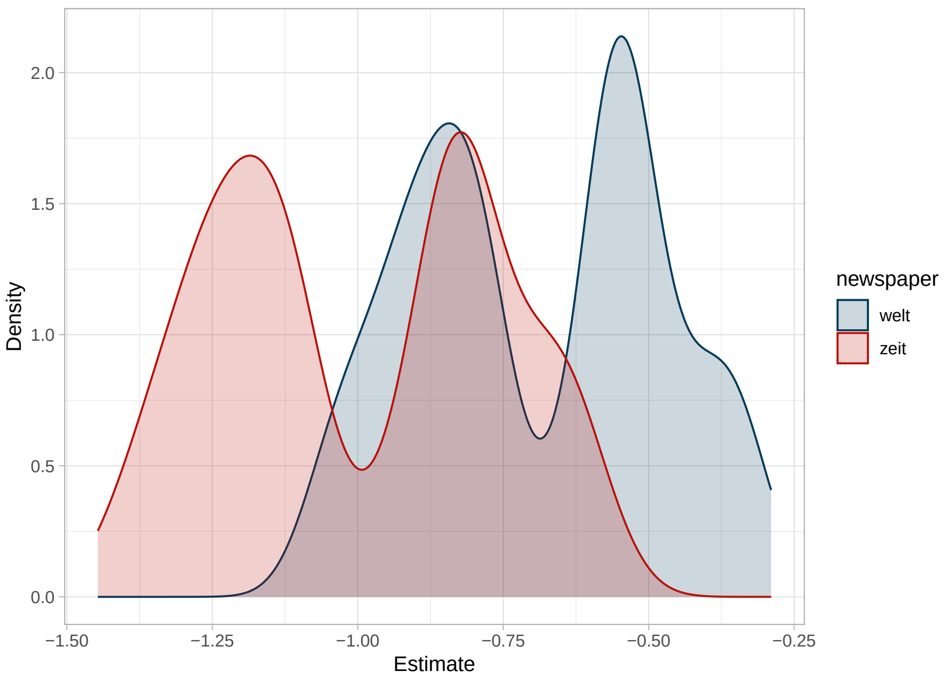

Figure A8: Distribution of estimates with differing control variables (OLS)

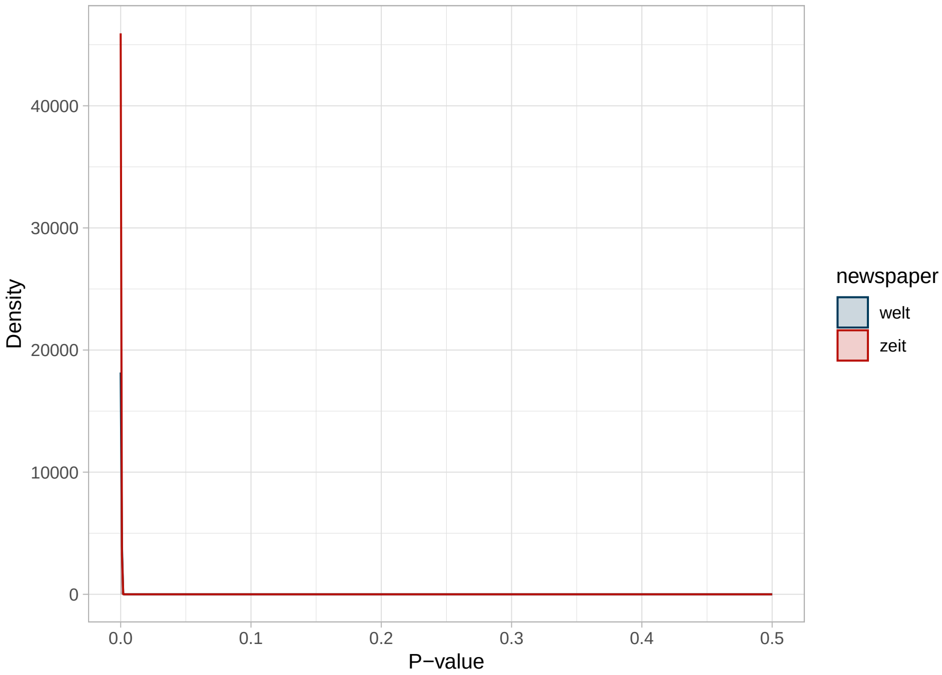

Figure A9: Distribution of p-values with differing control variables (OLS)

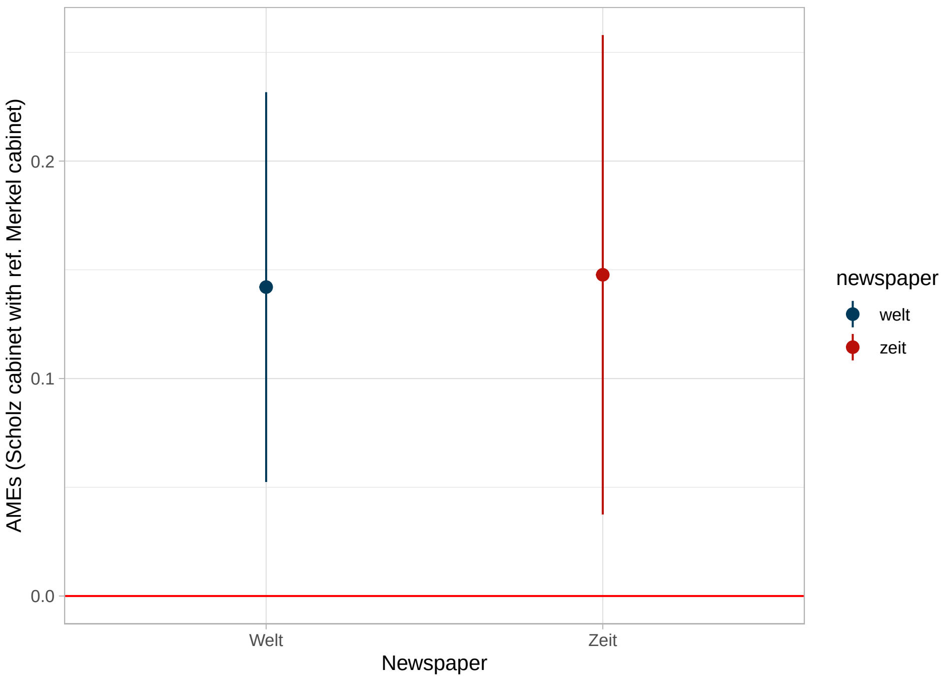

Figure A10: Average marginal effect of Scholz government (negative articles)

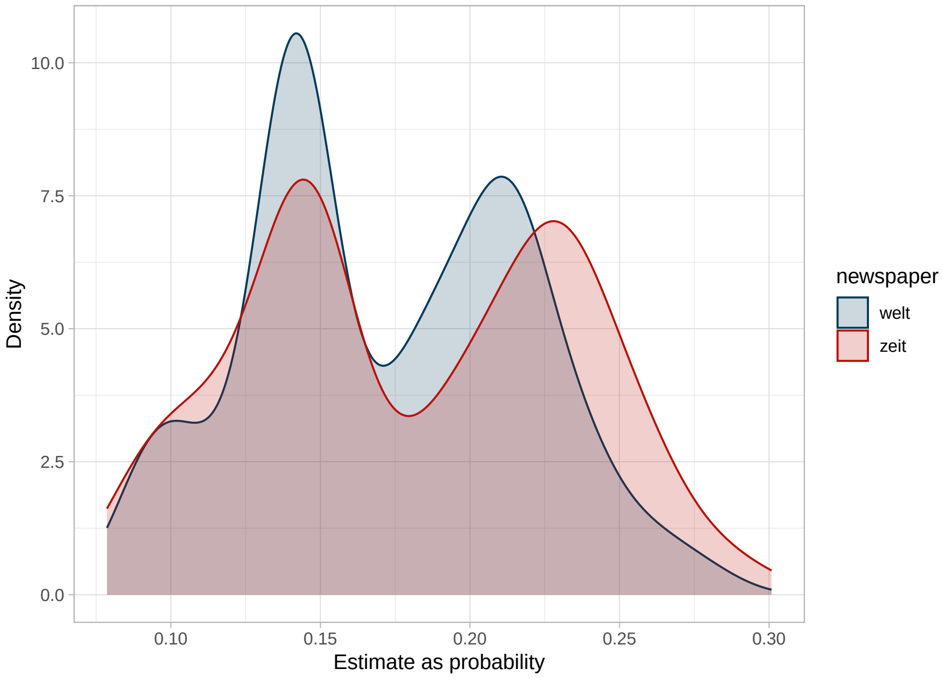

Figure A11: Distribution of estimates with differing control variables (logistic regression on negative articles)

