Density dependence without resource partitioning on an online petitioning platform

Nathan TeBlunthuis

A thesis submitted in partial fulfillment of the requirements for the degree of

Master of Arts

University of Washington

2017

Committee:

Benjamin Mako Hill

W. Lance Bennett

Program Authorized to Offer Degree:
Communication

University of Washington

Abstract

Density dependence without resource partitioning on an online petitioning platform

Nathan TeBlunthuis

Chair of the Supervisory Committee:
Assistant Professor Benjamin Mako Hill
Department of Communication

Online petitions are a collective action tactic that leverages digital affordances in pursuit of discursive opportunities. Prior efforts to explain why some petitions are more successful than others emphasize signer motivations, petition framing, social media, or resources from movement organizations. We advance a key insight of organizational ecology: population-level variables like density and concentration also constrain success. We use latent Dirichlet allocation (LDA) topic models to measure overlap density and frame specialization. We then model how ecological dynamics affect petition signature counts. We observe density dependence: a curvilinear relationship between overlap density and success. We anticipated resource partitioning: specialists enjoy competitive advantages under concentration, but we find no evidence for it. We discuss boundary conditions for ecological dynamics commonly found in organizational fields induced by the distinctive scope of e-tactic platforms. Platforms may produce concentration without advantages for specialists by lowering entry costs for generalists and specialists alike.

1.1 Introduction

People are increasingly turning to digital communication channels in pursuit of voice (DiMaggio, Hargittai, Neuman, & Robinson, 2001). This is an important development in ongoing social change. Social movements now commonly use digital channels to organize, mobilize protests, distribute cultural products, and pursue discursive opportunities (Klandermans, van Stekelenburg, Damen, van Troost, & van Leeuwen, 2014; Tarrow, 2011; Vasi, Walker, Johnson, & Tan, 2015). An expanding "e-tactic" repertoire, organized and performed primarily online, further leverages digital affordances to achieve widespread participation (Earl & Kimport, 2011; Rolfe, 2005). E-tactics include disruptive protests such as distributed denial of service attacks and also more conventional techniques like boycotts, email campaigns, and petitions.

Although opportunities for pursuing voice proliferate on the internet, obtaining a large audience remains difficult for most (DiMaggio et al., 2001; Goldhaber, 1997; Tufekci, 2013). Many speak. Few are heard. This is evident on e-tactic platforms. Hundreds of thousands run online petition campaigns, but only a small number obtain a great many signatures (Margetts, John, Hale, & Yasseri, 2015).

We aim to better understand inequalities of digital voice using the lens of organizational ecology theory (Hannan & Freeman, 1977; Carroll, 1985). Social movement scholars have fruitfully applied organizational ecology to understand how ecological dynamics relate to organizational success and growth and change in social movements (Soule & King, 2008; Minkoff, 1997, 1995; Olzak & Ryo, 2007; Olzak & Uhrig, 2001). Many petitions have similar targets, frames, and aims. We argue that they consequently compete over overlapping attention resources and that this influences their success. We test two classic organizational ecology theories: resource partitioning (Carroll, 1985) and density dependence (Carroll & Hannan, 1989; Hannan & Freeman, 1989).

E-tactics introduce a distinctive scope for organizational ecology. The ranges of density (the number of actors desiring a resource) and concentration (the inequality of the resource distribution) exceed that traditionally observed in the social movement sector. The industries previously studied using organizational ecology have high entry costs and durable organizations. In the domain of e-tactics, entry costs are very low, platforms provide non-rival resources, and informal and short-lived forms of organization are common.

We focus on online petitions, a popular e-tactic in movement repertoires (Earl & Kimport, 2011; Earl, 2006; Tarrow, 2011; Tilly, 1995). Political scientists recognize that petitions have long ranked among the most widespread of modern political tactics for recruitment, cadre development, and influence seeking (Carpenter & Moore, 2014; Carpenter & Schneer, 2015; Nall, Schneer, & Carpenter, 2017). While it is commonly observed that petitioning platforms, like many sites of online participation, are dense and concentrated environments, prior analyses study factors including attributes of potential signatories (Hale, Margetts, & Yasseri, 2013; Jungherr & Jürgens, 2010; Margetts et al., 2015; Yasseri, Hale, & Margetts, 2013), preferential attachment (van de Rijt, Akin, Willer, & Feinberg, 2016; van de Rijt, Kang, Restivo, & Patil, 2014), and textual attributes of individual petitions (Hagen et al., 2016). They do not consider ecological factors like specialization and competition.

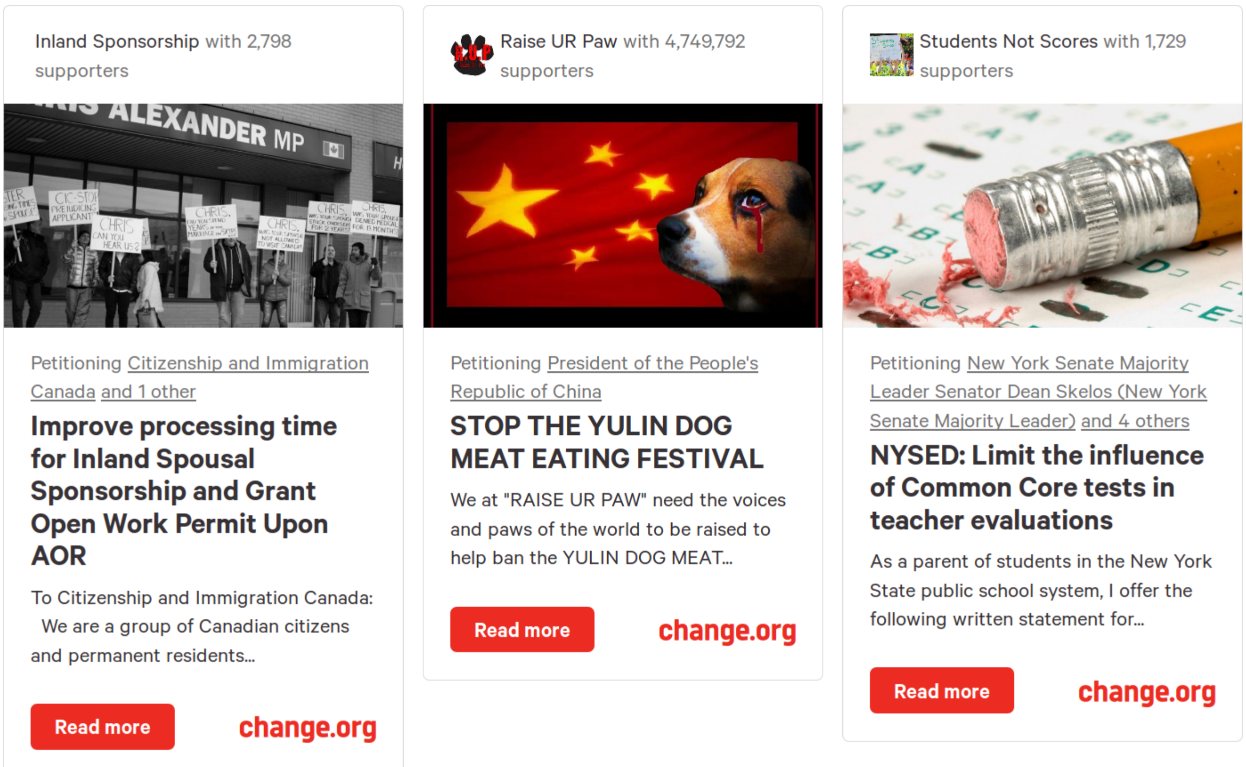
To test models of density dependence and resource partitioning in the domain of e-tactics, we must assign individual petitions to groups and measure petition specialization. Doing so, we encounter a familiar problem for students of organizational ecology. As Hannan and Freeman (1977) write: "unfortunately, identifying a population of organizations is no simple matter." Prior work generally uses either manual schemes for categorical assignment developed and applied by researchers and their assistants, or coarse proxies based on established industrial boundariesh.

We make a methodological contribution to population ecology by using latent Dirichlet allocation (LDA) topic models (Blei, Ng, & Jordan, 2003; DiMaggio, Nag, & Blei, 2013) to measure niches. LDA, an unsupervised machine learning algorithm, constructs topics by modeling distributions of words across petitions. The degree to which a petition's lexicon matches a topic's distribution of words determines the petition's membership in a given topic. These topics place petitions in a multidimensional space. Petitions near to one another in the space use similar language and are therefore likely to appeal to related sets of preferences, claims, and issues. We say that such petitions have higher levels of density. Figure 1.1 shows example petitions with varying levels of density according to our model.

Topic modeling also affords measuring a petition's "frame specialization" according to the inequality of its topic memberships (Soule & King, 2008; Cress & Snow, 2000). Algorithmic assignment of petitions to topics makes it possible to analyze a dataset of 442,109 Change.org petitions. This dataset is orders of magnitude larger than those used in prior empirical studies in organizational ecology. Our novel method has broad applications for ecological studies of e-tactics and other textual artifacts.

We find evidence of density dependence but not resource partitioning between generalists and specialists. As predicted by density dependence theory, we observe an ∩-shaped (inverse-U-shaped) relationship between the density of petitions in topic space (overlap density) and the expected level of success. The ∩-shaped relationship implies that as overlap density increases, mutualism is exchanged for competition. The importance of mutualism declines and the importance of competition grows. This suggests that early successes can signal that a tactic is effective and that exogenous factors like political opportunities jointly promote petition creation and increase the available resources for mobilizing participants. A large majority of petitions in our dataset have levels of overlap density where increasing density is associated with increasing participation. This suggests that the positive benefits of mutualism and shared opportunities typically outweigh the negative effects of competition.

We anticipate, following the classical resource partitioning model, that petitions are more successful in more concentrated populations (Carroll, 1985). However, we observe no relationship between concentration and specialist success. Three possible explanations for this result are that resource partitioning occurs but drives a small part of variance in concentration; that opportunities for specialists exist but systematically go unrealized; and that niches where specialists have competitive advantages are not plentiful. In our discussion, we suggest that future work should extent ecology analyses to incorporate resources provided by digital platforms into theoretical models. Such resources may be sufficient for production



Low Density: Spousal Immigration	Medium Density: Animal Cruelty	High Density: Education

Figure 1.1: Examples of Change.org petitions with varying levels of density.

of generalist and specialist petitions alike.

1.1.1 Population ecology of digital voice

Technological change provides social scientists the opportunity to study collective action in new contexts (DiMaggio et al., 2001). We consider the continuing and widespread adoption of computer mediated communication technologies which began in the late 20th century. How is the new media context shaping collective action for social change? One way is that e-tactics including petitions, boycotts, and other forms of direct action organized online are now a normal part of society, particularly as used by young people in democracies (Bakker & de Vreese, 2011; Zukin, 2006; Inglehart & Catterberg, 2002; Earl & Kimport, 2009). Strong claims that digital media are fundamentally empowering to social movements are hard to sustain, but a range of weaker claims are credible (Farrell, 2012). Digital media make it easier to reach broader audiences, to communicate asynchronously and remotely, and to reproduce messages. While far from revolutionary, such affordances are widely adopted by social movements and shape organizing structures, tactical repertoires, and framing processes (Bennett & Segerberg, 2013; Earl & Kimport, 2011; Vasi et al., 2015; Garrett, 2006).

