SOURCES OF POLITICIANS' EXPRESSED AFFECTIVE POLARIZATION: IDEOLOGICAL EXTREMITY, ELECTORAL INCENTIVES, AND LEGISLATION PERFORMANCE

By

Siyuan Ma

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Communication – Doctor of Philosophy

2023

ABSTRACT

Politicians' expressed affective polarization has been considered an important but less studied factor that reflects their ideology and further influences emotions in the general public. The current study has explored the changes in politicians' expressed affective polarization level from 2011 to 2022, and examined the influence of politicians' ideology extremity, electoral incentives, and legislative performance on their expressed affective polarization. By collecting Twitter data and applying supervised machine-learning models, the current study has depicted that politicians' expressed affective polarization level increased rapidly in the past 12 years, much faster than the increase of ideological polarization. By analyzing the two-way fixed-effects panel data (N = 796 elites, T = 6 terms in 12 years), the current study has found that ideological extremity positively influences expressed affective polarization, while legislative performance nearly has no influence. Electoral incentives negatively impact the overall expressed affective polarization but positively impact the in-party liking expressions. However, the ideological extremity, electoral incentives, and legislative performance effect sizes are small. Meanwhile, the overall social media expression, the 2016 presidential election, and the power of control in Congress have important influences on politicians' expressed affective polarization. The findings indicate that the media environment, political events, and politicians' status could have a larger impact than expected. Further study directions are also discussed in the dissertation.

ACKNOWLEDGEMENTS

This current study has been supported by the Dissertation Completion Fellowship awarded through the Michigan State University Graduate School.

This dissertation would not have been possible without the help, support, and patience of many people working with me. I would like to thank my advisor, Professor Tai-Quan Peng, and my co-advisor, Professor Daniel Bergan, who have offered me so much advice and encouragement during my entire studying experience at Michigan State University. I want to thank Professor Dustin Carnahan and Professor Kenneth Frank for your great suggestions to improve this project.

I also want to thank Mr. Haoxiang Zhang, Mr. Elijah Peterson, Mr. Yichao Wang, Mr. Shizhe Diao, Dr. Danding Wang, Professor Yi Zhu, Ms. Yue Zhang, and Ms Qi Zheng, for your help and thoughts on this project, especially on the human coding and machine learning parts.

I would also like to more broadly thank the faculty, staff, and colleagues at or outside of Michigan State University for their support, conversations, and suggestions over the past five years.

Finally, I want to express my thanks to my parents, Xiaobo Ma (马晓波) and Zhuo Wang (王卓), for your love and emotional support with me. We have gone through some difficult times together, and I believe our future life will be better.

TABLE OF CONTENTS

INTRODUCTION

In 1796, George Washington warned about "*hyper-partisanship*" in his Farewell Address: "It agitates the community with ill-founded jealousies and false alarms, kindles the animosity of one part against another, foments occasionally riot and insurrection. It opens the door to foreign influence and corruption." Politicians are certainly not prophets - when the founding fathers tried to warn US people about the dangers of polarization, they may not expect to foresee things happen in 200 years later. However, people are witnessing the five overly precise predictions: jealousies, false alarms, animosity, riots, and foreign influence fall to them one after another. In the past decade, the US people have experienced unreasonable filibusters, well-spread misinformation, hate speeches on social media, the capital attack in 2021, and interventions in presidential elections. According to Pew investigations in 2019 and 2022, polarization not only increases partisan hostilities but also relates to the general public's widespread disappointment in politics. 87% of Americans say political polarization is threatening their way of life (Pew Research Center, 2019), 72% of Republicans regard Democrats as more immoral, and 63% of Democrats say the same about Republicans (Pew Research Center, 2022).

Some scholars have proposed that politicians' polarization could not only 1) prompt political stalemate such as filibusters in a party-divided Congress to procrastinate legislation, but also 2) link to the general public's affective polarization, and impel uncivil statements and irrational decisions (e.g., Druckman et al., 2013; McCarthy, 2019; Herbst, 2010; Kim et al., 2021). Thus, political polarization may have the ability to fuel democracy backsliding, such as coups, rigged elections, and executive aggrandizements since 2010 (Boese et al., 2021), and might fuel the high tide of populism across both developing and developed countries (Kyle & Meyer, 2020).

However, the detailed mechanism under the umbrella concept of political polarization is less explored. Previous studies have found that the general public is affectively polarized (Iyengar et al., 2012), and the general public's affective polarization correlates with politicians' ideological polarization (Webster & Abramowitz, 2017). But at the same time, the general public is not familiar with politicians' ideologies and policy preferences (Iyengar et al., 2012; Somin, 2014). Thus, people could observe a phenomenon that the general public could link affective polarization with ideological polarization without knowing what the ideology exactly is. The phenomenon indicates that the general public's affective polarization is not directly influenced by the politicians' ideological polarization, but influenced through the politicians' expressed affective polarization.

Nevertheless, the level of politicians' expressed affective polarization, and the sources of expressed affective polarization lack exploration. This dissertation will first conduct an exploratory and descriptive analysis of the level of politicians' expressed affective polarization, then test sources contributing to their expressed affective polarization. The current study has proposed that the politicians' ideological extremity, electoral incentives, and legislation performance would contribute to their expressed affective polarization. The unit of analysis in the hypotheses is each politician (each individual), but the hypotheses model also considers the group-level influence. Previous studies tended to analyze from a group (party) level and revealed the change of a whole party (e.g., Fiorina et al., 2005; Abramowitz & Saunders, 2008; Lelkes, 2019). However, politicians do not always act as a group, and they also have different personal characteristics (e.g., Fenno, 2003; Mayhew, 2004; Jacobson & Carson, 2019). It is worthwhile to analyze politicians' polarization from an individual level.

To sum up, politicians' expressed affective polarization is important, but researchers lack an understanding of how and why it is expressed. Therefore, politicians' ideological extremity, electoral incentives, legislation performance, expressed affective polarization, and their relationships deserve further study.

See Figure 1 for details about the proposed mechanism.

There are some crucial implications of studying politicians' polarization. Theoretically, this study could contribute to constructing the ideology-affection polarization mechanism. Intuitively speaking, the findings would help understand how politicians' polarized communication practices on social media - which are thought to contribute to further affective polarization in mass audiences (e.g., Herbst, 2010; McCarthy, 2019) - come to be adopted and expressed. Methodologically, the current study could provide a succinct workflow of using social media resources and applying a supervised machine learning approach to analyze and describe changes in politicians' expressed affective polarization.

LITERATURE REVIEW

Fundamental Dimensions of Polarization: Ideological and Affective Polarization

The original concept of *polarization* is derived from *polarity*, which means the "state of having two opposite or contradictory tendencies, opinions, and aspects (Oxford English Dictionary)." Previous studies basically focus on two dimensions of polarization: ideological polarization and affective polarization.

Previous studies first focused on ideological polarization, which referred to the extent that two groups of people with opposing ideologies strengthen their original position (Stroud, 2010). Researchers analyzed US national investigation (ANES) results on people's preference (from liberal to neutral then to conservative) of seven major national policies. They found that the general public had gotten gradually polarized since the 1980s, but the overall polarization level may not be severe (Fiorina et al., 2005; Abramowitz & Saunders, 2008; Prior, 2013). However, by analyzing voting behaviors and bill co-sponsorships, previous studies found clear evidence that politicians, especially federal-level legislators, were polarizing rapidly in the past four decades (Layman et al., 2006; McCarty, 2019; Neal et al., 2020).

Subsequent studies found that many people, especially partisans in the general public, did not clearly understand ideology or policy. Instead, their polarization meant a sense of emotion, including liking the in-group and disliking the out-group (Iyengar et al., 2012), also called affective polarization. Iyengar et al. (2012) used a thermometer ranging from 0 (refers to cold) to 100 (refers to warm) to measure how partisans felt about in-party and out-party. Findings showed that the US partisans' in-party warm feelings did not change much over time, while both two parties' cold feelings towards the out-party had become more extreme over the past four decades (1970s-2010s). Scholars also used indirect questions to measure affective polarization.

They found that American partisans had become increasingly averse to the prospect of their child marrying someone from the opposing party (Iyengar et al., 2012), and partisans would avoid living with a member of the other party at roughly the same level as they were avoiding someone described as “not at all clean and tidy” (Druckman & Levy, 2021).

In summary, previous studies have found that while politicians are getting ideologically polarized, at the same time, the general public is getting more affectively polarized than ideologically polarized. However, scholars have limited knowledge about politicians’ affections. Thus, the current study aims to assess politicians’ expressed affective polarization.

Politicians’ Expressed Affective Polarization

Many previous studies have studied the general public’s ideological and affective polarization, as well as the politicians’ ideological polarization. However, the politicians’ affective polarization is less studied, partly due to the difficulties of investigating them.

Nowadays, the widespread usage of social media provides opportunities to access politicians’ daily activities and study their affective polarization from one specific perspective: the affective polarization expressed by politicians. The expressed affective polarization is defined as the extent to which politicians express a favorable sentiment toward the political in-group (in-party liking) or a negative sentiment toward the political out-group (out-party disliking). The in-party liking refers to emotional support towards 1) the party an individual belongs to, or 2) the political characters from the same party; and out-party disliking refers to an individual’s emotional disapproval of 1) the opposite party, or 2) the opposite-party political characters.

Expressed affective polarization is the publicly available part of affective polarization. This concept deserves deeper study because it is an important component for the politicians to communicate with the general public. Politicians need to efficiently communicate with the general public because 1) it is politicians' duty to report legislation progress and explain policies to the citizens they serve (e.g., Fenno, 2003; Mayhew, 2004; Jacobson & Carson, 2019), and 2) politicians need to gain constituents' support through communication (e.g., Fenno, 2003; Mayhew, 2004; Jacobson & Carson, 2019). Meanwhile, previous studies have found that the general public are not familiar with policies or ideologies, but are more familiar with political characters (Iyengar et al., 2012; Somin, 2014). Therefore, the politicians' expressed affective polarization has played an efficient role in the politician-public communication, because it contains the sentiment towards political characters that are easy for the general public to understand. In addition, previous studies have also found that people's affective polarization has relationship with their ideological polarization (e.g., Webster & Abramowitz, 2017; Abramowitz, 2021; Hobolt et al., 2020), which indicates that the politicians' expressed affective polarization may to some extent reflect their ideology. That makes the expressed affective polarization even more interesting. Note that the expressed affective polarization may embed with policy evaluations, but the fundamental part of expressed affective polarization is it must mention political characters.

The concept of expressed affective polarization is somewhat similar to some other concepts, such as polarizing language or uncivil statements. However, the expressed affective polarization has a unique meaning that best describes a broad group of emotional and extreme information that other concepts do not summarize.

First, the polarizing language cannot represent what the current study has measured. According to previous studies, the polarizing language, by definition, could aim at either the character or the issue (West, 2017; Ballard et al., 2022). But the expressed affective polarization, by definition, does not include tweets that solely aim at a political issue without mentioning any political characters. Here are the reasons for excluding solely issue-related polarizing tweets. First, previous studies measured politicians' ideological polarization by the bill co-sponsorship or the statement on a political issue (Rogowski & Sutherland, 2016; Webster & Abramowitz, 2017). Politicians' solely political-issue-based tweets are also a kind of statement. Thus, it is not surprising that politicians' expression of solely political-issue-based tweets correlates with their performance in Congress. Secondly, previous studies have also found that politicians' ideological polarization correlated with the general public's affective polarization (Rogowski & Sutherland, 2016; Webster & Abramowitz, 2017). Thus, it is also not surprising to predict that politicians' expression of solely political-issue-based tweets may influence the general public's affective polarization. Thirdly, previous studies have already done some general, descriptive analyses on polarizing tweets broadly considering issues or characters. For instance, previous studies have found that more ideologically extreme members, those from safer districts, and those who are not in the president's party are more likely to send polarizing tweets (Ballard et al., 2022). Also, the polarizing tweets gain more engagement and political antipathy from partisans in the general public (e.g., Hong & Kim, 2016; Ballard et al., 2022). Because previous studies have analyzed the sources and the impact of the overall polarizing language and the solely political-issue-based expressions, the current study wants to go more specific and test affective polarization that contains political characters.

In addition, the current study has also defined the expressed affective polarization differently from uncivil statements. The expressed affective polarization emphasizes the emotional attitude towards a political character (party or people). The emotion in expressed affective polarization could be either positive or negative, while the uncivil expressions could only capture intemperate and improper expressions. However, politicians do not always need, and in fact they are not encouraged, to express uncivil statements (Theocharis et al., 2020; Unkel & Kumpel, 2022). Politicians can attract people, make criticisms, by using civil words -- some expressions can be both polite and sarcastic.

To sum up, the concept of expressed affective polarization deserves much more attention because it is a crucial component for the politicians to communicate with the general public. It could also be a vivid exhibition of the politicians' ideology, and link to the general public's affective polarization. In order to know how politicians' expressed affective polarization is influenced by ideology and may further influence the general public's affective polarization, the current study has proposed that it is necessary to first conduct an exploratory and descriptive analysis of politicians' expressed affective polarization. In addition, polarization is defined more as a trend rather than a fixed status (Lelkes, 2016); the changing level of polarization over time itself could contribute to depicting society. Thus, the current study would analyze the level of expressed affective polarization by congressional terms and display the changes. The current study proposes the first research question:

RQ1. How do politicians' expressed affective polarization level change over time?

Sub-dimensions of Expressed Affective Polarization

To better depict and understand politicians' ways of communication, the expressed affective polarization should be analyzed in more detailed sub-dimensions.

First, the expressed affective polarization needs a detailed description from the group identity dimension. Previous studies have proposed that in-party liking and out-party disliking contribute differently to the overall affective polarization. For example, in-party liking contributes most to affective polarization in non-political settings, and out-party disliking contributes most to affective polarization in political settings (Rudolph & Hetherington, 2021; Amira et al., 2019). To re-examine this finding under a political-elite context, the current study should consider the in-party and out-party separately.

Second, the expressed affective polarization also needs a detailed description from the content evaluation dimension – to be specific, whether the affective polarization is expressed towards the people or policy. In a complete political spectrum, politicians from far-liberal and far-conservative sides have extremely different ideological preferences (Sznajd-Weron & Sznajd, 2005). However, those elites can work as colleagues in the same political system. For instance, communism-leaning politicians (far-liberal) and new-fascist-leaning politicians (far-conservative) have to work together in the European Parliament (The Political Groups of the European Parliament, 2019). Compared to Europe, previous studies have shown that US politicians have fewer ideological conflicts but have more emotional hostility toward their opponents (McCarty et al., 2016). It indicates that the expressed affective polarization of US politicians is not always embedded with logical policy debates or saying "*America's future*" (Martin & Burns, 2023) but is also related to solely attacks or boasts[1] on the political characters. However, a character-only expression is more harmful than a policy-embedded expression (Marchal, 2022): Policy-embedded expression could at least provide some logical clarifications about policies, so that provide clear and distinct electoral choices for constituents (Rogowski, 2014) and increase voter turnouts (e.g., Hetherington, 2008; Abramowitz, 2010). In contrast, character-only expression may not even have this positive contribution of policy clarification. Thus, it is also worthwhile to differentiate the content evaluation types and study how policy-embedded polarization expression and character-only polarization expression change over time

[1] Here boast means simply expressing positive attitude towards political characters without mentioning any policy reasons.

and how much they are influenced by ideological extremity, electoral incentives, and legislative performance.

The current study defines policy-embedded expressed affective polarization as the extent to which a politician has expressed in favor of the in-party liking or out-party disliking along with policy evaluations. Policy evaluation means mentioning the pros or cons, or neutrally mentioning the impact of a policy. The character-only expressed affective polarization is defined as the extent to which a politician has expressed in favor of the in-party liking or out-party disliking with solely emotional evaluations toward political characters. The character-only emotional evaluation could be further defined and broken up into 1) attacking or boasting party names and party-affiliated organizations and 2) attacking or boasting politicians.

To combine the group identity dimension and content evaluation dimension together, the current study proposes the following research questions:

RQ2a. How does politicians' policy-embedded in-party liking level change over time?

RQ2b. How does politicians' policy-embedded out-party disliking level change over time?

RQ2c. How does politicians' character-only in-party liking level change over time?

RQ2d. How does politicians' character-only out-party disliking level change over time?

