Threatening Event, National Identity and Network Dynamics of Motivated Information Communication:
Exploring Japanese Twitter during the Rise of Territorial Disputes, April through October 2012*

Gento Kato[†1], Takanori Fujiwara[2], Christian Collet[3], Tetsuro Kobayashi[4], and Takafumi Suzuki[5]

[1]*Department of Political Science, University of California, Davis*
[2]*Department of Computer Science, University of California, Davis*
[3]*Department of Politics and International Studies, International Christian University*
[4]*Department of Media and Communication, City University of Hong Kong*
[5]*Department of Media and Communications, Toyo University*

August 11, 2017

Abstract

Lab-experiment and survey based studies find that intergroup threat and group identity have significant implications on the formation of political attitudes, that those factors encourage individuals to approach information with directional motivation. On the other hand, unrealistic assumptions prevent those findings to be directly generalized to the real-world context. In this study, we use novel twitter network data during rising territorial disputes in Japan to capture real-world information communication process under intergroup threat. In our data, twitter users communicate information through the network of retweets and web-links. This "information communication network" is examined from both individual-level and society-level perspectives. We find partial support to the previous laboratory and survey based findings. In terms of retweets, it is confirmed that threatening territorial dispute incidents and salient national identity contribute to the increase in individual-level reliance on national identity holders. In the society level, after the incident, identity holders gain more influence in the network than no identity holders on average, but less influence in total. In terms of web-links, links to both opinionated and factual domains increase after the threatening incident, but differ in the timing and persistence of impact. While the impact on factual domains is immediate and short-lasting, the impact on opinionated domains is late and long-lasting. By observing dynamic behavior of inter-dependent individuals in the network, current findings give new insights to the study of intergroup threat, group identity, and motivated reasoning.

*Paper prepared for presentation at the 113th American Political Science Association Annual Meeting, San Francisco, CA, August 31, 2017.

†gento.badger@gmail.com

1 Introduction

In explaining attitudes toward different groups in society, threat and identity are two key factors. Numbers of lab-experiment and survey based studies find that rise in the perception of intergroup threat (Stephan, Diaz-Loving and Duran, 2000; Stephan and Mealy, 2011) and group identity (Tajfel, 1974; Tajfel and Turner, 1979) deteriorates ingroup members' attitudes toward outgroup. The theory and evidence from motivated reasoning (Kunda, 1990) further confirm those findings by suggesting that threat and identity encourage the directional processing of encountered information, which helps to form hostile attitudes toward outgroup.

While evidence is robust, lab-experiment and survey based approaches have limited ability to explain information process in the real-world society. Lab experiment may isolate the specific causal structure in the real-world process but are limited by the small sample size of participants. The mechanism revealed from small-scale samples may not function the same in large-scale population. An Individual-level survey does scope for a larger population, however, it often misses the aspect of *social interaction*. In the real society, it is difficult to assume that the information is processed independently, without communications between information senders and receivers.

To overcome limitations in previous findings, we utilize *Twitter* to analyze information process. First, Twitter provides the real-world data of large-scale population (i.e., twitter users). This nature makes it possible to analyze the society as a whole, not as a simulated small subset. Second, Twitter involves variables to capture information *communication* process. From retweets and web-links, we can draw the network structure of how twitter users interact with each other to communicate information and form attitudes[1]. In addition, Twitter is particularly useful in tracking *dynamic* process. In contrast to surveys or lab experiments which can only be conducted in short time-span and/or discrete points in time, time-stamped tweets provide data to construct long time-span, continuous process of information communication and attitude transformation.

In the current study, we collect Japanese tweets data on Takeshima and Senkaku territorial disputes. Japan-South Korea (Takeshima) and Japan-China (Senkaku) territorial disputes continue for more than 40 years. Disputes have been threats to the Japanese public and discussed to be sources of hostility toward South Korea and China. The full set of 4,086,539 tweets relevant to territorial disputes is collected from 2012 April through October. During this period, tensions in both disputes rose with the occurrence of major incidents over the disputed territory. To assess threat and identity, we capture the occurrence of threatening event (i.e., territorial dispute incidents) by detecting the jump in tweet frequency, and use machine learning technique to train and predict expression of national identity in user profiles. Our analysis explores twitter network patterns in relation to the timings of territorial issue incident and identity salience in user profiles.

The following sections proceed as follows. Next section describes the theoretical connections between threat, identity, and motivated reasoning. Then, section 3 explains the nature of our case. Section 4 introduces major variables of interest and the structure of twitter network data. Section 5

[1]It is important to note here that the provided data consist of *all* relevant tweets, not *sample* of tweets. Because, with a sample of data, network analysis may miss the critical piece of information.

generates specific hypotheses for this study, and section 6 and 7 provide results of analysis. Finally, section 8 summarizes findings and discusses the significance of the current project.

2 Theory: Threat, Identity, and Motivated Reasoning

To explain the dynamics of intergroup behavior, threat and identity are considered as critical factors to drive ingroup members to hold negative attitudes toward outgroup. To start with, a threat from outgroup comes in various forms. Intergroup threat theory (Stephan and Stephan, 2000; Stephan and Mealy, 2011) identify two broad categories, which they call *realistic* and *symbolic* threats. A realistic threat is represented by threats to materialistic power, resource, and welfare of a group or an individual, while a symbolic threat is described as threats to values, belief system, ideology, and worldview of a group or an individual. The theory predicts that perception (does not have to be real) of those threats induce ingroup members to develop hostile attitudes or behaviors toward the outgroups. More specifically, it is argued that symbolic threat leads more to reduced empathy, conformity to group value, and even vicious behavior such as genocide or torture, while realistic threat induces a more pragmatic response to cope with the threat (Stephan, Diaz-Loving and Duran, 2000). Lab experiment and survey based studies confirm expectations in various intergroup contexts such as White-Black race (Stephan et al., 2002), Christian-Muslims (Velasco González et al., 2008), domestic-international students (Charles-Toussaint and Crowson, 2010; Harrison and Peacock, 2010), native-immigrants (Stephan et al., 1998; Stephan, Diaz-Loving and Duran, 2000) and more others (for review, see Riek, Mania and Gaertner, 2006).

Group identity is a closely related but distinctive factor to influence intergroup attitudes. Social identity theory (Tajfel, 1974; Tajfel and Turner, 1979) claims that the social behavior of individuals with stronger group identity is determined more by their group attributes (e.g., nation, ethnicity) than by their personal attributes or preferences. The theory further implies that "the mere perception of belonging to two distinct group is sufficient to trigger intergroup discrimination favoring the in-group" (Tajfel, 1974, 38). In the context of national or ethnic identity, the direct negative relationship between identity strength and intergroup attitudes are confirmed in numbers of lab-experiments (e.g., Druckman, 1994; Yogeeswaran and Dasgupta, 2014) and survey based studies (e.g., Corkalo and Kamenov, 2003; Li and Brewer, 2004).

In addition to its direct influence, group identity has been discussed as both antecedent and moderator of intergroup threat. First, intergroup threat theories point out that those with strong ingroup identity are susceptible to outgroup threat (Riek, Mania and Gaertner, 2006; Stephan and Mealy, 2011). Verkuyten (2009) and Brylka, Mähönen and Jasinskaja-Lahti (2015) apply path model to the representative sample survey data of Finns and Dutch and confirm this expectation. They find that the impact of national identity on negative intergroup attitudes (i.e., anti-multiculturalism and anti-immigrant sentiments) are mediated by the perception of outgroup threat. Second, ingroup identity moderates the influence of threat on attitudes and behavior. Using the questionnaire on Israeli undergraduates, Bizman and Yinon (2001) find that the relationship between realistic threat perception and negative intergroup attitudes is statistically significant only among those who have strong Israeli identity[2]. In another study with the non-Asian American student and university staff sample in the United States, Morrison and Ybarra (2008) take an experimental approach to manipulate threat. In the experiment, participants are assigned to either threat-inducing question or non-threatening question on the perception of Asian-American. They find that only among strong racial group identifier, intergroup attitudes (i.e., social dominance orientation) under the threat condition is significantly more negative compared to the no-threat condition.

Motivated reasoning (Kunda, 1990) claims that individuals have either *accurate* or *directional* motivations to process new information and form attitudes. With accurate motivation, individuals process information to obtain "correct" understandings of the subject, thus acquire and evaluate information without bias. On the other hand, with directional motivation, individuals acquire more information consistent with their motivation (e.g., previous relevant attitudes), and spend more cognitive resource and time to devalue information inconsistent with their motivation. The expected process is confirmed in carefully-designed lab experiments (Taber and Lodge, 2006A; Lodge and Taber, 2013; Druckman, Fein and Leeper, 2012). Also, survey (Fischle, 2000) and computer simulation (Kim, Taber and Lodge, 2010; Lodge and Taber, 2013) based studies show that surveyed or simulated information communication and attitude formation process is more consistent with expectations from motivated reasoning model than with expectations from alternatives that assume no motivation. In the context of intergroup attitude formation, motivated reasoning provides a cognitive mechanism behind the connections between threat, identity, and negative intergroup attitudes. Threat and identity are supposed to give directional motivation for people to more frequently acquire negative outgroup information and devalue positive outgroup information.

In sum, numbers of lab-experiment and survey based studies find that intergroup threat and group identity are significantly related to negative intergroup attitudes. In addition, theory and evidence from motivated reasoning suggest that threat and identity are influencing intergroup attitudes through the biased processing of new information. However, as we already noted, there are substantive limitations to the approaches in previous studies. First, lab-experiment designs prevent the direct application of findings to the real-world. Society is more complex and larger than the specified simple and small-scale process in the lab. Even when the experiment is successful in isolating the factor of certain behavior, it does a poor job in explaining the phenomena occurring in real society. Second, survey based studies may solve the problem of generalizability, but often ignore the social interaction aspect of attitude formation[3]. By design, surveys assume that individuals encounter information from independent external sources, and process it in isolation. In reality, this is not the case. People live with social networks, and information sender and receivers are interdependent with each other.