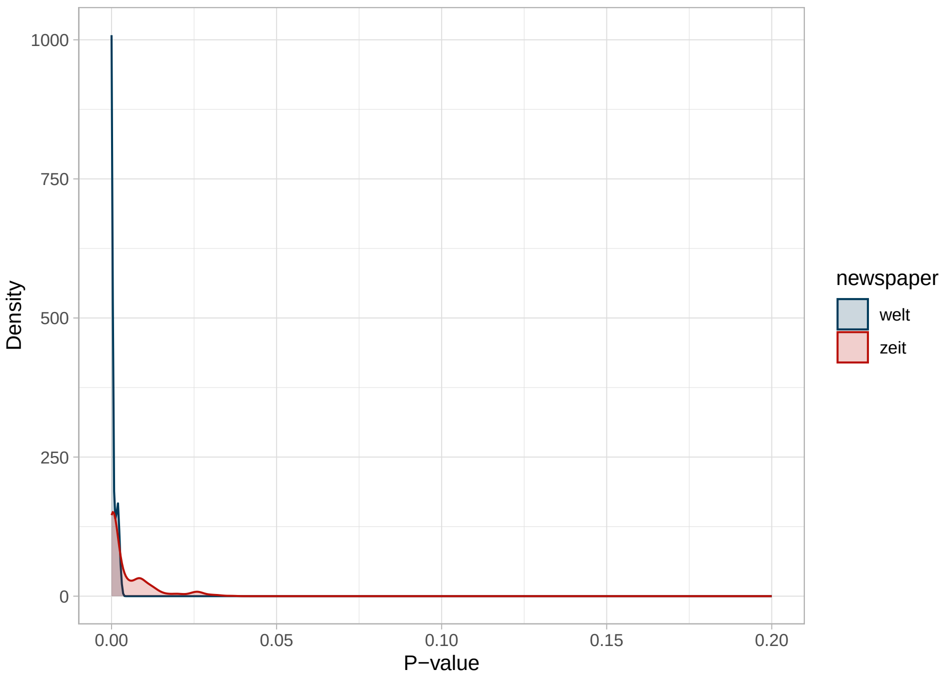

Figure A12: Distribution of p-values with differing control variables (logistic regression on negative articles)

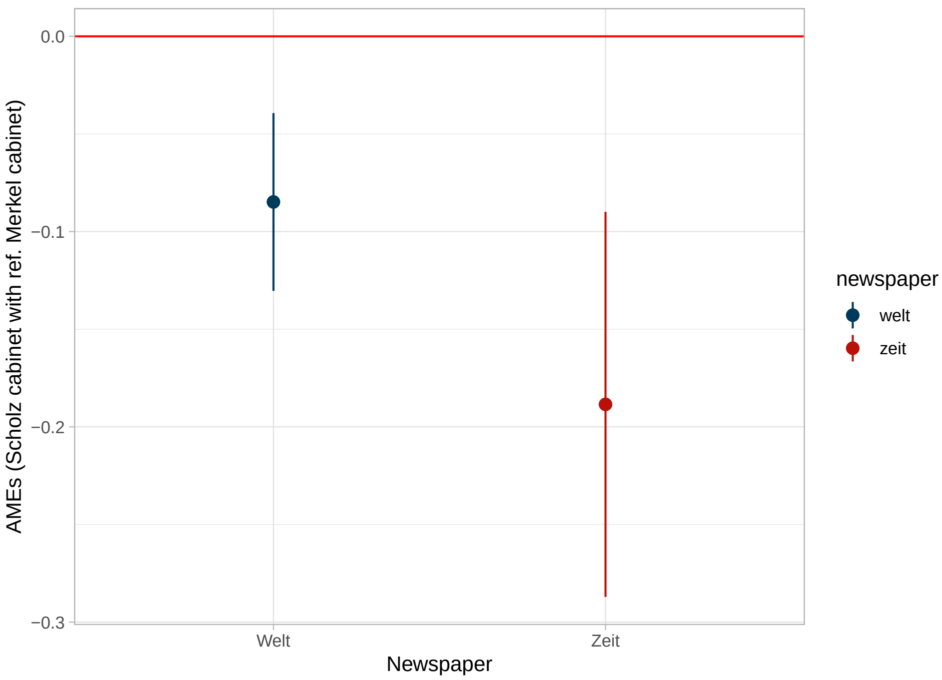

Figure A13: Average marginal effect of Scholz government (positive articles)

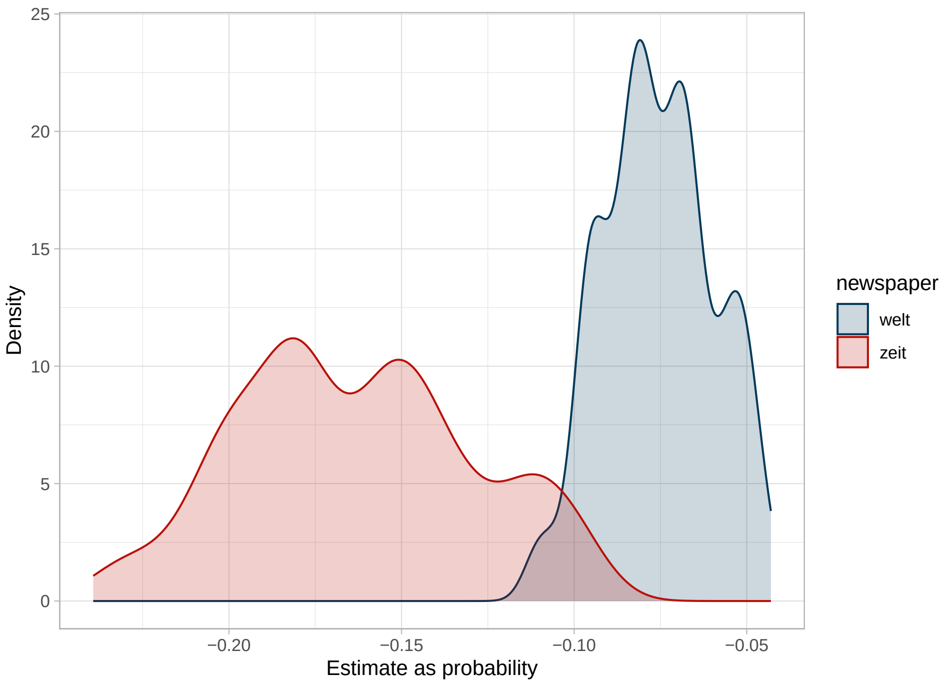

Figure A14: Distribution of estimates with differing control variables (logistic regression on positive articles)

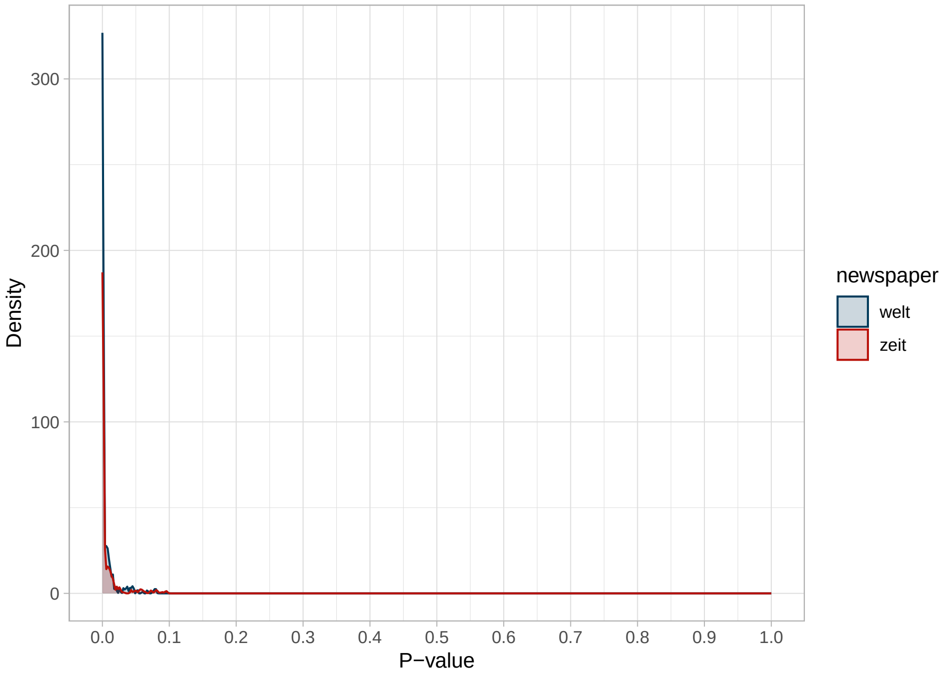

Figure A15: Distribution of p-values with differing control variables (logistic regression on positive articles)