Relationship between Ideological Extremity and Expressed Affective Polarization

To take one step further, understanding the roots of expressed affective polarization contributes to explicating the polarization mechanism of politicians. There are two different explanations for the mechanism of polarization. First, traditional studies believed that the affective polarization was originated from social identification. In the US political context, affective polarization is triggered by one of the crucial social identities, partisanship (Iyengar et al., 2012), and ideological extremity is not a necessary condition for affective polarization (Iyengar et al., 2012; Mason, 2015). However, recent but less-mentioned studies have proposed that affective polarization has its rational root and links tightly with ideological extremity, especially among people with more political knowledge (e.g., Webster & Abramowitz, 2017; Abramowitz, 2021; Hobolt et al., 2021). For instance, scholars want to analyze people's perceptions of a candidate who has proposed policies that they disagree with. While showing that the candidate was from an opposite party, respondents felt moderately "cold" about him/her. While showing that the candidate held an extremely different ideology, respondents felt extremely "cold" about him/her. In other words, respondents reacted far more strongly to ideology cues than party cues, especially if it was the ideology of the member of the out-party (Lelkes, 2019). Webster & Abramowitz (2017) also find evidence through an online experiment that feelings about the opposing party and its leaders are strongly related to social welfare policy preferences. The results of this online experiment show that there is a rational basis for voters' emotional responses to political leaders and groups—negative affect is based to a large extent on policy disagreement (Webster & Abramowitz, 2017). In addition, the relationship between ideology extremity and affective polarization is not only found in the US but also in European countries (Hobolt et al., 2021). For instance, the more people agree with one specific policy (e.g., Brexit, Catalan independence), the more they would like people who have similar policy preferences and dislike people who support the opposite policy.

Since expressed affective polarization is one unique perspective of affective polarization, we could expect there may be a connection between expressed affective polarization and ideological extremity. According to the rational assumption of politicians, especially some evidence showing elites are more rational than the general public, we should not expect politicians themselves to be "ideologically innocent" when they choose to express themselves in political settings (Amadae, 2021; Kinder & Kalmoe, 2017). Instead, we should expect politicians' expressed affective polarization to be influenced by their own ideology. Studies among politicians have also provided evidence that ideologically extreme politicians would reject the characteristics of the out-party or express more extreme words than the moderate elites (e.g., Pew Research Center, 2020; Ridge, 2020; Ballard et al., 2022).

Therefore, ideological extremity should positively influence politicians' expressed affective polarization. In order to examine this less-mentioned polarization mechanism explanation in the US politicians' context, it is necessary to study the politicians' relationship between ideological extremity and expressed affective polarization. The current study proposes the hypothesis:

H1: Politicians' ideological extremity will positively influence the expressed affective polarization on social media.

The current study has specified the expressed affective polarization in several detailed situations by combining and considering sub-dimensions of group identity and content evaluation. Therefore, the current study will test the relationship in more detail and proposes the hypotheses:

H1-a: Politicians' ideological extremity will positively influence the policy-embedded in-party liking on social media.

H1-b: Politicians' ideological extremity will positively influence the policy-embedded out-party disliking on social media.

H1-c: Politicians' ideological extremity will positively influence the character-only in-party liking on social media.

H1-d: Politicians' ideological extremity will positively influence the character-only out-party disliking on social media.

Relationship between Electoral Incentives and Expressed Affective Polarization

The electoral incentive is defined as the threat that a politician has received of **losing** an election (Auter & Fine, 2016; Kalla & Porter, 2020). Politicians' electoral incentive is another important perspective influencing expressed affective polarization. Under the system of US representative democracy, politicians (which refer to federal legislators in the current study) need to shape their communication patterns to fit constituents' demands so that they could get more votes and have a better chance of winning the election (e.g., Fenno, 2003; Mayhew, 2004; Jacobson & Carson, 2019). Along with the party-sorting tendency of constituents' voting behavior, politicians should be more polarized when they need more votes from their supportive constituents.

Scholars have proposed that Congress members pursue some principal goals: winning elections, influence within the House, and contributing to good public policy (Fenno, 2003). Winning the election is fundamental before influencing the House or making policy. The *election* includes the general election (mostly a bi-partisan election) and the primary election (competing for an in-party nomination).

The current study first discusses the general election. To win the election, Congress members must gain constituents' vote support and be responsive to constituents' policy appeals. Half a century ago, constituents concentrated more on local political issues than national ones and valued Congress members' credibility and service to the constituency (e.g., Fenno, 2003; Mayhew, 2004). Thus, in the 1970s, Congress members needed to cautiously maintain personal (homestyle) relationships with their constituencies and build high credibility by fighting for local interests. Although Congress members also needed to explain national policies and standpoints of parties, their individual characters played a much more important role than the party during the election. Thus, Mayhew (2004) and Fenno (2003) proposed that studies of the US Congress must focus on individuals rather than groups of individuals, such as parties.

However, the party has now played an increasing role in Congress members' elections (Lawrence et al., 2006; Curry & Lee, 2019) because constituents have focused more on national policies and voted by their party affiliation (Bartels, 2000; Jacobson & Carson, 2019). It does not mean that Congress members' characteristics, performance, resources, and domestic political issues do not influence elections, but Congress members' attitudes toward national policy and their party affiliation's influence on elections grow rapidly. For instance, literature has found that the better the economy performs, the better the congressional candidates of the president's party do in an election (Sides & Haselswerdt, 2019; Jacobson & Carson, 2019). Another study has found that the better the presidential candidate performs, the better the congressional candidates of the president's party do in an election (Sievert & McKee, 2018; Jacobson & Carson, 2019). National economy and presidential performance should not directly relate to local issues; however, they all impact Congress members' election vote share now.

Constituents' demands have changed, but Congress members' political pursuits do not change. They must still respond to constituents' new demands and explain national policies to gain more votes. To respond to constituents' recent-developed attention to national policies and party affiliation, Congress members must show their efforts in national policy-making. Especially, donors in the district – which are mostly strong partisan constituents - have the ability to demand policy responsiveness in exchange for vital primary election resources (Kujala, 2020). The tendency to show efforts on national policy is even strengthened by constituents' interests in national policies and partisanships. Because constituents' current interests indicate that Congress members' individual characteristics are less important, while the party background is more important during the election. Politicians are no longer unique but replaceable. To win the election, Congress members would rely more on the party's resources and support. Thus, party leaders could better control party members in Congress, leading to more cohesive and polarized voting in Congress (Jacobson & Carson, 2019).

To maintain constituents' vote support, Congress members must align their communication with constituents' new demands on national policy and party affiliation. This alignment is constructed by two basic tasks: build support for themselves and undermine their challengers' influence. The first task is insufficient without the second (Jacobson & Carson, 2019; Russell, 2021). In the beginning, challengers are more likely to attack the incumbents, while the incumbents are less likely to engage. Since the 1990s, the campaign has become more negative (e.g., Lau & Pomper, 2001; Geer, 2012). The in-office politicians also attack the challengers to win the seat in Congress. In summary, in response to constituents' preferences, Congress members' communication becomes negative and polarized.

Social media platforms like Twitter have recently inherited polarized and negative communication patterns from offline communication. Twitter is the most popular social media platform for congressional elections. Nearly all Congress members have at least one Twitter account (Pew Research Center, 2020). Congress members' communication pattern is considered as a"digital homestyle" (Russell, 2021). However, this "digital homestyle" does not allow Congress members to communicate differently and separately to different targeted groups - which the traditional homestyle communication often did - because Twitter is more suitable for broadcasting information than interpersonal communication (Jungherr, 2016; Vesko & Trilling, 2019). On this nationwide social media platform, Congress members cannot shape their expressions to different interest groups, and they are not only talking to people in the district but also people who treat them as legislative leaders across the nation. Therefore, Congress members' expressions on social media platforms have to focus more on national issues and become more negative and polarized to attract more constituents' vote support. The more the Congress members' electoral incentive (feel the threat of losing the election), the more they need to respond to constituents' preferences, and then the more they need to be polarized in expression.

Therefore, the current study proposes that politicians' general election incentive should positively influence their expressed affective polarization. The current study hypothesizes that:

H2: Politicians' general electoral incentive will positively influence the expressed affective polarization on social media.

The current study has specified the expressed affective polarization in several detailed situations by combining and considering sub-dimensions of group identity and content evaluation. Therefore, the current study will test the relationship in more detail and proposes the hypotheses:

H2-a: Politicians' general electoral incentive will positively influence the policy-embedded in-party liking on social media.

H2-b: Politicians' general electoral incentive will positively influence the policy-embedded out-party disliking on social media.

H2-c: Politicians' general electoral incentive will positively influence the character-only in-party liking on social media.

H2-d: Politicians' general electoral incentive will positively influence the character-only out-party disliking on social media.

Moreover, election incentive also exists in the primary election. Politicians' expressed affective polarization is also influenced by some "core" supportive constituents, such as strong partisan voters in the district. The conventional idea proposes that primary elections are platforms to exhibit the influence of core supportive constituents. The threat of a primary election drives Congress members to diverge from the general electorate to take positions closer to their loyal partisan constituency (McCarty et al., 2009; Stone & Simas, 2010). Therefore, previous studies have found that fifty-seven percent of nominees are more extreme than the average copartisan in their district, including 65% of Democrats and 49% of Republicans (Kujala, 2020). We can indicate that the more Congress members need support during the primary election, the more they will express affective polarization.

Similarly, the current study proposes that the politicians' primary election incentive should positively influence their expressed affective polarization. The current study hypothesizes that:

H3: Politicians' primary electoral incentive will positively influence the expressed affective polarization on social media.

The current study has specified the expressed affective polarization in several detailed situations by combining and considering sub-dimensions of group identity and content evaluation. Therefore, the current study will test the relationship in more detail and proposes the hypotheses:

H3-a: Politicians' primary electoral incentive will positively influence the policy-embedded in-party liking on social media.

H3-b: Politicians' primary electoral incentive will positively influence the policy-embedded out-party disliking on social media.

H3-c: Politicians' primary electoral incentive will positively influence the character-only in-party liking on social media.

H3-d: Politicians' primary electoral incentive will positively influence the character-only out-party disliking on social media.

To sum up, constituents' support is the foundation of Congress members to win the election and continue their careers, and constituents nowadays focus on national policies and party affiliations. Therefore, constituents would nominate candidates who can represent their party and policy positions. This tendency might be enhanced as candidates are increasingly aware that they win the election with a particular party and policy position. Thus, they should affectively express similar polarization levels as an instrumental response to maintain constituents' support.

Relationship between Legislation Performance and Expressed Affective Polarization

Politicians' legislation performance would be defined as their contribution to policy/law-making. Previous studies on governance and political support have indicated that 1) winning an election and 2) maintaining a good legislation performance reputation are major goals of politicians (Fenno, 2003; Linde & Dahlberg, 2021). Because winning the election guarantees a position in the democratic political system for politicians, a good legislation reputation could strengthen people's support and the political system's legitimacy. The current study has already hypothesized that electoral incentives could positively relate to expressed affective polarization because politicians must cater to constituents' demands. However, the legislation performance could be a calm power that contributes differently to expressed affective polarization. Thus, studying politicians' legislative performance as another indicator is worthwhile.

Russell (2021) has proposed three styles of politicians, policy wonks, constituent servants, and partisan warriors. Policy wonks care most about their national policy legislative performance, while partisan warriors care least about their legislative performance. Politicians who care about legislative performance, such as policy wonks, also value the art of compromise because they often need support from opposition party members (Russell, 2021). To compromise is a signal of making real efforts on legislation. Thus, politicians focusing more on legislation are motivated to reduce their affective polarization expression on social media because less polarization could show their reputation as policy wonks and exhibit their legislative efforts and transparency to the general public (McConnell & Hart, 2019; Porumbescu et al., 2017). Therefore, the current study proposes hypotheses:

H4: Politicians' legislative performance will negatively influence the expressed affective polarization on social media.

Similarly, the current study has specified the expressed affective polarization in several detailed situations by combining and considering sub-dimensions of group identity and content evaluation. Therefore, the current study will test the relationship in more detail and proposes the hypotheses:

H4a: Politicians' legislative performance will negatively influence the policy-embedded in-party liking on social media.

H4b: Politicians' legislative performance will negatively influence the policy-embedded out-party liking on social media.

H4c: Politicians' legislative performance will negatively influence the character-only in-party liking on social media.

H4d: Politicians' legislative performance will negatively influence the character-only out-party disliking on social media.

The hypotheses model is summarized. See Figure 2 for details.

METHODS AND MEASUREMENTS

The current study's analyses follow the listed procedure: 1) collect politician tweets to calculate each US politician's levels of expressed affective polarization, policy-embedded in-party liking, policy-embedded out-party disliking, character-only in-party liking, and character-only out-party disliking. 2) collect data on each politician's ideological extremity, electoral incentives, and legislation performance. 3) depict the expressed affective polarization, policy-embedded in-party liking, policy-embedded out-party disliking, character-only in-party liking, and character-only out-party disliking changes over time. 4) use the two-way fixed-effects panel data models to test the influence of ideological extremity, electoral incentives, and legislation performance on expressed affective polarization and the sub-dimensions.

Data Collection and Cleaning

The current study focuses on US politicians' expressed affective polarization on social media and its sources of ideological extremity, electoral incentives, and legislation performance. In order to provide analyses on research questions and test the hypotheses, the current study has collected Twitter data for the dependent variables and third-person platform data for the independent variables. Note that the politicians specifically refer to all US federal-level legislators once in office from 2011 to 2022. Politicians do not include US overseas territory legislators. Social media refers explicitly to the Twitter platform.

First, the current study applied the data scraping software provided on the third-person platform (Shaikh, 2021) to collect all original tweets, replies, and retweets with comments of 112^{th}-117^{th} US federal-level legislators from 2011 to 2022. This software only got public data available to the unauthenticated user and did not hold the capability to scrape anything private. The following procedure collected politicians' Twitter accounts: researchers started with the handles provided by the official or academic websites (e.g., https://ucsd.libguides.com/congress_twitter/senators; https://pressgallery.house.gov/member-data/members-official-twitter-handles), then searched each politician's full name on the Twitter platform. The verified accounts with blue check marks provided by the platform were collected. Archived or private accounts were not collected. Except for official accounts, some politicians also have personal accounts. Researchers also collected uncertified real-name personal accounts if more than three official accounts followed the unverified account. The current study retrieved tweets when the politicians were in office. The current study focused on the textual parts of the tweets; Emojis, picture links, and webpage links are excluded, while hashtags are included.

To sum up, the current research collected 818 politicians with 1452 Twitter accounts and 4,595,752 tweets in six terms (2011-2022). Each politician in each term is one data point. Among all data points (N=2865): 560 of them are Senators in their Terms, 2305 of them are Representatives in their Terms; 503 of them are First Incumbent legislators in their Terms, 2362 of them are Re-elected legislators in their terms; 1397 of them are Democrats in their Terms, 1468 of them are Republicans in their Terms.

For the independent variables, researchers collected ideology extremity scores from Lewis et al. (2023)'s Voteview project, collected general and primary electoral incentive scores from New York Times election report websites (e.g., New York Primary Results, 2010) and Secretary of State websites (e.g., Alabama Secretary of State | 2010 Election Information, 2010), and collected legislative performance scores and the control variables data from Volden and Wiseman (2018)'s Legislative Performance project. Measurements of these variables are listed below.

Measurements

Politicians' Expressed Affective Polarization

We could use the measurement of affective polarization (e.g., Iyengar et al., 2012; Ballard et al., 2022) for reference to measure expressed affective polarization. The thermometer measurement in a survey design is the widely applied way of affective polarization measurement. However, we cannot use survey investigations to measure the politicians' expressed affective polarization level. A possible way to measure elites' expressed affective polarization level is to use observational data, such as transcripts of their congress/parliament speeches or social media posts. Social media has increasingly become an important communication channel among politicians and between politicians and the general public. Therefore, the current study uses social media posts (Twitter posts) to measure expressed affective polarization.

Previous studies have explored three measurements with social media data that relate to expressed affective polarization (Yu et al., 2021; Borrelli et al., 2021; Ballard et al., 2022). The three measurements all started from tweet-level identification of polarization and synthesized the tweet-level findings to a higher-level polarization. However, their works still have space for improvement. First, researchers did not find a precise enough way to identify the tweet-level sentiment. For instance, Ballard et al. (2022) applied human coding as the gold standard for identifying sentiment, but the intercoder reliability of human coding was not very good. Borrelli et al. (2021) can save more time identifying tweet sentiment with a dictionary-based approach but does not contain a human-coding validation. Second, both Yu et al. (2021) and Borrelli et al. (2021) synthesized the sentiment of tweets to a group (team, party) level polarization instead of an individual (politician) level, and did not show the change in polarization level with a time sequence. (Appendix D describes the three measurements in detail.)

Therefore, the current study conducted a supervised machine-learning model to identify tweet-level polarization. The goal is to find a concise and precise way to measure tweet-level polarization with acceptable internal validity. Then the current study synthesized tweet-level polarization to individual-level expressed affective polarization by term and described the termly changes in politicians' expressed affective polarization from 2011 to 2022.

In this study, the individual-level expressed affective polarization is defined as the extent to which a politician expresses a positive attitude toward the political in-group (in-party) or a negative attitude toward the political out-group (out-party). It is measured by each legislator's number of tweets that expressed affective polarization in each term. This study has applied a supervised machine learning model to identify whether each tweet expressed affective polarization. The human coders' intercoder reliability and machine learning's F1 scores have been reported.

