Measuring Inter-party Communication:

A transfer learning approach

Anna Adendorf[*], Oke Bahnsen[†], Thomas Gschwend[‡], Lena Maria Huber[§], Simone Paolo Ponzetto[¶], Ines Rehbein[‖] and Lukas F. Stoetzer[**]

November 9, 2023

Abstract

Inter-party communication is crucial in representative democracies, enabling information exchange and dialogue among political parties. Despite its importance, research on this topic remains limited due to a lack of comprehensive conceptualization and challenges in large-scale measurement. This article proposes a holistic definition of inter-party communication as public communication by parties about others with a positive, neutral, or negative stance, focusing on collaboration, policy, or personal issues. To effectively measure inter-party communication, we introduce a novel transfer learning approach capable of automatically classifying large volumes of textual data. Two case studies on coalition signals in Germany and negative campaigning in Austria demonstrate its effectiveness. The study contributes to our understanding of political discourse and the dynamics of party competition. Our approach advances automatic text classification methodologies and opens new avenues for studying political communication.

Keywords: inter-party communication, transfer learning, transformer models, coalition signals, negative campaigning

[*]PhD Student University of Mannheim, E-mail: anna.adendorf@uni-mannheim.de
[†]PhD Student University of Mannheim, E-Mail: oke.bahnsen@mzes.uni-mannheim.de
[‡]Professor University of Mannheim. E-Mail: gschwend@uni-mannheim.de
[§]Post-doc University of Mannheim. E-Mail: lena.huber@uni-mannheim.de
[¶]Professor University of Mannheim. E-Mail: ponzetto@uni-mannheim.de
[‖]Post-Doc University of Mannheim. E-Mail: rehbein@uni-mannheim.de
[**]Professor Witten/Herdecke University. E-Mail: lukas.stoetzer@uni-wh.de

1. Introduction

Inter-party communication refers to the exchange of information, ideas, and arguments between political parties within a democratic system. In a representative democracy, political parties play a crucial role in representing the diverse interests of the electorate and facilitating dialogue and compromise among different groups. Effective inter-party communication is, therefore, essential for ensuring that the political process is responsive to the needs and concerns of the people. Observing communication dynamics between parties and their elites helps the public to monitor and evaluate how well political parties fulfil their representative functions and make informed electoral choices. Public communication with their opponents can also help parties achieve specific electoral and ideological goals.

Despite its importance for representative democracy, there is little research with a focus on inter-party communication. We see two possible explanations for this. First, there is no comprehensive conceptualization of inter-party communication that covers all of its specific aspects. Second, measuring and categorizing communication between political parties on a large scale is quite challenging.

In this article, we propose a comprehensive definition of inter-party communication. We define inter-party communication as public communication by political parties about other political parties, with either a positive, neutral, or negative stance, for the purpose of achieving specific objectives. This perspective is network-centric, in that parties are connected by their inter-party communication signals and the public observes the communication addresses, either directly or through the news media. Furthermore, it is dynamic as the directed and purposeful communication evolves over time. This conceptualization of inter-party communication brings together several aspects that were previously described in isolation, namely communication about collaboration, policy, and personal issues. Two prevalent types of purposeful communication between political parties are *negative campaigning* (see e.g. Skaperdas and Grofman, 1995) and *coalition signals* (see e.g. Meffert and Gschwend, 2011). Negative campaigning is often used in inter-party communication to try to sway voters away from the opposing parties. During campaigns, parties also often send coalition signals, clarifying with whom they do (not) want to govern.

This conceptual definition calls for a powerful measurement approach that is able to grasp the nuances and context of communication between parties. To collect data on inter-party communication, we develop a new measurement strategy, a transfer learning approach, to automatically classify large amounts of textual information as directed party elite communication. A 'transfer learning approach' refers to a machine learning methodology that leverages knowledge and patterns acquired from a large corpus of text to improve the classification of inter-party communication in specific data sources. Previous research relied primarily on hand-coding strategies (e.g., Best, 2015; Golder, 2005; Haselmayer, Meyer and Wagner, 2019) or dictionary approaches (Bowler, McElroy and Müller, 2022; Haselmayer and Jenny, 2017) to classify particular types of inter-party communication. In contrast, our transfer learning approach is resource-efficient and more fine-grained, able to distinguish the subtleties, context, and polarity within sentences, as it is pre-trained on a large corpus of text. This is of practical importance because determining whether "two parties try to form a coalition" has very different substantive implications than identifying that "two parties do *not* try to form a coalition". We provide a detailed outline of the steps required to implement a transfer learning approach for applications to different types of inter-party communication.

We then showcase our approach with two case studies on negative campaigning and coalition signals. In the first case study, we identify positive, neutral, or negative coalition signals in newspaper articles about recent German federal elections. The evaluation of our approach reveals the advantages of our method over existing dictionary and supervised approaches for the detection of coalition signals. Fine-tuning a pre-trained transformer-based language model gives higher recall and precision. A second case study on negative campaigning in Austrian party press releases underlines this finding. Our approach can be used to determine the party target and the stance towards this party in the data. The two case studies illustrate how scholars can apply our approach when interested in identifying targeted party elite communication about a specific subject in textual data.

The paper makes several contributions to existing research on political communication and party competition. Our conceptualization of inter-party communication allows for subsuming different types of communication under a common umbrella, which up to now have been mostly studied in isolation. Elections put parties into competition with each other, which naturally results in the need to publicly negotiate different issues such as coalition formation, policy positions and priorities, and the integrity of party elites. Our proposed concept and measurement approach thereby opens up new avenues for future research. The described method can be used by researchers to categorize inter-party communication from various textual sources. This makes it possible to study new data sources for testing hypotheses about inter-party communication. For example, under what conditions are parties more likely to signal their preferred coalition partners to the public (Golder, 2005)? Are media outlets more inclined to publish negative campaign messages (Haselmayer, Meyer and Wagner, 2019)? This study further advances the field of automatic text classification by demonstrating that fine-tuning transfer-learning models can improve upon classic approaches, like dictionaries, but also supervised machine-learning models (Barberá et al., 2021).

2. Conceptualizing Inter-party Communication

Public communication between various political parties is part of the larger phenomenon of political communication and represents an important characteristic of democracies. Following conventional definitions, we understand political communication as "purposeful communication about politics" (McNair, 2011, 4). Political communication can occur between or among three key actors: political elites (e.g., political parties), the media, and the citizens (e.g. Dumdum and Bankston, 2022; McNair, 2011; Zaller, 1999). It is driven by the goal of political parties and other political elites to gain public support, the interest of the media in maximizing their audience, and the motivation of citizens to hold political parties and other political elites accountable (Zaller, 1999). This article focuses on political communication among political elites, or between various political parties more specifically, what we call *inter-party communication*.

Previous studies have paid surprisingly little attention to inter-party communication. According to a comprehensive review of research articles published in leading communication and political science journals between 2000 and 2017, only 7.06 per cent of the articles examined communication among political elites (Dumdum and Bankston, 2022). Naturally, an even smaller proportion of articles focuses on inter-party communication. The scarcity of empirical research in this field is also reflected by the lack of conceptualizations of inter-party communication (for an exception see De Nooy and Kleinnijenhuis, 2013).

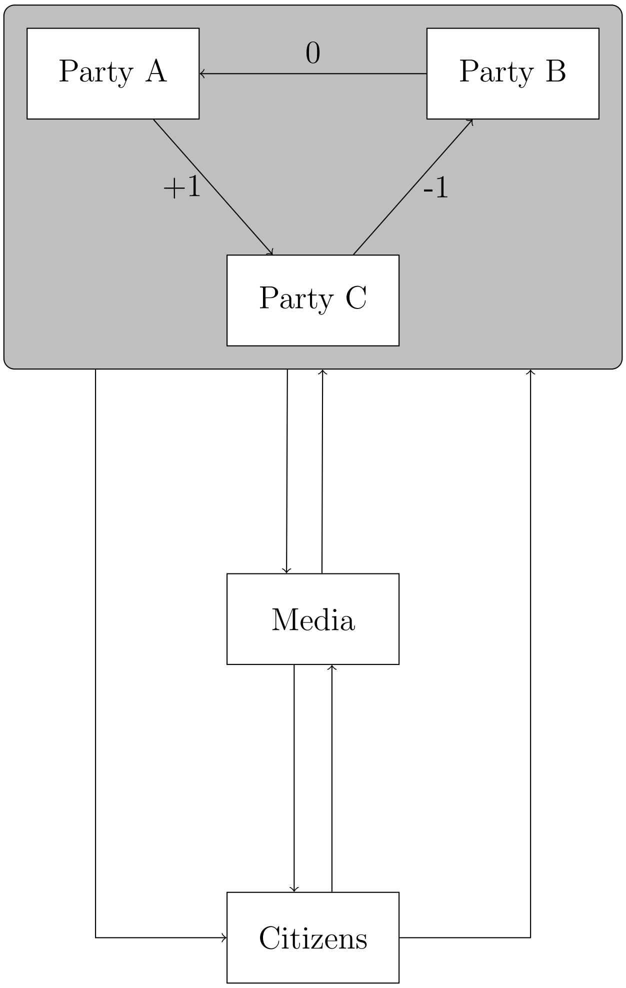
We define inter-party communication as purposeful communication between political parties. Specifically, following McNair (2011, p. 4), we describe inter-party communication as *all forms of communication undertaken by political parties about other political parties for the purpose of achieving specific objectives.* Each instance of inter-party communication involves a purposeful statement made by one party about another, with a specific positive, neutral, or negative stance.[1]

More formally, we propose a network-centric perspective on inter-party communication, building on the work by De Nooy and Kleinnijenhuis (2013) (see also Kleinnijenhuis and De Nooy, 2013; Song, Nyhuis and Boomgaarden, 2019, for a similar conceptualization). This perspective envisions inter-party communication as a dynamic network (Snijders, Van de Bunt and Steglich, 2010), consisting of various political parties (*nodes*) and statements about other parties (*ties*) with different stance polarities (positive, neutral, or negative). These ties are asymmetric, meaning each statement has a direction, and reciprocity is not required. The network is dynamic because the ties between the parties change over time.