Qualitative evidence shows that digital media afford discursive opportunities for movements by changing the structure of framing processes (Graeff, Stempeck, & Zuckerman, 2014; Vasi et al., 2015). Framing processes are the ways movements work to continuously construct meaning for both insiders and outsiders, and are important for producing collective identity, obtaining influence, and movement outcomes (Cress & Snow, 2000; Benford & Snow, 2000; Snow, Rochford, Worden, & Benford, 1986; Snow, 2004). E-tactics, online activist media, and social network sites can help movements circumvent the interface between social movement organizations and media gatekeepers that historically constrained access to broader audiences (Tufekci, 2013; Gamson & Wolfsfeld, 1993; Garrett, 2006).

That said, digital media cannot be expected to erode underlying structural "rules of access" that determine who obtains standing or what frames are repeated (Gamson & Wolfsfeld, 1993). Instead, those seeking voice through digital media encounter distinctive constraints.

Norms about what makes digital content viral or shareable may promote "personalized action" frames that interpretively flexible (Bennett, 2012) or "post-materialist particularism" frames that avoid ideological or redistributive claims (Hersh & Schaffner, 2015). Historically, movement organizations could select media representatives with intention. Today, digital representatives often emerge from pools of movement supporters with technology access and skills (Tufekci, 2013; Garrett, 2006). Such capabilities are distributed along lines of inequality following the digital divide that privilege the young, male, educated, and able-bodied (DiMaggio et al., 2001; Hargittai & Shaw, 2013, 2015). Furthermore, the very platforms widely used in attempts to create discursive opportunities are themselves powerful actors with their own politics, organizational interests, and rules of access (Gillespie, 2010; Karpf, 2016).

In addition to the above, we argue that activism campaigns in dense and concentrated digital media environments may support or constrain one anothers' chances for success through ecological dynamics akin to those studied in organizational ecology (Hannan & Freeman, 1977). A central insight of ecology is that success of a particular individual depends on the group of other individuals sharing the environment. Individuals in a competitive dynamic struggle with one another for needed resources and suffer decreased chances of growth and survival. In a mutualism dynamic individuals contribute common resources that improve the chances of growth and survival of others in the group. Organizational sociologists discovered that such dynamics occur not only in biological populations but also in groups of organizations (McPherson, 1983; Baum & Shipilov, 2006; Hannan & Freeman, 1984; Carroll, 1984; Swaminathan, 2001). We suggest that these dynamics also play out in groups of digital mobilizations because attempts to create discursive opportunities through e-tactic campaigns may mutually support (or compete over) beneficial resources like attention, networks of supporters, competency in digital culture, technology skills, and access to media gatekeepers.

Online petitions continue a long tradition in contentious politics spanning many social, political, and historical contexts. Petitioning is an established tactic for creating discursive opportunities and organization building (Karpf, 2016; Earl & Kimport, 2011). Historical examples include wide and impactful adoption by the British abolitionist movement in the 18th century (Tarrow, 2011; Tilly, 1995). Petitions against U.S. President Andrew Jackson were instrumental in the formation of the Whig party in the 1830s (Carpenter & Schneer, 2015). Woman abolitionists in the U.S canvassed petitions to spread their movement and developed a cadre of organizers who became leaders of the suffragist movement (Carpenter & Moore, 2014).

A important episode in the emergence of the "Black Lives Matter" movement shows how movements can use online petitions to create discursive opportunities (Graeff et al., 2014).[1] The news that George Zimmerman killed Trayvon Martin may have fallen off the media radar if Kevin Cunningham, then a law student at Howard University, had not seen the story and created a Change.org petition.[2] While major newspapers picked up the story only following street demonstrations, the petition provided material to activist online publications who issued stories marking signature count milestones (e.g. the 100,000$^{\text{th}}$ signature). Resources utilized to mobilize participants included celebrity tweets, support by social movement organizations, and coverage in activist media. Cunningham's petition obtained over a million signatures before the first protest (Graeff et al., 2014). Change.org subsequently became a popular platform for online petitioning by Black Lives Matter activists (Karpf, 2016).

Petitions, both online and offline, follow a conventional format with two components: a "prayer" and a signature list (Carpenter & Moore, 2014). The prayer typically expresses a grievance and an appeal for resolution and is addressed to target(s) who have the power to redress the grievance. Both the quantity and credibility of names on the list legitimize the petition to potential signers and persuade the targets to make concessions. Petitions are also a valuable tool for recruitment as organizations use them to collect names and contact information of signers (Nall et al., 2017).

Online petitioning adapts this established tactic to leverage affordances of internet com-

---

[1] One could argue that it is accurate to call Black Lives Matter a cycle of the civil rights movement.

[2] https://www.change.org/p/prosecute-the-killer-of-our-son-17-year-old-trayvon-martin

munication technology (Earl & Kimport, 2011). Affordances for recruitment partly explain the popularity of online petitioning with some contemporary social movement organizations (Karpf, 2016). That said, a minority of the petitions in our dataset are created by organizations. Platforms that allow users to easily create petitions and mobilize participants have helped routinize the tactic not only for movement organizations, but also for individuals, and even for actors acting outside of recognizable social movements.

Some petitions in our dataset concern popular culture and have little overt political component. They may petition television networks to renew programs for additional seasons or video game companies to change the rules to online games. These petition campaigns are not frivolous, rather they engage in issues that are meaningful to participants. They are a striking example of how routinized tactics can achieve widespread adoption in movement societies (Earl & Kimport, 2009; Earl & Schussman, 2008; Karpf, 2016).

It is clear that online petitions are a mainstream means for pursuing voice. Despite this popularity, only a tiny minority of online petitions garner many signatures (Margetts et al., 2015). To be heard by many, or to help an organization identify supporters, an online petition must obtain signatures. Yet mobilizing a large number of participants is a difficult collective action problem because the marginal contribution of individuals who sign, canvass, or disseminate a petition to obtaining the collective goods demanded in the prayer is relatively small (Olson, 1965).

Some prior efforts to explain the distribution of signatures view social information as a kind of selective incentive that when mediated by psychological factors can motivate participation (Margetts et al., 2015). Studies in this vein look to features of petition platforms like recommendation engines, effects of early signatures, information dissemination on social media, psychological factors (Yasseri et al., 2013; Hale et al., 2013; Margetts et al., 2015), and the contributions of a few highly active signers (Jungherr & Jürgens, 2010; Huang, Suh, Hill, & Hsieh, 2015). Celebrity endorsements in social media can be a major driver of signature growth (Graeff et al., 2014). Others seek to understand lexical patterns associated with petition success (Hagen et al., 2016). Still others find evidence of weak preferential attachment dynamics (van de Rijt et al., 2014; van de Rijt et al., 2016).

These studies identify resources important for petition success but they primarily seek to explain why individuals participate in e-tactics and view collective actions in isolation. They do not empirically consider how petition success may relate to competition, conditions promoting petition creation, or how concentration shapes what kinds of petitions are successful. We address these gaps by extending population ecology to study e-tactics.

Organizational ecology has been fruitful to social movement theory (Olzak & Uhrig, 2001; Olzak & Ryo, 2007; Minkoff, 1997, 1995; Soule & King, 2008). Resource mobilization theory, a hugely influential perspective on social movements, focuses on how social movement organizations provision selective incentives to enable mobilization, despite the problem of collective action (McCarthy & Zald, 1977, 2001; Olson, 1965). Like organizational ecology, resource mobilization theory sees contests over resources as central to organizational success. In particular, it predicts that organizations specialize to avoid competition but that specialization increases risk of failure (McCarthy & Zald, 1977; Zald & McCarthy, 1980). Soule and King (2008) employ resource partitioning theory, an ecological theory which instead predicts that specialists can be more likely to survive, but only in conditions of concentration. By synthesizing resource mobilization theory with resource partitioning theory, they formulate and validate a more complete theory of specialization for social movement organizations.

Minkoff (1997) applies density dependence: another theory of organizational ecology. She finds evidence that civil rights movement organizations popularized tactics and organizational forms subsequently used by the women's movement. She argues that organizational development, legitimization, and learning help drive protest cycles. Olzak and Ryo (2007) find that organizational density has curvilinear relationships with diversity of goals and diversity of tactics in the civil rights movement and Olzak and Uhrig (2001) use a density dependence argument to study the legitimation and diffusion of radical tactics in the women's movement in Germany in the mid 20th century.

We apply organizational ecology to consider online petitions. Online petitions need resources to mobilize participants like how firms and social movement organizations need resources to survive. In population ecology, such resources are conceived as distributed in pools called *niches* (Hannan & Freeman, 1977). Like organizations, online petitions have heterogeneous resource requirements. It is therefore useful to distinguish between an individual petition's *fundamental niche* and its *realized niche* (Baum & Shipilov, 2006; Dobrev, Kim, & Hannan, 2001; McPherson, 1983).[3] A fundamental niche is defined by the totality of resources an individual is capable of utilizing. However, an individual is often unable to control its entire fundamental niche. The subspace of the fundamental niche an individual is able to utilize is called its realized niche. Fundamental niches overlap when they contain some of the same resources (Baum & Singh, 1994). A individual's *overlap density* is the amount its fundamental niche overlaps with fundamental niches of others in the group (McPherson, 1983; Baum & Singh, 1994). Competition occurs when resources in niche overlaps are rival so it is intuitive that overlap density is positively associated with competition.

We argue that online petitions may compete over signatures. The act of signing a petition appears relatively cheap (Margetts et al., 2015). Yet the ease of clicking a button to sign a petition belies the resources needed to disseminate the petition and persuade potential participants which may be both scarce and rival. When many petitions espouse similar claims they appeal to similar and overlapping pools of supporters. These supporters have limited time and attention to devote to signing petitions. They also have limited capacities to canvass these petitions in their social networks. Advertisements, key endorsements by high status individuals, and placement in featured lists on platforms may be especially scarce and costly. Therefore, we expect a competitive dynamic to reduce the average levels of petition success when overlap density is great.

While competition can make success more elusive, a lack of competition may reflect a lack of opportunity. Fundamental niche size is positively related to overlap density because the number of petitions seeking a resource is related to its perceived utility. The causal direction of mechanisms for this relationship may go in either direction. Exogenous increases in the size

[3]Many studies use datasets that do not afford distinguishing between fundamental and realized niches or measuring niche overlap and so avoid such matters.

of the fundamental niche may increase overlap density because more resources are available for crafting and promoting petitions. Conversely, mutualism between petitions may increase the size of the fundamental niche.

More concretely, consider how the most successful petitions, like the Trayvon Martin petition, promote imitator petitions (Graeff et al., 2014; Karpf, 2016). Actors who do not see value in online petitioning are not likely to create many petitions. A movement may believe that its aims, frames or political opportunities are not amenable to online petitioning. Subsequent examples of successful petitions may counter such beliefs, legitimizing the tactic. The movement may then increase the resources devoted to petitioning and thereby the quantity and success of petitions. A movement may lack digital campaigning skills or media resources beneficial for disseminating petitions. Experience running petition campaigns may increase these capacities and thereby the size the fundamental niche.

The trade-off between beneficial mutualism and damaging competition implies that petitions with extremely dense or extremely sparse fundamental niches have worse chances than petitions in between. When overlap density is low, the positive relationship between overlap density and success is stronger than the negative relationship and increases in overlap density reflect a more hospitable environment. However, the marginal benefits of overlap density decline as expansion of the fundamental niche slows. At high levels of overlap density, competition dominates as additions to the population shrink realized niches more than they grow the fundamental niche. Based on the cumulative effect of the mutualistic and competitive relationships between overlap density and petition success, we predict the classic curvilinear result of density dependence theory (Carroll & Hannan, 1989). **H1: Petitions with intermediate overlap density are more successful on average than petitions with high or low overlap density.**

Density dependence theory helps us understand how ecological factors shape petition success, but does not consider causes or effects of extreme participation inequality. It does not explain how petition creators may effectively respond to competitive pressures. Resource partitioning theory says that specialization can be an effective response to competition when resources are concentrated (Carroll, 1985). In order for specialists to exist, the environment must be high-dimensional so that differentiation is possible. Generalists have large fundamental niches because they attempt to expand and control a large share of resources. Specialists are able to realize a smaller niche where they have competitive advantages. The co-existence of large generalists and small specialists produces concentration because relatively many specialists can succeed, but generalists hold the major share of resources. For example, Carroll and Swaminathan (2000) argue that the high dimensionality of consumer taste in beer, especially preferences for local beer, helped the emergence of specialist microbreweries in the United States, a market previously dominated by a small number of generalist macrobrewers.

