HARVARD | BUSINESS | SCHOOL

Ideology and Composition Among an Online Crowd: Evidence from Wikipedians

Shane Greenstein
Grace Gu
Feng Zhu

Working Paper 17-028

Ideology and Composition Among an Online Crowd: Evidence from Wikipedians

Shane Greenstein
Harvard Business School

Grace Gu
Boston College

Feng Zhu
Harvard Business School

Working Paper 17-028

Funding for this research was provided in part by Harvard Business School.

Ideology and Composition among an Online Crowd: Evidence from Wikipedians*

Shane Greenstein
Harvard Business School
Boston, MA 02163
sgreenstein@hbs.edu

Grace Gu
Boston College
Chestnut Hill, MA 02467
graceyuangu@gmail.com

Feng Zhu
Harvard Business School
Boston, MA 02163
fzhu@hbs.edu

Abstract

Online communities bring together participants from diverse backgrounds and often face challenges in aggregating their opinions. We infer lessons from the experience of individual contributors to Wikipedia articles about U.S. politics. We identify two factors that cause a tendency toward moderation in collective opinion: either biased contributors contribute less, which shifts the composition of participants, or biased contributors moderate their own views. Our findings show that shifts in the composition of participants account for 80% to 90% of the moderation in content. Contributors tend to contribute to articles with slants that are opposite of their own views. Evidence suggests that encountering extreme contributors with an opposite slant plays an important role in triggering the composition shift and changing views. These findings suggest that collective intelligence becomes more trustworthy when mechanisms encourage confrontation between distinct viewpoints. They also suggest, cautiously, that managers who aspire to produce content "from all sides" should let the most biased contributors leave the collective conversation if they can be replaced with more moderate voices.

* We thank Jana Gallus, Marco Iansiti, Gerald Kane, Karim Lakhani, Abhishek Nagaraj, Frank Nagle, Michael Norton, Huaxia Rui, Brian Silverman, anonymous reviewers, and seminar participants at many universities, the INFORMS Annual Meeting, the Conference on Open and User Innovation, the Academy of Management Annual Meeting, the Workshop on Information Systems and Economics, and the Interntional Industrial Organizations Conference . We thank Justin Ng and John Sheridan of HBS Research Computing Services for their research assistance. We also thank Alicia Shems and Kate Adams for their editorial assistance. We gratefully acknowledge financial support from the Division of Research of the Harvard Business School.

1. Introduction

The growth of online communities that blur the boundaries between readers and writers has upended our understanding on the generation and consumption of online content. Online communities bring together participants from disparate traditions, with different methods of expression, different cultural and historical opinion foundations, and, potentially, different facts (e.g., Arazy et al. 2011; Ransbotham and Kane 2011; Kane et al. 2014; Gallus 2016).

Despite the diversity of opinions, and sometimes due to it, the composition and opinions of participants evolve as they interact with alternative content and points of view other than their own. A crowd's opinion reflects the aggregation of participants' opinions. Hence, at any point in time and over time, a crowd renders its opinion accordingly. Although a number of studies have sought to understand the bias in a crowd's opinion or, broadly, the limit of collective intelligence (e.g., Galton 1907; Shankland 2003; Antweiler and Frank 2004; Lemos 2004; Surowiecki 2004; Giles 2005; Chesney 2006; Rajagopalan et al. 2011; Mollick and Nanda 2015; Greenstein and Zhu 2018), little research has examined the manner by which the behavior of participants influences or is influenced by the bias of a crowd. Understanding this question helps shed lights on whether managers of such communities should intervene. It also informs managers on how to design effective rules and algorithms to steer interactions in ways that reduce a crowd's bias.

This study seeks to answer this question by measuring participants' actions and viewpoints. Our evidence comes from one of the longest-running online conversations on Wikipedia. We trace all the participation on 66,389 English language articles about US political topics from the start of Wikipedia in 2001 to January 2011. These articles received more than 10 million edits from 2,887,140 unique contributors. We follow Greenstein and Zhu (2018)'s approach to employ an adaptation of the method developed by Gentzkow and Shapiro (2010) for rating newspaper editorials. In these ratings, *slant* denotes the degree of opinion along a continuous yardstick. It can take on extreme degrees of red (e.g., Republican), extreme degrees of blue (e.g., Democrat), and all the shades of purple in between. *Bias* is the absolute value from the zero point of this yardstick and thus denotes the strength of the opinion. We then use these measures to characterize the evolution of the bias and slant of each participant opinion over time. We also gain insights into which experience prompts biased participants to stay or leave and which experiences induce them to maintain or change their opinion and, consequently, how such changes affect a crowd's bias.

Contributor behavior on Wikipedia tends to move toward less biased and less segregated conversations on most topics, consistent with Wikipedia's aspiration to present a neutral point of view (NPOV) in its content, which is succinctly summarized as "Assert facts, including facts about opinions—but don't assert opinions themselves."[1] Although considerable heterogeneity is found in behaviors, more Wikipedia contributors participate in unsegregated than segregated conversations. For example, a slanted contributor is on average 8% more likely to edit an article with an opposite slant than an article with the same slant. This tendency is pervasive.

We find that biased contributors moderate their own views as they encounter extreme content of the opposite slant or receive pushback from other contributors. Moreover, the composition of the existing contributors changes. We do not find evidence of a major change in the composition of new participants but do find evidence that more biased contributors exit sooner. Exposure to extreme opposite views is associated with a higher likelihood of withdrawal from participation. Furthermore, exit is the most significant driver of Wikipedia's bias. Simulations suggest that exit is responsible for 80.25%–90.98% of the decline in the slant.

We examine a special circumstance, *mass edits*, where an article attracts an unusually high number of contributions in one day due to a sudden social event or breaking news about the topic. Such events are plausibly caused by factors exogenous to the Wikipedia community. During mass edits, articles experience more flips in slant in one day—from extremely blue/red to extremely red/blue. Consequently, contributors during mass edits are 11.8% more likely to be exposed to extreme content of both slants. As a result,

[1] Source: https://en.wikipedia.org/wiki/Wikipedia:Neutral_point_of_view, accessed November 2018.

contributors involved in mass edits demonstrate significantly faster reductions in slant than those involved in normal edits.

Our approach also allows us to analyze how fast someone changes his/her mind. For example, our estimates suggest that extreme Republicans take one year longer to become regular providers of neutral content than extreme Democrats. We trace this distinction to differences in the topics in which Democrats and Republican contributors participate in.

The findings offer important implications on the management of online communities. Past research typically focuses on various levers, such as social, psychological, and economical, available to managers of online communities, which they can use to maximize participation or minimize churn (e.g., Lerner and Tirole 2002; Wasko and Faraj 2005; Bagozzi and Dholakia 2006; Jeppesen and Frederiksen 2006; Moon and Sproull 2008; Nambisan and Baron 2010; Zhang and Zhu 2011; Gallus 2016; Nagle 2018). We find that one of the key mechanisms on how Wikipedia achieves an NPOV is by letting contributors with extreme viewpoints leave the communities. As long as new contributors continue to arrive, we see less reason for community managers to maximize participation or be overly concerned about the exit of participants. For other community managers, if they spire to draw on multiple opinions and achieve a balance between them in their online communities, they must insist that contributors also aspire to that goal and actively discourage participation from those who maintain extreme points of view.

We also identify a key feature of Wikipedia that facilitates the convergence to neutrality, that is, contributors are frequently exposed to the content of opposite slants. In practice, however, various communities often design algorithms to expose their contributors to content that aligns with their preferences. Although this strategy maximizes participants' satisfaction, as shown in our research, such practices are harmful in building a less-polarized and unbiased crowd.

The importance of Wikipedia in the modern society makes understanding its production interesting in its own right. Most reference information has moved online, and these online sources have displaced other sources of information in every developed country. Wikipedia is the top 20 site in several developed countries and, by far, the most popular and referenced online repository of comprehensive information in the developed world. The English language version of Wikipedia has received over eight billion page views per month and over 500 million unique visitors per month.[2] Many firms also utilize Wikipedia as an input. Amazon (Alexa), YouTube, and Google (search), among others, use Wikipedia as a free source for neutral "facts" and as an unrestricted source for vocabulary in different languages.[3]

2. Relationship with Prior Work

Considerable research has examined the property of online crowds. Although some studies show that collective decision making can be more accurate than experts' decision making (e.g., Antweiler and Frank 2004; Lemos 2004; Surowiecki 2004; Giles 2005; Rajagopalan et al. 2011), others find that a crowd can be more biased (e.g., McPherson et al. 2001; Sunstein 2001; Rector 2008; Gentzkow and Shapiro 2011; Park et al. 2013; Bail et al. 2018; Greenstein and Zhu 2018). For instance, Gentzkow and Shapiro (2011) find biases in online conversations about political content and other topics higher than the segregation of offline news consumptions**.** Greenstein and Zhu (2018) show that Wikipedia articles are on average more biased than those in Britannica, an encyclopedia authored by experts. Several studies have proposed new approaches to aggregate opinions from the crowds to minimize bias (e.g., Fan et al. 2005; Muchnik et al. 2013; Prelec et al. 2017; Wang et al. 2017). Although these studies acknowledge that the bias of a crowd reflects the aggregate of individual participants, to the best of our knowledge, no studies have tracked long-run changes in how extremists (moderates) participate and whether they change their expression to more moderate (extreme) views. We think that this gap arises, in part, because it is rare to observe an online crowd over a long time period and a wide array of topics. This study has an example of such case in Wikipedia, analyzing almost a decade of participation.

---

[2] "Wikipedia vs. the small screen." http://www.nytimes.com/2014/02/10/technology/wikipedia-vs-the-small-screen.html?_r=1, accessed June 2019.

[3] See, e.g., "YouTube May Add to the Burdens of Humble Wikipedia," https://www.nytimes.com/2018/03/19/business/media/youtube-wikipedia.html, accessed June 2019.

This concern becomes more relevant when participants confront *contested knowledge*—defined as topics involving subjective, unverifiable, or controversial information. Many observers are worried about the emergence of *segregated* conversations in the presence of contested knowledge in online crowds. Segregated conversation may become an "echo chamber" (EC) of like-minded views (e.g., Sunstein 2001; Van Alstyne and Brynjolfsson 2005; Carr 2008; Quattrociocchi et al. 2016; Shore et al. 2016; Sun et al. 2017). The opposite behavior, an *unsegregated* conversation, involves contributors with diverse ideas and opposing views (Benkler 2006). Many unsegregated conversations bring varying perspectives into a common view by accelerating a confrontation or discourse between contradictory facts and ideas. Generally, segregated conversations are blamed for many undesirable outcomes.[4]

In the case of contested knowledge in Wikipedia, prior research (Greenstein and Zhu 2012, 2016, 2018) shows that the slant and bias of content evolve and bias in Wikipedia articles slowly declines over time. Prior research does not identify the underlying mechanism other than to affiliate it with more revision. The shift in the slant and bias can be caused by many factors, such as the arrival of moderate contributors, withdrawal of extremists, or changes in contributors' own viewpoints. Without examining actual participants' behavior, it is difficult to draw the right managerial implications for community managers. For instance, Greenstein and Zhu (2018) find that bias in Wikipedia articles tends to slowly decrease with more revisions. One may infer from this result that community managers need to encourage more participation to reduce bias faster. Our research shows that the situation is not just about the number of contributions but about the identity of the contributors. More contributions reduce bias because over time these contributions mostly come from moderate contributors. Fundamentally, the change in the composition

[4] As already implied above, concerns about the health and tenor of political conversations have motivated prior works (e.g., Sunstein 2001; Carr 2008; Lawrence et al. 2010; Gentzkow and Shapiro 2011; Greenstein and Zhu 2012, 2016; Shore et al. 2016; Boxell et al. 2017). Closer to our study, Gentzkow and Shapiro (2011) focus on online conversations about political content and other topics, and Gentzkow and Shapiro (2010) start from the premise that ideological tendencies appear in the language of speakers. Segregation can facilitate the radicalization of some individuals and groups (Purdy 2015). See, for example, http://www.vice.com/read/we-asked-an-expert-how-social-media-can-help-radicalize-terrorists and http://www.rand.org/randeurope/research/projects/internet-and-radicalisation.html, accessed June 2017. Segregated conversation can also discourage interracial friendships, disconnect different social segments, and stimulate social isolation. In traditional media, ideological biases in news content affect the political language (e.g., DellaVigna and Kaplan 2007; Stone 2009; Chiang and Knight 2011; Durante and Knight 2012).

and ideology of the crowd drives the bias reduction. Hence, different from prior research, our study suggests that community managers should devote efforts into designing processes to encourage extremists to leave or convert themselves into more moderate contributors rather than to maximize participation.

In addition, prior studies show that people's beliefs become more reinforced when they encounter information that is aligned with their prior beliefs (e.g., Van Alstyne and Brynjolfsson 2005; Gilbert et al. 2009; Bakshy et al. 2015; Lee et al. 2015; Garimella et al. 2018). However, the manner by which people react to opinions that differ from their beliefs is unclear. For example, studies find that people may demonstrate a "pushback": they refute evidence that has a contrary effect on belief (e.g., Nyhan and Reifler 2010; Wood and Porter 2019). In some rare cases, we may observe a "backfire effect": given evidence against their beliefs, people can reject the evidence and believe even more strongly.[5] Our empirical results that many extreme contributors choose to leave after encountering opposite opinions provide support for the claim that changing people's beliefs is difficult. At the same time, a number of contributors become more moderate after encountering opposite opinions. Although such changes are responsible for a small fraction of overall bias reduction, the evidence restores our hope that communities, such as Wikipedia, can help reduce polarization in our society as they gradually work toward an NPOV.

This study is also related to the literature on platform design for user-generated.[6] Although much of this literature has examined how algorithms (un)intentionally nudge user behavior in one direction or another and how they may produce unanticipated aggregate outcomes because they often seek to match content with a user's taste to maximize a user's satisfaction or to retain users, algorithms play no role in

[5] See, for example, "The Backfire Effect," https://archives.cjr.org/behind_the_news/the_backfire_effect.php, accessed August 2019.