In the current study, we overcome limitations in previous studies by collecting twitter postings during the rise of territorial disputes in Japan. Before stepping into the analysis, following two sections start from describing the case (section 3) and the structure (section 4) of our dataset.

[2] Their study finds no moderation role of identity in the relationship between symbolic threat and negative intergroup attitudes.

[3] Snowballing survey is one way to overcome this limitation. However, it is still a sampled portion of the society, and one cannot reject the possibility of missing critical elements in the societal interaction.

3 Case: Takeshima and Senkaku Territorial Disputes in Japan, April-October 2012

Japan has been involved in the territorial dispute with South Korea and China for more than 40 years. With South Korea, it is a dispute over Takeshima (Korean name: Dokdo), and with China, it is a dispute over Senkaku (Chinese name: Diaoyu) islands. Our study focuses on April through October 2012, when those territorial disputes received significant attention from the Japanese public. During this six months period, several major incidents occurred around the disputed territories. For Takeshima, South Korean president Lee Myung-bak visited the disputed island in August. Even when Takeshima is practically controlled by the South Korean army, this is the first time for the current president of South Korea to land the territory. His action invites strong responses from the Japanese public. For Senkaku, in April, the governor of Tokyo metropolitan government announced his plan to officially acquire Senkaku islands (which had private Japanese owner at that time), which triggered the active reaction from China. While disputed islands are practically controlled by Japan Coast Guard, Chinese activists from Hong Kong attempted and succeeded to land disputed islands in August, and Chinese military boats repeatedly enter territorial sea around those islands. In September, Japanese national government decided to nationalize Senkaku islands by buying those territories from the previous owner.

During the six months, attitudes of Japanese people toward South Korea and China are significantly deteriorated. Figure 1 illustrates this tendency. It shows the result from monthly public opinion poll with the nationally representative sample of Japanese people, conducted by *Jiji Press*. The survey asks respondents to choose up to three countries they like/dislike the most, and scores in the figure indicate differences between proportions of people liking the country and disliking the country. Lines clearly show that the favorability scores toward South Korea and China declined significantly during April through October (the shaded area). Especially from July to October, scores drop by over 30% for South Korea and about 10% for China. In addition, through 2013, those dropped scores never return to the previous level in 2011. There have been a clear structural change in ingroup (i.e. Japanese) perceptions toward outgroups (i.e., South Korea and China).

There are two reasons as to why this case is an appropriate occasion to test theoretical implications of intergroup threat and group identity. First, given its zero-sum game nature, incidents over disputed territory can be considered as the *realistic* threat to Japanese people. Previous studies show that territorial conflict is more influential of people's attitudes than other types of conflict. Using World Value Surveys, Hutchison, Gibler, and others find that the past occurrence of militarized territorial disputes decreases the political tolerance towards unfavorable social group (Hutchison and Gibler, 2007) and strengthens national identity (Gibler, Hutchison and Miller, 2012), while other types of militarized interstate disputes do not.

On the other hand, Takeshima and Senkaku disputes never actually developed into real "conflict" involving casualty and the use of a weapon. Also, the dispute is occurring in small islands at the peripheral of the country. Here, Tanaka (2015) conducted survey experiment on Japanese people to assess if the distance from the disputed territory (i.e., Takeshima) influences people's attitudes.

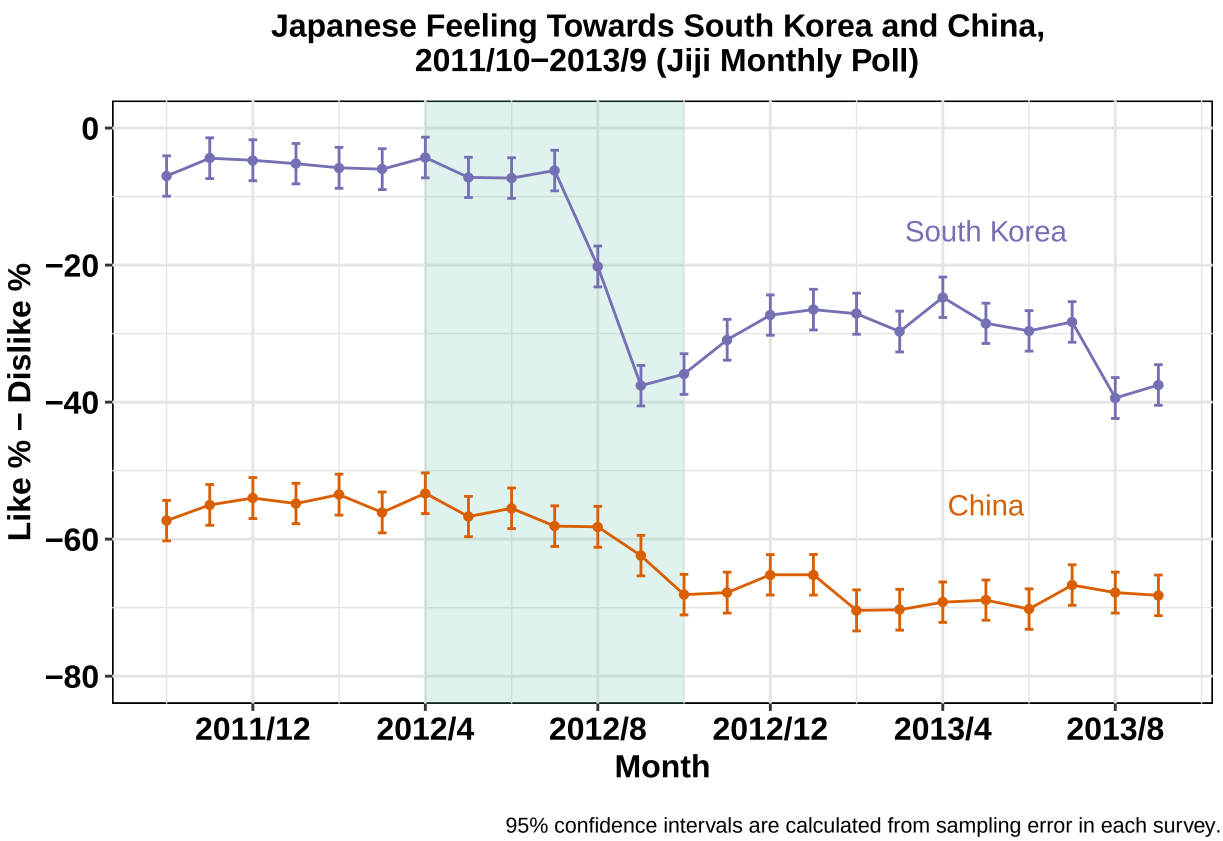

Figure 1: Over-time Feelings of Japanese People toward South Korea and China, Jiji Monthly Poll 2011-2013 (Shaded Area Indicates the Period of Data Collection)

He finds that those people living further away from Takeshima are generally less compromising on Takeshima issue, and more responsive to the nationalistic frame of the issue. In other words, people living away from the disputed territory are not being practical to resolve the dispute. Thus, for most of the Japanese people, the territorial threat may not be real but *symbolic*. Consistent with this conceptualization, historical studies of East Asian relations often suggest that Takeshima and Senkaku disputes are driven by "symbolic significance for domestic political conflict, rather than by a desire for natural resources" (Deans, 2000).

Either *realistic* or *symbolic*, Takeshima and Senkaku territorial disputes pose significant threat to Japanese people. Historical studies of Japanese national identity repeatedly note that Takeshima and Senkaku disputes are closely connected to the rise in Japanese nationalism and national identity (Sasada, 2006; Rozman, 2013; Suzuki, 2015). Then, in 2012, the rise in Takeshima and Senkaku disputes led to the dramatic decline in Japanese attitudes toward China and South Korea. The task of the current study is to explore the mechanism behind this attitude transformation. We use implications from motivated reasoning to explore the mechanism. More specifically, we focus on the network of information communication processes to reveal the role of threatening events (i.e., incidents in disputed territories) and strong national identity holders. The next section describes variables and the structure of twitter information communication network data.

4 Data

With the rise of Internet, it becomes common for ordinary people to express their political views through the web and share them with many and unspecified others. Established in 2006, Twitter is one of the major platforms for this new type of opinion reporting, which is known as "microblogging." By only allowing up-to 140 characters on each message (tweet), Twitter can be seen as the media to represent short summaries of users' political views. In addition to one's own view, twitter users actively spread information of others through *retweets* (replies)[4] and web-links. The retweet is the mechanism to respond to other twitter users, represented by @ followed by the user name. Twitter is considered to be the highly influential opinion channel in today's political sphere. Recent studies in the United States and Europe show that Tweet texts are predictive of real-world political change, such as election outcomes (Tumasjan et al. 2011; Bermingham and Smeaton 2011; Ceron, Curini and Iacus 2015 but Gayo-Avello 2012; Murthy 2015), the popularity of political leaders in representative surveys (Ceron et al., 2014) or even stock market (Bollen, Mao and Zeng, 2011).