There are many ways that organizations can specialize. For example, Soule and King (2008) apply resource partitioning to study how social movement organizations may specialize in goals or in tactics. They suggest that future work consider "frame specialization," an additional dimension along which movement organizations may specialize. We consider frame specialization for petitions because it is a salient conception of specialization for intrinsically discursive objects. We use LDA topic models to measure frame specialization according to the degree to which a petition belongs to few topics or to many (Blei et al., 2003). This aligns closely with Carroll's (1985) definition of a specialist as focused on a few resource domains. A frame specialized petition has a prayer that raises relatively few distinct grievances, issues, ideas, and frames. Consider a petition that makes claims both about environmentalism and animal rights. Such a petition is a generalist relative to a petition that makes claims only about one of these topics. It attempts to mobilize participants who sympathize with both the environmentalist and animal rights movements. In doing so, it competes in the domain of environmentalist petitions and in the domain of animal rights.

Carroll (1985) specifies a number of conditions for resource partitioning in addition to a high dimensional environment. These are economies of scale, limited organizational pliability and agency, finite resources for sustenance, and units of analysis defined by market boundaries. Not all of the conditions described by Carroll (1985) are clearly present in online petitioning. For one, it is not obvious that petitions exhibit economies of scale. That said, information diffusion processes (Johnson, Faraj, & Kudaravalli, 2014; Cheng, Adamic, Dow, Kleinberg, & Leskovec, 2014), and rich-get-richer effects (van de Rijt et al., 2016) can also provide cumulative advantages to generalists. Also, while petitioning resources are clearly finite, petitions depend on these for success rather than for survival. The resources petitions need to persist are typically provided by the platform and are not rival between petitions. Survival is relevant in Carroll's (1985) model because concentration is an outcome of selection driven by organizational survival. Our hypothesis instead assumes that perceived availability of petitioning resources helps select which petitions are created.

We argue that populations of online petitions exhibit resource partitioning. The resource space is high dimensional because there exists strong preferences for certain ways of framing issues. Making a broad set of claims may be a good strategy for appealing to potential signers with heterogeneous values. At the same time, specializing on a narrow set of claims may be a good strategy when the claims are known to be highly motivating to potential signers. When many potential signers strongly prefer certain claims, petitions emphasizing these claims should be successful. Specialists should have competitive advantages in their niches because a petition cannot easily make a broad set of claims while emphasizing strongly held ones. We expect activists to produce a greater quantity of specialized petitions when conditions favor specialists. Yet specialists will rarely obtain as many signatures as the most successful generalists. The co-existence of many smaller specialists with a few large generalists produces concentration: a highly unequal distribution of signatures. Therefore we predict the classic result of resource partitioning theory (Carroll, 1985). **H2: Specialized petitions will be more successful in more concentrated fundamental niches.**

1.2 Empirical Setting: Change.org

We analyze petitions hosted on Change.org, likely the most widely used and impactful "warehouse site" or platform that hosts petitions created by users (Earl & Kimport, 2011). Some petition sites, such as Moveon.org, belong to social movement organizations who host their own petitions and focus on appealing to their membership for signatures (Karpf, 2016). Some governments maintain platforms like Whitehouse.gov and Petition.parliament.uk and promise official standing to petitions that garner a sufficient number of signatures. Government petition sites have a specific national context, but general purpose sites like Care2.com, Moveon.org and Change.org can contain petitions on a wide range of topics, in many languages, created by people around the world (Earl & Kimport, 2011).

Change.org is a for-profit petitioning platform. It has grown to a great scale, with over 2,000,000 petitions and 100,000,000 users in 2015.[4] It is easy to see from browsing Change.org that many petitions exist with similar or overlapping sets of grievances. The most signed Change.org petition at the time we collected our data, with 4,431,005 signatures, condemns the Yulin dog meat festival in GuangXi, China.[5] However, this is by no means the only petition against dog meat consumption or even that specific festival (there are dozens). Petitions making claims about animal rights or ethical treatment of animals are even more numerous. The basic observation that people create many similar petitions suggests that ecological factors may shape petition success.

1.3 Data

Our dataset consists of 442,109 petitions from Change.org. We only include petitions that were active for at least one week, and that are at least 480 characters long. Petitions shorter than this are generally poorly written and have few signatures. LDA topic models do not perform well on short texts. Topic modeling for short texts is an active area of research in computer science (Yan, Guo, Lan, & Cheng, 2013).

We used the *langdetect* python package, which boasts a 99.1% accuracy on English language classification, to identify the language of petitions.[6] We included only those petitions

---

[4] https://www.change.org/press (Accessed February 12, 2015)

[5] https://www.change.org/p/stop-the-yulin-dog-meat-eating-festival

[6] https://pypi.python.org/pypi/langdetect

which could reliably be identified as written in English.[7] Some petitions on Change.org appear independent, but are actually translated interfaces to the same petition campaign. We discovered and removed 9196 such duplicates and include only the earliest English language version of each.

Popular online platforms for user generated content attract "spam", the mass dissemination of unsolicited commercial messages. The data collected from the Change.org API contains a significant quantity of obvious spam "petitions." We identified and removed spam using Akismet, a tool widely used to detect spam in comments on Wordpress blogs.[8] We validated that this spam filter is effective on our data. We hand coded 100 randomly selected English petitions with more than 480 characters as spam or not-spam. The API correctly identifies 20 of 25 spam petitions with 1 false positive. More details of the data collection and cleaning process are found in Appendix A

1.3.1 Topic Modeling Petitions

Operationalizing ecological niches for online petitions presents a challenging methodological problem. Entities like firms, Change.org petitions, and even organisms are not objectively demarcated into kinds with the same fundamental niche.[9] Ecological studies of firms often use established constructions such as industry or market and assign groups using objective criteria such as SEC code, categories of inputs or outputs, membership overlap, or protest attendance (Baum & Shipilov, 2006; McPherson, 1983; Dobrev et al., 2001).

Such external criteria are useful because they do not rely on the subjective judgment of researchers. Using human coding of petitions to identify groups is subject to the threat of bias from group demarcation and assignment. Human coding requires defining niche boundaries either *a priori* or inductively. The boundaries of niches demarcated in this way are therefore

---

[7] Some petitions still contain non-english text because they contain multiple translations.

[8] https://akismet.com

[9] Debates within biology about where one species of organism ends and another begins predate Darwin and stretch to the present day (Hey, 2006; Hannan & Freeman, 1977).

sensitive to the researcher's beliefs about how resources are distributed. These may be subjective or based on limited information. The need to obtain inter-coder reliability limits the complexity of grouping criteria to those communicable between coders. These threats also apply to methods utilizing supervised machine classification and the method of Hopkins and King (2010).

A challenge for both human coding and external criteria is to convert categorical measures into niche overlaps. Real world firms, social movement organizations, and e-tactic protests often involve multiple domains. Firms can participate in multiple industries, social movement organizations can be active in multiple movements, and petitions can make claims about many topics. A realistic method of assigning membership would place individuals in multiple groups proportionally to the individual's niche overlaps with other members in the group. In settings like online petitioning, where external criteria are unavailable or unreliable, doing so is empirically challenging.

We overcome these challenges in group demarcation and assignment using an LDA topic model to operationalize ecological niches (Blei et al., 2003). LDA is an unsupervised machine learning model that assigns petitions to multiple topics according to the distribution of words in the petitions. DiMaggio et al. (2013) argue that topic models have strengths that give them affinity with sociological perspectives on framing. LDA has "ability to capture polysemy," the multiple meanings of individual words. Any particular word can be assigned to multiple topics. However the meaning of the word will be different in each topical context. For example the word "race" should take a different meaning in a topic about electoral campaigns ("races") than in a topic about racially discriminatory policing or about racing sports. As a mixed membership model, LDA also assigns petitions to multiple topics according to the probability that a word in the petition is drawn from a topic. Since topics are statistically independent, LDA jointly clusters petitions and places them in a high dimensional topic space. We assume that the fundamental niches of two petitions overlap proportionally to their closeness in this topic space.

Measuring specialization presents some of the same challenges as measuring niche overlap. Bias can enter through subjective, *a priori* identifications of specialization. Moreover, specialization is often treated as a dichotomous variable. It is more realistic to model firms, social movement organizations, and collective actions as specializing in varying degrees. As a mixed membership model, LDA allows documents to exhibit heteroglossia, the speaking with multiple voices (DiMaggio et al., 2013). We argue that each of these "voices" are related to a resource pool and therefore that LDA makes it possible for us to operationalize Carroll's (1985) definition of specialization as drawing from many resource pools. We define frame specialization according to the inequality of petition topic memberships (Soule & King, 2008). A petition that draws from few topics has greater frame specialization than one that draws from many.

We divided our dataset into a training set with 420,003 petitions and a validation set of 22,105 petitions and fit an LDA topic model on the training set. We use the validation set to select hyper parameters and and to monitor convergence. We used package NLTK (Loper & Bird, 2002) to preprocess petition texts and the software package Gensim (Řehůřek & Sojka, 2010, May, 22) to fit the topic model.

The most epistemically significant choice in LDA topic modeling is choosing the number of topics. The problem of selecting the number of topics is analogous to choosing between broad or narrow criteria for including organizations in a social movement industry, which as McCarthy and Zald (1977) note can be a difficult task. If the number of topics is too great, then many topics will be overly narrow or incoherent. If the number of topics is too few then topics will be too broad and not validly capture the notion of a niche.

We report results from a model with 100 topics. Our substantive results are robust to models using 30, 50, and 150 topics. Choosing a model with more topics means assuming that petitions draw from narrower and more finely grained topics, but increases the proportion of topics that are difficult to interpret. The model with 100 topics provides more topics that convincingly represent distinct grievances and preferences characteristic of social movements. Not all of the topics unambiguously correspond to substantive frames. Several elude interpretation altogether or are associated with spam we are unable to remove.

The three authors and one additional volunteer independently evaluated the 100 topics and chose the 34 we agreed were meaningful and substantive. We report results using only these topics, but this does not affect our results. We fit our model using stochastic variational inference until the likelihood of a 5% holdout set converged. We estimate asymmetric concentration priors using maximum likelihood estimation (M. D. Hoffman, Blei, Wang, & Paisley, 2013; Wallach, Mimno, & McCallum, 2009). Further details on text preprocessing, topic modeling, validation, and the top words for each topic are available in Appendix B.

1.4 Measures

We define petition success as the signature count for the petition (S). Signature count is the dependent variable for both hypotheses. It directly measures the number of participants in a collective action, but breaks from ecological studies of organizations which typically define success in terms of survival. Survival is not an appropriate definition of success for petitions, or for mobilizations more generally, because they do not pursue survival.

