**Gaining The Competitive Edge**

xxx

Philipp Kläger

SNR: 2062105

March 25, 2022

Master Thesis Marketing Research

Master Program – M.Sc. Marketing Analytics

Marketing Department Tilburg School of Economics and Management

Tilburg University

Readers:

L.K. Deer

B. Deleersnyder



**Management Summary**

xxx

**Table of Contents**

[1. Introduction 1](#_Toc99053197)

[2. Literature Review 4](#_Toc99053198)

[3. German Political System 8](#_Toc99053199)

[3.1 Electoral System 8](#_Toc99053200)

[3.2 Institutional Environment 9](#_Toc99053201)

[3.3 2021 Federal Elections 11](#_Toc99053202)

[4. Model 12](#_Toc99053203)

[4.1 Framework 12](#_Toc99053204)

[4.1.1 Campaign Period 12](#_Toc99053205)

[4.1.2 Winners and Losers 12](#_Toc99053206)

[4.2 Conceptual Model 14](#_Toc99053207)

[4.3 Hypotheses 15](#_Toc99053208)

[5. Data 16](#_Toc99053209)

[5.1 Data Collection and Transformation 16](#_Toc99053210)

[5.1.1 Raw Data Collection 16](#_Toc99053211)

[5.1.2 Data Cleaning 16](#_Toc99053212)

[5.1.3 Data Preparation & Variable Operationalization 17](#_Toc99053213)

[5.2 Data Overview and Model-Free Evidence 17](#_Toc99053214)

[6. Methodology 19](#_Toc99053215)

[6.1 Wordscore Procedure 19](#_Toc99053216)

[6.2 Robustness Check of Estimates 23](#_Toc99053217)

[6.3 Measurements for Difference Testing 24](#_Toc99053218)

[7. Results 25](#_Toc99053219)

[7.1 Study 1 – Congruence 25](#_Toc99053220)

[7.2 Study 2 – Topic Shifts 30](#_Toc99053221)

[8. Discussion 33](#_Toc99053222)

[8.1 Summary 33](#_Toc99053223)

[8.2 Marketing Implications 34](#_Toc99053224)

[8.3 Limitations and Future Research 34](#_Toc99053225)

[References 36](#_Toc99053226)

[Appendices 43](#_Toc99053227)

# 1. Introduction

In few areas, the distinction between winning and losing is as essential and decisive as in politics. On the one side, the election winners gain governmental power get to set policies in favor of their electorate and their stakeholders. On the other side, the election losers enter the opposition and have few means to implement policies in favor of their electorate. This divisive nature is also pivotal for marketing spendings. In the immediate context of the election, only the winners are awarded with a profitable return on investment. In the worst case scenario, the losers are left empty-handed regarding their campaign spendings. Hillary Clinton is the most prominent, recent example for this scenario. During her 2016 presidential campaign, Clinton spent 768 million dollars on her marketing activities. She was eventually beat by her advisary Donald Trump – with no involvement in the new government (Moorthy, 2019).

Yet despite this nature, the importance of political marketing expressed in marketing spendings has never been evaluated higher. Political scientists have been documenting a near-exponential rise in marketing spendings over the last 30 years (Hoegg & Lewis, 2011). On the national level, marketing spendings for presidential races in the US have doubled from 2004 to 2016. On the regional level, campaign spendings for House and Senate races have increased by 675% between 1990 and 2016 (Lovett, 2019; Geer, 2012). According to Forbes, in the 2019-2020 election cycle alone, total spendings for political advertising were estimated at 8.5 billion dollars for television, radio and digital media. This estimate exceeds the spendings of the last election cycle by 108% (Homonoff, 2020).

Academic evidence highlights the importance of poltical marketing in winning votes and therefore directly influencing election outcomes. First, political marketing increases the probability that more voters of the party’s target group decide to exercise their voting right. Thus, the political marketing activities shape the partisan composition of the electorate in its favor (Spenkuch & Toniatti, 2018). Second, appearance-based inferences about candidates induce changes in election results. Candidates with an appearance of a high level of overall competence attain more electoral success (Hoegg & Lewis, 2011; Todorov et al., 2005). The perception of candidate appearances are arguably susceptible to changes induced by marketing activites. This line of argument is supported by recent empirical evidence: more consistent ad messages are correlated with increases in online word-of-mouth and voter preference for the candidate (Fossen et al., 2022). Third, xxx. undecided voters

Elections can subsequently be viewed as marketing contests, contended by the efforts of candidates and associated parties (Hoegg & Lewis, 2011).

importance of social media & twitter, media agency etc.

rise of political discourse on Twitter, very US-centric, one paper that links to Europe.

xxx

how parties communicate with the electorate

This thesis directly investigates how dispersion in political positions expressed on Twitter among candidates within a party correlates with a party's electoral success. Then a sentence on the importance of consistent messaging.

Existing studies ...

I extend this work by adopting a rich, state of the art text analytic approach that positions each candidate on a left-right ideological scale.

I use this position to construct a measure of how far each candidate's expressed ideology is away from their party's program.

This measure then feeds into an ... (whatever your statistical tests will be)

Our analysis examines the 2021 German Federal Elections. Based on the constitutional system as parliamentary republic, German elections are very similar to elections in other western countries. There is no empirical evidence that Twitter has a significantly different importance in Germany compared to other western countries. Therefore, our insights can also be applied to political environments of other western countries.

The data collection for this research followed two general steps. First, the official candidate list for the 2021 German Federal Elections is obtained (bpb, 2021a). This candidate list serves as seed for the first data set. Subsequently, the first data set represents the posts of candidate accounts from the six major parties within the campaign period (Dec. 2020 – Sep. 2021). Only verified Twitter accounts were included for the final data set.

Second, reference texts and reference scores were collected for the analysis. On one side, the party manifestos for the six major political parties were collected for the 2017 and 2021 German Federal Elections. On the other side, the 2019 Chapel Hill Expert Survey was obtained. The 2017 party manifestos were utilized as reference texts. The overall ideological score of the expert survey were used as the reference scores and assigned to the respective 2017 party manifestos. The overall ideological score refers to the known left-to-right positioning of the parties according to the expert evaluations. Subsequently, the 2017 party manifestos with their respective scores were utilized to predict the scores for the 2021 party manifestos and 2021 tweets.

We find results that indicate xxx (key results). xxx (study purpose, linking back to winners and losers). xxx (marketing implication).

research question mentioned in the introduction

# 2. Literature Review

Democratic representation can be understood as the preference congruence between voters, candidates and party. However, the relationship between candidates, parties and voters varies depending on the voting system. In countries with plurality voting (e.g., USA), the proportions of the parliament are constituted by the results of direct candidate mandates. Subsequently, marketing campaigns are more candidate-centric and subsequently less alignment with the party positions is required. Conversely, in countries with proportional voting systems (e.g., Germany), the proportions of the parliament are constituted by the results of the proportional party votings. Subsequently, marketing campaigns are more party-centric, causing the electoral success of candidates and parties to be more dependent of each other (Pedrazzani & Segatti, 2022). In line with this, empirical findings suggest that the voter-candidate congruence is best explained at party level in european voting systems (Belchior, 2012). Therefore, the dispersion between candidate and party messaging becomes important.

Advances in text analytics allow us to measure congruence between candidates and parties (see group C in table 1). The wordscore procedure enables the computational text scaling on a left-to-right scale based on word frequencies and expert evaluations on party ideology as initial reference scores (Benoit et al., 2022; Martin & Vanberg, 2007; Laver et al., 2003). Applying a similar approach, Slapin & Proksch (2008) developed the wordfish model that enables the computational text scaling of policy positions for time-series data (Slapin & Proksch, 2008). Even though both techniques are applicable for all kinds of political texts, so far they have been predominantly used for party manifesto data. However, these text analysis techniques could also be used to examine the political discussions on social media. Compared to party manifestos, social media is arguably the more frequent point of display for policy positions by candidates and parties towards voters. Addressing this vacancy, this research intends to contribute to current political marketing by applying the wordscore procedure to score candidate tweets. At this time, there is only one pertinent study that implemented a similar research objective. Using a linguistic similarity approach, Giavazzi et al. (2020) examined the changes in tweet texts of parties in the immediate aftermath of exogeneous events. Their results suggest that following exogeneous events that are perceived as threats, the language in tweet texts on average assimilates over time to the language of the right-wing party AfD (Giavazzi et al., 2020).

Achieving voter-candidate congruence at party level is an important factor for the success of electoral political marketing campaigns (see group A in table 1). More consistent advertising messages are correlated with increases in online word-of-mouth and voter preference (Fossen et al., 2022).

xxx

xxx

xxx

xxx

xxx

xxx

xxx

Two other focal concepts of political marketing campaigns should also be considered.

First, voter behavior, through the decision to vote or not and whom they vote for, is an important component as the equivalent to consumption in political marketing (Moorthy 2019).Voter behavior is generally modelled based on notions first presented by Newman (1999) that proposes five cognitive domains that influence voter behavior: political issues, social imagery, candidate personality, situational contingency, and epistemic value. Political issues refer to the policies the candidate indicates to introduce once elected. Social imagery and candidate personality reflect the usage of stereotypes to induce specific associations (e.g., with archetypes) to establish the desired image in the voter’s mind. Situational contingency utilizes the presentation of hypothetical events to illustrate why the opposing candidate is not fit for the office that he is in contention for. Epistemic value is a concept that is supposed to induce curiosity in choosing a particular candidate. In election campaigns, smaller parties usually utilize epistemic value to compensate for the discrepancy in credibility compared to bigger parties resulting from the lower levels of awareness of and knowledge about smaller parties (Newman, 1999; Newman & Perloff, 2004).

Second, political marketing should be considerate of the disseminating political brand (see group B in table 1). The question whether positive or negative messages are more effective is primarily dependent of the incumbent status of the party. Incumbent government parties use more positive sentiment to make their introduced policies over the last legislature period more palatable to voters. Conversely, incumbent opposition parties tend to use more negative sentiment to criticize the current government and emphasize the difference towards their own proposed policies (Crabtree et al., 2020). The power of using polarity regarding political brands is especially apparent in advertising. Negative advertising by opposing candidates has a stronger negative effect on politicians with strong brands, compared to policians with weaker brands (Fossen et al., 2019).

Recent empirical studies display that appearance-based inferences about candidates induce changes in election results. These inferences specifically include traits like competence and moral. Their impact is partially driven by trait associations at the party brand level (Hoegg & Lewis, 2011; Cwalina & Falkowski, 2016). Favorable advertisement and brand attitudes can be generated with cause-related marketing campaigns (Folse et al., 2014). However, voters evaluate candidate weaknesses with more weight than candidate strengths (Klein & Ahluwalia, 2005). Particularly populists exploit this, as their campaigns are more negatively oriented, contain more personal attacks and fear messages than campaigns of non-populist candidates (Nai, 2018).

This paper intends to contribute to current political marketing research in three dimensions. First, I utilize the wordscore procedure to measure the political ideology of parties and candidates in the 2021 German Federal Elections. Second, the wordscore outputs are used to construct a measure of candidate-party congruence. Finally, I test whether differences in candidate-party congruence among the political parties in Germany are associated with their relative electoral success in 2021 German Federal Elections.

