# Social Capital, Political Polarity, and Socioeconomic Status as Predictors of Political Incivility on Twitter: A Congressional District-Level Analysis

A pre-print of the article that appeared in *Social Science Computer Review, 35*(1), 10-32. https://doi.org/10.1177/0894439315602858

**Vargo, C.,** & Hopp, T. (2016).

***Abstract:*** Using 414,322 Tweets drawn from143,404 individual Twitter users located in all 435 U.S. congressional districts, this study employed big data and machine learning techniques to explore the degree to which social capital potential (i.e., the degree to which a congressional district has the potential for interconnected citizen networks), district-level socio-economic factors, and in-district partisan polarization were associated with election oriented incivility on Twitter during the 2012 presidential election. Broadly speaking, and with some exceptions, the results indicated that election oriented incivility on Twitter was highest in districts that had low levels of social capital potential, low socio-economic status indicators, and low levels of partisan polarity.

***Keyword list:*** incivility, big data, twitter, social capitol, 2012 general election, congressional districts, political polarity

***Total word count***: 13,537
***Body text word count***: 9,197

***Submitted to***: *Social Science Computer Review*

**Social Capital, Political Polarity, and Socioeconomic Status as Predictors of Political Incivility on Twitter: A Congressional District-Level Analysis**

Over the past two decades, the rapid proliferation of web-based, socially interactive digital platforms (i.e., "social media") has provoked intense exploration of its democratic potential (e.g., Freelon, 2013; Gil de Zúñiga, Jung, & Valenzuela, 2012; Papacharissi, 2004; Sunstein, 2001; Valenzuela, Kim, & Gil de Zúñiga, 2012). Perhaps unsurprisingly, this exploration has yielded fairly sharp debate (e.g., Dahlberg, 2001; Davis, 2009; Gladwell, 2011; Shirky, 2011; Stromer-Galley & Muhlberger, 2009). Perspectives framing the debate can be segmented into two categories. The first perspective holds that the accessible nature of the Internet will encourage a relative widening of the public sphere and in so doing, "pave the road for a democratic utopia" (Papacharissi, 2004, p. 260). Conversely, the alternate perspective suggests that despite its potential, digital technologies simply reinforce existing participatory gaps and thereby, "frequently induce fragmented, nonsensical, and enraged discussion" (p. 260) that possesses inherently limited democratic value.

In light of the debate over social media's true value to democracy, the present study probed the degree to which an online micro-blogging platform, Twitter, facilitates fruitful and expanded civic involvement or, instead, merely produces disengagement and incivility. Examination of the literature suggests that offline civic involvement has been previously associated with a host of contextual factors, including community heterogeneity, community in/stability, socio-economic conditions, and ideologically-driven partisan polarization (e.g., Costa & Kahn, 2003; Fiorina & Abrams, 2008; Kang & Kwak, 2003). Building upon this research, the current study set out to explore the relationship between contextual-level features associated with individual U.S.

congressional districts and aggregated (at the congressional level) political incivility on Twitter. Using 414,322 Tweets drawn from 143,404 individual Twitter users located in all 435 U.S. congressional districts, this study was specifically interested in the degree to which social capital potential (i.e., the degree to which a congressional district has the potential for interconnected citizen networks), in-district political polarization, and district socioeconomic status (SES) were related to civility on Twitter. Notably, and despite the obvious potential for the Internet to both widen and deepen the public sphere, relatively few studies have attempted to explore the degree to which web-based platforms such as social media sites *actually* facilitate meaningful participation throughout the citizenry. As such, this study analyzed *observed* citizen behavior using big data techniques, which offer novel means of addressing questions relevant to the relationship between social media, society, and democracy, particularly as they relate to the nature of political discussion.

### ***Incivility***

Drawn from the Latin word *incivilis* (which can be roughly translated as *unmannerly, unjust*, and, ultimately, *unbecoming of a citizen*), incivility refers to a wide array of behaviors that range from rudeness and name-calling to forcible theft and hooliganism. In its myriad of forms, civility, as the converse of incivility, is thought to be central to a well-functioning democracy. For instance, Boyd (2006) argued that adherence to the principles of civil engagement helps facilitate social interactions. Civility helps citizens "communicate respect for others and generate habits of moral equality in the everyday of life of a democracy" (p. 863). Specifically on the topic of civil communication, Coe, Kenski, & Rains (2014) noted that "commitment to civil discourse—the free and respectful exchange of ideas—has been viewed as a democratic ideal from the ancient Athenian forums to the mediated political debates of modern times" (p. 658). Grand (2014) summarized the importance of civility by simply noting, "democracy requires democrats" (p. 8).

Incivilities manifest in a number of behavioral forms. According to Boyd (2006), there are two, inter-related connotations of civility. The first of these connotations, referred to by Boyd as *formal civility*, speaks to the manner and tone in which interactions are carried out in everyday life. The second connation, understood as *substantive civility*, denotes a sense of membership in the surrounding social and political community. These dimensions are inherently related to one another, as the former (*formal civility*) can be understood as exerting a direct influence on the latter (*substantive incivility*). Specifically, Boyd notes: "civility is a kind of 'adverbial' restraint on the civic language we speak with one another. In the same way that one is enjoined to speak politely, modestly or temperately, the adverbial condition of civility modifies and qualifies conduct without specifying its content" (p. 864).

Modern society is perforated by concerns that the age of civility (if indeed such a thing ever existed) is in its twilight. Forni (2011), for instance, remarked:

> ***In today's America, incivility is on prominent display: in the schools, where bullying is pervasive; in the workplace, where an increasing number are more stressed out by co-workers than their jobs; on the roads, where road rage maims and kills; in politics, where strident intolerance takes the place of earnest dialogue; and on the Web, where many check their inhibitions at the digital door (par. 1).***

Partially in response to such concerns, communication scholars have increasingly begun to investigate the issue of incivility, particularly as it relates to political incivility online (e.g., Coe, Kenski, & Rains, 2014; Hmielowski, Hutchens, & Cicchirillo, 2014; Papacharissi, 2004; Rowe, 2015; Santana, 2014). Operating under the assumption that computer-mediated civil political discourse is a meaningful component of 21st century political involvement, these studies have focused on the individual user and addressed topics such as the role of usage patterns relative to behavior (e.g., Hmielowski, Hutchens, & Cicchirillo, 2014; Papacharissi, 2004), the relationship between media content/structure and discussion tone (e.g., Borah, 2014; Coe, Kenski, & Rains, 2014), and the effects of uncivil discussion on issue perceptions (Anderson, Brossard, Scheufele, Xenos, & Ladwig, 2014). Focus on individual-level attributes through experiments and surveys helps build a comprehensive understanding on the social effects of incivility. It generally does not address incivility as a phenomenon that is rooted in larger surrounding social factors (e.g., such as the socioeconomic factors in which an individual lives). Moreover, these studies tend to address citizen behavior prospectively rather than when it actually occurs.

Accordingly, the goal of this work was to explore the degree to which contextual factors associated with social life are associated with aggregated patterns of computer-mediated political incivility. In so doing, this study attempted to empirically investigate the degree to which factors commonly associated with *substantive* incivility (i.e., membership in the surrounding social and political community) can also be used to describe patterns of *formal* incivility (i.e., the manner and tone in which social interactions are carried out on Twitter).

***Social Capital***

The theory of social capital has a long, profound, and varied intellectual history that includes, but is certainly not limited to, eminent and dynamic thinkers such as de Tocqueville, Durkheim, and Bourdieu. At its core, the theory holds that associations in networks of citizens help "to sustain civil society and community relations in a way that generates trust and cooperation between citizens and a high level of civic engagement and participation" (Newton, 2001, p. 201). The theory of social capital presupposes that a civil and engaged society is based, to no small degree, on individual citizens' access to "network ties of goodwill, mutual support, shared language, shared norms, social trust, and a sense of mutual obligation that people can derive value from" (Huysman & Wulf, 2004, p.1; cf Ellison, Steinfeld, & Lampe, 2006, p. 7). Broadly speaking, social capital is an elastic concept that is commonly thought of as both a *cause* and an *effect* (Ellison, Steinfeld, & Lampe, 2007).

Social capital can be compartmentalized into two distinct types: *bridging* capital and *bonding* capital. As defined by Gittell and Vidal (1998), bridging social capital refers to ties, primarily weak in nature, that bring together groups of people previously unfamiliar with one another. Conversely, bonded social capital refers to the type of social capital, generally predicated upon the existence of strong ties, "that brings closer together people who already know each other" (p. 15; cf Yuan & Gay, 2006, p.1067). Despite the semantic association with "weak" and "strong" ties, both types of social capital play an instrumental role in the maintenance of behaviors (i.e., open exchange of ideas, task-related research exchange, new knowledge generation, distribution of social support) necessary for a civil and productive society (Granovetter, 1973; Yuan & Gay, 2006).

Specifically, bonding social capital facilitates in-group cohesion, which facilitates interpersonal trust and knowledge sharing (Krackhardt, 1992). Alternately, bridging capital encourages connections between otherwise disconnected networks of people and helps both “facilitate the exchange of information between distinct groups” and expedites “the flow of ideas among groups” (Kavanaugh, Reese, Carroll, & Rosson, 2005).

Prior research has measured social capital on both individual and aggregate levels. On the individual level, prior research has operationalized social capital in terms of the quantity and frequency respondents engage in pro-social civic behaviors (e.g., Blanchard & Horan, 1998; Onyx & Bullen, 2000), political and organizational participation (e.g., Wellman, Haase, Witte, & Hampton, 2001), degree of interpersonal trust (e.g., Lee & Lee, 2010; Shah, 1998), and life contentment (e.g., Shah, Kwak, & Holbert, 2001). On the aggregate level, researchers have conceptualized high levels of community-based social capital as manifest in aggregated composites of individual-level indicators of social capital (e.g., Hendryx, Ahern, Lovrich, & McCurdy, 2002), fraction of the eligible participation who voted in recent elections (e.g., Chamlin & Cochran, 1995; Rosenfeld, Messner, & Baumer, 2001), and structural opportunities for community action (Israel, Beaulieu, & Hartless, 2001).

***Social Media & Social Capitol***

Others have studied social media and the extent to which it generates social capitol or incivility. Facebook pages can also promote social activism, and increase public awareness and optimism about social projects (e.g. such as a preservation attempt) (Marinov & Schimmelfennig, 2015). In a broad sense, digital media use has bolstered the amount of political talk that those with lower levels of political interest (Bimber et al., 2015). Those that are generally unmotivated to talk about politics do seem to talk more when they use digital media to consume news. However, some debate has sparked as to where that "talk" takes place, and the degree to which it is beneficial to society.

In a study of newspaper comments during the 2012 general election, Coffey Kohler and Granger (2015) found that the majority of online discussion was uncivil. The commentators used "shocking" and "disappointing" language that "contained not just insults about the candidates but designations of anyone who supported a candidate" (pg. 262). The authors did find thoughtful arguments amongst the comments, but were "somewhat pessimistic" as the degree to which social media can generate social capitol. They concluded that newer forms of social media (e.g. Twitter) might be better at generating participatory and deliberative democracy. Johannessen echoes this sentiment in concluding that, "online social media platforms, can contribute to increased social capital and increased political debate among citizens" (pg. 2573).