We could alternatively define success in terms of obtaining concessions from the target. However, this seems inconsistent with the fact that many petitions, even those with many signatures, cannot reasonably be expected to persuade the target. Defining petition success in terms of concessions would view the most signed petition on Change.org at the time we collected our data, a petition against the Yulin dog meat festival, as ineffective. However, the campaign against the dog meat festival has grown widespread and obtained significant media coverage.[10] Equating success with concessions neglects the recruitment function of petitions (Nall et al., 2017). A long signature list demonstrates that a petition reached and resonated with sympathizers. Furthermore, the signature list may reflect power to convert movement adherents (whose beliefs align with the movement) into constituents (who participate in the movement) (McCarthy & Zald, 1977). Our use of signature count as a measure of success follows other studies that analyze signatures as a measure of participation (Margetts et al.,

[10]"http://www.nytimes.com/2015/06/24/world/asia/dog-eaters-in-yulin-china-unbowed-by-global-derision.html"

2015; Hale et al., 2013; Jungherr & Jürgens, 2010).

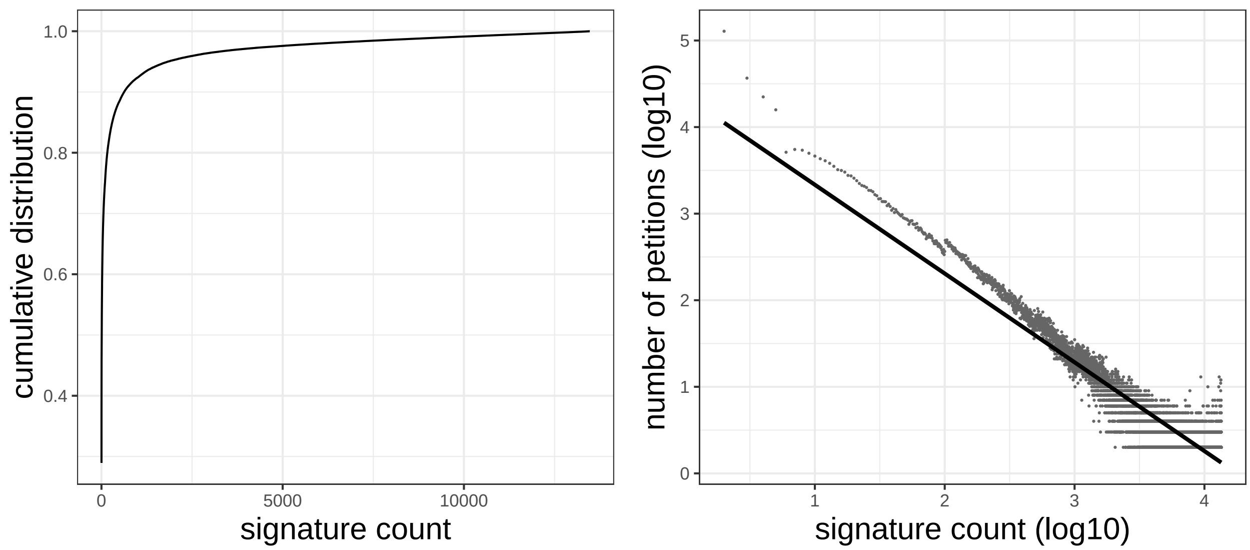


Figure 1.2: Signatures are distributed according to a power law. The left chart shows the proportion of petitions with fewer than a given number of signatures. The right chart shows a log-log plot.

As demonstrated in Figure 1.2, signatures are distributed among petitions according a to power-law such that few petitions gain the majority of signatures. As with the German and UK petition sites (Hale et al., 2013; Jungherr & Jürgens, 2010), the vast majority of petitions on Change.org fail to obtain many signatures. More than 52% of petitions fail to garner even 10 signatures and only the top 24% garner 100. The median number of signatures is 8 and the mean is 1,291.

Overlap density (D) is the independent variable for hypotheses **H1**. We identify our continuous measure of overlap density using the mixed membership assignments provided by the LDA topic model. We group petitions by the months in which they are created. We then assign topic memberships such that a petition belongs to a topic according to the probability that a word in the petition is drawn from the topic. Next we calculate topic-level density by aggregating membership in the topic over all petitions created in the time

$$D_i = \sum_{t \in \mathbf{T}} \left(m_{i,t} \sum_{j \in \mathbf{P_i}} m_{j,t} \right)$$

Equation 1.1: Measure of topic density for petition i. T is the set topics memberships and $m_{i,t}$ is the membership in a given topic t. $\mathbf{P_i}$ is the set of all petitions created in the same month as petition i. Finally, $m_{j,t}$ is the membership of petition j in the topic t.

$$F_i = \sum_{t \in \mathbf{T}} m_{i,t}^2$$

Equation 1.2: Measure of frame specialization for petition i. This is simply the Herfindahl-Hirschman index of the petition's topic memberships. This gives a measure of inequality in the range $[1/\mathbf{N}(\mathbf{T}), 1]$, where $\mathbf{N}(\mathbf{T})$ is the number of topics.

window. Finally we compute petition-level overlap density by taking the sum of the topic-level densities weighted by the petition's topic memberships. Equation 1.1 shows how we calculate overlap density.

The independent variable of **H2** is the interaction between frame specialization (F) and concentration (C). Frame specialization is simply the degree to which a petition focuses on few topics or on many. We measure frame specialization as the inequality of topic memberships. A petition about a few topics is more specialized than one about many. Equation 1.2 shows how we calculate frame specialization using the Herfindahl-Hirschman index (HHI), a commonly used inequality statistic,.

Concentration is defined by the inequality of the distribution of signatures. A topic where petitions have more equal levels of success is less concentrated than one with great levels of inequality. As with density, we group petitions by month and assign them to topics according

$$C_{i,t} = \sum_{j \in \mathbf{P}_i} m_{j,t} \left(\frac{S_j}{\sum_{k \in \mathbf{P}_i} m_{k,t} S_k} \right)^2$$

$$C_i = \sum_{t \in \mathbf{T}} m_{i,t} C_{i,t}$$

Equation 1.3: Measure of concentration for petition i. $C_{i,t}$ is the concentration of topic t for the month where petition i was created. The number of signatures obtained by petition j is represented by S_j.

to the posterior probability of the topic for the petition. We then measure the inequality of each topic, again using the HHI. Finally, we obtain our measure of the concentration of a petition's fundamental niche by taking the sum of concentration of each topic weighted by the petition's membership in each topic. Equation 1.3 shows how we compute concentration.

We include a control variable indicating whether or not a petition is created by an organization. Petition creators on Change.org may specify an organizational association and there are good reasons to think that organizational backing may influence petition success. Social movement organizations play an important role in mobilizing participants to sign petitions (McCarthy & Zald, 1977). Organizations play an important role in developing frames that resonate with publics (Snow, 2004). Therefore, organizations may have advantages at composing petitions. In our dataset, about 9% of petitions are created by organizations.

Change.org has changed significantly over time. Fundamental niches and political opportunities also shift over time. To account for this, we model random intercept variance terms for each monthly time window (Gelman & Hill, 2007).

Some topics are more popular than others so we include a continuous variable to account for membership each topic. This is the probability that a word in a petition is drawn from the topic as estimated by the topic model.

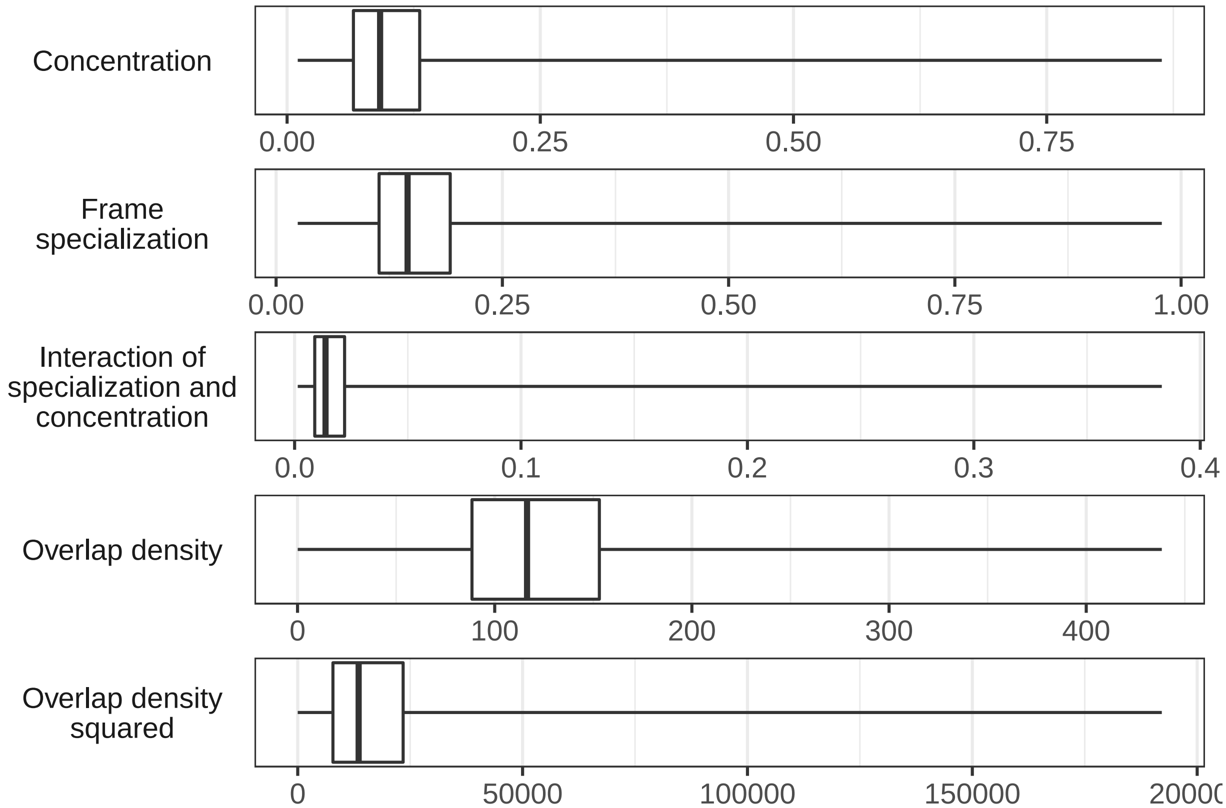


Figure 1.9: Box plots showing the range, median, upper, and lower quantiles of our indepedent variables.

Some petitions creators create several petitions. We account for this by modeling random intercept variance terms for each petition creator.

1.5 *Analytic Plan*

We test our hypotheses using multi-level negative binomial regression (Gelman & Hill, 2007). We use negative binomial regression because our dependent variable, signature count, is a highly skewed count variable. The mean of the log of signature count ($\overline{\log(S)} = 2.9$) is less than its variance ($\mathrm{Var}(\log(S)) = 5.8$)), so poisson regression would be over-dispersed. Equation 1.4 shows the full specification of our model. We estimate our model using maximum

$$y_i = B_0 + B_1 D_i + B_2 D_i^2 + B_3 C_i + B_4 F_i + B_5 C_i F_i + \mathbf{BZ} + \alpha_{p[i]} + \gamma_{m[i]} + \varepsilon_i$$

$$\alpha_p = \alpha_0 + \alpha_1 \overline{D}_p + \alpha_2 \overline{D_p^2} + \alpha_3 \overline{C}_p + \alpha_4 \overline{F}_p + \alpha_5 \overline{C_p F_p} + \varepsilon_p$$

$$\gamma_m = \gamma_0 + \gamma_1 \overline{D}_m + \gamma_2 \overline{D_m^2} + \gamma_3 \overline{C}_m + \gamma_4 \overline{F}_m + \gamma_5 \overline{C_m F_m} + \varepsilon_m$$

Equation 1.4: Specification of our model. D is overlap density. C is concentration. F is frame specialization. $\mathbf{Z}$ are controls for organizational backing and topic memberships. α_p is the intercept modeled for petition creators. γ_m is the intercept modeled for creation in a month. Bars over variables indicate group-level means.

likelihood estimation and the glmer.nb function in the R package lme4 (Bates, Mächler, Bolker, & Walker, 2015).