In Japan, Twitter is one of the most popular platforms of microblogging, or more broadly, social networking service. In 2012, the public opinion survey shows that 15.7% of whole public use Twitter, and this proportion rose to 37.3% for the age of 20s. In addition to young users, compared to other platforms of social networking service (e.g., Facebook, LINE), Twitter is more widely acquired by older generations (Ministry of Internal Affairs and Communications, 2012, 61). Twitter community covers the significant portion of political opinion sphere in Japan. To capture the

[4]Technically, "retweet" and "reply" are separate terms. Given the similarity in their usage, however, both are treated as "retweet" in this study.

political opinion flow in Japanese twitter community under threat, we obtain the total of 4,086,539 Japanese tweets from *Twitter inc.* during the time of rising in territorial disputes in 2012. The data consist of *full* set of tweets with the keywords of 竹島 (*Takeshima*) and 尖閣 (*Senkaku*) posted between April 15th and October 14th 2012[5]. In further considerations, all tweets with words 竹島, 独島 (*Dokdo/Dokuto*: Korean name for Takeshima), 韓国 (*Kankoku*: South Korea) and 朝鮮 (*Chosen*: Korea) are treated as relevant to Takeshima issue, and all tweets with words 尖閣, 魚釣 (*Diaouyu/Uotsuri*: Chinese name for Senkaku) and 中国 (*Chugoku*: China) are treated as relevant to Senkaku issue. Note that there may be an overlap between two issues if a tweet includes keywords from both issues.

Tweets come with both tweet and user level variables, including: *tweet texts*, *time of tweet post*, *user name*, *time of user creation* and *user profile texts*. In this section, we use those variables to capture threatening event, measure national identity, identify web-domain characteristics, and construct twitter network.

4.1 Territorial Dispute Incident: Capturing Threat through Tweets Frequency

As the perception of threat increases, individuals are expected to respond more strongly (Stephan and Mealy, 2011). Therefore, we expect the frequency of territorial dispute relevant tweets to increase as the threat perception increases. Figure 2 confirms the expectation. Peaks in tweet frequency correspond almost perfectly with dates of major incidents occurred in disputed territories. For Takeshima, the peak in the tweet frequency starts on August 10th. On this day, South Korean president Lee Myung-bak visits the disputed territory, which is the most important incident for Takeshima dispute in this period. For Senkaku, the frequency has at least two major peaks. The first, highest peak starts on August 15th, when Hong Kong activists successfully land the disputed territory. The timing of the second peak is somewhat ambiguous at around September 18th, but it roughly corresponds with the time when groups of Chinese fishing boats are reported to enter territorial sea around disputed islands.

In the subsequent analysis, we mainly focus on the most significant territorial dispute incident – one comes with the most significant change in tweet frequency. Here, we use *changepoint* package (Killick and Eckley, 2014) in R to detect exact timing (in minute level) of the single changepoint in the tweet frequency. It captures the point in distribution with the largest change in variance. The detected timings are shown as vertical purple lines in Figure 2. Specifically, the most significant *territorial dispute incident* for Takeshima is 12:05 am August 10th, for Senkaku is 3:58 pm August 15th. While the detected timing may not be the unique incident in the period, we expect the most significant change in information communication pattern to occur in this timing.

[5] More specifically, tweets are extracted from Sunday, April 15th 0:00 GMT through Saturday, October 13th 24:00 GMT (equivalent to Sunday, April 15th 9:00 JST through Sunday, October 14th 9:00 JST.)

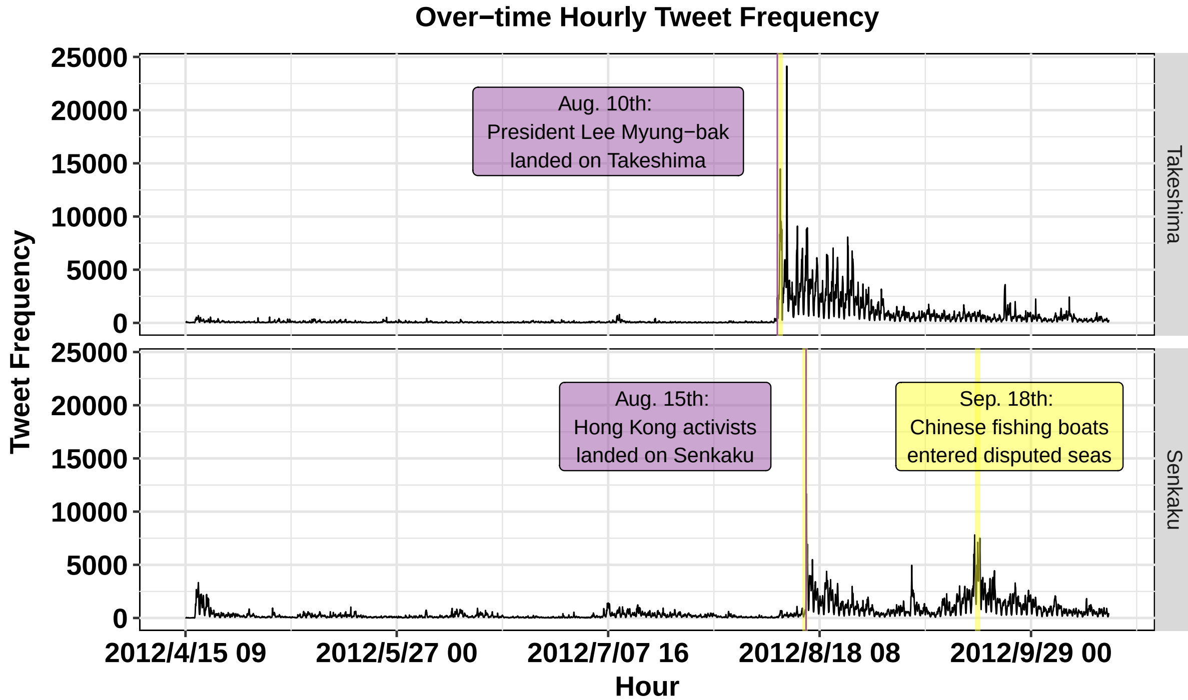

Figure 2: Hourly Frequency of Tweet Post and the Timing of Territorial Dispute Incident

4.2 National Identity: Machine-Learning of Patriotic and Right-wing User Profiles

In addition to a threat, another theoretically important variable is a group identity (i.e., Japanese national identity in the context of this study). Here, twitter data come with informative user profile texts, but data are too big to manually code every profile. Therefore, we utilize machine learning method to automatically classify all profiles[6] using sampled set of manually coded profiles.

As the first step, Three political science major graduate school students are hired as coders, and manually code randomly sampled 1000 user profiles. Here, national identity is formally described as the "pervasive sense of subjective attachment to the nation" (Huddy and Khatib, 2007, 65), and there are carefully designed multi-question methods to measure it through a survey. On the other hand, there is no established method to extract a national identity from user profile texts. In general, national identity is often discussed with two related but distinctive aspects of national attachment: *patriotism* and *nationalism* (Druckman, 1994; Karasawa, 2002; Li and Brewer, 2004). First, patriotism is expressed as the love and proud of the nation[7]. Studies find that the measure

[6]Note that some profiles in the data are empty. Those blank profiles are considered in the separate category in the subsequent analysis.

[7]Note that there are different ways to define patriotism. The definition here is often described as *symbolic* patriotism.

of national pride sentiments correlate strongly with the measure of pure attachment to the nation, but considered independent from political ideology (Huddy and Khatib, 2007). To capture this first aspect, coders are asked to code "the expression of a sense of belonging to Japan." Second, nationalism is the concept of national attachment that is closely connected to political ideology. In Japan, *right-wing nationalism* is often described as a broader concept that embraces national identity (Mccormack, 2000; Rose, 2000), while Karasawa (2002) finds that "attachment to the ingroup and ethnocentrism" belong to "separate dimensions" of factor analysis (645). Thus, nationalism is a related but a distinctive concept of patriotism. Coders are instructed to code this second aspect by "the expression of right-wing ideology." Krippendorf's alpha (Hayes and Krippendorff, 2007) for manual-codes are 0.789 for patriotism and 0.797 for right-wing nationalism[8]. Both scores go above the recommended threshold of 0.7, thus considered to be reliable.

Then, to machine-learn patriotism and right-wing nationalism in user profiles, this study uses random forest (RF) classifier (Breiman, 2001). This method was initially utilized in the field of bioinformatics (e.g. Cutler and Stevens, 2006) but recently been applied to texts. Even when applications are not many, for Japanese texts, Jin and Murakami (2007) suggests that performance of RF is generally better than other popular machine-learning methods to classify authorships of texts. In addition, RF also has an ability to calculate each variable's level of contribution to the classification, which cannot be produced by other methods. The RF classification proceeds as follows. First, in the training data with 1000 sampled profiles, rows represent profiles and columns represent uni-grams (i.e., dummy appearance of words) in profiles[9]. Then, we start with boot-strapping the original data matrix $M_{i,j}$ 500 times with replacement. From those bootstrapped samples, two-thirds are used for classification, and one-third are kept for test the model (out-of-bag test). Then, from each sample to be used for classification, we extract random subsets of $\sqrt{j}$ variables (uni-grams). Next, by the Gini index shown in below, we construct unpruned decision tree in each of replicated data matrix with reduced uni-grams:

$$GI = 1 - \sum_{c=1}^{n} [p(c|x)]^2 \tag{1}$$

In the above equation, $p(c|x)$ indicates the probability of x (a text with reduced uni-grams) belongs to c (class) (Suzuki, 2009). With the majority of votes by $p(c|x)$ of all x, new classifications is given to each text.

To understand the contribution of each variable in classification, this study utilize variable importance measure denoted as VI_{acu} for each variable m. It is calculated by the following formula

[8] Nationalism is originally a five category code of ideology, but due to low variability and low frequency of left-wing ideology, it is recoded to the dummy variable of representing right-wing ideology/nationalism or not.

[9] Words in profiles are identified by Japanese morphological analysis system, *MeCab*. The morphological analysis is conducted by *RMeCab* (`http://rmecab.jp/wiki/index.php?RMeCab`), developed by Motohiro Ishida.