Policy-Embedded In-Party Liking

Policy-embedded in-party liking is defined as the extent to which a politician has expressed in favor of the in-party liking along with policy evaluations. In-party liking indicates that the politician's tweet contains a positive attitude towards same-party characters. Policy embed indicates the politician's tweet has mentioned bills, executive orders, administrative plans, or thoughts on social problems. Policy-embedded in-party liking is operationalized as each legislator's number of tweets showing a positive attitude toward their own party along with comments on policies in each term.

For instance, a tweet saying, "I like the POTUS because B3W benefits the world's development and democracy," is an example of policy-embedded in-party liking. A tweet saying "*** is by no means the best president" is not an example of policy-embedded in-party liking.

Character-Only In-Party Liking

Character-only in-party liking is defined as the extent to which a politician has expressed in favor of the in-party liking with pure emotional evaluations towards political characters (e.g., party, people). Character-only in-party liking is operationalized as each legislator's number of tweets showing a positive attitude towards their own party without mentioning any policy in each term.

For instance, a tweet saying, "*** will be the most suitable legislator for our district and I love him" is an example of character-only in-party liking. A tweet saying, "I fully respect POTUS because he acted quickly to solve the broader crisis," is not an example of character-only in-party liking.

Policy-Embedded Out-Party Disliking

Policy-embedded out-party disliking is defined as the extent to which a politician has expressed in favor of the out-party disliking along with policy evaluations. Out-party disliking indicates that the politician's tweet has contained a negative attitude towards opposite-party characters. Policy-embedded out-party disliking is operationalized as each legislator's number of tweets showing a negative attitude towards their opposite party along with comments on policies in each term.

For instance, a tweet saying, "The POTUS is terrible because B3W would increase 100 billion deficits in 2023" is an example of policy-embedded out-party disliking. A tweet saying, "*** is the worst president I have ever seen," is not an example of policy-embedded out-party disliking.

Character-only Out-Party Disliking

Character-only in-party liking is defined as the extent to which a politician has expressed in favor of the out-party disliking with pure emotional evaluations towards political characters (e.g., party, people). Character-only out-party disliking is operationalized as each legislator's number of tweets showing a negative attitude towards their opposite party without mentioning any policy in each term.

For instance, a tweet saying, "Joe Biden is by no means the worst president I have ever seen, and he must resign," is an example of character-only out-party disliking. In this example, the poster attacked a political character but did not mention any bills, executive orders, administrative plans, or thoughts on social problems.

Ideological Extremity

Ideological extremity refers to the extent that an individual has identified their position on ideology. If aggregating the individual-level ideological extremity to a societal level, we would observe a picture of ideological polarization among people (Stroud, 2010). The ideological extremity is measured by the first-dimension DW-nominate score. Each politician (refers to the US legislator in the current study) has one specific DW-nominate score each term. The first-dimension DW-nominate score is an estimate of ideology that places lawmakers on a Democratic (-1) vs. Republican (+1) scale based on their voting decisions. Since the current study examines the extent of extremity instead of the ideology stance, it will take the absolute value of first-dimension DW-nominate scores as the ideological extremity.

General Electoral Incentive and Primary Electoral Incentive

The electoral incentive is defined as the threat that a politician has received of **losing** an election (Auter & Fine, 2016; Kalla & Porter, 2020). A politician who faces the danger of losing the election will receive more electoral incentives. The current study uses a reversed value of vote shares to measure the general and primary electoral incentives. For the general electoral incentive, the vote share data are collected from Volden and Wiseman (2018)'s legislative performance project at Vanderbilt University. For the primary electoral incentive, the vote share data are collected from New York Times election report websites from 2010 to 2020 and the Secretary of State Offical websites. Some politicians did not experience primary elections because they were nominated by the Governor or temporarily recruited by their party. The current study excluded these data points. The less vote share that a politician has received, the more electoral incentive they would feel.

Legislation Performance

Politicians' legislation performance would be defined as their contribution to policy/law-making. The current study plans to calculate each politician's level of legislation performance by the Legislative Effectiveness Score (LES) developed by Volden and Wiseman (2018). The LES draws on fifteen indicators that collectively capture the proven ability of a legislator to advance her agenda items through the legislative process and into law. More specifically, 1) the LES identifies the number of bills that each member of the House of Representatives sponsored (BILL) and the number of those bills that received any action in committee (AIC), or action beyond committee (ABC) on the floor of the House. For those bills that received any action beyond committee, LES also identifies how many of those bills subsequently passed the House (PASS) and how many became law (LAW). 2) the LES categorizes all bills in order of importance as being either commemorative (C), substantive (S), or substantive and significant (SS). For each of these three categories of bills, the researchers relied on the five important stages of the legislative process (above) to produce the final set of fifteen indicators.

Researchers have already calculated the LES of each legislator from the 112th to 117th Congress. The current study will apply the LES as an interval variable from 0 to 11.

Control Variables

According to previous literature, the current study has proposed several control variables for the model. These variables may not directly influence politicians' expressed affective polarization but have been considered influential on politicians' congressional behavior and online social media expressions. These control variables include 1) congressional identity, which literature has found that senators are more likely to express extreme attitudes online than House representatives (e.g., Haber, 2011) because they do not need to consider some parts of the local constituents too much. The current study has coded Senator =1 and House Representative =0. 2) power of control, which literature has found that minority party members use social media more frequently because they get less attention from traditional media platforms (e.g., Lassen & Brown, 2011; Russell, 2018). The current study has coded the Majority party =1 and the Minority party =0. 3) seniority, which literature has found that first-term members are more aggressive online and more likely to use social media to communicate instead of emails (e.g., Evans et al., 2014; Blum et al., 2022). The current study has coded Re-elected politician =1 and the First incumbent politician =0. 4) Overall social media expression, indicating that as the total number of tweets increases, so does the number of polarized tweets. The overall social media expression is measured by each politician's total number of tweets per term.

Some other variables are also worth controlling, including 5) gender, which literature has found women legislators communicate more on average than their male counterparts (e.g., Evans et al., 2014; Blum et al., 2022) and embrace more partisan behaviors than expected (e.g., Evans et al., 2014; Clark & Evans, 2020). 6) party affiliation, which literature has found democrats much prefer Twitter, and republicans are more extreme in online expression (e.g., Russell, 2018; Blum et al., 2022). 7) ethnicity, which literature has found less represented ethnicity groups, such as African American or Latino members, communicate more than average (e.g., Tillery, 2019; Blum et al., 2022). However, variables 4, 5, and 6 do not vary across time. In the two-way fixed-effect panel data model, the influence of the three variables has been controlled by the individual effect α_i, and the time effect λ_t. Thus, control variables 5, 6, and 7 are not shown in the following hypotheses model.

Content Analysis of Tweets: Human Coding Step

The current study needs to conduct a content analysis of the tweets to identify tweets that contribute to expressed affective polarization and then specify if these tweets belong to policy-embedded in-party liking, policy-embedded out-party disliking, character-only in-party liking, and character-only out-party disliking. The current study recruited human coders to construct a training dataset and then applied a machine-learning model to label all the collected tweets.

In the human coding step, the current study designed a coding scheme to help identify tweets contributing to expressed affective polarization, policy-embedded in-party liking, policy-embedded out-party disliking, character-only in-party liking, or character-only out-party disliking.

Appendix C shows details of the coding scheme.

After designing the coding scheme, two researchers with Communication and Political Science backgrounds worked together to code the tweets. Two coders randomly selected one tweet from one account each time and finally got 35 tweets from each Twitter account to construct a dataset. For each round of training, coders randomly selected 100 tweets from the constructed dataset to code. From the first round of training to the fifth round, two coders' intercoder reliability gradually increased. In the sixth round of training, two coders selected 300 tweets and reached good consistency in intercoder reliability. The intercoder reliability (Cohen's Kappa) threshold is above 0.6 (Landis & Koch, 1977), but the current study's Cohen's Kappa of expressed affective polarization = 0.86 (N=300), Cohen's Kappa of policy-embedded in-party liking = 0.81 (N=300), Cohen's Kappa of policy-embedded out-party disliking = 0.88 (N=300), Cohen's Kappa of character-only in-party liking = 0.87 (N=300), Cohen's Kappa of character-only out-party disliking = 0.66 (N=300). Then the two coders coded the constructed dataset together to build the training set for further machine-learning models. Because many tweets did not contain any kind of expressed affective polarization and might influence the performance of machine learning, the current study excluded some tweets that did not contain any polarization from the training set. Thus, the current study artificially enhanced the percentage of polarization tweets. Finally, the current study got a training set with N =39,814 tweets, including 7.4% policy-embedded in-party liking tweets, 7.4% policy-embedded out-party disliking tweets, 5.5% character-only in-party liking tweets, and 19.1% overall expressed affective polarization tweets.

In addition, the coders have found that the number of character-only out-party disliking tweets is much sparser than predicted, which contained less than 0.4% among all Twitter posts. Methodologically, both the human coding and machine learning results would be unstable with such sparse data. Because the character-only out-party disliking tweets are rare, even though we could find that the ideological extremity, electoral incentives, and legislative performance have influences on it, the influences would have limited contribution to the four sources' overall influences on expressed affective polarization. Therefore, the current study will report and discuss these unexpected findings in the result part but will not further test the related hypotheses.

Content Analysis of Tweets: Machine Learning Step

The current study applied the Twitter-RoBERTa model for supervised machine learning (Barbieri et al., 2020). The BERT model is a cutting-edge natural language processing model initially developed by Google, and the current study adopted a finetuned version from Cardiff University, the Twitter-RoBERTa, to deal with the Twitter political context. Researchers split the human coding dataset into 95% VS 5% as the training set and evaluation set, then used K-fold cross-validation (K=10) to evaluate the tweets-coding performance, and finally used two test datasets to examine the machine learning models' performance. The first test dataset contained 2000 tweets stratified and randomly sampled from all tweets[2]. The second test dataset contained 2000 tweets constructed by randomly selecting one tweet from one account each time. The current study reported the F1 score and accuracy scores for the evaluation dataset and two test datasets, with all F1 scores (label=1) above 0.7, weighted F1 scores above 0.90, and accuracy scores above 0.90. The results have shown that the machine-learning models have good performances.

See Table 1 for details.

The current study applied the PolarComm project (Zhang & Ma, 2023) in the third-person platform for tuning the Twitter-RoBERTa model (Barbieri et al., 2020).

Because the machine learning model performs well in coding tweets that contribute to expressed affective polarization, policy-embedded in-party liking, policy-embedded out-party disliking, and character-only in-party liking, the current study has applied the model to identify all politician tweets. After this step, the current study could get the data for dependent variables.

[2] Note that Two human coders examined intercoder reliability again in the second dataset, and we still have a good consistency. Cohen's Kappa of expressed affective polarization = 0.88, Cohen's Kappa of policy-embedded in-party liking = 0.73, Cohen's Kappa of policy-embedded out-party disliking = 0.92, Cohen's Kappa of character-only in-party liking = 0.79, and Cohen's Kappa of character-only out-party disliking = 1.00 (N=200).

Analytic Plans

The current study will construct two-way fixed-effect panel data models to analyze the effect of ideological extremity, electoral incentives, and legislation performance on the overall and three sub-dimensions of expressed affective polarization. The expressed affective polarization and the three sub-dimensions are calculated from each politician's number of related tweets in each term. All IVs and DVs are measured termly (every two years). IVs in one specific term would match the DVs in the same term. For example, ideological extremity and legislation performance in term 112 (2011-2012) and electoral performances for term 112 (2011-2012) would match the expressed affective polarization in term 112 (2011-2012). The congressional identity (senator =1 and house representative = 0), power of control (majority party = 1 and minority party = 0), seniority (re-elected = 1 and newcomer = 0), and the level of overall social media expression (number of total tweets by each person each term) variables are the control variables. The four control variables are represented by $\gamma_1 i$ to $\gamma_4 i$. The individual effect is represented by α_i, and the time effect is represented by λ_t. The current study defines ExpAP as short for expressed affective polarization, InPlyLik as short for policy-embedded in-party liking, OutPlyDislik as short for policy-embedded out-party disliking, and InCharLik as short for character-only in-party liking. IE is short for ideological extremity. LP is short for politicians' legislative performance. GE is short for general electoral incentive. PE is short for primary electoral incentive. For each individual *i* at year *t*, we have:

$$ExpAP_{it} = \alpha_i + \lambda_t + \beta_1 IE_{it} + \beta_2 LP_{it} + \beta_3 GE_{it} + \beta_4 PE_{it} + (\gamma_{1\,it} \ldots + \cdots \gamma_{4\,it}) + e_{it}$$
$$InPlyLik_{it} = \alpha_i + \lambda_t + \beta_1 IE_{it} + \beta_2 LP_{it} + \beta_3 GE_{it} + \beta_4 PE_{it} + (\gamma_{1\,it} \ldots + \cdots \gamma_{4\,it}) + e_{it}$$
$$OutPlyDislik_{it} = \alpha_i + \lambda_t + \beta_1 IE_{it} + \beta_2 LP_{it} + \beta_3 GE_{it} + \beta_4 PE_{it} + (\gamma_{1\,it} \ldots + \cdots \gamma_{4\,it}) + e_{it}$$
$$InCharLik_{it} = \alpha_i + \lambda_t + \beta_1 IE_{it} + \beta_2 LP_{it} + \beta_3 GE_{it} + \beta_4 PE_{it} + (\gamma_{1\,it} \ldots + \cdots \gamma_{4\,it}) + e_{it}$$

RESULTS

Descriptive Analyses

The current study has collected each politician in each term's level of ideological extremity (short as dw1-absolute), general election incentive (short as general election), primary election incentive (short as primary election), legislative performance (short as LES), expressed affective polarization (short as overall-polar), policy-embedded in-party liking (short as policy-liking), policy-embedded out-party disliking (short as policy-disliking), character-only in-party liking (short as character-liking), overall social media expression (short as total tweets), congressional identity, power of control, and seniority[3].

Among all data points (N=2865): 560 of them are Senators in their Terms, 2305 of them are Representatives in their Terms; 503 of them are First Incumbent legislators in their Terms, 2362 of them are Re-elected legislators in their terms; 1397 of them are Democrats in their Terms, 1468 of them are Republicans in their Terms. Some data points are missing because some politicians do not have public Twitter accounts, especially in early terms. As time goes on, people could observe that more politicians begin to have their accounts, and more politicians get re-elected. The observation fits election poll data because only a small percentage of seats could be replaced by new legislators, and over 80% (even over 90% in the Senate) of the seats remain in the same people until they retire, pass away, or move to another office. The numbers of Democrats and Republicans are close in Congress, so there is often a weak majority or minority in the past 12 years.

See Table 2 for details.

[3] The dissertation will use these shortened names in the following paragraphs, tables, and figures.

Among all data points, each politician per term has posted 1579 tweets, including 312 expressed affective polarization tweets, 98 policy-liking tweets, 209 policy-disliking tweets, and 40 character-liking tweets on average. A small percentage of politicians posted most of the tweets, so the distributions of tweets, overall expressed affective polarization tweets, policy-liking tweets, policy-disliking tweets, and character-liking tweets are negatively skewed.

The current study has also analyzed the percentage of polarized tweets. Among all data points, each politician per term has posted 17.5% expressed affective polarization tweets, 5.8% policy-liking tweets, 11.1% policy-disliking tweets, and 2.7% character-only tweets. A small percentage of politicians posted most of the tweets, so the distributions of tweets, overall expressed affective polarization tweets, policy-liking tweets, policy-disliking tweets, and character-liking tweets are also negatively skewed.

See Table 3 for details.

Each politician per term has an average score of 0.44 for ideology extremity, 0.99 for legislative performance, and on average, wins 64% vote shares in the general election and 79% in the primary election. As time goes on, politicians have greatly increased posting tweets, especially polarization-related tweets. However, the ideology extremity and legislation performance scores have slightly increased by 4.1% and 9.6%. In addition, the general election and primary election vote shares remain stable.

See Table 3 for details.

Research Questions Analyses

The current study focuses on how politicians' expressed affective polarization, and subdimensions of policy-liking, policy-disliking, and character-liking levels change over time.

According to the results, the current study has found that as time goes on, politicians have greatly improved posting tweets, especially affective-polarization-related tweets.

The current study has analyzed the raw number of tweets. From 2011 to 2022, the number of total posted tweets has increased by 140%, the number of expressed affective polarization tweets has increased by 414%, the number of policy-liking tweets has increased by 547%, the number of policy-disliking tweets has increased by 427%, and the number of character-liking tweets has increased by 124%.

Policy-disliking tweets are the major contributor to overall expressed affective polarization tweets, which account for 67.1% of all expressed affective polarization tweets. Policy-liking tweets account for 31.4%, and character-liking tweets only account for 12.9%. Because some tweets contain both liking and disliking attitudes, the sum of the three subdimensions exceeded 100%.