The first and second order terms for overlap density (D) are obviously highly correlated. We therefore orthogonalize the second order polynomial for overlap density when fitting our model. We normalize frame specialization (F), overlap density (D), and concentration (C) by subtracting the mean and dividing by the standard deviation. We reverse these transformations when interpreting model coefficients.

Models with group-level variance terms can produce biased coefficients when groups are correlated with predictor variables (Bafumi & Gelman, 2007). However this bias is easily corrected by including the group-level predictor means in the model. Therefore, we include creator-level and month-level means for the terms for density, focus, concentration, and the interaction between concentration and focus in the model.

1.6 Results

Table 1.2 shows the coefficients and standard errors for our analytic variables. We find evidence for **H1**, that overlap density has a ∩-shaped relationship success on average ($B = -30$, $SE = 3.6$). The negative sign on this coefficient indicates a ∩-shaped relationship between overlap density and signature counts. The adjusted R^2 statistic for the model is 0.16, indicating that our model is able to explain an estimated 16% of the variance in signature counts. For reference, Hagen et al. (2016) model petition signatures on Whitehouse.gov using a wide array of natural language processing techniques to analyze petition texts and obtain an adjusted R^2 statistic of 0.35. Much of their model's explanatory power comes from including the number of signatures obtained in the first 24 hours.

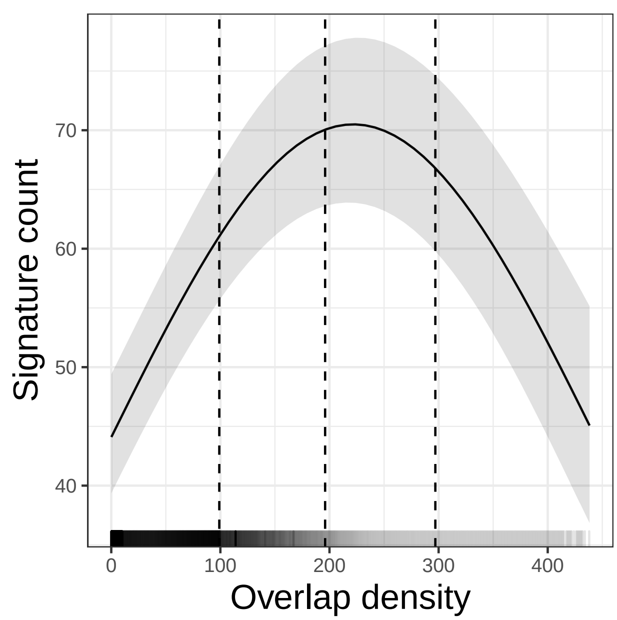


Figure 1.18: Marginal effects plots showing the model predicted relationship between topic denisty and median signature count with the 95% confidence interval. Lines at the bottom of the graph show that most petitions are in the lower range of overlap density. Vertical lines show the density values of the example petitions from Figure 1.1.

Intercept	4.30 (0.04)***
Density (D)	56.64 (6.01)***
Density squared (D^2)	-29.51 (3.62)***
Organization created	0.91 (0.01)***
Concentration (C)	4.65 (0.13)***
Frame specialization (F)	-5.66 (0.07)***
mean(Density) per month ($\overline{D_m}$)	0.09 (0.09)
mean(Density squared) per month ($\overline{D_m^2}$)	-0.20 (0.11)
mean(Concentration) per month ($\overline{F_m}$)	-0.34 (0.14)*
mean(Specialization) per month ($\overline{F_m}$)	-0.12 (0.11)
mean(Sepecialization:concentration) per month ($\overline{C_m F_m}$)	0.06 (0.15)
Concentration:frame specialization (CF)	-0.20 (0.61)
AIC	4805134.35
BIC	4805673.32
Log Likelihood	-2402518.18
Num. groups: creator.id	31477
Num. groups: month.block	108
Var: creator.id (Intercept)	4.03
Var: month.block (Intercept)	0.08

Table 1.2: Negative Binomial Regression. The unit of analysis is the petition. The dependent variable is signature count.

Since our model uses orthogonalized polynomials for overlap density, direct interpretation of the coefficient is not straightforward. To simply communicate our results, we provide Figure 1.18, a marginal effects plot showing the relationship between overlap density (D) and signature count (S) as predicted by the model. A great majority of Change.org petitions have levels of overlap density such that increases in overlap density are associated with increases in success. The optimal range of overlap density is near 220, the 96th percentile. A typical petition with overlap density in this range mobilizes about 21 more participants than a comparable petition with overlap density around 27, the 5th percentile. A small minority of petitions occupy niches where increasing density is associated with decreasing levels of participation, but the amount of decrease in participation at high density is comparable to the amount of decrease in participation at low density. A typical petition with overlap density around 440, the 99th percentile, mobilizes near 25 fewer participants than a comparable petition in the optimal range. Although the magnitudes of these relationships are small compared to the amount of variance in signature counts, they are substantial compared to the median signature count, which is 8.

Our finding that most petitions occupy niches where increasing density is associated with increasing levels of participation suggests that mutualism and shared opportunities are important factors for petition success. Many petition campaigns that struggle may be unusual or may not conform to petitioning norms. As overlap density (D) increases, the strength of mutualism declines and the strength of competition increases. About 17,684 petitions in our dataset have high overlap density and fare worse than comparable petitions of moderate overlap density. This suggests that petitions may compete over rival resources needed for success, but that competition only exceeds mutualism in cases of unusually high density.

We do not find evidence of resource partitioning between generalists and specialists on Change.org. We hypothesized that frame specialization (F) might be an advantageous strategy in more concentrated niches. This would have been indicated by a positive coefficient on the interaction term (CF) between concentration and frame specialization. However, we observe no such relationship ($B = -0.2$, $SE = 0.61$).

Our dataset contains many observations and sufficient variation so it is doubtful that a lack of statistical power is driving this result. Concentration (C) and frame specialization (F) are virtually uncorrelated (Spearman's $\rho = -0.0066$). We use post-hoc power analysis to evaluate confidence in rejecting **H2**. Given the observed variance in signature counts and our sample size, we would detect a relationship so weak as to account for the difference between 1000 and 1015 signatures (S) at significance level 0.01 with probability 0.96 ($t = 4.1$). We interpret this as evidence against the existence of strong resource partitioning between specialists and generalists.

We also observe a negative regression coefficient for frame specialization ($B = -5.7$, $SE = 0.073$). This validates our measure of frame specialization since it suggests that generalists have advantages on average. It also helps us rule out one explanation for the negative result for **H2**. If we had found a positive coefficient for specialization, that would be evidence that conditions favor specialists regardless of concentration. We observe a positive regression coefficient for concentration ($B = 4.7$, $SE = 0.13$). This is evidence that petitions on average do better in concentrated niches regardless of specialization. Conditions of concentration reflect opportunity, on average, but not advantages specific to specialists.

Finally, we note a substantial positive regression coefficient for organizational creation ($B = 0.91$, $SE = 0.013$). This indicates that petitions created by organizations obtain an average of about 2.5 times as many signatures as petitions not created by organizations. This is an intuitive result that confirms the importance of social movement organizations for mobilization in the context of e-tactics (Karpf, 2016, 2012).

Robustness checks and coefficients for topics are available in Appendix C.

1.7 Threats to Validity

Interpretation of our results should consider the limitations of our data and methods. Signature count is an incomplete measure of petition success. For example, not all signatures are created equal. The signature of a celebrity or official may be more influential other signatures. Some issues, like those of local or regional interest, might not require as many signatures to gain concessions, recruit allies, or create discursive opportunities. Despite these imperfections, signature count provides the most readily available piece of information associated with a petition's ability to gain concessions from targets and to influence potential signatories. Despite these limitations, we follow other studies of online petition success use signature count as a measure of success.

Topic modeling has inherent limitations. Topic modeling begins by reducing documents to "bags-of-words," a representation that ignores syntax and grammatical structure (Shannon, 1948). Interpreting the meaning of topic models can be challenging because topics reflect latent patterns that do not necessarily correspond with intuitive or theoretical distinctions. That said, in many case topics can be usefully interpreted by presenting the most probable words for each topic (DiMaggio et al., 2013). Although, interpreting specific topics is not a central concern of our study, the interpretability of topics is important for validating that topics learned by the model relate to identifiable social movements, contentious issues, preference sets, or claims.

Despite their limitations, topic models are useful as an objective criteria for grouping large numbers of documents. Methods relying on human judgment require establishing groups *a priori* and rely on inherently subjective interpretation of texts. Making reductionist decisions to render the complexity and nuance of the social world analyzable is done in all social science. We took steps to validate our topic models and to show that our results are not sensitive to technical decisions in the modeling process or to the choice of topics to include in our measures of overlap density, concentration, and specialization.

Our collected data contain a significant amount of spam. We take measures to remove this spam, but no spam filter is perfect. Some valid petitions are removed and some spam remains in our dataset. Our approach attempts to minimize bias from data loss at the cost of including some spam in our dataset.

Another threat stems from the fact that population ecology theorizes competition but does not test for it directly. We claim that our finding for **H1** that overlap density (D) has a ∩-shaped relationship with petition success implies the existence of forces of legitimacy and competition in online petitioning. However, we are not able to measure mechanisms of competition or legitimacy directly or to show causal direction.

Population ecology hypotheses are usually formulated in terms of survival rather than success. We think that our adaptation of organizational ecology to online petitions is resilient to this transformation. While survival is not a generally salient trait for petitions campaigns, it may be possible to operationalize survival for petitions according to the duration over which a petition collects signatures. Although, testing our hypotheses using survival analysis on such an outcome would be a useful robustness check, this is not possible using the publicly available data from Change.org.

A final threat is that our study is inherently limited in scope to online petitioning on Change.org. We only test population ecology theories for one kind of e-tactic. Our results may not even generalize to other petition platforms, let alone other kinds of e-tactics, or offline collective actions.

1.8 Discussion

Our work contributes to the study of collective action online by extending population ecology to study online mobilizations using e-tactics. We conclude that ecological factors can be a boon or a barrier to obtaining voice in digital media. From density dependence theory, we hypothesized a ∩-shaped relationship between overlap density and success. Our data and analysis support this prediction. For most of the petitions in our data, increasing overlap density is related to increased participation. This suggests that resources needed for creating petitions and mobilizing signers overlap. When these pools are small then few actions are initiated and they are less successful. However, where resources are abundant, more actions are initiated and they are more successful on average. Furthermore, successes that legitimize an e-tactic may be important so that organizers and activists come to see the tactic as useful and devote resources to disseminating petitions. Successful petitions can recruit constituents, grow sentiment pools, and provide information about what kinds of petitions can be effective.

Finally, exogenous events that drive awareness or political opportunity may in turn increase the resources available for both petition creation and petition signing.

The exemplar petitions from Figure 1.1 illustrate the curvilinear relationship between density and signature count. The petition about spousal immigration processing times is to the left of the maximum of the curve in Figure 1.18. This petition concerns legitimate issue, but one that is fairly niche among immigration petitions in its focus on Canadian visas for spouses. According to our model, if the amount of petitions related to Canadian immigration processing times was greater than it is, that would reflect increased opportunities for this petition. On the other hand, the petition against the common core in New York is to the right of the maximum. Education, and the common core in particular, is a very popular topic for Change.org petitions. Our results suggest that these petitions may compete with one another over signatures. Finally, the petition against the Yulin dog meat festival is near the maximum. Some of the most popular petitions on Change.org concern the Yulin dog meat festival and our model suggests that the number of such petitions is near that associated with optimal petition success.