This potential however is not uniform and unconditional. Others have noted that online participation varies greatly from person-to-person and from topic to topic (Lutz, Hoffmann & Meckel, 2014). Researchers have noted that those who contribute to political discussion on Facebook are not very different in terms of gender, ethnicity, citizenship, and educational level of the students' mothers (Vissers & Stolle, 2014). However, participation did positively correlate with Internet news consumption.

***Social Capitol Potential***

As the current methodological approach limited our ability to directly measure social capital, this study was instead primarily concerned with the *potential* for social capital formation. Social capital is, centrally speaking, concerned with individual citizens' access to networks built around social trust and norms of reciprocity. It can thus be reasoned that the presence (or absence) of certain social factors can serve as a conduit for the formation and maintenance of social capital (in both *bridged* and *bonded* forms). At least three factors, existing on the contextual level, should encourage the formation of social capital. First, prior research has shown that racial heterogeneity hampers relationship building and, thus, is negatively associated with social capital formation (e.g., Costa & Kahn, 2003; Leigh, 2006; Yuan & Gay, 2006). Second, research has shown that neighborhood stability allows citizens to form and maintain the networks necessary for the production of social capital, which therefore acts as a vanguard against incivil behavior. Third, we reasoned that familiarity with the broader cultural and social norms that govern local social behavior should encourage social capital formation and lessen the prevalence of incivil online communication behavior. Each of these factors is reviewed in the paragraphs below.[1]

Racial diversity simultaneously offers benefits and imposes costs on communities. Guren, Nagda, and Lopez (2004), for instance, found that campus-wide racial heterogeneity was associated with a number of positive outcomes related to on-campus democratic engagement while Costa and Kahn (2003) argued that racial diversity helps communities retain "epicurean variety" and resist so-called "urban sprawl" (p. 103). Perhaps unfortunately, however, racial heterogeneity has been consistently associated with low levels of community-wide social capital (e.g., Alesina & Ferra, 2002; Costa & Kahn, 2003; Rupasingha, Goetz, & Freshwater, 2006; Vigdor, 2004). According to Yuan and Gay (2006), the tendency for racial diversity to be associated with low levels of social capital formation can be explained through a number of inter-related sociopsychological theories, including *self-categorization theory* (e.g., Turner, 1987), the *similarity-attraction hypothesis* (e.g., Byrne, 1971), and the *theory of homophily* (McPherson & Smith-Lovin, 1987). Put most simply, these theories hold that “similarity breeds connections” (McPherson, Smith-Lovin, & Cook, 2001, p. 415) and that society has embedded within it a “natural aversion to heterogeneity” (Alesina & Ferrara, 2002, p. 225). Building upon this logic, the current study suggested that communities with high levels of racial fragmentation have lower potential for social capital production as citizens are likely to exercise natural preferences for homogenization by restricting their social interactions to densely similar subgroups. We further suggested that such diminished social capital production potential would be associated with increased levels of incivility, resulting in the following hypothesis:

**H1: High levels of racial heterogeneity will be positively associated with high levels of political incivility on Twitter.**

Scholars of community integration and attachment have long noted the negative relationship between residential mobility and social capital production (Kang & Kwak, 2003). Referred to as the *systems model of community attachment* (Kasarda & Janowitz, 1974; cf from Kang & Kwak, 2003), there exists fairly robust empirical support for the notion that length of residence is positively correlated with interest in local/community affairs (e.g., Oropesa, 1992), participation in local government/affairs (e.g., Kasarda & Janowitz, 1974), and the maintenance of local friendships (e.g., Liu, Ryan, Aurbach, & Besser, 1998). As noted by Kan (2007), “rapid inflows and outflows of residents in a neighborhood lead to neighborhood instability” (p. 437), which, in turn, results in social capital deficits. Broadly speaking, the relationship between social capital and residential stability is multidirectional. Residential outflow results in social capital reduction while social capital dearths are, themselves, an impetus for neighborhood abandonment. This study predicted that districts featuring higher levels of residential stability would be more likely to feature high levels of social capital and, therefore be likely to have diminished levels of incivility.

**H2: High levels of neighborhood stability will be negatively related to political incivility on Twitter.**

The production of social capital is predicated on the notion that individual actors have access to an array of social networks that are meaningfully situated around shared norms and a sense of mutual obligation upon which social trust can be built (Huysman & Wulf, 2004). Access to networks of trust and support is reliant, to a non-ignorable degree, on individual citizens' ability to understand the informal norms that parameterize social participation (e.g., Coleman, 1988). Social norms specify "what actions are regarded by a set of persons as proper or correct, or improper and incorrect " (Coleman, 1990, p. 243). Non-compliance with social norms results in a range of negative outcomes ranging from diminished social trust to social ostracization. The potential for social capital production can be meaningfully linked to individuals' ability to adeptly and meaningfully decode and apply relevant social norms. Prior research suggests that such understanding of social norms is rooted, at least partially, in geographic mobility (e.g., Kan, 2006; Welch & Baltzell, 1984; Hagan, MacMillan, & Wheaton, 1996). Communities that feature high levels of inward/outward migration often lack common understanding of social norms. Such collective agreement is necessary for trust building (David, Janiak, & Wasmer, 2010; Schiff, 1992). Thus, we predicted that high levels of district mobility would be positively associated with incivility:

**H3: High levels of geographic mobility will be positively associated with political incivility on Twitter.**

### *Partisan Polarity*

The past two decades have seen increased public and scholarly concern over partisan polarization in the United States (e.g., Brady & Han, 2006; Campbell, Rockman, & Rudalevige, 2008; DiMaggio, Evans, & Bryson, 1996; Fiorina & Abrams, 2008). Much of this research has focused on partisan polarization of elites, specifically relative to the relationship between elite polarization and the presumed polarization of the electorate (e.g., Hetherington, 2001). A smaller but nonetheless substantive body of research has explored *mass polarization*. More often than not, this research has targeted outcomes such as public opinion and voting behaviors (e.g., Klinkner, 2004; Klinkner & Hapanowicz, 2005; Nunn & Evans, 2006; Glaeser & Ward, 2006). Perhaps echoing the broader debate on partisan polarization, there does not seem to be an identifiable consensus on if and how contextual polarization influences civic and political behaviors (Fiorina & Abrams, 2008). Broader examination of the factors that motivate/inhibit social behavior does, however, provide some useful clues for parsing this contradictory evidence, especially in relation to the present research objectives. For the purposes of the present work, we center our discussion on the notion that that hotly contested districts (i.e., *swing* or *battleground* districts) "represent more heterogeneous environments, where people are more likely to encounter counter-attitudinal messages" (Wolak, 2006, p. 354) both in the form of interpersonal interactions and political advertisements. At the individual-level, exposure to counter-attitudinal messages has been previously associated with negative reactions, particularly among those who are highly involved with a particular issue (Krosnick & Petty, 1995). Other research has suggested the existence of a *stimulation effect* whereby conflict-inducing political advertisements (i.e., negative and attack ads) mobilize voter involvement (e.g., Goldstein & Freedman, 2002; Lau, Sigelman, Heldman, & Babbit, 1999). Given that prior scholarship has suggested a relationship between the tone used in political advertising and the tone used by citizens (e.g., Cho, 2013) and that political advertising has become increasingly negative in nature (Farnam, 2012; Rubenstein, 2014), it stands to reason that citizens in battleground districts are likely to engage in comparatively higher levels of incivil discourse.

**H4: High levels of partisan polarity will be negatively associated with political incivility on Twitter.**

***Socioeconomic Status***

Socioeconomic status (SES) has long been understood as a key predictor of social cohesion and civic behavior (Boardman & Robert, 2000; Larsen et al, 2004). As summarized by Oliver (1999), prior research has consistently indicated that civic participation is higher in more affluent settings. Specifically, the "underlying theme throughout these works is that people who are surrounded by more participators (i.e., the educated and affluent) feel more social pressure and are given more opportunities to participate themselves" (p. 190). The influence of socioeconomic factors is thought to pervade all aspects of social life. According to Glascok (2014), "the overall economic status of a neighborhood might have direct effects through the quality of public schools, informal job networks, the monitoring of teenage behavior, and offering of positive role models" (p. 94).

A number of theoretical perspectives support the notion that the surrounding socioeconomic climate influences social and political behavior. *Social disorganization* *theory* (Shaw & McKay, 1942), for instance, holds that socio-economic conditions play an instrumental role in a given community's ability to come together around common goals and, in so doing, cooperatively address chronic problems afflicting the community (Aiyar, Zimmerman, Morrel-Samuels, & Reischl, 2014). Specifically, community factors such as high poverty and unemployment rates limit a community's ability to control and organize behavior, subsequently resulting in diminished levels of social trust. Relatedly, classic theoretical approaches centered on *political socialization* (e.g., Hyman, 1959; Settle, Bond, & Levitt, 2011) hold that more affluent communities allow parents to devote more resources to child-rearing, particularly as it relates to shaping worldview constructions, political knowledge sets, and political/civic behaviors (Jennings & Niemi, 1974; McIntosh, Hart, & Youniss, 2007; Meirick & Wackerman, 2005). Comparatively affluent areas also are more likely to possess primary and secondary school systems that have more experienced teachers, offer increased opportunities for extracurricular participation, and feature curricular approaches that include civics education (e.g., Wener, 1991; McFarland & Thomas, 2006; Settle, Bond, & Levitt, 2011). Still other models, such as those describing chronic stress imposed by poverty (e.g., Baum, Garofalo, & Yali, 1999; Hay, 1988; Steptoe & Feldman, 2001) hold that low SES status negatively affects physical and social wellbeing, which in in turn is associated with feelings of powerlessness and poorer coping skills. Finally, as pointed out Brady, Verba, & Schlozman (1995), a lack of available resources may inhibit the socioeconomically disadvantaged from perceiving themselves as stakeholders in the surrounding democratic society and thus be comparatively more inclined to engage in incivil behavior.

Relative to communication behaviors, previous research has found a direct link between contextual socioeconomic status and incivil communication practices. In a study of 400 undergraduate students, Glascock (2014) found that a composite variable describing (perceived) neighborhood quality was significantly and negatively linked to the propensity to use verbally aggressive/abusive language. In Brady, Verba, & Scholzman's (1995) explication of their *political resource model*, the authors found that resource paucity was associated with a relative lack of the communication and organizational necessary to facilitate effective civic participation. Relatedly, Dahlgren's notion of *civic culture* (2000; 2003; 2005) suggests that the facets of the socio-cultural world that constitute pre-conditions for democratic participation influence all forms of civic and political practice, including those practices related to citizen engagement in the online public sphere. Specifically, this suggests that online behavior is inextricably linked to citizens' lived experiences and personal resources (Dahlgren, 2005).

If it is indeed the case that citizens who exist in low SES contexts feel comparatively excluded from the social and political mainstream, it logically follows that they would be more likely to express sentiment that deviates from what is broadly considered civil. Interpreted relative to the objectives of the current study, the above literature suggests that citizens in low SES districts may be less likely to meaningfully participate in online discussions related to civic/political life and, thus, more likely to engage in less-than-civil discussion online. As such, the following hypothesis was posited:

**H5: Socioeconomic status will be negatively associated with political incivility on Twitter.**

Notably, SES is reflected in an array of measures representing various facets of social life. While there is not universal agreement on the most precise indicators of socio-economic status, researchers have typically used measures such as annual household income, education level, and occupational status (Conger & Donnellan, 2007; Glascock, 2014). Following this research, we employed four primary measures of each district's relative socioeconomic status. First, reasoning that low SES districts would be associated with diminished opportunities for political involvement, lower levels of general political involvement, and heightened feelings of disenfranchisement, we suggested that mean household income would be negatively related to political incivility on Twitter. Second, we suggested that districts with high unemployment rates would be associated with increased levels of political incivility as residents may feel increasingly distressed by their surrounding economic conditions. Third, we predicted that the education level of the district would be negatively associated with incivility. This assumption was rooted in the perception that college-educated adults are more likely to possess knowledge relevant to social and political process (and their importance), and thus be more likely to meaningfully and productively engage in civic activities. Fourth, building on emerging body of literature that suggests that access to healthcare is a reliable indicator of socioeconomic status (e.g., Baum, Garofalo, & Yali, 1999; Heck & Parker, 2002), we predicted that the percentage of respondents with health insurance would be negatively related to political incivility on Twitter. These hypotheses are formally explicated below.