Figure 1 provides an example of an inter-party communication network in a three-party system at a given point in time. Here, Party A sends a positive message about Party C, Party B makes a neutral statement about Party A, while Party C sends a negative

[1] As we understand it, *stance* is different from *sentiment*: The stance of an inter-party communication event refers to the attitude towards a party, while the sentiment regards the overall tone of the statement (see Bestvater and Monroe, 2023).

Figure 1: Inter-party communication network in a three-party system



Note: The arrows within the gray shaded square illustrate purposeful statements of positive (+1), negative (-1), or neutral (0) stance polarity between parties.

message about Party B. In this example, there is no reciprocity. This illustration also emphasizes the interdependence of all three key actor groups in political communication, including the media and the citizens. It shows that communication of parties reaches citizens either directly or via the media, that media coverage influences political parties (e.g., via agenda-setting) and that the interests of citizens shape media reporting, as well as the communication between parties (see, e.g. McNair, 2011, Chapter 1).

Building on this understanding, we propose a comprehensive conceptualization of inter-party communication, distinguishing between statements that convey a positive, negative, or neutral stance with regard to collaboration, policy, and personal issues.

Collaboration refers to statements made by one party with the intention of expressing willingness or reluctance to cooperate with other parties. Positive communication may include discussions about potential alliances or coalitions, negative communication could involve rejections or refusal of collaboration, while neutral communication comprises cases where a party declares that it does not seek, nor rules out, cooperation with another party.

Inter-party communication on policy pertains to statements about other parties' policy positions, issue emphasis, prospective policy pledges, or retrospective record. Positive communication may involve support or agreement on certain policies advocated by another party, negative communication may criticize or highlight perceived flaws in rival parties' policies, while neutral communication expresses a neutral stance towards the policies of other parties.

Personal issues relate to statements about individual representatives of rival parties, including party leaders or candidates. Positive communication may praise skills, experience and achievements of individuals from other parties, while negative communication may involve criticism, character attacks, or questioning the competence, integrity, or suitability of individuals for public office. Again, neutral communication conveys a neutral stance towards individuals of other parties.

It follows from this definition that inter-party communication has three inherent characteristics. First, inter-party communication is *purposeful*: It serves the parties' primary goal of maximizing votes in elections. Accordingly, statements made by parties about other parties seek to influence the voting calculus of citizens, for instance, by providing negative information about rival parties (e.g. Lau and Rovner, 2009), or by signalling coalition preferences (e.g. Gschwend, Meffert and Stoetzer, 2017). As citizens learn about party communications mainly through media coverage, parties have an incentive to make newsworthy statements to attract media attention (Strömbäck, 2008).

Second, inter-party communication is *directed*: Each instance of inter-party communication involves a sender party and one or more addressee parties. While reciprocity is not required (see above), empirical evidence suggests that attacks from one party often lead to counterattacks from the other party, following a tit-for-tat strategy (e.g., Dolezal, Ennser-Jedenastik and Müller, 2016).

Third, inter-party communication is *dynamic*: It is continuously ongoing and can change over time. This is because the logic of the electoral cycle influences public statements by parties about other parties. For instance, government parties may speak positively about each other throughout the legislative term to promote government unity but become more critical during election campaigns to win votes. Moreover, parties respond to prior attacks or support from other parties, resulting in an evolving communication network (De Nooy and Kleinnijenhuis, 2013).

Regardless of whether one considers collaboration, policy, or personal issues, inter-party communication tends to be subtle and diffuse. This is due, in part, to the directionality (lack of reciprocity) and dynamic nature (change over time) of inter-party communication. An important requirement for the empirical measurement of inter-party communication is, therefore, to be able to grasp the nuances and context of communication between parties.

3. Existing approaches to measure inter-party communication

Previous studies on specific types of inter-party communication focus on both, mediated and unmediated communication channels. For example, researchers collect mediated information from the news media, such as newspaper articles or TV news (De Nooy and Kleinnijenhuis, 2013; Lau and Pomper, 2002). Other studies examine unmediated communication channels, e.g., parties' campaign ads (Damore, 2002; Walter and van der Brug, 2013), press releases (Dolezal, Ennser-Jedenastik and Müller, 2016, 2017), election manifestos (Curini, 2011; Dolezal et al., 2018), or social media data (Auter and Fine, 2016; Gross and Johnson, 2016).

One common method for studying inter-party communication is manual coding, where human coders assess word choices and polarity to gain insights into how parties interact (see for example Best, 2015; Curini and Martelli, 2010; Druckman, Kifer and Parkin, 2009; Elmelund-Præstekær, 2010). Manual coding allows for a thorough and systematic understanding of how parties communicate about each other. Researchers can develop specific coding instructions tailored to their research questions, enabling them to gather precise information about their area of interest.

Yet, there are several drawbacks to manual coding. First, it is a labour- and resource-intense practice that requires multiple coders in order to ensure the reliability of the results. Second, the data obtained through manual coding reflect only what the original researchers intended to capture. This makes it challenging to apply the coded data to study other research questions or to analyze different cases. As a result, the reuse of manually coded data for new analyses is not straightforward.[2]

Hence, to efficiently collect data on inter-party communication, such a labour-intensive and time-consuming approach may not be the ideal solution. This is especially true when studying phenomena that recur in every election, such as negative campaigning or pre-electoral coalition signals. It is, therefore, critical to develop new methods for collecting these data in an automated manner that can be readily applied to new cases.

A promising new approach is the use of dictionaries, such as the one developed by, Bowler, McElroy and Müller (2022) to detect pre-electoral coalition signals. Although this dictionary currently identifies only positive signals between parties (indicating a willingness to form a government coalition together), it represents a significant step towards automating the study of inter-party communication as it can be applied to other cases than the one under immediate study.

However, the use of dictionary approaches raises questions about their ability to meet the requirements for a comprehensive measurement of inter-party communication. This is because dictionaries cannot fully grasp the nuances and context of communication between parties. Since the nature of political discourse is highly complex, understanding the underlying meaning and implications of statements demands a deeper level of analysis.

In the next section, we propose a transfer learning classifier as an alternative approach that addresses these challenges. Unlike dictionaries, transfer learning classifiers can cap-

[2]While manual coding is the most common way to investigating inter-party communication, there are other approaches such as unsupervised topic models(Grimmer, 2010), supervised learning (Fowler et al., 2020), or surveys among voters or party elites (Ketelaars, 2019).

ture the nuances in party rhetoric, including the varying degrees of positivity, neutrality, or negativity in statements, as well as the context in which they are made. This enables a more fine-grained analysis of inter-party communication. Moreover, the transfer learning approach enhances efficiency by reducing the labour-intensive process of manual coding. It can rapidly process vast amounts of data from different sources, making it ideal for studying inter-party communication across various elections and political landscapes.

4. A transfer learning approach for the classification of inter-party communication

In this section, we will present our transfer learning approach and illustrate how it can facilitate the collection and analysis of different types of inter-party communication. The method we propose makes use of recent advances in the area of Natural Language Processing (NLP), specifically, transfer learning based on pre-trained transformer-based language models (Vaswani et al., 2017; Devlin et al., 2019). Instead of training a network entirely from the ground up for a specific learning objective, these models are already pre-trained on large amounts of data using a general learning objective.

Typical tasks for pre-training are the prediction of masked words in the text (Masked Language Modelling) and the so-called "next sentence prediction". For the first task of Masked Language Modelling, around 15% of the tokens in the input text are randomly selected and are "masked" so that the model cannot see them. The model is then trained to predict probable slot fillers for the masked words. For the second task, the model is presented with pairs of sentences, S_1 and S_2, and has to predict whether S_2 is a probable continuation for S_1 in a natural language text. Through these tasks, the model learns which words are more probable to appear in the context of other words and which sentences are more likely to continue a certain text than others.

During pretraining, the model acquires general knowledge about natural language structure and meaning. This knowledge becomes highly valuable when the model is applied as a new measurement strategy for inter-party communication. Because the model has acquired general language knowledge through pretraining, it requires less annotated data for the actual task. This efficiency in data usage results in a more resource-effective approach. This approach is referred to as *transfer learning,* where knowledge learned through training on one task is transferred to solve another task, by using the model parameters of the first model as *a priori* information when learning the parameters of the second model.

To enhance the models' performance on a specific downstream task, we can add a fine-tuning step. In this step, we further refine the model parameters through supervised learning, using a small set of annotated data. The prediction of any type of elite communication is conceivable. Below, we showcase this for (i) the prediction of coalition signals in German newspaper articles and (ii) the prediction of opinion targets in negative campaigning in Austrian party press releases. Political communication takes place through diverse channels, including newspapers and TV news (mediated) and social media platforms (unmediated). To ensure the validity and applicability of its findings in real-world political scenarios, it is essential to demonstrate that the classification works for both mediated and unmediated forms of communication.

Transformer-based models (e.g., Devlin et al., 2019; Liu et al., 2019; Yang et al., 2019) have been applied to different areas of NLP so far and have been shown to increase prediction accuracy for many applications. For our task, we apply and evaluate a well-known transformer model, called BERT (Devlin et al., 2019).[3] BERT is an auto-encoder language model that learns to predict masked words in a text during pretraining. In addition, the BERT pretraining includes the task of predicting whether a sentence is a likely continuation of another sentence. Both tasks help the model to acquire knowledge about syntactic and semantic relations between the words and sentences in a large corpus. Our implementation is based on the Huggingface transformers library (Wolf et al., 2020) and PyTorch (Paszke et al., 2017). For more details on the training process, downsampling and hyperparameter settings, please refer to Appendix A.