[6] Prior research has examined the importance of contributor motivation for a variety of tasks, such as software design, entrepreneurial finance, and engineering (e.g., Kogut and Metiu 2000, 2001; Rothaermel and Sugiyama 2001; Von Krogh and Von Hippel 2006; Chesbrough 2006; Roberts et al. 2006; Yang et al. 2009; Ramasubbu and Balan 2012; Ransbotham et al. 2012; Kane et al. 2014; Xu et al. 2015; Gallus 2016; Nagaraj 2017; Qiu and Kumar 2017). Most empirical studies have examined how online organizations aggregate contributions to solve collective problems (e.g., Kogut and Zander 1992; Lee and Cole 2003; Hargadon and Bechky 2006; Kuk 2006; Tucci and Villarroel 2007; Xu and Zhang 2009, 2013; Faraj et al. 2011; Ransbotham and Kane 2011; Afuah and Tucci 2012; Chen et al. 2012; Pierce 2012; Bassamboo et al. 2015).

our setting. Wikipedia employs an architecture that gives participants considerable discretion in achieving platform-wide ideals and aspirations.[7] Other studies have examined segregation of participants in social networks, such as Twitter. Shore et al. (2016) study sharing links on Twitter and examine whether participants share with others who are like-minded. Bail et al. (2018) study a field experiment about following opinion leaders on Twitter and examine whether exposure to opposite viewpoints changes users' ideologies over time. This scenario is similar in our study, although the exposure in their study is monetary incentivized. Both Shore et al. (2016) and Bail et al. (2018) examine whether social interactions reinforce segregated conversation, but they reach different conclusions. We regard our setting as an opportunity to understand user behavior in the absence of algorithms and social networking features. Our study suggests that the platform at risk of losing users may actually provide the optimal solution.

3. Empirical Setting

Founded in 2001, Wikipedia positions itself as "the free encyclopedia that anyone can edit" or, in other words, as an online encyclopedia entirely written and edited through user contributions. Topics are divided into unique pages, and users can select any page to revise. It has become the world's largest "collective intelligence" experiment and one of the largest human projects ever to bring information into one source.

Contributions come from tens of millions of dedicated contributors who participate in an extensive set of formal and informal roles.[8] Some roles entail specific responsibilities in editing tasks; however, the Wikimedia Foundation employs a limited set of people and does not generally command its volunteers. Instead, it develops mechanisms to govern the volunteer co-production process (Kane and Fichman 2009;

---

[7] Similar with other online communities, Wikipedia has adopted explicit aspirations, rules, norms, policies (Forte et al. 2009; Jemielniak 2014; Schroeder et al. 2012), and quality assurance procedures (Stvilia et al. 2008), which shape contributors' behavior. Many online communities have adopted privilege access schemes that formally define roles (Arazy et al. 2015; Burke et al. 2008; Collier et al. 2008; Forte et al. 2012), and Wikipedia has performed this as well. This initiative has led to a myriad of coordination mechanisms (Kittur et al. 2007a; Kittur and Kraut 2008; Kittur et al. 2007b; Schroeder and Wagner 2012), social interactions (e.g., Halfaker et al. 2011; Forte et al. 2012), and behaviors aimed at conflict resolution (Arazy et al. 2011).

[8] See https://en.wikipedia.org/wiki/Wikipedia:User_access_levels, accessed June 2017.

Te'eni 2009; Zhang and Zhu 2011; Hill 2017). All these voluntary contributors are considered editors on Wikipedia. The organization relies on contributors to discover and fix passages that do not meet the site's content tenets. However, no central authority tells contributors how to allocate their editorial effort.

The reliance on volunteers has many benefits and drawbacks. Among the latter, there is a long-standing concern that interested parties attempt to rewrite Wikipedia to serve their own parochial interests. Despite the persistence of such concerns, little systematic evidence has pointed in one direction or another. The available evidence on conflicts suggests that contributors who frequently work together do not get into as many conflicts as those who do not, nor do their conflicts last as long (Piskorski and Gorbatâi 2017). Although such behavior can lead to edits from contributors with different points of view, no direct evidence shows that it leads to more content that finds compromises between opposite viewpoints.

Although the Wikipedia attempts to attract a large and diverse community of contributors, it also invites many slanted and biased views, and the openness of Wikipedia's production model (e.g., allowing anonymous contributions) is subject to sophisticated manipulations of content by interested parties. Hence, there is a widespread acceptance of the need for constant vigilance and review.

A key aspiration for all Wikipedia articles is an NPOV (e.g., Majchrzak 2009; Hill 2017). To achieve this goal, "conflicting opinions are presented next to one another, with all significant points of view represented" (Greenstein and Zhu 2012). When multiple contributors make inconsistent contributions, other contributors devote considerable time and effort debating whether the article's text portrays a topic from an NPOV. Because Wikipedia articles face few realistic limits regarding their number or size[9] (due to the absence of any significant storage costs or any binding material expense), conflicts can be addressed by

[9] Over time, a de facto norm has been developed that tends to keep most articles under 6,000 to 8,000 words. This guideline has arisen as editorial teams have debated and discussed the article length necessary to address the topic of the page. Of course, some articles grow to enormous length, and editor contributors tend to reduce this length by splitting them into sub-topics. A prior work (Greenstein and Zhu 2016) finds that the average Wikipedia article is shorter than this norm (just over 4,000 words), but the sample includes a few longer articles (the longest is over 20,000 words).

adding more points of view to articles instead of eliminating them (e.g., Stvila et al. 2008). In general, most disputes are settled without interventions from Wikipedia administrators.[10]

4. Data and Summary Statistics

4.1. Measuring Contributor Slant and Bias

We extend the approach of Greenstein and Zhu (2018) to measure Wikipedia contributors' slants and biases. This approach relies on the modification of an existing method, developed by Gentzkow and Shapiro (2010), for measuring slants and biases in newspapers' political editorials.[11] For example, Gentzkow and Shapiro (2010) find that Democratic representatives are more likely to use phrases, such as "war in Iraq," "civil rights," and "trade deficit," whereas Republican representatives are more likely to use phrases, such as "economic growth," "illegal immigration," and "border security."[12] Similarly, we compute an index for the slant of each article from each source, tracking whether articles employ words or phrases that appear to slant toward either Democrats or Republicans.[13]

Initially, we assume that a contributor's slant is constant throughout the years and define a contributor's slant as the average slant of all the contributions that the person made in our sample. Then, we allow a contributor's slant to evolve over time. The measure of a contributor's slant is computed based on the contributions in each year instead of throughout the sample period. A contributor's bias is the absolute value of the slant in both cases.

---

[10] Similar with all matters at Wikipedia, contributors have discretion to settle disputes on their own. The organization offers a set of norms for the dispute resolution processes, which can be quite elaborate, including the three-revert edit war rule and rules for the intervention of arbitration committees and mediation committees. Administrators can also decide to freeze a contentious article.

[11] Gentzkow and Shapiro (2010) characterize how newspapers also use such phrases to speak to constituents who favor one political approach over another.

[12] Several studies have applied their approach in analyzing political biases in the online and offline content (e.g., Greenstein and Zhu 2012; Jelveh et. al. 2014; Shore et al. 2016). In addition, although Budak et al. (2016) use alternative approaches to measure ideological positions of news outlets, their results are consistent with those of Gentzkow and Shapiro (2010).

[13] An article's slant changes only when code phrases are added and/or dropped.

To construct our sample, we focus on broad and inclusive definitions of U.S. political topics, including all Wikipedia articles that include the keywords "Republican" or "Democrat." We start by gathering a list of 111,216 relevant entries from the online edition of Wikipedia on January 16, 2011. Eliminating the irrelevant articles and those concerning events in countries other than the U.S.[14] reduces our sample to 70,305. Our sample includes topics that are highly debated, such as abortion, gun control, foreign policy, and taxation, and less disputed ones relating to minor historical and political events and biographies of regional politicians. We then collect the revision history data from Wikipedia on January 16, 2011, which yields 2,891,877 unique contributors.

Our key dependent variable is *Contributor Slant*. This measure is developed in two steps. First, every article on Wikipedia has a revision history that, for every edit, records pre-edit and post-edit versions. We compute the slant index for the pre- and post-edit article versions, take the difference between the two, and use this difference as the *slant change* for an edit. We obtain the slant change of every edit. For sequential edits from the same contributor that happened consecutively and without anyone else editing between them, we treat the sequence of edits as one single edit.[15]

To analyze participant behaviors, we exclude the first version of all articles in our sample (or if the article has only one version, then the whole article) as we do not have a prior article slant and cannot observe the EC or non-EC effect for such contributions. We also exclude contributors who made more than 950 edits in any one year (top 0.01%), since these contributors could be bots that regularly maintain Wikipedia or contributors who created many articles when Wikipedia was first founded. These procedures reduce the

[14] The words "Democrat" and "Republican" do not appear exclusively in entries about U.S. politics. If a country name shows up in the title or category names, we then check whether the phrase "United States" or "America" shows up in the title or category names. If yes, then we keep this article. Otherwise, we search the text for "United States" or "America." We retain articles in which these phrases show up more than three times. This process allows us to keep articles on issues, such as the "Iraq war," but excluded articles related to political parties in non-U.S. countries.

[15] These consecutive edits tend to be highly correlated, or they can be several parts of a complete contribution, such as where the contributors saved their work several times. As a robustness check, we exclude deletions from a contributor's edits if the deletion does not bring an article's slant from left/right leaning to right/left leaning or from less to more extreme. Accordingly, deleting biased content to make an article more neutral will not be considered a biased edit. Accordingly, all of our results still hold.

number of observations in the sample to 9,585,443, the number of articles to 65,361, and the number of unique contributors to 2,886,830. Unless otherwise specified, we use this analysis sample throughout the paper.

Next, we focus on individual contributors. We identify and measure the types of changes that each contributor makes to Wikipedia articles. We assign each edit to each contributor and assign a slant value for each edit. Under the assumption that every contributor has one fixed type of slant, we compute the *Contributor Slant* as the average value of the slant index of this contributor. A zero value of *Contributor Slant* means that the user's edits either contain a balanced set of Republican/Democratic words (weighted by their cardinal values) or do not include any of the slanted phrases. A negative or positive value of *Contributor Slant* means that the contributor is Democrat or Republican leaning, respectively. Accordingly, the absolute value of a contributor's slant is equal to the contributor's bias. In our sample, 92.6% of the contributors have a zero slant, whereas the remaining 225,000 contributors make at least one slanted contribution. As it turns out, the majority (57.5%) of contributions to Wikipedia come from contributors with a measurable slant or bias.

Table 1 presents the distribution of contributor types over a 10-year period. When computing the number of Democratic, Republican, and neutral contributors to Wikipedia each year, we count each contributor only once, even if the contributor contributes many times in a year. We summarize the distribution of the contributors' total number of edits over the 10 years in Figure 1. Our sample reflects the well-known skewness of contributions to Wikipedia. More than 75% of the contributors in our sample contributed only once in the entire 10-year period, whereas 97.5% of the contributors contributed fewer than 10 times, averaging less than one contribution per year. Only 1% of the contributors contributed more than 30 times in our sample. We also show the number of edits, number of contributors, and average number of edits per contributor by the contributors' years of experience in Figures 2–4, respectively. Although contributors with four to five years of experience comprise the large part of our sample in terms of the number of contributors and the total number of edits, the average number of edits per contributor does not vary much with years of experience, except for the 0.18% of contributors who joined in January 2011.[16]

We define a contributor as *core* if his or her total number of edits is distributed in the top 10% of all contributors' total number of edits, which in this case is equal to a total of no less than three contributions. Core contributors make 74% of the contributions in the entire sample. In other words, 84.2% of the edits for each article are from core contributors and 15.8% edits are from peripheral contributors. Furthermore, although the number of neutral contributors who contribute each year is more than 10 times the number of contributors who have a slant, the proportion of core contributors in the neutral slant group (15.9% for the full sample) is much smaller compared with that in the other two groups (63.8% and 65.5% for the full sample, respectively). In summary, slanted contributors are more core than neutral contributors, and much of the slanted content comes from contributors making many edits.

4.2. Composition or Ideological Shift?

A few simple graphs illustrate the evolution of contributors' biases. Figure 5 is a bar chart of the average bias of any contributor who contributed at least once in that given year. In addition, in the graphs, we do not plot the observation in 2001 because contributors in the founding year of Wikipedia tend to be different from contributors in later years, and they only account for 0.03% of the full sample.[17] The average bias of the contributors declines over the years.

Two types of changes may have contributed to the decline in bias: a change in the composition of the contributors and/or an ideological shift. That is, new contributors with a moderate bias may join Wikipedia each year, and relatedly, the existing extreme contributors edit less over time or gradually stop participating. Alternatively, contributors can become less biased in their contributions, in which case our assumption of a fixed slant for each contributor requires modification.

[16] Excluding this group of contributors does not change the qualitative results.
[17] Excluding all contributors who joined in 2001 does not change the qualitative results.

First, we consider the changes in the biases of people joining Wikipedia in different years. We compute the average bias of contributors entering in different years and plot the results in a bar chart in Figure 6. We do not observe an obvious pattern across years. In general, except in 2002, contributors who entered earlier are not systematically more biased compared with those who entered later.[18]

Next, we consider existing contributors on Wikipedia. Figure 7 displays the average number of contributions of the extreme contributors each year. A contributor is considered "extreme" if his/her slant across all years is more than two standard deviations away from the center. Although the number of edits from the contributors' first year to their second year on Wikipedia seems to increase, this is followed by a declining pattern as they stay past two years. In other words, as these extreme contributors stay longer, they become less active over time. This pattern indicates the exit of extreme contributors as a potential explanation for the overall bias decline.

For extreme contributors who keep contributing over the years, we then ask whether we can observe a decline in the biases of their contributions. Figure 8 plots the average contributor bias each year for those who are considered "extreme." If we redefine the slant and bias each year (based on their changes in that year), then we observe a constant declining pattern in the biases from these extreme contributors. The pattern suggests that extreme editors become less slanted over time.