(Suzuki, 2009):

$$VI_{acu} = \frac{\text{Mean}(C_{oob} - C_{per})}{\text{Standard Error}} \tag{2}$$

C_{oob} indicates the number of votes cast for the correct class in out-of-bag samples using all variables and C_{per} shows the number of votes cast for the correct class when m variables are randomly permuted in the out-of-bag samples. Intuitively, VI_{acu} represents the importance of the permuted variable (uni-gram/word). It shows the degree to which the classification lose its accuracy when word m is replaced by another word.

Table 1 presents the result for the machine learned classification using RF method. To evaluate the results, we conduct RF classification for 100 times with 950 cases of training data and 50 cases of test data. The table shows the median values of precision (the proportion of machine-learned codes coinciding with original codes), recall rates (the proportion of original codes coinciding with machine-learned codes) and F_1 values [10]. The result indicates that RF predictors are highly effective in predicting patriotism and right-wing nationalism in user profiles. The number of codes for right-wing nationalism is slightly underestimated (high precision but low recall rate), but for patriotism, the prediction performs well for both precision and recall rate.

Table 1: Results of Classification Experiment of User Profiles

Variable Name	Precision	Recall	F_1
Patriotism	0.8333	0.8944	0.8496
Right-wing Nationalism	1.0000	0.6000	0.7143

Table 2 shows the 10 most important words in the classification[11]. While two lists are similar, differing patterns confirm the original intentions of coding. Attachment-oriented words like 誇り (*hokori*: pride) and 文化 (*bunka*: culture) are being important only for patriotism classification, and ideology-oriented words such as 保守 (*hoshu*: conservative) are being important only for right-wing nationalism classification.

Finally, using predictions from trained RF classifier, all profiles in the dataset are machine-coded. Three RF predictions are generated for each of the variables, the final code is the one which got the majority in three predictions[12]. As a result, out of 3702966 tweets with non-empty user profiles[13], 15.56% (576198) are coded as holding strong patriotism, and 9.85% (364643) are coded as having strong right-wing nationalism (note that two identities are *not* mutually exclusive). In the analysis,

[10] F_β value is calculated by following equation: $\frac{(\beta^2+1) * Precision * Recall}{\beta^2 * Precision + Recall}$.

[11] The results are extracted from the first of three classifications with full 1000 cases used as training set. As explained later, those three classifications are used to predict codes in full dataset.

[12] Some user profiles consist only of special characters (i.e. emoji (picture characters), languages other than Japanese or English), that cannot be analyzed by MeCab. Those profiles are coded as holding no identity.

[13] 383573 profiles are empty, thus excluded from the machine-learning target.

Table 2: Top 10 Important Words in Classification of User Profiles

	Patriotism			Right-wing Nationalism		
	Word	Translation	VI_{acu}	Word	Translation	VI_{acu}
1	日本	Japan	0.0866	日本	Japan	0.0162
2	日本人	Japanese	0.0227	民主党	Democratic Party of Japan	0.0142
3	国	country/nation	0.0062	保守	conservative	0.0101
4	誇り	pride	0.0054	日本人	Japanese	0.0086
5	反日	anti-Japan	0.0053	反日	anti-Japan	0.007
6	文化	culture	0.0031	韓	*Kan* abbreviation for S.Korea	0.0059
7	国土	national territory	0.0031	政権	political administration	0.0057
8	民主党	Democratic Party of Japan	0.0028	勢力	power/influence	0.0041
9	危機	crisis	0.0023	国土	national territory	0.004
10	感謝	gratitude/thanks	0.0019	国	country/nation	0.003

a user is considered to have an identity if at least one profile is coded as having identity[14]. Also, the results for two types of identity are presented separately to make comparisons.

4.3 Factual vs. Opinionated: Characteristics of Web-Link Domains

In web-link URL, website domain (e.g., the part "`http://www.asahi.com`" in the URL "`http://www.asahi.com/articles/...`") contains important characteristics of the information provided by the specified source. In this study, domains are categorized by expected factual-opinionated nature of information in the website. By *factual*, we mean that those websites tend to provide information that is fact or neutral source based. By *opinionated*, we mean that those websites tend to provide information that is impression/opinion based. Figure 3 illustrates four categories of domains following this factual-opinionated spectrum. First, official websites of quality newspapers and TV news stations are coded as *hard news*, which is expected to be the most factual source of information relative to other domains. Second, on the other extreme of the spectrum is *personal media*. This category includes domains of personal websites, blogs, social networking services, and internet forums. In those websites, information comes in the form of impression or opinion towards the issue, and often the case, not backed up by validated factual evidence. Remaining two categories are considered to be in between the above two. *Soft news* includes domains of news media with web-only outlets and entertainment newspapers. These websites may provide more rigid and evidence based arguments than personal media, but often more opinionated or emotional than those websites included in the hard news category. *Curated news* represents domains of news curation websites (e.g., Yahoo! or Google News). These websites do not provide their own articles but collect news from both hard news and soft news websites. Therefore, their news characteristics are expected to be in between hard news and soft news.

[14]We only consider profiles created before the territorial issue incident to capture user-level identity status. This procedure code those users who change their profiles to identity-salient after the incident to be no identity holder. Since we want to capture consistently salient status of identity, it eliminates the possibility that our identity measurement is induced by the threatening incident.

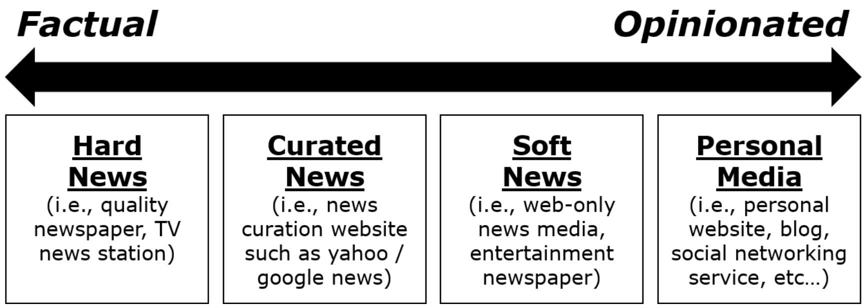

Figure 3: Categorization of Web-Domains

Using the above coding scheme, we manually categorized domains. To focus on the significant information sources, only the most frequently linked 500 domains are coded, which cover 95.06% of all cumulative URLs appeared in the full dataset (The full set of domain codes are provided in Appendix A)[15]. Out of 500 domains, 55 are hard news (cover 28.05% of all cumulative URLs from 500 domains), 30 are curated news (cover 13.61%), 40 are soft news (cover 8.95%), and 296 are personal media (cover 35.10%)[16].

The substantive meaning of the factual-opinionated categorization lies in how helpful the information is in forming attitudes. Factual sources may provide unbiased, neutral information, but one needs to make extra cognitive efforts to form attitudes. Opinionated sources provide more ready-made advice to form attitudes. By definition, the information from opinionated source often includes more opinions than facts. Given the limited cognitive resource and time, opinionated information is a useful heuristics (Tversky and Kahneman, 1974) for people to form or strengthen attitudes. On other hand, heuristics may involve bias that prevents individuals from making "correct" judgment. Here, the trade off between cognitive load and bias is important. The high reliance on opinionated information source may stabilize and strengthen attitudes, but at the same time, it can lead to biased judgments in forming attitudes.

4.4 Twitter Network: Retweets and Web-links

Lastly, we generate variables of network-attributes by extracting information from raw tweet texts. First is the information on *retweets*. We identify retweets by considering whether the tweet text includes "@" or not, which is the common format to reply to or quote tweets/users[17]. Then,

[15] There are 9582 domains in total, but most of them appear only once or twice. We choose 500 most frequent domains to increase efficiency in the coding process. Remaining 9082 domains remain un-coded, but since they are rarely referenced, it can be said that those information sources are not significant in explaining network characteristics. Still, a future study may explore those minor domains to see if the result changes.

[16] The number or percentages do not add up to 500 or 100%, because 79 domains do not fall into four categories. Those extra domains are represented by commercial sites for advertisement and unidentifiable domains (e.g., closed sites or non-converted shortened URL) that involve insufficient information to specify their characteristics.

[17] The original data provide variables for retweets and retweeted user name, which is identified by whether the tweet uses "official" retweet feature in Twitter. However, in Japan during 2012, it was not common to use this official feature to respond to or quote past tweets. Users often manually insert "@" to represent retweets.

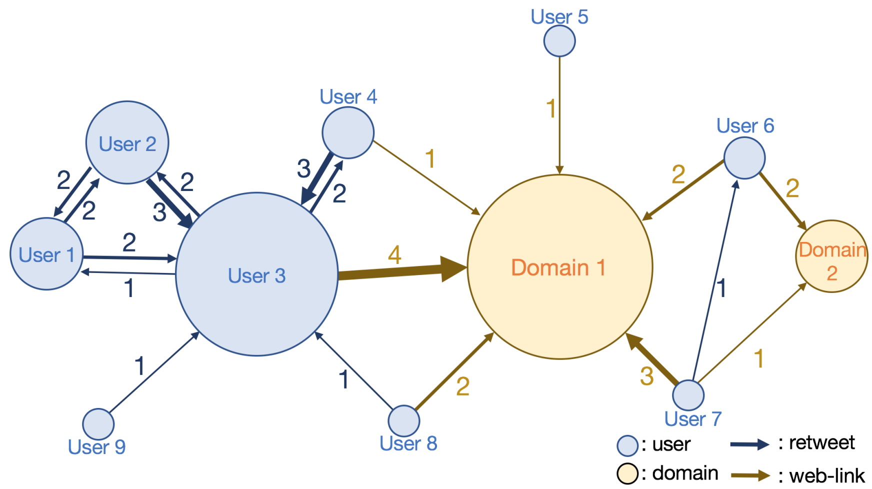

Figure 4: Sample Image of Twitter Network

retweeted user is specified by extracting the word following "@", which indicates the user name of which the tweet is replying to. Second is the information on *web-links*. Here, we extract all URLs appeared in tweet texts, and in the case of shortened URLs, expand to original URLs.