We find that petitions are less successful on average when overlap density is very great. This is evidence for competition between e-tactic actions. Although the highly unequal distribution of participation suggests strong competition, we find that competition exceeds mutualism only for petitions with very high levels of overlap density. This suggests that adherence to petitioning norms, connection to established frames, and opportunity structures may be stronger barriers than competition for most petitions. Petition creators usually find more success when following the herd than when striking out alone.

That said, a sizable number, though a low proportion, of petitions in our dataset face levels of overlap density reflecting an environment so competitive that success is lower on average. This suggests that petition campaigns may compete over important factors like attention of constituents, space in activist media, access to gatekeepers and celebrities, and placement in the featured lists on petition platforms.

Although we predicted that frame specialization would be more advantageous for petitions in concentrated populations, our evidence suggests this does not occur. Online petitions do not seem to exhibit resource partitioning between generalists and specialists. Previous work has shown that social movement organizations, like firms in many industries, can specialize to avoid competition with large generalists (Soule & King, 2008). Our evidence suggests that online petitions cannot. What explains this surprising result? We offered this hypothesis assuming that conditions for specialists would promote the production of specialized petitions, but this assumption might be wrong. Perhaps opportunities for specialists still exist under concentration, but petition creators do not take advantage of them. Petition creators may not have any way to identify conditions favorable to specialization.

A second explanation for our result is that some resource partitioning occurs, but is not responsible for much of the variation in concentration we observe. Concentration that arises from preferential attachment, diffusion processes, or other mechanisms found in computer mediated communication may not confer advantages to specialists. Future work should attempt to identify resource partitioning while controlling for such mechanisms.

We offer a third explanation that is consistent with our observation that concentration is related to increased signature count regardless of specialization. Platforms may prevent resource partitioning by providing non-rival resources for creating and sustaining petitions, but not for mobilizing signatures. Under resource partitioning, specialists are able to find narrow niches where they have a competitive advantage, but generalists can only succeed by drawing from a diverse set of resources. By providing non-rival resources to all new petitions, platforms promote the production of generalists and specialists alike. Future work to disentangle these explanations should formally analyze dynamics between specialists and generalists in concentrated environments that provide limited non-rival resources.

Lack of resource partitioning in online petitioning has implications for how framing processes are structured, and for what kinds of frames develop and spread. Bennett (2012) finds that movements that partially substitute digitally connected and loosely structured activist networks for formal organizations often adopt "personalized action frames" as opposed to "collective action frames." Personalized action frames are interpretively flexible and can be easily appropriated by heterogeneous movements. Collective action frames are protest frames that emphasize collective identity. Perhaps movements utilizing collective action frames do not see e-tactics like online petitions as instrumental for achieving their aims. It may be that they are committed to relatively radical frames that are not amenable to online petitioning. Hersh and Schaffner (2015) make a similar observation about successful petitions on Whitehouse.gov with a style of framing they call "post-materialist particularism" that emphasizes grievances specific to particular circumstances and avoids ideological claims about wealth and resource distribution.

Particularistic and personalized frames tend to have low frame specialization in our operationalization. For example, the Trayvon Martin petition was instrumental in the development of the black lives matter movement but makes few explicit claims about race. These are limited to describing Trayvon Martin as "a young black man." With a frame specialization score of 0.19, this unambiguously successful petition is on the low end of the observed range of frame specialization. Instead of focusing on a broad ideological frame about racial injustice, this petition makes specific claims about a particular event. It demonstrates the injustice of Trayvon Martin's killing on the basis of his personality, character, and significance to his family. It also invokes conventional norms of criminal justice in how it criticizes the police and calls for the prosecution of George Zimmerman. Making a number of moral claims with broad credibility may have helped this petition mobilize support from diverse resource pools.

Petitions with ideological and collective identity frames have smaller fundamental niches. Without resource partitioning, specialists face competition both from other specialists and from generalists. Competition between petitions may undermine their collective goals—a fact that is in tension with online petitioning environments that encourage the production of many particularistic petitions. By providing non-rival resources for petition creation, petitioning platforms like Change.org may structure competition in ways that make it difficult for frame specialists to find a niche. Future designers of e-tactics should consider how to mitigate competition through affordances for cooperative framing work.

Finally, we showed how LDA can enable innovative applications of population ecology.

Assigning individuals to populations according to objective criteria removes the need for *a priori* assumptions about which groups are relevant and how they should be defined. As a mixed-membership model, LDA provides a realistic model of group membership for cases where boundaries between groups are not crisply defined. Finding evidence for density dependence validates that statistical analysis of symbol usage can capture overlaps in resource environments. We also showed how frame specialization, drawing resources from multiple pools aligned with frames, can be measured as inequality in topic membership (Soule & King, 2008; Carroll, 1985). Our application of LDA topic models to operationalize population ecology theory opens a range of possibilities for for exciting future work. Ecological dynamics in domains far beyond the realm of e-tactics can likely be measured through computational textual analysis. Future applications may include popular culture, social movement organizations, internet memes, and agenda setting in the news media.

Acknowledgments

Portions of this work were produced in collaboration with Benjamin Mako Hill and Aaron Shaw.

BIBLIOGRAPHY

Bafumi, J. & Gelman, A. (2007). *Fitting multilevel models when predictors and group effects correlate*. Working Paper. Rochester, NY. Retrieved June 16, 2017, from https://papers.ssrn.com/abstract=1010095

Bakker, T. P. & de Vreese, C. H. (2011). Good news for the future? Young people, Internet use, and political participation. *Communication Research*, *38*(4), 451–470. doi:10.1177/0093650210381738

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1). doi:10.18637/jss.v067.i01

Baum, J. A. C. & Shipilov, A. V. (2006). Ecological approaches to organizations. In *Sage Handbook for Organization Studies* (pp. 55–110). Rochester, NY: Sage. Retrieved August 27, 2016, from http://papers.ssrn.com/abstract=1017085

Baum, J. A. C. & Singh, J. V. (1994). Organizational niches and the dynamics of organizational founding. *Organization Science*, *5*(4), 483–501. doi:10.1287/orsc.5.4.483

Benford, R. D. & Snow, D. A. (2000). Framing processes and social movements: An overview and assessment. *Annual Review of Sociology*, *26*, 611–639. doi:10.1146/annurev.soc.26.1.611

Bennett, W. L. (2012). The personalization of politics political identity, social media, and changing patterns of participation. *The Annals of the American Academy of Political and Social Science*, *644*(1), 20–39. doi:10.1177/0002716212451428

Bennett, W. L. & Segerberg, A. (2013). *The logic of connective action: Digital media and the personalization of contentious politics*. New York, New York: Cambridge University Press.

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. *The Journal of Machine Learning Research*, *3*, 993–1022. Retrieved December 3, 2015, from http://dl.acm.org/citation.cfm?id=944937

Carpenter, D. & Moore, C. D. (2014). When canvassers became activists: Antislavery petitioning and the political mobilization of American women. *American Political Science Review*, *108*, 479–498. doi:10.1017/S000305541400029X

Carpenter, D. & Schneer, B. (2015). Party formation through petitions: The Whigs and the Bank War of 1832–1834. *Studies in American Political Development*, *29*, 213–234. doi:10.1017/S0898588X15000073

Carroll, G. R. (1984). Organizational ecology. *Annual review of Sociology*, *10*(1), 71–93. doi:10.1146/annurev.so.10.080184.000443

Carroll, G. R. (1985). Concentration and specialization: Dynamics of niche width in populations of organizations. *American Journal of Sociology*, *90*(6), 1262–1283. doi:10.1086/228210

Carroll, G. R. & Hannan, M. T. (1989). Density dependence in the evolution of populations of newspaper organizations. *American Sociological Review*, *54*(4), 524. doi:10.2307/2095875

Carroll, G. R. & Swaminathan, A. (2000). Why the microbrewery movement? Organizational dynamics of resource partitioning in the U.S. brewing industry. *American Journal of Sociology*, *106*(3), 715–762. doi:10.1086/318962

Cheng, J., Adamic, L., Dow, P. A., Kleinberg, J. M., & Leskovec, J. (2014). Can cascades be predicted? In *Proceedings of the 23rd International Conference on World Wide Web* (pp. 925–936). WWW '14. New York, NY, USA: ACM. doi:10.1145/2566486.2567997

Cress, D. M. & Snow, D. A. (2000). The outcomes of homeless mobilization: The influence of organization, disruption, political mediation, and framing. *American Journal of Sociology*, *105*(4), 1063–1104.

DiMaggio, P., Hargittai, E., Neuman, W. R., & Robinson, J. P. (2001). Social implications of the Internet. *Annual Review of Sociology*, *27*(1), 307–336. doi:10.1146/annurev.soc.27.1.307

DiMaggio, P., Nag, M., & Blei, D. (2013). Exploiting affinities between topic modeling and the sociological perspective on culture: Application to newspaper coverage of U.S. government arts funding. *Poetics*, *41*(6), 570–606. doi:10.1016/j.poetic.2013.08.004

Dobrev, S., Kim, T., & Hannan, M. (2001). Dynamics of niche width and resource partitioning. *American Journal of Sociology*, *106*(5), 1299–1337. doi:10.1086/320821

Earl, J. (2006). Pursuing social change online: The use of four protest tactics on the Internet. *Social Science Computer Review*, *24*(3), 362–377. doi:10.1177/0894439305284627

Earl, J. & Kimport, K. (2009). Movement societies and digital protest: Fan activism and other nonpolitical protest online. *Sociological Theory*, *27*(3), 220–243. doi:10.1111/j.1467-9558.2009.01346.x

Earl, J. & Kimport, K. (2011). *Digitally enabled social change: Activism in the Internet age*. Cambridge, MA: MIT Press.

Earl, J. & Schussman, A. (2008). Contesting cultural control: Youth culture and online petitioning. In *Civic life online: Learning how digital media can engage youth* (pp. 71–95). The John D. and Catherine T. MacArthur Foundation series on digital media and learning. Cambridge, MA: MIT Press.

Farrell, H. (2012). The consequences of the Internet for politics. *Annual Review of Political Science*, *15*(1), 35–52. doi:10.1146/annurev-polisci-030810-110815

Gamson, W. A. & Wolfsfeld, G. (1993). Movements and media as inacting systems. *Annals of the American Acadamy of Political and Social Science*.

Garrett, R. K. (2006). Protest in an information society: A review of literature on social movements and new ICTs. *Information, Communication & Society*, *9*(2), 202–224. doi:10.1080/13691180600630773

Gelman, A. & Hill, J. (2007). *Data analysis using regression and multilevel / hierarchical models*. New York, NY: Cambridge University Press.

Gillespie, T. (2010). The politics of 'platforms'. *New Media & Society*, *12*(3), 347–364. doi:10.1177/1461444809342738

Goldhaber, M. H. (1997). The attention economy and the net. *First Monday*, *2*(4). doi:10.5210/fm.v2i4.519

Graeff, E., Stempeck, M., & Zuckerman, E. (2014). The battle for 'Trayvon Martin': Mapping a media controversy online and off-line. *First Monday*, *19*(2). doi:10.5210/fm.v19i2.4947

Hagen, L., Harrison, T. M., Uzuner, Ö., May, W., Fake, T., & Katragadda, S. (2016). E-petition popularity: Do linguistic and semantic factors matter? *Government Information Quarterly*. doi:10.1016/j.giq.2016.07.006

Hale, S. A., Margetts, H., & Yasseri, T. (2013). Petition growth and success rates on the UK No. 10 Downing Street website. In *Proceedings of the 5th Annual ACM Web Science Conference* (pp. 132–138). ACM. doi:10.1145/2464464.2464518

Hannan, M. T. & Freeman, J. (1977). The population ecology of organizations. *American Journal of Sociology*, *82*(5), 929–964. doi:10.2307/2777807

Hannan, M. T. & Freeman, J. (1984). Structural inertia and organizational change. *American Sociological Review*, *49*(2), 149. 07708 bibtex*:HannanStructuralInertiaOrganizational1984. doi:10.2307/2095567

Hannan, M. T. & Freeman, J. (1989). *Organizational Ecology* (1st ed.). Cambridge, MA: Harvard University Press.