**Table 1. Synopsis of prior work analyzing media and panel study data in political contexts.**

This table presents an overview for the most relevant, previous scientific findings for the purpose of this research. The findings are categorized in three groups to gain an easier understanding for similarities between the presented studies.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Author(s)** | **Area** | **Methods** | **Sample Size** | **Source** | **Key Findings and Contributions** |
| **Group A: Political Advertising** | | | | | |
| Berger & Milkman (2012) | USA | quantitative | 6,956 | Journal of Marketing Research | Ad virality is partially propelled by physiological arousal.  Content that evokes high-arousal emotions is more viral. |
| Berman et al. (2019) | USA | quantitative | 9,5 million | Journal of Marketing Research | During debates, attention is divided and success is found  in succinct and detail-oriented tweets. After debates,  attention is undivided and success is found in abstract  tweets rich in visuals and emotions. |
| Clarke et al. (2021) | USA | qualitative | 4,361 | Journal of Political Marketing | In the 2010 US House and Senate elections, the  Democrats displayed themselves as more ideologically  liberal in areas where the support for their president  Barack Obama was strongest. |
| Nai (2018) | worldwide | quantitative | 195 | Journal of Political Marketing | Populist campaigns are 15% more negative, contain 11%  more personal attacts and 8% more fear messages than  campaigns of non-populist candidates. |
| **Group B: Political Brands** | | | | | |
| Cwalina & Falkowski (2016) | Poland | qualitative | 132 | Journal of Political Marketing | Voters recognize moral and competence in candidates.  Candidates that are perceived as moral and competent  are more liked than those wihout these traits. |
| Hoegg & Lewis (2011) | USA | qualitative | 153 | Journal of Marketing Research | Appearance-based inferences about candidates induce  changes in election results. Their impact is partially  driven by trait associations at the party brand level. |
| Klein & Ahluwalia (2005) | USA | qualitative | 2,485 | Journal of Marketing | Voters evaluate candidate weaknesses with more weight  than candidate strengths. This effect is robust for voters  who dislike the candidate. |
| Slothuus & Bisgaard (2020) | Denmark | quantitative | 2,902 | American Journal of Political Science | If a party switches its policy positions, its partisans  change their policy opinions instantly and substantially.  This effect also applies when the new positioning goes  against the partisans‘ previous beliefs. |
| **Group C: Linguistic Similarity & Wordscore Methods** | | | | | |
| Giavazzi et al. (2020) | Germany | quantitative | 178,271 | NBER Working Paper Series | Following exogeneous events that are perceived as  threats, the language in tweet texts on average  assimilates to the language of the right-wing party AfD. |
| Laver et al. (2003) | England,  Germany, Ireland | quantitative | 22 | American Political Science Review | Development of the wordscore procedure. Reference  texts with known a priori scores can be used to score  new texts based on wordscores. |

# 3. German Political System

The following sections provide a brief introduction structure of the German Political System and a summary of the 2021 electoral race in Germany. These sections provide the necessary background to digest the empirical aspects and marketing implications in later chapters.

## 3.1 Electoral System

The *Bundestag* is the parliament and subsequently major legislative body of the Federal Republic of Germany. Its members are regularly elected every four years by the German people as their representatives through the means of general elections. The German Federal Elections are organized with proportional representation voting. Each eligible voter has two votes. With the first vote, the voter determines which candidate is supposed to directly represent the electoral constituency the voter is registered. The candidate with the highest number of votes wins this direct mandate. With the second vote, the voter determines which party is supposed to represent him or her in the parliament. The results of the second vote are operationalized with hierarchical candidate lists, made by each party respectively. Subsequently, if a party receives 30% of the second votes in the federal elections, this party earns 30% of the seats in the parliament for the upcoming legislative period. Additionally, the first 30% of the people in the party’s candidate list is elected to be part of the parliament. Hence, the second vote is usually regarded as the more impactful vote (bpb, 2021b). Because of the resulting candidate-party dependency regarding election results, the second vote establishes the necessity to aggregate candidate level data to party level to appropriately account for their relationship.

There are a couple of noteworthy peculiarities in the electoral system of Germany. Firstl, if a party wins more direct mandates with the first votes than they technically earned seats with the second votes, the winning direct candidates are still allowed to be members of the new parliament (*Überhangmandate*). In this event, these mandates are balanced with additional mandates to the other parties according to the proportions of the second vote results (*Ausgleichsmandate*). Thus, the additional mandates restore the actual proportions of the second votes, while also accounting for the first vote results. Second, a party only receives their seats in the parliaments if at least one of two conditions are satisfied. They either need to account for at least five percent of the second votes (*Sperrklausel*) or win at least three direct mandates (*Grundmandatsklausel*) (Federal Returning Officer, 2021b; bpb, 2021b). Parties that represent minority populations are explicitly excluded from the necessary regulations to be part of the parliament. In 2021, this rule solely applies to the Southern Schleswig Voters’ Association (SSW), representing the Danish minority in the state Schleswig-Holstein (SSW, 2021).

## 3.2 Institutional Environment

For the 2021 German Federal Elections, 53 parties were deemed eligible for voting by the responsible government institution, the federal returning officer (Federal Returning Officer, 2021a). At the start of this research, six parties were projected to satisfy the necessary conditions of being included in the new parliament. This projection was confirmed by the election results. Subsequently, these six parties will be referred to as the "major" parties in the following and introduced briefly (Forsa 2022). The party introductions are based on the dossiers of the impartial state institution German Federal Agency for Civic Education to exclude potential biases.

**Alternative for Germany (AfD)**

Founded in response to measures to combat the European currency crisis, the Alternative has most notably fought for restrictive immigration policy, a conservative social policy, and an anti-establishment orientation. Now typically identified as a right-wing populist party, the Alternative has been successful in all elections since 2014. The Alternative became the first right-wing party in post-war Germany to earn representatives in each of the 16 state parliaments and the Bundestag (Decker, 2020a; Lewandowsky, 2018; Decker, 2016).

**Christian Democrats (CDU/CSU)**

The Christian Democrats (CDU/CSU) represent a conservatively oriented party of the Christian-bourgeois camp in Germany. Typically associated with anti-socialist orientation, its program is characterized by conservative, liberal and Christian-social positions. As the only major twin party in Germany, the union consists of technically two different parties, as the CDU acts as its representative in 15 of 16 German states. At the same time, the CSU only exists in Bavaria instead of the CDU (Decker, 2021c; Bösch, 2018; Zolleis & Schmid, 2013).

**Liberal Democrats (FDP)**

The Liberal Democrats (FDP) were the only minor party formed at the beginning of the Federal Republic to permanently secure its parliamentary existence aside from the big parties of CDU/CSU and SPD. Occupying a pivotal role in forming coalitions, they have been the favorite partner in coalitions by both CDU/CSU and SPD. With their liberal economic positions and restrictive stance on refugee and European policy, they have an above-average reach among higher-income voters (Decker, 2021a; Treibel, 2018).

**The Greens (Bündnis 90 / Die Grünen)**

The Greens were the most successful newly founded party in the history of the Germany until the emergence of the Alternative (AfD). They were founded in protest against environmental destruction, the use of nuclear energy, and nuclear armament. Since the 1980s, they have established themselves permanently as the fourth force in the party system alongside the CDU/CSU, SPD, and FDP (Decker, 2020b; Probst, 2015; Probst, 2013).

**The Left (Die Linke)**

The Left Party assembles the remnants of the East German Socialist Union Party (*SED*), previously organized in its immediate successor organizations Party of Democratic Socialism (*PDS*) and later also the Labor and Social Justice Party (*WASG*) since the German reunification in 1990. It thus has its roots both in the pro-labor-union environment and the protest against the social policies of the 2000s (Decker, 2021b; Neu, 2018; Pfahl-Traughber, 2013).

**Social Democrats (SPD)**

The Social Democrats (SPD) embody the oldest German party. Giving up its adherence to Marxism with the Godesberg Program in 1959, the SPD began to open up as a party of the people in the late 1950s, becoming the typical representation of its traditional clientele of industrial workers, as well as a large proportion of middle-class voters. Its program can be characterized by the values of freedom, justice, and solidarity (Decker, 2021c; Krell & Woyke, 2015).

## 3.3 2021 Federal Elections

**Figure 1. Poll data for the 2021 German Federal Elections (own rep.; based on Forsa 2022).**

Note: Each monthly estimate represents the last poll of each month. The used estimates are derived from weekly conducted polls by the German opinion research institute Forsa.

As figure 1 displays, three parties in the Conservatives (CDU/CSU), Greens and Social Democrats (SPD) were leading the weekly polls during the last year before the elections. Over the last two months before the election, the Social Democrats managed to attain roughly 10% of additional voter support. Three parties (SPD, Greens and FDP) were able to win votes compared to the previous elections (see figure 2). Particularly in regards to the second votes, the Social Democrats (+2.41 million) and Greens (+2.69 million) were able to gain the support of over two million voters respectively (table 2). In contrast, the Conservatives (-4.14 million) and the Left Party (-2.03 million) lost several million voters. In the aftermath of the elections, the three parties of Social Democrats (SPD), Greens and Liberal Democrats (FDP) formed a government coalition (Forsa, 2022; Federal Returning Officer, 2018; Federal Returning Officer, 2022b). A more elaborative summary of the elections and the campaigning can be found in Appendix C.

**Figure 2. Comparison of results for the 2017 and 2021 German Federal Elections. Percentual shares are based on the second vote proportions (own rep.; based on Federal Returning Officer, 2018; Federal Returning Officer, 2022b).**

# 4. Model

The following sections expounds the framework (chapter 4.1) and conceptual model (4.2) of this research. The framework elaborates the necessary definitions and assumptions for the model construction. Lastly, the hypotheses (chapter 4.3) illustrate the proposed relationships.

## 4.1 Framework

### **4.1.1 Campaign Period**

There is no legal regulation in the German constitution that determines election campaigns' specific beginning. The prevailing opinion is that the determination of the election date can be viewed as the starting point for the election campaigns of parties and candidates (Federal Returning Officer, 2021a). On December 8th, 2020, president Frank-Walter Steinmeier announced that the election would take place on September 26th, 2021 (Federal Returning Officer, 2021b). Hence, this research defines the campaign period as the time period between December 8th, 2020, and September 26th, 2021. This research utilizes the campaign period as the time limits for the data collection process.

### **4.1.2 Winners and Losers**

As table 3 displays, the framework conceptualizes the distinction in three party groups. The group *Winners* refers to parties that won voter support both in regards to the previous election and the campaign period start. Parties that match these criteria can be characterized to have had a universally successful electoral campaign. Conversely, the group *Losers* refers to parties that lost voter support in regards to both reference points. Parties that match these criteria can be characterized to have had a universally unsuccessful electoral campaign. A possible alternative criterion for the distinction between winners and losers is the status of either government or opposition. However, this criterion is not directly applicable to the election results. The strongest parties do not necessarily form a government coalition. This is evident in the results of the 2021 German Federal Elections: the Conservatives (CDU/CSU) represent the second strongest faction in the new parliament, yet the new government coalition consists of the first, third and fourth strongest factions (SPD, Greens & FDP) (Federal Returning Officer, 2022b).

The group *Inconclusive* refers to parties that won voter support regarding one of the reference points, but lost voter support regarding the other reference point. This group is created for the two parties Alternative (AfD) and the Greens, as their campaigns cannot be definitively assessed as either wins or losses. The Alternative (AfD) lost parliament seats compared to the previous elections, but won voter support during the campaign term. Additionally, the Alternative (AfD) was able to win more direct mandates compared to the previous elections. Conversely, the Greens achieved their best electoral result in history with 14.8% by winning 5.9% of second votes compared to the previous elections. However, during the campaign period the Greens lost considerable voter support. During the first months of the campaign period, the Greens were able to gain the lead in the polls, yet they eventually came in as third strongest faction – losing 4.2% of voter support compared to the campaign start (Forsa, 2022; Federal Returning Officer, 2022b).

**Table 2. Categorization of six major political parties in three groups. The categorization is based on their change in voter support compared to the previous federal elections and the campaign period start.**

Campaign Success (Group)

Winners (SPD & FDP) won voter support in comparison to both pre-measures.

Inconclusive (AfD & Greens) won voter support in comparison to only one of two pre-measures.

Losers (CDU/CSU & The Left) lost voter support in comparison to both pre-measures.