We now describe the different steps of our general approach. Various types of inter-

[3]We also tested two other transformer models, RoBERTa (Liu et al., 2019) and XLNet (Yang et al., 2019), but obtained higher $F1$ scores on our data using BERT.

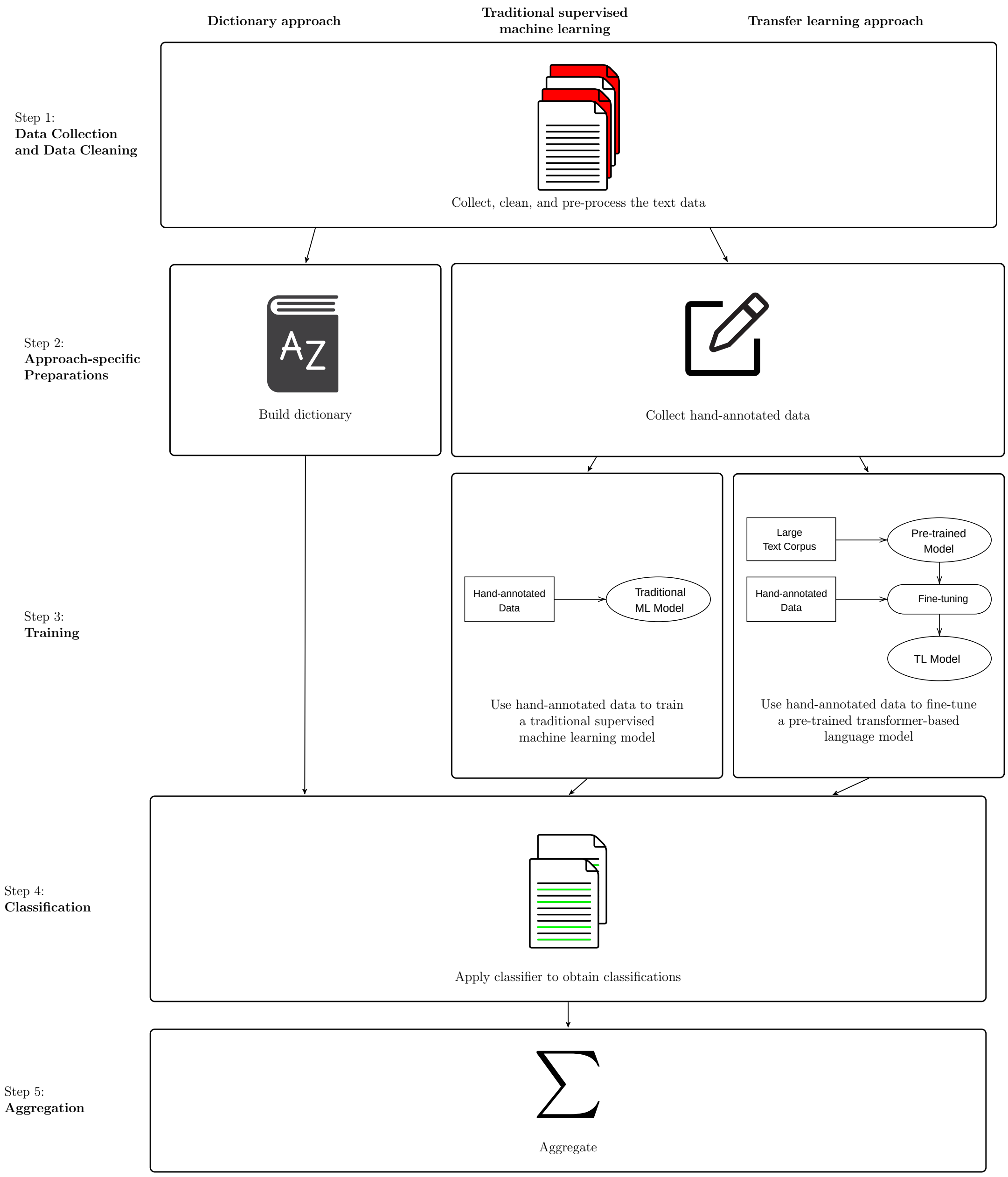
party communication can be modelled similarly. First, we need to identify the target of the communicated message (i.e., the target of negative campaigning or the recipient of the coalition signal). Second, we need to determine the stance of the message towards the target (i.e., whether the message takes a positive or negative stance towards the target or whether the coalition signal is positive or negative).

To model this, we decompose the problem into two separate tasks and train a classifier for each task. The first step consists of identifying sentences that contain a probable *target* and predicting the target name (i.e., party name or coalition option). In the second step, we learn to predict the *polarity* of the message (positive, neutral, negative). The motivation for this decomposition is that modelling both tasks in one step would result in numerous sparse label combinations of party name/coalition option + polarity that are infeasible to learn for any machine learning model. Breaking down the task into two steps reduces the number of labels and enhances the efficiency of training data usage. Polarity determination in a message or signal should not rely on the specific target, making this approach more effective.

We use the same transfer-based text classification model for both tasks where we present the model with a sentence or text sequence and, in the first step, learn to predict the target and, in the second step, to predict the polarity of the text sequence. We compare our approach to a traditional supervised machine learning system, and we also include a dictionary approach as a baseline for the performance evaluation. Figure 2 illustrates the step-by-step workflow for the dictionary approach (left), the traditional ML classifier (middle), and the transfer learning approach (right).

We now apply and test our new approach in two case studies for both mediated and unmediated communication channels. In Case Study 1, we try to predict pre-electoral coalition signals in German newspaper articles and in Case Study 2, we evaluate our method on predicting negative campaigning in Austrian party press releases. We focus on coalition signals and negative campaigning, as these are the most prominent types of communication between political parties (see for example De Nooy and Kleinnijenhuis, 2013; Vowe, 2009).

Figure 2: Three approaches to classify elite communication



5. Case study 1: Identifying coalition signals in the news

In our first case study, we are interested in pre-electoral *coalition signals*. Following Gschwend, Meffert and Stoetzer (2017), we define a coalition signal as a statement about a party's preference towards a possible coalition in which the party itself would be a member. These signals can be observed in parliamentary democracies, where the formation of coalitions after the election to form a government is typically required. As such, it may be regarded as an inter-party communication event that focuses on collaboration with either a positive, neutral or negative stance polarity.[4]

A coalition signal has three key components. These components include the sender, which is the party expressing the signal, the addressee, representing the potential coalition being addressed, and the statement's polarity, which can be positive, neutral, or negative.

For a statement to be identified as a coalition signal, it is not required that the sender directly speaks about the party's preferences concerning a specific coalition option, but it might also include the citation of (direct and indirect) quotes by a journalist. Sometimes, the addressee might be left unspecified (e.g., when a party rules out a coalition with "any left party"). In addition, we also consider a text segment to be a coalition signal if it is not completely clear whether the identified statement about a potential coalition was originally made by a representative of an involved party or merely by an observer, such as a journalist, expert or another political actor. This includes cases when the statement seems to be coming from a party as a whole, but it is not clear when or where that statement has been made, i.e., when it is not possible to identify the source.

5.1. Data Creation and Prediction of Coalition Signals

To identify coalition signals in newspaper articles, we rely on three different approaches: the dictionary approach of Bowler, McElroy and Müller (2022), a traditional ML approach, and our transfer learning approach.

As the first step in the process, we collect and clean the data (see Step 1 in Figure 2). The data we use in this study are German newspaper articles published prior to the German Bundestag elections in 1998, 2002, 2005, 2009, 2013, and 2017. We include all articles from two German daily newspapers[5] that were published in a period of up to four weeks (28 days) before each election (not including the actual election day). We first identify relevant articles with coalition signals based on a keyword search (see Step 1 in Figure 2). Therefore, we used the following case-insensitive search query: `*election* AND (*coalition* OR *pact* OR *collaboration* OR *alliance*)`.[6] All relevant articles

[4]This is in contrast to the conceptualization of De Nooy and Kleinnijenhuis (2013), who only consider pre-electoral coalition signals with positive stance polarity.

[5]These are the *Süddeutsche Zeitung* (SZ) and *Frankfurter Allgemeine Zeitung* (FAZ), which are considered to be two of the most important and influential German daily newspapers.

[6]The German translation of the search query is `*wahl* UND (*koalition* ODER *bündnis*`

need to include a mention of a party or a coalition term, as well as a mention of a federal or state election.

At the second step in the process, we set up the approach-specific preparations (see Step 2 in Figure 2). Bowler, McElroy and Müller (2022) set up the dictionary that they use to identify coalition signals in newspaper articles. This dictionary was designed to identify the presence of a coalition signal but does not capture the polarity of the signal. The authors use a dictionary with party names and terms indicating cooperation or coalition to identify sentences that might include coalition signals. If a sentence includes a term signalling cooperation *and* a reference to a coalition, it is classified as a coalition signal. To be considered as a coalition reference, the sentence must include at least two mentions of different parties *or* a coalition option that clearly identifies the parties participating in the coalition (e.g., "grand coalition" stands for a coalition between CDU/CSU and SPD or a "Jamaica coalition" refers to a coalition between the CDU/CSU, FDP and Greens).

In order to create a gold standard for training our classifier, we manually annotated newspaper articles. The annotators first read the whole article and highlight all signals in the article.[7] After this initial step, all highlighted signals in the article are coded in chronological order. For each identified coalition signal, we code the sender(s) and addressee(s) of the signal, so that each can be attributed to one or more specific parties. In cases where this information is not explicitly specified in the text, the instance is coded as *unspecified*. Moreover, the annotators identify and code the polarity of each signal (negative, neutral or positive).

Then we create a dataset where we map sender and addressee information to undirected coalition options (e.g., both, signals with SPD as sender and FDP as addressee as well as signals with FDP as sender and SPD as addressee are mapped onto SPD + FDP). We find 742 instances of coalition signals in our corpus of newspaper articles.[8] We observe 26 different coalition options, of which only 15 appear at least 10 times in our data (see Table A2 for an overview). When we also consider the polarity of the coalition, the number of different outcomes (coalition option × polarity) increases to 58, rendering the prediction task infeasible.