**$H5_a$: District-wide mean annual household income will be negatively related to political incivility on Twitter.**

**H5b: District-wide unemployment rates will be positively associated with political incivility on Twitter.**

**H5c: The percent of citizens in each district with a bachelor's degree or higher will be negatively associated with political incivility on Twitter.**

**H5d: The percent of citizens in each district with health insurance will be negatively associated with political incivility on Twitter.**

## Method

### *Retrieving the Data from Twitter*

Version 1.0 of Twitter's API was called to download relevant Tweets during the election period. Specifically, the Streaming API call was used to download public messages from Twitter that mentioned the terms "Obama" or "Romney." The data was collected at a large scale, with a total of 70 million Tweets being collected in all, which suggests that the sample size was representative (Morstatter et. al, 2013). The collection started on August 1st, 2012 and ended on the Election Day, November 6th. In all, 465,582 Tweets were collected with GPS coordinates. Here users disclosed their location at the time the Tweet was sent.

Version 3 of the Sunlight Foundation's Congress API was used to resolve GPS coordinates (latitude and longitude) to correct congressional district in which the message came from (Sunlight Congress API, 2014). In total, 414,322 messages from 143,404 Twitter users were successfully resolved to the congressional district in which that message was sent.

### *The Use of Geotagged Messages from Twitter*

A limitation of this paper is its use of geo-tagged data from Twitter. At the time of data collection, for a Tweet to be included in our dataset, a Twitter user must have enabled geo-tagging of messages. Twitter prompts users to make a decision on this, typically once. It can then be optionally toggled at anytime on a per tweet basis. Researchers have shown that while the percentage of users that opt-in to geo-tagging continues to rise; it was approximately only 1% of all Twitter messages at the time of collection (Mahmud, Nichols & Drews, 2014).

In our dataset of the general election candidates, we measured a concentrated user group. As such, we measured a sub-population of Twitter, those who wanted their location to be broadcasted alongside their message. Other studies have survived this same limitation and predicted larger general phenomena successfully, such as infectious disease transmission and breaking news stories (e.g., Sadilek, Kautz & Silenzio, 2012; Meyer et al., 2011).

Still, questions exist as to how representative this subsample is. Luckily, we did not just collect geotagged tweets from the API at the time. As such, we can compare the entire collection (77 million) to the subsample. While those opting out of geotagging might be slightly more civil ($M$ = .19 versus $M$ = .24) such a difference is very small (0.05). A tweet with one incivil word would equal 1. For the incivility variable, the full-dataset seems to be distributed show greater skewness and kurtosis ($S$ = 8.98, $K$ = 142.22) to the small dataset ($S$ = 6.26, $K$ = 80.01). However, a visual inspection shows that the likely cause of these differences is outliers far away from the mean with very high incivility scores. For visual view of how similar the distributions appear to be around the mean see Table 2.

One possible reason for a difference and the greater presence of outliers in incivility in the full dataset is Twitter bots (or flame bots). A study of the 2011 Russian election concluded 25,860 bots, “swarmed the hashtags that legitimate users were using to communicate in an attempt to control the conversation and stifle search results related to the election" (Grier & Paxson, 2012). Researchers have noted key advantages of using geotagged data as it pertained to reaching representative samples of public opinion. First, the vast majority of "spam bots" that plague Twitter do not "opt-in" to geotagging of their Tweets (Guo & Chen, 2014; Grier & Paxson, 2012). Instead they only provide a minimum amount of metadata alongside their tweet. If a bot were to do so and accurately report it, they might reveal their location in an emerging country, where spam bots are prevalent (Grier & Paxson, 2012). The vast majority of spam on Twitter comes from other countries, due to economical reasons (Rao & Reiley, 2012). No clear reasons exist as to why a bot would "spoof" a specific location inside of the US, as it requires extra work (i.e. setting metadata that is optional using command line machine code) to do so, and is not required to send messages (Grier & Paxson, 2012). While bots do use proxys and VPNs to access Twitters from IP addresses inside of the U.S., they do not set the metadata to match the locations in which the intermediary servers are located. Locking a spam bot into set locations (e.g. where a VPN server was) may make them easier to detect.

The differences in the geotagged dataset could merely be the difference of Twitter spam bots, media and other organizations (all of which don't use GPS) that were removed from the GPS dataset. For instance the Twitter spambot @preciousliberty tweeted 21,179 times often with positive incivility scores (the bot appeared to have an anti-Obama agenda). Just as a representative polling company makes efforts to verify participants identity, using GPS, we are able to increase the degree to which we were "polling" real

Twitter users. If representative public opinion is what we are after, we may be better suited with verifiable locations.

A look at Table 1 shows the top 25 most active user accounts for each dataset. On the left, the average number of tweets per day range from 181 to 94 and on the right the range is 10 to 3. The left is behavior much more indicative of what others have found to spam bot behavior in general elections (Grier & Paxson, 2012), and the left is much more feasible for a human to actually broadcast. In an analysis of these 50 accounts, 26% percent of users from the non-geotagged dataset appeared to be spam while only 4% of the geotagged dataset appeared as such. This preliminarily evidence actually suggests that the GPS dataset may contain less noise per capita and be more representative of actual people than the raw dataset.

We suspected that the geotagging by area would differ, as the population does from district to district. To test this, we took a closer look at the distribution of tweets and the population distribution from the 2012 census.

The standard deviation of the distribution of the number of geotagged tweets by district was relatively half of the mean ($M$ = 329.67, $SD$ = 158.14) the same was true for the actual population of the districts ($M$ = 725,247.00, $SD$ = 366, 205.96). The skewness of the tweet sample ($S$ = 2.11) and the actual population ($S$ = .5) were both skewed to the left (.50). Kurtosis values were highly positive in the tweet sample ($K$ = 15.1) and the ($K$ = 8.11). Both samples suffer from similar skewness and kurtosis effects, however the tweet sample does appear to suffer to a greater degree.

As we've shown above, the distributions do not appear to be very different. Moreover, we truly believe that they should not be identical, or that we should be using data that has more spam and data from other countries (Guo & Chen, 2014). While it might not be the "biggest" data, here we made the conscious quality versus quantity choice. Overall, we feel our districts were robust enough to achieve a representative subsample. The smallest district in our sample still had 82 Twitter users (Wisconsin, District 7) and a total population of 158,667.

### ***Pivoting to User***

The data was then pivoted by username. For each user in the dataset an average location was derived. If a user sent messages from more than one district, the district in which the majority of messages originated from was treated as that user's home district. For each user, an average incivility score was calculated. This was done to limit any one user in a district from inflating district level results. Finally, all users for each district were then averaged, and average incivility scores were created for each district.

### ***Census Data***

Data from the 2013 American Community Survey (1-Year estimates) was used to obtain SES data points on each congressional district (United States Census Bureau, 2014). This data was collected in 2012, during the time of the election period.

### ***Measures***

This study aggregated and analyzed all data at the district level. The below paragraphs provide an overview of the specific procedures used to construct/define all of the variables used in the present study.

*Incivility*. Incivility was the criterion measure in the current study. Given the size of the data, the corpus could not be manually annotated. The researchers employed computer-assisted content analysis to derive incivility scores using Python. A script was written to process the wordlists and to develop incivility scores for the unit of analysis. It processed the list and detected the presence of words in the lists in the Tweets. It did so using regular expressions, windowed and exact matching. For some words that were ambiguous in nature, spaced word windows were used to make sure false positives were not counted. We followed Zamith and Lewis's (2015) suggestions for computational-coders. They reassert that computational methods work best when variables are "readily identifiable" in texts. They also echo Conway (2006) in stressing the importance of exhaustive word dictionaries and that contain all possible variants of words and their relevant words for the phenomena being studied (e.g. incivility). Further, they start to etch a workflow that is reflexive of emergent mass communication research that utilizes computer-assisted content analysis (Citation Withheld for review, 2014; Citation Withheld for review, 2015).

Not unlike inter-coder agreement with manual content analysis, "algorithms and dictionaries must often be repeatedly revised and tweaked to improve their performance" (Zamith & Lewis; pg. 4). The iterative process only concludes when the analysis yields a satisfactory level of construct validity. This is assessed when the researcher evaluates the algorithmic coder against the same coding decisions humans. In this case, two coders both must agree with each other to establish "gold standard" data. Then, the computer and the humans two agree at an acceptable level. Here, given the blatant nature of incivility in Twitter messages, human-to-human agreement of the two coders produced no errors.

Finally, the researchers draw attention to the point that once a computer has been verified to be valid, it will be also be reliable, as computers are persistent and consistent and not prone to human error. Thus, calculating reliability is not necessary (Riffe, Fico and Lacy, 2014; Zamith & Lewis, 2015).

As mentioned, incivility can have varying definitions (Coe, Kenski, & Rains, 2014; Herbst, 2010). For instance, Papacharissi (2004) differentiated between *conversational impoliteness* and *collective impoliteness,* noting that the former is frequently a spontaneous reflection of emotions while the latter represents conscious disregard for democratic consequences and is therefore substantively more damaging to civil society. Coe, Kenski, and Rains (2014), argued that incivility, as a conversational artifact, is structured around the notions of *disrespect* and *irrelevance*. Operating under the presumption that incivility is directed toward perceived oppositional forces, the authors proposed that incivility manifests in one or more of the following forms: name-calling, aspersion, lying, vulgarity, and pejorative speech. Similarly, Santana (2012; 2014), defined incivility as having nine key components: (1) name calling; (2) threats; (3) vulgarities; (4) abusive or foul language; (5) xenophobia; (6) hateful language, epithets or slurs; (7) racist or bigoted sentiments; (8) disparaging comments on the basis of race/ethnicity; and (9) use of stereotypes.

In the current study, we adopted Santana's definition (with some modifications) of incivility as it broadly encompassed the conceptual approaches used by Papacharissi (2004) and Coe, Kenski, and Rains (2014) while also including constructs potentially relevant to the 2012 presidential election. An initial, manual scan of 800 Tweets revealed that 8 of the 9 concepts were present in Tweets. Xenophobia was not present and did not appear to be relevant to the election. The researchers then leveraged two wordlists that are regularly used to detect concepts (2), (3), (4), (6), (7) and (8) on short informal text on the web. Google's "bad word list" and ClueBot's "insult list" have been used by dozens of machine learning data scientists seeking to automatically detect insults in social commentary (Mueller, 2012). "Strong words" have been shown by others to be the most effective search terms. Taking from the literature you provided us, "Many of the comment threads displayed levels of incivility that were at times shocking and generally disappointing… Terms such as 'idiots' and 'morons' were commonly used" (pg. 262, Coffey Kohler and Granger, 2015).