Information of *retweeted user*, and *web-link* can be used to draw twitter network (Kwak et al., 2010; Suh et al., 2010). Figure 4 presents the sample image. In this network, the fundamental unit (called *vertex/node*) is twitter user or web domain, and information communication (i.e., retweet or web-link) is represented by arrows (called *edges/links*) connecting vertices. Edges are *directed* and *weighted*. For example, the edge directed from user 3 to Domain 1 indicates the web-links from user 3 to domain 1. Similarly, the edge directed from user 2 to user 3 represents retweets from user 2 to user 3. The thickness of edge indicates weight to the network tie, which is measured by the frequency of retweets or web-links (higher frequency is represented by the thicker line).

For each user, the total information communication level can be captured by the combined thickness of edges coming in and out of vertices, called *degree*. There are two types of degree. The one for edges coming out of a vertex is *out-degree*. This captures individual-level activity to access information of others. Note that there is no out-degree for web domains because there is no mechanism in this network for websites to reference twitter users. The one for edges coming into a vertex is *in-degree*. This captures one aspect of how much the user or domain is influential in the network because higher in-degree implies that the user or domain is referenced more by other users. In Figure 4, user 2 and domain 1 have particularly high in-degree compare to other users/domains, implying that they are more influential in the network than others.

5 Hypothesis: Territorial Dispute Incidents, National Identity Holder, and Motivated Twitter Communication

In this section, we formulate hypotheses specific to our data. The theories of intergroup threat, group identity, and motivated reasoning generate various expectations on the roles of territorial dispute incidents and national identity holder in the twitter network. In the current study, we focus on the user/domain specific characteristics of information sources (i.e., identity status of a twitter user and factual/opinionated nature of a web domain) to test those theoretical expectations. This approach helps to isolate the nature of information communication process in twitter user/domain level network[18].

To start with, intergroup threat theory implies national identity holders to have a higher threat perception to out-group (i.e., South Korea and China) and thus higher sensitivity to territorial dispute issue. Therefore, in the context of our twitter data, the first hypothesis can be constructed as follows:

H1. Compare to those without national identity, national identity holders posts more territorial dispute relevant tweets.

Next three hypotheses are relevant to information communication patterns in the individual-level network. First, at all time, we expect strong national identity holders to have a stronger directional motivation to strengthen and protect their identity than those without strong national identity. In our twitter network data, this tendency can appear in following two ways:

H2a. Compare to those without national identity, national identity holders are more likely to retweet fellow national identity holders than no national identity holders.

H2b. Compare to those without national identity, national identity holders are more likely to link opinionated web-domains but less likely to link factual web-domains.

H2a is based on the expectation that the information from the strongly identified member of the group is more useful in strengthening one's identity. H2b is based on the expectation that the opinionated information source is more helpful in strengthening attitudes than the factual information source.

Then, the occurrence of territorial dispute incidents should make individuals have directional than accuracy motivation. Before the incident, we expect individuals to have accurate motivations, thus they process information from different sources equally. After the incident, however, the rise in threat perception should make individuals gain directional motivation, thus process information in a biased way to protect group identity. In parallel with H1, in our twitter network data, this pattern can be presented in two ways:

[18] Here, we do not use tweet specific variables (e.g., positive/negative sentiments in tweets) to test hypotheses. Since characteristics of tweet/web-link are not static within individuals, it is difficult to draw individual-level implications from tweet level variables. Still, this omission may limit the implications from our analysis, so the future study can extend the analysis by incorporating tweet specific variables.

H3a. Compare to before the territorial dispute incident, after the incident, individuals are more likely to retweet national identity holders than no national identity holders.

H3b. Compare to before the territorial dispute incident, after the incident, individuals are more likely to link opinionated web-domains but less likely to link factual web-domains.

In addition, previous studies suggest that group identity plays a moderating role to strengthen the relationship between threat and intergroup attitudes. Therefore, the moderation hypothesis is formulated as follows:

H4. The relationship between territorial dispute incident and retweet/web-link patterns (H3) is stronger among national identity holders than among no national identity holders.

Fifth hypothesis is relevant to the roles of individual actors (i.e., twitter users and web domains) in the society-level information communication network. If the individual-level mechanism of H3 is true, national identity holders and opinionated domains are more likely to be retweeted/linked under threat than under no threat. This leads to the following hypothesis regarding roles of national identity holders and opinionated domains in the network:

H5a. Compare to before the territorial dispute incident, after the incident, national identity holders are more influential/central than non-identity holders in the society-level information communication network.

H5b. Compare to before the territorial dispute incident, after the incident, opinionated web-domains are more influential/central than factual web-domains in the society-level information communication network.

6 Analysis 1: Patterns in Individual-Level Network

In this section, we analyze patterns in information communication behavior *from* individual twitter users. In particular, we focus on frequency of tweets and out-degree in twitter network. The first part compares the patterns between users with different level of national identity. The second part assesses the impact of threatening incidents by observing the time-series patterns of information communication.

6.1 The Impact of Strong Identity

We start the analysis with comparing the simple tweet frequency between twitter users with and without a national identity. Figure 5 presents the main result. It compares the mean number of tweets for the entire period in each group. Left two panels consider national identity as *patriotism*, and right two panels consider national identity as *right-wing nationalism*. Machine learning

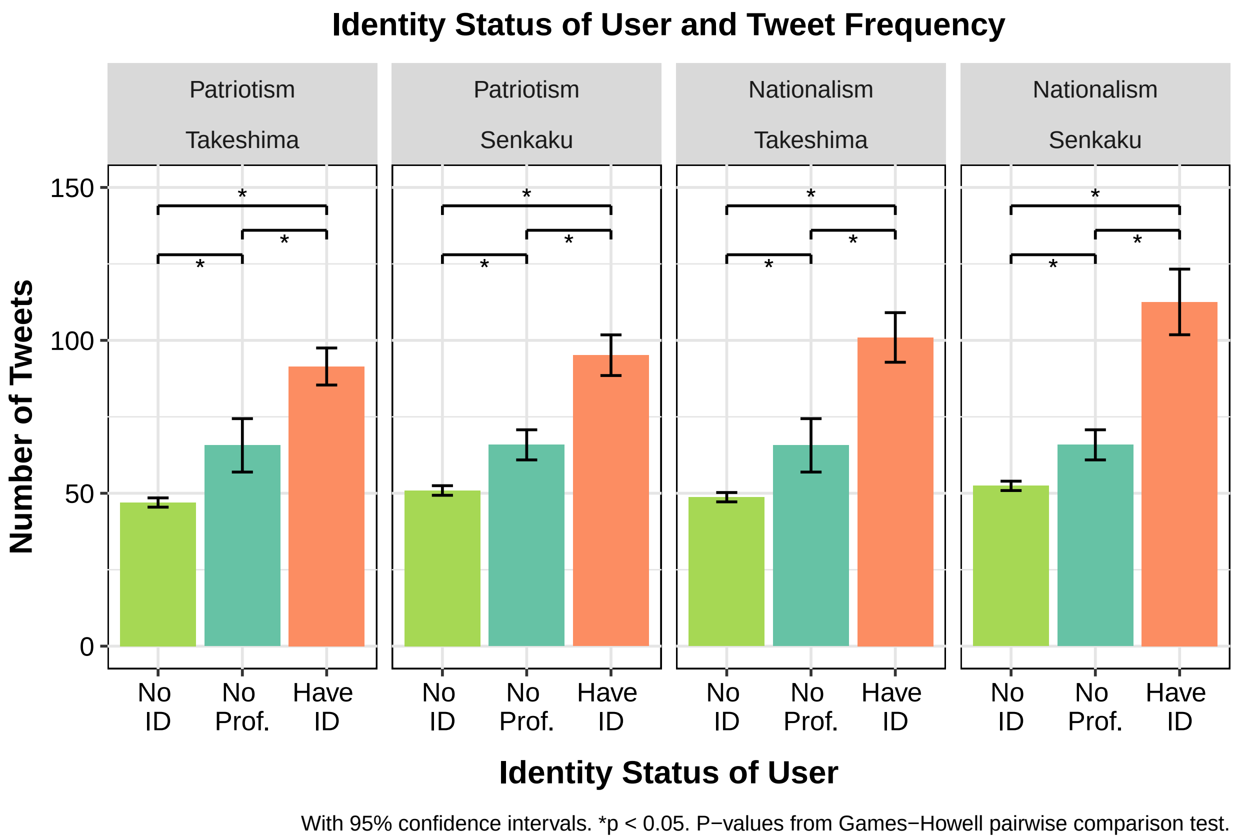

Figure 5: Identity Holders are More Likely than No Identity Holders to Tweet

method of the identity classification is described in section 4.2. In addition to those with and without identity, we also calculate the scores for those with empty user profiles. No profile users are expected to have mixed characteristics of those with and without identity[19].

All panels in Figure 5 show clear tendency that strong national identity holders, either as patriotism or nationalism, tend to tweet more frequently than users without strong national identity, thus confirming H1. For example on Takeshima issue, average patriotic identity holder tweet 91 times in 182 days period, while those without patriotic identity tweet only 47 times on average during the same period. The tendency for the users without a profile is in the middle of those with and without identity, which confirms the initial expectation. Since distributions are heteroscedastic and do not follow a normal distribution, we apply the Games-Howell pairwise comparison test (Games and Howell, 1976) to assess inter-group differences. The result indicates all inter-group differences in the mean number of tweets to be statistically significant ($p < 0.05$).