Incumbent Status status resulting from the 2017 German Federal Elections

Elective Status status resulting from the 2021 German Federal Elections

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Party** | **Pre-Measures** | | | **Post-Measures** | | **Deltas** | | **Campaign Success** |
|  | *Election I* | *Incumbent Status* | *Campaign Start* | *Election II* | *Elective Status* | *Legislature Period* | *Campaign Period* |  |
| Alternative  (AfD) | 12.6% | opposition | 9.0% | 10.3% | opposition | -2.3% | +1.3% | Inconclusive |
| Conservatives (CDU/CSU) | 33.0% | government | 36.0% | 24.1% | opposition | -8.9% | -11.9% | Losers |
| Liberal  Democrats (FDP) | 10.8% | opposition | 6.0% | 11.5% | government | +0.7% | +5.5% | Winners |
| Alliance 90 / The Greens (Greens) | 8.9% | opposition | 19.0% | 14.8% | government | +5.9% | -4.2% | Inconclusive |
| The Left Party  (Left) | 9.2% | opposition | 7.0% | 4.9% | opposition | -4.3% | -2.1% | Losers |
| Social  Democrats (SPD) | 20.5% | government | 16.0% | 25.7% | government | +5.2% | +9.7% | Winners |

## 4.2 Conceptual Model

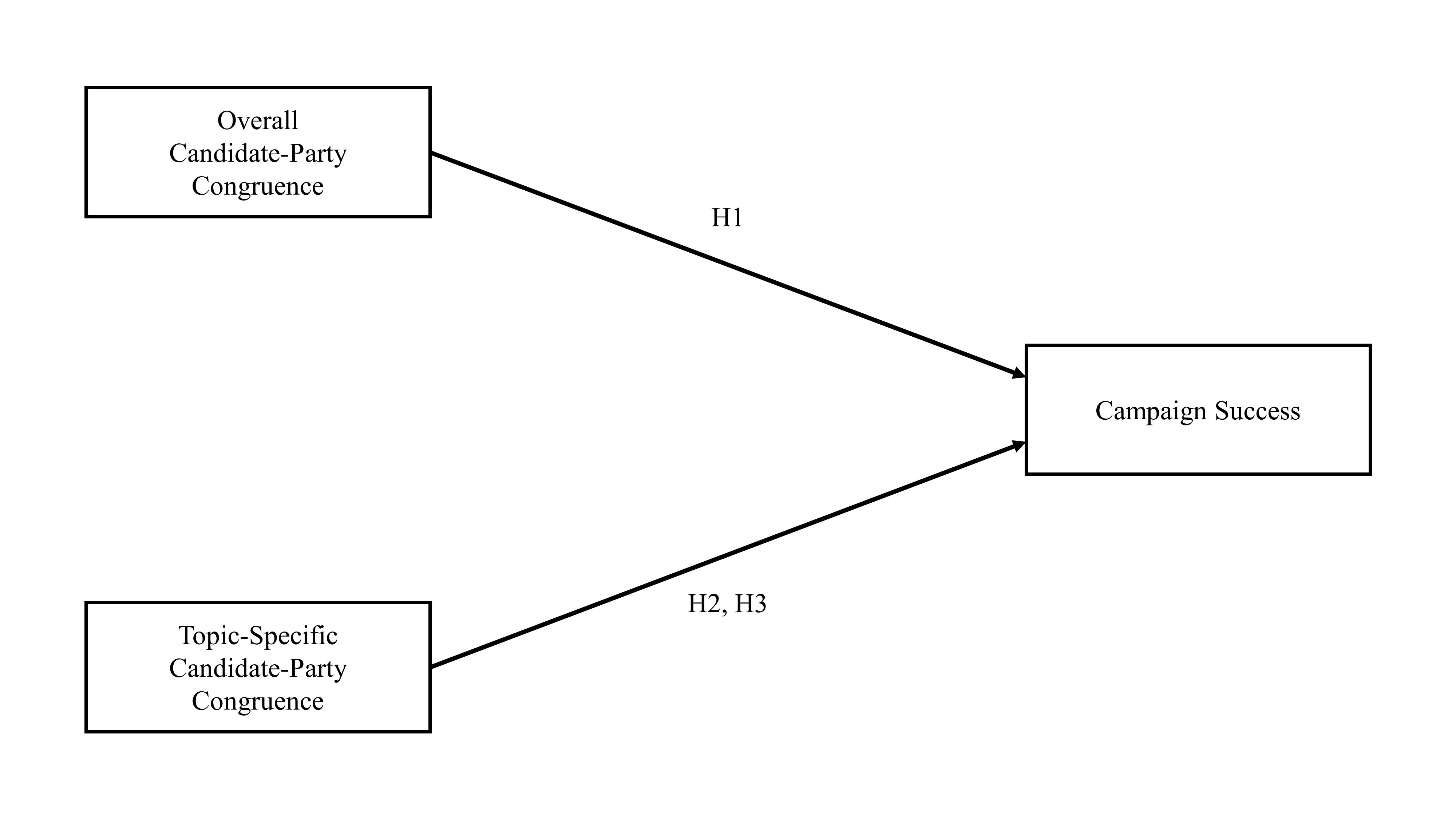
We want to test whether differences in candidate-party congruence among the political parties in Germany are associated with their relative electoral success in 2021 German Federal Elections. More consistent advertising messages are correlated with increases in online word-of-mouth and voter preference (Fossen et al., 2022).

tying back to lit review

why the topics. one lit source

**Figure 3. Conceptual Model**

This conceptual model. The variables of interest are defined below. Their proposed relationship are elaborated in the hypotheses (chapter 4.3).



**Overall Candidate-Party Congruence.**

measures the deviation of candidate tweets within each party from the overall party position based on their left-to-right wing positioning.

**Topic-Specific Candidate-Party Congruence.**

measures the deviation of topic-specific candidate tweets from their overall candidate tweet position.

**Campaign Success.**

categorizes the six major political parties in Germany in three campaign groups based on their electoral campaign results (see chapter 4.1.2). We distinguish between ‘winners’ (SPD & FDP), ‘inconclusive’ (AfD & Greens) and ‘losers’ (CDU/CSU & The Left)

## 4.3 Hypotheses

**Pre-Test.**

*The deviation between the 2021 party and candidate tweet left-to-right positioning differs between parties.*

The intent to test for differences in left-to-right positioning between all parties may appear self-explanatory at first sight. However, before we test for differences between campaign winners and losers and therefore three groups of parties, it is reasonable to first test for differences among parties. If differences among all parties individually cannot be substantiated, it is unlikely the necessary aggregation in campaign groups for our analysis can yield any significant results. Hence, this research deploys this pre-test before the hypothesis testing.

**Study 1. Congruence**

*H1. The deviation between the 2021 party and candidate tweet left-to-right positioning differs between campaign winners and losers.*

We expect a higher level of candidate-party congruence displayed in the candidate tweets for the campaign winners compared to the campaign losers. Therefore, we expect a significantly lower deviation between the 2021 party and candidate tweet left-to-right positioning for the campaign winners.

**Study 2. Topic Shifts**

*H2. The deviation between the overall candidate tweet and corona-specific candidate tweet left-to-right positioning differs between campaign winners and losers.*

*H3. The deviation between the overall candidate tweet and environment-specific candidate tweet left-to-right positioning differs between campaign winners and losers.*

We expect a higher deviation from the overall positioning towards the topic-specific positioning in candidate tweets for the campaign winners compared to the campaign losers.

# 5. Data

In this chapter, a summarizing overview of the utilized data sets of this research shall be presented. To gain an easier understanding, the explanations for Twitter and Party Program Data are displayed separately because they differ from each other considerably.

## 5.1 Data Collection and Transformation

### **5.1.1 Raw Data Collection**

*Twitter Data*

First, the official candidate list for the 2021 German Federal Elections is obtained (bpb, 2021a). This candidate list serves as seed for the first data set. Subsequently, the first data set represents the posts of candidate accounts from the six major parties within the campaign period (Dec. 2020 – Sep. 2021). Only verified Twitter accounts are included for the final data set. The Twitter data is collected using the Twitter Search API and an R download script. In this download script, several data partitions are scraped. Each data partition represents the collected candidate tweets within one party. All data partitions are merged during the cleaning stage. Every scraped data set is exported as a CSV file in UTF-8 encoding.

*Party Program Data*

The texts of the party programs for 2017 and 2021 were collected directly from the official websites of the major political parties as PDF files. Unfortunately, the 2017 SPD program was not available for download anymore. Instead, the text was obtained from the ManifestoR database (Burst et al., 2021). Lastly, the downloaded PDF files were converted into TXT files.

### **5.1.2 Data Cleaning**

*Twitter Data*

First, the data is adjusted by dropping irrelevant columns for this analysis. Only the variables ‘date’ (time of each tweet), ‘screen\_name’ (Twitter handle of each candidate) and ‘text’ (tweet text) are kept. Second, a time adjustment is made by excluding tweets outside the campaign term between December 8th, 2020, and September 26th, 2021. Subsequently, the ‘time’ variable is also dropped. Third, the ‘text’ variable is cleaned by removing unwanted special characters, smilies, numbers, HTML codes, URLs, and excess white spaces using regular expressions and the dplyr manipulation language. Lastly, NA values are handled by replacing them with ‘None’ for character variables and ‘0’ for numeric variables. After the time subset for the campaign period and the cleaning of tweet texts, the final tweet data set contains 218,256 observations.

*Party Program Data*

To make the text data amenable to natural processing tools, punctuations, numbers, symbols, URLs, and German and English stopwords are removed. To avoid the unwanted dropping of German words that contain special characters (e.g., ä, ö, ü, ß), an additional regular expression is implemented.

### **5.1.3 Data Preparation & Variable Operationalization**

First, an id variable for parties is created to identify each tweet's party allegiance quickly. The resulting categorical variable has six levels representing the respective party: ‘AfD’, ‘CDU’, ‘FDP’, ‘Gruene’, ‘Linke’, and ‘SPD’. Second, a party group variable is created. Based on the classification in the model, it displays the levels ‘winners’ (SPD & FDP), ‘losers’ (CDU & Linke) and ‘inconclusive’ (AfD & Gruene). Third, the tweet texts are extracted from the data and converted into documents. This is operationalized by aggregating all tweet texts of each individual candidates into one document. The aggregation is accomplished by the collapsing of rows of each candidate and subsuming the contents of the text column into one object with the delimiter of a space. Subsequently, all tweet texts of one candidate are treated as one document. Analogous to the tweets, each manifesto is saved as a document.

## 5.2 Data Overview and Model-Free Evidence

**Figure 4. Total volume of candidate tweets. Grouped per party (own rep.).**

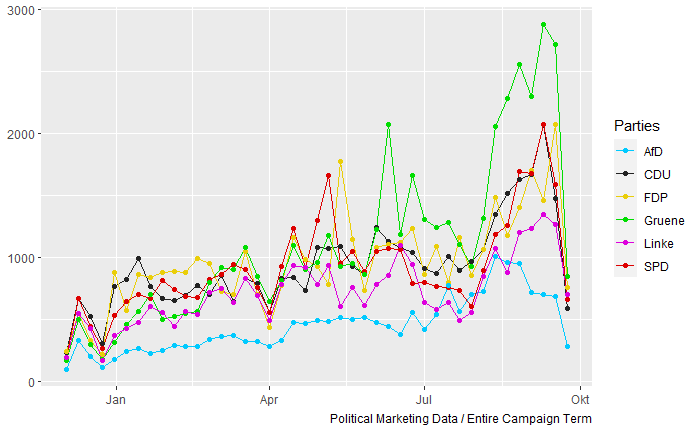
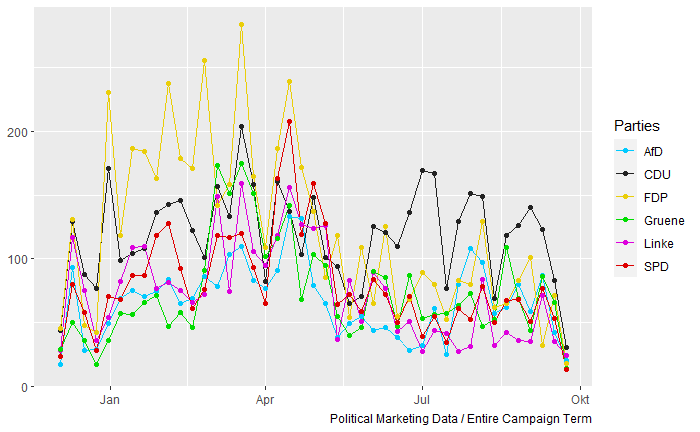
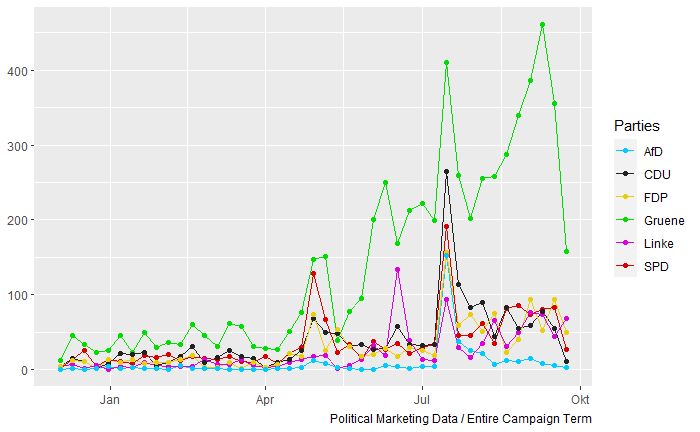


Figure 4 displays the total volume of tweets by each party during the campaign term. The tweet volumes of all parties steadily increased until the election date on September 26th 2021. In the closing months of the campaign period, the Greens released substantially more tweets than their advisories. Figures 5 and 6 display topic-specific tweet volumes for the topics Corona and Environment. They indicate a responsive relationship between tweet volumes and campaign-relevant events. Figure 4 exhibits roughly two different phases, as there is visible low point in tweet volumes of all parties in the end of April 2021. The first phase (Dec. 2020 – Apr. 2021) roughly matches the period of existent, corona-related lockdown measures and restrictions in Germany. The second phase (May 2021 – Sep. 2021) is characterized with considerably reduced corona-specific tweet volumes by all parties. Conversely, the importance of the topic Environment expressed in tweet volumes drastically increased in late April 2021 (figure 5). Additionally, a spike in tweet volumes of environment-specific tweets around the time of the floods in western Europe is visible (July 13th – July 15th 2021).