A.6. Second robustness analysis

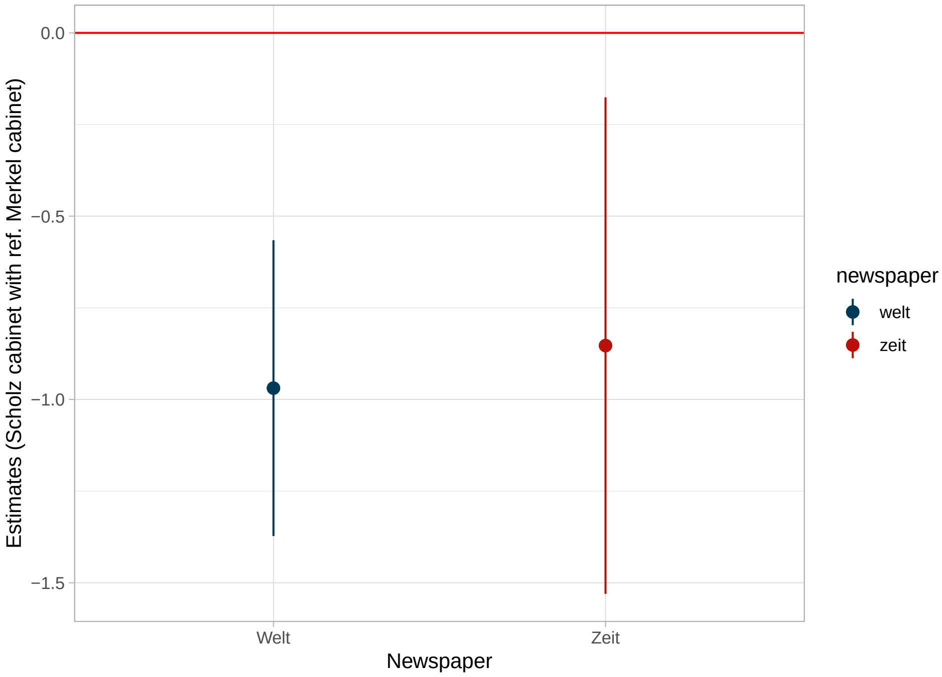

Figure A16: Marginal effects of Scholz government (OLS model)

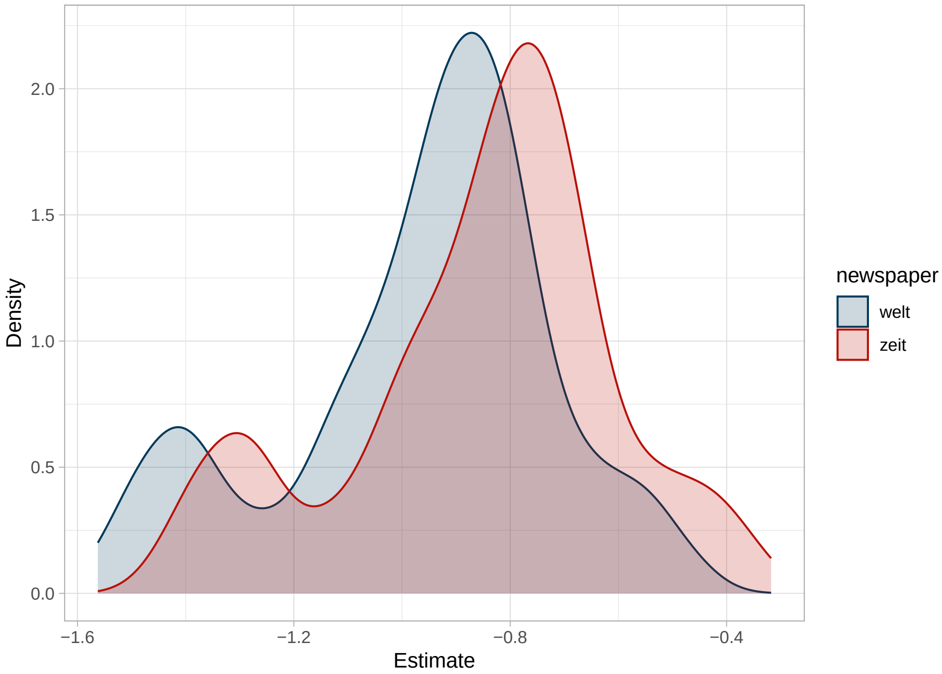

Figure A17: Distribution of estimates with differing control variables (OLS)

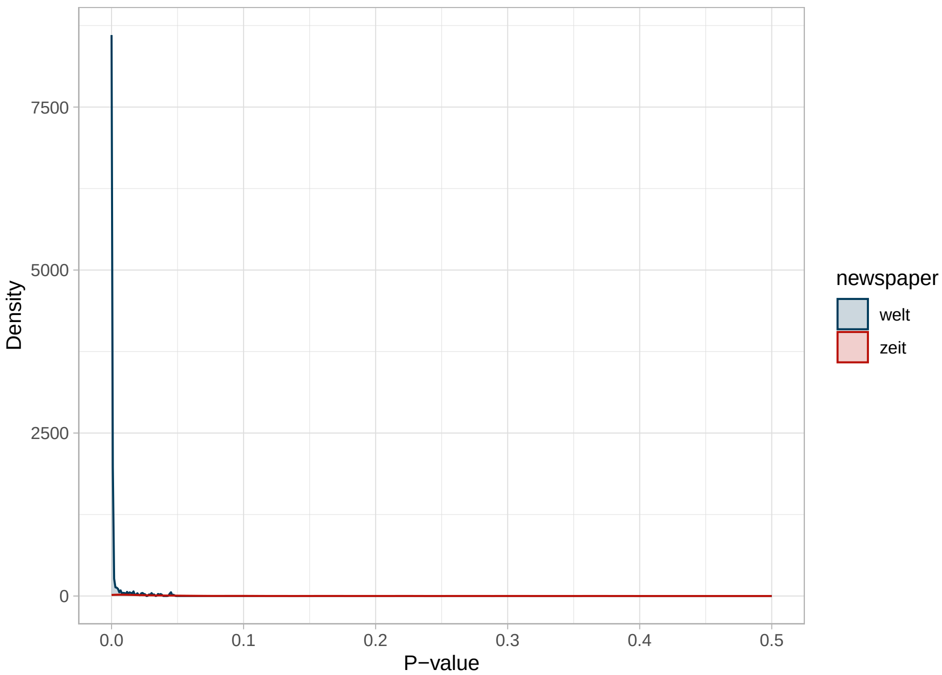

Figure A18: Distribution of p-values with differing control variables (OLS)

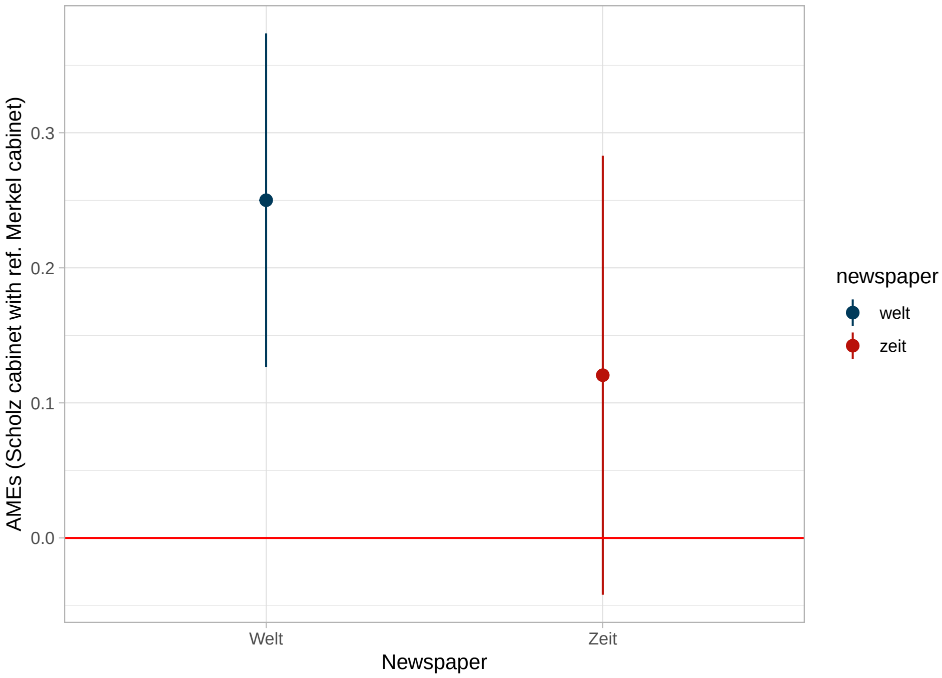

Figure A19: Average marginal effect of Scholz government (negative articles)