Hargittai, E. & Shaw, A. (2013). Digitally savvy citizenship: The role of Internet skills and engagement in young adults' political participation around the 2008 presidential election. *Journal of Broadcasting & Electronic Media*, *57*(2), 115–134. doi:10.1080/08838151.2013.787079

Hargittai, E. & Shaw, A. (2015). Mind the skills gap: The role of internet know-how and gender in differentiated contributions to Wikipedia. *Information, Communication & Society*, *18*(4), 424–442. doi:10.1080/1369118X.2014.957711

Hersh, E. D. & Schaffner, B. F. (2015). *Post-materialist particularism: What petitions can tell us about biases in the policy agenda*. Working Paper. New Haven, Connecticut. Retrieved July 13, 2016, from http://www.eitanhersh.com/uploads/7/9/7/5/7975685/petitions_v4_3.pdf

Hey, J. (2006). On the failure of modern species concepts. *Trends in Ecology & Evolution*, *21*(8), 447–450. doi:10.1016/j.tree.2006.05.011

Hoffman, M. D., Blei, D. M., Wang, C., & Paisley, J. (2013). Stochastic variational inference. *The Journal of Machine Learning Research*, *14*(1), 1303–1347. Retrieved May 14, 2016, from http://dl.acm.org/citation.cfm?id=2502622

Hoffman, M., Bach, F. R., & Blei, D. M. (2010). Online learning for latent dirichlet allocation. In *Advances in neural information processing systems* (pp. 856–864). Retrieved May 16, 2016, from http://papers.nips.cc/paper/3902-online-learning-for-latentdirichlet-allocation!

Hopkins, D. J. & King, G. (2010). A method of automated nonparametric content analysis for social science. *American Journal of Political Science*, *54*(1), 229–247. doi:10.1111/j.1540-5907.2009.00428.x

Huang, S.-W., Suh, M. (, Hill, B. M., & Hsieh, G. (2015). How activists are both born and made: An analysis of users on Change.org. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (pp. 211–220). CHI '15. New York, NY: ACM. doi:10.1145/2702123.2702559

Inglehart, R. & Catterberg, G. (2002). Trends in political action: The developmental trend and the post-honeymoon decline. *International Journal of Comparative Sociology*, *43*, 300–316. doi:10.1177/002071520204300305

Johnson, S. L., Faraj, S., & Kudaravalli, S. (2014). Emergence of power laws in online communities: The role of social mechanisms and preferential attachment. *Management Information Systems Quarterly*, *38*(3), 795–808. Retrieved April 26, 2017, from http://aisel.aisnet.org/cgi/viewcontent.cgi?article=3193&context=misq

Jungherr, A. & Jürgens, P. (2010). The political click: Political participation through e-petitions in Germany. *Policy & Internet*, *2*(4), 131–165. doi:10.2202/1944-2866.1084

Karpf, D. (2012). *The MoveOn effect: The unexpected transformation of American political advocacy*. New York, NY: Oxford University Press.

Karpf, D. (2016). *Analytic Activism*. New York, NY: Oxford University Press.

Klandermans, B., van Stekelenburg, J., Damen, M.-L., van Troost, D., & van Leeuwen, A. (2014). Mobilization without organization: The case of unaffiliated demonstrators. *European Sociological Review*, *30*(6), 702–716. doi:10.1093/esr/jcu068

Loper, E. & Bird, S. (2002). NLTK: The natural language toolkit. In *Proceedings of the ACL-02 Workshop on Effective Tools and Methodologies for Teaching Natural Language Processing and Computational Linguistics - Volume 1* (pp. 63–70). ETMTNLP '02. Stroudsburg, PA, USA: Association for Computational Linguistics. doi:10.3115/1118108.1118117

Margetts, H., John, P., Hale, S., & Yasseri, T. (2015). *Political turbulence: How social media shape collective action*. Princeton, NJ: Princeton University Press.

McCarthy, J. D. & Zald, M. N. (1977). Resource mobilization and social movements: A partial theory. *The American Journal of Sociology*, *82*(6), 1212–1241.

McCarthy, J. D. & Zald, M. N. (2001). The enduring vitality of the resource mobilization theory of social movements. In J. H. Turner (Ed.), *Handbook of sociological theory* (pp. 533–565). Handbooks of Sociology and Social Research. Boston, MA: Springer. doi:10.1007/0-387-36274-6_25

McPherson, M. (1983). An ecology of affiliation. *American Sociological Review*, *48*(4), 519–532. doi:10.2307/2117719

Minkoff, D. C. (1995). *Organizing for equality: The evolution of women's and racial-ethnic organizations in America, 1955-1985*. Arnold and Caroline Rose book series of the American Sociological Association. New Brunswick, NJ: Rutgers University Press.

Minkoff, D. C. (1997). The sequencing of social movements. *American Sociological Review*, *62*(5), 779–799. doi:10.2307/2657360

Nall, C., Schneer, B., & Carpenter, D. P. (2017). Paths of recruitment: Rational social prospecting in petition canvassing. *American Journal of Political Science*. doi:10.1111/ajps.12305

Olson, M. (1965). *The logic of collective action: Public goods and the theory of groups*. Cambridge, MA: Harvard University Press.

Olzak, S. & Ryo, E. (2007). Organizational diversity, vitality and outcomes in the civil rights movement. *Social Forces*, *85*(4), 1561–1591. doi:10.1353/sof.2007.0076

Olzak, S. & Uhrig, S. C. N. (2001). The ecology of tactical overlap. *American Sociological Review*, *66*(5), 694. doi:10.2307/3088954

Řehůřek, R. & Sojka, P. (2010, May, 22). Software framework for topic modelling with large corpora. In *Proceedings of the LREC 2010 Workshop on New Challenges for NLP Frameworks* (pp. 45–50). Valletta, Malta: ELRA. Retrieved from http://is.muni.cz/publication/884893/en

Rolfe, B. (2005). Building an electronic repertoire of contention. *Social Movement Studies*, *4*(1), 65–74. doi:10.1080/147428305000051945

Shannon, C., E. (1948). A mathematical theory of communication. *The Bell System Technical Journal*, *27*, 379–423, 623–656. doi:10.1145/584091.584093

Snow, D. A. (2004). Framing processes, ideology, and discursive fields. In D. A. Snow, S. A. Soule, & H. Kriesi (Eds.), *The Blackwell Companion to Social Movements* (pp. 380–412). Maiden, MA: Blackwell Publishing Ltd. doi:10.1002/9780470999103.ch17

Snow, D. A., Rochford, E. B., Worden, S. K., & Benford, R. D. (1986). Frame alignment processes, micromobilization, and movement participation. *American Sociological Review*, *51*(4), 464–481. doi:10.2307/2095581

Soule, S. A. & King, B. G. (2008). Competition and resource partitioning in three social movement industries. *The American Journal of Sociology*, *113*(6), 1568–1610. doi:10.1086/587152

Swaminathan, A. (2001). Resource partitioning and the evolution of specialist organizations: The role of location and identity in the U.S. wine industry. *Academy of Management Journal*, *44*(6), 1169–1185. doi:10.2307/3069395

Tarrow, S. G. (2011). *Power in movement: Social movements and contentious politics*. Cambridge; NY: Cambridge University Press.

Tilly, C. (1995). *Popular contention in Great Britain, 1758-1834*. Cambridge, MA: Harvard University Press.

Tufekci, Z. (2013). "Not this one": Social movements, the attention economy, and microcelebrity networked activism. *American Behavioral Scientist*, *57*(7), 848–870. doi:10.1177/0002764213479369

van de Rijt, A., Akin, I. A., Willer, R., & Feinberg, M. (2016). Success-breeds-success in collective political behavior: Evidence from a field experiment. *Sociological Science*, *3*, 940–950. doi:10.15195/v3.a41

van de Rijt, A., Kang, S. M., Restivo, M., & Patil, A. (2014). Field experiments of success-breeds-success dynamics. *Proceedings of the National Academy of Sciences*, *111*(19), 6934–6939. doi:10.1073/pnas.1316836111

Vasi, I. B., Walker, E. T., Johnson, J. S., & Tan, H. F. (2015). "No fracking way!" Documentary film, discursive opportunity, and local opposition against hydraulic fracturing in the United States, 2010 to 2013. *American Sociological Review*, *80*(5), 934–959. doi:10.1177/0003122415598534

Wallach, H. M., Mimno, D. M., & McCallum, A. (2009). Rethinking LDA: Why priors matter. In *Advances in neural information processing systems* (pp. 1973–1981). Retrieved February 17, 2016, from http://papers.nips.cc/paper/3854-rethinking-lda-why-priors-matter

Yan, X., Guo, J., Lan, Y., & Cheng, X. (2013). A biterm topic model for short texts. In *Proceedings of the 22nd International Conference on World Wide Web* (pp. 1445–1456). New York, NY: ACM. doi:10.1145/2488388.2488514

Yasseri, T., Hale, S. A., & Margetts, H. (2013). *Modeling the rise in Internet-based petitions*. arXiv preprint. Retrieved October 26, 2015, from http://arxiv.org/abs/1308.0239

Zald, M. N. & McCarthy, J. D. (1980). Social movement industries: Cooperation and conclict amongst social movement organizations. In L. Kriesberg (Ed.), *Research in Social Movements, Conflict, and Change* (Vol. 3, pp. 1–20). Greenwich, CT: JAI Press. Retrieved June 21, 2017, from https://deepblue.lib.umich.edu/bitstream/handle/2027.42/50975/201.pdf?sequence=1

Zukin, C. (2006). *A new engagement? Political participation, civic life, and the changing American citizen*. New York, NY: Oxford University Press.

Appendix A

DATA COLLECTION AND VALIDATION APPENDIX

We downloaded a complete set of petitions from the Change.org application programming interface (API) using a script that ran from December 22nd 2015 to December 24th 2015. Some petitions were available on the API, but had not been published. We only include published petitions. We removed one petition with an invalid end date.

We identified translated interfaces to the same petition campaign on the basis that such petitions have different IDs but identical signature counts and creators. They share another identifier that was not available in the API, but was embedded in HTML of the petitions' web page. We obtained these identifiers using a web scraper and used them to merge these petitions.

We observe that the spam petitions displays search engine optimization tactics like spam frequently found in comments on Wordpress blogs and other websites that accept user generated content. We use the Akismet[1] API to identify spam petitions. Akismet is a widely used service mainly used to detect spam in comments on Wordpress blogs. The accuracy of spam detection algorithms is platform sensitive so we necessarily validate that this spam filter is effective on our data. On a set of 100 randomly selected English petitions with more than 480 characters coded which we coded as spam or not-spam, the API correctly identifies 20 of 25 spam petitions with 1 false positive. Although Akismet is not a perfect spam filter, it effectively reduces the proportion of spam in our dataset from about 25% to about 5% with little risk of removing a significant portion of valid petitions.