During the past 12 years / 6 terms, the number of total tweets increased steadily. However, the number of expressed affective polarization tweets did not increase steadily. Polarization tweets increased mildly from term 112 to term 114 (2011-2016), increased sharply in term 115 (2017-2018), and then increased relatively slowly from term 116 to term 117 (2019-2022). Among the three subdimensions, policy-liking and policy-disliking tweets shared a similar trend with the overall expressed affective polarization tweets. However, the character-liking tweets only contained a small percentage of the overall expressed affective polarization tweets and increased slower than the other sub-dimensions.

The current study has also analyzed the percentage of polarized tweets. From 2011 to 2022, the percentage of expressed affective polarization tweets has increased by 123%, the percentage of policy-liking tweets has increased by 193%, the percentage of policy-disliking tweets has increased by 139%, and only the percentage of character-liking tweets has decreased 12.4%.

During the past 12 years / 6 terms, the percentage of polarized tweets has shared a similar trend with the raw number of polarized tweets. Polarization tweets percentage increased mildly from term 112 to term 114 (2011-2016), increased sharply in term 115 (2017-2018), and then increased relatively slowly from term 116 to term 117 (2019-2022). In term 112 (2011-2012), the number of expressed affective polarization tweets accounted for 12.0% of the total tweets, while in term 117 (2019-2020), the number of expressed affective polarization tweets accounted for 25.5% of the total tweets. Among the three subdimensions, policy-liking and policy-disliking tweets percentages shared a similar trend with the overall expressed affective polarization tweets percentage. However, the character-liking tweets percentage has remained low percentage and even slightly decreased in term 117 (2021-2022).

See Figure 3 for details.

The current study has split the tweets' raw number data into Democratic and Republican subsets and found that Democrats and Republicans have different patterns of affective polarization expression. Democrats expressed affective polarization tweets increased sharply in term 115 (2017-2018), maintained a high level in term 116 (2019-2020), and slightly dropped in term 117 (2021-2022). Republicans expressed affective polarization tweets did not increase sharply until term 117 (2021-2022). Democrats are more likely to express policy-liking than Republicans, while both party members like to express policy-disliking and do not often express character-liking.

The current study has also split the tweets percentage data into Democratic and Republican subsets and found that Democrats and Republicans have different patterns of affective polarization expression. The tweets percentage and tweets raw number shared a similar trend. Democrats expressed affective polarization tweets percentage increased sharply in term 115 (2017-2018, and slightly dropped in terms 116 and 117 (2019-2022). Republicans expressed affective polarization tweets percentage did not increase sharply until term 117 (2021-2022). Democrats are slightly more likely to express policy-liking than Republicans, while both party members like to express policy-disliking and do not often express character-liking.

See Figure 4 for details.

The current study needs to specifically report the finding of character-only out-party disliking tweets. According to the human coding results, US politicians seldom tweet solely attacking tweets without mentioning any policies. The finding is consistent with previous studies in the traditional media age and early social media age. According to previous studies on political expressions, researchers have found politicians mostly post policy-embedded out-party disliking information but rarely express character-only out-party disliking to the public (e.g., Geer, 2006; Auter & Fine, 2016). This phenomenon exists not only in the traditional media age (Geer, 2006) but also in the social media age (Auter & Fine, 2016). For example, one study has found that only 4.5% of Facebook posts were non-policy-related attacks during the 2010 election (Auter & Fine, 2016). Auter & Fine's (2016) investigation was conducted during the campaign, while the current study focuses on the whole in-office time. Considering politicians are more active and aggressive during the campaign than the routine time (Schmuck & Hameleers, 2020), the percentage of character-only out-party disliking expressions in the current study may be even lower than the reported 4.5%.

In addition, previous literature has also found that challengers or people lacking supporters or resources were more likely to express character-only out-party disliking a “desperation strategy” (Lau & Pomper, 2001; Auter & Fine, 2016), but those candidates are unlikely to win the election. The current study focuses on in-office politicians (federal Congress legislators), so they are relatively unlikely to express character-only out-party disliking.

To sum up, the finding that current US politicians seldom post character-only out-party disliking tweets may be different from the general public’s current perception of politicians. However, the investigation across all US federal legislators has depicted that most of the in-office politicians in most of the time are still policy-focused. It is a positive finding for both US politicians and the public.

Hypotheses Testing

The current study tested the influence of ideological extremity, general election incentive, primary election incentive, and legislative performance on expressed affective polarization, as well as three subdimensions (policy-embedded in-party liking, policy-embedded out-party disliking, and character-only in-party liking). Because the expressed affective polarization and its three subdimensions were negatively skewed, they were transformed by taking the $\log_{10}$ of (1+x). In addition, the election incentives were reverse-coded by the vote shares that politicians received. The higher vote shares they got, the lower they had electoral incentives.

The dataset has 818 politicians with 6 terms. We could not expect that some unobservable characteristics of politicians have no covariance with the other explanatory variables, such as their ideology or political status. And we also could not expect that some unobservable characteristics of time have no covariance with the other explanatory variables, such as the vote share or Congress dominance. Thus, the current study has proposed two-way fixed-effects (TWTE) panel data models to test the hypotheses. The current study has four dependent variables, which indicates four models. The current study examined the four models’ heteroscedasticity, first-order auto-correlation, and cross-sectional dependence. All four models have heteroscedasticity, first-order auto-correlation, and cross-sectional dependence.

See Table 4 for details.

The current study used the modified Hausman test to examine the model choices. All four models showed significant differences between the two-way fixed and random-effect models, indicating that the current study's choice of two-way fixed-effect models was correct. Meanwhile, because the data has heteroscedasticity, first-order auto-correlation, and cross-sectional dependence, the current study should use the Driscoll-Kraay modified regression[4] to test the models (Hoechle, 2007).

See Table 5 for details.

According to the results, the current study has found that ideology extremity positively influences expressed affective polarization (B=0.31, β=0.08, $^{*}p$<0.05). In subdimensions, ideology extremity positively influences policy-liking (B=0.17, β=0.05, $^{*}p$<0.05), policy-disliking (B=0.47, β=0.10, $^{**}p$<0.01), and character-liking (B=0.29, β=0.09, $^{*}p$<0.05). Therefore, H1, H1a, H1b, and H1c are supported. But to analyze from the standardized coefficient β, the effect sizes are small. The findings indicate that politicians' ideology and expression are correlated but not as strongly as expected.

The current study has found that the general electoral incentive has a very small and negative influence on expressed affective polarization (B=0.00, β=-0.05). In subdimensions, the general electoral incentive negatively influences policy-disliking (B=0.01, β=-0.11, ^{+}p<0.1) but positively influences policy-liking (B=0.00, β=0.01) and character-liking (B=0.00, β=0.03).

[4] R code package command: xtscc

Therefore, H2a and H2c are supported, while H2 and H2b are not supported. The findings indicate that politicians with low vote shares are more likely to be cautious about expressing polarization, especially on criticisms such as policy-embedded out-party disliking. However, politicians with low vote shares are more likely to call for in-party support, such as increasing expressions of in-party liking. However, the effect sizes of general electoral incentives on expressed affective polarization and its subdimensions are small. It indicates that people should not over-explain the general electoral incentive's influence on expressed affective polarization.

The current study has found similarly that the primary electoral incentive nearly has no influence on expressed affective polarization (B=0.00, β=0.00). In subdimensions, the primary electoral incentive negatively influences policy-disliking (B=0.00, β=-0.01) but positively influences policy-liking (B=0.001, β=0.03, $^{**}p$<0.01) and character-liking (B=0.001, β=0.02, $^{**}p$<0.01). Therefore, H3a and H3c are supported, while H3 and H3b are not supported. The findings demonstrate that primary and general electoral incentives share similar patterns. The low vote share indicates less expressed affective polarization, especially out-party disliking expressions, but more in-party liking expressions. However, the effect sizes of primary electoral incentives on expressed affective polarization and its subdimensions are even smaller than the general election. It indicates that the primary electoral incentive has a very limited influence on expressed affective polarization.

The current study has found that legislative performance has nearly no influence on expressed affective polarization (B=0.00, β=0.00). In subdimensions, the legislative performance positively influences policy-liking (B=0.01, β=0.02) but negatively influences character-liking (B=0.00, β=-0.01) and nearly has no influence on policy-disliking (B=0.00, β=0.00). Therefore, H4c is all supported, but H4, H4a, and H4b are not supported. The effect sizes of legislative performance on expressed affective polarization and its subdimensions are also small.

Some control variables have an important influence on expressed affective polarization. Congressional identity (B=-0.15, β=-0.10, $^{**}p<0.01$), Power of control (B=-0.09, β=-0.07, $^{*}p<0.05$), and Seniority (B=-0.01, β=-0.01) share a similar pattern: the higher a person's political status, the less they would express affective polarization. Senators, majority party members, and re-elected politicians are less likely to express affective polarization than House representatives, minority party members, and first incumbent politicians. In the subdimensions, Senators, majority party members, and re-elected politicians are also more "calm," which means they slightly increase policy-liking expressions but try to avoid attacking others (e.g., policy-disliking) or providing unreasonable flatteries (e.g., character-liking expressions).

The overall social media expression greatly influences expressed affective polarization and its three subdimensions. The findings indicate that the amount of polarized tweets increases when the total tweet expressions increase. Time also has an important contribution to expressed affective polarization. The model set term 112 (2011-2012) as the baseline and found that terms 113 (2013-2014) and 114 (2015-2016) even had less affective polarization expression. However, starting from term 115 (2017-2018), expressed affective polarization increased rapidly. The results indicate that the presidential campaign during 2015 and 2016 was indeed a watershed moment in American politics.

See Table 6 for details.

The current study also tested the influence of ideological extremity, general election incentive, primary election incentive, and legislative performance on the percentage of expressed affective polarization, as well as three subdimensions (policy-embedded in-party liking, policy-embedded out-party disliking, and character-only in-party liking). Because the percentage expressed affective polarization and its three subdimensions were negatively skewed, they were transformed by taking the $\log_{10}$ of (1+x). In addition, the election incentives were reverse-coded by the vote shares that politicians received. The higher vote shares they got, the lower they had electoral incentives.

Similarly, the current study applied two-way fixed-effects (TWTE) panel data models to test the hypotheses. Because the data has heteroscedasticity, first-order auto-correlation, and cross-sectional dependence, the current study should use the Driscoll-Kraay modified regression[5] to test the models (Hoechle, 2007).

See Tables 7 and 8 for details.

According to the results, the current study has found that ideology extremity positively influences the percentage of expressed affective polarization tweets (B=0.31, β=0.15, $^{*}p<0.05$). The ideology extremity also positively influences the percentage of policy-liking (B=0.17, β=0.07, $^{*}p<0.05$), policy-disliking (B=0.47, β=0.14, $^{**}p<0.01$), and character-liking tweets (B=0.28, β=0.12, $^{*}p<0.05$).

The general electoral incentive has small and negative influences on the percentage of expressed affective polarization tweets (B=0.00, β=-0.09) and the percentage of policy-disliking tweets (B=-0.01, β=-0.16, $^{+}p<0.1$). And the general electoral incentive has small and positive influences on the percentage of policy-liking tweets (B=0.00, β=0.02) and the percentage of character-liking tweets (B=0.00, β=0.03).

[5] R code package command: xtscc

The primary electoral incentive also has small and negative influences on the percentage of expressed affective polarization tweets (*B*=0.00, β=-0.01) and the percentage of policy-disliking tweets (*B*=0.00, β=-0.02). And the primary electoral incentive has small and positive influences on the percentage of policy-liking tweets (*B*=0.001, β=0.05, $^{**}p<0.01$) and the percentage of character-liking tweets (*B*=0.001, β=0.04, $^{**}p<0.01$).

The legislative performance has nearly no influence on the percentage of expressed affective polarization tweets (*B*=0.00, β=0.00). And the legislative performance has positive influences on the percentage of policy-liking (*B*=0.01, β=0.02) and policy-disliking tweets(*B*=0.00, β=-0.01) but has a negative influence on the percentage of character-liking tweets (*B*=-0.01, β=-0.02).

Some control variables have an important influence on expressed affective polarization. Congressional identity (*B*=-0.15, β=-0.19, $^{**}p<0.01$), Power of control (*B*=-0.09, β=-0.14, $^{*}p<0.05$), and Seniority (*B*=-0.01, β=-0.01) share a similar pattern: the higher a person's political status, the less they would express affective polarization. Senators, majority party members, and re-elected politicians are less likely to express affective polarization than House representatives, minority party members, and first incumbent politicians. Time also has an important contribution to expressed affective polarization. The model set term 112 (2011-2012) as the baseline and found that terms 113 (2013-2014) and 114 (2015-2016) even had less affective polarization expression. However, starting from term 115 (2017-2018), expressed affective polarization increased rapidly.

See Table 9 for details.

The current study has also tested the models with time lags. In the time-lagged models, politicians' ideological extremity and legislative performance in time 1 would be used to predict expressed affective polarization in time 2. The expressed affective polarization is measured by either $\log_{10}$ (number of raw tweets) or $\log_{10}$ (percentage of tweets). The time-lagged models also show similar findings that ideological extremity has a small positive effect on expressed affective polarization, while legislative performance has a small negative effect on expressed affective polarization. Electoral incentives have small negative effects on the overall expressed affective polarization but have small positive effects on the in-party liking expressions. Meanwhile, the overall social media expression, the 2016 presidential election, politicians' identity in Congress, and the power of control in Congress have important influences on politicians' expressed affective polarization.

In addition, the time-lagged models (with the percentage of tweets as the DV) have shown that when people post more tweets, they would also post a higher percentage of policy-embedded out-party disliking tweets but post a lower percentage of in-party liking tweets. Overall speaking, when people post more tweets, they would slightly post a higher percentage of overall expressed affective polarization tweets.

See Table 10 and Table 11 for details.

The current study treated expressed affective polarization and its subdimensions as raw numbers and percentages. The findings between the two ways of treatment are similar. Ideological extremity has a positive influence on expressed affective polarization, but the effect size is small. While electoral incentives and legislative performance even have more limited influence on expressed affective polarization.

DISCUSSION

The current study explores US politicians' dynamic of affective polarization expression on social media from 2011 to 2022 by terms. This study also examines a theoretical assumption that politicians' ideology and political pursuits could influence their online expression to the general public (e.g., Fenno, 2003; Mayhew, 2004; Kinder & Kalmoe, 2017; Jacobson & Carson, 2019; Amadae, 2021). Specifically, the current study examines a model of how politicians' ideology extremity, legislative performance, and electoral incentives could influence their expressed affective polarization on social media.

Study Findings and Theoretical Implications

Sources: Ideology Extremity, Electoral Incentives, and Legislative Performance

The findings demonstrate that US politicians have greatly increased the total expression and the affective polarization expression on social media in the past 12 years, especially the policy-embedded in-party liking and policy-embedded out-party disliking expressions. This study also finds that politicians' ideology extremity positively influences the expressed affective polarization, while legislative performance nearly has no influence on the expressed affective polarization. The general and primary electoral incentives have mixed influences on the expressed affective polarization, which negatively influence the expression of overall affective polarization, especially the policy-embedded out-party disliking, but positively influence the expression of policy-embedded and character-only in-party likings. Most of the hypotheses are supported, however, with limited effect sizes. In addition, politicians' expressed affective polarization has other influential sources, such as the overall social media expression, the 2016 election, Senators in Congress, and the power of control in Congress.

In sum, politicians' expressed affective polarization level has grown rapidly in the past 12 years, but its sources are complex. Politicians' ideology and political pursuits have contributed to their expressions but do not contribute as effectively as people have expected. Meanwhile, the media environment and rule-breaking political events could be more influential than expected.

Affection More Than Ideology: Politicians Share Same Pattern with the General Public

Previous studies have shown that the general public has a visible increase in affective polarization but does not consistently increase in ideological polarization (e.g., Iyengar et al., 2012; Fiorina et al., 2005; Abramowitz, 2010). Meanwhile, politicians have significantly increased ideological polarization in the past four decades (e.g., McCarty, 2019; Neal, 2020). The current study has contributed to the last piece of the puzzle and found that politicians' expressed affective polarization has increased much more rapidly than their ideological polarization in the past 12 years. From congressional terms 112 to 117 (from 2011 to 2022), US legislators' average absolute DW-nominate score has increased by 4.1%, with Democrats increasing by 0.1% and Republicans increasing by 8.5%. Compared to studies showing four decades' difference, the 12-year growth rate is not astonishing. However, US legislators' expressed affective polarization on social media has increased by 414%, with Democrats increasing 654% and Republicans increasing 448%, far beyond their ideological polarization's growth rates. In addition, unlike Democrats have remained stable in ideology, both Democrats and Republicans have increased their level of expressed affective polarization. The increasing trend of Democrats even goes earlier and faster than Republicans. Therefore, the current study proposes that politicians share the same pattern with the general public: their affective polarization levels increase much faster than ideological polarization. Just like a citizen who can not tolerate opposite-party friends may not be familiar with in-party policies, politicians actively criticizing others may not always support extreme bills. What people have expressed online does not always fit what they do in reality.