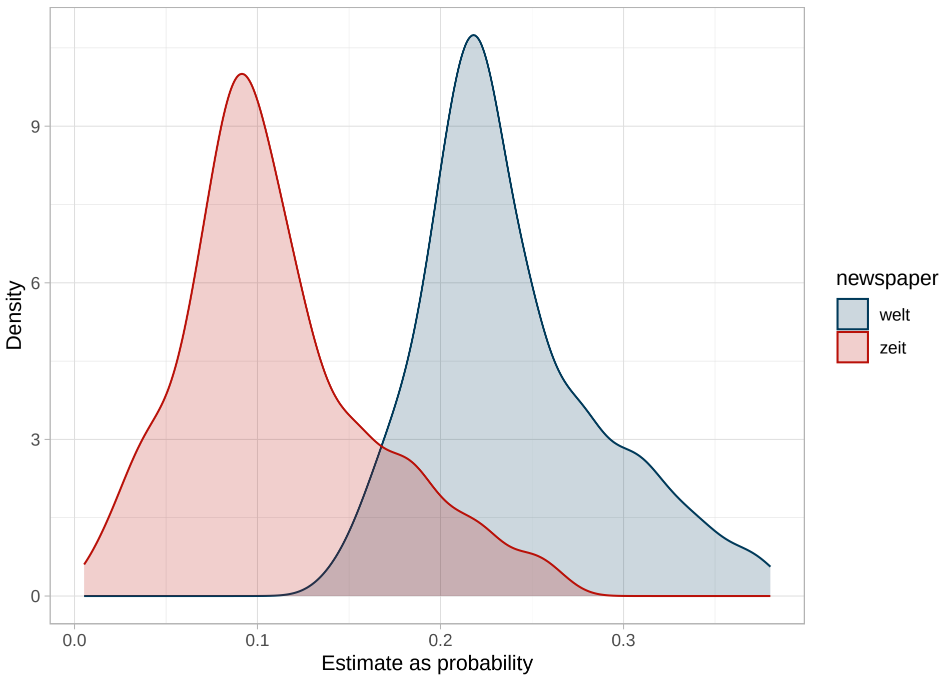

Figure A20: Distribution of estimates with differing control variables (logistic regression on negative articles)

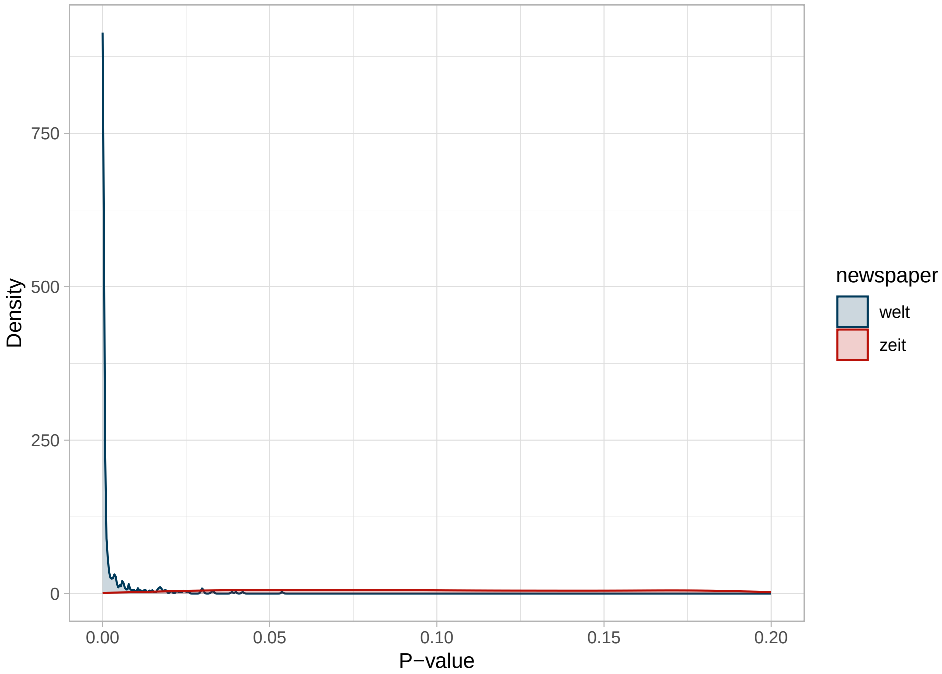

Figure A21: Distribution of p-values with differing control variables (logistic regression on negative articles)

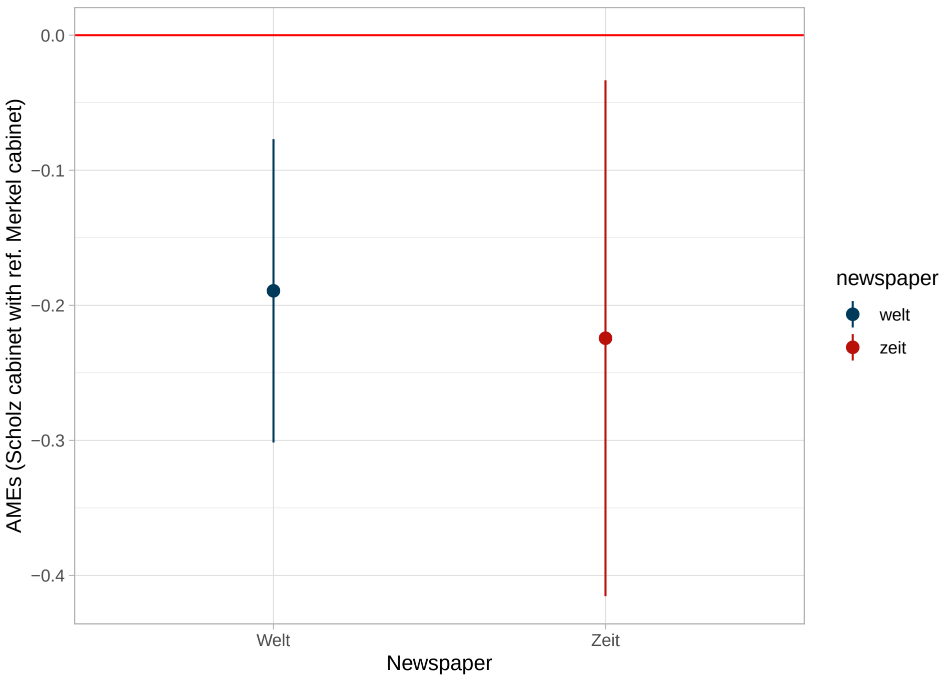

Figure A22: Average marginal effect of Scholz government (positive articles)

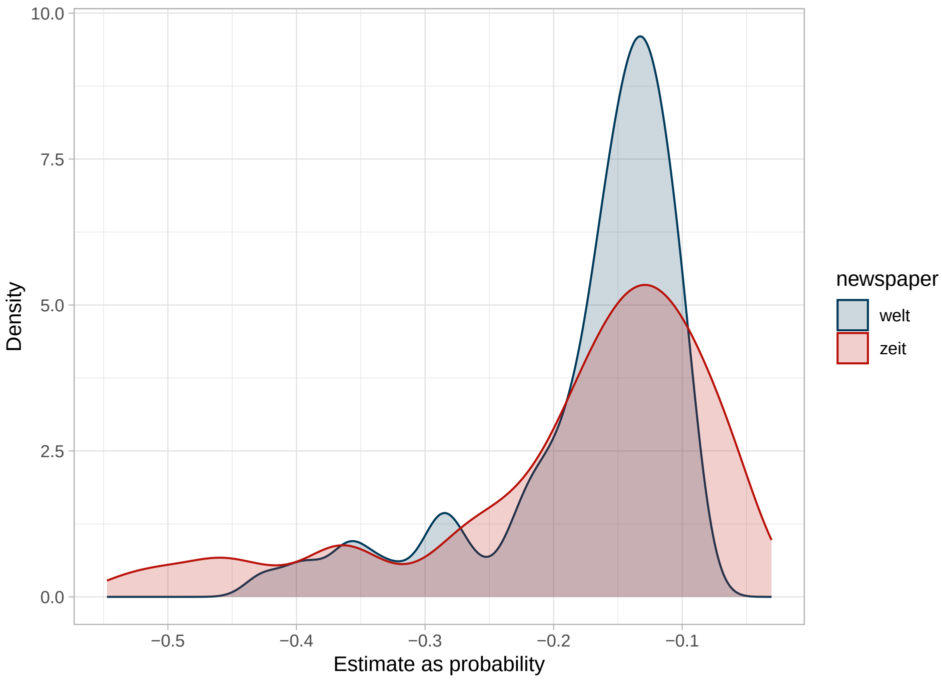

Figure A23: Distribution of estimates with differing control variables (logistic regression on positive articles)