Overall, these graphs suggest a change in the composition of extreme contributors and their participation mostly due to the exit of participants and an ideological shift favoring more neutral contributions. The arrival of new participants with neutral views appears to play little role. Do these patterns survive more careful statistical scrutiny? What factors drive the observed behavior patterns? We next define variables in preparation for investigating these questions.

[18] The contributors who entered in 2002, the second year of Wikipedia's existence, are only 0.13% of the full sample. Prior work has shown that the earliest Wikipedian's tended to be extreme Democrat-leaning contributors and composititiion quickly changed to more moderate participants on average and years before Wikipedia became popular. See Greenstein and Zhu (2016).

4.3. Variables

Contributor Slant assumes that contributors have the same slant over their lifetimes in our sample. We next define *Contributor Yearly Slant*, which divide contributors' edits by year, and for each year use the same calculation as for *Contributor Slant* (i.e., we compute the average slant change of all the edits that a contributor has made within a given year). If a contributor's numeric value for slant remains unchanged throughout the years, then his or her *Contributor Yearly Slant* is equal to *Contributor Slant*.[19] Relatedly, *Contributor Yearly Bias* is the average absolute value of the slants (i.e., the biases of a contributor's edits in a given year). The purpose of this variable is to better capture a contributor's bias.

Prior Article Slant denotes an article's slant before a given edit. This variable is essential for analyzing an article's (mis)match with a contributor's slant.

To measure a contributor's experience in interacting with different types of content, we count the contributor's number of edits in a given time period targeting extreme opposite-slant articles, divided by the contributor's total number of edits during that time, and label it as *Opposite-Slant Article Edits Fraction.* Similarly, the proportion of a contributor's edits targeting extreme same-slant articles is labeled as *Same-Slant Article Edits Fraction*, which captures the amount of extreme content that he or she has interacted with. For a contributor who only made neutral edits in a given year, that is, the contributor is considered neutral, *Opposite Slant Article Edits Fractions* and *Same Slant Article Edits Fractions* is equal to zero as either extreme-left or extreme-right content should be seen as "opposite to" or "same as" a neutral person's ideology.

Pushbacks may shape contributors. An extreme example of a pushback is the *revision war*, where a contributor's edit is immediately reverted by another contributor. In this case, the original contributor edits back the same contribution. We count the edits that a contributor makes during such revision wars, divided by the contributor's total number of edits, as *Revision War Edits Fraction*.

[19] If a contributor does not make any contribution in a given year, then his or her *Contributor Yearly Slant* has a missing value in that year.

Throughout the study, unobservable features of articles are a central concern. We add additional measures that may have attracted editors and otherwise had a spurious correlation with the slant or bias of an article. We measure the length of the articles using the number of words in an article prior to a certain edit and label it as *Prior Article Length.* We measure the number of the article's external references and label it as *Prior Refs*. Articles that are long may incorporate more viewpoints, which in turn tends to attract more contributors. In addition, Wikipedia requires citations from major third-party sources as references for its article content (often listed at the bottom of the page), so articles with more references are also more likely to incorporate more outside arguments or controversial views at the time.

In a similar vein, additional controls measure editors' unobservable features. One such variable is *Number of Edits*, which is the total number of *yearly* edits that the contributor has made on Wikipedia in a given year. Another variable is *Starting Number of Edits*, which is equal to the total number of edits that the contributor made in his or her first two years after joining Wikipedia.

To examine the determinants of composition, we define a dummy variable *Stay* at the contributor level. *Stay* is equal to 1 if the contributor made at least one edit in the last year in our sample; otherwise, it is equal to 0.[20]

Table 2 provides the summary statistics of all the variables used in our analysis. The average *Contributor Slant* in our sample is negatively close to zero, indicating that Democrat-leaning contributors are, on average, more slanted than Republican-leaning contributors. Moreover, the article versions in our sample exhibit similar absolute values of extreme slant on both ends. For control variables *Prior Article Length*, *Prior Refs*, and *Number of Edits*, a substantial variation is found across article versions for each of the measures, and we use the logarithm of these control variables in our models because they are highly skewed.[21]

---

[20] We did not construct *Stay* as a time-varying dependent variable by year because only 1.95% of the contributors in our sample was inactive (i.e., "exited") in one year and came back to edit again in another year.

[21] For contributor–year-level variables, observations include only contributor–year combinations where the contributor made at least one edit in the previous year, because the independent variables for the composition shift analysis are calculated based on the past year's edits.

5. Analyzing Contributor Slant

Motivated by the pattern in the raw data, this section quantifies how contributors' contributions change with their edits and analyze whether (and how) their editing experience affects their slant decline.

5.1. Contributors' Participation Pattern on Wikipedia

We first investigate the type of content contributors interact with on Wikipedia. For every edit in our sample, we estimate the following regression model:

$$Contributor\ Slant_j = \alpha_0 + \alpha_1 Prior\ Article\ Slant_{ijt} + X_{ijt}B + \sigma_i + \eta_t + \varepsilon_{ijt}. \qquad (1)$$

In this baseline specification, we set the contributor slant to a fixed value, even though we can observe the same contributor multiple times. Coefficient α_1 identifies whether the average contribution follows EC or non-EC, as noted earlier. To address concerns about unobservable factors influencing the choice, we include X_{ijt}, a vector of the article's characteristics and control variables after article *i* is edited by contributor *j* at time *t*, and σ_i, an article-fixed effect, to control for any fixed differences among articles, and η_t, a year fixed effect, to control for any common media/macroeconomic shocks or Wikipedia policy changes that may differentially affect articles from different years. We note that the key exogenous variable is measured with considerable noise, which can induce attenuation bias in the estimate.[22] Hence, we view the result, at best, as an underestimate. In an alternative approach that mitigates the noise, we create two categorical variables. On the basis of *Contributor Slant*, we create *Contributor Category*, which takes the value of -1, 0, or 1, representing contributors with a slant two standard deviations below the mean, in between, and above the mean, respectively. *Prior Article Category* is the categorical version of *Prior Article*

[22] This is a standard concern with poorly measured exogenous variables. See, for example, Draper and Smith, 1998, p. 19.

Slant. We use *Contributor Category* as the dependent variable, with *Prior Article Category* as the explanatory variable, to estimate standard models for categorical choice (see Appendix Table A1).

In Table 3, we report the estimation results of Equation (1) using ordinary least squares (OLS) regressions. Models (1) to (3) use *Contributor Slant* as the dependent variable. Model (1) includes only *Prior Article Slant* as the explanatory variable. Model (2) adds the control variables *Log (Prior Article Length)* and *Log(Prior Refs).* Model (3) replicates Equation (1), with article and year fixed effects included. The coefficients on *Prior Article Slant* are negative and significant in all three models.[23] This finding indicates that an increase in the article's slant is associated with a decrease in the slant of its next contributor. That is, when the article is more Republican leaning, it tends to attract a more Democrat-leaning user as its next contributor. This pattern is consistent with the non-EC behavior.

5.2. Ideological Shift: How Does Editing Experience Change the Contributions from Contributors?

Why do biased contributors become moderate? What effect does opposite content have? We estimate the following:

$$Contributor\ Yearly\ Bias_{jt} = \beta_0 + \beta_1 Number\ of\ Edits_j +$$

$$\beta_2 Opposite\ Slant\ Article\ Edits\ Fraction_{j,t-1} + \beta_3 Same\ Slant\ Article\ Edits\ Fraction_{j,t-1} +$$

$$\beta_4 Revision\ War\ Edits\ Fraction_{j,t-1} + \mu_j + \epsilon_{jt}. \qquad (2)$$

The unit of analysis is contributor–year, which allows us to observe how each contributor's ideology evolves over time. Observations in the sample include contributor–year combinations where the contributor made at least one edit in the previous year as the independent variables are calculated with reference to the past year's edits. The dependent variable is *Contributor Yearly Bias*, which captures the average bias of a

[23] As a robustness check, we added the slant of the last edit on the article as a control variable and clustered the standard erorr at the article level. All results continue to hold.

contributor's edits in a given year. The contributor's yearly total number of edits, *Number of Edits*, controls for how active the contributor was in the past year.

To capture a contributor's experience in interacting with different types of content, we count the contributor's number of edits in a given time period targeting extreme opposite-slant articles, divided by the contributor's total number of edits during that time, as *Opposite-Slant Article Edits Fraction*, and the proportion of a contributor's edits targeting extreme same-slant articles, as *Same-Slant Article Edits Fraction*. Moreover, to investigate how pushbacks may shape contributors, we measure a form of pushback called the *revision war*, where a contributor's edit is immediately reverted by another contributor and the original contributor edits back the same contribution. We count the edits that a contributor makes during such revision wars, divided by the contributor's total number of edits, as *Revision War Edits Fraction*. In this regression model, we include the lagged *Opposite-Slant Article Edits Fraction*, *Same-Slant Article Edits Fraction*, and *Revision War Edits Fraction* to test how a contributor's experience in the past year with different types of extreme content and pushbacks affects his or her likelihood of adding more bias to existing content. μ_j is a contributor fixed effect to control for any fixed differences among contributors.

Table 4 reports the regression results. In Model (1), we observe that although encountering extreme content with the same slant reinforces the contributor's own ideology, interacting with opposite-slant extreme content causes the contributor's slant to become moderate. We do not find a significant effect from receiving pushbacks on the contributor's average bias. When we use an alternative dependent variable, *Contributor Yearly Maximum Bias*, in Model (2), computed the same as *Contributor Yearly Bias* but taking the contributor's maximum bias instead of average bias in the year, we find a significant negative effect of receiving pushbacks on the contributor's maximum bias in that year.[24] In summary, encountering extreme content of the opposite slant (rather than the same slant) or receiving pushback from other contributors reduces the contributor's own bias to some extent.

[24] We also added year fixed effects and clustered the standard erorr at the contributor level as a robustness check. All results contitute to hold.

Interpreting the above coefficients is difficult, so we use a Markov matrix to illustrate how the slant composition of contributors evolves. This matrix, reported in Table 5, is constructed as follows. First, we divide the time that a contributor has been on Wikipedia in half. Then, we divide the direction of this contributor's edits by attaching values (-1, 0, 1) to negative, zero, and positive slant edits. On the basis of the sum of these values for the first and second halves of the contributor's activity, we categorize the contributor as Democrat, Neutral, or Republican. If the sum of all edits in one half is negative (positive), then the contributor is a Democrat (Republican), and if the sum of all edits in this half is zero, then the contributor is neutral. We perform this step for each half of every contributor's activity on Wikipedia and accumulate them to get the overall transition probabilities in the entire community. For the Democratic- and Republican-leaning contributors in the first half, there is more than a 70% chance that they will move to neutral in the second half of their contribution life span.

Although the community of participants has a tendency of moving toward neutral, Table 5 does not provide any sense of whether this change is more or less pronounced than the composition shift. We first provide evidence for a causal explanation. Then we characterize the composition shift and compare the two shifts.

5.3. Mass Edits: Causal Evidence

The ideal design to establish causality is to employ some exogenous shocks and observe how contributors' slants change before and after these shocks and compare these actions with those that did not receive shocks. We operationalize this idea with the special circumstances of *mass edits*, where an article attracts an unusually high number of contributions in one day due to a sudden social event or breaking news about the topic. On such occasions, the article usually receives a large volume of searches online. However, social events or breaking news is unpredictable, and so are the mass edits.

We define mass edits using online search volumes from the Google Trends website.[25] Specifically, we use the article's title as the search keyword(s) in Google Trends, collect the global daily Google Search Index (GSI) for a two-week window around the potential "mass edit" date, and compare whether the average GSI of the three-day window around the search date (days -1, 0, and 1) is greater than the three-day average GSI before the mass edit date (days -4, -3, and -2) and the three-day average GSI after the mass edit day (days 2, 3, and 4). We then aggregate all the contributions to the article–date level and define an article as experiencing a "mass edit event" if it meets the following conditions: 1) the article receives more than 10 contributions on that date and 2) its title has an abnormal search peak in GSI during the three-day window, as described earlier. We also combine consecutive mass events that happened to the same article on multiple consecutive days as a single mass edit event. Finally, we focus on the top mass edit events whose number of edits is above the 99th percentile of all events as they best represent such shocks. In this way, we identify the top 35 mass edit events.

To estimate the effect of experiencing mass edits on a contributor's later slant, we use a treatment/control approach. We consider the contributors who are exposed to a given mass edit the *Treated* contributors. We use a propensity score matching method to construct a control group of contributors. Specifically, for a given mass edit event, we identify all the contributors who made at least one edit during the event. Then, for each of the treated contributors, we find another contributor who did not edit during the event to create a "matched pair" in our data set. We perform matching based on their previous slant and editing behavior before the mass edit event. Once a contributor is matched, we exclude him or her from future matching, and we repeat the process for the next mass edit event. Altogether, we identify 5,148 treated contributors who experienced mass edits and another 5,148 control contributors as their matched pairs.

[25] For descriptions of the Google Trends website and Google Search Index, see https://support.google.com/trends/?hl=en#topic=6248052, accessed May 2018.

First, we examine whether mass edits produce more biased content, which may lead to contributor slant change. We use a t-test to examine whether the articles being edited during mass edit events are more biased compared with those under normal edits. The results show that, on average, articles (“article versions”) after mass edits are more biased than those experiencing normal edits, that is, the average absolute article slant for articles after mass edits (0.1346) is significantly greater than the average absolute article slant after normal edits (0.0839), with $p = 0.000$. Indeed, mass edit events are more likely to produce biased content.

Next, we use OLS regressions to examine how the treated contributor’s average bias changes compared with the control contributors due to the mass edits. Table 6 reports the results. The treated contributors’ average biases after experiencing the mass edit event are significantly lower compared with those of their pairs who were not exposed to mass edits.