Growing Polarization: More with the Flow Rather Than Deliberation

Previous studies have a theoretical assumption that politicians are more rational than the general public and propose that their political behaviors are likely to be guided by rationality (e.g., Fenno, 2003; Mayhew, 2004; Kinder & Kalmoe, 2017; Jacobson & Carson, 2019; Amadae, 2021). However, the current study has found that politicians' ideology, legislative performance, and electoral incentives do have influences with expected directions on expressed affective polarization but with limited effect sizes (all standardized coefficients $\beta < 0.15$).

Meanwhile, the level of overall social media expression has an important influence on expressed affective polarization (standardized coefficient $\beta > 0.5$). The findings indicate that politicians express more affective polarization partly because they use social media more actively and emerge in an environment of unlimited expression than before. Politicians could - relatively - post as many as they wanted and receive few restrictions compared with television, newspaper, or magazine reports (e.g., Gainous & Wagner, 2014). Thus, the more politicians have expressed, the more disclosure they may have on some specific dimensions, such as affective polarization.

It is not surprising that the overall media use could influence one specific direction of expression. However, it is interesting to find that the influence of media use and the influence of ideology and political pursuits are out of proportion. The findings indicate that politicians express themselves by following the flow of social media use rather than being influenced by some designed political goals. In other words, ideology extremity, legislative performance, and electoral incentives could change politicians' affective polarization expression. However, the expressed affective polarization may not be a deliberate response to these concerns.

2016 Presidential Election: A Watershed for Both Parties

The current study has also found that the 2016 presidential election has a crucial influence on politicians' expressed affective polarization. All four models, including the overall expressed affective polarization and three subdimension models, have shown that the politicians in terms 113 and 114 (2013-2016) expressed even less affective polarization than people in term 112 (2011-2012). However, in terms 115, 116, and 117 (2017-2022), politicians expressed much more affective polarization than people in term 112 (2011-2012). The findings indicate that some events between terms 114 and 115 (2015-2018) have greatly changed the political environment. The most possible answer could be the 2016 presidential election. During this election, Donald Trump used Twitter as the base of his campaign and broke the rule of political expression. Many of his expressions could be defined as affective polarization, such as using unsubstantiated information to criticize his Democratic opponents or using exaggerated words and tones to praise his Republican friends (e.g., Bacon, 2016; Brenner, 2021; Coll, 2017).

The 2016 Trump election may have a different influence on the Democrats and Republicans. Typically, people would expect the minority party would be more likely to post negative tweets (Lee, 2016). However, it is the Democrat Party that first got "stimulated" by Trump and increased out-party disliking and the overall affective polarization expressions. According to the research question results, Democratic legislators increased policy-disliking expressions right after the 2016 election, and this high-level policy-disliking expression continued until Donald Trump was defeated in late 2020. Compared with Democrats, Republicans between 2016 and 2020 did not increase out-party disliking and overall polarization expressions that much. After Joe Biden took office in 2021, Republicans may have learned from Trump's example and increased the policy-disliking expressions sharply. The trend lasted until the end of 2022 and might be continued. The results indicate that the 2016 election has largely influenced both parties. If we try to depict a potential picture - Democrats were irritated in 2017, while Republicans were enlightened in 2021. The 2016 Presidential Election was a watershed for American politics for both parties, and largely influenced politicians' expressed affective polarization in the following several terms.

More Expressed Affective Polarization from the Minority Party

The hypotheses models have shown that the Minority party members in Congress (whatever House or Senate) would be more likely to express affective polarization. According to models in Table 6, the Majority party members would significantly reduce the expressed affective polarization (B=-0.09, β=-0.07, p<0.05) and especially reduce the policy-embedded out-party disliking expression (B=-0.23, β=-0.16, p<0.10). Meanwhile, the results are similar to those additional models in Tables 9, 10, and 11[6].

The findings that Minority party members are more likely to express affective polarization and policy-embedded out-party disliking have reinforced previous findings that the Minority party is more likely to make confrontations in legislation and campaigns (Lee, 2016). Starting from the late 1980s, the US federal level Congress fell into an "insecure majority" situation (Lee, 2016). It means no party could maintain a stable Majority party status, so both Democratic and Republican Parties feel insecurity and have to join an endless competition that

[6] The additional models include : 1) treating expressed affective polarization (DV) as a percentage; 2) a time-lagged model that calculating ideological extremity and legislation performance at $time_{(t-1)}$; 3) a time-lagged model that calculating ideological extremity and legislation performance at $time_{(t-1)}$ and treating expressed affective polarization (DV) as a percentage.

re-occurs every two years. When the control of Congress hangs in the balance, party members will prioritize how they might affect their party's common fate, and invest more heavily in party organization and partisan collective action (Arbour, 2014; Lee, 2016). Moreover, because the lack of power is a freedom from the sense of governing responsibility, the minority party members would be more motivated to criticize the Majority party and highlight the party difference (Mann & Ornstein, 2016; Lee, 2016). This phenomenon existed in the traditional media age in which television campaigns dominated elections, while the current findings indicate that it still remains in the social media age.

Practical Implications

The current findings could have several practical implications. First, the findings indicate that appealing to ideology, electoral incentives, or legislative performance does not effectively change politicians' expressed affective polarization. In other words, people's criticisms of politicians' working style, or even votes, may not be enough to change their expression on social media. The crucial points are social media use and the political environment. If people or the platform could reward politicians who are less polarized, they may reduce the expression of affective polarization. In addition, the whole society should continue to be aware of the negative impact of "Trumpian" campaigns. Otherwise, the expressed affective polarization may spiral upward along with the rotation of ruling parties (Bennett & Livingston, 2018; Bennett & Berenson, 2020).

However, the current study should also acknowledge that changing social media use or the political environment is long-term and hard work. For instance, rewarding some specific kinds of politicians on social media may challenge fundamental values such as freedom of speech (e.g., Macedo, 2022). Meanwhile, politicians who benefit from Trumpian campaigns may not have the motivation to change their expressions. Considering the large percentage of Trump supporters in society, it is also unlikely to reach a consensus in the general public that Trumpian campaigns are strategic burdens of American politics. One study cannot answer all the questions; thus, researchers must conduct more studies to provide more practical tools for individuals.

Limitations and Future Study Directions

Here are some limitations of the study. First, the current study's machine learning model performances on predicting policy-embedded in-party liking and character-only in-party liking are just marginally acceptable (with F1=0.7). Researchers could do more work on improving the performance of machine learning models. Secondly, the two-way fixed-effect model has received some recent critiques. Scholars have concerns about factors that could have affected mean levels of the outcome that are not captured in linear trends, heterogeneity of treatment effects, and the choice of the functional form of the trend when multiple time points are observed prior to the introduction of treatment. These concerns can affect estimates, standard errors, and inferences (Callaway & Sant'Anna, 2021; Kahn-Lang & Lang, 2020). Thirdly, the current study only focused on the Twitter platform. However, politicians also use other social media platforms like Facebook, Instagram, or Youtube. The level of expressed affective polarization might be different among different platforms. In addition, the current study mainly focuses on the sources of politicians' expressed affective polarization but does not have space to test the expressed affective polarization's effect on the general public. The study's contribution could be more considerable if it constructed a whole cause-and-effect mechanism of expressed affective polarization.

The current study also has great potential for future studies. First, future studies could work on the methodology part of improving the machine learning model performance. One of the crucial restrictions of the Twitter-RoBERTa model is that it does not have much pre-trained information on policy names, politicians, and their party affiliations. Modifying the machine learning model with discipline-specific information requires detailed and tedious work but could substantially improve the performance in one research discipline, such as political communication.

Second, the current study also provides an opportunity to test different operationalizations of polarization. Previous polarization studies have proposed similar but different ideological and affective polarization operationalizations, such as the thermometer measurement (e.g., Iyengar et al., 2012; Lelkes & Westwood, 2017), the in-direct "comfortable neighbor, friend, and marriage" measurement (Iyengar et al., 2012), the social media posts number-counting measurement (Yu et al., 2022), and the ambivalence measurement (Basinger & Lavine, 2005). With the complete political social media dataset, future studies could cross-validate those different operationalizations and create systematic reviews of polarization studies.

Thirdly, as the limitation also mentioned, future studies should focus on constructing the cause-and-effect mechanism of expressed affective polarization and explain how politicians' thoughts and behaviors finally impact the general public. In addition, future studies could also test the influence of politicians' social networks on expressed affective polarization. These studies could be regarded as detailing the influence of social media use. Since politicians' ideology and political pursuits may not contribute most to their expressions, peer pressure and social network could be more influential than expected.

CONCLUSION

The current study has explored the dynamic of US politicians' level of expressed affective polarization from 2011 to 2022. By collecting Twitter posts from US federal-level legislator accounts and applying supervised machine learning models to identify expressed affective polarization tweets, the current study has found that US politicians' expressed affective polarization level has increased rapidly in the past 12 years, so do its two subdimensions including policy-embedded in-party liking expressions and policy-embedded out-party disliking expressions. The increasing trend is much faster than the increasing trend of ideological polarization.

The current study has also focused on the same group of politicians and used panel data models to examine the influences of ideology extremity, electoral incentives, and legislative performance on expressed affective polarization. The findings demonstrate that ideology extremity positively influences expressed affective polarization and legislative performance nearly has no influence on expressed affective polarization. The electoral incentives negatively influence the overall expressed affective polarization but positively influence in-party liking expressions. However, the effect sizes of these sources are small, indicating that politicians' ideology, electoral incentives, and legislative performance have contributed to their expressions but do not contribute as effectively as people expected.

Meanwhile, the overall social media expression, the 2016 presidential election, and the power of control in Congress are more influential than expected. The majority party members are less likely to express affective polarization. However, when the overall use of social media increases or rule-breaking political events have gotten political rewards, politicians are more likely to express affective polarization later on. The findings indicate that politicians' expressed affective polarization is also largely influenced by the media environment, rule-breaking political events, and the elites themselves' political status. Further studies on the media environment, social networks, peer pressures, and modifying machine learning performances are also worth trying in the future.

REFERENCES

Abramowitz, A. I., & Saunders, K. L. (2008). Is Polarization a Myth? *The Journal of Politics*, *70*(2), 542–555. https://doi.org/10.1017/s0022381608080493

Abramowitz, A. I. (2010). Transformation and polarization: The 2008 presidential election and the new American electorate. *Electoral Studies*, *29*(4), 594–603. https://doi.org/10.1016/j.electstud.2010.04.006

Abramowitz, A. I. (2021). It's Only You and Me and We Just Disagree: The Ideological Foundations of Affective Polarization. *The Forum*, *19*(3), 349–364. https://doi.org/10.1515/for-2021-0019

Alabama Secretary of State | 2010 Election Information. (2010). https://www.sos.alabama.gov/alabama-votes/voter/election-information/2010/

Amadae, S. M. (2021, March 11). *rational choice theory | Definition, Examples, & Facts*. Encyclopedia Britannica. https://www.britannica.com/topic/rational-choice-theory

Amira, K., Wright, J. C., & Goya-Tocchetto, D. (2019). In-Group Love Versus Out-Group Hate: Which Is More Important to Partisans and When? *Political Behavior*, *43*(2), 473–494. https://doi.org/10.1007/s11109-019-09557-6

Arbour, B. (2014, December 9). All Politics is Local? Not anymore. *The Washington Post*. https://www.washingtonpost.com/news/monkey-cage/wp/2014/12/09/all-politics-is-local-not-anymore/

Auter, Z. J., & Fine, J. A. (2016). Negative Campaigning in the Social Media Age: Attack Advertising on Facebook. *Political Behavior, 38*(4), 999–1020. https://doi.org/10.1007/s11109-016-9346-8

Bacon J. (2016 October 3). Donald Trump's election is a victory for emotion-led marketing. *MarketingWeek*. Retrieved from https://www.marketingweek.com/donald-trumps-rise-is-a-victory-for-emotion-led-marketing/

Ballard, A. O., DeTamble, R., Dorsey, S., Heseltine, M., & Johnson, M. (2022). Dynamics of Polarizing Rhetoric in Congressional Tweets. *Legislative Studies Quarterly*. https://doi.org/10.1111/lsq.12374

Barbieri, F., Camacho-Collados, J., Neves, L., & Espinosa-Anke, L. (2020). Tweeteval: Unified benchmark and comparative evaluation for tweet classification. *arXiv preprint arXiv:2010.12421*.

Bartels, L. M. (2000). Partisanship and voting behavior, 1952-1996. *American Journal of Political Science*, 35-50. https://doi.org/10.2307/2669291

Basinger, S. J., & Lavine, H. (2005). Ambivalence, information, and electoral choice. *American Political Science Review*, *99*(2), 169-184.

Bennett, B., & Berenson, T. (2020 November 7). How Donald Trump lost the election. *Time*. Retrieved from https://time.com/5907973/donald-trump-loses-2020-election/

Bennett, W. L., & Livingston, S. (2018). The disinformation order: Disruptive communication and the decline of democratic institutions. *European Journal of Communication*, *33*(2), 122-139. https://doi.org/10.1177/0267323118760317

Blum, R., Cormack, L., & Shoub, K. (2022). Conditional Congressional communication: how elite speech varies across medium. *Political Science Research and Methods*, 1-8. https://doi.org/10.1017/psrm.2022.28

Boese, V. A., Lindberg, S. I., & Lührmann, A. (2021). Waves of autocratization and democratization: a rejoinder. *Democratization*, *28*(6), 1202–1210. https://doi.org/10.1080/13510347.2021.1923006

Borrelli, D., Iandoli, L., Ramirez-Marquez, J. E., & Lipizzi, C. (2021). A Quantitative and Content-Based Approach for Evaluating the Impact of Counter Narratives on Affective Polarization in Online Discussions. *IEEE Transactions on Computational Social Systems*, 1–12. https://doi.org/10.1109/tcss.2021.3094775

Brenner, I. (2021). Disinformation, disease, and Donald Trump. *International Journal of Applied Psychoanalytic Studies*. https://doi.org/10.1002/aps.1694

Burns, A., & Martin, J. (2022, July 12). *McCarthy Says He Would Urge for Trump Resignation After Jan 6 in Recording*. The New York Times. https://www.nytimes.com/2022/04/21/us/politics/trump-mitch-mcconnell-kevin-mccarthy.html

Callaway, B., & Sant'Anna, P. H. (2021). Difference-in-differences with multiple time periods. *Journal of econometrics*, *225*(2), 200-230. https://doi.org/10.1016/j.jeconom.2020.12.001

Clark, J. H., & Evans, H. K. (2020). Let's talk about sex: Examining the factors influencing congressional response to# MeToo on Twitter. *PS: Political Science & Politics*, *53*(1), 51-56. https://doi.org/10.1017/S1049096519001124

Coll, S. (2017, December 11). Donald Trump's "fake news" tactics. *The New Yorker*. Retrieved from https://www.newyorker.com/magazine/2017/12/11/donald-trumps-fake-news-tactics

Curry, J. M., & Lee, F. E. (2019). Non-party government: Bipartisan lawmaking and party power in congress. *Perspectives on Politics*, *17*(1), 47-65. https://doi.org/10.1017/S1537592718002128

Druckman, J., & Levy, J. (2021). Affective Polarization in the American Public. *Northwestern Institute for Policy Research*.

Druckman, J. N., Peterson, E., & Slothuus, R. (2013). How Elite Partisan Polarization Affects Public Opinion Formation. *American Political Science Review*, *107*(1), *57–79*. https://doi.org/10.1017/s0003055412000500

Evans, H. K., Cordova, V., & Sipole, S. (2014). Twitter style: An analysis of how house candidates used Twitter in their 2012 campaigns. *PS: Political Science & Politics*, *47*(2), 454-462. https://doi.org/10.1017/S1049096514000389

Fenno, R. (2003). *Home style: House members in their districts*. Pearson College Division.

Fiorina, M. P., Abrams, S. J., & Pope, J. C. (2005). Culture war. *The myth of a polarized America*, *3*.

Gainous, J., & Wagner, K. M. (2014). *Tweeting to power: The social media revolution in American politics*. Oxford University Press.

Geer, J. G. (2006). *In defense of negativity: Attack ads in presidential campaigns*. Chicago: University of Chicago Press.

Haber, S. (2011). The 2010 US Senate elections in 140 characters or less: an analysis of how candidates use Twitter as a campaign tool. *WRLC Digital Repository*.

Herbst, S. (2010). *Rude democracy: Civility and incivility in American politics*. Temple University Press.

Hetherington, M. J. (2008). Turned off or turned on? How polarization affects political engagement. *Red and blue nation? Consequences and correction of America's polarized politics*, *2*, 1-33.