As described above, we decompose the task of automated coalition signal detection in newspaper articles into two subtasks. In the first step, we identify sentences that contain a coalition

`ODER *zusammenarbeit* ODER *allianz*).`

[7]All three annotators were students of political science who received extensive training and followed the instructions of a detailed codebook.

[8]When we distinguish between sender and addressee, the overall number of signals slightly increases.

signal and predict the coalition option (e.g., SPD + Greens). In the second step, we learn to predict the *polarity* of the signal (positive, neutral, negative). By decomposing the task into two subtasks, we can reduce the number of labels the model has to learn.

We model both subtasks as a sequence classification problem where we present the model with a sentence and let it learn, i.e., predict the coalition option (including 'none') in the first step and the polarity (positive, neutral, negative) of the respective sentence in the second step (Step 3 in Figure 2). As we only have a few instances for training, we use an n-fold cross-validation (e.g., Neunhoeffer and Sternberg, 2019) setting, where we train the model on five of the six elections and test the trained model on the unseen sixth election. We repeat this procedure six times in order to obtain predictions for each of the six elections.

As an additional baseline of comparison, we follow the same cross-validation setup as described above and train a supervised, feature-based machine learning classifier on our data. Specifically, we use a Support Vector Machine (SVM) algorithm, which has been shown to obtain good results for text classification problems.[9] We pre-process our input data by tokenizing the data and removing stop words. Then we create the feature vectors that are the input to our SVM. For feature extraction, we use a "bag of words" approach and weigh the extracted features, based on the term frequency–inverse document frequency (TF-IDS) weighting scheme.

During training, we perform Bayesian optimization over hyperparameters on the training set in a stratified 10-fold setup, using Bayes search. Then we use the best parameter setting for each training fold to predict coalition signals in our unseen test set. In the end, we collect all predictions for the targets and the respective polarity for the six test folds and evaluate them against the ground truth annotations (see below for the results of this evaluation).

We then apply both classifiers to the newspaper articles to provide us with sentence-level predictions of coalition signals (see Step 4 in Figure 2). The dictionary method by Bowler, McElroy and Müller (2022) does not encode the polarity of coalition signals. In contrast, the transfer learning approach is able to differentiate between positive, negative, and neutral coalition signals. Both methods, however, cannot distinguish between the sender and addressee of a coalition signal.

After obtaining sentence-level classifications of coalition signals, it is necessary for many applications to generate a summary of coalition signals in a particular election campaign or

[9]Previous assessments validated that the SVM algorithm consistently outperformed other supervised classification methods in terms of performance for our application.

text (see Step 5 in Figure 2). One potential aggregation measure of coalition signals was introduced by Bowler, McElroy and Müller (2022) and indicates how often a coalition option c is mentioned compared to all other coalition options before a particular election e. This simple measure is defined as the quotient of the number of sentences that contain coalition signals concerning coalition option c, $n_{e,c}$, and the total number of sentences that contain coalition signals (irrespective of the coalition signalled), N_e: $\theta_{e,c} = \frac{n_{e,c}}{N_e}$. As a relative frequency, this measure ranges from 0 to 1. $\theta^1_{e,c} = 0$ implies that there is not a single coalition signal concerning coalition option c before election, e while $\theta^1_{e,c} = 1$ means that all coalition signals are about coalition option c. As this measure does not account for the polarity of coalition signals, we refer to $\theta^1_{e,c}$ as *salience* measure of coalition signals.[10]

5.2. Evaluation

We assess the performance of our transfer learning approach based on the hand-coded German newspaper articles, as outlined earlier. For our annotated dataset, we gather sentence-level predictions from our transfer learning method, as well as the dictionary approach by Bowler, McElroy and Müller (2022) and the SVM classifier, and evaluate these predictions against the manual annotations. We use the precision, recall and F1-Score to evaluate the performance.[11] The transfer learning predictions result from a 6-fold cross-validation setup, generating out-of-sample predictions for each of the six test folds.

We present the results for four evaluation settings (A, B, C, and D), summarized in Table 1. In setting (A), we compare the three approaches at the sentence level, assessing the correct prediction of coalition signals without considering coalition option or polarity. The results show that our transfer learning strategy (40.9) outperforms the dictionary approach (32.1) and the SVM (33.0). Although the dictionary approach has higher precision than the transfer learning strategy, it definitely has lower recall (28.9% vs. 48.7%). The results show a higher $F1$ score

[10] An alternative aggregation measure that also considers the polarity of coalition signals could, for example, be based on the log-RILE scale introduced by Lowe et al. (2011) for measuring party positions: $\theta_{e,c} = \log \frac{n^{pos}_{e,c}+0.5}{n^{neg}_{e,c}+0.5}$. This measure is defined as the logged quotient of the number of sentences that contain a *positive* coalition signal concerning a coalition option c before election e, $n^{pos}_{e,c}$, and the number of sentences that contain *negative* coalition signals concerning this coalition, $n^{neg}_{e,c}$.

[11] The precision is the ratio of true positive predictions to all positive predictions made by the model. The recall is the ratio of true positive predictions to all actual positive instances. The F1 score is the harmonic mean of precision and recall, providing a balance between the two metrics. A higher F1 score indicates a model that performs well both in terms of precision and recall.

Table 1: Results for the prediction of coalition signals in newspaper articles

	(A) Signal (yes/no)			(B) Coalition			(C) Coalition, polarity			(D) Aggregation
	Prec	**Rec**	**F1**	**Prec**	**Rec**	**F1**	**Prec**	**Rec**	**F1**	**RMSE**
Dictionary	36.0	28.9	32.1	81.0	37.7	51.4	46.3	21.0	28.9	0.0015
SVM	30.9	35.5	33.0	64.9	36.4	46.6	41.2	23.4	29.9	0.0024
Transfer	35.2	48.7	40.9	65.5	49.2	56.2	42.6	32.7	37.0	0.0010

Note: (A) Prediction of coalition signals on the sentence-level (signal: yes/no); (B) prediction of the coalition option on the signal-level; (C) prediction of the coalition option *and* polarity on the signal-level; and (D) prediction of the aggregation measure $\theta^1_{e,c}$.

for our transfer learning approach (40.9) than for the dictionary approach (32.1) and the SVM (33.0). Although the dictionary approach has a higher precision than the transfer learning approach, this clearly comes at the cost of recall (28.9% versus 48.7%).

In settings (B) and (C), we shift the focus to the signal level. However, challenges arise because hand annotations frequently span multiple sentences per signal, while predictions are at the sentence level. By mapping predictions of individual sentences to corresponding signals, we are able to evaluate the performance on the signal level. Setting (B) evaluates predicted coalition options regardless of polarity while setting (C) considers both coalition option and polarity.[12] The $F1$ score in setting (B) is higher for our transfer learning approach (56.2) than for the competing dictionary approach (51.4) where our approach again has a lower precision but a higher recall. Results for the SVM are even below the dictionary baseline (46.6). In setting (C), the accuracy as measured by the $F1$ score is again higher for our approach (37.0) than for the dictionary approach (28.9) and the SVM (29.9). The transfer learning approach accurately predicts 243 out of the 742 coalition signals in the data, whereas the dictionary and SVM approaches only detect 156 (dictionary) and 174 signals (SVM), respectively.

Setting (D) extends the evaluation beyond the sentence- and signal-level by taking into account coalition signal aggregations. We compute the salience values for both the ML approaches and the dictionary approach and measure the difference between the predicted and the true salience values using the root-mean-square error (RMSE).[13] The RMSE for our transfer learn-

[12] Arguably, setting (B) is a fairer comparison than setting (C) as the dictionary approach does not model the polarity of coalition signals.

[13] The root-mean-square error is defined as follows: $\sqrt{\frac{\sum_{e,c}(\theta^{1,hand}_{e,c} - \theta^{1,pred}_{e,c})^2}{\sum_{e,c} 1}}$.

ing approach is lower than that for the dictionary approach; yet, both values are fairly similar. As a result, the salience values for election-coalition combinations predicted by our approach are slightly closer to the true values than the salience values by the dictionary approach. The error for the SVM, however, is notably higher despite obtaining a slightly higher $F1$ (Setting C). This might reflect the lower precision of the model, as compared to the dictionary approach. The null hypothesis for this test states that the two marginal probabilities for each outcome (dictionary, transfer learning) are the same (i.e., there is no significant difference between the predictions of our two models).

In sum, we find that our transfer-learning approach outperforms both the dictionary approach of Bowler, McElroy and Müller (2022) and the SVM classifier on each of the evaluation settings described above, as indicated by the $F1$ scores and the RMSE.

6. Case study 2: Identifying negative campaigning in party press releases

In the second case study, we evaluate our approach for predicting negative campaigning in press releases by political parties in Austria during election campaigns. Negative campaigning occurs when a party's "campaign will concentrate on the perceived weaknesses of their opponent's policy proposals, prior policy failures, and/or personal peccadilloes" (Lau and Rovner, 2009, p. 286). Hence, negative campaigning can be thought of as inter-party communication about policy or personal issues with a negative stance polarity.

6.1. Data Creation and Prediction of Negative Campaigning

To be able to study negative campaigning in the context of Austrian elections, we need to identify whether a press release addresses a specific party and, in addition, whether the author's stance towards this party is positive, negative or neutral. Again, we compare the results for our proposed transfer learning approach to a traditional feature-based SVM baseline and a simple dictionary approach, as described below.

To do that we rely on a data set of party press releases provided by the Austrian National Election Study (AUTNES) (Müller et al., 2021), which includes press releases issued by Austrian parties within a time frame of six weeks before the national elections in 2002, 2006, 2008, 2013 and 2017. The data contain the title, subtitle and the content of the press release, together with some meta-information (such as the author of the release and their party affiliation).