Why do the biases of contributors decrease after experiencing mass edits? In our proposed mechanism, encountering extreme opposite-slant content reduces contributors’ bias. Contributors may be more frequently exposed to extreme opposite-slant content during mass edits, which leads to their ideological shift. To test this mechanism, we perform a simple t-test mean comparison to compare the average article slant during mass edits vs. normal edits. The results show that the article versions after each mass edit are significantly *more biased* than those after each normal edit ($p < 0.0000$). This finding means that contributors are exposed to more extreme content during mass edits.

Next, we define a “flip” of an article’s slant when the slant changes from extremely left/right leaning (i.e., more than two standard deviations away left/right from neutral) to extremely right/left leaning. We estimate the likelihood that an article has at least one slant flip on mass edit days and compare it to normal edit days. The logit regressions are shown in Table 7. Aggregated to the article–date level, *Flip* is equal to 1 if an article has at least one flip on a given day and 0 if the article has no slant flip on that day. After converting the estimated coefficients, an article is 11.8% more likely to experience slant flips during mass edits than during normal edits.

The analyses provide additional evidence to support our proposed mechanism. Contributors encountering mass edits are more likely to be exposed to content of both extremes. Because the impact of the opposite-slant extreme content on a contributor's slant is greater than that of the same-slant extreme content (Table 4), bias decreases more in the contributors who encounter mass edits compared with those who do not. The exogenous nature of mass edit events allows us to obtain a causal inference of such effects.

5.4. Composition Shift: Why Do Extreme Contributors Leave Over Time?

To examine each contributor's exit decision, we employ the following Logit regression to examine each contributor's likelihood of staying/exiting:

$$\begin{aligned} Stay\_Dummy_j \; = \; & \theta_0 + \; \theta_1 Starting\ Number\ of\ Edits_j + \\ & \theta_2 OppositeSlant\ Article\ Edits\ Fraction_j + \; \theta_3 SameSlant\ Article\ Edits\ Fraction_j + \\ & \theta_4 Revision\ War\ Edits\ Fraction_j + \; Vintage\ Dummies_j + \; \omega_j, \end{aligned} \tag{3}$$

where the unit of analysis is each contributor. We include a similar set of explanatory variables as shown in Table 4. The contributor's number of edits in the first two years, *Starting Number of Edits,* controls for how active the contributor was when joining Wikipedia. We also control for the contributor's year of entry to capture the vintage effect.

Table 8 presents the results. Model (1) includes all contributors who joined before 2010, whereas Model (2) includes only the core contributors. The coefficients for *Opposite-Slant Article Edits Fraction* and *Revision War Edits Fraction* are negative and statistically significant. Calculating the average marginal effects shows that experiencing the opposite-slant extreme content or revision wars reduces a contributor's likelihood of staying on Wikipedia by 2.7% and 25.3%, respectively, compared with a contributor who encounters no opposite-slant content or revision wars. The effect of interacting with same-slant content, shown by the coefficient of *Same-Slant Article Edits Fraction*, goes in the other direction, which increases the likelihood of a contributor staying. The results in Tables 4, 6, and 8 show that encountering more extreme opposite-slant content or pushbacks from others leads contributors to become either more moderate or more likely to leave.

5.5. Which Effect Contributes More to Changes in the Overall Bias?

We compare the size of each effect using various five-year subsamples to identify which effect contributes more to overall trends. Below, we take the five-year period of 2004–2009 to illustrate the simulation, and we repeat this process for all other five-year periods.

The goal is to calculate the proportion of the change in overall bias that is due to the ideological shift and that due to the composition shift. To begin, we compute the average bias—the absolute value of slant—of any contributor who contributed at least once in the first and last years. We denote the former crowd of contributors as the *starting crowd* and the latter crowd of contributors as the *ending crowd*. The average bias of the starting crowd in 2004 is 0.00606, whereas the average bias of the ending crowd in 2009 is 0.00421, a decline of 30.5%.

Next, we identify the composition of each crowd. In the starting crowd, 2,866 (4.34%) contributors are the staying contributors who still edit after five years, with an average bias of 0.00539, and 63,124 (95.66%) are the leaving contributors who did not edit after five years and whose average bias in the first year is 0.00609. The average bias of the starting crowd is 0.00606 = 0.00539 × 4.34% + 0.00609 × 95.66%. In the ending crowd, the 2,866 staying contributors have an average bias of 0.00346. The remaining 522,985 (99.45%) are new contributors who joined during the five-year period, whose average bias in 2009 is 0.00421. Thus, we compute the proportion of the bias decline due to each effect.

To consider the importance of the composition effect, we simulate the possible scenario if only the composition effect shaped the outcome. If some contributors leave after five years but the remaining contributors have the same slant as in the beginning, then the ending crowd's bias would be equal to the average of the staying contributors' beginning bias (0.00539) and the new contributors' bias (0.00421), yielding 0.00539 × 0.55% + 0.00421 × 99.45% = 0.00422. Accordingly, the five-year bias decline from the composition effect will be 0.00606 - 0.00422 = 0.00184.

For comparison, we simulate the possible scenario if only the ideological shift shaped the outcome. To estimate the effect of ideological shift, we observe all contributors who are still edit after five years but their slant changes over this time. This is the crowd whose ideologies shift. If we assume that such ideological shift happened to all contributors in the starting crowd instead of just those who remain active, then the ending crowd's average bias would be equal to the average of the staying contributor's decreased bias (0.00346) and the new contributors' bias (0.00421), yielding 0.00346 × 11.2% + 0.00421 × 88.8% = 0.00413.[26] Accordingly, the five-year bias decline from the ideological shift only will be 0.00608 - 0.00413 = 0.00195.

The ideological effect is not common enough to have a large effect on the overall results. It involves as few as only 4.34% of the starting crowd in the above calculation. With an actual crowd bias decline of 0.00185 in the raw data, a simple equation solving for the proportion of each effect yields the estimate that roughly 88.4% of the change is due to the composition effect and the remaining 11.6% is due to the ideological shift effect. Although this is a simple back-of-the-envelope calculation, it provides an indicator that the decline in contributor slant is largely due to the composition effect.

We repeat the above simulation for all five-year periods in our sample and report the effect sizes in Table 9. Panel A uses the same sample as in the main analysis, which excludes the first edits of each article that creates the article, and Panel B uses the full sample containing all the raw edits. Apart from the first five-year period, both sets of results show a consistent pattern: composition shift accounts for 80.25%–90.98% of the effect, whereas ideology shift accounts for the remaining proportion of the effect ranging from 9.02% to 19.75%. The effect sizes are robust across different five-year periods except for the first five years (i.e., 2001 to 2006). A possible reason is that in 2001, when Wikipedia was first founded, the earlier contributors contain a larger portion of extreme participants than later years, as is also shown by the

[26] 11.2% = (2,866 + 63,124) / (2,866 + 63,124 + 522,985).

percentages in Table 1. The total number of contributors that year is also only 0.03% of the full sample. As this crowd has a large portion of extreme contributors to begin with and with the extreme contributors becoming neutral over time, a greater ideology shift is observed in the first five-year period compared with later years. The later subsamples have more participants and examine Wikipedia in its most "developed" state, where the participants are more familiar with all the norms and rules.

6. Discussion

6.1. Rate of Slant Change: How Long Will It Take for Contributors to Become Neutral?

We estimate how long it takes for a contributor's slant to gradually become neutral if this tendency continues. In this simulation, we observe the slant at an aggregated level, and we simulate the slant change of Wikipedians as a crowd, regardless of which part of the change is due to the ideological or composition shift.

We use a Markov chain process to simulate the evolution. Although a contributor's slant exhibits a long-term trend over the years, it frequently fluctuates, and this fluctuation should be accounted for. We divide the slant into different bins and investigate how a contributor's slant changes from one bin to another. *Contributor Yearly Slant* is divided into seven bins and divided by the ± 0.5, ± 1.5, and ± 2.5 standard deviation intervals. The middle bin represents a neutral slant, whereas the first and last bins represent the extreme slants. We then compute a transition matrix for contributor slant based on our empirical data. For each year, we compute the proportion of contributors whose yearly slant moves from one slant bin to another and fill the probabilities in the transition matrix for this year. Averaging the transition matrices across all years gives us the final transition matrix we use in our simulation (reported in Table 10).

In this transition matrix, the rows denote the starting bins and the columns denote the ending slant. Bin 4 represents a neutral slant, defined as a slant index ranging from -0.5 to 0.5 standard deviations away from the mean. We find that: (1) the probabilities on the diagonal are large. As expected, contributors tend to have a high chance of staying near their original slant; and (2) the farther the end bins are from the start bins, the smaller the probabilities. These findings indicate that the contributor slant change is a gradual and cumulative process, and it is not likely that the contributor's slant would suddenly jump from one extreme to another.

Next, we use the transition matrix to simulate the contributor slant change process over time. We compute the time it takes for a contributor to have a greater than 50% probability of moving to neutral (see Table 11). As expected, the length of time depends on the contributor's original slant: Extremely slanted contributors spend a longer time moving to neutral than slightly slanted contributors. Surprisingly, we find that, on average, it takes one more year for Republicans to become neutral than for Democrats.

We test for possible reasons why Republican contributors tend toward a neutral slant more slowly than Democratic contributors. First, do Republican contributors display more EC behavior than Democratic contributors? The regression results of Equation (1) using the two groups separately do not support this explanation. Republican contributors show less EC behavior than Democratic contributors.

Second, Republican contributors may choose to edit less extreme articles compared with Democratic contributors, and so they are less influenced during their interaction with the online content. However, no statistically significant difference was found between the level of content extremeness for the articles edited by Republicans or Democrats. The distributions contain similar bias and variance.

A third possible reason may stem from the contributors' number of edits. Republican contributors make fewer edits than Democrats, so their experience has less of an effect on the overall tendency and may differ in some way. Summary statistics provide evidence for this explanation. In our sample, the total number of edits from Democratic contributors is approximately 1.5 times that of Republican contributors.

6.2. Is the Measure of Contributor Slant Representative of Ideologies?

One may be concerned about whether the measure of slant in Wikipedia is representative of contributors' real-world political ideologies. In addition, a neutral article in our sample can either be interpreted as having no slanted words at all or as having equal numbers of very slanted words. These concerns may lead to questioning the external validity of the slant measure.

To address these concerns, we use an alternative measure of the slant and bias. We match the voting data from the 2004 presidential election to locations affiliated with contributors' IP addresses.[27] We restrict our sample to contributors who are not logged in when editing the articles because Wikipedia only reveals IP addresses for contributors without user IDs. We also drop contributors with IP addresses located outside the U.S.. We then test the relationship between the voting record and *Prior Article Slant* using OLS regressions. Note that this analysis analyzes the behavior of a different population of contributors than the contributors we have examined thus far.[28] This regression is valid under the assumption that a contributor has, on average, the political preferences of the region in which he or she lives.

Appendix Table A2 presents the results. *RepPerc* denotes the percentage of Republican votes in the contributor's county. As we use positive values in the slant index to indicate Republican-leaning ideologies for Wikipedia users and articles, the negative and statistically significant coefficient of *Prior Article Slant* suggests that a contributor from a county with a higher percentage of Republican votes tends to target a Democratic-leaning article when he or she contributes on Wikipedia. The results show a non-EC pattern in the contributing process and are qualitatively similar to the prior estimates. This finding also supports the notion that the measure of contributors' slant reflects the contributors' real political ideologies.

We also collect talk pages for articles, which are used by contributors to discuss edits and achieve consensus. The total size of an article's talk pages has a correlation of 0.22 with the average bias of the article over time, suggesting that our bias measure does capture how contested an article is.

[27] The data on the geolocation of IP comes from MaxMind. We match up the county records.

[28] The identities of contributors are known after they register and when they edit after logging on. An anonymous edit comes from either an unregistered contributor or from an editor who chose not to log on before editing. Hence, the samples can possibly include some of the same contributors, but identifying the fraction is impossible.

6.3. What Else Could Be Driving the Non-EC Behavior?

The effect of non-EC in contributors' voluntary editing behavior indicates that contributors are more likely to edit articles with the opposite slant. This can also be due to the revision war among contributors, which may have little to do with the article's slant. We address this concern by including only the initial edits of every contributor when they revise an article for the first time. This rules out revision wars or any possible correcting behavior later in the edits.

Appendix Table A3 shows that the signs and statistical significance of the estimated coefficients do not change, and the magnitude of the coefficients becomes even larger, indicating an even stronger non-EC effect than that when investigating all edits. The results further strengthen the robustness of the effect.

7. Conclusions

Wikipedia has a long record of bringing opposing opinions into the same conversation. Over time, Wikipedia has become less biased. Our study finds that this change is partly due to an ideological shift, where contributors contribute less slanted content over time. It is also due to a composition shift, where extremely biased contributors contribute less content over time. Contributors interact with opposite viewpoints and do so more frequently than participating in ECs. Extreme contributors either become more moderate or tend to leave after interacting with an opposite-slant content or encountering pushbacks from others.

This study offers an approach for identifying the mechanisms contributing to (un)segregated conversations. It identifies the factors that alter the composition of the crowd and cause a contributor's viewpoint to evolve. Nothing in this approach presumes the results; the approach can flexibly measure contributions to (un)segregated conversations in a variety of settings.

The findings inform open questions for two important managerial issues. Wikipedia's stewards, the Wikimedia Foundation, face an important question about how to encourage the emergence of an NPOV.

Our findings suggest that retaining existing contributors with moderate opinion is, by far, the most important factor in maintaining an NPOV on its articles. Wikipedia should continue to expose all extremes to the opposite opinion as it tends to arise from normal practice, which leads to the exit of the most extreme contributors and moderation of views among those who stay. This approach will work as long as the entry of new contributors continues to draw on a diverse set of opinions as it has in the past.