Hobolt, S. B., Leeper, T. J., & Tilley, J. (2020). Divided by the Vote: Affective Polarization in the Wake of the Brexit Referendum. *British Journal of Political Science*, *51*(4), 1476–1493. https://doi.org/10.1017/s0007123420000125

Hoechle, D. (2007). Robust standard errors for panel regressions with cross-sectional dependence. *The stata journal*, *7*(3), 281-312. https://doi.org/10.1177/1536867X0700700301

Hong, S., & Kim, S. H. (2016). Political polarization on twitter: Implications for the use of social media in digital governments. *Government Information Quarterly*, *33*(4), 777-782. https://doi.org/10.1016/j.giq.2016.04.007

Iyengar, S., Sood, G., & Lelkes, Y. (2012). Affect, Not Ideology: A social identity perspective on polarization. *Public Opinion Quarterly*, *76*(3), 405–431. https://doi.org/10.1093/poq/nfs038

Jacobson, G., & Carson, J. (2019). *The Politics of Congressional Elections* (Tenth). Rowman & Littlefield Publishers.

Jungherr, A. (2016). Twitter use in election campaigns: A systematic literature review. *Journal of information technology & politics*, *13*(1), 72-91. https://doi.org/10.1080/19331681.2015.1132401

Kahn-Lang, A., & Lang, K. (2020). The promise and pitfalls of differences-in-differences: Reflections on 16 and pregnant and other applications. *Journal of Business & Economic Statistics*, *38*(3), 613-620. https://doi.org/10.1080/07350015.2018.1546591

Kalla, J. L., & Porter, E. (2020). Correcting Bias in Perceptions of Public Opinion Among American Elected Officials: Results from Two Field Experiments. *British Journal of Political Science*, *51*(4), 1792–1800. https://doi.org/10.1017/s0007123419000711

Kim, J. W., Guess, A., Nyhan, B., & Reifler, J. (2021). The Distorting Prism of Social Media: How Self-Selection and Exposure to Incivility Fuel Online Comment Toxicity. *Journal of Communication*, *71*(6), 922–946. https://doi.org/10.1093/joc/jqab034

Kinder, D. R., & Kalmoe, N. P. (2017). Neither liberal nor conservative. In *Neither Liberal nor Conservative*. University of Chicago Press.

Kujala, J. (2020). Donors, primary elections, and polarization in the United States. *American Journal of Political Science*, *64*(3), 587-602. https://doi.org/10.1111/ajps.12477

Kyle, J. & Meyer, B. (2020, March 10). *High Tide? Populism in Power, 1990–2020. Institute for Global Change*. https://institute.global/policy/high-tide-populism-power-1990-2020

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 159-174.

Lassen, D. S., & Brown, A. R. (2011). Twitter: The electoral connection?. *Social Science Computer Review*, *29*(4), 419-436. https://doi.org/10.1177/0894439310382749

Lau, R. R., & Pomper, G. M. (2001). Negative campaigning by US Senate candidates. *Party Politics*, *7*(1), 69-87.

Lawrence, E. D., Maltzman, F., & Smith, S. S. (2006). Who wins? Party effects in legislative voting. *Legislative Studies Quarterly*, *31*(1), 33-69. https://doi.org/10.3162/036298006X201724

Layman, G. C., Carsey, T. M., & Horowitz, J. M. (2006). PARTY POLARIZATION IN AMERICAN POLITICS: Characteristics, Causes, and Consequences. *Annual Review of Political Science*, *9*(1), 83–110. https://doi.org/10.1146/annurev.polisci.9.070204.105138

Lee, F. E. (2016). *Insecure majorities: Congress and the perpetual campaign*. University of Chicago Press.

Lelkes, Y. (2016). Mass polarization: Manifestations and measurements. *Public Opinion Quarterly*, *80*(S1), 392-410.

Lelkes, Y. (2019). Policy over party: comparing the effects of candidate ideology and party on affective polarization. *Political Science Research and Methods*, *9*(1), 189–196. https://doi.org/10.1017/psrm.2019.18

Lelkes, Y., & Westwood, S. J. (2017). The limits of partisan prejudice. *The Journal of Politics*, *79*(2), 485-501. https://www.jstor.org/stable/26551007

Lewis, J. B., Poole, K., Rosenthal, H., Boche, A., Rudkin, A., & Sonnet, L. (2023). Voteview: Congressional roll-call votes database. Retrieved from https://voteview.com

Linde, J., & Dahlberg, S. (2021). Quality of Government and Political Support. *The Oxford Handbook of the Quality of Government*, 228.

Macedo, S. (2022). Lost in the marketplace of ideas: Towards a new constitution for free speech after Trump and Twitter?. *Philosophy & Social Criticism*, *48*(4), 496-514. https://doi.org/10.1177/01914537221089363

Mann, T. E., & Ornstein, N. J. (2016). *It's even worse than it looks: How the American constitutional system collided with the new politics of extremism*. Basic Books.

Marchal, N. (2022). “Be Nice or Leave Me Alone”: An Intergroup Perspective on Affective Polarization in Online Political Discussions. *Communication Research*, *49*(3), 376-398.

Martin, J., & Burns, A. (2023). *This Will Not Pass: Trump, Biden, and the Battle for America's Future*. Simon and Schuster.

Mason, L. (2015). “I disrespectfully agree”: The differential effects of partisan sorting on social and issue polarization. *American Journal of Political Science, 59*(1), 128-145. https://doi.org/10.1111/ajps.12089

Mayhew, D. R. (2004). *Congress: The Electoral Connection*. Yale University Press.

McCarty, N., Poole, K. T., & Rosenthal, H. (2016). *Polarized America: The dance of ideology and unequal riches*. MIT Press.

McCarty, N. (2019). *Polarization: What everyone needs to know*. Oxford University Press.

McConnell, A., & Hart, P. (2019). Inaction and public policy: understanding why policymakers ‘do nothing.’ *Policy Sciences*, *52*(4), 645–661. https://doi.org/10.1007/s11077-019-09362-2

Neal, Z. P. (2020). A sign of the times? Weak and strong polarization in the US Congress, 1973–2016. *Social Networks*, *60,* 103-112. https://doi.org/10.1016/j.socnet.2018.07.007

New York Primary Results - Election 2010. (2010). https://www.nytimes.com/elections/2010/results/primaries/new-york.html

Pew Research Center. (2019, October 10). *Partisan Antipathy: More Intense, More Personal*. https://www.pewresearch.org/politics/2019/10/10/partisan-antipathy-more-intense-more-personal/

Pew Research Center (2020, May 27). *The partisan landscape and views of the parties*. https://www.pewresearch.org/politics/2019/10/10/the-partisan-landscape-and-views-of-the-parties/

Pew Research Center. (2022, August 9). *As Partisan Hostility Grows, Signs of Frustration With the Two-Party System*. https://www.pewresearch.org/politics/2022/08/09/as-partisan-hostility-grows-signs-of-frustration-with-the-two-party-system/

Porumbescu, G., Bellé, N., Cucciniello, M., & Nasi, G. (2017). Translating policy transparency into policy understanding and policy support: Evidence from a survey experiment. *Public Administration*, *95*(4), 990–1008. https://doi.org/10.1111/padm.12347

Prior, M. (2013). Media and political polarization. *Annual Review of Political Science*, *16*, 101-127. https://doi.org/10.1146/annurev-polisci-100711-135242

Ridge, H. M. (2020). Enemy Mine: Negative Partisanship and Satisfaction with Democracy. *Political Behavior*. https://doi.org/10.1007/s11109-020-09658-7

Rogowski, J. C., & Sutherland, J. L. (2016). How Ideology Fuels Affective Polarization. *Political Behavior*, *38*(2), 485–508. https://doi.org/10.1007/s11109-015-9323-7

Rudolph, T. J., & Hetherington, M. J. (2021). Affective Polarization in Political and Nonpolitical Settings. *International Journal of Public Opinion Research*, *33*(3), 591–606. https://doi.org/10.1093/ijpor/edaa040

Russell, A. (2018). US senators on Twitter: Asymmetric party rhetoric in 140 characters. *American Politics Research*, *46*(4), 695-723. https://doi.org/10.1177/1532673X17715619

Russell, A. (2021). *Tweeting is leading: how senators communicate and represent in the age of Twitter*. Oxford University Press.

Schmuck, D., & Hameleers, M. (2020). Closer to the people: A comparative content analysis of populist communication on social networking sites in pre-and post-election periods. *Information, Communication & Society*, *23*(10), 1531-1548. https://doi.org/10.1080/1369118X.2019.1588909

Shaikh, S. (2021). Twitter-scraper-selenium. Retrieved from https://github.com/shaikhsajid1111/twitter-scraper-selenium

Sides, J., & Haselswerdt, J. (2019). Campaigns and elections. *New Directions in Public Opinion*, 297-314.

Sievert, J., & McKee, S. C. (2019). Nationalization in US Senate and gubernatorial elections. *American Politics Research*, *47*(5), 1055-1080. https://doi.org/10.1177/1532673X18792694

Somin, I. (2014). The Problem of Political Ignorance. *The Regulatory Review*. https://www.theregreview.org/2014/02/03/03-somin-political-ignorance/

Stroud, N. J. (2010). Polarization and Partisan Selective Exposure. *Journal of Communication, 60*(3), 556–576. https://doi.org/10.1111/j.1460-2466.2010.01497.x

Sznajd-Weron, K., & Sznajd, J. (2005). Who is left, who is right?. *Physica A: Statistical Mechanics and its Applications*, *351*(2-4), 593-604.

Tillery, A. B. (2019). Tweeting racial representation: How the congressional Black Caucus used Twitter in the 113th congress. *Politics, Groups, and Identities*. https://doi.org/10.1080/21565503.2019.1629308

The Political groups of the European Parliament. (2019). European Parliament. https://www.europarl.europa.eu/about-parliament/en/organisation-and-rules/organisation/political-groups

Theocharis, Y., Barberá, P., Fazekas, Z., & Popa, S. A. (2020). The Dynamics of Political Incivility on Twitter. *SAGE Open*, *10*(2), 215824402091944. https://doi.org/10.1177/2158244020919447

Unkel, J., & Kümpel, A. S. (2022). Patterns of Incivility on U.S. Congress Members' Social Media Accounts: A Comprehensive Analysis of the Influence of Platform, Post, and Person Characteristics. *Frontiers in Political Science, 4*. https://doi.org/10.3389/fpos.2022.809805

Volden, C., & Wiseman, A. E. (2018). Legislative effectiveness in the united states senate. *The Journal of Politics*, *80*(2), 731-735. https://doi.org/10.1086/697121

Webster, S. W., & Abramowitz, A. I. (2017). The Ideological Foundations of Affective Polarization in the U.S. Electorate. *American Politics Research*, *45*(4), 621–647. https://doi.org/10.1177/1532673x17703132

West, D. M. (2017). *Air wars: Television advertising and social media in election campaigns, 1952-2016*. CQ Press.

Yu, X., Wojcieszak, M., & Casas, A. (2021). Affective polarization on social media: In-party love among American politicians, greater engagement with out-party hate among ordinary users. (*Manuscript*).

Zhang, H, & Ma, S. (2023). PolarBert. Retrieved from https://github.com/PolarComm/PolarBert

APPENDIX A: TABLES

Table 1

Machine Learning Performance on Evaluation and Test Datasets

	Evaluation Dataset (K-Fold=10)			Test Dataset 1			Test Dataset 2		
	Accuracy	F1(label =1)	F1(Weighted)	Accuracy	F1(label =1)	F1(Weighted)	Accuracy	F1(label =1)	F1(Weighted)
(Overall) Expressed affective polarization	0.90	0.85	0.90	0.93	0.80	0.92	0.91	0.77	0.90
Policy-embedded in-party liking	0.95	0.68	0.95	0.97	0.70	0.97	0.96	0.71	0.98
Policy-embedded out-party disliking	0.95	0.81	0.95	0.95	0.81	0.97	0.95	0.82	0.95
Character-only in-party liking	0.97	0.70	0.96	0.98	0.71	0.99	0.97	0.70	0.98

Table 2

Descriptive Analysis of Congress Members on Their Party, Congressional Identity, Power of Control, and Seniority

Term	Congressional Identity		Power Of Control		Seniority		Party	
	House	Senate	Minority	Majority	First Incumbent	Re-elected	Democrat	Republican
112	330	84	188	226	104	310	194	220
113	375	99	220	254	102	372	235	239
114	383	98	217	264	65	416	217	264
115	394	91	228	257	69	416	228	257
116	414	95	230	279	96	413	270	239
117	409	93	249	253	67	435	253	249

Table 2 (cont'd).

Total (N)	2305	560	1332	1533	503	2362	1397	1468

Table 3

Descriptive Analysis of IVs and DVs

	Term 112-117 (2011-2022)						Term 112	Term 113	Term 114	Term 115	Term 116	Term 117
Variable	Mean	S.D.	Min	Max	Skewness	Kurtosis	Mean	Mean	Mean	Mean	Mean	Mean
Policy-liking	98	134	0	1669	3.458	19.191	30	39	48	89	168	194
Policy-disliking	209	367	0	4452	4.163	25.697	66	111	85	290	319	348
Character-liking	40	60	0	870	5.443	49.135	21	28	43	45	53	47
Overall-polar	312	437	0	5063	3.508	18.948	104	158	159	387	481	529
Total tweets	1579	1493	1	19732	2.586	13.569	865	1257	1407	1678	2037	2077
Policy-liking percentage	.058	.047	0	.36	1.624	4.010	.032	.033	.036	.057	.088	.094
Policy-disliking percentage	.111	.112	0	.66	1.562	2.337	.072	.085	.062	.128	.135	.172
Character-liking percentage	.027	.037	0	1	11.469	233.443	.026	.023	.032	.030	.029	.023
Overall-polar percentage	.175	.121	0	1	1.234	2.016	.116	.126	.117	.194	.224	.258
General vote-share	64	13	0	100	.689	2.842	63	63	66	66	64	63
Primary vote-share	79	22	15	100	-.750	-.496	81	78	80	79	78	80
Legislative Effectiveness Score	.989	1.053	0	1.3	2.764	12.468	.931	.966	.973	1.008	1.022	1.021
DW1-Absolute	.443	.159	0	1	.568	.596	.435	.437	.444	.448	.440	.452
People in each term (N)				2865			414	474	481	485	509	502

Table 4

The Results of Heteroscedasticity, Auto-correlation, and Cross-sectional Dependence

Model Name	Heteroscedasticity test	Auto-correlation test	Cross-sectional Dependence test
Overall-polar model	$\chi^2(796) = 4.9E+32^{***}$	$F=237.57^{***}$	YES

Table 4 (cont'd).

Policy-liking model	χ^2 (796) =2.9E+33***	F=444.04***	YES
Policy-disliking model	χ^2 (796) =3.8E+30***	F=356.10***	YES
Character-liking model	χ^2 (796) =4.1E+35***	F=42.58***	YES

Note. *** *p*<0.001, ** *p*<0.01, * *p*<0.05, $^{+}$ *p*<0.1.

Table 5

Modified Hausman Test

	Hausman test (sigmamore)	Hausman test (sigmaless)
Overall-polar model	χ^2(13) =50.47***	χ^2(13) =51.00***
Policy-liking model	χ^2(13) =95.71***	χ^2(13) =97.81***
Policy-disliking model	χ^2(13) =50.96***	χ^2(13) =51.52***
Character-liking model	χ^2(13) =57.95***	χ^2(13) =58.70***

Note. *** *p*<0.001, ** *p*<0.01, * *p*<0.05, $^{+}$ *p*<0.1.

Table 6

The Influence of Ideological Extremity, Legislation Performance, and Electoral Incentives on Expressed Affective Polarization

	Model 1: $\log_{10}$ (Overall Expressed Affective Polarization)		Model 2: $\log_{10}$ (Policy-liking)		Model 3: $\log_{10}$ (Policy-disliking)		Model 4: $\log_{10}$ (Character-liking)	
IV	Coef. (S.E.)	β	Coef. (S.E.)	β	Coef. (S.E.)	β	Coef. (S.E.)	β
DW-Nominate	.31* (.10)	.08	.17* (.06)	.05	0.47** (.09)	0.10	0.29* (.09)	0.09
Legislative Effectiveness Score	.00 (.00)	.00	.01 (.01)	.02	0.00 (.01)	0.00	0.00 (.01)	-0.01
General Election	.00 (.00)	-.05	.00 (.00)	.01	-0.01^{+} (.00)	-0.11	0.00 (.00)	0.03
Primary Election	.00 (.00)	.00	.001** (.00)	.03	0.00 (.00)	-0.01	0.001** (.00)	0.02
Senator=1	-.15** (.04)	-.10	.06 (.08)	.04	-0.18* (.06)	-0.10	-0.20*** (.01)	-0.15
Majority=1	-.09* (.03)	-.07	.03 (.06)	.02	-0.23^{+} (.11)	-0.16	0.06* (.02)	0.06
Seniority=1	-.01 (.02)	-.01	.06^{+} (.03)	.04	-0.04 (.06)	-0.02	-0.02 (.03)	-0.01
Number of Tweets	.99*** (.02)	.79	.77*** (.03)	.62	1.09*** (.05)	0.72	0.66*** (.01)	0.60

Table 6 (cont'd).