[19]We also considered if the result is affected by the automated account of tweets called "BOT". We identify BOT accounts by whether the user name or user profile include relevant keywords such as "bot" or "news." BOT accounts tweet significantly more than other accounts, but separate consideration of them does not change the general result on identity status. Therefore, in the paper, we present the results without the consideration of BOT account status.

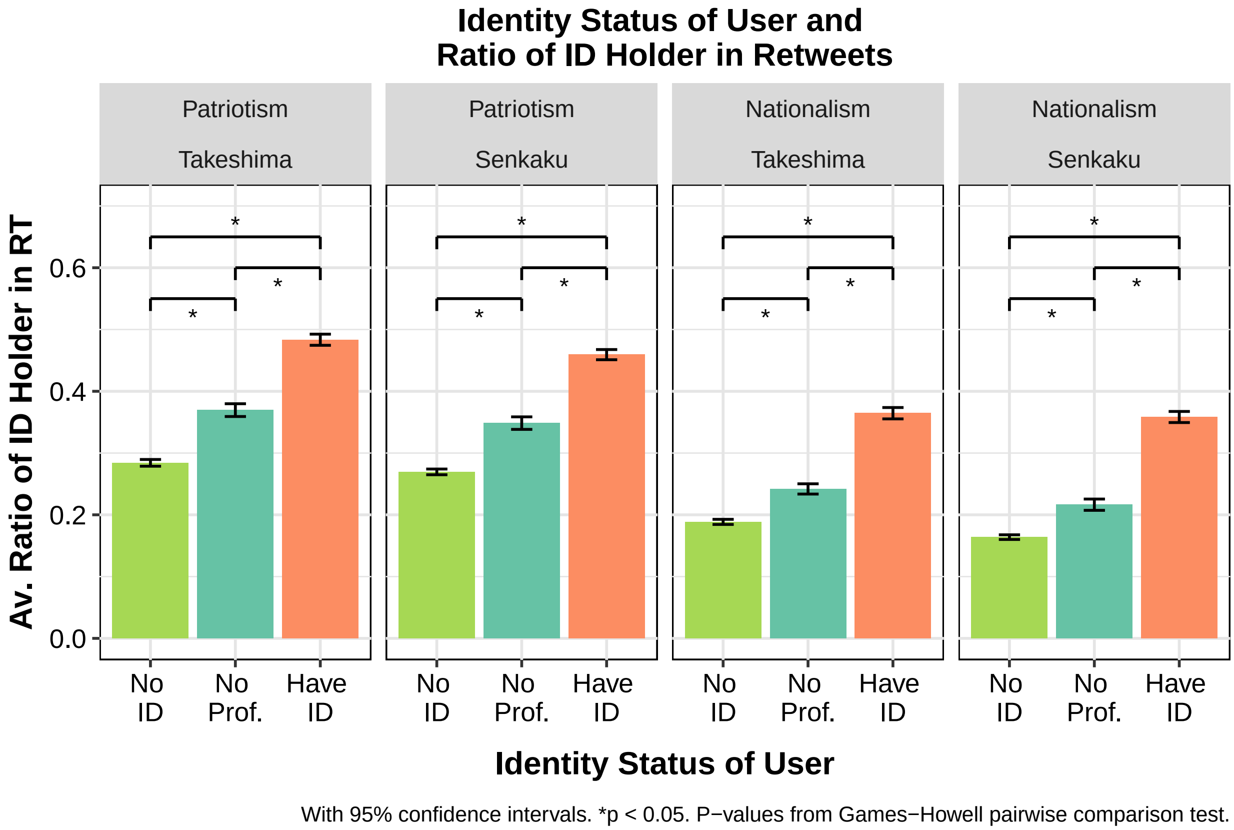

Figure 6: Identity Holders are More Likely than No Identity Holders to Retweet Identity Holders

Second, we evaluate retweet patterns to assess H2a. For each twitter user, we calculate the proportion of retweets to identity holders in the total number of retweets[20], and take an average score among groups of users with identity, without identity, and without a profile. Figure 6 presents the result of the analysis. It indicates that the ratio of national identity holders in retweet is higher for national identity holders – either in terms of patriotism or nationalism – than for no national identity holders. For example on Takeshima issue, 48.3% of average patriotic identity holder's retweets are consisted from patriotic identity holders, while only 28.4% of average no identity holder's retweets consist from patriotic identity holders. As the same as for tweet frequency, users with no profile shows the tendency in the middle of identity holders and no identity holders. The result confirms H2a, and all differences are statistically significant ($p < 0.05$).

Third, we evaluate H2b by comparing the average frequency of links to domains in each category[21].

[20] Users with no profile are excluded from the calculation since empty profile provides no information on the identity of users. Also, to avoid the ratio to be extremely biased by the behavior of users who posted only a few times, we remove users who tweeted less than 10 times. Lastly, we remove users who tweeted only before the corresponding territorial disputes incident to eliminate the potential impact occurred after the incident.

[21] Domains with no coding (ones not included in 500 most frequently linked domains) are excluded from the calculation, since those domains involve no information on factual-opinionated nature of domains.Also, to avoid the ratio