**Figure 5-6. Comparison of tweet volumes of topic-specific tweets.**

**Corona (upper) vs. Environment tweets (lower) (own rep).**



**

# 6. Methodology

## 6.1 Wordscore Procedure

A key component of this thesis is to measure political candidates ideology as expressed on

Twitter, and compare them to their political party's ideology. To implement this, I utilize a text analytic procedure, known as wordscores. The following sections outline the wordscore procedure and how it is applied in this thesis.

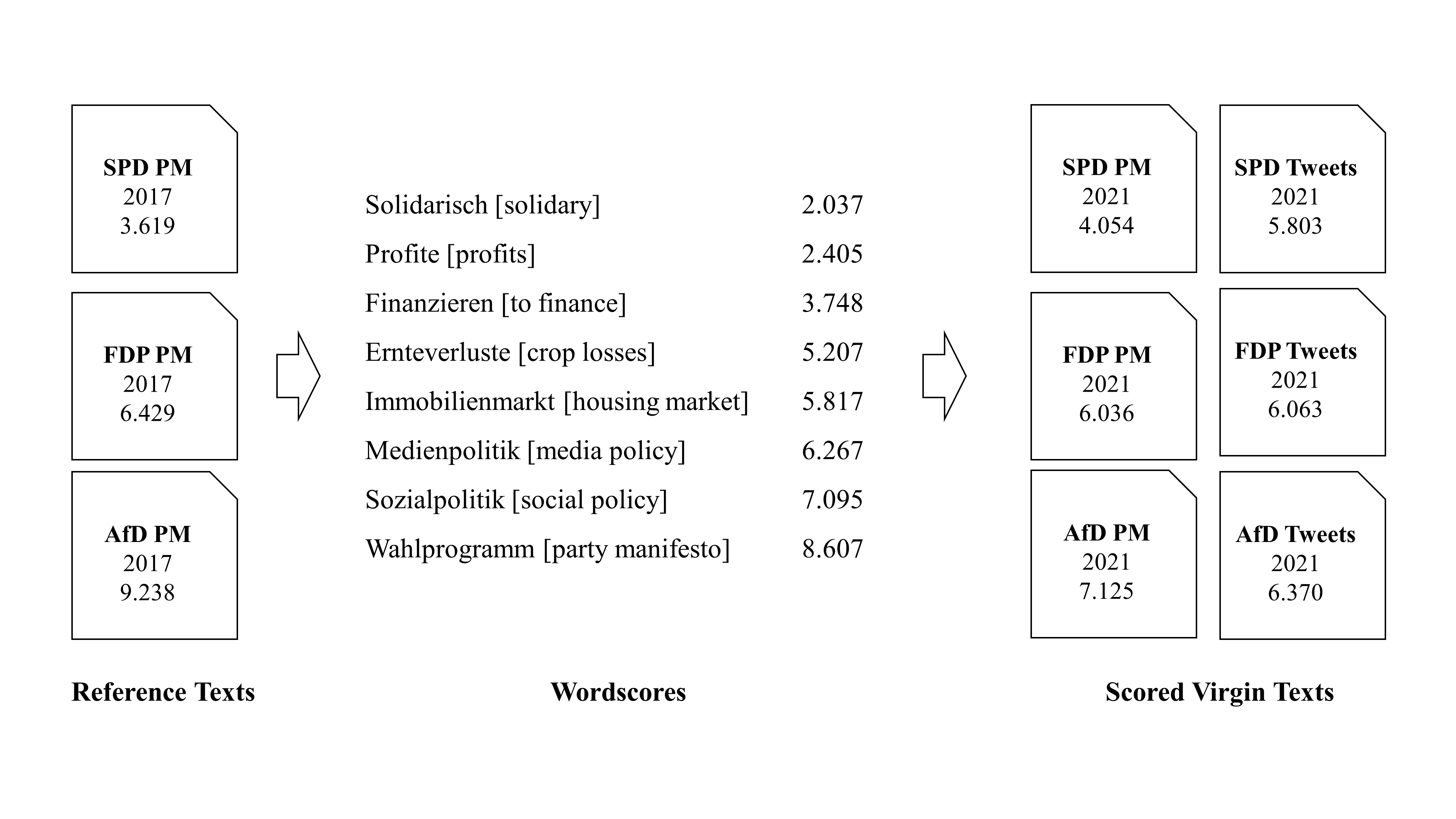
The wordscore procedure was first introduced by Laver et al. (2003) and is designed to map words and texts based on their ideological position. To construct this mapping, the procedure first needs to learn how individual words correpond to ideological positions. First, I use political manifestos from 2017 German Federal Elections for each party that contain their political ideology. These ideologies are then mapped to a left-to-right positioning index using expert evaluations from the 2019 Chapel Hill Expert Survey. Thus, each manifesto is scored in its ideological position. Second, scores for each unique word and subsequently each of the manifesto texts are computed. Third, with this mapping between words and positions now learned by the statistical model, I can use it to predict the ideological positions of new, as yet unscored texts. These first three steps are analoguous to the original procedure proposed by Laver et al. (2003). However, because our virgin texts are tweets rather than manifestos, an additional rescaling step is needed to account for the inaccuracies induced by the change in platform (manifestos vs tweets). I adopt two rescaling procedures proposed by Laver et al. (2003) and Martin & Vanberg (2007) who argue this transformation leads to more robust predictions. These four steps are shown in table 3, and the following sections provide further details on the implementation of each.

**Table 3. Overview of the Wordscore Procedure (own rep.)**

|  |  |  |
| --- | --- | --- |
|  | **Methodological Procedure** | **Application of this Research** |
| Step 1. | A Priori Model Calibration   * collecting reference texts * assigning reference scores to reference texts * collecting virgin texts | * 2017 party manifestos * 2019 Chapel Hill Expert Survey * 2021 party manifestos and 2021 campaign period tweets |
| Step 2. | Computing Wordscores & Reference Textscores |  |
| Step 3. | Computing Virgin Textscores |  |
| Step 4. | Robust Transformation of Scores |  |

**Figure 7. Exemplary excerpt of the application of the wordscore procedure in this research (own rep.)**

Note: PM refers to party manifesto



The main computational operation in this thesis is the application of the wordscore procedure. To illustrate the conceptual idea of wordscores, an exemplary excerpt of the procedure in its general steps is displayed in figure 7.

The wordscore procedure estimates policy positions by comparing two sets of political texts with scores in a numerical positional scale. The dimensions of the positional scores need to be well-delineated and defined before the start of the analysis. Like Laver et al. (2003), this research utilizes a positional scale that indicates left-wing to right-wing positioning. On the left-hand side, reference texts refer to a set of texts whose positions can be assigned a priori based on reliable independent sources or expert evaluations. On the right-hand side, virgin texts refer to a set of texts whose scores are not yet known and to be examined in the analysis. It is important to distinguish between the terms wordscores and textscores. Wordscores refer to the expected policy position of any text, given that we only look at a particular word. Textscores refer to the expected policy position of texts based on the computation of wordscores (Laver et al., 2003).

**Selection of Reference Texts & Reference Textscores**

The selection of an appropriate set of reference texts is essential for the validity of the wordscore procedure. Laver et al. (2003) propose a guideline in choosing suitable reference texts. The “hard-and-fast rule” requires reliable estimates of, or assumptions about, the respective positions in the examined policy dimensions. In Germany, the proposed policies of each party are regularly expressed in party manifestos approaching the elections. The 2019 Chapel Hill Expert Survey contains left-wing-to-right-wing evaluations of party manifestos (1999-2019) of 277 European parties on political ideology. The 2019 Chapel Hill Expert Survey is chosen because it is the most comprehensive expert survey on political parties in Europe. The 2019 Chapel Hill Expert Survey operationalizes with a scale of 0 to 10. In this scale, the score 0 represents holistically left, and the score 10 represents holistically right (Jolly et al., 2019). As displayed in table 4, the left-wing-to-right-wing scores of the 2019 Chapel Hill Expert Survey are applied as reference scores for the party manifestos. These party manifestos are utilized as reference texts.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Table 4. Reference Scores**  Scores are derived from the overall ideological score of the 2019 Chapel Hill Expert Survey In this scale, the score 0 represents holistically left, and the score 10 represents holistically right. Subsequently, all scores below 5 are more left-wing oriented and conversely all scores higher than 5 are more right-wing oriented (own rep.; based on Jolly et al., 2019). | | | | | |
| **The Left** | **The Greens** | **SPD** | **FDP** | **CDU/CSU** | **AfD** |
| 1.429 | 3.238 | 3.619 | 6.429 | 6.524 | 9.238 |

This research conducts the selection of reference texts and reference textscores in R. This approach requires the usage of corpora and document-feature matrices.

A corpus represents the storage library for the UTF-8 encoded text derived from the original documents. It is the data source for a document-feature matrix (DFM), in which the documents represent the rows and the features (e.g., words, syntactic dependencies etc.) are displayed by the columns. Subsequently, each row displays one unit of analysis, and each column represents one analysis variable. The texts in document-feature matrices must be prepared by tokenization, lower-casing texts, and removing punctuation and stopwords. Tokenization refers to the extraction of singular, unique words as ‘tokens’ (Benoit et al., 2022).

**Generating Wordscores from Reference Texts**

The set of reference texts is represented by in this research. The textscores express the estimated a priori position of each reference text on the scoring dimension . First, we observe the relative frequencies of tokens in the reference texts. This variable expresses the proportion of the total number of occurrences for each token to the total number of words in texts. Computed for all reference texts, the results are stored in a matrix of relative word frequencies. Second, we compute the conditional probability matrix based on the matrix of relative word frequencies. For each peculiarity of and , this matrix expresses the probability that we are reading reference text , given the occurrence of word (Laver et al., 2003):

(1)

Given this probability matrix, the wordscores for the reference texts can be generated. With this measure at hand, we attain a left-to-right score for each unique word. The wordscores are computed as the mean of textscores , weighted by the probabilities :

(2)

**Scoring Virgin Texts with Textscores**

With the individual words now scored using the party manifestos, the next step is to predict the ideological positions of unseen texts. These unseen texts are called virgin texts. The set of virgin texts is represented by in this research. Analogous to the wordscores of reference texts, we first observe the relative frequencies of tokens in the virgin texts. Second, the textscores are computed as the mean dimension score of all the scored words , weighted by the relative frequencies of the scored words (Laver et al., 2003):

(3)

This score infers the expected position of the virgin text on the a priori dimension of analysis. In this research, the a priori dimension of analysis is the left-to-right positioning. This inference operates with the necessary assumption that the relative frequencies of words are related to policy positions uniformly in reference and virgin texts. This assumption only holds, if the reference texts are chosen appropriately.

## 6.2 Robustness Check of Estimates

**Uncertainty of Estimates**

The wordscore procedure calculates wordscores and textscores as precise point estimates. Taking the uncertainty of point estimates into account, a variance and standard error for each estimate can be calculated additionally. The variance expresses the deviation of each wordscore from the virgin textscore, weighted by the frequency of the given word in the virgin text:

(4)

Given this variance, a standard deviation for each estimate can be computed with the square root of the variance. The term refers to the total number of scored virgin words:

(5)

**Robustness Check 1 – LBG Transformation (Laver et al., 2003)**

Words that occur in all or most of reference texts tend to be scored near the mean overall score of all reference texts. Subsequently, raw virgin textscores are often situated more clustered than the reference textscores and the dispersion of the virgin text scores is scored on a smaller scale. To adjust for this, we need to convert the scores of the virgin texts to attain the same dispersion metric as the reference texts.

(6)

refers to the average textscore of the virgin texts. and are sample standard deviations of reference and virgin textscores respectively. This transformation maintains the relative position of the virgin textscores and adjusts their variance to the reference text variance. Subsequently, reference textscores and virgin textscores are directly comparable (Laver et al., 2003).

**Robustness Check 2 – MV Transformation (Martin & Vanberg, 2007)**

Instead of using standard deviations, the authors rescale by utilizing the extreme points of reference scores (, ) and raw textscores (, ). Given any raw textscore , the robust textscore is:

(7)

Comparing both approaches, the MV transformation has a key advantage. The LBG transformation utilizes standard deviations to rescale the raw textscores of virgin texts. This approach is highly sensitive to the selection of virgin texts. However, the MV transformation rescales all raw textscores – both reference and virgin texts – according to the original reference scores. This approach is insensitive to the selection of virgin texts and produces more reliable estimates than the LBG transformation. Hence, we use the MV transformation (Martin & Vanberg, 2007).