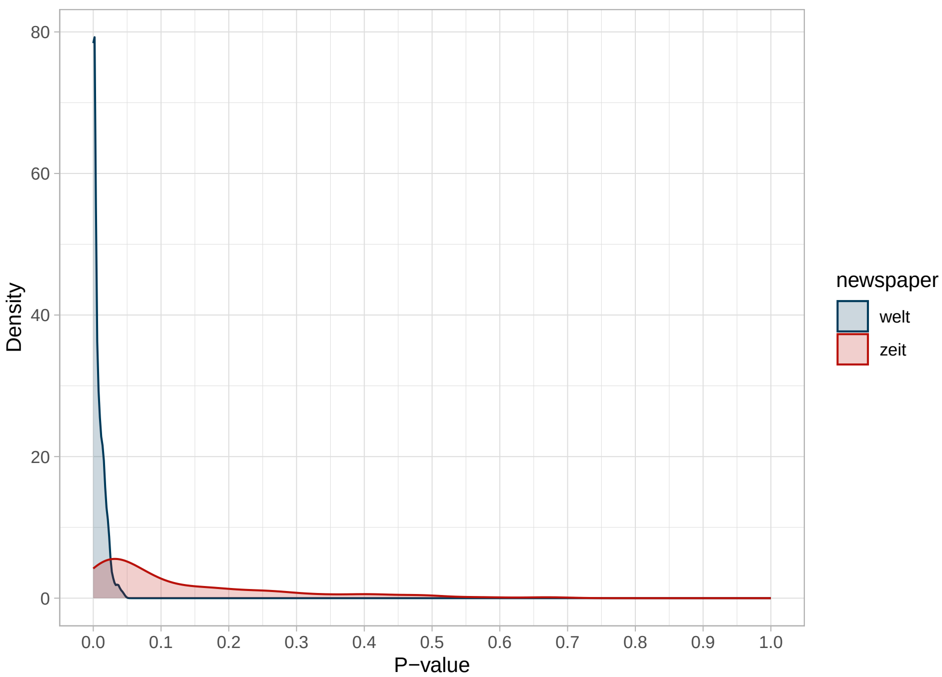

Figure A24: Distribution of p-values with differing control variables (logistic regression on positive articles)

A.7. Third robustness analysis

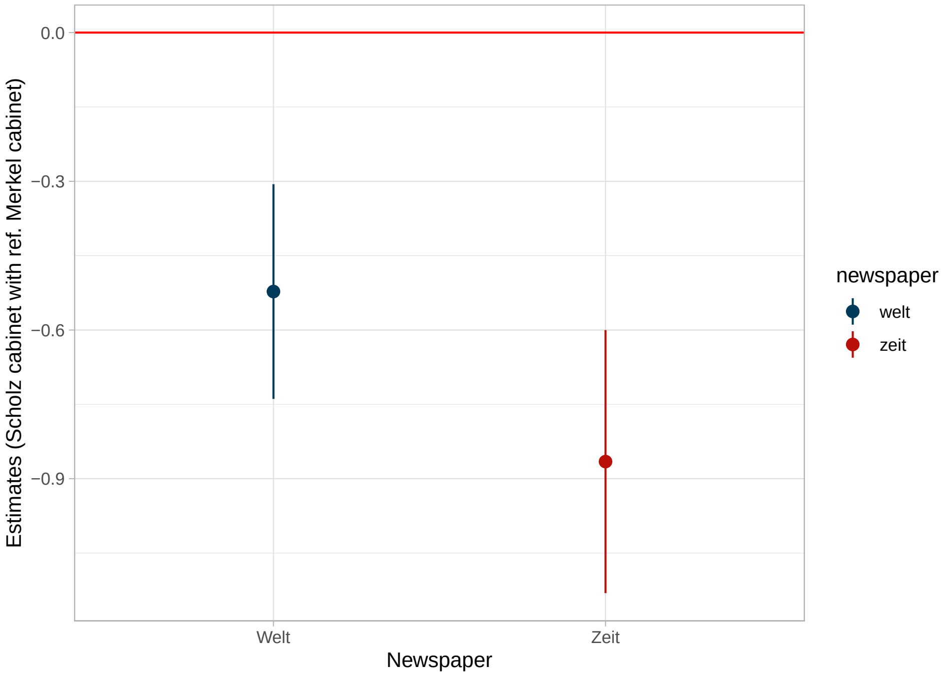

Figure A25: Marginal effects of Scholz government (OLS model)

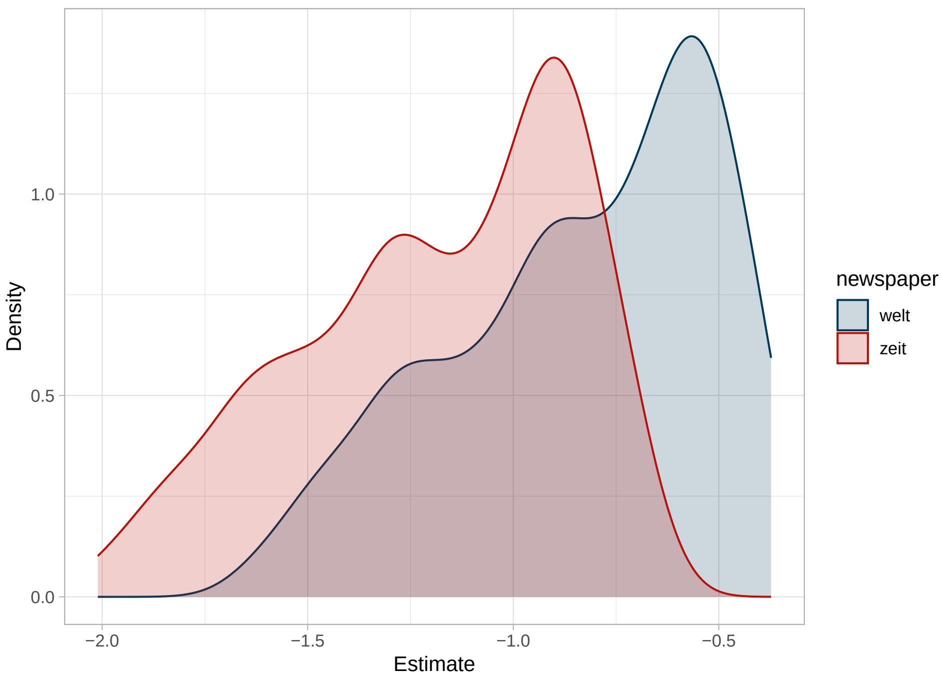

Figure A26: Distribution of estimates with differing control variables (OLS)

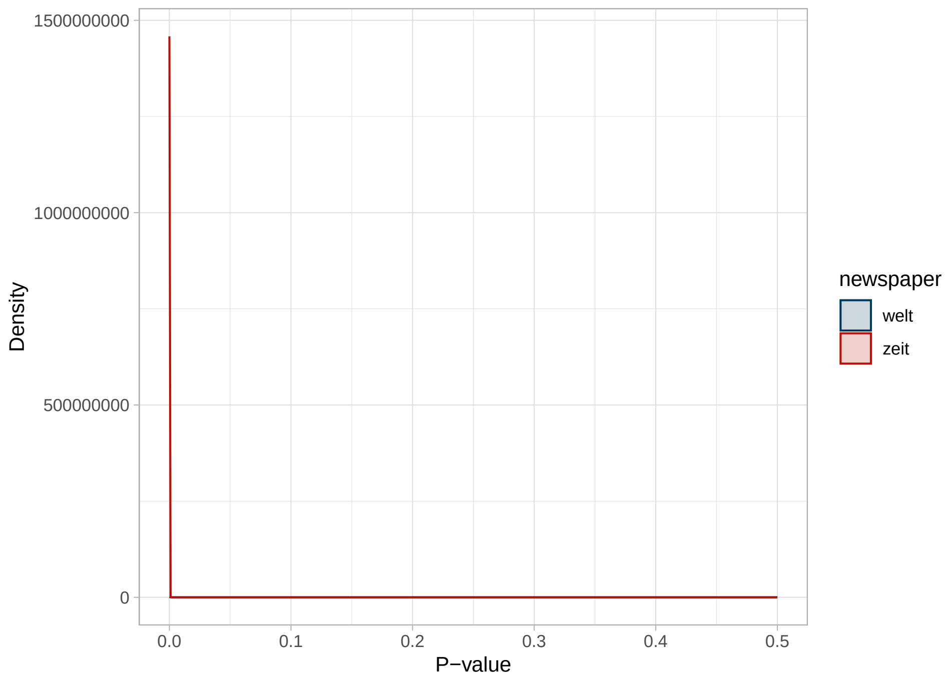

Figure A27: Distribution of p-values with differing control variables (OLS)