No clear wordlists existed for concepts (1) and (5). As such they were derived during the manual content analysis.

Starting with over 600 words that been linked to concepts encompassed by incivility, the researchers proceeded with four rounds of manual content analysis. 800 random tweets were pulled from the dataset. Each tweet was scored for incivility using the initial wordlists. A python script took stemmed words and looked for a combination of windowed and unwindowed matches. The score was calculated by adding the number of words that appeared in that Tweet. In addition, when the matching word was in uppercase, or when the message contained an exclamation point, Tweets were boosted by an additional point, as most affective boosting tools similarly do (Thelwall, 2010). The researchers read each message and (1) verified that the words flagged were indeed incivil (2) scanned the message to see if any additional words in the Tweet were incivil, but undetected and (3) verified that the overall score was correct. In cases where (1) was incorrect, the researchers altered the keywords (i.e., adding a word window around "hell" so "hello" was not detected). When (2) was incorrect, the researchers added the appropriate keyword. In all cases (3) was satisfied. This process was performed four times; each time the keyword list was adjusted. “Bad words” are not the only type of incivility that can be found in text. In fact, out of the box, the wordlists were not sufficient. To make sure that our word lexicons were appropriate to the domain, we manually trained the wordlists per the guidelines in Zamith and Lewis (2015). In testing for false positives and false negatives, we made sure the detection was complete.

After each iteration the incivility measure improved across (1) and (2). The final percent agreement with the manual annotations was 98.5% for (1) and 98% for (2). In all, the researchers were left with a list of 650 keywords that addressed all 8 concepts that are encompassed by incivility. While this wordlist is powerful on the corpus at hand, its validity is likely limited to the context of the 2012 election, specifically short informal text that mentioned either Obama or Romney.

As a result of these procedures, a single index was created. Each user was assigned an average incivility score. These scores ranged from 0 (completely civil) to 23.00 (highly incivil). Using the individual scores assigned to each user, we then calculated a mean incivility score for each congressional district. Figure 1 provides a visual representation of the average incivility scores associated with each U.S. congressional district.

INSERT FIGURE 1 ABOUT HERE

*Social capital potential indicators.* Three factors relevant to each district’s potential for social capital were employed: racial heterogeneity, residential tenure, and geographic mobility. Residential tenure was conceptualized as the percent of the population with more than 12 years of tenure in their current residence while geographic mobility was operationalized as the percentage of the population in each district that was born in-state. Racial heterogeneity was calculated at the district level using the index previously employed by Costa and Kahn (2003). The measure's scale ranged from 0 (complete homogeneity) and 1 (complete heterogeneity). The measure was calculated as

$$Heterogeneity = 1 - \sum_{k} s_{ki}^{2}$$

where *k* represents the number of racial categories recorded in the Census dataset (white, black, American Indian, Asian, other, and more than one race) and $s_{ki}$ represents the share of each racial category in voting district *i* .

The structure of the Census data framed the operationalization of some of these variables. For instance, the response categories associated with residence tenure were "Moved in 2010 or later," "Moved in 2000 to 2009," "Moved in 1990 to 1999" and so on. Judging that the formation of meaningful social capital with one's neighbors would be an involved process that could, conceivably speaking, take more than four years, we used a somewhat conservative approach to formation of the residence tenure variable. Likewise, it may have been desirable to measure geographic mobility specifically on the district level; however, the nature of the available Census data forced us to operationalize the variable relative to in-state births. This method is justifiable as those born in their current district of residence are also, obviously, born in their current state of residence.

*District polarity*. District polarity was assessed using a modified version of Cook's Partisan Voting Index (PVI; Wasserman, 2013). The PVI is based upon the district's voting behavior in the previous two presidential elections (2008 and 2012). Although the Twitter data was harvested before the 2012 election, we nonetheless used the results of the 2012 election as part of this measure because, we reasoned, it provided an accurate depiction of the polarization climate within the district at the time of data collection. Without modification, the PVI describes the direction (i.e., R or D) and magnitude of partisan polarity in each district relative to the nation as a whole. The PVI was calculated as

$$PVI \ = \left(\frac{pA + pB}{2}\right) - \left(\frac{PA + PB}{2}\right)$$

where $pA$ represents the percentage of the two-way presidential vote that Obama received in the district in 2008, $pB$ represents the percentage of the two-way vote that Obama received in the district in 2012, $PA$ represents the percentage of the two-way national vote that Obama received in 2008, and $PB$ represents the percentage of the two-way national vote that Obama received in 2008. A large positive value suggested that the district was polarized in favor of the Democratic Party while a large negative value suggested polarization in favor of the Republican Party. Values near 0 were indicative of a relative lack of partisan polarity. As the current study was interested in assessing partisanship magnitude *irrespective* of party affiliation, we took the absolute value of each district's PVI value. The measure thus returned a value that described district-level polarity irrespective of the party towards which such polarity trended toward.

*Socioeconomic status indicators.* SES factors of interest in the current study included mean annual household income in each congressional district, the percent of individuals in each district with a bachelor's degree or higher, district unemployment rate (represented as a percentage of the total workforce), and percent with of the district with health insurance.

*Control variables*. All hypotheses tests controlled for a number of potentially confounding factors. First, as there are substantial differences in district population sizes across the US, we controlled for number of inhabitants in each district. The median age of each district was controlled for, as Twitter use is heavier among younger users. Similarly, the percent of urban inhabitants was similarly controlled for, as Twitter users tend to reside in urban areas (Duggan, Ellison, Lampe, Lenhart, & Maden, 2015). Finally, the effects of number of users and average Tweets per user in each district were accounted for as a means of ensuring that the incivility measure was not upwardly biased by districts that housed a comparatively small number of infrequent but highly incivil Twitter users. [2]

Descriptive statistics for each measure are provided in Table 1. Zero-order correlations for all of the variables included in the hypotheses tests are shown in Table 3.

INSERT TABLE 3 ABOUT HERE

INSERT TABLE 4 ABOUT HERE

## ***Results***

Hypotheses H1 – H5 were tested using hierarchical OLS regression. Classes of independent variables (i.e., social capital potential indicators, social economic status indicators, and district-wide partisan polarity) were examined both individually and in conjunction with one another.

As seen in Table 5, ethnic heterogeneity was positively and significantly related to incivility on Twitter when examined both individually ($\beta$ = .43, $p$ < .001) and after controlling for the effects of all other independent variables ($\beta$ = .35, $p$ < .001); therefore, H1 was supported. H2 suggested that residential tenure would be negatively related to incivility on Twitter. In the model consisting only of the control variables and social capital indicators, residential tenure measure was, indeed, negatively and significantly related to incivility on Twitter, $\beta$ = -.18, $p$ < .01. However, when included in the all-entry model, residential tenure was no longer significantly related to incivility on Twitter. As a result, H2 received only partial support. H3 suggested there would be a similarly negative relationship between geographical mobility and incivility on Twitter. Contrary to expectations, there was a strong, *positive* relationship between geographic mobility and Twitter incivility. This was true in both the initial model consisting of only the social capital potential indicators and in the all-entry model ($\beta$ =.54, $p$ < .001, $\beta$ =.44, $p$ < .001, respectively). H3, accordingly, was not supported.

H4 suggested that districts with lower levels of partisan polarity would yield higher levels of uncivil Twitter activity. As seen in the polarity-only model, the relationship between partisanship and incivility was not significant at the $p$ < .05 level. However, upon controlling for the influence of the social capital potential variables and SES indicators, there was a negative and significant relationship between district partisanship and political incivility on Twitter, $\beta$ =-.17, $p$ < .001. This result was generally supportive of H4.

Finally, $H5_{a-d}$ were broadly concerned with the relationship between district SES status and Twitter incivility. The data indicated that the mean household income was a positive predictor of incivility on Twitter ($\beta$ =.25, $p$ < .01) in the model including only the SES indicators; however, this relationship disappeared in the all-entry model. Given that we predicted a negative relationship between annual household income and incivility on Twitter, no support for $H5_a$ was observed. Unemployment rate was positively and significantly related to incivility in the SES-only model ($\beta$ =.20, $p$ < .001). However, this relationship disappeared in the all-variable model. Thus, only partial support was found for $H5_b$. Percent of users in each district with a bachelor's degree or higher was a negative and significant predictor of incivility on Twitter in both the SES-only and all-variable models ($\beta$ =-.62, $p$ < .001 and $\beta$ =-.37, $p$ < .001, respectively), thus support was observed for $H5_c$. Finally, the percent of the district with healthcare insurance was not related to Twitter incivility in either the SES indicator model or in the all-entry models. Accordingly, $H5_d$ was not supported. Table 5 provides a full report of the hypotheses tests. [3]

INSERT TABLE 5 ABOUT HERE

Notably, the relationship between geographic mobility and incivility was both strong and in the opposite direction of what was hypothesized (H3). To better explore this result, we investigated the potentiality that the relationship between geographic mobility and incivility on Twitter was moderated by SES status. Prior research has broadly indicated that a complex relationship between SES status and mobility exists as the factors motivating population in and outflows are markedly dissimilar in high SES regions when compared to low SS regions (e.g., Bowles, 1970). Given the exploratory nature of this investigation, we generated a composite index of SES status using exploratory factor analysis (EFA). Included in this analysis were the mean household income, unemployment rate, education level, and healthcare coverage indicators. Factor extraction was accomplished using principal axis factoring. All factors with eigenvalues ≥ 1.00 were extracted. The resulting solution was unidimensional in nature and explained 56.10% of the total variance (extracted $h^2$ values ranged from .29 to .86). The factor itself was generated using regression computation (see DiStefano, Zhu, & Mindrila, 2009 for review of available approaches for factor computation). The resultant measure of district-wide

SES status had standardized values ranging from -2.03 to 3.69 (M = 0.00, SD = 0.95; Skew = 0.80, Kurtosis = 0.82) with higher values being indicative of higher SES status.

This SES factor was subsequently used in an exploratory moderation analysis in which we examined the influence of geographic mobility at various SES values. As seen in Table 6, SES status indeed exerted a moderating effect on the relationship between geographic mobility and incivil behavior on Twitter. The nature of this moderation effect was probed using the Johnson-Neyman technique (Johnson & Fay, 1950), which assesses regions of significance across the range of the moderator.[4] As shown in Table 7, there was a strong positive relationship between percentage of the district born in-state and Twitter incivility in very low and low SES districts. However, in moderate and high SES districts, this relationship disappeared.

INSERT TABLE 6 ABOUT HERE

INSERT TABLE 7 ABOUT HERE

## Discussion

The present study drew upon a corpus of nearly 70 million Tweets posted around the time of the 2012 presidential election to explore the relationship between physical context and online behavior. Specifically, using a subsample consisting of 414, 322 Tweets drawn from 143,404 individual Twitter users, we investigated the relationship between district-level social capital potential, partisan polarity, and SES characteristics and the average amount of Twitter incivility that emanated from each U.S. congressional district. With some exceptions, the results suggested that low levels of social capital potential, low levels of polarity, and low SES status were associated with heightened levels of incivility.