## 6.3 Measurements for Difference Testing

We construct a measure of difference for the candidate tweets. It captures the candidate-party congruence as the deviation of robust candidate tweet textscores from the measure for overall party positioning for each candidate in party :

(8)

With this measure at hand, we implement an ANOVA test to formally examine differences. An ANOVA test requires roughly normally distributed, independent samples with homogeneity of variances. Thus, we implement a Levene Test. It operates with the null hypothesis that the variances can be considered equal. If the null hypothesis is rejected, the standard ANOVA results are not directly applicable. We can instead use the Welch ANOVA, which provides robust results, if variance homogeneity is violated. Because ANOVA tests are omnibus procedures, they only identify general differences between groups – not between which particular groups. Hence, we deploy the Tukey post-hoc test, which precisely determines the nature of differences between specific groups (Montgomery, 2013).

# 7. Results

## 7.1 Study 1 – Congruence

The goal of study 1 is to compute the candidate-party congruence displayed in the tweet texts and examine it for differences between campaign winners and losers. To implement this, the wordscores and textscores for all manifesto and candidate tweet documents are computed. Then, the raw textscores are rescaled using the MV transformation to attain robust point estimates for all documents.

**Table 5. Comparison of 2021 Party Manifesto and Average Tweet Positioning.**

With the usage of the wordscore procedure, the raw textscore estimates for the 2021 party manifesto and candidate tweets are computed. To operationalize this, the aggregated tweet texts by each singular candidate are treated as one document. Therefore, each candidate attains one textscore. For the final estimates, each raw textscore is rescaled according to the MV transformation (Martin & Vanberg, 2007). Lastly, the overall party position is computed as the unweighted average of all candidate textscores within each party. The delta expresses the difference between the MV estimates of the 2021 manifesto and tweet textscores.

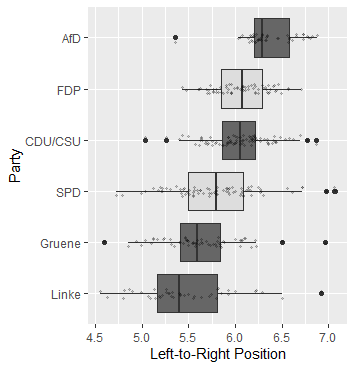
|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Party | Reference Texts | | Virgin Texts | |  |
|  | 2017 party manifesto  reference scores | 2017 party manifesto  MV estimates  [standard errors] | 2021 party manifesto  MV estimates  [standard errors] | 2021 campaign tweets  MV estimates  [standard errors] | Delta |
| The Left | 1.429 | 1.429  [0.030] | 2.811  [0.030] | 5.472  [0.065] | +2.611 |
| The Greens | 3.238 | 3.148  [0.022] | 4.137  [0.023] | 5.626  [0.056] | +1.489 |
| Social Democrats (SPD) | 3.619 | 3.915  [0.028] | 4.054  [0.039] | 5.803  [0.098] | +1.749 |
| Liberal Democrats (FDP) | 6.429 | 6.535  [0.032] | 6.036  [0.033] | 6.063  [0.068] | +0.027 |
| Conservatives (CDU/CSU) | 6.524 | 5.797  [0.043] | 5.086  [0.028] | 6.044  [0.104] | +0.958 |
| Alternative (AfD) | 9.238 | 9.238  [0.067] | 7.125  [0.048] | 6.370  [0.117] | -0.755 |

Note: MV is an abbreviation for the authors Martin & Vanberg.

Table 5 displays the robust textscores for the 2021 party manifestos and the unweighted average of all candidate tweet textscores within each party. We can observe multiple interesting peculiarities here. As expected, the virgin texts are scored more clustered than the reference texts. Additionally, the Conservatives (CDU/CSU) and Liberal Democrats (FDP) have a switched relative positioning in the virgin texts. Whereas the Conservatives are scored as the second most right-wing positioned party based on their 2017 party manifesto, the Liberal Democrats occupy this role for both 2021 party manifestos and the campaign tweets. Comparing the MV estimates of 2021 party manifestos and campaign tweets, it is salient that, except for the Alternative (AfD), all other five parties are scored with a more right-wing oriented textscore. The candidates of the Left Party (+2.611) deviate the most from their 2021 party manifesto positioning, while the Liberal Democrats (+0.027) have the most aligning manifesto and candidate tweet positionings.

**Figure 8. Boxplot of candidate tweet textscores in ascending median order. Grouped by party. Winners (SPD & FDP) are depicted in grey. Other parties are depicted in black.**

The boxplot displays the median value and IQR range for the candidate textscores in each party. The individual candidate textscores are depicted by the grey jitter dots across the boxplot. Outliers in the party distributions are displayed by the round dots outside the boxplot rectangle. We operationalize with a scale of 0 to 10. In this scale, the score 0 represents holistically left, and the score 10 represents holistically right (Jolly et al., 2019). Subsequently, all scores below 5 are more left-wing oriented and conversely all scores higher than 5 are more right-wing oriented.



Color Legend: Winners (grey), other parties (black)

Figure 9 examines the distribution of candidate tweet textscores within each party. There are no visible differences in the sample distributions of the winner parties (SPD & FDP) compared to the other parties. However, based on the range of candidate tweet positionings, the more right-wing oriented parties (AfD, FDP & CDU/CSU) have denser candidate tweet textscore distributions than the more left-wing oriented parties (SPD, Greens & The Left).

**Table 6. Differences between party textscores and candidate tweet textscores.**

For the pre-test, the candidate tweet textscores are grouped based on their party affiliation (six political parties). For the testing of hypothesis 1, the candidate tweet textscores are grouped based on their campaign group affiliation (winners, losers & inconclusive). Analagous to the computation of the estimates in table 6, the raw textscores are rescaled according to the MV transformation to attain robust estimates (Martin & Vanberg, 2007).

Two different measures of spread are utilized in the difference testing. M1 implements the 2021 party manifesto textscore as reference point estimate. Therefore, the measure of spread in the M1 iteration refers to the squared difference between the 2021 party manifesto textscore and the candidate tweet textscore. M2 implements the mean of all candidate textscores within each party as reference point estimate. Therefore, the measure of spread in the M2 iteration refers to the squared difference between the mean candidate textscore per party and the candidate tweet textscore.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **Pre-Test: Party**  Differences between all six parties  (AfD, CDU/CSU, FDP, Greens, Left, SPD) | | **H1: Campaign Success**  Differences between campaign groups  (winners vs. losers vs. inconclusive) | |
|  | M1 | M2 | M1 | M2 |
| **Levene Test** |  |  |  |  |
| df | 5 | 5 | 2 | 2 |
| F value | 30.795 | 4.410 | 13.452 | 0.476 |
| p value | <0.001 (\*\*\*) | <0.001 (\*\*\*) | <0.001 (\*\*\*) | 0.622 |
| **ANOVA** |  |  |  |  |
| df | 5 | 5 | 2 | 2 |
| Sum Sq | 1761.8 | 1.627 | 169.7 | 0.077 |
| Mean Sq | 352.4 | 0.325 | 84.87 | 0.039 |
| F value | 178.4 | 4.872 | 13.83 | 0.549 |
| p value | <0.001 (\*\*\*) | <0.001 (\*\*\*) | <0.001 (\*\*\*) | 0.578 |
| **Welch ANOVA** |  |  |  |  |
| df | 5 | 5 | 2 | 2 |
| F value | 183.44 | 153.82 | 10.503 | 0.603 |
| p value | <0.001 (\*\*\*) | <0.001 (\*\*\*) | <0.001 (\*\*\*) | 0.548 |

A formal test of the above notion utilizes an ANOVA test. The null hypothesis states that the deviation from the overall party positioning to the candidate tweet positioning does not differ among parties (H1) or party groups (H2). Conversely, the alternative hypothesis states that there are significant differences. To implement this test, we need to formalize a measure of spread between party positioning and candidate tweet positionings. One alternative is to utilizing the 2021 party manifesto as reference textscore for the individual candidate tweet textscores. This research refers to this measure of spread as M1. However, comparing manifesto and tweet texts directly can be viewed as biased, as the language in a document for political ideology does not necessarily correspond with the language candidates use to address voters on social media. Hence, this research constructs an alternative measure of spread called M2, that instead uses the mean candidate tweet textscore as the reference point. This way, tweets are directly compared to tweets instead of manifestos. The results of the ANOVA tests are displayed in table 6.

**Pre-Test**

The Levene Test results for candidate-party difference at party level indicate that the variances of candidate tweet textscores are significantly unequal among all six parties (F(5)=30.795, p<0.001). Hence, the assumption of variance homogeneity does not hold and we need to interpret the Welch ANOVA. As expected, we observe significant differences for the candidate-party differences among parties for both measures of spread (M1: F(5)=183.44, p<0.001; M2: F(5)=153.82, p<0.001). This result indicates that the congruence between the party positioning and candidate tweet positioning significantly differs among parties, regardless of the deployed measure of spread. Therefore, we find support for the pre-test.

**Hypothesis 1**

The Levene Test results for candidate-party difference at campaign group level indicate that the variances of candidate tweet textscores are significantly unequal among the three campaign groups (winners, losers & inconclusive) for the M1 measure of spread (F(2)=13.452, p<0.001). However, the assumption of variance homogeneity does hold for the M2 measure of spread (F(2)=0.476, p=0.622). Subsequently, we need to interpret Welch statistic for M1 and can directly interpret the ANOVA results for the M2 measure of spread. As expected, we observe significant differences for the candidate-party differences among campaign groups for the M1 measure of spread (F(2)=10.503, p<0.001). However, these significant differences cannot be substantiated for the M2 measure of spread (F(2)=0.549, p=0.578). These results enable two inferences. First, the differences between 2021 party manifesto and candidate tweet textscores significantly differs among campaign groups. Second, the differences between mean candidate tweet textscore and candidate tweet textscores does not significantly differ among campaign groups. Therefore, we find partial support for hypothesis 1.

Based on the results for the hypothesis 1 variance analysis, we can deploy a Tukey HSD post-hoc test to analyze the size and orientation of the significant differences between 2021 party manifesto and candidate tweet textscores among campaign groups. The results of the Tukey HSD test are presented in table 7.

**Table 7. Tukey post-hoc test results for oriented hypotheses 1 testing based on the previous ANOVA results. M1 measure of spread is deployed.**

Based on the model specifications, the following three campaign groups are applied:

Winners . Social Democrats (SPD) & Liberal Democrats (FDP)

Inconclusive. Alternative (AfD) & Greens

Losers. Conservatives (CDU/CSU) & The Left

|  |  |  |  |
| --- | --- | --- | --- |
|  | Difference | 95% Confidence Interval | p value |
| loser-inconclusive | 1.510 | [0.752; 2.269] | <0.001 |
| winner-inconclusive | 0.197 | [-0.541; 0.935] | 0.806 |
| winner-loser | -1.314 | [-2.018;-0.609] | <0.001 |

We observe two highly significant between-group differences. First, the candidates of the winner parties deviate significantly less (-1.314) from the 2021 party manifesto positioning than the loser parties (p<0.001). Second, the candidates of the loser parties deviate significantly more (+1.510) from the 2021 party manifesto positioning than the inconclusive parties (p<0.001). These results enable two inferences. First, the candidates of winner parties display a significantly higher level of congruence towards the language used in the 2021 party manifestos. Second, the candidates of the loser parties showcase low amounts of congruence towards the language utilized in their 2021 party manifestos.

## 7.2 Study 2 – Topic Shifts

The goal of study 2 is to compute the deviation between the overall candidate tweet textscores and topic-specific candidate tweet scores and then examine whether these deviations significantly differ among campaign groups (winners, losers & inconclusive). We conduct this analysis to examine potential, latent heterogeneity of textscores regarding different topics. To implement this, we first generate the topic-specific candidate tweet textscores for the central election topics Corona and Environment.