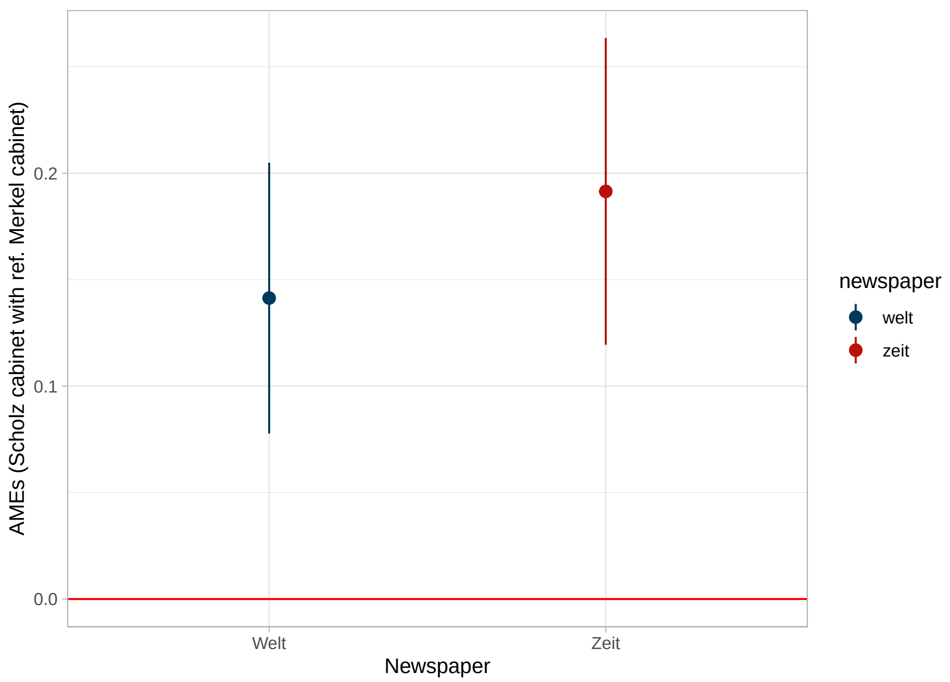

Figure A28: Average marginal effect of Scholz government (negative articles)

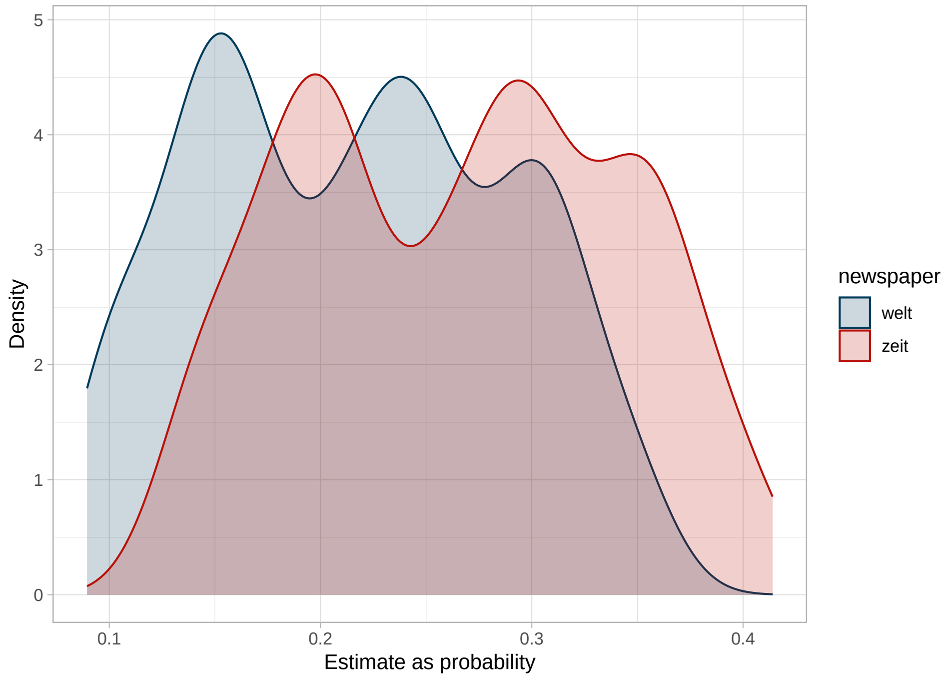

Figure A29: Distribution of estimates with differing control variables (logistic regression on negative articles)

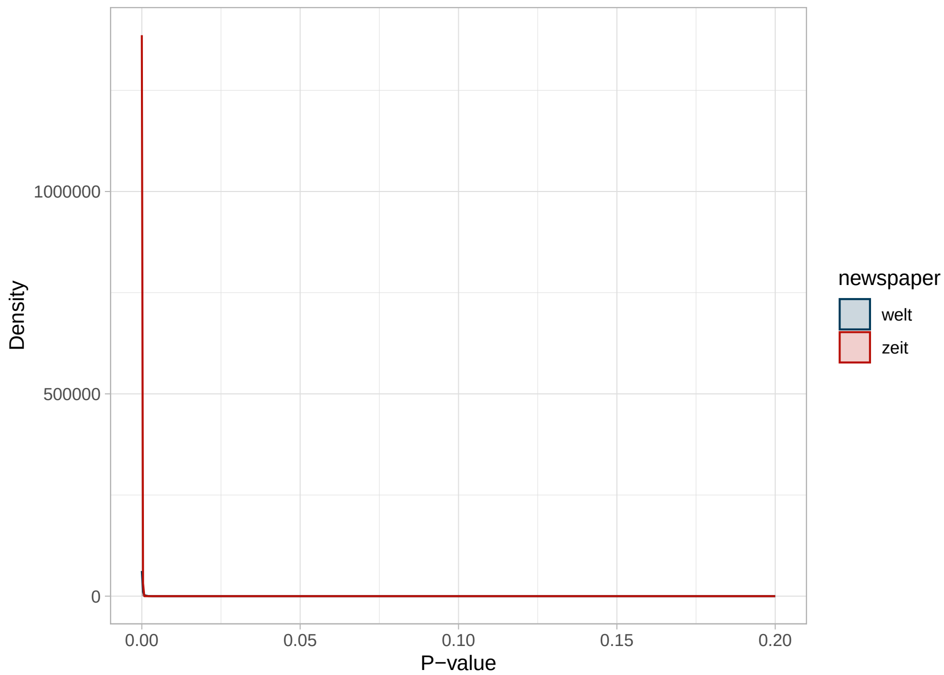

Figure A30: Distribution of p-values with differing control variables (logistic regression on negative articles)

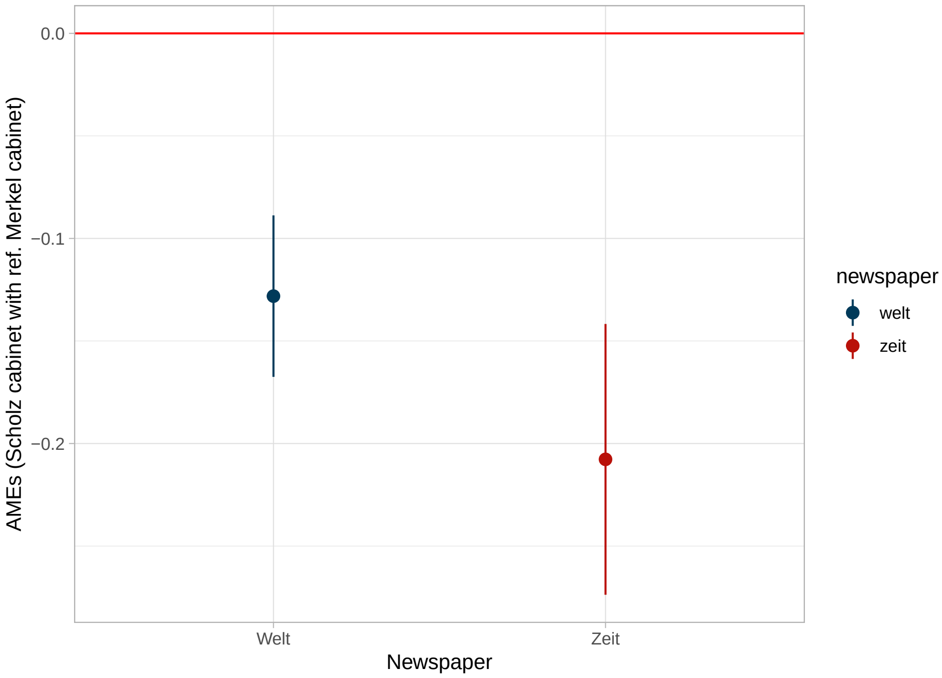

Figure A31: Average marginal effect of Scholz government (positive articles)