There are a number of implications that stem from the current results. First, initial evidence here suggests that the tenor of online political discussion is broadly reflective of the broader social, cultural, and partisan contexts within which users physically reside (e.g., Dahlgren, 2005). Specifically, our analyses suggested factors that have been previously associated with offline incivility (i.e., racial heterogeneity and low levels of college education among the population), can similarly be correlated with online incivility. Understood in its broadest sense, our findings seem to offer support for theoretical perspectives that connect online behavior with lived experience. Moreover, in regards to the online potential of social/new media, these findings mirror others (e.g., Albrecht, 2006; Baek, Wojcieszak, & Delli Capini, 2011) in their suggestion that online deliberation cannot, alone, overcome the systematic inequalities that have traditionally structured civic involvement. If we are truly this reflexive of our lived experience online, what other theories may predict our discourse? Certainly social capitol is not the only theory tied to socioeconomic factors of our surroundings. Researchers should use these geographically contextual variables as possible explanatory variables for other analysis of discourse.

Furthermore, the present results pose questions in regards to the overall value of *civil* political discourse. Such questions are not novel to this work, and have been raised by a number of scholars in regards to both online and offline discourse (e.g., Dahlberg, 2007; Herbst, 2010; Papacharissi, 2004; Price, Nir, & Capella, 2006). As Schudson (1997) argued, “democratic talk is not necessarily egalitarian but it is essentially *public*, and if this means that democratic talk is talk among people of different backgrounds and values, it is also profoundly *uncomfortable*” (p. 299, emphasis from the author).

Nonetheless, sentiments such as “conversation is the soul of democracy” (Kim, Wyatt, & Katz, 1999, p. 362) and “civil discussion lies at the heart of democracy” (Rowe, 2015, p. 121) color, if not define, our current collective understanding of the necessary ingredients for a well-functioning democracy. However, in the current case, our results suggested that incivil discourse was highest in districts that were characterized, in part, by factors traditionally thought to be indicative of a healthy and diverse democracy (i.e., low levels of partisan polarity and high levels of racial diversity).

Clearly, these findings do not suggest that society is forced to choose between *either* a civil democracy *or* a well functioning one. High levels of unconstrained and unfocused incivility are unlikely to yield the type of social cohesion and ideological compromise necessary for the continued and virile existence of the democratic state. That said, the present results suggest that there may be something of a misalignment between the idealized notion of democracy and its functional enactment. Future studies may investigate incivility and its effect on predicting the current political environment. Do governments perform better when digital and social media incivility is lower? Are these findings generalizable to other countries’ social media discourse, or do other countries behave differently in relation to their socio-economic environments? Does the role of government type or style act as a mediator?

Notably, we failed to either fully or partially support a number of our hypotheses. In some cases, such as those hypotheses related to SES status, it could be that a single indicator (i.e., district education level) served as the optimal representative of the on-hand phenomenon and, therefore, explained a bulk of the variance. In other cases, such as the relationship between geographic mobility and incivil deliberation, it could be that variable inter-relationships are highly conditional in nature. For instance, as seen in Table 6 and Table 7, the influence geographic mobility on uncivil discussion waned at moderate and high SES status levels. Future research should continue the exploration of online incivility by systematically investigating factors that have the potential to exert moderating influences. Additional expansions could also investigate the relationship between different types of social capital (i.e., bonding or bridging) and the prevalence on incivility. Finally, as this study primarily explored incivil behavior on the (aggregated) district level, future research could employ multi-level modeling techniques (MLM; e.g., Park, Eveland, & Cuedeck, 2008; Raudenbush & Bryk, 2001) to better understand the relationship between contextual influences and individual behavior.

At least a few factors limit the current findings. First, the nature of the data should limit the scope into which the findings are generalizable. The source of data here, Twitter, may very well be a proxy of public opinion as some have shown (i.e., Vargo, 2011), but it also is invariably different from survey data. Instead, Twitter messages are, at best, an instantaneous measure of behavior, not a durable measure of emotion or feelings (Vieweg, 2010). Moreover, the corpus here was controlled to a specific event, the 2012 general election. The messages gathered in this analysis were also directed at a specific political candidate (e.g., Obama and/or Romney). This prohibits making generalizations about the levels of incivility regarding a specific region outside of the context of the political candidates, during the election. While the findings still yield important conclusions toward discourse, democracy and general elections, they do not necessarily scale to grandiose findings of our nation and incivility as a whole.

In spite of these limitations, this study shows that behaviors on social/digital media platforms can be modeled using a broad range of social science theories traditionally used to explain offline behaviors. As such, these findings could be extrapolated to predict how congressional districts might act in future elections. Using public data such as the census, for instance, campaign observers could reasonably expect to forecast and predict civil discussion using the methods discussed in this paper.

**Citations**

Aiyer, S. M., Zimmerman, M. A., Morrel-Samuels, S., & Reischl, T. M. (2014). From broken windows to busy streets: A community empowerment perspective. *Health Education & Behavior,* 1-11.

Albrecht, S. (2006). Whose voice is heard in online deliberation? A study of participation and representation in political debates on the Internet. *Information, Communication & Society, 9,* 62 – 82.

Alesina, A., & Ferrara, E. L. (2000). Participation in heterogeneous communities. *Quarterly Journal of Economics, 115,* 847 – 904.

Anderson, A. A., Brossard, D., Scheufele, D. A., Xenos, M. A., & Ladwig, P. (2014). The 'nasty effect:' Online incivility and risk perceptions of emerging technologies. *Journal of Computer-Mediated Communication, 19,* 373 – 387.

Baek, Y. M., Wojcieszak, M., Delli Capini, M. X. (2011). Online versus face-to-face deliberation: Who? why? what? With what effects? *New Media & Society, 14,* 363 – 383.

Baum, A., Garofalo, J. P., & Yali, A. M. (1999). Socioeconomic status and chronic stress: does stress account for SES effects on health? *Annals of the New York Academy of Sciences, 896,* 131 – 144.

Bimber, B., Cunill, M. C., Copeland, L., & Gibson, R. (2015). Digital Media and Political Participation The Moderating Role of Political Interest Across Acts and Over Time. *Social Science Computer Review*, *33*(1), 21-42.

Blanchard, A., & Horan, T. (1998). Virtual communities and social capital. *Social Science Computer Review*, 16, 293 – 307.

Boardman, J. D., and S. A. Robert. 2000. Neighborhood socioeconomic status and perceptions of self-efficacy. Sociological Perspectives, 43, 117- 136.

Borah, P. (2014). Does it matter where you read the news story? Interaction of incivility and news frames in the political blogosphere. *Communication Research, 41,* 809 – 827.

Bowles, S. (1970). Migration as investment: Empirical tests of the human investment approach to geographical mobility. *The Review of Economics and Statistics, 52,* 356 – 362.

Boyd, R. (2006). The value of civility? *Urban Studies, 43,* 863 – 878.

Brady, H. E., Verba, S., & Schlozman, K. L. (1995). Beyond SES: A resource model of political participation. *American Political Science Review, 89,* 271 – 294.

Brady, D. W., & Han, H. C. (2006). Polarization then and now: A historical perspective. In P. S. Nivola & D. W. Brady (eds.), *Red and blue nation, Vol. 1* (pp. 119 – 174). Baltimore, MD: The Brookings Institution.

Byrne, D. E. (1971). *The attraction paradigm*. New York: Academic Press.

Campbell, C, Rockman, B. A., & Rudalevige, A. (2008). *The George Bush legacy.* Washington, D.C.: CQ Press.

Chamlin, M. B., & Cochran, J. K. (1995). Assessing Messner and Rosenfeld's institutional anomie theory: A partial test. *Criminology, 33,* 411 – 429.

Cho, J. (2013). Campaign tone, political affect, and communicative engagement. *Journal of Communication, 63,* 1130 – 1152.

Coe, K., Kenski, K., & Rains, S. A. (2014). Online and uncivil? Patterns and determinants of incivility in newspaper website comments. *Journal of Communication, 64,* 658 – 679.

Coffey, D. J., Kohler, M., & Granger, D. M. (2015). Sparking Debate Campaigns, Social Media, and Political Incivility. *Controlling the Message: New Media in American Political Campaigns*, 245.

Coleman, J. S. (1988). Social capital in the creation of human capital. *American Journal of Sociology, 94,* 95 – 120.

Coleman, J. (1990). *Foundations of social theory*. Cambridge, Mass.: Harvard University Press.

Conger, R. D., & Donnellan, M. B. (2007). An interactionist perspective on the socioeconomic context of human development. *Annual Review of Psychology, 58,* 175–199.

Costa, D. L., & Kahn, M. E. (2003). Civic engagement and community heterogeneity: An economist's perspective. *Perspectives on Politics, 1,* 103–111.

Dahlberg, L. (2001). The Internet and democratic discourse: Exploring the prospects of online deliberative forums extending the public sphere. *Information, Communication & Society, 4,* 615–633.

Dahlberg, L. (2007). Rethinking the fragmentation of the cyberpublic: From consensus to contestation. *New Media & Society, 9,* 827 – 847.

Dahlgren, P. (2000). The Internet and the democratization of civic culture. *Political Communication, 17,* 335 – 340.

Dahlgren, P. (2003). Reconfiguring civic culture in the new media milieu. In J. Corner & D. Pels (eds.), *Media and political style: Essays on representation and civic culture* (pp. 151 – 170). London: Sage.

Dahlgren, P. (2005). The Internet, public spheres, and political communication: Dispersion and deliberation. *Political Communication, 22,* 147 – 162.

David, Q., Janiak, A., & Wasmer, E. (2010). Local social capital and geographical mobility. *Journal of Urban Economics, 68,* 191 – 204.

Davis, R. (2009). *Typing politics: The role of blogs in American politics.* New York: Oxford University Press.

DiMaggio, P., Evans, J., & Bryson, B. (1996). Have Americans' social attitudes become more polarized? *American Journal of Sociology*, *102*, 690 – 755.

Dimock, M., Doherty, C., Kiley, J., & Oates, R. (2014, June 12). *Political polarization in the American public: How increasing ideological uniformity and partisan antipathy affect politics, compromise and everyday life.* Retrieved from Pew Research center: http://www.people-press.org/files/2014/06/6-12-2014-Political-Polarization-Release.pdf

DiStefano, C., Zhu, M., & Mindrila, D. (2009). Understanding and using factor scores: Considerations for the applied researcher. *Practical Assessment, Research, & Evaluation*, *14*, 1 – 11.

Duggan, M., Ellison, N. B., Lampe, C., Lenhart, A., & Madden, M. (2015, Jan 9). *Demographics of key social networking platforms.* Retrieved from Pew research Internet project: http://www.pewinternet.org/2015/01/09/demographics-of-key-social-networking-platforms-2/

Freelon, D. (2013). Discourse architecture, ideology and democratic norms in online political discussion. *New Media & Society.* Advance online publication. doi: 10.1177/1461444813513259

Guo, D., & Chen, C. (2014). Detecting Non-personal and Spam Users on Geo-tagged Twitter Network. *Transactions in GIS*, *18*(3), 370-384.

Ellison, N. B., Steinfeld, C., & Lampe, C. (2006, June). Spatially bounded online social networks and social capital: The role of Facebook. Paper presented at the Annual Meeting of the International Communication Association. Dresden, Germany.

Ellison, N. B., Steinfeld, C., & Lampe, C. (2007). The benefits of Facebook “friends:” Social capital and college students’ use of online social network sites. J*ournal of Computer-Mediated Communication, 12,* 1143 – 1168.