**Table 8. Comparison of Average Overall Tweet Positioning and Topic-Specific Tweet Positioning**

**(Corona, Environment).**

Analogous to the computation of table 6, the wordscores and textscores are computed and subsequently rescaled according to according to the MV transformation (Martin & Vanberg, 2007). The overall party position is computed as the unweighted average of all candidate textscores within each party. The deltas express the differences between the overall candidate tweet positioning and the topic-specific candidate tweet positionings for corona and environment respectively.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Party | Overall Tweet Positioning  MV estimates  [standard errors] | Corona Tweet Positioning  MV estimates  [standard errors] | Env. Tweet Positioning  MV estimates  [standard errors] | Corona  Delta | Environment Delta |
| The Left | 5.472  [0.065] | 5.679  [0.543] | 4.639  [0.827] | +0.207 | -0.833 |
| The Greens | 5.626  [0.056] | 5.812  [0.573] | 4.580  [0.549] | +0.186 | -1.046 |
| Social Democrats (SPD) | 5.803  [0.098] | 5.901  [0.558] | 4.810  [0.768] | +0.098 | -0.993 |
| Liberal Democrats (FDP) | 6.063  [0.068] | 5.950  [0.467] | 5.628  [0.775] | -0.113 | -0.435 |
| Conservatives (CDU/CSU) | 6.044  [0.104] | 5.988  [0.459] | 5.049  [0.643] | -0.056 | -0.995 |
| Alternative (AfD) | 6.370  [0.117] | 5.901  [0.464] | 5.729  [0.936] | -0.469 | -0.641 |

Note: MV is an abbreviation for the authors Martin & Vanberg.

Table 8 displays the unweighted average candidate tweet textscores per party, itemized for the overall, corona-specific and environment specific positioning. The deltas express the differences between the overall and topic-specific positioning for each party. We observe that the topic Corona is discussed with very similar word usages of all parties, as the topic-specific positionings are scored much more condensed than the overall tweet positioning. The topic Environment is discussed coherently with more left-wing oriented word usages by all parties. The Greens (-1.046) deviate the most from their overall positioning, whereas the Alternative

(-0.641) deviates the least from their overall positioning.

Figures 7 and 8 displays the distribution of topic-specific candidate tweet textscores within each party. We observe that the median candidate tweet textscore for Corona is almost identical across all parties. A similar observation can be made about the three more left-wing oriented parties (SPD, Greens, Left Party) for the topic Environment.

**Figures 9-10. Boxplots of topic-specific candidate tweet textscores in ascending median order. Grouped by party. Winners (SPD & FDP) are depicted in grey. Other parties are depicted in black.**

The boxplot displays the median value and IQR range for the candidate textscores in each party. The individual candidate textscores are depicted by the grey jitter dots across the boxplot. Outliers in the party distributions are displayed by the round dots outside the boxplot rectangle. We operationalize with a scale of 0 to 10. In this scale, the score 0 represents holistically left, and the score 10 represents holistically right (Jolly et al., 2019). Subsequently, all scores below 5 are more left-wing oriented and conversely all scores higher than 5 are more right-wing oriented.

|  |  |
| --- | --- |
|  |  |
| **Corona Tweets**  Color Legend: Winners (grey), other parties (black) | **Environment Tweets** |

**Table 9. Difference between overall candidate tweet textscores and topic-specific candidate tweet textscores.**

For the testing of hypothesis 2 and 3, the candidate tweet textscores are grouped based on their party group affiliation (winners, losers & inconclusive). Analagous to the computation of the estimates in table 6, the raw textscores are rescaled according to the MV transformation to attain robust estimates (Martin & Vanberg, 2007). The deviations between overall and topic-specific positioning are computed for each candidate.

|  |  |  |
| --- | --- | --- |
|  | **H2: Corona Shifts** | **H3: Environment Shifts** |
| **Levene Test** |  |  |
| df | 2 | 2 |
| F value | 0.281 | 0.290 |
| p value | 0.972 | 0.749 |
| **ANOVA** |  |  |
| df | 2 | 2 |
| Sum Sq | 0.100 | 5.700 |
| Mean Sq | 0.075 | 2.870 |
| F value | 0.040 | 0.355 |
| p value | 0.961 | 0.702 |

For the hypothesis test, we deploy an ANOVA test that applies a similar setup as study 1. The null hypothesis states that the deviation from the overall candidate tweet positioning to the topic-specific candidate tweet positioning does not differ among party groups. Conversely, the alternative hypothesis states that there are significant differences among party groups. The results of the ANOVA test are displayed in table 9.

**Hypothesis 2**

The Levene Test confirms the homogeneity of variances, so the output of the ANOVA is directly interpretable (F(2)=0.281, p=0.972). We do not observe any significant differences in the corona-specific deviations among campaign groups (F(2)=0.040, p=0.961). Therefore, we do not find support for hypothesis 2.

**Hypothesis 3**

The Levene Test confirms the homogeneity of variances, so the output of the ANOVA is directly interpretable (F(2)=0.290, p=0.749). We do not observe any significant differences in the corona-specific deviations among campaign groups (F(2)=0.355, p=0.702). Therefore, we do not find support for hypothesis 3.

# 8. Discussion

## 8.1 Summary

The goal of this research is to identify differences in social media marketing activities between electoral campaign winners and losers in the context of the 2021 German Federal Elections. For this goal, we collected two large sets of data. First, we collect all electoral candidate tweets of the six major political parties in Germany during the entire campaign period. Second, we collect the party manifestos of the six major political parties for the 2017 and 2021 Federal Elections. We deploy advanced text analysis methods for this research. The wordscore procedure allows us to map texts on a left-to-right wing scale based on the occurrence of words in the texts and expert evaluations on political ideology (Laver et al., 2003). We adjust this method with a transformation that converts the raw scores into more robust estimates (Martin & Vanberg, 2007). Using this methodology, we are able to map all collected party manifestos and candidate tweets on a left-to-right wing scale for our analysis. To account for a potentially latent heterogeneity of topics in the tweets, we also examine central election topics (Corona, Environment) individually in addition to the overall tweet positioning.

We observe several interesting results. First, we find support for a correlation between a higher level of candidate-party congruence for the campaign winners compared to the campaign losers. Second, we observe that the topic Corona is discussed with very similar word usages of all parties, as the topic-specific positionings are scored much more condensed than the overall tweet positioning. However, we cannot substantiate support for differences among campaign groups in the corona-specific deviations from the overall positioning. Third, the topic Environment is discussed coherently with more left-wing oriented word usages by all parties. However, similarly to the topic Corona, we cannot substantiate support for differences among campaign groups in the environment-specific deviations from the overall positioning.

Our results present novel implications for academic research and emphasize the need to further study and build robust measures of candidate-party congruence. An interesting outline for this would be to link these measures to actual voter behavior. Recent examples for empirical studies like this focused on the inclusion of voter preferences and survey data (Fossen et al., 2022; Costello et al., 2022).

Our findings have implications beyond the 2021 German Federal Elections. They are applicable to other German federal and state elections, because they have the same political system and at least partially the same voters. Outside of German politics, the importance of candidate-party congruence applies to other countries with a multi-party system and electoral systems with proportional representation such as Austria, Belgium, Denmark, Sweden and the Netherlands.

## 8.2 Marketing Implications

The results of this research present empirical evidence for a correlation between a higher candidate-party congruence in candidate tweets and the campaign winners (SPD & FDP) of the 2021 German Federal Elections. The candidates of campaign winners conveyed the most congruent messaging on social media, thus creating a more consistent public image to potential voters. This is arguably in line with current research, as more consistent advertising messages are correlated with increases in online word-of-mouth and voter preference (Fossen et al., 2022).

Because we utilize the wordscore procedure to attain these results, we can derive multiple conclusions from this. Parties and candidates should closely work together to ensure a congruent messaging on social media. This can be accomplished by utilizing common social media managers and agencies for timely coordination, as well as consistency in style and content dimensions. Additionally, parties can deploy focus group studies in which they identify which slogans or common phrases are best received by their electorate and target groups of potential voters. Using these slogans and common phrases consistently across candidates would also achieve a high level of candidate-party congruence.

## 8.3 Limitations and Future Research

This research utilizes ANOVA tests for its hypothesis testing. Therefore, the research design only enables inferences for correlations as a statistical indicator of the relationship between variables. However, it does not allow to make any inferences for cause-and-effect relationships between variables. In line with this, the omitted variable bias should be considered. There might be other variables that influence campaign success and candidate-party congruence, that are not explicitly modelled or accounted for. Campaign and electoral success depend on voter behavior, yet the influences on this variable are difficult to be directly measured and isolated from other influences. Additionally, the dissemination of information on social media likely does not reflect the entirety of information dissemination in the political discourse, as not every piece of information or opinion is directly shared on social media by voters.

While the wordscore procedure is a powerful tool, its inherent inefficiencies in measurements have also been covered in political marketing research. First, if party manifestos are analyzed with wordscores, the inherent assumption is made that each mentioned topic or policy has the same importance for the parties’ stakeholders. However, this is arguably not true: some policies are more dispensable than others, as the results of coalition negotiations display. Second, words that occur in all or most of reference texts tend to be scored near the mean overall score of all reference texts. Subsequently, raw virgin textscores are often situated more clustered than the reference textscores and the dispersion of the virgin text scores is scored on a smaller scale. However, this inaccuracy is accounted for as best as possible with the utilized transformation methods in this research (Laver et al., 2003; Martin & Vanberg, 2007). Additionally, the degree of condensation differs considerably across countries, party types, and party manifestos that they are based on, which reaffirms the validity of the wordscore procedure (Bräuninger et al., 2013).

Three outlines for future research can be derived from the results of this research. First, the conceptual idea of wordscores can be extended with recent advances in text analysis. A possible approach for this is the concept of topical word embeddings, which takes context and word orders into account. This technique iteratively assigns latent associations for words as topics. An example for this would be association of the topics ‘food’ and ‘information technology’ to the word ‘apple’ due its nature as fruit and brand name for the IT company Apple (Liu et al., 2015; Kusner et al., 2015). Building on this, word embeddings could also be connected to sentiments and emotions (Rudkowsky et al., 2018). Second, an algorithm to filter the tweets for their content could be applied in the data cleaning process. As social media accounts of politicians are not exclusively aligned to electoral campaign activities, some tweets might not be necessarily relevant for analysis, thus causing a bias of estimates. Third, the results of this research can be validated with similar research designs in European countries with similar political and voting systems.

# References

Belchior, A. M. (2012). Explaining Left–Right Party Congruence Across European Party

Systems. *Comparative Political Studies*, *46*(3), 352–386.

https://doi.org/10.1177/0010414012453695

Benoit, K., Watanabe, K., Wang, H., Nulty, P., Obeng, A., Müller, S., Matsuo, A., Lowe, W., &

European Research Council. (2022, March 1). *Quick Start Guide*. Quanteda. Retrieved March 9, 2022, from

https://quanteda.io/articles/quickstart.html

Berger, J., & Milkman, K. L. (2012). What Makes Online Content Viral? *Journal of Marketing*

*Research*, *49*(2), 192–205.

https://doi.org/10.1509/jmr.10.0353

Berman, R., Melumad, S., Humphrey, C., & Meyer, R. (2019). A Tale of Two Twitterspheres:

Political Microblogging During and After the 2016 Primary and Presidential Debates. *Journal of Marketing Research*, *56*(6), 895–917.

https://doi.org/10.1177/0022243719861923

Bräuninger, T., Debus, M. & Müller, J. (2013). Estimating Policy Positions of Political Actors

Across Countries and Time, Mannheim : 2013, Arbeitspapiere – Mannheimer Zentrum

für Europäische Sozialforschung, 153, 1-28.

Bundeszentrale für politische Bildung (bpb) (2021a). Kandidatinnen und

Kandidaten [Candidates]. Bundeszentrale für politische Bildung [Federal Agency for

Civic Education]. Retrieved on September 15, 2021, from https://www.bpb.de/nachschlagen/zahlen-und-fakten/bundestagswahlen/339917/kandidierende

Bundeszentrale für politische Bildung (bpb) (2021b). Bundestag.

Bundeszentrale für politische Bildung [Federal Agency for

Civic Education]. Retrieved on September 15, 2021, from https://www.bpb.de/nachschlagen/lexika/pocket-politik/16362/bundestag

Burst, T., Krause, W., Lehmann, P., Lewandowski, J., Matthieß, T., Merz, N., Regel, S. &

Zehnter, L. (2021). Manifesto Corpus. Version 2021-1. Berlin: WZB Berlin Social Science Center. Retrieved on September 15, 2021, from

https://manifesto-project.wzb.eu/information/documents/corpus

Bösch, F. (2018). *Christlich Demokratische Union Deutschland (CDU)* [Christian Democratic

Union of Germany (CDU)], in: Frank Decker / Viola Neu: Handbuch der deutschen Parteien, Vol. 3, Wiesbaden, 242-261.

Clarke, A. J., Gray, T. R., & Hughes, A. G. (2018). Advertising Ideology: Using Crowd-Sourced

Measurement to Understand Campaigns. *Journal of Political Marketing*, *20*(2), 198–218.

https://doi.org/10.1080/15377857.2018.1486771

Costello, R., Toshkov, D., Bos, B., & Krouwel, A. (2022). Congruence between Voters and

Parties: The Role of Party-Level Issue Salience, p.1-27.