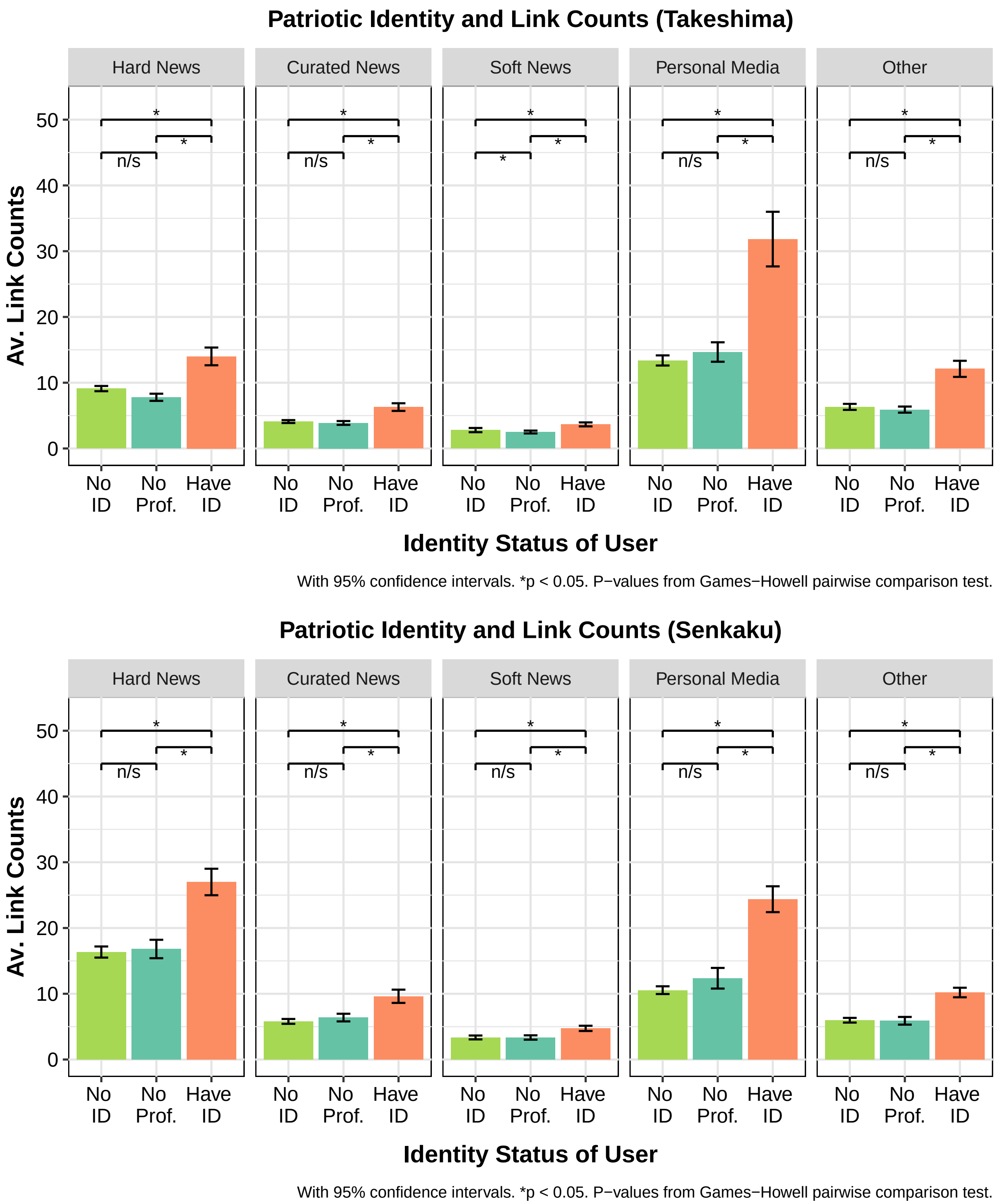

Figure 7: Patriotic Identity and Link Frequency by Domain Categories

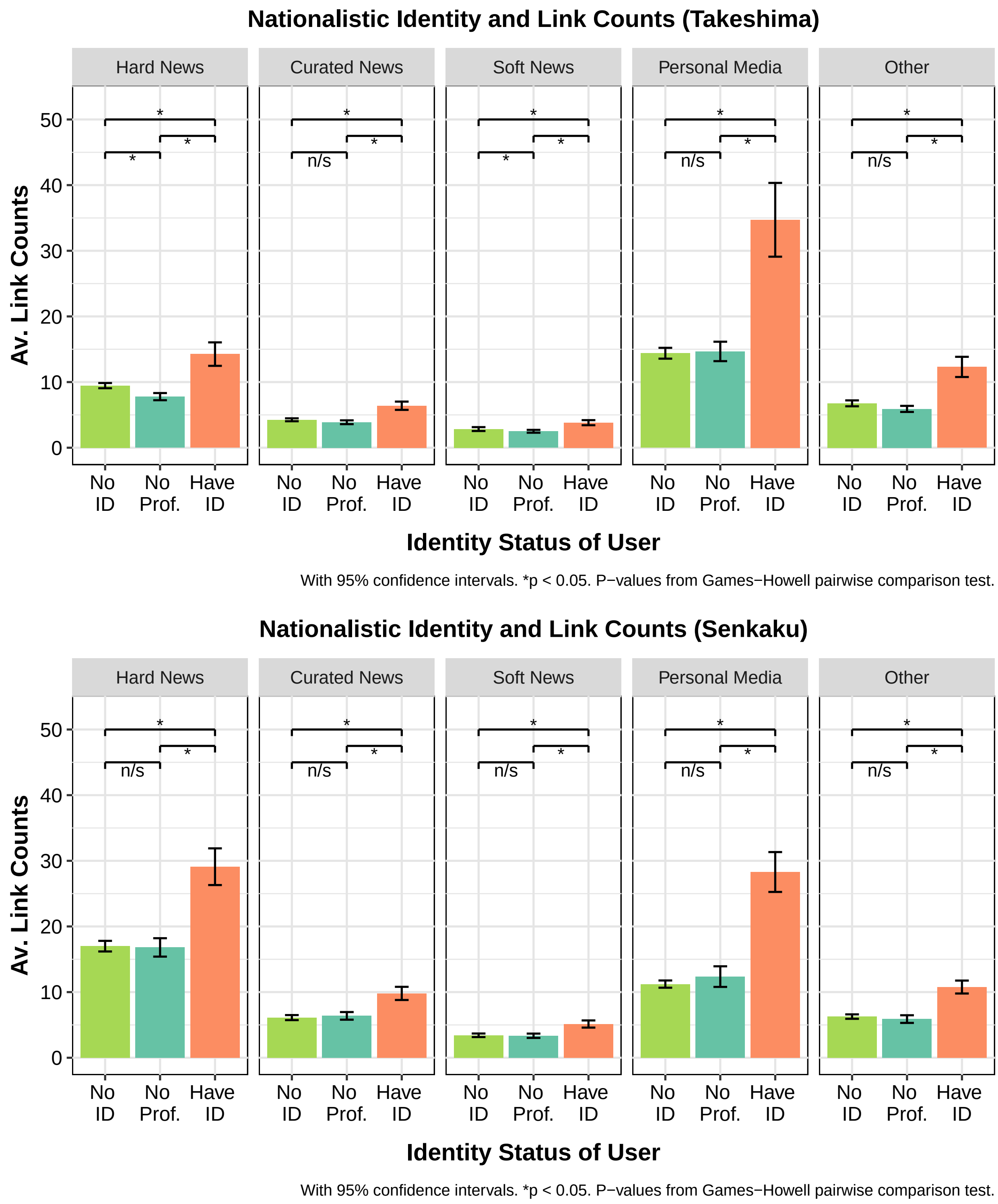

Figure 8: Nationalistic Identity and Link Frequency by Domain Categories

Results are shown in Figure 7 (for patriotic identity) and Figure 8 (for nationalistic identity). The results show partial support to H2b. First, there is a general tendency that identity holders link any category of domains more frequently than no identity holders. All differences are statistically significant ($p < 0.05$). This tendency is roughly the same across two types of identity, while is slightly stronger for nationalistic identity. It confirms our expectation for more opinionated domains (i.e., personal media, soft news), but does not support our expectation for more factual domains (i.e., hard news, curated news). Second, comparing across categories of domains, links to personal media show the largest and the most consistent (across both identity type and issue) identity-no identity difference in frequency. For example, on patriotic identity for Takeshima issue, identity holders link personal media for 32 times on average, while no identity holders link it for only 13 times. Hard news also records large inter-identity status difference, while the tendency differs across issues. The difference is large for Senkaku issue – identity holders' average hard news link frequency exceeds no identity holders by 11-12 counts – but relatively small for Takeshima issue – identity holders' average link frequency exceeds no identity holders by only 4-5 counts. The inter-identity status difference for curated news and soft news are significantly smaller than the above two categories. Third, the tendency for no profile users generally fall in between identity holders and no identity holders, but often closer to no identity holders. For most of the comparisons, the differences between no profile users and no identity holders are not statistically significant ($p > 0.05$).

In sum, the comparisons between national identity holders and no national identity holders partially support our expectations. To start, we observe that national identity holders simply tweets more than no identity holders on territorial disputes (H1 is supported). Then, when accessing to information from other twitter users, national identity holders retweets more to identity holders than no identity holders (H2a is supported). On referencing information from external websites, identity holders more frequently link any types of domains, but the largest and the most consistent difference is found among personal media – the most opinionated domains (H2b is partially supported).

6.2 The Impact of Threatening Incident

In this part, we visualize the daily time-series trend in retweet and web-link communication patterns to evaluate H3 and H4.

6.2.1 On Identity Holder Ratio in Retweets

To assess H3a and H4, the initial result for retweeting patterns is shown in Figure 9. The pattern is presented separately for users with and without identity, because previous section shows the significant difference between retweeting patterns of identity holders and no identity holders. We use LOESS (Cleveland, 1979) to draw smoothing lines, and 95 percent confidence intervals are

to be extremely biased by the behavior of users who posted only a few times, we remove users who tweeted less than 10 times. Lastly, we remove users who tweeted only before the corresponding territorial disputes incident to eliminate the potential impact occurred after the incident.

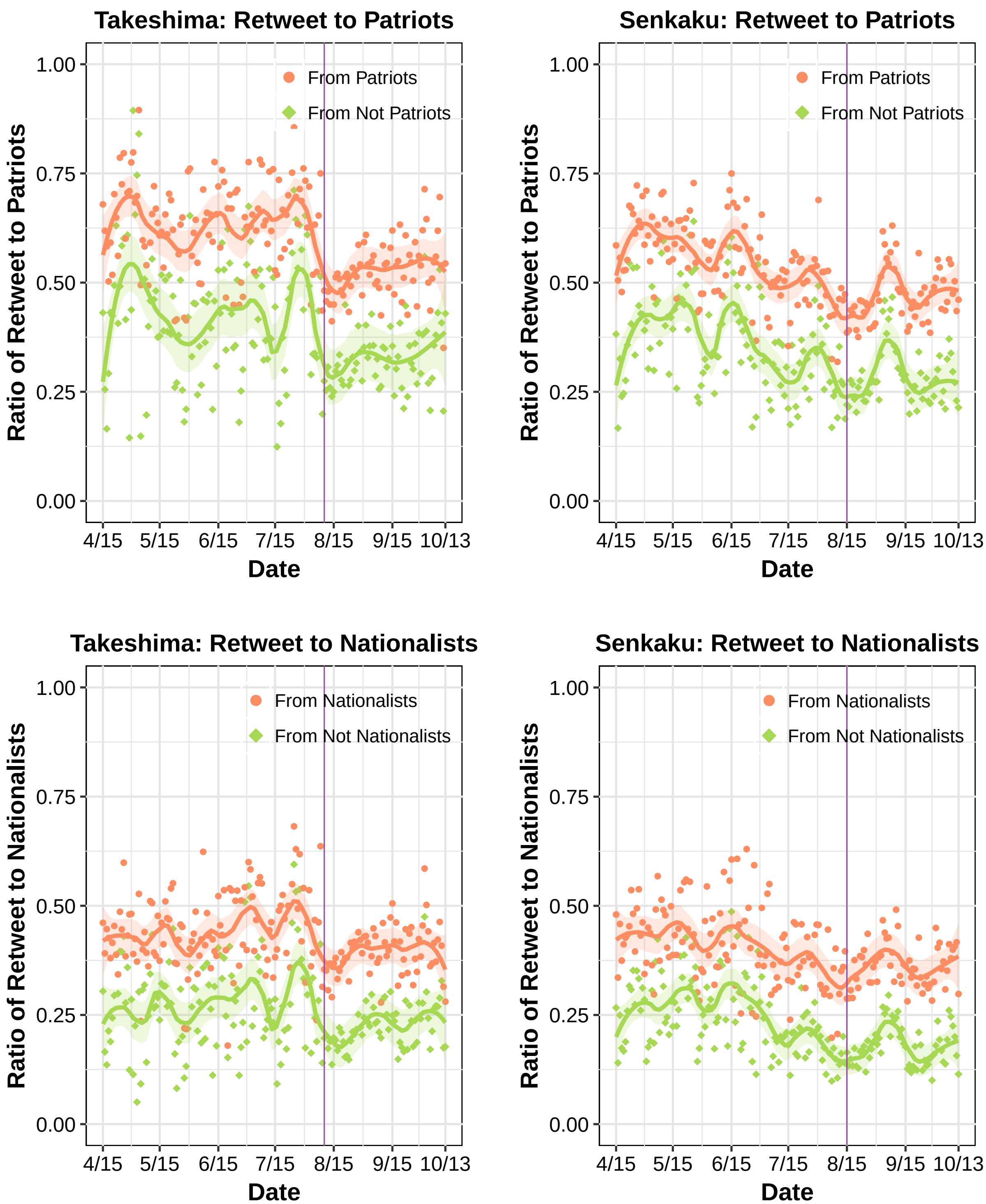

Figure 9: ID Holder Ratio in Retweet and Incident (Takeshima)

displayed with semitransparent colors around the lines. The timing of the most significant incident, as discussed in section 4.1, is shown as vertical purple line in each graph. From the first look, we find the tendency that contradicts with our expectation in H3a. For both issues and for both users with or without identity, the average proportion of identity holders in retweets seem to decrease after the incident. Taking the averages of entire periods before and after the incident, the identity holder ratio decreases by 0.05 to 0.1 in the post-incident period compared to the pre-incident period (see Appendix B).