Farnam, T. W. (2012, Feb 12). Study: Negative campaign ads much more frequent, vicious than in primaries past. T*he Washington Post.* Retrieved from http://www.washingtonpost.com/politics/study-negative-campaign-ads-much-more-frequent-vicious-than-in-primaries-past/2012/02/14/gIQAR7ifPR_story.html

Fiorina, M. P., & Abrams, S. J. (2008). Political polarization in the American public. *Annual Review of Political Science, 11,* 563 – 588.

Forni, P. M. (2011, Jan 28). Why civility is necessary for society’s survival. *The Dallas Morning News*, Retrieved from http://www.dallasnews.com/opinion/sunday-commentary/20100723-p.m.-forni-why-civility-is-necessary-for-society_s-survival.ece

Gil de Zúñiga, H., Jung, N., & Valenzuela, S. (2012). Social media use for news and individuals’ social capital, civic engagement and political participation. *Journal of Computer-Mediated Communication, 17,* 319 – 336.

Gittell, R., & Vidal, A. (1998). *Community organizing: Building social capital as a development strategy.* Thousand Oaks, CA: Sage.

Gladwell, M. (2011). From innovation to revolution. Do social media make protests possible? *Foreign Affairs, 90*, 153.

Glascock, J. (2014). Contribution of demographics, sociological factors, and media usage to verbal aggressiveness. *Journal of Media Psychology, 26*, 92 – 102.

Goldstein, K., & Freedman, P. (2002). Campaign advertising and voter turnout: New evidence for a stimulation effect. *The Journal of Politics, 64*, 721 – 740.

Grand, S. R. (2014). *Understanding Tahir Square: What transitions can teach us about the prospects for Arab democracy.* Washington D.C.: The Brookings Institution.

Granovetter, M. S. (1973). The strength of weak ties. *American Journal of Sociology, 78,* 1360 – 1380.

Gurin, P., Nagda, B. A., & lopez, G. E. (2004). The benefits of diversity in education for democratic citizenship. *Journal of Social Issues, 60*, 17 – 34.

Hagan, J., MacMillan, R., & Wheaton, B. (1996). New kid in town: Social capital and the life course effects of family migration on children. *American Sociological Review, 61*, 368 – 385.

Hay, D. I. (1988). Socioeconomic status and health status: A study of males in the Canada. *Social Science & Medicine, 27,* 1317- 1325.

Hayes, A. F., 2013. *An introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. New York: Guilford Press.

Heck, K. E., & Parker, J. D. (2002). Family structure, socioeconomic status, and access to healthcare for children. He*alth Services Research, 37,* 171 – 184.

Hendryx, M. S., Ahern, M. M., Lovrich, N. P., & McCurdy, A. H. (2002). Access to health care and community social capital. *Health Services Research, 37,* 85 – 101.

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

Hmielowski, J. D., Hutchens, M. J., Cicchirillo, V. J. (2014). Living in an age of online incivility: Examining the conditional indirect effects of online discussion on political flaming. *Information, Communication & Society*, 17, 1196 – 1211.

Huysman, M., & Wulf, V. (Eds.). (2004). *Social capital and information technology.* Cambridge, MA: MIT Press.

Hyman, H. (1959). *Political socialization*. Glencoe, IL: Free Press

Israel, G. D., Beaulieu, L. J., & Hartless, G. (2001). The influence of family and community social capital on educational achievement. *Rural Sociology, 66,* 43 – 68.

Johnson, P. O., & Fay, L. C. (1950). The Johnson-Neyman technique, it's theory and application. *Psychometrica, 15,* 349 -367.

Kan, K. (2007). Residential mobility and social capital. *Journal of Urban Economics, 61,* 436 – 457.

Kang, N., & Kwak, N. (2003). A multilevel approach to civic participation: Individual lengths of residence, neighborhood residential stability, and their interactive effects with media use. *Communication Research, 30, 80 – 106.*

Kavanaugh, A. L., Reese, D. D., Carroll, J. M., & Rosson, M. B. (2005). Weak ties in networked communities. *The Information Society, 21*, 119 – 131.

Krackhardt, D. (1992). The strength of strong ties: The importance of philos in organizations. In N. Nohria & R. Eccles (Eds.), *Networks and Organizations: Structure, Form and Action* (pp. 216–239). Boston: Harvard Business School Express.

Krosnick, J. A., & Petty, R. E. (1995). Attitude strength: An overview. In R. E. Petty and J. A. Krosnick (Eds.), *Attitude strength: Antecedents and consequences* (pp. 1-22). Hillsdale, NJ: Erlbaum.

Larsen, L., Harlan, S. L., Bolin, B., Hackett, E. J., Hope, D., Kirby, A., Nelson, A., Rex, R. T., & Wolf, S. (2004). Bonding and bridging: Understanding the relationship between social capital and civic action. *Journal of Planning Education and Research, 24,* 64 – 77.

Lau, R., Sigelman, L., Heldman, C., & Babbit, P. (1999). The effects of negative political advertisements: A meta-analytical assessment. *American Political Science Review, 93,* 851–976.

Lee, J. & Lee, H. (2010). The computer-mediated communication network: Exploring the linkage between online community and social capital. *New Media & Society, 12,* 71 – 727.

Leigh, A. (2006). Trust, inequality, and ethnic heterogeneity. *Economic Record, 82,* 268 – 280.

Liu, Q. A., Ryan, V., Aurbach, H., & Besser, T. (1998). The influence of local church participation on rural community attachment. *Rural Sociology, 63,* 432-450.

Mahmud, J., Nichols, J., & Drews, C. (2012, May). Where Is This Tweet From? Inferring Home Locations of Twitter Users. In *ICWSM*.

Marinov, N., & Schimmelfennig, F. (2015). Does Social Media Promote Civic Activism? Evidence from a Field Experiment, Center for Comparative and International Studies, ETH Zürich.

McFarland, D., & Thomas, R. (2006). Bowling young: How youth voluntary associations influence adult political participation. *American Sociological Review, 71,* 401-425.

McIntosh, H., Hart, D., & Youniss, J. (2007). The influence of family political discussion on youth civic development: Which parent qualities matter? *PS: Political Science and Politics, 40,* 495-499.

McPherson, M., & Smith-Lovin, L. (1987). Homophily in voluntary organizations: Status distance and the composition of face-to-face groups. American Sociological Review, 52, 370-389.

McPherson, M., Smith-Lovin, L., & Cook, J. M. (2001). Birds of a feather: Homophily in social networks. *Annual Review of Sociology, 27,* 415–444.

Meirick, P., & Wackman, D. (2004). Kids voting and political knowledge: Narrowing gaps, informing votes. *Social Science Quarterly, 85,* 1161-1177.

Meyer, B., Bryan, K., Santos, Y., & Kim, B. (2011). TwitterReporter: Breaking News Detection and Visualization through the Geo-Tagged Twitter Network. Paper presented at ISCA 26th International Conference on Computers and Their Applications (pp. 84-89).

Morstatter, F., Pfeffer, J., Liu, H., & Carley, K. M. (2013). Is the Sample Good Enough? Comparing Data from Twitter's Streaming API with Twitter's Firehose. Paper presented at ICWSM (pp. 400-408). Boston: AAAI.

Mueller, A. (2012, Sept 21). Detecting Insults in Social Commentary. Retrieved from https://www.kaggle.com/c/detecting-insults-in-social commentary/forums/t/2744/what-did-you-use/15951

Newton, K. (2003). Trust, social capital, civil society, and democracy. *International Political Science Review, 22,* 201 – 214.

O'Brien, R. M. (2007). A caution regarding rules of thumb for variance inflation factors. *Quality & Quantity, 41, 673* – 690.

Oliver, J. E. (1999). The effects of metropolitan segregation on local civic participation. *American Journal of Political Science, 43,* 186 – 212.

Onyx, J., & Bullen, P. (2000). Measuring social capital in five communities. *The Journal of Applied Behavioral Science, 36,* 23 – 42.

Oropesa, R. S. (1992). Social structure, social solidarity and involvement in neighborhood improvement associations. *Social Inquiry, 62*, 107-118.

Park, H. S., Eveland, W. P., & Cuedeck, R. (2008). Multilevel modeling: Studying people in contexts. In A. F. Hayes, M. D. Slater, & L. B. Snyder (Eds.*), Advanced data analysis methods for communication research* (pp. 219 – 246). Thousand Oaks, CA: Sage.

Papacharissi, Z. (2004). Democracy online: Civility, politeness, and the democratic potential of online political discussion groups. *New Media & Society, 6,* 259 – 283.

Park, H. S., Eveland, W. P., & Cuedeck, R. (2008). Multilevel modeling: Studying people in contexts. In A. F. Hayes, M. D. Slater, & L. B. Snyder (Eds.*), Advanced data analysis methods for communication research* (pp. 219 – 246). Thousand Oaks, CA: Sage.

Price, V., Nir, L., & Cappella, J. N. (2006). Normative and informational influences in online political discussions. *Communication Theory, 16,* 47 – 74.

Rao, J. M., & Reiley, D. H. (2012). The economics of spam. *The Journal of Economic Perspectives*, 87-110.

Raudenbush, S. W., & Bryk, A. S. (2001). *Hierarchical linear models: Applications and data analysis methods.* Thousand Oaks, CA: Sage.

Riffe, D., Lacy, S. & Fico, F. (2014). *Analyzing media messages : using quantitative content analysis in research*. New York: Routledge/Taylor & Francis Group.

Rosenfeld, R., Messner, S. F., & Baumer, E. P. (2001). Social capital and homicide. *Social Forces*, *80,* 283 – 309.

Ross, C. E., & Mirowsky, J. (2011). Neighborhood disadvantage, disorder and health. *Journal of Health and Social behavior, 42,* 258 – 276.

Rowe, I. (2015). Civility 2.0: A comparative analysis of incivility in online political discussion. *Information, Communication & Society, 18,* 121 – 138.

Rubenstein, L. (2014, Sept 16). 2014 General Election Advertising Opens Even More Negative than 2010 or 2012. Wesleyan Media Project. Retrieved from http://mediaproject.wesleyan.edu/releases/2014-general-election-advertising-opens-even-more-negative-than-2010-or-2012/

Rupasingha, A., Goetz, S. J., & freshwater, D. (2006). The production of social capital in US counties. *The Journal of Socio-Economics, 35,* 83 – 101.

Sadilek, A., Kautz, H. A., & Silenzio, V. (2012). Predicting Disease Transmission from Geo-Tagged Micro-Blog Data. Paper presented at the Twenty-Sixth AAAI Conference on Artificial Intelligence (pp. 136-142).

Santana, A. D. (2012). Civility, anonymity and the breakdown of a new public sphere (doctoral dissertation). University of Oregon, Eugene, Oregon.

Santana, A. D. (2014). Virtuous or vitriolic: The effect of anonymity on civility in online newspaper reader comment boards. *Journalism Practice, 8,* 18 – 33.

Schiff, M., 1992. Social capital, labor mobility, and welfare: the impact of Uniting States. *Rationality and Society*, *4,* 157–175.

Schudson, M. (1997). Why conversation is not the soul of democracy. *Critical Studies in Mass Communication 14,* 1–13.

Settle, J. E., Bond, R., & Levitt, J. (2011). The social origins of adult political behavior. *American Politics Research, 39,* 239 – 263.

Shah, D. V. (1998). Civic engagement, interpersonal trust, and television use: An individual level assessment of social capital. *Political Psychology, 19,* 469 – 496.

Shah, D. V., Kwak, N., & Holbert, L. (2001). “Connecting” and “disconnecting” with civic life: Patterns of Internet use and the production of social capital. *Political Communication, 18,* 141 – 162.