After cleaning the data (see Step 1 in Figure 2), human annotators coded the author of the release, their party affiliation, the issues or actors discussed in the release and the author's stance towards this issue or actor (support, rejection, neutral) (Step 2 in Figure 2). The data cover nearly 10,000 press releases (see Table A7), coded for 425 different issues or actors, including 12 political parties. In our work, we focus on those press releases that target the four political parties that competed in all elections during the respective time frame (SPÖ, ÖVP, FPÖ, Greens) and assign all press releases that are not about one of these parties to the OTHER class. As a result, we end up with a set of five target labels that we want to predict.

Once we know the target, we want to predict whether a press release conveys a positive, negative or neutral stance toward this particular target. The prediction of stance, however, is only meaningful for press releases with a target, i.e., that include (positive or negative) statements about a rival party. We remove all press releases without a target (OTHER class) from the training and validation set, assuming that there is no useful information for the classifier to learn. For the prediction, however, we keep these instances in the test set and include them in the evaluation, mapping their label to the negative class. This means that the results for stance prediction are unrealistically low, which is not a problem for us as we are mostly interested in the detection of negative campaigning. For the final evaluation, we rely on the predictions of the target classifier to identify press releases that belong to the OTHER class and only use the stance predictions for the remaining instances.

For the dictionary approach, we create a simple dictionary with party names to identify relevant targets in the press releases (for details, see Table A8 in the Appendix). To predict the author's stance towards the target, we use Sentimerge (Emerson and Declerck, 2014), a large sentiment dictionary with more than 15,000 entries. To prepare the resources for later use, we remove all entries from Sentimerge where the sentiment score is zero. This reduces the number of entries to 14,050. We also lowercase the text in the title and subtitle of the press releases and remove stopwords.

Our setup for training follows the approach described above (Step 3 in Figure 2). Again, we decompose the task into two subtasks, (1) predicting the target (SPÖ, ÖVP, FPÖ, Greens, OTHER) and (2) the author's stance toward the target. We use the same model architectures as before and train a traditional feature-based SVM classifier and a transformer-based transfer learning model (see Subsection B in the Appendix for details on feature extraction and training). As the data includes press releases for five elections, we run a 5-fold cross-validation where the data for each election year is once used as the out-of-sample test set.

In the classification step (Step 4 in Figure 2), we apply our trained classifiers to the data in the test set, following the 5-fold cross-validation regime. For the dictionary baseline, we use the party name dictionary to look for relevant mentions in the title or subtitle of each press release and label it as the target. If we find more than one party name in the release, we greedily assign the target label to the first candidate.[14] If no party name is found, we label this instance as OTHER.

For polarity prediction, we iterate over the SentiMerge dictionary and look for entries that occur in the press release. We then sum up the (positive and negative) sentiment scores for those entries and normalize the result by the number of dictionary terms found in the text. Please note that we count each dictionary term only once even if it occurs multiple times in the text. As a result, we obtain a score for each press release, based on the sum of all sentiment terms in the release. To predict the author's stance towards the target, we determine a threshold for positive and negative polarity. Therefore, we use a bootstrapping process where we repeatedly draw 1,000 data samples with replacement, which gives us a distribution of the sample means. We then take the 2.5 and 97.5 percentile of the distribution as the confidence interval and label every press release with a polarity score below the 2.5 percentile as negative and every release with a score higher than the 97.5 percentile as an instance with positive polarity.

6.2. Evaluation

Table 2 shows results for target prediction for the dictionary baseline (1), the SVM (2), and the transfer learning approach (3). The last column presents the averaged F1 score over the five folds for each election year. The dictionary approach outperforms the SVM in the first two elections (2002, 2006) but results decrease over time. On average, the two baselines obtain results in the same range (0.60 F1). For the transfer learning approach, we report averaged

[14] We also ran experiments where we included the content of the press release, however, results on the development set decreased for this setting, due to a higher number of false positives.

Table 2: Results for negative campaigning in Austrian press releases

	2002	2006	2008	2013	2017	Total
Target						
Dictionary	0.60	0.63	0.63	0.55	0.56	0.60
SVM	0.56	0.57	0.64	0.59	0.65	0.60
Transfer	0.73± 0.05	0.73± 0.04	0.76± 0.05	0.66± 0.05	0.70± 0.05	0.70
TransferE	0.76	0.78	0.80	0.73	0.75	0.76
Target + Stance						
Dictionary	0.54	0.56	0.56	0.48	0.48	0.53
SVM	0.52	0.53	0.59	0.56	0.61	0.56
Transfer	0.67± 0.04	0.67± 0.04	0.69± 0.05	0.68± 0.05	0.68± 0.05	0.68
TransferE	0.73	0.74	0.76	0.69	0.72	0.73

Note: Results for the transfer learning approach are averaged over five independent runs with different initialisations (± reports standard deviation). TransferE reports results for an ensemble classifier with majority vote.

results over five different initializations. Our method consistently outperforms both baselines on each election, with an average F1 score of 0.70. However, we notice a high standard deviation of 0.05 between models from different initializations. To address this issue, we apply a well-known ML technique, i.e., ensemble learning (Dasarathy and Sheela, 1979; Dong et al., 2020), where we consider the five models as an ensemble of classifiers and determine the predicted labels by taking the majority vote over the classifier ensemble (TransferE).

Given that we already trained the individual models to report averaged results for different initializations, all we have to do now is to collect the predictions made by the five models and assign each instance the label that has been predicted by the majority of the classifiers. We then evaluate the final labels against the manually assigned labels and report results in Table 2 (TransferE). As has been noted before, "an ensemble is often more accurate than any of the single classifiers in the ensemble" (Opitz and Maclin, 1999, 169). This is also true for our setting where we observe an improvement over the individual models for each of the elections, with an average F1 of 0.76%.

Results for predicting both target and stance, are shown in Table 2, rows (5)–(8).[15] To obtain

[15]Please note that we cannot report results for stance prediction alone as we removed the press releases without a target from the training data and therefore need the combined labels to identify the NONE class (press releases with no target).

the final labels, we merge the target and stance predictions into one atomic label and evaluate these labels against the manually coded classes. The dictionary approach performs poorly, with an F1 in the range of 0.48 to 0.56%. The SVM yields slightly better results (0.52% – 0.61% micro-F1). Again, the transformer-based approach outperforms both baselines on each individual election year, showing that the context-sensitive semantic representations are superior to symbolic, word-count-based representations (i.e., the tf-idf features used by the SVM). Combining the predictions of the classifier ensemble achieves best results and improves the averaged results from 0.68 to 0.73% micro-F1, thus increasing the robustness of our prediction model.

7. Discussion

In this article, we explore a crucial but often overlooked facet of representative democracies: inter-party communication. By highlighting its pivotal role in facilitating dialogue, information exchange, and compromise among political parties, we underscore its significance in fostering democratic processes. Despite its importance, inter-party communication has remained relatively understudied, partly due to the lack of a comprehensive conceptualization and the challenges associated with large-scale measurement.

Addressing these gaps, we propose a refined understanding of inter-party communication that encompasses diverse dimensions, including collaboration, policy, and personal issues. Specifically, we define inter-party communication as public communication by political parties about others, pursuing specific objectives, that conveys either a positive, neutral, or negative stance. Thereby we bring together different strands of the literature that rarely speak to one another, including spatial and issue competition, negative campaigning, personalization, and coalition signals. This enriched conceptual framework provides a dynamic, network-centric view of purposeful communication between parties, deepening our understanding of political interactions and their evolution over time.

Our second contribution lies in the development of a novel transfer learning approach to effectively measure inter-party communication at a large scale. By harnessing the capabilities of transformer-based language models, we surpass conventional methods like manual coding and dictionary-based techniques in capturing the subtleties, context, and polarity within sentences. This approach is particularly valuable as it enables the classification of intricate concepts without necessarily demanding an excessive volume of training data.

The effectiveness of our approach was demonstrated with two case studies on negative campaigning in party press releases and coalition signals in newspaper articles, which represent particularly frequent and important types of elite communication. The two applications show that our approach offers a powerful tool for researchers to analyse textual data on inter-party communication from mediated as well as unmediated communication channels.

Our study has far-reaching implications for the study of political communication and beyond. By providing a unified framework for diverse forms of communication, we can now subsume various types of communication that were previously studied in isolation under a common umbrella. This conceptual clarity allows for a more accurate assessment of parties' communication dynamics and enables us to uncover the broader patterns and strategies that drive party competition, collaboration, and negotiation dynamics.

Moreover, it allows for a more accurate assessment of the impact of political discourse on public opinion, electoral outcomes, and democratic processes. By shedding light on the full spectrum of different forms of inter-party communication, we provide researchers with a more nuanced view of the information environment of voters for making well-informed electoral choices. This richer understanding of party dynamics includes information on policy positions, political alliances, and the strategic underpinnings of inter-party communication.

In addition, the versatility of our measurement approach has broader implications for the field of automatic text classification, advancing beyond traditional methods like dictionaries and supervised machine learning models. The transfer learning approach for inter-party communication can be applied to diverse data sources from various communication platforms, including press releases, social media postings, speeches, and others. The method can be extended to explore a range of different research questions, such as the impact of inter-party communication on voter behaviour and public opinion, the evolution of party strategies over time, and the cross-national variations in party communication tactics.

Overall, this study bridges a critical gap in political communication research by enhancing our understanding of inter-party communication in representative democracies. It illuminates the importance of studying inter-party communication and offers valuable insights into how parties strategically engage with each other. As the field of political communication continues to evolve, our work provides a solid foundation for further investigation of the complex interactions that shape democratic systems.

Data Availability Statement

The data and replication files supporting the results and analyses presented in the paper will be available upon request from the corresponding author following the publication of the research findings.

References

Auter, Zachary J and Jeffrey A Fine. 2016. "Negative campaigning in the social media age: Attack advertising on Facebook." *Political Behavior* 38(4):999–1020.