For webmasters of crowds, we draw a related lesson from Wikipedia's experience. If a webmaster aspires to draw on multiple opinions and achieve near-neutrality in the content produced by their online communities, then the experience at Wikipedia does not suggest a passive approach to managing contested knowledge. Simply maximizing participation, regardless of the opinion, is also a mistake. Webmasters must articulate their aspirations for an NPOV and insist that contributors also aspire to that goal while recruiting a diversity of opinion. If they are successful at recruitment, then actively discouraging participation from those who maintain extreme points of view is reasonable. Indeed, our findings suggest, cautiously, that one way to achieve moderate outcome is to encourage extremists to leave. We stress the cautious nature of the advice because Wikipedia samples from wide and multiple points of view in spite of changes in the composition and that a wide sample should not be taken for granted in other settings.

The findings also raise a subtle question: How does Wikipedia transform controversial topics into arguments that include many points of view and sustain the community over time? We speculate that Wikipedia's success in this regard arises from the institutions that help overcome the challenges affiliated with aggregating contested knowledge. First, the aspiration of achieving an NPOV directs attention to the principle that no side can claim exclusive rights to determine the answer. Second, norms allow every contributor to add another paragraph if it reduces tension by giving voice to dissent. Reducing disputes this way costs little: miniscule storage and transmission costs reduce the cost of listing another view on a web page. Our results also suggest that the conflict resolution mechanisms and the mix of informal and formal norms at Wikipedia play an essential role in encouraging a community that works toward an NPOV. This finding is consistent with theories suggesting that articles go through a lifecycle and settle into a consensus, which contributors subsequently "defend" (see, e.g., Kane et al. 2014). We also note a new open question: Although our findings suggest that Wikipedia's mechanisms work as desired, our findings raise questions about which specific norms, other than the declaration of principles, also contribute.

These findings also raise concerns on the platform design literature. We speculate that some simple design differences may have profound consequences for (un)segregating conversations. For example, on Facebook, an algorithm selects content for users, and its design increases the chance that the participants read and write content only in a community of like-minded people. By contrast, Wikipedia contributors can select to examine whichever content they desire and add and remove material or refine the content in myriad ways. Contributors on Facebook/Twitter can only add additional content on top of what is already there. Allowing the removal or editing of anyone's contributions can change how the reader and writer choose to direct the conversations, resulting in contributions from different points of view. Some platforms also aggregate contributions in ways that shape the prevalence of segregation. For example, on Yelp (e.g., rating restaurants) or Rotten Tomatoes (e.g., rating movies), additional materials can be added without limit. These platforms provide a numerical summary that can direct conversations between readers and reviewers. Our results prompt questions about whether a numerical summary motivates others with views that differ from the summary or attracts more reviews from those who agree with it, and how such a process makes the summary more valuable.

References

Afuah, A., and Tucci, C.L. 2012. Crowdsourcing as a solution to distant search. *Academy of Management Review* 37(3): 355–375.

Antweiler, W., and Frank, M.Z. 2004. Is all that talk just noise? The information content of internet stock message boards. *Journal of Finance* 59(3): 1259–1294.

Arazy, O., Nov, O., Patterson, R., and Yeo, L. 2011. Information quality in Wikipedia: The effects of group composition and task conflict. *Journal of Management Information Systems* 21(4): 71–98.

Arazy, O., Ortega, F., Nov, O., Yeo, L., and Balila, A. 2015. Functional roles and career paths in Wikipedia. *Computer Supported Cooperative Work (CSCW):* 1092–1105.

Bail, C.A., Argyle, L.P., Brown, T.W., Bumpus, J.P., Chen, H., Hunzaker, M.F., Lee, J., Mann, M., Merhout, F., and Volfovsky, A. 2018. Exposure to opposing views on social media can increase political polarization. *Proceedings of the National Academy of Sciences* 115(37): 9216–9221.

Bagozzi, R.P., and Dholakia, U.M. 2006. Open source software user communities: A study of participation in Linux user groups. *Management Science* 52(7): 1099–1115.

Bakshy, E., Messing, S. and Adamic, L.A. 2015. Exposure to ideologically diverse news and opinion on Facebook. *Science* 348(6239):1130-1132.

Bassamboo, A., Cui, R., and Moreno, A. 2015. The wisdom of crowds in operations: Forecasting using prediction markets. Working paper. SSRN: https://ssrn.com/abstract=2679663.

Benkler, Y. 2006. *The Wealth of Networks: How Social Production Transforms Markets and Freedom.* Yale University Press.

Boxell, L., Gentzkow, M., and Shapiro, J.M. 2017. Is the internet causing political polarization? Evidence from demographics. *National Bureau of Economic Research* Working paper No. 23258.

Budak, C., Goel, S., and Rao, J. M. 2016. Fair and balanced? Quantifying media bias through crowdsourced content analysis. *Public Opinion Quarterly* 80(S1): 250–271.

Burke, M., and Kraut, R. 2008. Mopping up: Modeling Wikipedia promotion decisions. *Proceedings of the 2008 ACM Conference on Computer Supported Cooperative Work (CSCW)*: 27–36.

Carr, N. 2008. *The Big Switch: Rewiring the World, from Edison to Google*. Norton, New York.

Chen, S.S., Chuang, Y.W., and Chen, P.Y. 2012. Behavioral intention formation in knowledge sharing: Examining the roles of KMS quality, KMS self-efficacy, and organizational climate. *Knowledge-Based Systems* 31(7): 106–118.

Chesbrough, H.W. 2006. *Open Innovation: the New Imperative for Creating and Profiting from Technology*. Harvard Business Press.

Chesney, T. 2006. An empirical examination of Wikipedia's credibility. *First Monday* 11(11).

Chiang, C.F., and Knight, B. 2011. Media bias and influence: Evidence from newspaper endorsements. *Review of Economic Studies* 78(3): 795–820.

Collier, B., Burke, M., Kittur, N., and Kraut, R. 2008. Retrospective versus prospective evidence for promotion: The case of Wikipedia. *2008 Meeting of the Academy of Management.*

DellaVigna, S., and Kaplan, E. 2007. The Fox News effect: Media bias and voting. *Quarterly Journal of Economics* 122(3): 1187–1234.

Draper, N.R., and Smith, H. 1998. *Applied Regression Analysis (3rd ed.).* John Wiley.

Durante, R., and Knight, B. 2012. Partisan control, media bias, and viewer responses: Evidence from Berlusconi's Italy. *Journal of the European Economic Association* 10(3): 451–481.

Fan, M., Tan, Y., and Whinston, A.B. 2005. Evaluation and design of online cooperative feedback mechanisms for reputation management. *IEEE Transactions on Knowledge and Data Engineering* 17(2): 244–254.

Faraj, S., Jarvenpaa S.L., and Majchrzak, A. 2011. Knowledge collaboration in online communities. *Organization Science* 22(5): 1224–1239.

Forte, A., Kittur, N., Larco, V., Zhu, H., Bruckman, A., and Kraut, R.E. 2012. Coordination and beyond: Social functions of groups in open content production. *Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work:* 417–426.

Forte, A., Larco, V., and Bruckman, A. 2009. Decentralization in Wikipedia governance. *Journal of Management Information Systems* 26(1): 49–72.

Gallus, J. 2016. Fostering public good contributions with symbolic awards: A large-scale natural field experiment at Wikipedia. *Management Science* 63(12): 3999–4015.

Galton, F. 1907. Vox populi (the wisdom of crowds). *Nature* 75(7): 450–451.

Garimella, K., De Francisci Morales, G., Gionis, A. and Mathioudakis, M. 2018. Political discourse on social media: Echo chambers, gatekeepers, and the price of bipartisanship. In *Proceedings of the 2018 World Wide Web Conference*.

Gentzkow, M., and Shapiro, J.M. 2003. Media, education, and anti-Americanism in the Muslim world. *Journal of Economic Perspectives* 18(3): 117–133.

Gentzkow, M., and Shapiro, J.M. 2010. What drives media slant? Evidence from U.S. daily newspapers. *Econometrica* 78(1): 35–71.

Gentzkow, M., and Shapiro, J.M. 2011. Ideological segregation online and offline. *Quarterly Journal of Economics* 126(4): 1799–1839.
Gilbert, E., Bergstrom, T., and Karahalios, K. 2009. Blogs are echo chambers: Blogs are echo chambers. In *42nd Hawaii International Conference on System Sciences*.
Giles, J. 2005. Internet encyclopaedias go head to head. *Nature* 438(12): 900–901.
Greenstein, S., and Zhu, F. 2012. Is Wikipedia biased? *American Economic Review (Papers and Proceedings)* 102(3): 343–348.
Greenstein, S., and Zhu, F. 2016. Open content, Linus' Law, and neutral point of view. *Information Systems Research* 27(3): 618–635.
Greenstein, S., and Zhu, F. 2018. Do experts or collective intelligence write with more bias? Evidence from Encyclopædia Britannica and Wikipedia. *MIS Quarterly* 42(3): 945–959.
Halfaker, A., Kittur, A., and Riedl, J. 2011. Don't bite the newbies: How reverts affect the quantity and quality of Wikipedia work. *Proceedings of the 7th International Symposium on Wikis and Open Collaboration (WikiSym'11)*.
Hargadon, A.B., and Bechky, B.A. 2006. When collections of creatives become creative collectives: A field study of problem solving at work. *Organization Science* 17(4): 484–500.
Hill, M. 2017. Almost Wikipedia: Eight early Encyclopedia projects and the mechanisms of collective action. Working paper.
Jelveh, Z., Kogut, B., and Naidu, S. 2014. Political language in economics. *Columbia Business School Research*, Paper No. 14–57.
Jemielniak, D. 2014. *Common Knowledge?: An Ethnography of Wikipedia*. Palo Alto, CA: Stanford University Press.
Jeppesen, L.B., and Frederiksen, L. 2006. Why do users contribute to firm-hosted user communities? The case of computer-controlled music instruments. *Organization Science* 17(1): 45–63.
Kane, G.C., and Fichman, R.G. 2009. The shoemaker's children: Using Wikis for information systems teaching, research, and publication. *MIS Quarterly* 33(1): 1–17.
Kane, G.C., Johnson, J., and Majchrzak, A. 2014. Emergent life cycle: The tension between knowledge change and knowledge retention in open online coproduction communities. *Management Science* 60(12): 3026–3048.
Kittur, A., Chi, E., Pendleton, B.A., Suh, B., and Mytkowicz, T. 2007. Power of the few vs. wisdom of the crowd: Wikipedia and the rise of the bourgeoisie. *World Wide Web* 1(2): 19.
Kittur, A., and Kraut, R. 2008. Harnessing the wisdom of crowds in Wikipedia: Quality through coordination. *ACM Conference on Computer Supported Cooperative Work:* 37–46.
Kittur, A., Suh, B., Pendleton, B.A., and Chi, E.H., 2007. He says, she says: Conflict and coordination in Wikipedia. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems:* 453–462.
Kogut, B., and Metiu, A. 2000. The emergence of E-Innovation: Insights from open source software development. *Reginald H. Jones Center Working Paper*.
Kogut, B., and Metiu, A. 2001. Open-source software development and distributed innovation. *Oxford Review of Economic Policy* 17(2): 248–264.
Kogut, B., and Zander, U. 1992. Knowledge of the firm, combinative capabilities, and the replication of technology. *Organization Science* 3(3): 383–397.
Kuk, G. 2006. Strategic interaction and knowledge sharing in the KDE developer mailing list. *Management Science* 52(7): 1031–1042.
Lawrence, E., Sides, J., and Farrell, H. 2010. Self-segregation or deliberation?: Blog readership, participation, and polarization in American politics. *Perspectives on Politics* 8(1): 141–157.
Lee, G.K., and Cole, R.E. 2003. From a firm-based to a community-based model of knowledge creation: The case of the Linux kernel development. *Organization Science* 14(6): 633–649.
Lee, Y-J., Hosanagar, K., and Tan, Y. 2015. Do I follow my friends or the crowd? Information cascades in online movie ratings. *Management Science* 61(9): 2241–2258.

Lemos, R. 2004. Security research suggests linux has fewer flaws. *CNET News*, December 13, http://news.cnet.com/Security-research-suggests-Linux-has-fewer-flaws/2100-1002_3-5489804.html.

Lerner, J., and Tirole, J. 2002. Some simple economics of open source. *Journal of Industrial Economics* 50(2): 197–234.

Majchrzak, A. 2009. Where is the theory in Wikis? *MIS Quarterly* 33(1): 18–20.

McPherson, M., Smith-Lovin, L., and Cook, J.M. 2001. Birds of a feather: Homophily in social networks. *Annual Review of Sociology* 27(1): 415–444.

Mollick, E., and Nanda, R. 2015. Wisdom or madness? Comparing crowds with expert evaluation in funding the arts. *Management Science* 62(6): 1533–1553.

Moon, J.Y., and Sproull, L.S. 2008. The role of feedback in managing the Internet-based volunteer work force. *Information Systems Research* 19(4): 494–515.

Muchnik, L., Aral, S., and Taylor, S.J. 2013. Social influence bias: A randomized experiment. *Science* 341(6146): 647–651.

Nagaraj, A. 2017. Does copyright affect reuse? Evidence from Google Books and Wikipedia. *Management Science* 64(7): 3091–3107.

Nagle, F. 2018. Learning by contributing: gaining competitive advantage through contribution to crowdsourced public goods. *Organization Science* 29(4): 569–587.

Nambisan, S., and Baron, R.A. 2010. Different roles, different strokes: Organizing virtual customer environments to promote two types of customer contributions. *Organization Science* 21(2): 554–572.

Nyhan, B., and Reifler, J. 2010. When corrections fail: The persistence of political misperceptions. *Political Behavior* 32(2): 303–330.

Park, J., Konana, P., Gu, B., Kumar, A., and Raghunathan, R. 2013. Information valuation and confirmation bias in virtual communities: Evidence from stock message boards. *Information Systems Research* 24(4): 1050–1067.

Pierce, L. 2012. Organizational structure and the limits of knowledge sharing: Incentive conflict and agency in car leasing. *Management Science* 58(6): 1106–1121.

Piskorski, M.J., and Gorbatâi A. 2017. Testing Coleman's social-norm enforcement mechanism: Evidence from Wikipedia. *American Journal of Sociology* 122(4): 1183–1222.