Shaw, C. R., & McKay, H. D. (1942). *Juvenile delinquency and urban areas.* Chicago, IL: University of Chicago Press.

Shirky, C. (2011). The political power of social media: Technology, the public sphere, and political change. *Foreign Affairs, 90,* 28 – 41.

Steptoe, A., & Feldman, P. J. (2001). Neighborhood problems as sources of chronic stress: Development of a measure of neighborhood problems, and associations with socioeconomic status and health. *Annals of Behavioral Medicine, 23,* 177 – 185.

Stromer-Galley, J., & Muhlberger, P. (2009). Agreement and disagreement in group deliberation and its consequences: Satisfaction, engagement, and opinion reevaluation. *Political Communication, 26,* 173–192.

Sunlight Congress API. (2014, Jan 12). Official API Documentation. Retrieved from https://sunlightlabs.github.io/congress

Sunstein, C. (2001). *Republic.com*. Princeton, NJ: Princeton University Press

Thelwall, M., Buckley, K., Paltoglou, G., Cai, D., & Kappas, A. (2010). Sentiment strength detection in short informal text. *Journal of the American Society for Information Science and Technology*, 61, 2544-2558.

Thomas, K., Grier, C., & Paxson, V. (2012, April). Adapting social spam infrastructure for political censorship. In *Proceedings of the 5th USENIX conference on Large-Scale Exploits and Emergent Threats* (pp. 13-13). USENIX Association.

Turner, J. C., (1987). *Rediscovering the social group: A self-categorization theory*. Oxford: basil Blackwell.

United States Census Bureau. (2014, Dec 4). 2013 ACS Data Release. Retrieved from http://www.census.gov/acs/www/data_documentation/2013_release/

Valenzuela, S., Kim, Y., & Gil de Zúñiga, H. (2012). Social networks that matter: Exploring the role of political discussion for online political participation. *International Journal of Public Opinion Research, 24,* 163 – 184.

Vissers, S., & Stolle, D. (2014). The Internet and new modes of political participation: online versus offline participation. *Information, Communication & Society*, *17*(8), 937-955.

Wegner, G. (1991). Review: Political socialization and schooling. *History of Education Quarterly, 31*, 389-391.

Wolack, J. (2012). The consequence of presidential battleground strategies for citizen engagement. *Political Research Quarterly, 59,* 353 – 361.

Vargo, C. (2011). Twitter as Public Salience: An Agenda-Setting Analysis. Paper presented AEJMC, St. Louis, Mo.

Vieweg, S. (2010). The ethics of Twitter research. Paper presented at the Computer Supported Cooperative Work 27. (pp. 1-3).

Vigdor, J. L. (2004). Community composition and collective action: Analyzing initial mail response to the 2000 Census. *The Review of Economics and Statistics, 86,* 303 – 312.

Wasserman, D. (2013). Introducing the 2014 Cook Political Report Partisan Voter Index. Accessed December 12, 2014, from http://cookpolitical.com/story/5604.

Welch, M. R., & Baltzell, J. (1984). Geographic mobility, social integration, and church attendance. *Journal for the Scientific Study of Religion, 23,* 75 – 91.

Wellman, B., Haase, A. Q., Witte, J., & Hampton, K. (2001). Does the Internet increase, decrease, or supplement social capital? Social networks, participation, and community commitment. *American Behavioral Scientist, 45*, 436 – 455.

Yuan, Y. C., & Gay, G. (2006). Homophily of network ties and bonding and bridging of social capital in computer-mediated distributed teams. *Journal of Computer-Mediated Communication, 11*, 1062 – 1084.

## Endnotes

1. *Potential* could be interpreted in a number of ways. For instance, high levels of racial heterogeneity could present the *potential* for citizens to build ties with members of a number of different racial groups. However, in the present study, we interpreted potential in terms of factors previously shown to *empirically* affect the production and maintenance of social capital on the community/contextual level.

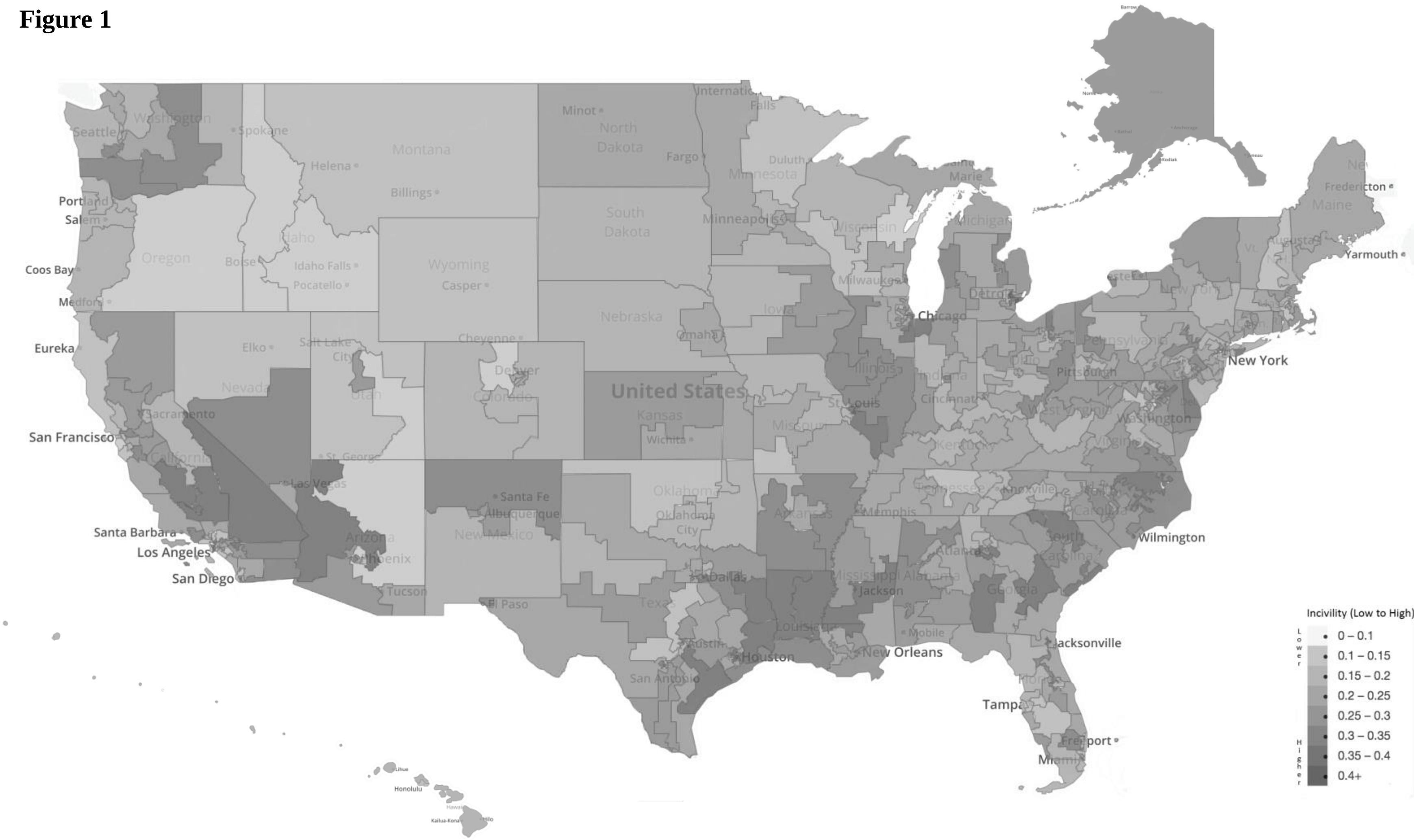
2. Before aggregation, we examined the descriptive statistics associated with Twitter usage on the individual level. Number of analyzed Tweets per user ranged from 1 to 743 (M = 2.83, SD = 11.09).

3. Due to the strong correlation between household income and education measures (*r* = .88; see Table 4), the SES-only and all variables models (Table 5) possessed relatively high levels of multicollinearity. Removal of the household income measure from these models did not meaningfully impact either the patterns of significance (including direct and relationship strength) or the amount of variance explained in the criterion variable. Moreover, the highest observed VIF coefficient was well below the commonly employed heuristic of 10.00 (e.g., O'Brien, 2007). As such, the models were presented in the initially hypothesized form.