Term 112	dummy	dummy	dummy	dummy
Term 113	.03*** (.00)	.02** (.00)	0.02^{+} (.01)	-0.04** (.01)
Term 114	.01** (.00)	.04*** (.00)	-0.06*** (.01)	0.12*** (.01)
Term 115	.22*** (.00)	.27*** (.00)	0.09*** (.01)	0.16*** (.01)
Term 116	.31*** (.01)	.47*** (.01)	0.28*** (.01)	0.12*** (.01)
Term 117	.36*** (.01)	.49*** (.01)	0.31*** (.01)	0.07** (.02)
_cons	-1.19 (.05)	-.93 (.12)	-2.04 (.16)	-0.72 (.05)
R-squared	.7894	.7262	.5093	.4715
F-statistic	F(13,5)=3148.81	F(13,5)=310.71	F(13,5)=164.63	F(13,5)=300.47
Prob(F)	<0.001	<0.001	<0.001	<0.001
N(obs)	2753	2753	2753	2753
N(groups)	796	796	796	796

Note. *** *p*<0.001, ** *p*<0.01, * *p*<0.05, $^{+}$ *p*<0.1.

Table 7

The Results of Heteroscedasticity, Auto-correlation, and Cross-sectional Dependence (DVs as Percentages)

Model Name	Heteroscedasticity test	Auto-correlation test	Cross-sectional Dependence test
Overall-polar model	$\chi^2(796) = 5.7E+33^{***}$	F=230.02***	YES
Policy-liking model	χ^2 (796) =1.0E+34***	F=361.79***	YES
Policy-disliking model	χ^2 (796) =1.8E+35***	F=376.78***	YES
Character-liking model	χ^2 (796) =8.1E+33***	F=43.38***	YES

Note. *** *p*<0.001, ** *p*<0.01, * *p*<0.05, $^{+}$ *p*<0.1.

Table 8

Modified Hausman Test (DVs as Percentages)

	Hausman test (sigmamore)	Hausman test (sigmaless)
Overall-polar model	$\chi^2(13) = 49.04^{***}$	$\chi^2(13) = 49.45^{***}$
Policy-liking model	$\chi^2(13) = 65.60^{***}$	$\chi^2(13) = 65.69^{***}$
Policy-disliking model	$\chi^2(13) = 44.18^{***}$	$\chi^2(13) = 44.50^{***}$
Character-liking model	$\chi^2(13) = 31.96^{**}$	$\chi^2(13) = 31.85^{**}$

Note. *** *p*<0.001, ** *p*<0.01, * *p*<0.05, $^{+}$ *p*<0.1.

Table 9

The Influence of Ideological Extremity, Legislation Performance, and Electoral Incentives on Expressed Affective Polarization (DVs as Percentages)

	Model 5: log_{10} (Overall Expressed Affective Polarization_percent)		Model 6: log_{10} (Policy-liking_percent)		Model 7: log_{10} (Policy-disliking_percent)		Model 8: log_{10} (Character-liking_percent)	
IV	Coef. (S.E.)	β	Coef. (S.E.)	β	Coef. (S.E.)	β	Coef. (S.E.)	β
DW-Nominate	.31* (.10)	.15	.17* (.04)	.07	.47** (.10)	.14	.28* (.08)	.12
Legislative Effectiveness Score	.00 (.00)	.00	.01 (.01)	.02	.00 (.01)	.01	-.01 (.01)	-.02
General Election	.00 (.00)	-.09	.00 (.00)	.02	-.01^{+} (.00)	-.16	.00 (.00)	.03
Primary Election	.00 (.00)	-.01	.001** (.00)	.05	.00 (.00)	-.02	.001** (.00)	.04
Senator=1	-.15** (.04)	-.19	.04 (.08)	.04	-.17* (.01)	-.13	-.23*** (.01)	-.25
Majority=1	-.09* (.03)	-.14	.02 (.06)	.02	-.23^{+} (.11)	-.22	.05 (.03)	.06
Seniority=1	-.01 (.02)	-.01	.06^{+} (.02)	.06	-.04 (.05)	-.03	-.01 (.03)	-.01
Term 112	dummy		dummy		dummy		dummy	
Term 113	.03*** (.00)		-.03*** (.00)		.03*** (.00)		-.10*** (.01)	
Term 114	.01^{+} (.00)		.00 (.01)		-.05* (.01)		.05** (.01)	
Term 115	.22*** (.01)		.21*** (.01)		.11** (.02)		.07** (.01)	
Term 116	.31*** (.01)		.40*** (.01)		.31*** (.01)		.01 (.02)	
Term 117	.36*** (.01)		.41*** (.01)		.34*** (.02)		-.04 (.02)	
_cons	-1.22 (.05)		-1.56 (.08)		-1.79 (.15)		-1.65 (.06)	
R-squared	.3672		.3972		.1729		.0725	
F-statistic	F(12,5)=1809.03		F(12,5)=9924.91		F(12,5)=154.18		F(12,5)=2228.89	
Prob(F)	<0.001		<0.001		<0.001		<0.001	
N(obs)	2753		2753		2753		2753	
N(groups)	796		796		796		796	

Note. *** p<0.001, ** p<0.01, * p<0.05, $^{+}$ p<0.1.

Table 10

The Influence of Ideological Extremity, Legislation Performance, and Electoral Incentives on Expressed Affective Polarization (Time-lagged Models)

	Model 9: $\log_{10}$ (Overall Expressed Affective Polarization_lag1)		Model 10: $\log_{10}$ (Policy-liking_lag1)		Model 11: $\log_{10}$ (Policy-disliking_lag1)		Model 12: $\log_{10}$ (Character-liking_lag1)	
IV	Coef. (S.E.)	β	Coef. (S.E.)	β	Coef. (S.E.)	β	Coef. (S.E.)	β
DW-Nominate$_{(t-1)}$	.06 (.07)	.02	$.12^{+}$ (.05)	.03	-.05 (.12)	-.01	$.26^{**}$ (.03)	.08
Legislative Effectiveness Score$_{(t-1)}$	$-.03^{+}$ (.01)	-.04	$.04^{+}$ (.01)	.06	-.08 (.04)	-.11	.00 (.01)	.00
General Election	.00 (.00)	-.04	.00 (.00)	.03	$-.01^{+}$ (.00)	-.11	.00 (.00)	.02
Primary Election	$-.001^{+}$ (.00)	-.03	$.001^{**}$ (.00)	.04	$-.002^{+}$ (.00)	-.05	.00 (.00)	.02
Senator=1	$-.23^{***}$ (.02)	-.15	-.08 (.10)	-.05	$-.26^{**}$ (.06)	-.14	$-.26^{**}$ (.03)	-.19
Majority=1	$-.09^{+}$ (.03)	-.07	.03 (.06)	.02	-.23 (.11)	-.16	$.07^{+}$ (.03)	.07
Seniority=1	-.01 (.02)	-.01	.01 (.03)	.00	-.07 (.08)	-.04	.05 (.04)	.03
Number of Tweets	1.04^{***} (.04)	.83	$.74^{***}$ (.02)	.59	1.19^{***} (.08)	.79	$.68^{***}$ (.01)	.63
Term 113	dummy		dummy		dummy		dummy	
Term 114	$-.03^{***}$ (.00)		$.03^{***}$ (.00)		$-.11^{***}$ (.00)		$.16^{***}$ (.00)	
Term 115	$.19^{***}$ (.00)		$.24^{***}$ (.00)		$.07^{***}$ (.01)		$.19^{***}$ (.01)	
Term 116	$.27^{***}$ (.00)		$.45^{***}$ (.01)		$.23^{***}$ (.01)		$.13^{***}$ (.00)	
Term 117	$.32^{***}$ (.00)		$.45^{***}$ (.01)		$.28^{***}$ (.01)		$.10^{***}$ (.01)	
_cons	-1.16 (.13)		-.63 (.01)		-2.04 (.28)		-.91 (.07)	
R-squared	.7632		.6879		.4857		.4269	
F-statistic	F(12,4)=10114.78		F(12,4)=31.49		F(12,4)=22.99		F(12,4)=181.89	
Prob(F)	<.001		<.01		<.01		<.001	
N(obs)	1969		1969		1969		1969	
N(groups)	644		644		644		644	

Note. $^{***}p$<0.001, $^{**}p$<0.01, $^{*}p$<0.05, ^{+}p<0.1.

Table 11

The Influence of Ideological Extremity, Legislation Performance, and Electoral Incentives on Expressed Affective Polarization (Time-lagged Models and DVs as Percentages)

	Model 13: $\log_{10}$ (Overall Expressed Affective Polarization_percent_lag1)		Model 14: $\log_{10}$ (Policy-liking_percent_lag1)		Model 15: $\log_{10}$ (Policy-disliking_percent_lag1)		Model 16: $\log_{10}$ (Character-liking_percent_lag1)	
IV	Coef. (S.E.)	β	Coef. (S.E.)	β	Coef. (S.E.)	β	Coef. (S.E.)	β
DW-Nominate$_{(t-1)}$	.06 (.07)	.03	.12^{+} (.05)	.05	-.05 (.12)	-.02	.26** (.03)	.11
Legislative Effectiveness Score$_{(t-1)}$	-.03^{+} (.01)	-.08	.04^{+} (.01)	.11	-.08 (.04)	-.15	.00 (.01)	.00
General Election	.00 (.00)	-.07	.00 (.00)	.06	-.01^{+} (.00)	-.15	.00 (.00)	.03
Primary Election	-.001^{+} (.00)	-.05	.001** (.00)	.07	-.002^{+} (.00)	-.07	.00 (.00)	.03
Senator=1	-.23*** (.02)	-.29	-.08 (.10)	-.09	-.26** (.06)	-.20	-.26** (.03)	-.28
Majority=1	-.09^{+} (.03)	-.14	.03 (.06)	.04	-.23 (.11)	-.23	.07^{+} (.03)	.10
Seniority=1	-.01 (.02)	-.01	.01 (.03)	.01	-.07 (.08)	-.05	.05 (.04)	.05
Number of Tweets	.04 (.04)	.06	-.26*** (.02)	-.36	.19^{+} (.08)	.18	-.32*** (.01)	-.42
Term 113	dummy		dummy		dummy		dummy	
Term 114	-.03*** (.00)		.03*** (.00)		-.11*** (.00)		.16*** (.00)	
Term 115	.19*** (.00)		.24*** (.00)		.07*** (.01)		.19*** (.01)	
Term 116	.27*** (.00)		.45*** (.01)		.23*** (.01)		.13*** (.00)	
Term 117	.32*** (.00)		.45*** (.01)		.28*** (.01)		.10*** (.01)	
_cons	-1.16 (.13)		-.63 (.08)		-2.04 (.28)		-.91 (.07)	
R-squared	.3895		.4736		.2166		.1868	
F-statistic	F(12,4)=16.61		F(12,4)=31.49		F(12,4)=22.99		F(12,4)=181.89	
Prob(F)	<.01		<.01		<.01		<.001	
N(obs)	1969		1969		1969		1969	
N(groups)	644		644		644		644	

Note. *** $p<0.001$, ** $p<0.01$, * $p<0.05$, $^{+}$ $p<0.1$.

APPENDIX B: FIGURES

Figure 1

Current Proposed Mechanism on Expressed Affective Polarization

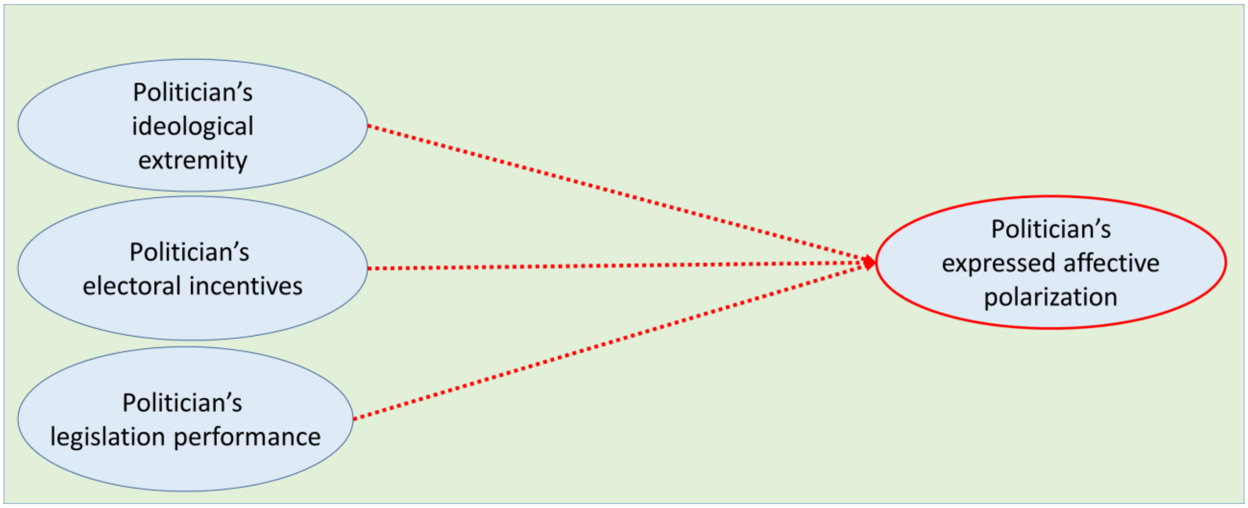


Figure 2

Hypotheses Model

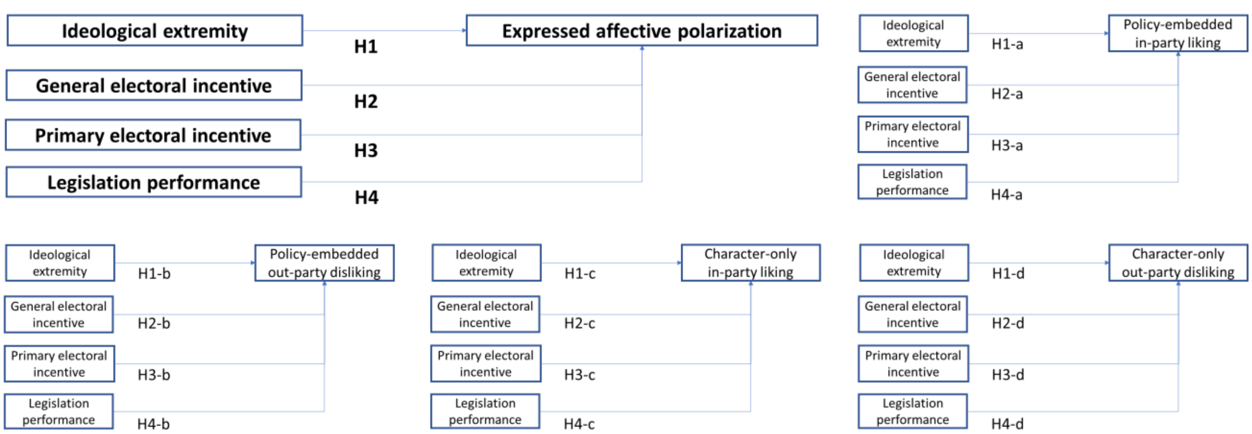
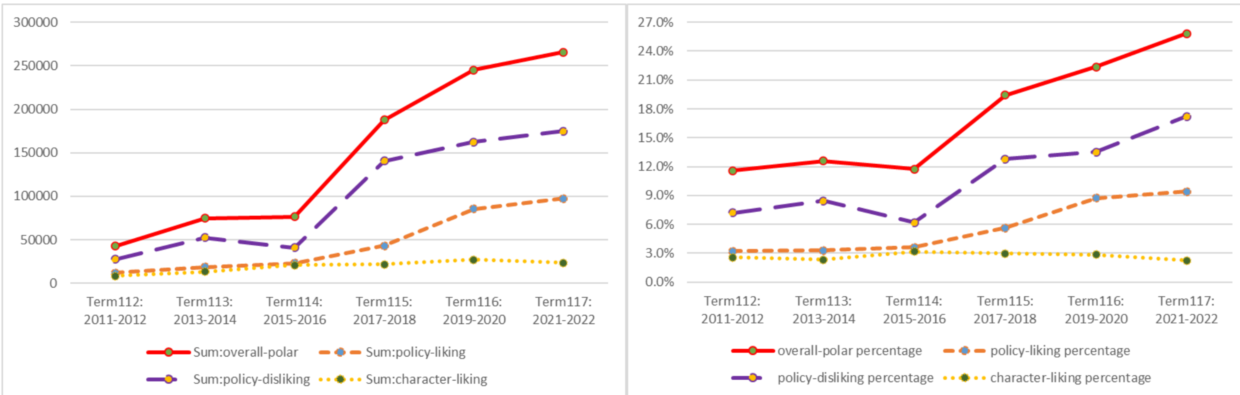