Barberá, Pablo, Amber E Boydstun, Suzanna Linn, Ryan McMahon and Jonathan Nagler. 2021. "Automated text classification of news articles: A practical guide." *Political Analysis* 29(1):19–42.

Best, Volker. 2015. *Koalitionssignale bei Landtagswahlen Eine empirische Analyse von 1990 bis 2012*. Nomos.

Bestvater, Samuel E and Burt L Monroe. 2023. "Sentiment is not stance: Target-aware opinion classification for political text analysis." *Political Analysis* 31(2):235–256.

Bowler, Shaun, Gail McElroy and Stefan Müller. 2022. "Voter expectations of government formation in coalition systems: The importance of the information context." *European Journal of Political Research* 61(1):111–133.

Curini, Luigi. 2011. "Negative Campaigning in No-Cabinet Alternation Systems: Ideological Closeness and Blames of Corruption in Italy and Japan Using Party Manifesto Data." *Japanese Journal of Political Science* 12(3):399–420.

Curini, Luigi and Paolo Martelli. 2010. "Ideological proximity and valence competition. Negative campaigning through allegation of corruption in the Italian legislative arena from 1946 to 1994." *Electoral Studies* 29(4):636–647.

Damore, David F. 2002. "Candidate Strategy and the Decision to Go Negative." *Political Research Quarterly* 55(3):669–685.

Dasarathy, Belur V and Belur V Sheela. 1979. "A composite classifier system design: Concepts and methodology." *Proceedings of the IEEE* 67(5):708–713.

De Nooy, Wouter and Jan Kleinnijenhuis. 2013. "Polarization in the Media During an Election Campaign: A Dynamic Network Nodel Predicting Support and Attack Among Political Actors." *Political Communication* 30(1):117–138.

Devlin, Jacob, Ming-Wei Chang, Kenton Lee and Kristina Toutanova. 2019. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, NAACL-HLT 2019, Minneapolis, MN, USA, June 2-7, 2019, Volume 1 (Long and Short Papers)*, ed. Jill Burstein, Christy Doran and Thamar Solorio. Association for Computational Linguistics pp. 4171–4186.

Dolezal, Martin, Laurenz Ennser-Jedenastik and Wolfgang C Müller. 2016. "Negative Campaigning and the Logic of Retaliation in Multiparty Competition." *The International Journal of Press/Politics* 21(2):253–272.

Dolezal, Martin, Laurenz Ennser-Jedenastik and Wolfgang C Müller. 2017. "Who will attack the competitors? How political parties resolve strategic and collective action dilemmas in negative campaigning." *Party Politics* 23(6):666–679.

Dolezal, Martin, Laurenz Ennser-Jedenastik, Wolfgang C Müller, Katrin Praprotnik and Anna Katharina Winkler. 2018. "Beyond salience and position taking: How political parties communicate through their manifestos." *Party Politics* 24(3):240–252.

Dong, Xibin, Zhiwen Yu, Wenming Cao, Yifan Shi and Qianli Ma. 2020. "A Survey on Ensemble Learning." *Front. Comput. Sci.* 14(2):241–258.

Druckman, James N, Martin J Kifer and Michael Parkin. 2009. "Campaign Communications in U.S. Congressional Elections." *American Political Science Review* 103(3):343–366.

Dumdum, Omar O and Levi Bankston. 2022. "The Interplay of Actors in Political Communication: The State of the Subfield." *Political Communication* 39(2):266–279.

Elmelund-Præstekær, Christian. 2010. "Beyond American negativity: Toward a general understanding of the determinants of negative campaigning." *European Political Science Review* 2(1):137–156.

Emerson, Guy and Thierry Declerck. 2014. SentiMerge: Combining Sentiment Lexicons in a Bayesian Framework. In *Proceedings of Workshop on Lexical and Grammatical Resources for Language Processing*. Dublin, Ireland: Association for Computational Linguistics and Dublin City University pp. 30–38.

Fowler, Erika F, Michael M Franz, Gregory J Martin, Zachary Peskowitz and Travis N Ridout. 2020. "Political advertising online and offline." *American Political Science Review* 115(1):130–149.

Golder, Sona N. 2005. "Pre-electoral coalitions in comparative perspective: A test of existing hypotheses." *Electoral Studies* 24(4):643–663.

Grimmer, Justin. 2010. "A Bayesian Hierarchical Topic Model for Political Texts: Measuring Expressed Agendas in Senate Press Releases." *Political Analysis* 18(1):1–35.

Gross, Justin H and Kaylee T Johnson. 2016. "Twitter Taunts and Tirades: Negative Campaigning in the Age of Trump." *PS Political Science & Politics* 49(4):748–754.

Gschwend, Thomas, Michael F. Meffert and Lukas F. Stoetzer. 2017. "Weighting Parties and Coalitions: How Coalition Signals Influence Voting Behavior." *Journal of Politics* 79(2):642–655.

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

Haselmayer, Martin, Thomas M Meyer and Markus Wagner. 2019. "Fighting for attention: Media coverage of negative campaign messages." *Party Politics* 25(3):412–423.

Ketelaars, Pauline. 2019. "Position, Preference and Personality: A Microlevel Explanation of Negativity in Day-To-Day Politics." *Political Psychology* 40(5):1019–1038.

Kleinnijenhuis, Jan and Wouter De Nooy. 2013. "Adjustment of issue positions based on network strategies in an election campaign: A two-mode network autoregression model with cross-nested random effects." *Social Networks* 35(2):168–177.

Lau, Richard R and Gerald M Pomper. 2002. "Effectiveness of Negative Campaigning in U.S. Senate Elections." *American Journal of Political Science* 46(1):47–66.

Lau, Richard R and Ivy Brown Rovner. 2009. "Negative campaigning." *Annual review of political science* 12:285–306.

Liu, Yinhan, Myle Ott, Naman Goyal, Jingfei Du, Mandar Joshi, Danqi Chen, Omer Levy, Mike Lewis, Luke Zettlemoyer and Veselin Stoyanov. 2019. "RoBERTa: A Robustly Optimized BERT Pretraining Approach." *CoRR* abs/1907.1.

Lowe, Will, Kenneth Benoit, Slava Mikhaylov and Michael Laver. 2011. "Scaling policy preferences from coded political texts." *Legislative studies quarterly* 36(1):123–155.

McNair, Brian. 2011. *An introduction to political communication*. Routledge.

Meffert, Michael F and Thomas Gschwend. 2011. "Polls, coalition signals and strategic voting: An experimental investigation of perceptions and effects." *European Journal of Political Research* 50(5):636–667.

Müller, Wolfgang C., Anita Bodlos, Martin Dolezal, Nikolaus Eder, Laurenz Ennser-Jedenastik, Christina Gahn, Elisabeth Graf, Martin Haselmayer, Teresa Haudum, Lena Maria Huber, Matthias Kaltenegger, Thomas M. Meyer, Katrin Praprotnik, Verena Reidinger and Anna Katharina Winkler. 2021. "AUTNES Content Analysis of Party Press Releases: Cumulative File (SUF edition).".

Neunhoeffer, Marcel and Sebastian Sternberg. 2019. "How cross-validation can go wrong and what to do about it." *Political Analysis* 27(1):101–106.

Opitz, David and Richard Maclin. 1999. "Popular Ensemble Methods: An Empirical Study." *J. Artif. Int. Res.* 11(1):169–198.

Paszke, Adam, Sam Gross, Soumith Chintala, Gregory Chanan, Edward Yang, Zachary DeVito, Zeming Lin, Alban Desmaison, Luca Antiga and Adam Lerer. 2017. Automatic differentiation in PyTorch. In *NIPS Autodiff Workshop*.

Skaperdas, Stergios and Bernard Grofman. 1995. "Modeling negative campaigning." *American Political Science Review* 89(1):49–61.

Snijders, Tom AB, Gerhard G Van de Bunt and Christian EG Steglich. 2010. "Introduction to stochastic actor-based models for network dynamics." *Social networks* 32(1):44–60.

Song, Hyunjin, Dominic Nyhuis and Hajo Boomgaarden. 2019. "A network model of negative campaigning: the structure and determinants of negative campaigning in multiparty systems." *Communication Research* 46(2):273–294.

Strömbäck, Jesper. 2008. "Four phases of mediatization: An analysis of the mediatization of politics." *The international journal of press/politics* 13(3):228–246.

Vaswani, Ashish, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser and Illia Polosukhin. 2017. Attention is all you need. In *Advances in Neural Information Processing Systems*. pp. 5998–6008.

Vowe, Gerhard. 2009. Feldzüge um die öffentliche Meinung. In *PR-Kampagnen*, ed. Ulrike Röttger. Wiesbaden: VS Verl. für Sozialwissenschaften pp. 69–86.

Walter, Annemarie S. and Wouter van der Brug. 2013. "When the gloves come off: Inter-party variation in negative campaigning in Dutch elections, 1981–2010." *Acta Politica* 48(4):367–388.

Wolf, Thomas, Lysandre Debut, Victor Sanh, Julien Chaumond, Clement Delangue, Anthony Moi, Pierric Cistac, Tim Rault, Rémi Louf, Morgan Funtowicz, Joe Davison, Sam Shleifer, Patrick von Platen, Clara Ma, Yacine Jernite, Julien Plu, Canwen Xu, Teven Le Scao, Sylvain Gugger, Mariama Drame, Quentin Lhoest and Alexander M Rush. 2020. Transformers: State-of-the-Art Natural Language Processing. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing: System Demonstrations, EMNLP 2020 - Demos, Online, November 16-20, 2020*, ed. Qun Liu and David Schlangen. Association for Computational Linguistics pp. 38–45.