Of the 778,764 English petitions with more than 480 characters. The Akismet algorithm flags 188,935 petitions, all of which we remove. We then identify Change.org users that created at least 1 petition that Akismet flags as spam and remove an additional 41,766 petitions by these users. Based on our set of 100 coded petitions, we estimate that 25% of petitions were spam. After the above spam removal steps we removed 230,701 petitions. This is in line with the expected prevalence of spam. We believe our results are robust to bias introduced by the spam filter. We replicated our results on a set of petitions collected in 2015, all of which were linked from tweets. This dataset while smaller and arbitrarily sampled has a much lower level of spam.

[1] https://akismet.com

Appendix B

TOPIC MODELING DETAILS

All Change.org petitions have two relevant textual fields. The first is an overview that describes the petition and its purpose to potential signers. The second is the prayer to the targets of the petition. We represent petitions by concatenating these two texts into a single document. To prepare petition prayers for topic modeling we first remove punctuation, numbers, and stop words from petition texts. We then lemmatize texts using the WordNet lemmatizer from the *nltk*[1] python package (Loper & Bird, 2002) and remove terms which occur in fewer than 50 petitions. Next we remove terms that occur less in less than 2 documents. We then remove petitions that contain no terms after this process. We use the implementation of LDA, the simplest topic model, from the *gensim*[2] python package (Řehůřek & Sojka, 2010, May, 22). We use stochastic variational inference to algorithm to efficiently optimize the LDA model (M. D. Hoffman et al., 2013; M. Hoffman, Bach, & Blei, 2010) on our large corpus.

Variational inference is an alternative to markov-chain-monte-carlo (mcmc) methods like Gibbs sampling. Rather than sampling from the posterior, which can is expensive when data is large and can be slow to converge, variational inference uses convex optimization to fit an approximate model. *Stochastic* variational inference updates the model in steps based on samples from the data. This makes it possible toobtain a good solution in few passes over a large dataset. We use a stepsize of 0.95. Figure B.1 shows that the posterior likelihood of the 50% holdout set converges on 100 passes over the data.

Fitting an LDA model requires selecting a number of hyperparameters. The most important hyperparameter is the number of topics. We discussed our choice for this parameter in the main body of the paper. Another important decision for LDA topic modeling is choosing values for the concentration parameters, α and β (Wallach et al., 2009). The document-topic concentration parameter, α, sets the probability that a given document will draw from one topic or many. Similarly, β, the word-topic concentration parameter sets the

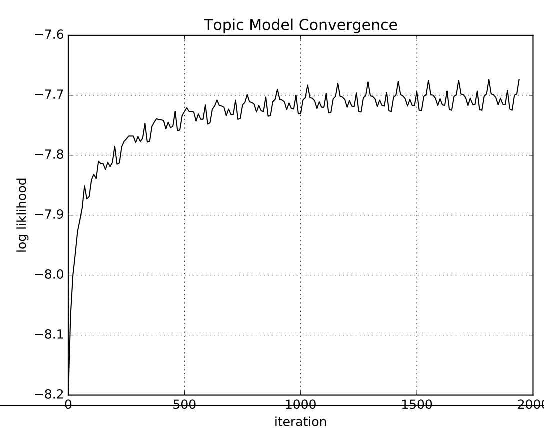


Figure B.1: LDA Convergence

[1] http://www.nltk.org/

[2] https://radimrehurek.com/gensim/

probability that a given word will be drawn from one topic or many. Analysts often set these parameters to fixed values for all topics. Doing this is called a symmetric prior. However, we are interested in the degree to which petitions words are drawn from a few topics or many. Therefore, using a symmetric priors would determine our results. Fortunately, it is possible to use asymmetric priors for α and β. Asymmetric priors assign a different concentration to each topic. We use maximum liklihood estimation to optimize asymmetric α and β. This technique can also increase the overall quality of topics (Wallach et al., 2009).

Other hyper-parameters concern optimizing the model using stochastic variational inference. These inference parameters include the learning decay, the batch size, the learning offset. These also include the seed of the pseudo-random number generator used to shuffle documents for stochastic inference. Fortunately, unlike choices of number of topics or topic concentration. Choices for these inference parameters do not imply any assumptions about the model. Instead they influence the rate at which the inference algorithm converges. Though they do not have a substantial influence on our results, choices for these parameters can subtly influence the quality of topics. To choose these parameters. We fit 96 LDA models. Each has different permutation of the learning parameters. We select the one with the greatest log-liklihood computed on a 10% hold-out set. Table B.1 shows the values of the learning parameters selected model.

parameter	value
learning offset	1024
learning decay	0.5
batch size	4096

Table B.1: Learning Parameters

Appendix C
REGRESSION TABLE

Table C.1: Table of fitted Negative Binomail Regression Models. The petition is the unit of analysis. The dependent variable is log signature count

Intercept	4.30 (0.04)***
density^1	56.64 (6.01)***
density^2	−29.51 (3.62)***
organization created	0.91 (0.01)***
topic 32	−1.40 (0.09)***
topic 36	−2.12 (0.07)***
topic 38	2.14 (0.07)***
topic 40	0.46 (0.08)***
topic 57	1.18 (0.17)***
topic 61	−2.50 (0.10)***
topic 73	−0.90 (0.05)***
topic 77	3.62 (0.06)***
topic 97	0.21 (0.08)*
topic 2	2.31 (0.12)***
topic 4	−0.75 (0.06)***
topic 17	−0.59 (0.09)***
topic 22	−0.06 (0.13)
topic 23	0.39 (0.15)**
topic 24	−1.31 (0.07)***
topic 25	−1.45 (0.07)***
topic 26	−0.19 (0.09)*
topic 27	−1.30 (0.07)***

topic 35	-2.81 (0.07)***
topic 41	-3.51 (0.24)***
topic 42	-1.26 (0.11)***
topic 49	1.73 (0.08)***
topic 51	-1.57 (0.08)***
topic 52	-7.64 (0.11)***
topic 56	-0.58 (0.06)***
topic 58	2.00 (0.06)***
topic 65	-2.41 (0.10)***
topic 66	-0.25 (0.15)
topic 74	2.29 (0.10)***
topic 81	5.41 (0.06)***
topic 87	0.37 (0.08)***
topic 89	1.23 (0.06)***
topic 94	-3.11 (0.16)***
topic 98	0.78 (0.12)***
concentration	4.65 (0.13)***
topic focus	-5.66 (0.07)***
mean.density.month	0.09 (0.09)
mean.density.sq.month	-0.20 (0.11)
mean.concentration.month	-0.34 (0.14)*
mean.focus.month	-0.12 (0.11)
mean.focus.by.conc.month	0.06 (0.15)
concentration:topic focus	-0.20 (0.61)
AIC	4805134.35
BIC	4805673.32
Log Likelihood	-2402518.18
Num. groups: creator.id	31477
Num. groups: month.block	108
Var: creator.id (Intercept)	4.03
Var: month.block (Intercept)	0.08

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

Appendix D

TOPIC TERMS APPENDIX

Topic:0	Topic:1	Topic:2	Topic:3
senior	community	cover	military
lawyer	support	fraud	war
aged	program	agent	government
65	help	collection	ha
lincoln	organization	digital	people

Topic:4	Topic:5	Topic:6	Topic:7
horse	right	image	street
christmas	human	beauty	emergency
milk	freedom	distribution	wall
cow	people	republic	911
slaughter	religious	pornography	ex

Topic:8	Topic:9	Topic:10	Topic:11
line	team	’s	girl
cell	v	’	young
base	football	florida	boy
tower	live	stand	poor
foot	game	ground	rich

Topic:12	Topic:13	Topic:14	Topic:15
hall	health	village	body
shark	study	fly	code
russia	use	boundary	wear
daniel	disease	fighter	size
russian	research	edge	uniform

Topic:16	Topic:17	Topic:18	Topic:19
6	home	tree	hill
beat	house	change	ride
premier	property	global	2006
queensland	housing	climate	rider
quarter	family	earth	silent

Topic:20	Topic:21	Topic:22	Topic:23
dog	st	war	m
animal	island	church	b
pet	little	john	c
shelter	charity	death	d
cat	donate	catholic	s

Topic:24	Topic:25	Topic:26	Topic:27
child	year	marriage	music
family	old	gay	song
parent	age	discrimination	band
kid	2012	gender	radio
mother	2013	equal	paul

Topic:28	Topic:29	Topic:30	Topic:31
farm	state	service	facebook
farmer	law	internet	page
bridge	governor	phone	website
factory	texas	customer	site
operation	north	data	post

Topic:32	Topic:33	Topic:34	Topic:35
american	european	company	energy
united	battle	oil	power
state	europe	industry	air
u	ball	corporation	green
president	france	price	plant

Topic:36	Topic:37	Topic:38	Topic:39
health	medium	california	canada
medical	news	san	disability
care	story	disney	canadian
patient	report	brown	disabled
marijuana	article	chicago	special

Topic:40	Topic:41	Topic:42	Topic:43
time	crime	member	water
change	justice	board	land
current	court	director	project
issue	prison	executive	river
policy	judge	mark	lake

Topic:44	Topic:45	Topic:46	Topic:47
government	dr	road	movie
uk	king	car	online
australia	martin	vehicle	watch
australian	joe	bus	download
country	jr	traffic	hd

Topic:48	Topic:49	Topic:50	Topic:51
food	book	event	abuse
product	robert	national	victim
store	grand	honor	sexual
sale	jones	award	sex
market	amazon	race	bullying

Topic:52	Topic:53	Topic:54	Topic:55
office	ha	art	law
washington	public	history	right
james	fact	michael	act
dc	action	beach	state
box	issue	culture	court

Topic:56	Topic:57	Topic:58	Topic:59
night	drug	firm	wa
sleep	super	reward	did
dark	addiction	recall	ha
inside	gang	associate	said
direction	happiness	ed	told

Topic:60	Topic:61	Topic:62	Topic:63
money	day	apple	people
pay	hour	london	n't
bank	time	chicken	make
fee	week	garden	want
business	ticket	pro	u

Topic:64	Topic:65	Topic:66	Topic:67
loan	violence	police	video
financial	domestic	officer	youtube
private	role	black	google
debt	model	gun	homeless
aid	pledge	law	dead

Topic:68	Topic:69	Topic:70	Topic:71
jersey	fan	kill	–
autism	tv	african	open
chris	series	bear	•
jim	channel	killing	close
nj	character	wolf	it's

Topic:72	Topic:73	Topic:74	Topic:75
sam	la	white	point
dean	y	red	restaurant
matt	e	blood	table
cafe	que	·	scale
demon	en	color	establishment

Topic:76	Topic:77	Topic:78	Topic:79
woman	india	2	park
men	indian	1	space
baby	government	3	area
female	minister	4	center
male	prime	5	facility

Topic:80	Topic:81	Topic:82	Topic:83
world	ban	mr	new
country	waste	strike	york
nation	plastic	na	michigan
international	bag	resign	ny
million	station	richard	city

Topic:84	Topic:85	Topic:86	Topic:87
vote	credit	club	download
party	card	language	free
election	information	sport	pc
political	license	field	game
candidate	application	english	file

Topic:88	Topic:89	Topic:90	Topic:91
school	pool	agreement	job
student	dolphin	deal	work
education	whale	nuclear	worker
university	sea	japan	employee
college	marine	zone	pay

Topic:92	Topic:93	Topic:94	Topic:95
city	adult	animal	game
council	single	stop	play
community	damage	cruelty	player
county	repair	human	release
resident	dating	cruel	nintendo

Topic:96	Topic:97	Topic:98	Topic:99
tax	life	veteran	specie
million	god	duty	forest
cost	man	r	fish
year	love	toy	natural
cut	heart	va	wildlife