3. Hayes' (2013) PROCESS macro was used to estimate the coefficients for the Johnson-Neyman technique.

# Figures and Tables

**Figure 1**

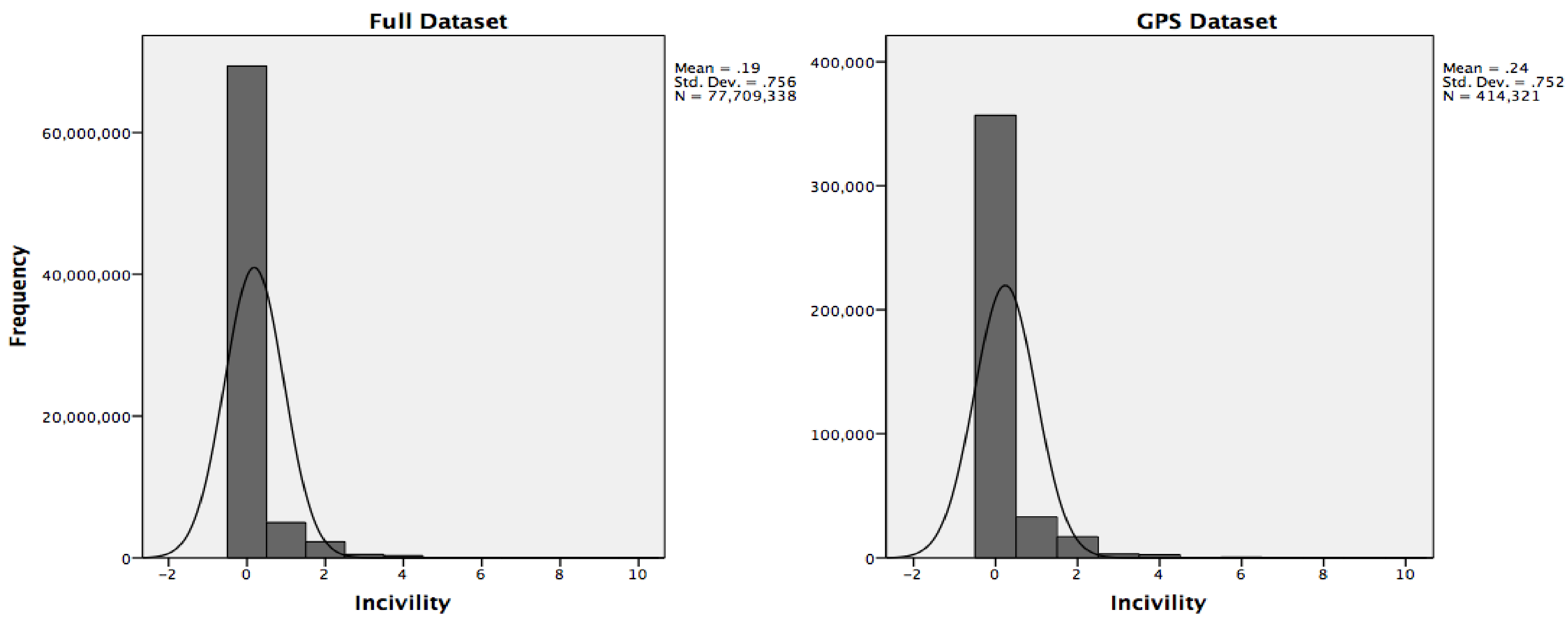



Visual representation of Twitter incivility in each U.S. congressional district

**Table 1**

| Non-Geotagged Dataset | | | | | Geo-tagged Dataset | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Username | # Tweets | Tweets/Day | Bot? | Susp./Del? | Username | # Tweets | Tweets/Day | Bot? | Susp./Del? |
| preciousliberty | 21179 | 181.02 | Yes | No | salsonthejob | 1194 | 10.21 | Yes | Yes |
| skew11 | 19487 | 166.56 | No | No | jidts07 | 754 | 6.44 | No | No |
| ProLiveDesign | 16054 | 137.21 | Yes | No | johnrogers411 | 743 | 6.35 | No | No |
| missb62 | 15398 | 131.61 | No | No | ganggreenie36 | 737 | 6.30 | No | No |
| PennyAgain | 14681 | 125.48 | Yes | Yes | uselephants | 732 | 6.26 | No | No |
| word_34 | 14411 | 123.17 | No | No | _pascuzzo | 706 | 6.03 | No | No |
| Maggyw519 | 14330 | 122.48 | Yes | No | armanwalker | 673 | 5.75 | No | No |
| redostoneage | 14114 | 120.63 | Yes | No | gayle_ryan | 661 | 5.65 | No | No |
| CelebVolger | 14013 | 119.77 | Yes | Yes | 418829 | 648 | 5.54 | No | No |
| PatsyRoussel | 13711 | 117.19 | No | No | Linda_Davis | 644 | 5.50 | No | No |
| Norsu2 | 13456 | 115.01 | No | No | duncan2665 | 559 | 4.78 | No | No |
| myvotefactor | 12631 | 107.96 | Yes | No | MauriceMichael | 551 | 4.71 | No | No |
| WiseUpLibs | 12172 | 104.03 | Yes | Yes | Llauger1964 | 543 | 4.64 | No | No |
| PatDollard | 11900 | 101.71 | No | No | WebAntOnYouTube | 537 | 4.59 | Yes | Yes |
| slone | 11805 | 100.90 | No | No | RRRoxy | 515 | 4.40 | No | No |
| obamaB0T | 11654 | 99.61 | Yes | Yes | phyrefyter | 501 | 4.28 | No | No |
| budholly59 | 11620 | 99.32 | No | No | luckyseas1 | 484 | 4.14 | No | No |
| RomneyTruth | 11584 | 99.01 | Yes | Yes | Islandmike13 | 475 | 4.06 | No | No |
| ThisisTrixx | 11549 | 98.71 | No | No | pbump | 471 | 4.03 | No | No |
| Clickman8 | 11509 | 98.37 | Yes | Yes | conservJ | 464 | 3.97 | No | No |
| MdlMurray | 11479 | 98.11 | No | No | AlanStump1 | 447 | 3.82 | No | No |
| OBAMA_GAMES | 11400 | 97.44 | Yes | No | luckyboy2007 | 443 | 3.79 | No | No |
| WashingtonDCTea | 11348 | 96.99 | No | No | hebert_jim | 432 | 3.69 | No | No |
| romneyobama | 11244 | 96.10 | Yes | No | Pogo407 | 426 | 3.64 | No | No |
| NancyWonderful | 11026 | 94.24 | No | No | frankieciv608 | 417 | 3.56 | No | No |
| Average | 13350.2 | 114.10 | 0.26 | 0.12 | Average | 590.28 | 5.05 | 0.04 | 0.04 |

**Table 2**



Full Dataset
Mean = .19
Std. Dev. = .756
N = 77,709,338
Frequency
60,000,000
40,000,000
20,000,000
0
−2
0
2
4
6
8
10
Incivility
GPS Dataset
Mean = .24
Std. Dev. = .752
N = 414,321
400,000
300,000
200,000
100,000
0
−2
0
2
4
6
8
10
Incivility

**Table 3**

| Indicator | Range | M | SD | Skew | Kurtosis |
|---|---|---|---|---|---|
| Average Incivility | 0.09 – 0.53 | 0.28 | 0.08 | 0.37 | 0.16 |
| Racial Heterogeneity | 0.06 - 0.75 | 0.38 | 0.16 | 0.11 | - 1.00 |
| Residence Tenure (%) | 13.84 – 50.70 | 31.94 | 6.30 | 0.00 | - 0.31 |
| Geographic Mobility (%) | 21.82 – 85.66 | 58.96 | 13.86 | - 0.44 | - 0.42 |
| Household Income ($) | 36,886.00 - 155,093.00 | 73,564.20 | 19,989.20 | 1.36 | 2.17 |
| Unemployment Rate (%) | 2.60 – 19.10 | 8.54 | 2.48 | 0.83 | 0.12 |
| % with Bachelor Degree/Higher | 8.30 – 69.16 | 29.26 | 10.18 | 0.87 | 0.79 |
| % with Healthcare Coverage | 61.53 – 97.39 | 85.48 | 5.71 | 0.12 | 1.52 |
| Partisan Polarity | 0.02 – 44.00 | 12.50 | 8.96 | 0.91 | 0.42 |
| District Population | 523,170 – 1,015,165 | 725,246.87 | 36,620.59 | 0.48 | 15.09 |
| Median Age | 27.70 – 52.30 | 37.82 | 3.65 | 0.20 | 0.76 |
| % Urban | 23.51 – 100.00 | 80.72 | 19.01 | -0.75 | - 0.53 |
| Number of Users | 82 -1,436 | 329.66 | 158.13 | 2.11 | 8.16 |
| Avg. Tweets Per User | 1.53 – 7.08 | 2.84 | 0.76 | 1.96 | 5.65 |

Descriptive statistics for items included in analyses

**Table 4**

| | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **10** | **11** | **12** | **13** | **14** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Avg. Incivility (1) | | .21*** | .08 | .32*** | -.34*** | .39*** | -.42*** | -.24*** | .03 | -.11* | -.20*** | -.05 | .00 | -.12* |
| Racial Heterogen.(2) | | | -.34*** | -.34*** | .14** | .47*** | .13** | -.37*** | .45*** | .15** | -.50*** | .59*** | .29*** | -.06 |
| Residence Tenure (3) | | | | .62*** | .03 | -.10* | -.09 | .50*** | -.19*** | -.26*** | .48*** | -.39*** | -.15** | -.16** |
| Geog. Mobility (4) | | | | | -.30*** | -.03 | -.35*** | .31*** | -.10* | -.17** | .04 | -.45*** | -.09 | -.22*** |
| Household Inc. (5) | | | | | | -.44*** | .88*** | .43*** | -.16** | .13** | .16** | .42*** | .09 | .08 |
| Unemp. Rate (6) | | | | | | | -.46*** | -.44*** | .31*** | -.10* | -.21*** | .17** | .12* | -.11* |
| % w/ Bach. Deg./+ (7) | | | | | | | | .47*** | -.07 | .15** | .14** | .45*** | .29*** | .07 |
| % w/ Healthcare Cov. (8) | | | | | | | | | -.31*** | -.06 | .50*** | -.18*** | .04 | .04 |
| Partisan Polarity (9) | | | | | | | | | | . 04 | -.37*** | .16** | .28*** | .03 |
| District Population (10) | | | | | | | | | | | -.17** | .14** | .12* | .00 |
| Median Age (11) | | | | | | | | | | | | -.34*** | -.16** | .11* |
| % Urban (12) | | | | | | | | | | | | | .25*** | .01 |
| Number of Users (13) | | | | | | | | | | | | | | -.05 |
| Avg. Tweets/User (14) | | | | | | | | | | | | | | |

*Note:* * $p < .05$, ** $p < .01$, *** $p < .001$

Zero-order correlations between items

**Table 5**

| **Variable Name** | **Controls Only** | **Social Capital** | **SES Indicators** | **Political Polarity** | **All Variables** | **VIF (All Variables Model)** |
|---|---|---|---|---|---|---|
| District Population | -.14** | - .11* | -.07 | -.14** | -.07 | 1.13 |
| Median Age | -.25*** | .02 | -.13* | -.25*** | .07 | 2.55 |
| % Urban | -.11* | -.10 | .05 | -.11* | .02 | 2.35 |
| Number of Users | .00 | - .06 | .10* | .01 | .07 | 1.46 |
| Avg. Tweets Per User | -.09 | -.02 | -.06 | -.09 | .01 | 1.14 |
| Ethnic Heterogeneity | | .43*** | | | .35*** | 2.89 |
| Residence Tenure | | - .18** | | | -.10 | 2.99 |
| Geographic Mobility | | .51*** | | | .44*** | 2.95 |
| Household Income | | | 25** | | .12 | 5.92 |
| Unemployment Rate | | | .20*** | | .12 | 2.43 |
| % Bachelor Degree/Higher | | | -.62*** | | -.37*** | 8.00 |
| % with Healthcare Coverage | | | .10 | | -.12 | 2.97 |
| Partisan Polarity | | | | -.04 | -.17*** | 1.48 |
| $R_2$ | .08 | .26 | .27 | .07 | .38 | |
| $\Delta R_2$ (Rel. to *Controls Only*) | | .18*** | .19*** | .00 | .30*** | |
| *F* | (5, 429)= 7.41 $p < .001$ | (8, 426) = 18.43 $p < .001$ | (9, 425) = 17.45 $p < .001$ | (6, 428) = 6.28 $p < .001$ | (13, 421) = 19.98 $p < .001$ | |

Note: * $p < .01$; ** $p < .01$, *** $p < .001$

Standardized coefficients describing relationship between social capital potential indicators, SES status indicators, partisan polarity, and Twitter incivility

**Table 6**

| **Variable** | ***β*** |
|---|---|
| % Born in State | .22*** |
| SES Status | -.43*** |
| % Born in-State x SES Status | -.11* |
| | |
| $R^2$ | .24 |
| *F* | (3, 431) = 45.69, $p < .001$ |

*Note:* * $p < .05$; ** $p < .01$, *** $p < .001$

Exploratory regression model exploring effect on % born in-state on Twitter incivility at different SES levels

**Table 7**

| Standardized SES Value | $\beta$ | *p* value |
|---|---|---|
| -2.03 (Lowest possible SES value) | .45 | < .001 |
| -1.74 | .42 | < .001 |
| -1.46 | .39 | < .001 |
| -1.17 | .35 | < .001 |
| -0.89 | .32 | < .001 |
| -0.60 | .29 | < .001 |
| -0.32 | .26 | < .001 |
| -0.03 | .22 | < .001 |
| 0.26 | .19 | < .001 |
| 0.54 | .16 | < .01 |
| 0.83 | .13 | < .05 |
| 0.89 | .12 | ≤ .05 |
| 1.11 | .09 | > .05 |
| 1.40 | .06 | > .05 |
| 1.69 | .03 | > .05 |
| 1.97 | -.01 | > .05 |
| 2.26 | -.04 | > .05 |
| 2.54 | -.07 | > .05 |
| 2.83 | -.10 | > .05 |
| 3.11 | -.13 | > .05 |
| 3.40 | -.17 | > .05 |
| 3.69 (Highest possible SES value) | -.20 | > .05 |

Johnson-Neyman analysis showing standardized relationship between percentage of population born in-state and incivility on Twitter at various levels of district SES