Prelec, D., Seung, H.S. and McCoy, J. 2017. A solution to the single-question crowd wisdom problem. *Nature* 541(7638): 532.

Purdy, W. 2015. Radicalization: Social media and the rise of terrorism. *Testimony presented before the House Committee on Oversight and Government Reform's Subcommittee on National Security*, https://oversight.house.gov/wp-content/uploads/2015/10/10-28-2015-Natl-Security-Subcommittee-Hearing-on-Radicalization-Purdy-TRC-Testimony.pdf.

Qiu, L., and Kumar, S. 2017. Understanding voluntary knowledge provision and content contribution through a social-media-based prediction market: A field experiment. *Information Systems Research* 28(3): 529–546.

Quattrociocchi, W., Scala, A., and Sunstein, C.R. 2016. Echo chambers on Facebook. Working paper. SSRN: https://ssrn.com/abstract=2795110.

Rajagopalan, M. S., Khanna, V. K., Leiter, Y., Stott, M., Showalter, T. N., Dicker, A. P., and Lawrence, Y. R. 2011. Patient-oriented cancer information on the internet: a comparison of Wikipedia and a professionally maintained database. *Journal of Oncology Practice* 7(5): 319–323.

Ramasubbu, N., and Balan, R.K. 2012. Overcoming the challenges in cost estimation for distributed software projects. *Proceedings of the 34th International Conference on Software Engineering*: 91–101.

Ransbotham, S., and Kane, G.C. 2011. Membership turnover and collaboration success in online communities: Explaining rises and falls from grace in Wikipedia. *MIS Quarterly* 35(3): 613–627.

Ransbotham, S., Kane, G.C., and Lurie, N.H. 2012. Network characteristics and the value of collaborative user-generated content. *Marketing Science* 31(3): 387–405.

Rector, L. H. 2008. Comparison of Wikipedia and other encyclopedias for accuracy, breadth, and depth in historical articles. *Reference Services Review* 36(1): 7–22.

Roberts, J.A., Hann, I.H., and Slaughter, S.A. 2006. Understanding the motivations, participation, and performance of open source software developers: A longitudinal study of the Apache projects. *Management Science* 52(7): 984–999.

Rothaermel, F.T., and Sugiyama, S. 2001. Virtual Internet communities and commercial success: Individual and community-level theory grounded in the atypical case of TimeZone.com. *Journal of Management* 27(3): 297–312.

Schroeder, A., and Wagner, C. 2012. Governance of open content creation: a conceptualization and analysis of control and guiding mechanisms in the open content domain. *Journal of the American Society for Information Science and Technology* 63(10): 1947–1959.

Shankland, S. 2003. Study lauds open-source code quality. *CNET News*, February 19. http://news.cnet.com/Study-laudsopen-source-code-quality/2100-1001_3-985221.html.

Shore, J., Baek, J., and Dellarocas, C. 2016. Network structure and patterns of information diversity on Twitter. Working paper.

Stone, D.F. 2011. Ideological media bias. *Journal of Economic Behavior & Organization* 78(3): 256–271.

Stvilia, B., Twidale, M. B., Smith, L. C., and Gasser, L. 2008. Information quality work organization in Wikipedia. *Journal of the American Society for Information Science and Technology* 59(6): 983–1001.

Sun, T., Viswanathan, S., and Zheleva, E. 2017. Creating social contagion through firm mediated message design: Evidence from a randomized field experiment. *Robert H. Smith School Research Paper No. RHS 2543864.*

Sunstein, C.R. 2001. *Echo Chambers: Bush v. Gore, Impeachment, and Beyond.* Princeton University Press.

Surowiecki, J. 2004. The wisdom of crowds: Why the many are smarter than the few and how collective wisdom shapes business. *Economies, Societies and Nations* 296.

Te'eni, D. 2009. Comment: The Wiki way in a hurry—The ICIS anecdote. *MIS Quarterly* 33(1): 20–22.

Tucci, C.L., and Villarroel, J.A. 2007. Open Source Corporate Strategy (OSCS): Reaching outside the firm's boundaries to tap upon global talent. In *Academy of Management Annual Meeting* (No. CSI-CONF-2007-001).

Van Alstyne, M., and Brynjolfsson, E. 2005. Global village or cyber-balkans? Modeling and measuring the integration of electronic communities. *Management Science* 51(6): 851–868.

Von Krogh, G., and Von Hippel, E. 2006. The promise of research on open source software. *Management Science* 52(7): 975–983.

Wang, Y., Goes, P., Wei, Z., and Zeng, D.D. 2017. Production of online word-of-mouth: Peer effects and the moderation of user characteristics. Working paper. SSRN: https://ssrn.com/abstract=2899452.

Wasko, M.M., and Faraj, S. 2005. Why should I share? Examining social capital and knowledge contribution in electronic networks of practice. *MIS Quarterly* 29(1): 35–57.

Wood, T., and Porter, E. 2019. The elusive backfire effect: Mass attitudes' steadfast factual adherence. *Political Behavior* 41(1): 135–163.

Xu, L., Niculescu, M.F., and Wu, D.J. 2015. When to use the open business model for software products under network effects? Working paper.

Xu, S.X., and Zhang, X.M. 2009. How do social media shape the information environment in the financial market? *ICIS 2009 Proceedings*: 56.

Xu, S.X., and Zhang, X.M. 2013. Impact of Wikipedia on market information environment: Evidence on management disclosure and investor reaction. *MIS Quarterly* 37(4): 1043–1068.

Yang, Y., Chen, P.Y., and Pavlou, P. 2009. Open innovation: An empirical study of online contests. *ICIS 2009 Proceedings*: 13.

Zhang, X.M., and Zhu, F. 2011. Group size and incentives to contribute: A natural experiment at Chinese Wikipedia. *American Economic Review* 101(4): 1601–1615.

Table 1: Distribution of Different Types of Contributors over Years

Year	Democrat Contributors	Core Democrat Contributors	Republican Contributors	Core Republican Contributors	Neutral Contributors	Core Neutral Contributors	Total # of Contributors Contributed in the Year
2001	26.4%	18.1%	20.0%	12.5%	53.6%	9.9%	800
2002	9.9%	7.5%	9.6%	7.4%	80.4%	17.6%	4,364
2003	8.5%	6.5%	8.8%	6.9%	82.6%	18.3%	14,951
2004	7.8%	5.7%	7.7%	5.9%	84.5%	17.3%	66,867
2005	7.0%	4.7%	6.7%	4.6%	86.3%	15.6%	242,121
2006	5.7%	3.6%	5.7%	3.6%	88.6%	14.7%	584,438
2007	5.3%	3.2%	5.2%	3.3%	89.5%	13.8%	706,195
2008	5.2%	3.1%	5.3%	3.2%	89.5%	13.9%	640,871
2009	4.7%	3.1%	4.7%	3.2%	90.5%	14.1%	526,255
2010	4.2%	2.8%	4.2%	2.9%	91.6%	13.2%	461,663
2011	9.5%	8.5%	10.8%	9.9%	79.6%	19.4%	26,886

Notes: "Democrat/Republican/Neutral Contributors" shows the percentage of contributors with negative/zero/positive *Contributor Slant* among all contributors who contribute in that year to the articles in our sample. "Core Democrat/Republican/Neutral Contributors" shows the percentage of that year's "Democrat/Republican/Neutral Contributors" whose total number of edits is distributed in the top 10% of all contributors' total number of edits. Final year, 2011, is sampled in January, which accounts for the low numbers in that year.

Table 2: Summary Statistics of Variables Used in the Main Analyses

Variable	Unit	Mean	Std. dev.	Median	Min	Max
For Participation Pattern Analysis:						
Contributor Slant	Contribution	-0.0001	0.024	0	-1.229	0.998
Prior Article Slant	Contribution	-0.056	0.208	0	-0.605	0.624
Prior Article Length	Contribution	4,118.030	3,785.242	3173	1	1,963,441
Prior Refs	Contribution	34.193	60.715	7	0	1,636
For Ideological Shift Analysis:						
Contributor Yearly Bias	Contributor-Year	0.004	0.028	0	0	0.819
Opposite-Slant Article Edits Fraction	Contributor-Year	0.006	0.051	0	0	1
Same-Slant Article Edits Fraction	Contributor-Year	0.069	0.207	0	0	1
Revision War Edits Fraction	Contributor-Year	0.014	0.056	0	0	1
Number of Edits	Contributor-Year	12.337	45.157	2	1	949
For Composition Shift Analysis:						
Stay	Contributor	0.030	0.170	0	0	1
Opposite-Slant Article Edits Fraction	Contributor	0.004	0.053	0	0	1
Same-Slant Article Edits Fraction	Contributor	0.075	0.251	0	0	1
Revision War Edits Fraction	Contributor	0.008	0.056	0	0	0.944
Starting Number of Edits	Contributor	0.321	0.693	0	0	9.568

Notes: Number of observations in this table is: 9,585,443 for contribution level; 381,135 for contributor-year level where the contributor made at least one edit in the year before; and 2,482,095 for contributor level where the contributor joined Wikipedia before 2010.

Table 3: OLS Regressions on the Relationship between Contributor Slant and Prior Article Slant

Model	(1)	(2)	(3)
Dependent Variable	Contributor Slant	Contributor Slant	Contributor Slant
Prior Article Slant	-0.0086*** [0.0001]	-0.0086*** [0.0001]	-0.0187*** [0.0004]
Log(Prior Article Length)		0.0006*** [0.0000]	0.0009*** [0.0001]
Log(Prior Refs)		-0.0004*** [0.0000]	-0.0010*** [0.0001]
Observations	9,585,443	9,585,443	9,585,443
Adjusted R-squared	0.006	0.006	0.007
Year FE	No	No	Yes
Article FE	No	No	Yes
Number of Articles	65,361	65,361	65,361

Notes: The unit of analysis is each edit in our analysis sample. Robust standard errors in brackets. *significant at 10%; ** significant at 5%; *** significant at 1%.

Table 4: OLS Regressions on Contributor Slant and Content the Contributor Interact With

Model	(1)	(2)
Dependent Variable	Contributor Yearly Bias	Contributor Yearly Maximum Bias
Log(Number of Edits)	0.0003***	-0.0001
	[0.0001]	[0.0003]
Opposite-Slant Article Edits Fraction	-0.0536***	-0.1599***
	[0.0042]	[0.0081]
Same-Slant Article Edits Fraction	0.0007*	0.0040***
	[0.0004]	[0.0012]
Revision War Edits Fraction	-0.0005	-0.0084*
	[0.0013]	[0.0048]
Observations	381,135	381,135
R-squared	0.0003	0.001
Year FE	Yes	Yes
Contributor FE	Yes	Yes

Notes: The unit of analysis is each contributor-year. Observations include contributor-years where the contributor made at least one edit in the year before. *Contributor Yearly Bias* is the average absolute value of the slants, i.e., the biases, of a contributor's edits that year. *Contributor Yearly Maximum Bias* is the maximum bias of a contributor's edits that year. A *Revision war* is defined as a contributor's edit being reverted immediately by another contributor, and then is immediately followed by the original contributor editing back the same contribution, as a "fight back." The "fractions" in this table are lagged by one year; for example, *Revision War Edits Fraction* in this table equals the number of edits that the contributor made in the past year during such pushback situations, divided by the contributor's total number of edits in the past year. Robust standard errors in brackets, clustered at the contributor level. *significant at 10%; ** significant at 5%; *** significant at 1%.

Table 5: OLS Regressions on How Contributor Slant Changes during Mass Edits

Model	(1)
Dependent Variable	Contributor Bias
Treated	-0.0015*** [0.0005]
Observations	10,296
Adjusted R-squared	0.001

Notes: Observations are the contributors who experienced mass edit events and their matched pairs. *Contributor Bias* is the contributor's average absolute value of their edits after the mass edit event. Robust standard errors in brackets. *significant at 10%; ** significant at 5%; *** significant at 1%.

Table 6: Logit Regressions between Article Slant Flips and Mass Edit Events

Model	(1)	(2)
Dependent Variable	Flip Dummy	Flip Dummy
Mass Edits Dummy	1.3893*** [0.1151]	2.0174*** [0.1161]
Prior Article Slant		-0.5597*** [0.0657]
Log(Prior Refs)		-0.0833*** [0.0116]
Log(Prior Article Length)		-0.4635*** [0.0076]
Observations	5,804,714	5,804,714
Pseudo R-squared	0.001	0.042

Notes: Observations in this panel are at the article-date level. The dependent variable *Flip Dummy* equals 1 if the article experiences at least one slant flip during that day. *Mass Edits Dummy* represents whether the article is receiving mass edits on the given day. Robust standard errors in brackets. *significant at 10%; ** significant at 5%; *** significant at 1%.

Table 7: Transition Matrix of Contributor Slant Change in Wikipedia

		First half of activity		
		Democratic Type	Neutral	Republican Type
Second half of activity	Democratic Type	0.1407	0.0328	0.1145
	Neutral	0.7451	0.9333	0.7416
	Republican Type	0.1142	0.0339	0.1439

Notes: The sample is constructed by dividing every contributor's time in half. Then divide the direction of his or her edits, i.e. attach values (-1, 0, 1) to negative, 0, positive slant edits. Sum up the edits' values for the first half and the second half of his or her activity. If the sum of all edits in this half is negative, the contributor is a Democrat Type in this half. If the sum of all edits in this half is zero, the contributor is Neutral in this half. If the sum of all edits in this half is positive, the contributor is Republican Type in this half.