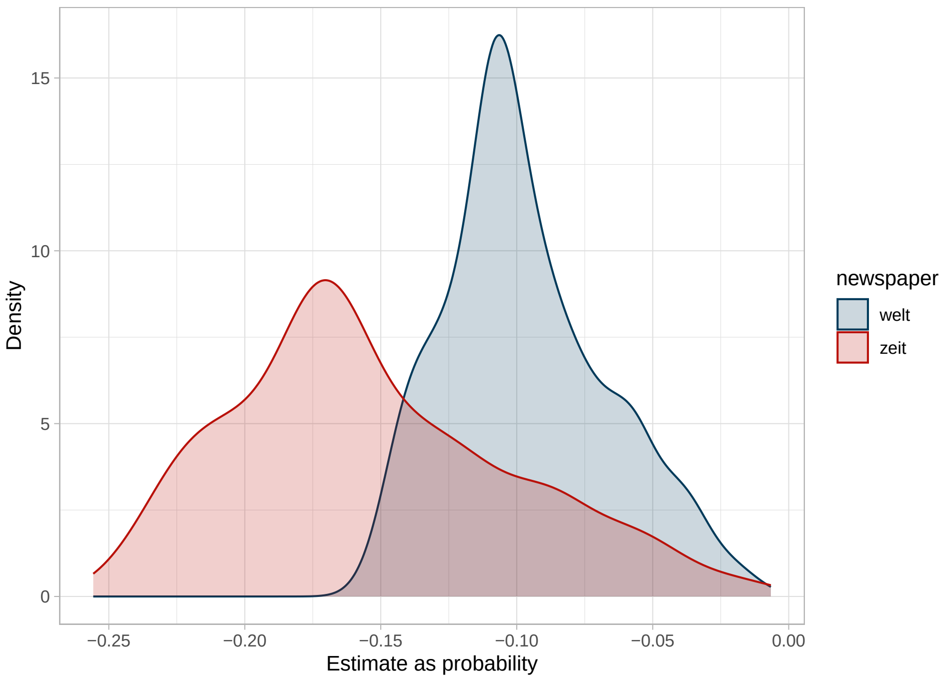

Figure A32: Distribution of estimates with differing control variables (logistic regression on positive articles)

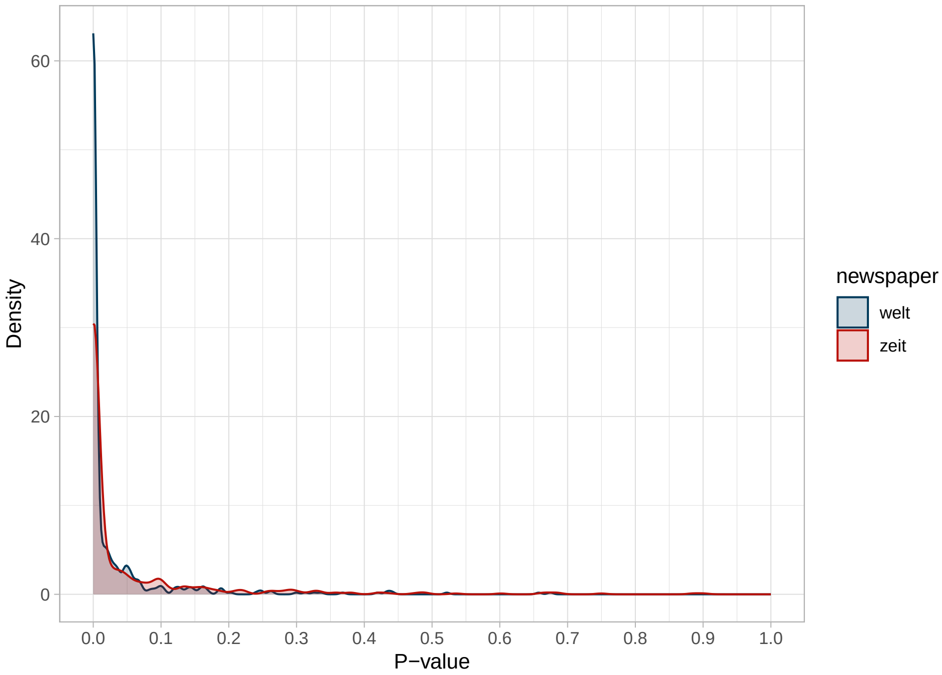

Figure A33: Distribution of p-values with differing control variables (logistic regression on positive articles)

A.8. Quoting patterns

Table A7: Validation of party quotes detection

Party	Precision	Recall	F1
AfD	1.00	1.00	1.00
CDU	0.81	0.81	0.81
FDP	0.80	0.89	0.84
Grüne	1.00	0.80	0.89
Linke	0.78	1.00	0.88
SPD	0.87	0.93	0.90