Crabtree, C., Golder, M., Gschwend, T. & Indridason, I.H. (2020). It Is Not Only What You Say,

It Is Also How You Say It: The Strategic Use of Campaign Sentiment. *The Journal of*

*Politics, 82(3)*, 1-17.

Cwalina, W., & Falkowski, A. (2016). Morality and Competence in Shaping the Images of

Political Leaders. *Journal of Political Marketing*, *15*(2–3), 220–239.

https://doi.org/10.1080/15377857.2016.1151121

Decker, F. (2016). *Die “Alternative für Deutschland” aus der vergleichenden Sicht der*

*Parteienforschung* [The "Alternative for Germany" from the Comparative Perspective of Party Research], in: Alexander Häusler, Die Alternative für Deutschland, Wiesbaden, 7-23.

Decker, F. (2020a). *Kurz und bündig: Die AfD* [In a nutshell: The AfD], in: Dossier. Parteien in

Deutschland, Bundeszentrale für politische Bildung. Universität Bonn, Institut für Politische Wissenschaft und Soziologie. Retrieved on October 18, 2021, from

https://www.bpb.de/politik/grundfragen/parteien-in-deutschland/afd/211108/kurz-und-buendig

Decker, F. (2020b). *Kurz und bündig: Die Grünen* [In a nutshell: The Greens], in: Dossier.

Parteien in Deutschland, Bundeszentrale für politische Bildung. Universität Bonn, Institut für Politische Wissenschaft und Soziologie. Retrieved on October 18, 2021, from

https://www.bpb.de/politik/grundfragen/parteien-in-deutschland/gruene/42149/kurz-und-buendig

Decker, F. (2020c). *Kurz und bündig: Die SPD* [In a nutshell: The SPD], in: Dossier. Parteien in

Deutschland, Bundeszentrale für politische Bildung. Universität Bonn, Institut für Politische Wissenschaft und Soziologie. Retrieved on October 18, 2021, from

https://www.bpb.de/politik/grundfragen/parteien-in-deutschland/spd/42080/kurz-und-buendig

Decker, F. (2021a). *Kurz und bündig: Die FDP* [In a nutshell: The FDP], in: Dossier. Parteien in

Deutschland, Bundeszentrale für politische Bildung. Universität Bonn, Institut für Politische Wissenschaft und Soziologie. Retrieved on October 18, 2021, from

https://www.bpb.de/politik/grundfragen/parteien-in-deutschland/fdp/42106/kurz-und-buendig

Decker, F. (2021b). *Kurz und bündig: Die Linke* [In a nutshell: The Left], in: Dossier. Parteien in

Deutschland, Bundeszentrale für politische Bildung. Universität Bonn, Institut für

Politische Wissenschaft und Soziologie. Retrieved on October 18, 2021, from

https://www.bpb.de/politik/grundfragen/parteien-in-deutschland/die-linke/42128/kurz-und-buendig

Decker, F. (2021c). *Kurz und bündig: Die CDU* [In a nutshell: The CDU], in: Dossier. Parteien

in Deutschland, Bundeszentrale für politische Bildung. Universität Bonn, Institut für

Politische Wissenschaft und Soziologie. Retrieved on October 18, 2021, from

https://www.bpb.de/politik/grundfragen/parteien-in-deutschland/cdu/42058/kurz-und-buendig

Federal Returning Officer. (2018). *Results Germany 2017 - The Federal Returning*

*Officer*. Bundeswahlleiter [Federal Returning Officer]. Retrieved March 3, 2022, from

https://www.bundeswahlleiter.de/en/bundestagswahlen/2017/ergebnisse/bund-99.html

Federal Returning Officer. (2021a): Wahlkampfphase [Campaign period]. Bundeswahlleiter

[Federal Returning Officer]. Retrieved October 7, 2021, from

https://www.bundeswahlleiter.de/service/glossar/w/wahlkampfphase.html

Federal Returning Officer. (2021b): Wahltermin zum 20. Deutschen Bundestag ist der 26.

September 2021. [The election date for the 20th German Bundestag is

September 26th, 2021]. Bundeswahlleiter [Federal Returning Officer].

Retrieved October 7, 2021, from

https://www.bundeswahlleiter.de/mitteilungen/bundestagswahlen/2021/20201214-wahltermin.html

Federal Returning Officer. (2022a). *2021 Bundestag Election: greater turnout of*

*young voters - The Federal Returning Officer*. Bundeswahlleiter [Federal Returning Officer]. Retrieved March 3, 2022, from https://www.bundeswahlleiter.de/en/info/presse/mitteilungen/bundestagswahl-2021/01\_22\_ergebnis-rws.html

Federal Returning Officer. (2022b). *Results Germany 2021 - The Federal Returning*

*Officer*. Bundeswahlleiter [Federal Returning Officer]. Retrieved March 3, 2022, from

https://www.bundeswahlleiter.de/en/bundestagswahlen/2021/ergebnisse/bund-99.html

Folse, J. A. G., Grau, S. L., Moulard, J. G., & Pounders, K. (2014). Cause-Related Marketing:

Factors Promoting Campaign Evaluations. *Journal of Current Issues & Research in Advertising*, *35*(1), 50–70. https://doi.org/10.1080/10641734.2014.866847

Forsa. (2022, March 1). Sonntagsfrage – Forsa (Wahlumfragen zur Bundestagswahl) [Sunday

Surveys - Forsa (electoral surveys regarding the German Federal Elections)]. Wahlrecht.de / Forsa Institute. Retrieved March 3, 2022, from https://www.wahlrecht.de/umfragen/forsa.htm

Fossen, B. L., Schweidel, D. A., & Lewis, M. (2019). Examining Brand Strength of Political

Candidates: a Performance Premium Approach. *Customer Needs and Solutions*, *6*(3–4), 63–75. https://doi.org/10.1007/s40547-019-0092-z

Fossen, B. L., Kim, D., Schweidel, D. A., & Thomadsen, R. (2022). The role of slant and

message consistency in political advertising effectiveness: evidence from the 2016

presidential election. *Quantitative Marketing and Economics*.

https://doi.org/10.1007/s11129-021-09246-x

Geer, J. G. (2012). The News Media and the Rise of Negativity in Presidential Campaigns. *PS:*

*Political Science & Politics*, *45*(03), 422–427.

https://doi.org/10.1017/s1049096512000492

Giavazzi, F. (2020, March 9). Terrorist Attacks, Cultural Incidents and the Vote for Radical

Parties: Analyzing Text from Twitter. National Bureau of Economic Research (NBER) Working Paper 26825. Retrieved February 14, 2022, from https://www.nber.org/papers/w26825

Hager, A., & Hilbig, H. (2020). Does Public Opinion Affect Political Speech? *American Journal*

*of Political Science*, *64*(4), 921–937.

https://doi.org/10.1111/ajps.12516

Hoegg, J., & Lewis, M. V. (2011). The Impact of Candidate Appearance and Advertising

Strategies on Election Results. *Journal of Marketing Research*, *48*(5), 895–909. https://doi.org/10.1509/jmkr.48.5.895

Homonoff, H. (2021, December 10). *2020 Political Ad Spending Exploded: Did It Work?*

Forbes.

https://www.forbes.com/sites/howardhomonoff/2020/12/08/2020-political-ad-spending-exploded-did-it-work/

Jolly, S., Bakker, R., Hooghe, L., Marks, G., Polk, J., Rovny, J., Steenbergen, M. &

Vachudova, M.A. (2019). Chapel Hill Expert Survey Trend File, 1999-2019. Electoral Studies. Retrieved February 14, 2022, from

https://www.chesdata.eu/2019-chapel-hill-expert-survey

Klein, J. G., & Ahluwalia, R. (2005). Negativity in the Evaluation of Political Candidates.

*Journal of Marketing*, *69*(1), 131–142.

https://doi.org/10.1509/jmkg.69.1.131.55509

Krell, C. & Woyke, M. (2015). *Die Grundwerte der Sozialdemokratie. Historische Ursprünge*

*und politische Bedeutung* [The fundamental values of social democracy. Historical origins and political significance], in: Christian Krell / Tobias Mörschel, Werte und Politik, Wiesbaden, 93-137.

Kusner, M., Sun, Y., Kolkin, N. & Weinberger, K. (2015). From Word Embeddings To

Document Distances. Proceedings of the 32nd International Conference on Machine Learning, in Proceedings of Machine Learning Research 37, 957-966.

Laver, M., Benoit, K., & Garry, J. (2003). Extracting Policy Positions from Political Texts Using

Words as Data. *American Political Science Review*, *97*(02). https://doi.org/10.1017/s0003055403000698

Lewandowsky, M. (2018). *Alternative für Deutschland (AfD)* [Alternative for Germany (AfD)],

in: Frank Decker / Viola Neu, Handbuch der deutschen Parteien, Vol. 3, 161-170.

Liu, Y., Liu, Z., Chua, T. S., & Sun, N. (2015). *Topical Word Embeddings*. AAAI Conference on

Artificial Intelligence. Twenty-Ninth AAAI Conference on Artificial Intelligence. Retrieved February 14, 2022, from

https://www.aaai.org/ocs/index.php/AAAI/AAAI15/paper/view/9314/0

Lovett, M. J. (2019). Empirical Research on Political Marketing: a Selected Review. *Customer*

*Needs and Solutions*, *6*(3–4), 49–56.

https://doi.org/10.1007/s40547-019-00100-6

Martin, L. W., & Vanberg, G. (2007). A Robust Transformation Procedure for Interpreting

Political Text. *Political Analysis*, *16*(1), 93–100.

https://doi.org/10.1093/pan/mpm010

Montgomery, D.C (2013). *Experiments with a Single Factor: The Analysis of Variance*, in:

Design and Analysis of Experiments. 10th Edition. 55-79.

Moorthy, S. (2019). On Marketing Strategy in Electoral Politics. *Customer Needs and Solutions*,

*6*(3–4), 57–62.

https://doi.org/10.1007/s40547-019-00102-4

Nai, A. (2018). Fear and Loathing in Populist Campaigns? Comparing the Communication Style

of Populists and Non-populists in Elections Worldwide. *Journal of Political Marketing*, *20*(2), 219–250.

https://doi.org/10.1080/15377857.2018.1491439

Neu, V. (2018). *Die Linke* [The Left Party], in: Frank Decker / Viola Neu, Handbuch der

deutschen Parteien, Vol. 22, Wiesbaden, 384-401.

Newman, B.I. (1999): *A predictive model of voter behavior*, in: B.I. Newman, Handbook of

political marketing. Thousand Oaks, CA: Sage.

Newman, B.I. & Perloff, R.M. (2004): *Political Marketing: Theory, Research, and Applications,*

*in: Lawrence Erlbaum Associates*, Handbook of Political Communication Research / edited by Lynda Lee Kaid, Mahwah, New Jersey, London, 17-45.

Pedrazzani, A., & Segatti, P. (2022). Responsiveness When Parties Are “Weak”: A Candidate-

Based Analysis of Voter-Party Congruence in Europe. *Party Politics*, *28*(1), 149–162.

https://doi.org/10.1177/1354068820968091

Pfahl-Traughber, A. (2013). *Die Partei des Demokratischen Sozialismus (PDS), DIE LINKE*

[The Party of Democratic Socialism (PDS), DIE LINKE], in: Oskar Niedermayer, Handbuch Parteienforschung, Wiesbaden, 541-562.

Probst, L. (2013). *Bündnis 90/Die Grünen (GRÜNE)* [Alliance 90/The Greens (GRÜNE)], in:

Oskar Niedermayer, Handbuch Parteienforschung, Wiesbaden, 509-540.

Probst, L. (2015). *Bündnis 90/Die Grünen: Absturz nach dem Höhenflug* [Bündnis 90/Die

Grünen: Crash after soaring high], in: Oskar Niedermayer, Die Parteien nach der

Bundestagswahl 2013, 135-158.

Rudkowsky, E., Haselmayer, M., Wastian, M., Jenny, M., Emrich, S., & Sedlmair, M. (2018).