Table 8: Logit Regressions on Contributors' Likelihood of Staying on Wikipedia

Model	(1)	(2)
Sample	All Contributors	Core Contributors Only
Dependent Variable	Stay	Stay
Log(Starting Number of Edits)	1.6167*** [0.0049]	0.6514*** [0.0051]
Opposite-Slant Article Edits Fraction	-1.2748*** [0.1254]	-0.5965*** [0.0795]
Same-Slant Article Edits Fraction	0.0682*** [0.0156]	0.0583** [0.0268]
Revision War Edits Fraction	-11.7968*** [0.2894]	-9.3117*** [0.1258]
Observations	2,482,095	337,670
Pseudo R-squared	0.316	0.126
Vintage Dummies	Yes	Yes

Notes: The unit of analysis is each contributor. Observations include contributors who joined before 2010. Model (2) includes only core contributors, i.e., contributors who made at least 3 edits in the sample. The dependent variable *Stay* equals 1 if the contributor made at least 1 contribution in 2010 or 2011. *Starting Number of Edits* is the total number of contributions that a contributor made in the first two years after he or she joined Wikipedia. *Vintage* is the year in which the contributor joins Wikipedia, or the vintage that the contributor belongs to. The "fractions" are defined the same as in Table 4, but for contributor's edits in all time instead of the year before. Robust standard errors in brackets. *significant at 10%; ** significant at 5%; *** significant at 1%.

Table 9: Simulation Results Comparing Ideology Shift and Composition Shift over Time

Panel A – Percentages of the Effect Sizes over Time, Using Analysis Sample

Five-Year Period	Composition Shift	Ideology Shift
2006-2011	83.56%	16.44%
2005-2010	83.95%	16.05%
2004-2009	88.36%	11.64%
2003-2008	90.98%	9.02%
2002-2007	89.38%	10.62%
2001-2006	48.03%	51.97%

Panel B – Percentages of the Effect Sizes over Time, Using Full Sample

Five-Year Period	Composition Shift	Ideology Shift
2006-2011	81.72%	18.28%
2005-2010	81.62%	18.38%
2004-2009	87.70%	12.30%
2003-2008	89.07%	10.03%
2002-2007	80.25%	19.75%
2001-2006	50.58%	49.42%

Notes: Percentages reported are from the simulation in Section 5.5 comparing the effect sizes of Composition Shift and Ideology Shift. Panel A uses the same sample as in the main analysis, which excludes the first edit of each article that creates the article. Panel B uses the full sample containing all the raw edits.

Table 10: Transition Matrix of Contributor Slant Change over Time

			Start Slant						
			bin1	bin2	bin3	bin4	bin5	bin6	bin7
			[-1.229, -0.059)	[-0.059, -0.035)	[-0.035, -0.012)	[-0.012, 0.012)	[0.012, 0.035)	[0.035, 0.059)	[0.059, 1.000)
End Slant	bin1	[-1.229, -0.059)	0.8298	0.0139	0.0024	0.0011	0.0013	0.0008	0.0015
	bin2	[-0.059, -0.035)	0.0717	0.7242	0.0044	0.0020	0.0103	0.0019	0.0007
	bin3	[-0.035, -0.012)	0.0591	0.1745	0.7438	0.0055	0.0040	0.0149	0.0029
	bin4	[-0.012, 0.012)	0.0323	0.0713	0.2286	0.9795	0.2089	0.0531	0.0277
	bin5	[0.012, 0.035)	0.0036	0.0128	0.0177	0.0060	0.7545	0.1867	0.0624
	bin6	[0.035, 0.059)	0.0008	0.0014	0.0015	0.0033	0.0052	0.7222	0.0757
	bin7	[0.059, 1.000)	0.0028	0.0019	0.0018	0.0025	0.0158	0.0203	0.8291

Notes: *Contributor Yearly Slant* is split by the ± 0.5, ± 1.5, and ± 2.5 standard deviations intervals. The middle bin represents neutral slant; the first/last bin represents extreme slant.

Table 11: Time Needed for a Contributor to Have > 50% Probability of Moving to Neutral Slant

Starting Contributor Slant	**Number of Years**
Extremely Democratic	10
Democratic	6
Slightly Democratic	3
Neutral	0
Slightly Republican	4
Republican	7
Extremely Republican	11

Notes: Number of years calculated based on the Markov Chain Process. *Neutral* state includes contributor slant 0.5 standard deviation away from 0. *Slightly Democratic (Republican)* state includes contributor slant between 0.5 and 1.5 standard deviations below (above) 0. *Democratic (Republican)* state includes contributor slant between 1.5 and 2.5 standard deviations below (above) 0. *Extremely Democratic (Republican)* state includes contributor slant more than 2.5 standard deviations below (above) 0. On average, after about 30 years, the probabilities in all articles' end state reach stationary distribution, with the probability of contributor slant moving to *Neutral* being 87.4%.

Figure 1: Distribution of Contributors' Total Number of Edits

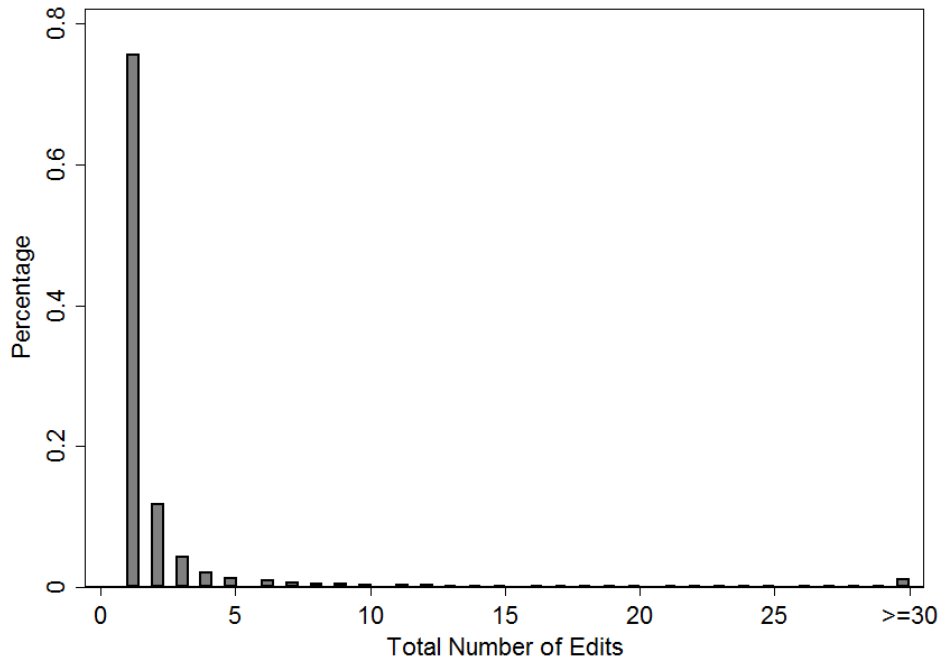


Figure 2: Distribution of All Edits in the Sample by Contributors' Years of Experience

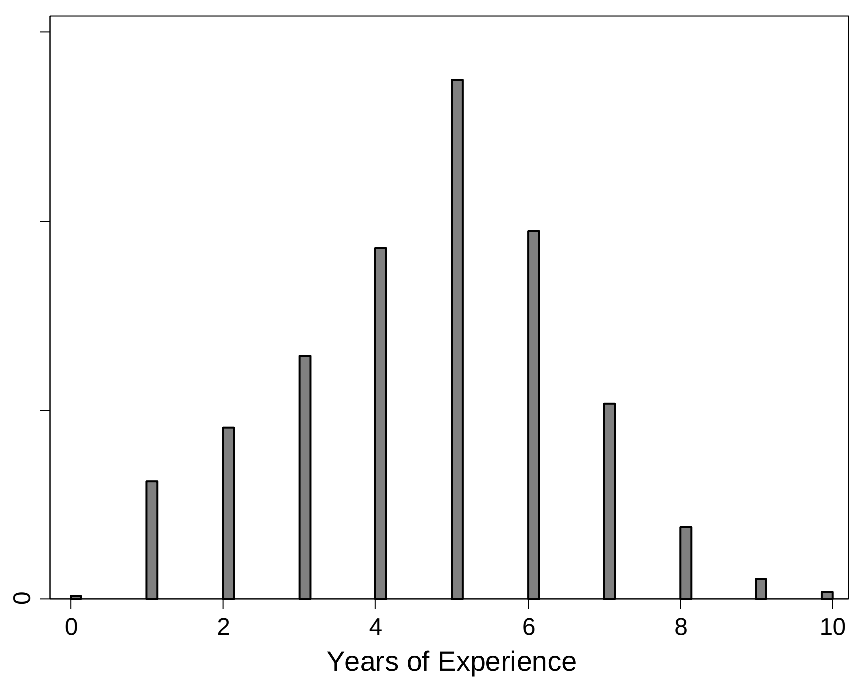


Figure 3: Distribution of Number of Contributors by Years of Experience

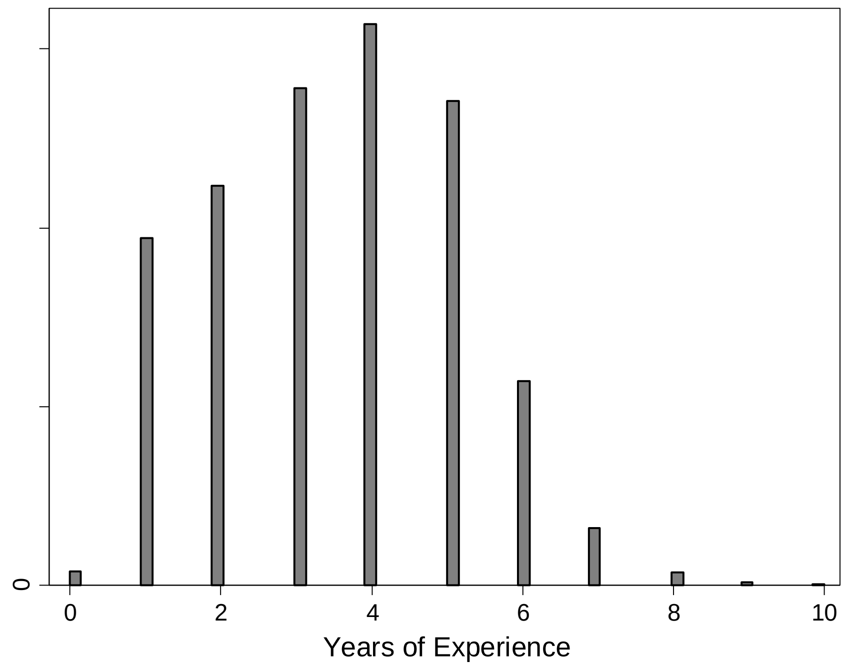


Figure 4: Distribution of Average Number of Edits per Contributor by Years of Experience