Yang, Zhilin, Zihang Dai, Yiming Yang, Jaime G Carbonell, Ruslan Salakhutdinov and Quoc V Le. 2019. XLNet: Generalized Autoregressive Pretraining for Language Understanding. In *Advances in Neural Information Processing Systems 32: Annual Conference on Neural Information Processing Systems 2019, NeurIPS 2019, December 8-14, 2019, Vancouver, BC, Canada*, ed. Hanna M Wallach, Hugo Larochelle, Alina Beygelzimer, Florence D'Alché-Buc, Emily B Fox and Roman Garnett. pp. 5754–5764.

Zaller, John. 1999. *A Theory of Media Politics: How the Interests of Politicians, Journalists, and Citizens Shape the News.* University of Chicago Press.

Appendix

Measuring Inter-party Communication: A transfer learning approach

Table of Contents

A. Coalition Signals

Data

After cleaning the data, we end up with a corpus of 15,735 sentences from 342 newspaper articles, of which each article includes at least one coalition signal (see Table A1).

Table A1: Summary of our data on the German federal elections 1998, 2002, 2005, 2009, 2013 and 2017

Year	Newspaper	Number of articles	Number of sentences
1998	SZ	17	620
	FAZ	54	2,516
2002	SZ	8	234
	FAZ	21	684
2005	SZ	27	1,289
	FAZ	34	1,476
2009	SZ	25	1,607
	FAZ	44	2,109
2013	SZ	25	1,461
	FAZ	34	1,746
2017	SZ	17	784
	FAZ	18	1,209
Total		324	15,735

Note: (SZ: *Süddeutsche Zeitung*, FAZ: *Frankfurter Allgemeine Zeitung*).

Table A2 indicates the most common coalition options. The top three identified coalition options are in fact the governments that resulted from the German federal elections between 1998 and 2017.

Training Details

Hyperparameters We used the following hyperparameter settings in our experiments. All our models were trained with a maximum sequence length of 128 tokens per input sentence and a batch size of 8. The batch size is a hyperparameter of our optimization algorithm, *gradient descent*, that specifies the number of training instances to be presented to the model before each parameter update. The coalition signal detection model was trained for 10 epochs. The number of epochs is another hyperparameter of our optimization algorithm and specifies how many passes through the training data the model has to complete during training. Due to its smaller size, the polarity prediction model was trained for 6 epochs only, to avoid overfitting. The learning rate was set to 1e-5. The learning rate determines the step size that the optimization algorithm takes towards minimizing the loss function during training. We trained all our models on a single GPU with 11GB RAM (NVIDIA RTX 2080Ti).

Downsampling The number of sentences in our dataset with and without coalition signals is highly imbalanced. This is because the vast majority of sentences do not include any coalition signal and are therefore labelled as negative examples for the phenomenon that we would like to learn. Consequently, we use downsampling (also known as *choice-based* sampling) to obtain a smaller dataset where the number of sentences with and without coalition signals is more

Table A2: Distribution of coalition options that are mentioned at least 10 times in our data (sentence level counts)

Frequency	Coalition option
352	CDU/CSU + FDP
341	CDU/CSU + SPD
173	SPD + Greens
172	SPD + FDP + Greens
89	CDU/CSU + FDP + Greens
73	SPD + LEFT
67	CDU/CSU + Greens
66	SPD + LEFT + Greens
23	CDU/CSU + other
22	SPD + FDP
18	SPD + unspecified
13	FDP + unspecified
10	LEFT + Greens
10	FDP + Greens
10	CDU/CSU

evenly distributed. For each training fold, we randomly select negative instances, i.e., sentences without a coalition signal, until the ratio of negative instances is $N\times$ the number of positive instances, i.e., sentences that contain a coalition signal. We experimented with different options for N and decided to set it to $N = 2$, as this resulted in a higher $F1$ score on a held-out subset of the training data.

For the second step, we only trained our polarity prediction model on the ground truth signals, i.e., the coalition signals detected by our annotators, ignoring all instances that are not signals, to learn to predict whether the signal is positive, negative or neutral. The test data includes all instances from the respective fold. To obtain the final prediction for each sentence in our test set, we only consider sentences where our first model detects a coalition signal. Only for those sentences, we also use the polarity prediction from the second model. As a result, we obtain predictions that encode a coalition option *and* polarity for each sentence (including 'none' for sentences where our first model did not predict a coalition signal) and evaluate those predictions against the ground truth.

Data Augmentation

A well-known property of supervised machine learning is that, in order to work well, it requires that the data points in the training and test sets come from a similar distribution. From Table A3, however, we can see that the results for the transfer learning models show considerable variation between different elections. Therefore, we now have a closer look at the results for individual elections.

Figure A1 plots $F1$ scores for the three models for the individual elections. For each election, we took 100 random samples (with replacement), where each sample includes 50% of the coalition signals for this election. We then evaluate each model on each of the 100 samples, which results in 2x100 data points for each election year. Based on these inputs, we draw a line plot with a 95% confidence interval.

Figure A1 shows that our approach achieves substantial improvements over the baseline for three elections (2002, 2005, 2009) while for the remaining elections (1998, 2013, 2017) the results

Table A3: Evaluation of coalition signal prediction ($F1$ scores for coalition type + polarity) on the signal level for each election (six-fold cross-validation setup)

Election	**Dictionary**	**SVM**	**Transfer**
1998	30.04	29.88	30.45
2002	17.98	24.24	48.21
2005	27.12	37.25	41.96
2009	32.96	39.38	37.93
2013	36.26	21.65	40.17
2017	18.32	7.55	19.23
Total	28.89	29.90	37.01

for baseline and transfer learning are actually very close. The gap in performance between the dictionary-based method and the transfer learning model is particularly large for the 2002 election, where the dictionary-based method predicted too many signals for a coalition between CDU/CSU and SPD, while the actual data only contained six instances for this particular coalition. Below are some examples where the model was wrong. It is striking that most of the examples include the term *Große Koalition* which automatically triggers the prediction of the aforementioned coalition type. Our ML-based models, which are sensitive to context, manage to predict at least some of those examples correctly.

For the 1998 election, $F1$ scores for the baseline and the transfer learning model are also very close. Here the baseline again overpredicts a coalition between CDU/CSU and SPD, as the data contains many references to the *Große Koalition*.

For the elections in 2013 and 2017, the results for the two models are closer. This is most probably due to changes in the distribution of coalition types in the data, which reflect the emergence of new coalition forms that did not exist before. The so-called Jamaica coalition, for example, only received newspaper coverage before the 2009 and 2017 elections. This is a well-known problem for supervised machine learning approaches, for which we propose a solution based on data augmentation, where we create synthetic data points for infrequent labels and add them to the training data.

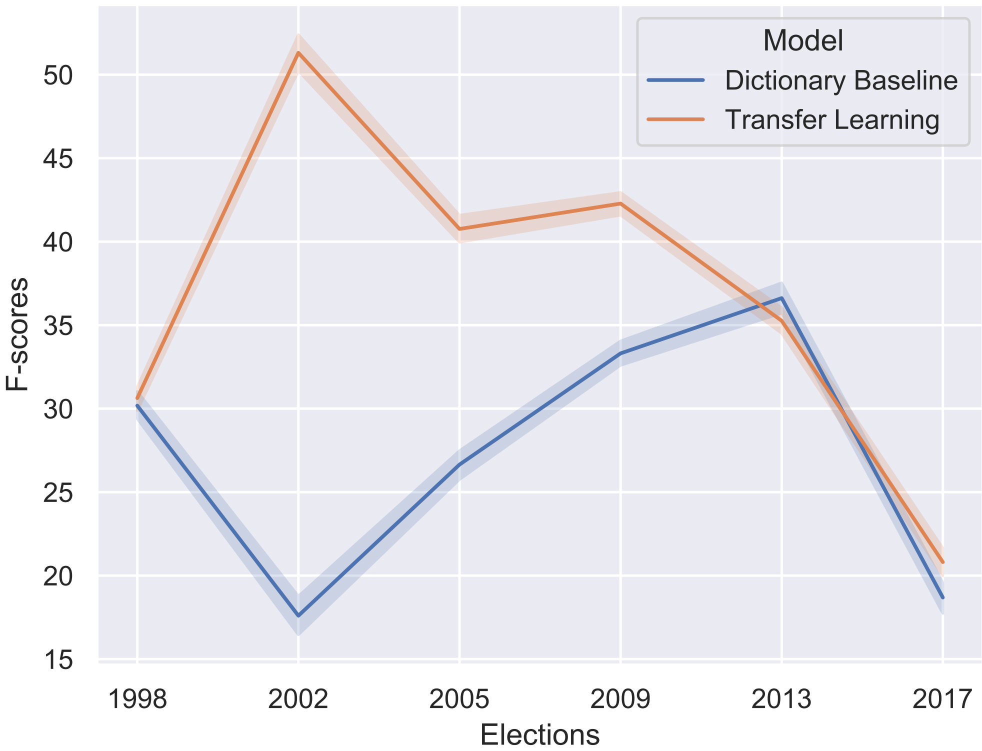
Table A4: Example for training data augmentation based on heuristics and a predefined dictionary

Step	**Coalition type**	**Example**
1	*SPD + FDP + Greens*	Einzig eine Ampel-Koalition mit SPD und Grünen hat die FDP ausgeschlossen.
2	*CDU/CSU + FDP + Greens*	Einzig eine Jamaika-Koalition mit SPD und Grünen hat die FDP ausgeschlossen.
3	*CDU/CSU + FDP + Greens*	Einzig eine Jamaika-Koalition mit CSU und Grünen hat die FDP ausgeschlossen.

Our approach works as follows. We start with seen data instances in the training data and define a list of coalition types that we want to predict (*SPD + FDP*, *CDU/CSU + Greens*, *SPD + LEFT + Greens*, *SPD + LEFT*, *CDU/CSU + FDP + Greens*, *SPD + FDP + Greens*, *SPD + Greens*, *CDU/CSU + SPD*, *CDU/CSU + FDP*, *CDU/CSU + SPD + FDP*). Please note that the list can also include new, unseen coalition types.