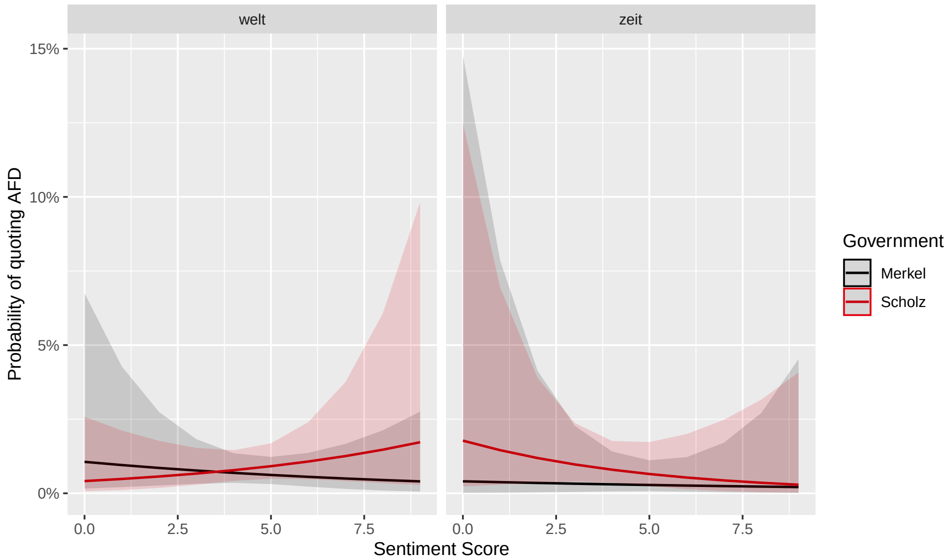

Figure A34: Probability of quoting AfD

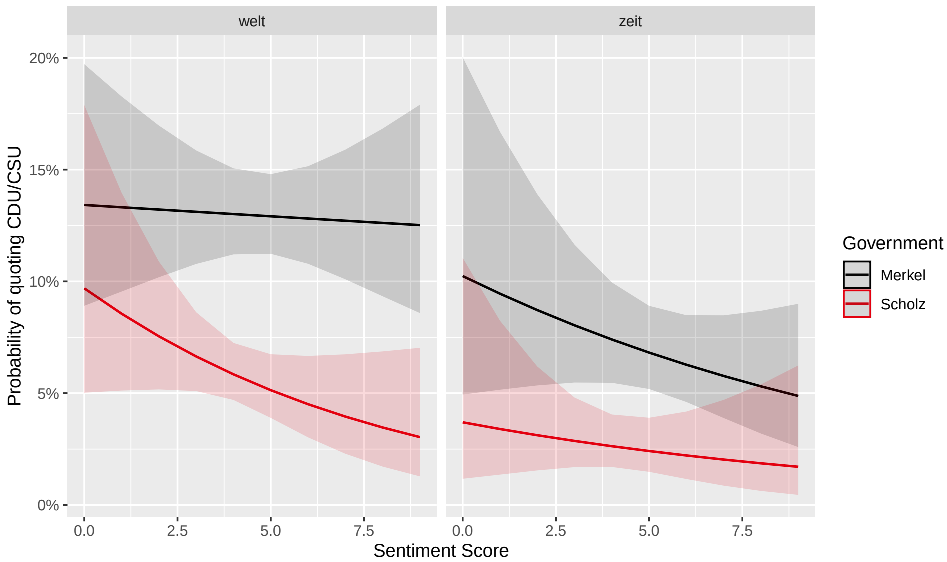

Figure A35: Probability of quoting CDU/CSU

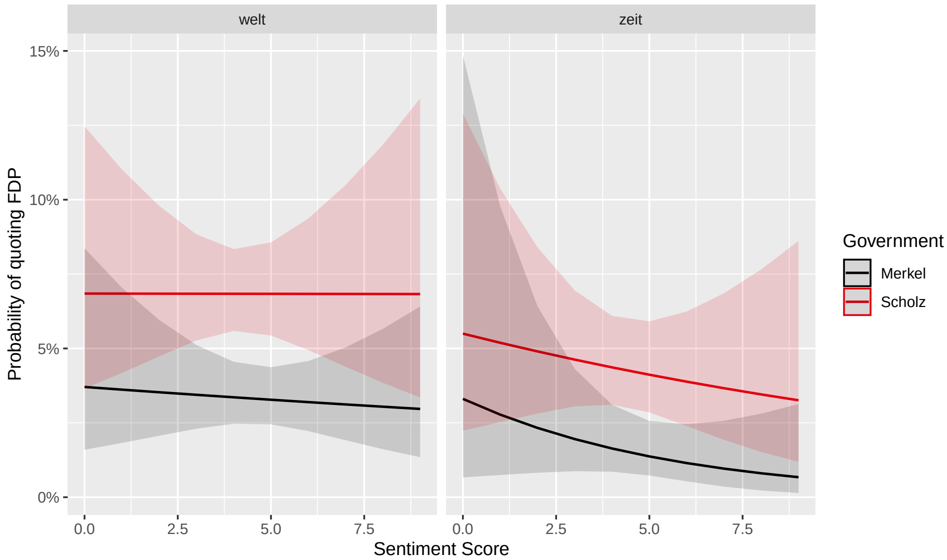

Figure A36: Probability of quoting FDP

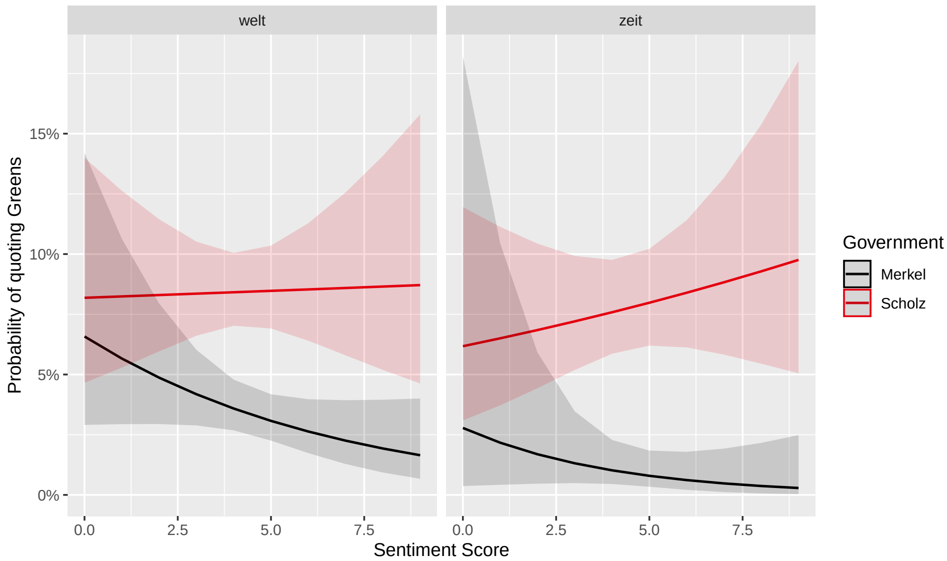

Figure A37: Probability of quoting Greens

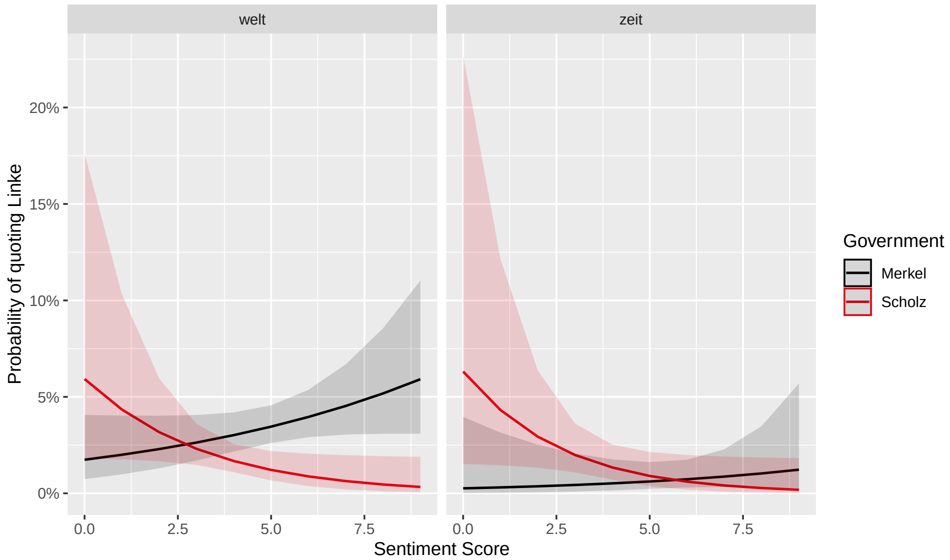

Figure A38: Probability of quoting The Left

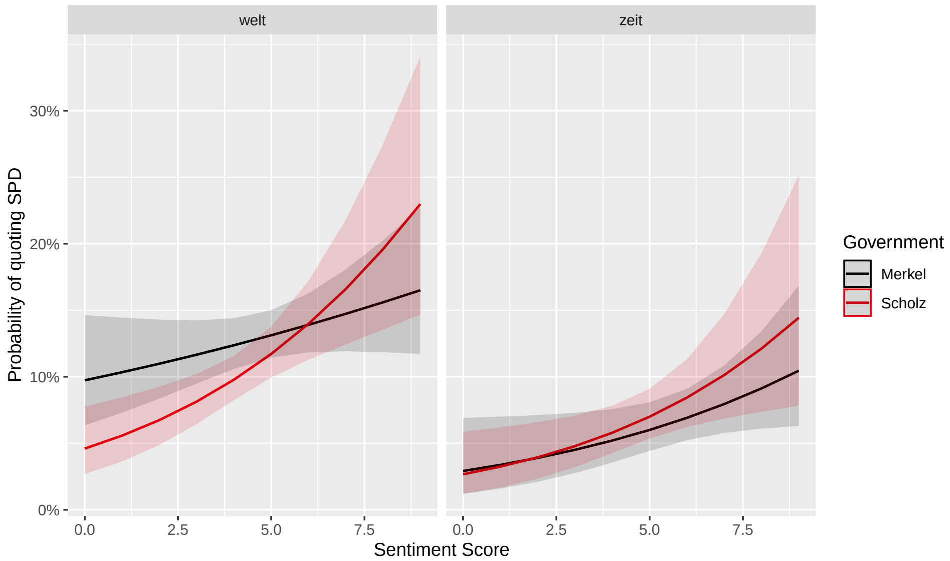

Figure A39: Probability of quoting SPD

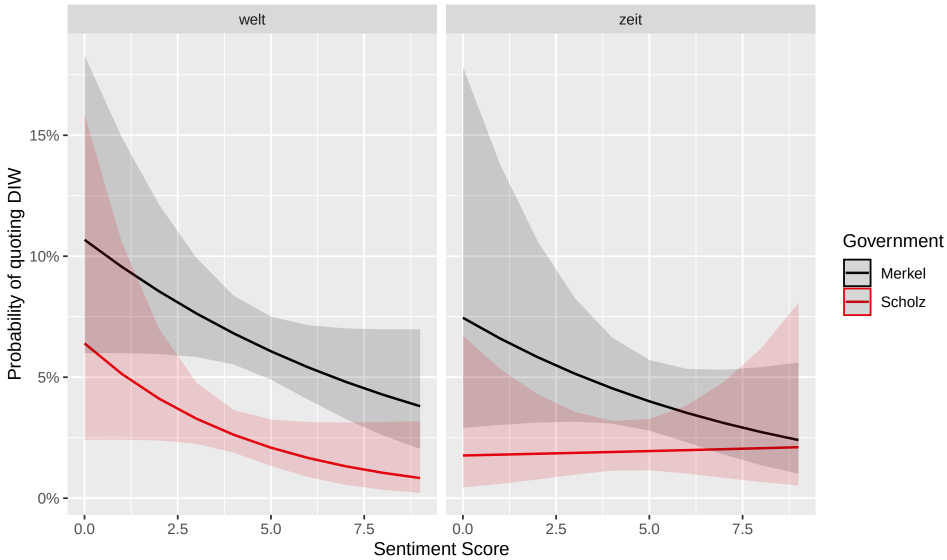

Figure A40: Probability of quoting DIW

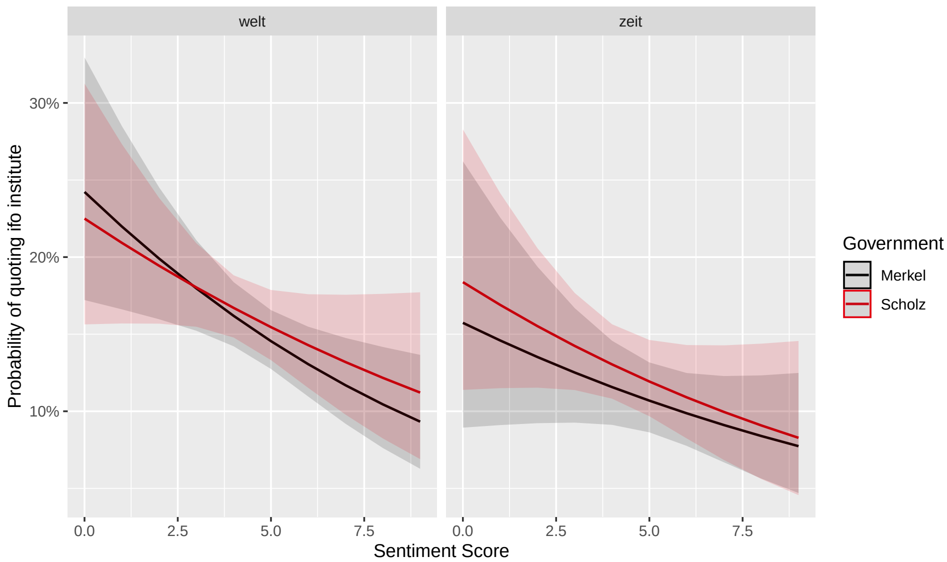

Figure A41: Probability of quoting ifo institute