More than Bags of Words: Sentiment Analysis with Word Embeddings. *Communication Methods and Measures*, *12*(2–3), 140–157.

https://doi.org/10.1080/19312458.2018.1455817

Slapin, J. B., & Proksch, S. O. (2008). A Scaling Model for Estimating Time-Series Party

Positions from Texts. *American Journal of Political Science*, *52*(3), 705–722. https://doi.org/10.1111/j.1540-5907.2008.00338.x

Slothuus, R., & Bisgaard, M. (2020). How Political Parties Shape Public Opinion in the Real

World. *American Journal of Political Science*, *65*(4), 896–911. https://doi.org/10.1111/ajps.12550

Spenkuch, J. L., & Toniatti, D. (2018). Political Advertising and Election Results\*. *The*

*Quarterly Journal of Economics*, *133*(4), 1981–2036.

https://doi.org/10.1093/qje/qjy010

Südschleswigscher Wählerverband (SSW) (2021). Über den SSW [About the

SSW]. Retrieved February 14, 2022, from

https://www.ssw.de/die-partei/ueber-den-ssw

Todorov, A., Mandisodza, A. N., Goren, A., & Hall, C. C. (2005). Inferences of Competence

from Faces Predict Election Outcomes. *Science*, *308*(5728), 1623–1626.

https://doi.org/10.1126/science.1110589

Treibel, J. (2018). *Freie Demokratische Partei (FDP)* [Free Democratic Party (FDP)], in: Frank

Decker / Viola Neu, Handbuch der deutschen Parteien, Vol. 3, Wiesbaden, 319-331.

Trilling, D. (2014). Two Different Debates? Investigating the Relationship Between a Political

Debate on TV and Simultaneous Comments on Twitter. *Social Science Computer Review*, *33*(3), 259–276.

https://doi.org/10.1177/0894439314537886

Zolleis, U. & Schmid, J. (2013). *Die Christlich Demokratische Union Deutschlands (CDU)* [The

Christian Democratic Union of Germany (CDU)], in: Oskar Niedermayer, Handbuch

Parteienforschung, Wiesbaden, 415-437.

# Appendices

**Appendix A. Data**

**Appendix B. Results**

**Appendix C. 2021 Federal Elections**

**GitHub Repository**

<https://github.com/philippklaeger/Gaining-The-Competitive-Edge>

**Data Sets**

**Candidate Tweets Data**

<https://drive.google.com/file/d/1kH7gd1iNSgWOnBj3fBfeY-oKexz_3p_8/view?usp=sharing>

**Party Manifesto Data**

<https://drive.google.com/file/d/1Sv7rNSH3eYUhfOmhQKGBU6MxvwKKz0gl/view?usp=sharing>

**Wordscore / Textscore Data**

<https://drive.google.com/file/d/182zpfeFUFarKkbjkqH90kpCs-iQsNO5u/view?usp=sharing>

**Appendix A. Data**

**Data Set – Tweets**

**Table A1. Summary Statistics for the Candidate Tweet Data.**

218,256 observations

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Party** | **SPD** | **CDU/CSU** | **Greens** | **FDP** | **AfD** | **Left** |
| **Observations** | 39,435 | 40,133 | 47,116 | 41,364 | 19,361 | 30,847 |
| Most Used Words | 1. spd (2987)  2. geht (1841)  3. cdu (1761)  4. menschen (1648)  5. mal (1505)  6. schon (1423)  7. danke (1393)  8. innen (1373)  9. dann (1354)  10. laschet (1250) | 1. deutschland (2166)  2. cdu (1974)  3. corona (1899)  4. geht (1743)  5. menschen (1699)  6. spd (1557)  7. wollen (1423)  8. danke (1351)  9. gibt (1285)  10. mio (1251) | 1. klimaschutz (3617)  2. menschen (2728)  3. geht (2208)  4. cdu (2201)  5. btw (1994)  6. laschet (1985)  7. bundesregierung (1923)  8. wollen (1915)  9. bundestag (1899)  10. innen (1856) | 1. fdp (3225)  2. vielzutun (1924)  3. corona (1852)  4. deutschland (1688)  5. geht (1654)  6. menschen (1638)  7. dann (1551)  8. immer (1514)  9. schon (1488)  10. gibt (1461) | 1. afd (7674)  2. deutschland (1619)  3. bundestag (1564)  4. corona (1301)  5. btw (925)  6. deutschland-abernormal (842)  7. merkel (830)  8. kinderkongress (793)  9. schon (758)  10. cdu (735) | 1. menschen (2137)  2. linke (1838)  3. bundesregierung (1493)  4. cdu (1461)  5. bundestag (1439)  6. btw (1361)  7. geht (1254)  8. corona (1206)  9. deutschland (1173)  10. gibt (1127) |

**Figures A2. Wordclouds of all candidate tweets.**



**Figures A3-A8. Wordclouds of all candidate tweets. Grouped by party.**

|  |  |
| --- | --- |
| Social Democrats (SPD) | Conservatives (CDU/CSU) |
| Greens | Liberal Democrats (FDP) |
| Alternative (AfD) | The Left |

**Table A2. Reference Scores (2019 Chapel Hill Expert Survey)**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **SPD** | **CDU/CSU** | **Greens** | **FDP** | **AfD** | **Left** |
| **Overall Ideological Score**  left vs. right | 3.619 | 6.524 | 3.238 | 6.429 | 9.238 | 1.429 |
| Economic Issues Score  left vs. right | 3.714 | 6.143 | 3.810 | 7.905 | 7.000 | 1.286 |
| Social & Cultural Values Score  libertarian vs. traditional | 3.381 | 6.571 | 1.095 | 3.429 | 9.524 | 2.810 |
| Immigration Policy Score  liberal vs. restrictive | 4.150 | 6.850 | 1.700 | 5.950 | 9.900 | 2.700 |
| Environmental Score  sustainability vs. econ. growth | 4.700 | 6.125 | 1.550 | 7.526 | 8.450 | 4.250 |
| State Spendings Score  public services vs. tax reduction | 3.053 | 6.395 | 3.158 | 8.895 | 7.286 | 1.000 |
| Identity Score  cosmopolitanism vs. nationalism | 3.421 | 6.211 | 1.105 | 3.895 | 9.684 | 2.895 |
| Trade Orientation Score  liberalization vs. protectionism | 4.222 | 3.778 | 3.882 | 1.611 | 6.688 | 6.222 |
| Civil Score  civil liberties vs. law & order | 4.050 | 7.150 | 1.850 | 3.450 | 9.500 | 2.421 |

Reference scores for the wordscore algorithm. The values for each party were taken from the Chapel Hill Expert Survey 2019 (n:21), an expert evaluation of stances based on party programs in 2019, used as reference scores for the 2017 party programs in the prediction of wordscores (left to right; scale 0-10).

**Appendix B. Results**

**Table B1. Left-Wing vs. Right-Wing Words**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **Most Left-Wing Words** | **Wordscore** |  | **Most Right-Wing Words** | **Wordscore** |
| 1. | Linke  [Left] | (1.570) | 1. | AfD  [AfD] | (8.857) |
| 2. | Solidarische  [solidary] | (2.247) | 2. | Wahlprogramm  [electoral manifesto] | (8.386) |
| 3. | Profite  [profits] | (2.497) | 3. | Bundestagswahl  [federal elections] | (8.247) |
| 4. | Beschäftigten  [employees] | (2.510) | 4. | Volk  [people] | (7.543) |
| 5. | Armut  [poverty] | (2.636) | 5. | Willkommenskultur  [welcoming culture] | (7.410) |
| 6. | Löhne  [wages] | (2.692) | 6. | Bundesbank  [federal bank] | (7.191) |
| 7. | Kämpft  [fights] | (2.736) | 7. | Islam  [Islam] | (7.124) |
| 8. | Gewerkschaften  [unions] | (2.742) | 8. | Zuwanderung  [immigration] | (7.113) |
| 9. | Sozial  [social] | (2.763) | 9. | Parteienfinanzierung  [party funding] | (7.063) |
| 10. | Druck  [pressure] | (2.790) | 10. | Schweizer  [Swiss] | (7.016) |

**Appendix C. 2021 Federal Elections**

**Poll History**

**Figure C1. Poll data for the 2021 German Federal Elections (own representation; based on Forsa 2022).**

Note: Each monthly estimate represents the last poll of each month. The used estimates are derived from weekly conducted polls by the German opinion research institute Forsa.

As figure 1 displays, three parties in the Conservatives (CDU/CSU), Greens and Social Democrats (SPD) were leading the weekly polls during the last year before the elections. Despite starting into the year with a considerable surplus, the Conservatives were unable to retain their lead in the elections in late September. The Greens appeared to be their fiercest rival for most of the campaign. However, it was ultimately the Social Democrats rose victorious. Over the last two months before the election, the Social Democrats managed to attain roughly 10% of additional voter support. In the aftermath of the elections, the three parties of Social Democrats (SPD), Greens and Liberal Democrats (FDP) formed a government coalition. Referring to the parties’ colors, this coalition is also known as the “traffic light coalition” (Forsa 2022).

**Election Results**

**Figure C2. Comparison of results for the 2017 and 2021 German Federal Elections. Percentual shares are based on the second vote proportions (own representation; based on Federal Returning Officer, 2018; Federal Returning Officer, 2022b).**

\* Note: The Left Party failed to win at least 5% of the second votes. However, they satisfied the other necessary condition of being included in the parliament by winning three direct mandates.

**Table C2. Comparison of vote counts and seats for the 2017 and 2021 German Federal Elections (own representation; based on Federal Returning Officer, 2018; Federal Returning Officer, 2022b).**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **2021** | | | **Difference to 2017** | | |
|  | **First Votes**  **(Candidate)** | **Second Votes**  **(Party)** | **Seats** | **First Votes**  **(Candidate)** | **Second Votes**  **(Party)** | **Seats** |
| Alternative (AfD) | 4,695,611 | 4,803,902 | 83 | -621,888 | -1.074,213 | -11 |
| Conservatives (CDU/CSU) | 13,239,572 | 11,178,298 | 197 | -4,046,666 | -4,139,046 | -49 |
| Liberal Democrats (FDP) | 4,042,951 | 5,319,952 | 92 | +793,713 | +320,503 | +12 |
| The Left | 2,307,536 | 2,270,906 | 39 | -1,659,101 | -2,026,364 | -30 |
| The Greens | 6,469,081 | 6,852,206 | 118 | +2,751,159 | +2,693,806 | +51 |
| Social Democrats (SPD) | 12,234,690 | 11,955,434 | 206 | +805,459 | +2,416,053 | +53 |

As figure 2 displays, three parties (SPD, Greens and FDP) were able to win votes compared to the previous elections (Forsa 2022). Particularly in regards to the second votes, the Social Democrats (+2.41 million) and Greens (+2.69 million) were able to gain the support of over two million voters respectively (table 2). In contrast, the Conservatives (-4.14 million) and the Left Party (-2.03 million) lost several million voters (Federal Returning Officer, 2018; Federal Returning Officer, 2022b). Additionally, the proportion of small parties that did not meet the entry requirements to be included in the parliament increased (Forsa 2022). Subsequently, for the first time in 16 years, the strongest faction in the parliament is represented by the Social Democrats and not the Conservatives.

**Voter Demographics**

The 2021 German Federal Elections are characterized with a record share of postal voters (47.3%). This characteristic could be attributed to the Covid-19 pandemic and the associated regulations. The voter turnout in 2021 was 76.6%. The age group of 50- to 59-year-olds (80.2%) represents the highest voter turnout per age group (Federal Returning Officer, 2021a). The Greens, Social Democrats (SPD) and Conservatives (CDU/CSU) were elected predominantly by women. Conversely, the Alternative (AfD), Liberal Democrats and the Left Party were preponderately elected by men (Federal Returning Officer, 2021c).

The Conservatives and Social Democrats are both reliant of older voters. The Conservatives realized above-average voter potential among voters aged 70 or older with a share of 38.4%. Only 8.2% of first voters and 10.3% of voters aged 25 to 34 elected the Conservatives. Similarly, the voters of the Social Democrats represent the strongest faction among voters aged 60 to 69 (31.7%), and the second strongest faction among voters aged 70 or older (34.2%). Unlike the Conservatives, the Social Democrats performed considerably better among first voters (15.6%) and voters aged 25 to 34 (17.4%) (Federal Returning Officer, 2021c).

The Greens and Liberal Democrats are particularly successful among young voters. The Greens represent the strongest faction among first voters (24.0%) and voters aged 25 to 34 (22.9%). Appositely, the Liberal Democrats also realized above-average voter potential among first voters (20.5%) and voters aged 25 to 34 (14.9%). Similarly, the Left Party draws its main support from first voters (7.8%) and voters aged 25 to 34 (6.6%). The Alternative (AfD) performed best among the voters aged 25 to 34 (14.9%) and voters aged 35 to 44 (12.4%) (Federal Returning Officer, 2021c).

Regarding all voters, the Alternative (10.3%) represents the second smallest faction in the new parliament. However, the Alternative has noticeably higher voter success in the states of the former German Democratic Republic. In the states Saxony (24.6%) and Thuringia (24.0%), the Alternative had the highest share of second votes of all parties. Except for Berlin, the Alternative represents the second or third strongest faction in all other four eastern states (Federal Returning Officer, 2021c).