For each coalition in the list, we create a dictionary that includes a set of coalition terms that are used to refer to each of the coalition types, and also add the parties that belong to each

Figure A1: Results for each model and election, plotted with a 95% confidence intervals



coalition (e.g., for the coalition type *CDU/CSU + FDP + Greens*, we include the coalition terms 'Jamaika' and 'schwarz-gelb-grün' and party names: CDU/CSU, FDP, Greens in the dictionary; for *CDU/CSU + FDP + Greens*, we include: 'Deutschlandkoalition', 'schwarz-rot-gelb' and the parties: CDU/CSU, SPD, FDP).

Next, we iterate over the training data and check each instance for the presence of the predefined keywords. If we find a keyword, we randomly select another coalition type from our dictionary and replace the original keyword with one that corresponds to the new coalition term (Steps 1 and 2 in Table A4). This allows us to create a new instance for a different label, based on the original training example. However, such created examples are not always meaningful, as the new keyword might be in conflict with the context of the sentence. In our example, the new coalition term does not fit the party names. To fix this, we add a post-processing step where we check whether the party names in the sentence match the ones in our predefined dictionary and, if we encounter conflicts, we replace them with the correct parties (Table A4, Step 3).

Using this approach, we create an additional 1,260 synthetic training instances that help to ameliorate the problem described above by adding data points for sparse coalition types. We refer to this approach as Transfer-augmented.

Tables A5 and A6 show results for the dictionary baseline and for our transfer learning approach without and *with* data augmentation. Overall, results increase from 37.0% to 40.0% $F1$, now outperforming the dictionary baseline by 11 percentage points. Crucially, when looking at results for individual elections (Table A6 and Figure A2), we were able to improve results even for the elections where our original model struggles, due to sparse data for new coalition types.

Table A5: Results for the prediction of coalition signals in newspaper articles

	(A) Signal (yes/no)			(B) Coalition			(C) Coalition, polarity			(D) Aggregation
	Prec	**Rec**	**F1**	**Prec**	**Rec**	**F1**	**Prec**	**Rec**	**F1**	**RMSE**
Dictionary approach	36.0	28.9	32.1	81.0	37.7	51.4	46.3	21.0	28.9	0.0015
Transfer learning	35.2	48.7	40.9	65.5	49.2	56.2	42.6	32.7	37.0	0.0010
Transfer augmented	31.6	54.2	39.9	66.2	55.2	60.2	43.4	37.1	40.0	0.0009

Note: (A) Prediction of coalition signals on the sentence-level (signal: yes/no); (B) prediction of the coalition option on the signal-level; (C) prediction of the coalition option *and* polarity on the signal-level; and (D) prediction of the aggregation measure $\theta^1_{e,c}$.

Table A6: Evaluation of coalition signal prediction ($F1$ for coalition type + polarity) on the signal level for each election (six-fold cross-validation setup)

Election	Dictionary	Transfer	Transfer-augmented
1998	30.04	30.45	40.60
2002	17.98	48.21	46.55
2005	27.12	41.96	41.23
2009	32.96	37.93	43.55
2013	36.26	40.17	40.16
2017	18.32	19.23	25.45
Total	28.89	37.01	40.00

Note: Transfer-exp refers to the transfer learning approach where we trained the BERT model on an expanded dataset with artificially created sentences.

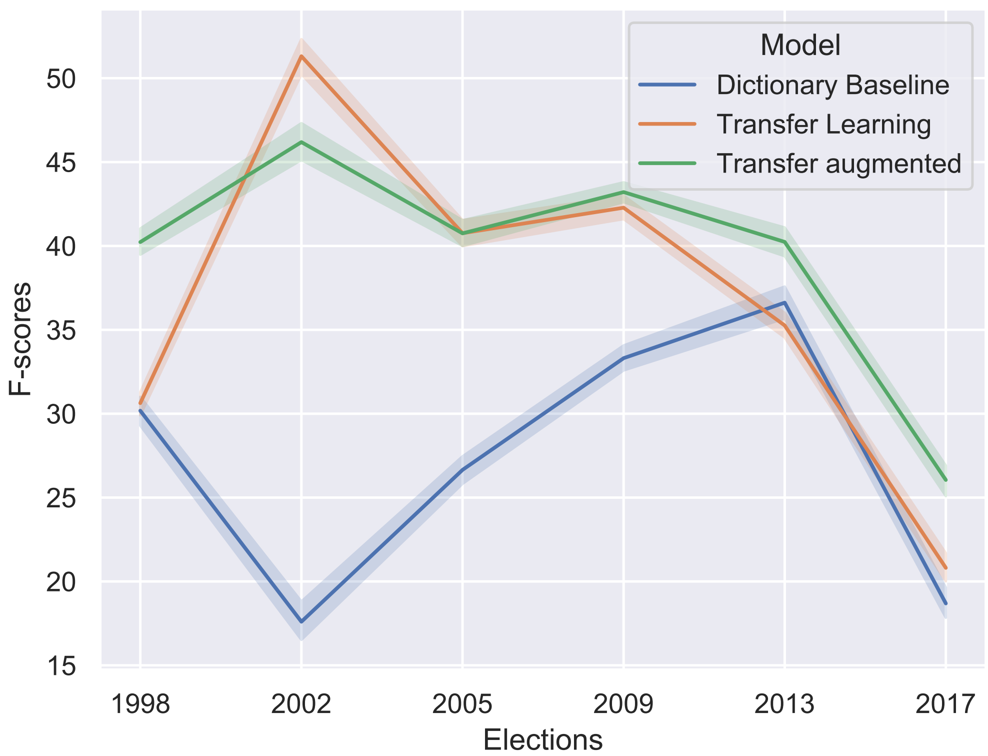
B. Negative Campaigning

Data

Dictionary for target prediction baseline

Table A8 shows the dictionary used for our rule-based target prediction baseline. We iterate over each text sequence in the data and search for dictionary entries and, if we find a match, we consider this as our target and assign the respective party label to the text sequence. We do a greedy search where, in case of multiple party mentions, we always use the first hit, based on the assumption that the target of the press release will appear in a prominent position early on in the text.

Figure A2: Results for each model and election, plotted with a 95% confidence interval



Training details

SVM baseline for target prediction We tokenize the data and extract bag of words features from the press releases, including the title, subtitle and text body of the data. We weigh the extracted features, based on their term frequency–inverse document frequency (TF-IDS). The best experimental settings have been determined, based on a grid search on the development set. We keep stop words and do not lemmatize the data. Then we apply χ^2 feature selection and keep the best 10,000 features for model training.

SVM baseline for stance prediction For stance detection, we apply the same approach as above, with the following changes. We do remove stop words and lemmatize the data before feature extraction. As before, we do χ^2 feature selection and keep the best 10,000 features for model training.

Transfer model for target prediction We use the pre-trained `bert-base-german-cased` model from Huggingface and train 5 models with different initializations on our data. The model learns to predict 5 labels (SPÖ, ÖVP, FPÖ, Greens, NONE). We extract features from the title and subtitle of the press releases but do not use the text body of the releases. We use the following hyperparameter settings (Table A9). We use early stopping, i.e., we evaluate our model on the development set and stop the training process if we do not observe any improvements in loss for 5 subsequent iterations (as improvement, we define a reduction in loss by at least 0.01).

Table A7: Distribution of press releases in the AUNTES data for different elections

year	# texts	w/target	OTHER
2002	2,088	1329	759
2006	2,234	1403	831
2008	2,594	1627	967
2013	1,954	1191	763
2017	1,073	618	455
Total	9,943	6168	3775

Table A8: Dictionary for rule-based target prediction baseline

ID	Party	Keyword
1	SPÖ	SPÖ
2	SPÖ	Sozialdemokratische Partei Österreichs
3	SPÖ	Sozialdemokratischen Partei Österreichs
4	ÖVP	ÖVP
5	ÖVP	Österreichische Volkspartei
6	ÖVP	Österreichischen Volkspartei
7	ÖVP	Neue Volkspartei
8	FPÖ	FPÖ
9	FPÖ	Freiheitliche Partei Österreichs
10	FPÖ	Freiheitlichen Partei Österreichs
11	GREEN	die Grünen
12	GREEN	Die Grünen
13	GREEN	die Grüne Alternative
14	GREEN	Die Grüne Alternative
15	GREEN	Grüne
16	GREEN	Grünen

Transfer model for stance prediction Our setup for stance prediction closely follows the one for target prediction, with the following modifications. We train the model to predict 3 labels (positive, negative, neutral). For feature extraction, we use the title, subtitle and text body of the press releases. The hyperparameters are listed in Table A9.

Table A9: Hyperparameter settings for the transfer learning approach (negative campaigning) for target and stance prediction

Feature	**Target**	**Stance**
batch size	8	8
max. seq length	128	512
max. no. train epochs	50	50
learning rate	1e-4	1e-4
max seq length	128	512
early stopping	yes	yes
early stopping delta	0.01	0.01
early stopping patience	5	5

Table A10: Results for negative campaigning (target prediction) on the AUNTES data set (transfer learning approach without ensemble classification)

Target	**2002**	**2006**	**2008**	**2013**	**2017**	**avg.**	# support
			Transfer learning approach				
SPÖ	0.61	0.65	0.76	0.66	0.51	0.64± 0.09	1,633
ÖVP	0.72	0.75	0.71	0.55	0.70	0.68± 0.08	1,619
FPÖ	0.39	0.41	0.58	0.54	0.53	0.49± 0.08	375
Greens	0.66	0.57	0.65	0.45	0.38	0.54± 0.12	236
OTHER	0.81	0.78	0.80	0.74	0.79	0.78± 0.03	6,078
Total	0.73	0.73	0.76	0.66	0.70	0.70	9,941
SD	± 0.05	± 0.04	± 0.05	± 0.05	± 0.05		

Note: All results are averaged over three runs. Support specifies the number of instances for each class in the test data.