Figure 3

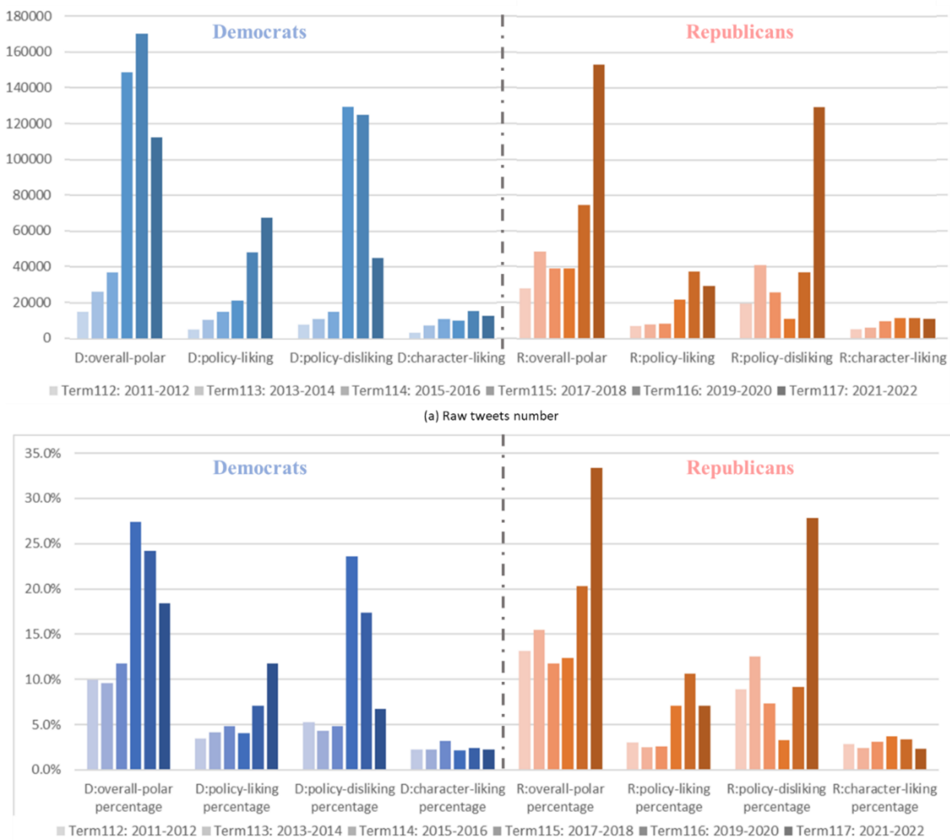
The Raw Number and Percentage of Expressed Affective Polarization Tweets From 2011-2022



(a) Raw tweets number (b) Percentage of tweets

Figure 4

Numbers and Percentages of Expressed Affective Polarization Tweets Split by Party



(a) Raw tweets number

(b) Percentage of tweets

Figure 5

The Society-Level Expressed Affective Polarization Dynamic During Utd-City Team Conflicts

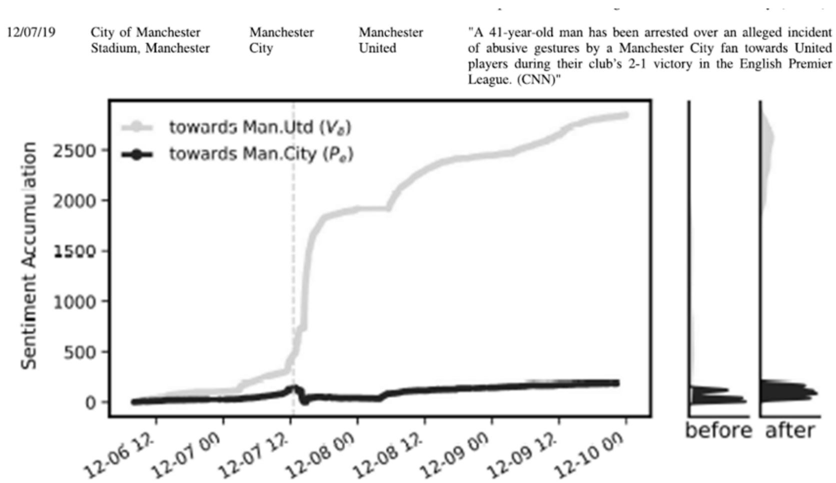


Figure 6

Examples of Tweets That Do and Do Not Contain Expressed Affective Polarization

Sender	Tweet Text	Classification
Rep. Alexandria Ocasio-Cortez (D-NY 14)	The GOP Senate needs to stop playing games and pass the 9/11 Victim's Compensation Fund.	Polarizing
Rep. Brendan Boyle (D-PA 13)	Republicans are choosing party over country. Plain and simple. History will not be kind to those putting blind partisanship ahead of the truth.	Polarizing
Rep. Adam Kinzinger (R-IL 16)	You know Ds are in trouble when @HillaryClinton their "reasonable" candidate is getting crushed by a 74 yr old socialist.	Polarizing
Rep. Bradley Byrne (R-AL 1)	This makes me angry but it is honestly just so sad. The Democrats are actually wishing death on people who step up to protect the unborn.	Polarizing
Sen. Dean Heller (R-NV)	Tax relief is not a partisan issue, it's common sense.	Not Polarizing
Rep. Antonio Delgado (D-NY 19)	The first task for Congress will be putting aside political games and re-opening the government. I hope that Democrats and Republicans alike can work together to do so.	Not Polarizing
Sen. Amy Klobuchar (D-MN)	After a decade of progress, the rate of uninsured children is now rising. This is very troublesome. Every child needs access to quality, affordable healthcare.	Not Polarizing
Rep. Beto O'Rourke (D-TX 16)	If you're angry or upset, if you want to do something, you can make a difference. Here's how you can help the families affected by the raids in Mississippi	Not Polarizing

APPENDIX C: CODING SCHEME

Coding Protocol for Identifying the Tweets Contributing to Expressed Affective Polarization

[Read through the *tweets* listed in the file. The *tweet* has two types, (1) the politicians' original post only or (2) the politicians' retweeting comment on other people's post. Code 1 if the tweet fits the situation; code 0 if it does not. When it is hard to identify the situation of the retweeting comment, we can use other people's original post (if we have it) for reference.]

1. Policy-embedded in-party liking tweet: It refers to tweets that politicians have expressed positive attitudes toward (1) their own party's name, (2) the organization name affiliated with their own party, or (3) the politician from their own party. The tweets should also contain evaluations of policy or social problems, whether the evaluation is supportive, critical, or neutral. A policy means a bill, bill draft, executive order, or administrative plan. A social problem means that the issue has a broad negative or controversial impact on US citizens (e.g., Immigrants, The kidnapping of Americans, Massive gun shootings). Calling for concrete political actions for ideologies (e.g., protecting freedom and democracy) is also considered to be a social problem. Simply attacks or boasts toward parties, party-affiliated organizations, and politicians are not social problems, even though some of them are wildly spread (e.g., "Trump in the bunker", "Let's go Brandon").

All likings and evaluations should be clear-stated literally. Coders should not try to extend the hidden meaning of the tweet. Tweets containing policy-embedded in-party liking should be coded as 1, or else they should be coded as 0.

2. Policy-embedded out-party disliking: It refers to tweets that politicians have expressed negative attitudes toward (1) their opposite party's name, (2) the organization name affiliated with their opposite party, or (3) the politician from their opposite party. The tweets should also

contain evaluations of policy or social issues, whether the evaluation is supportive, critical, or neutral. Tweets containing policy-embedded out-party disliking should be coded as 1, or else they should be coded as 0.

3. Character-only in-party liking: It refers to tweets that politicians have expressed positive attitudes toward (1) their own party's name, (2) the organization name affiliated with their own party, or (3) the politician from their own party. The tweets should not contain evaluations of policy or social problems, whether the evaluation is supportive, critical, or neutral. Tweets containing character-only in-party liking should be coded as 1, or else they should be coded as 0.

Congratulate political characters for private issues (politician's birthday, Christmas) without any policy evaluations should be considered character-only in-party liking.

4. Character-only out-party disliking: It refers to tweets that politicians have expressed negative attitudes toward (1) their opposite party's name, (2) the organization name affiliated with their opposite party, or (3) the politician from their opposite party. The tweets should not contain evaluations of policy or social problems, whether the evaluation is supportive, critical, or neutral. Tweets containing character-only out-party disliking should be coded as 1, or else they should be coded as 0.

Examples:

Policy-embedded in-party liking – evaluating social problems and mentioning in-party politicians. [After months of imprisonment, Brittney Griner is free and returning home to her family and loved ones. Thank you to President Biden and everyone who worked tirelessly to make this happen.]

Policy-embedded in-party liking – evaluating policy and mentioning in-party politicians. [Congratulations to @BethDoglio on advancing to the general election. She will fight for a Green New Deal and health care for all. Let’s do everything we can to put her in Congress.]

Policy-embedded out-party disliking – [I’m LIVE on the floor of the U.S. Senate to ask a simple question: Are any Republicans prepared to stand with rail workers who have ZERO paid sick days or are they instead going to back the outrageous greed of the rail industry?]

Character-only in-party liking – [Thank you President Trump, for continuing to be the leader of the Republican party and helping our conference unite. We are ready to get to work to Make America Great Again!]

Character-only in-party liking – [Happy Birthday, Rep. @JimLangevin!]

Character-only out-party disliking – [The plan is to stop the Democrats!!! Wake up to political reality and remember who the enemy is.]

NONE – policy yes, but no clear out-party disliking [I support @BCTGM Ingredion workers in their 20th week on strike in Cedar Rapids. If Ingredion can return $250 million this year to shareholders and pay its CEO $10 million last year, it can afford a fair contract that does not increase health costs, decrease wages, or cut jobs.]

NONE – evaluating policy but no in-party liking. It is a Democrat party infighting. [Congratulations to the 750 environmental groups who successfully defeated Sen. Manchin‘s Big Oil Side Deal tonight. Yes. We need to improve our transmission capabilities, but not for fossil fuels. In order to save the planet, we need to rapidly move to sustainable energy.]

NONE – no policy and no party. [Wishing a peaceful and merry Christmas from myself, Jane, and our entire family to all those who are celebrating.]

NONE – USNavy is not a partisan organization [Happy Birthday @USNavy and thank you to the great women and men in the Office of Naval Intelligence America's longest serving intelligence agency]

NONE – republican party infighting [⋯⋯⋯⋯In these trying times, we must fight for everything sacred. ⋯⋯⋯⋯ Alas without a Speaker, we are powerless..]

NONE – sent by Democrats, a bi-party cooperation [Had a great time exploring @ArchesNPS with @SenatorRomney. While I'm partial to @HarpersFerryNPS and @NewRiverNPS, my Great American Outdoors Act will help protect and preserve public lands across the country and encourage more people to visit our Wild and Wonderful parks.]

5. If the tweet itself is hard to understand literally, it should be coded as 0 for all categories.

Examples:

NONE – we do not try to extend the hidden meaning. [To watch my full remarks, please click here.]

Add-ons:

Tweets calling for bi-party cooperation will not be considered as any category of policy-embedded in-party liking, policy-embedded out-party disliking, character-only in-party liking, or character-only out-party disliking.

The user himself/herself (the politician) should also be considered as part of their party. So, paraphrasing other people saying something nice about the user him/herself should be considered as an in-party liking (e.g., I got support from the president, and he spoke highly of my legislation performance). Simply self-praise does not contribute to in-party liking (e.g., Last year, I made great efforts to support the railway workers to get paid properly). And, simply endorsement like “I got support from another in-party member” without clear praise does not contribute to the in-party liking.

A policy-embedded in-party liking tweet only means that the tweet mentions both in-party liking and the policy. The policy does not necessarily be the reason for in-party liking. So do the out-party disliking tweets.

The opposite party criticizing the POTUS, or “the current government as a whole,” should be considered as contributing to out-party disliking.

If you see some celebrities, local politicians, or organizations’ names, use Wikipedia to identify their party affiliation. If Wikipedia does not have enough information to identify their party affiliation, use Balletpedia. If both sources do not have results, the tweet will not be considered any category of expressed affective polarization.

If the tweet evaluates policies that are literally named for a party or a politician (e.g., TrumpCut, ObamaCare), we should code it as policy-embedded in-party liking or policy-embedded out-party disliking.

When you feel the whole tweet really means an out-party disliking (or in-party liking), but cannot find a keyword to identify this feeling, try to exclude the possibility of other explanations. If the whole tweet does not have polysemy, we still consider it as an out-party disliking.

Supreme Court starting from 2019 to 2022, and Kavanaugh and Barrett are considered Republicans. Starting from FDR, even though the politician has passed away, we still consider his/her party affiliation.

Simply “thank you” will not be considered as in-party liking.

Urging somebody to do something is not necessary to be a criticism. Normative prescriptions of actions are not polarizing unless there is also negative language.

Conservative is treated as the synonym of Republican. Liberal is treated as the synonym of Democrat. MAGA, Communist, and Socialist are not considered as Republican or Democrat.

APPENDIX D: THREE PREVIOUS MEASUREMENTS RELATED TO EXPRESSED AFFECTIVE POLARIZATION

Previous studies have explored some measurements related to expressed affective polarization with social media data.

Measurement 1: Identify sentiments by supervised machine learning in each post and use the sentiment results to integrate a higher-level expressed affective polarization. Yu et al. (2021) have proposed details of this measurement for a group-level expressed affective polarization. Step 1. Researchers collected all Twitter posts of 116th Congress members in the US, the 2000-2020 Presidential candidates, and the Trump cabinet members. Retweets without comments were removed. Step 2. Researchers selected tweets mentioning either party or key politicians. Step 3. Two trained coders manually annotated a random sample of tweets (n = 10,000, with 50% of tweets targeting each party) for whether they were negative, neutral, or positive toward the targeted party. Step 4. Researchers used a supervised machine learning approach to predict the tone of the remaining tweets. Step 5. Researchers synthesized tweet sentiments to a group level, and found that among all elite tweets discussing the two parties and key politicians, most were suggestive of affective polarization: 33% of the tweets were positive toward the in-party, and 28% were negative toward the out-party. Almost a quarter was neutral toward either the in-party (17%) or out-party (6%). Only on rare occasions did politicians criticize the in-party (6%) and praise the out-party (10%).

Measurement 2: Identify sentiments by dictionary-based approach in each post and use the sentiment results to integrate a higher-level expressed affective polarization. Borrelli et al. (2021) used social media posts during European soccer matches to measure society-level expressed affective polarization. Step 1. Typically, there were some conflicts during the match. Two teams in the match would be treated as one perpetrator and one victim. Researchers identified official Twitter accounts that proposed conflicting attitudes towards the match and named them Group A (perpetrator) and Group B (victim). Step 2. Researchers collected tweets replying to Group A and tweets replying to Group B. Step 3. Researchers calculated the society-level cumulated sentiment with VADER dictionary-based approach. Step 4. Researchers calculated the expressed affective polarization based on the cumulated sentiment. They found the whole society felt more positive about the victim and remained similarly negative about the perpetrator after the conflicts. (see Formula 1 and Figure 5 for details in Steps 4 and 5.)

$$CS_t^{V_e} = \int_{t_0}^{t} Sen_t^{V_e} dt; CS_t^{P_e} = \int_{t_0}^{t} Sen_t^{P_e} dt; p_t^{V_e,P_e} = \frac{CS_t^{V_e}}{N_t^{V_e}} - \frac{CS_t^{P_e}}{N_t^{P_e}} \quad (1)$$

(*Note. Pe* refers to the sentiment of replies to group A. *Ve* refers to the sentiment of replies to group B. *Nt* refers to the number of tweet replies.)

Measurement 3: Identify expressed affective polarization by supervised machine learning in each post but not integrate it into higher-level polarization. Ballard et al. (2022) used this measurement to analyze all tweets from 111-116 Congress members between 2009 and 2020 in the US. Step 1. Researchers hand-coded training data. Four research assistants coded a sample of 4000 tweets for whether they contained expressed affective polarization. Expressed affective polarization is defined as if the language used in the tweet matched our definition of being for the purpose of creating division between the speaker and another group via in-group/out-group identification. (See Figure 6 for details.) Overall, coders agreed 85% of the time, with a Cohen's kappa of 0.61 and Cronbach's alpha of 0.94. Step 2. Researchers used supervised machine learning to code tweets. In their project, they utilized RoBERT. Researchers used 3200 tweets as the training set and 800 tweets as the testing test and found that the model was highly accurate: the out-of-sample predictive accuracy was 90%, with a weighted F1 score of 0.9.