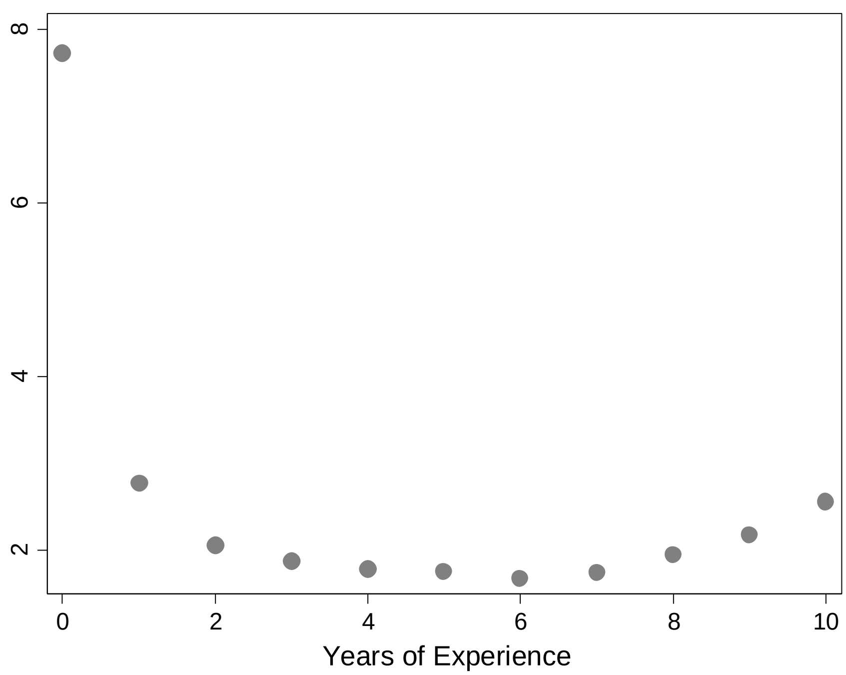


Figure 5: Average Contributor Bias over the Years

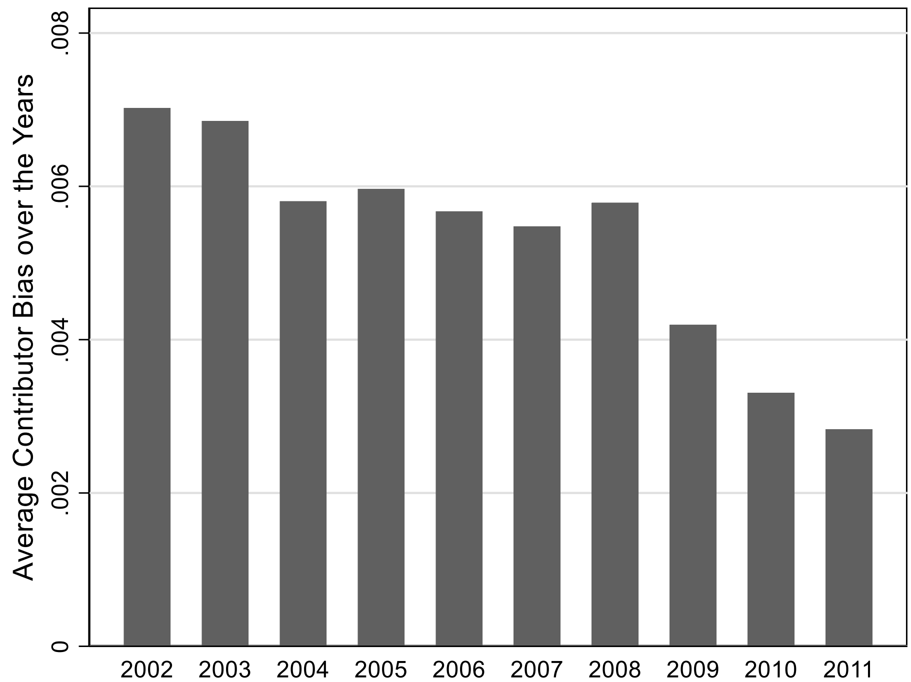


Figure 6: Vintage Effect for Contributors Entering in Different Years

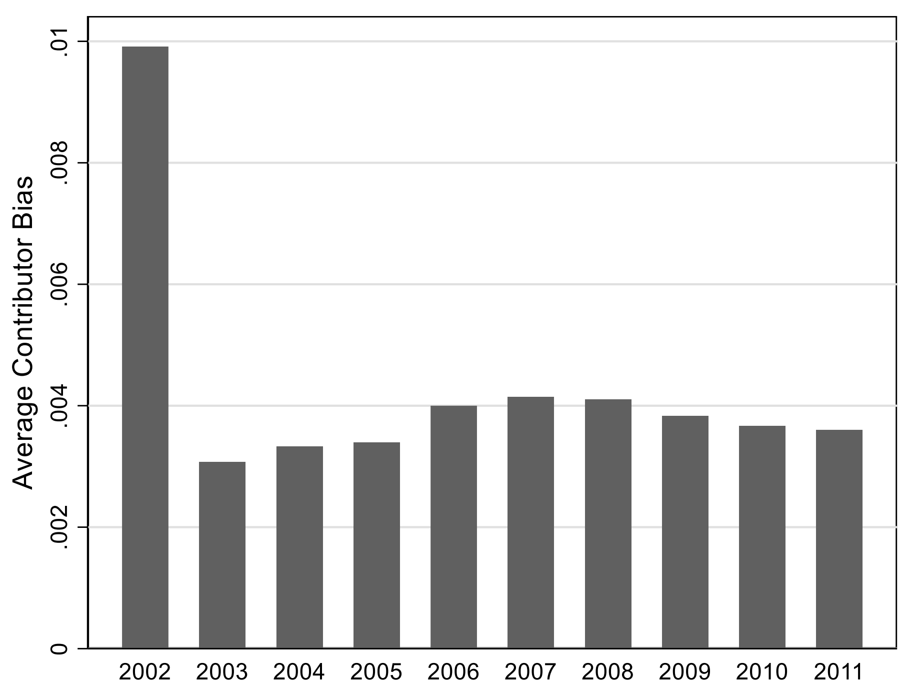


Figure 7: Average Number of Edits over Contributors' Years on Wikipedia, Extreme Contributors Only

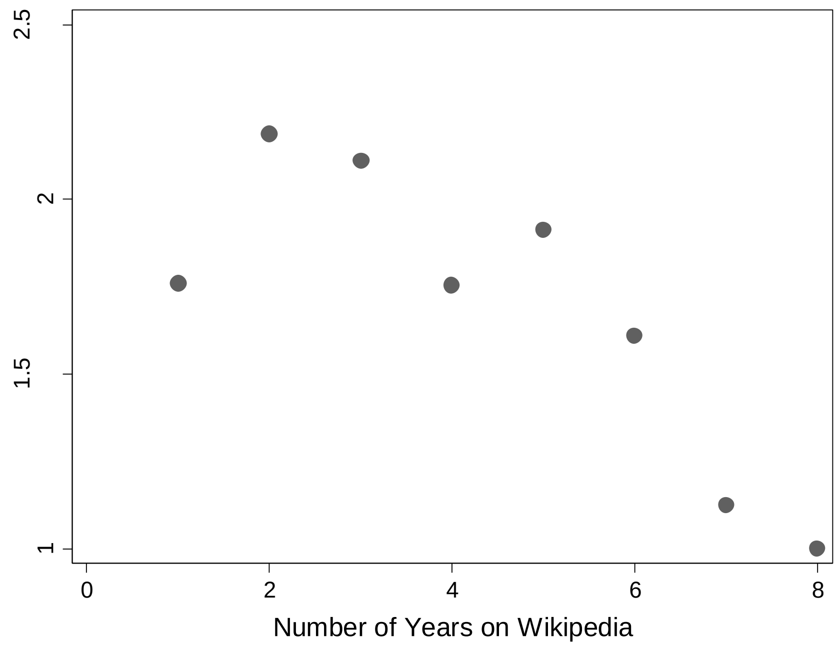
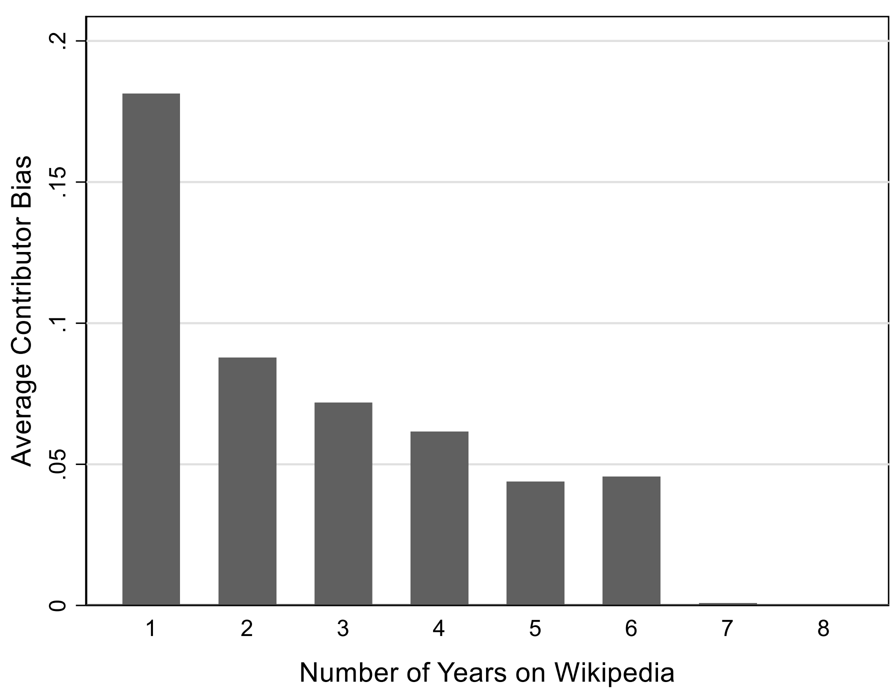


Figure 8: Average Contributor Bias Each Year over Contributors' Years on Wikipedia, Extreme Contributors Only



Notes: In the bar chart, the average bias in year seven is 0.0007 and in year eight is 0.

Online Appendix for "Ideology and Composition among an Online Crowd: Evidence from Wikipedians"

Table A1: Logit Regressions on the Relationship between Contributor Category and Prior Article Category

Model	(1)		(2)		(3)	
Dependent Variable	Contributor Category=-1	Contributor Category=1	Contributor Category=-1	Contributor Category=1	Contributor Category=-1	Contributor Category=1
Prior Article Category	2.04146*** [0.0266]	-2.4396*** [0.0136]	2.0562*** [0.0270]	-2.3682*** [0.0133]	2.0830*** [0.0271]	-2.3137*** [0.0132]
Log(Prior Article Length)			-0.0569*** [0.0045]	0.1245*** [0.0052]	-0.0320*** [0.0051]	0.1639*** [0.0058]
Log(Prior Refs)			-0.2184*** [0.0032]	-0.3067*** [0.0030]	-0.2908*** [0.0042]	-0.4112*** [0.0040]
Year FE	No		No		Yes	
Article FE	No		No		Yes	
Observations	9,585,443		9,585,443		9,585,443	
Pseudo R-squared	0.021		0.039		0.045	

Notes: The sample is the same as the main analysis sample in Table 3. *Contributor Category* is the categorical version of *Contributor Slant*, which takes the value of -1, 0, or 1, representing contributors with a slant two standard deviations below mean, in between, and above mean, respectively. *Prior Article Category* is the categorical version of *Prior Article Slant*, which takes the value of -1, 0, or 1 representing articles with a slant two standard deviations below mean, in between, and above mean, respectively. Again, we find that the coefficients for the categorical explanatory variable *Prior Article Category* is negative and significant in all cases, suggesting that the slant category of the next contributor is significantly negatively correlated with the slant category of the prior article. Robust standard errors in brackets. *significant at 10%; ** significant at 5%; *** significant at 1%.

Table A2: Regressions on the Relationship between Percentage of Republican in the Area and Prior Article Slant

Model	(1)	(2)
Dependent Variable	RepPerc	RepPerc
Prior Article Slant	-0.0009**	-0.0010**
	[0.0004]	[0.0004]
Log(Prior Article Length)		0.0037***
		[0.0001]
Log(Prior Refs)		0.0005***
		[0.0001]
Observations	2,438,628	2,438,628
Adjusted R-squared	0.000	0.001

Notes: Robust standard errors in brackets. *significant at 10%; ** significant at 5%; *** significant at 1%.

Table A3: Relationship between Contributor Slant and Prior Article Slant, First Edits Only

Models	(1)	(2)
Dependent Variables	Contributor Slant	Contributor Slant
Prior Article Slant	-0.0100***	-0.0236***
	[0.0001]	[0.0005]
Log(Prior Article Length)	0.0007***	0.0011***
	[0.0000]	[0.0001]
Log(Prior Refs)	-0.0005***	-0.0012***
	[0.0000]	[0.0001]
Observations	6,560,812	6,560,812
R-squared	0.008	0.008
Year FE	No	Yes
Article FE	No	Yes
Number of Articles	65,361	65,361

Notes: Robust standard errors in brackets. *significant at 10%; ** significant at 5%; *** significant at 1%. Observations in this panel only include every contributor's first edit of an article.

A4: Additional Robustness Checks on EC vs. Non-EC Effect

We also conduct several additional robustness checks to make sure the Non-EC effect is not driven by alternative explanations. First, our slant index is measured on the basis of frequently used phrases, or code phrases, favored by party representatives. It may be the case that longer articles tend to contain more code phrases and are therefore more measurable. In this case, long articles could drive our results. To rule out this explanation, we eliminate outlying long articles from our full sample, that is, articles that are more than two standard deviations above the mean article length. We obtain similar results.

Second, the articles whose titles contain code phrases might tend to show greater biases in our sample simply because these code phrases are more likely to be used repetitively in the article content. To check our findings against this concern, we exclude from our sample all articles whose title contains code phrases, which is 1.77% of all articles. Again, we find a significant Non-EC effect from the results.

Third, it is possible that certain code phrases are chosen simply because these words do not have other commonly-used synonyms that are neutral or of the opposite slant. In this case, as our measure captures the contributor's choice of words describing the same concept for a given topic, one's contribution may be slanted merely because he or she could not find neutral substitutes of the code phrases to choose from. We rely on the experiences of a legal and copyediting professional to identify these instances in our dictionary and leave only code phrases with natural substitutes. After re-measuring the slant index for articles and contributors, we repeat our analyses and find no significant change in our results. Therefore, the Non-EC effect is not driven by instances where contributors do not have a choice for substitute phrases.

Fourth, because contributors' edits to popular articles tend to have greater impact than those to less popular ones, their political slants measured from these popular articles could carry more weight. Therefore, we use articles' page views as weights when computing the average contribution slant and repeat our analysis using the weighted contributor slant. We continue to find significant Non-EC patterns.

We are also concerned that contributors blocked by Wikipedia administrators may affect our results.[29] These contributors may create extremely biased content initially and drop out of the dataset after being blocked. As a result, contributors overall may become more neutral over time. This problem is mitigated by our approach of assigning missing values to *Contributor Yearly Slant* when a contributor makes no edits in a year. As a robustness check, we repeat our analysis after dropping all 56,329 contributors who have ever been blocked (temporarily or permanently) and the associated 480,960 edits from our sample. Again, the results remain unchanged.

[29] Blocks are used to prevent damage or disruption to Wikipedia. Contributors may be blocked for reasons such as vandalism and edit warring. See https://en.wikipedia.org/wiki/Wikipedia:Blocking_policy for the detailed policy, accessed August 2017.

Finally, we test if the Non-EC effect is driven only by extremely slanted articles. We eliminate from our full sample articles with slant index two standard deviation points away from the mean. Changing this threshold to articles without slant in the top and bottom 10% does not differ qualitatively in results. The estimated coefficients with subsamples have the same signs but larger absolute values. We also conduct a robustness check that includes only contributors whose slant is not zero, and we continue to observe a Non-EC pattern among them.

A5: Procedure for Computing Slant Index

In G&S, for each congressperson *c*, they observe its ideology y_c and phrase frequency f_{pc}, the number of times phrase *p* appears in congressperson *c*'s speech, for each phrase *p*. For each phrase *p*, G&S regress the relative frequency $\overline{f_{pc}}$, where $\overline{f_{pc}} = f_{pc} / \sum_{p \in P} f_{pc}$, on y_c, and obtain the intercept and slope parameters a_p and b_p, for each phrase *p*.[30]

The 1,000 phrases exhibit heterogeneous slant. To mitigate the effect of outlier phrases (e.g., "African American" and "illegal immigration"), we set the parameter values for the 9 most left-leaning phrases and 9 most right-leaning phrases to be the same as the 10th most left-leaning phrase and the 10th most right-leaning phrase, respectively.

For each Wikipedia article n, we regress $\overline{f_{pn}} - a_p$, where $\overline{f_{pn}}$ is the relative frequency of phrase *p* in the article, on b_p for the 1,000 phrases to obtain the slope estimate $\overline{\mathrm{Y}_\mathrm{n}} = \frac{\sum_{p \in P} b_p (\overline{f_{pn}} - a_p)}{\sum_{p \in P} b_p^2}$. When an article has none of the 1,000 phrases, $\overline{\mathrm{Y}_\mathrm{n}}$ is 0.4975. We denote $\mathrm{Y}_\mathrm{n} = \overline{\mathrm{Y}_\mathrm{n}} - 0.4975$ and use Y_n as our bias index for article n.

[30] The parameter values, together with the 1,000 phrases, are available at http://www.icpsr.umich.edu/icpsrweb/ICPSR/studies/26242, accessed March 2019.