However, the above results could be affected by the change in a ratio of tweets from the national identity holders in total tweets. In other words, Figure 9 may be explained by the possibility that, after the incident, there are more users without a national identity to retweet-to, not that one is more likely to choose no identity holders to retweet. Figure 10 shows the tendency consistent with this suspicion. It visualizes the total number of tweets per day and the ratio of tweets posted by identity holders. The blue line indicates the proportion of identity holders' tweets in the total number of tweets posted on each day. It can be seen that the line drops significantly on the day of the incident, for both issues and both types of identity. This means that, just after the incident, even when one randomly chooses who to retweet from those who tweeted on the same day, there are fewer identity holders than no identity holders to retweet.

To cope with the change in the user population to retweet-to, we design the model of a retweet information communication process. The model assumes that a user chooses who to and how much to retweet from the pool of tweets that are available to him/her before he/she retweets. Let $r_i(t)$ be the ratio (probability) of identity holder to be retweeted in one tweet from user i at time t[22]. If we assume that user i randomly pick who to retweet from the pool of available tweets, the expected value of $r_i(t)$ can be calculated by an equation below.

$$E\left[r_i\left(t\right)\right] = \frac{N_{\mathrm{ID}}(t,\tau)}{N_{\mathrm{ID}}(t,\tau) + N_{\mathrm{No}}(t,\tau)} = R(t,\tau) \tag{3}$$

Here, users have a period τ, defined as the period of tweets to be remained in their memory. Whenever to retweet, users are assume to select someone who posted tweets during $[t-\tau, t]$. Then, $N_{\mathrm{ID}}(t,\tau)$ and $N_{\mathrm{No}}(t,\tau)$ represent a total number of tweets posted during $[t-\tau, t]$, by national identity holders and no identity holders, respectively. To shorten equations, $R(t,\tau)$ denotes the ratio calculated with $N_{\mathrm{ID}}(t,\tau)/N_{\mathrm{ID}}(t,\tau) + N_{\mathrm{No}}(t,\tau)$, which defines the proportion of identity holders' tweets in the pool of tweets that a user *can* retweet to.

Two major factors can explain the non-random selection of who to retweet from $R(t,\tau)$. First, each user may have a general tendency in the reaction level to national identity holders, as we evaluated in section 6.1 that national identity holders tend to have a higher likelihood than no identity holders to retweet fellow identity holders. We denote this tendency as a weight w_i and assume that w_i does not change over time. Second, each user could be affected by a sequential event occurring at each time point. We denote the impact size of event itself as $e(t)$, and formulate an effect received by each user with function $f_i(e(t))$. $e(t)$ indicates that the sequential event spreads its impact

[22]Note that multiple users can be retweeted in one tweet. $r_i(t)$ is a probability if only one user is retweeted, but a ratio if two or more users are retweeted.

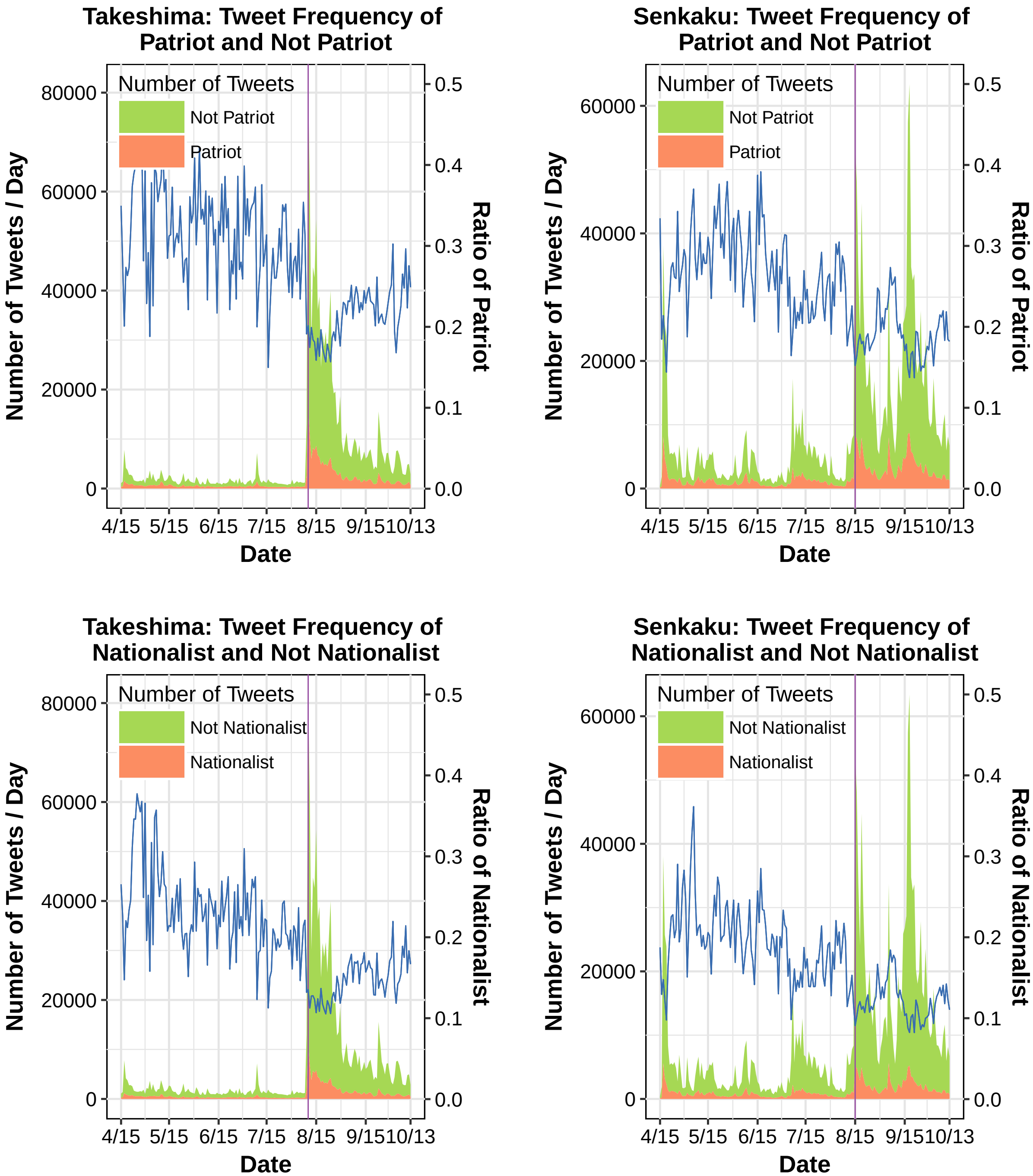

Figure 10: ID Holder Ratio in All Tweets and Incident (Takeshima)

uniformly, and $f_i(\cdot)$ implies that the impact received by the user is different across each individual. Adding the above two factors, Equation 3 can be modified to an equation below.

$$E\left[r_i\left(t\right)\right] = R(t,\tau)\left(w_i + f_i\left(e\left(t\right)\right)\right) \tag{4}$$

Since we want to know the event impact on users, by rearranging Equation 4, the following equation can be obtained.

$$w_i + f_i\left(e\left(t\right)\right) = E\left[r_i\left(t\right)\right]/R(t,\tau) \tag{5}$$

By taking an average across the users:

$$\bar{w} + \bar{f}\left(e\left(t\right)\right) = E\left[\bar{r}\left(t\right)\right]/R(t,\tau) \tag{6}$$

Now, we can estimate $\bar{w} + \bar{f}\left(e\left(t\right)\right)$ by adjusting observed proportion of identity holders in retweets (which we already have shown in Figure 9) by $R(t,\tau)$ (illustrated in Figure 10). We use one day (24 hours) as τ to calculate $N_{\text{ID}}(t,\tau)$ and $N_{\text{No}}(t,\tau)$ for each retweet at time t[23].

The impact of events and personal characteristics on retweets to the identity holders $(\bar{w} + \bar{f}\left(e\left(t\right)\right))$ is visualized in Figure 11. If this score is 1, an individual is choosing who to retweet purely randomly from the pool of tweets posted in the previous 24 hours. If the score is higher than 1, individuals are more likely than random to retweet identity holders. If the score is lower than 1, individuals are more likely than random to retweet no identity holders. From the initial observation, it should be noted that the score is generally higher than 1 for both identity holders and no identity holders. This implies that it is more likely than random for any twitter user to retweet identity holders. Individuals have a general tendency to select identity holders when they access territorial dispute information from other twitter users.

Then, the time-series movement in Figure 11 show different patterns than Figure 9. The score tends to increase at the point of threatening incident. On Senkaku issue for national identity holders – either in terms of patriotism or nationalism – there is a significant jump in the score at around the time point of threatening incident (August 15th). After the incident, compared to before the incident, users with national identity do choose more identity holders when they retweet. The impact of Senkaku event seems to be weaker among no identity holders, given that the size of change at around threatening incident is smaller for no identity holders. Taking the averages of entire periods before and after the incident, event impact scores for national identity holders on Senkaku issue increase by 0.40 to 0.81 in the post-incident period compared to the pre-incident period, while increases are limited to 0.10 to 0.14 for no identity holders (see Appendix B).

The above tendencies for Senkaku issue confirm our expectations in H3a and H4. Threatening incident comes with the increasing reliance on identity holders in retweets (H3a), and this tendency is strong among identity holders but weak among no identity holders (confirms H4)[24].

[23] This is obviously an arbitrary selection, but given the fast-moving nature of social networking service, we don't expect users to retweet something posted multiple days before. Shorter time-range can be a possibility. We tested with $\tau =$ one hour, yielding almost identical result as $\tau =$ one day.

[24] The further considerations by the difference in differences in event impacts yield the same implication (see Appendix C).

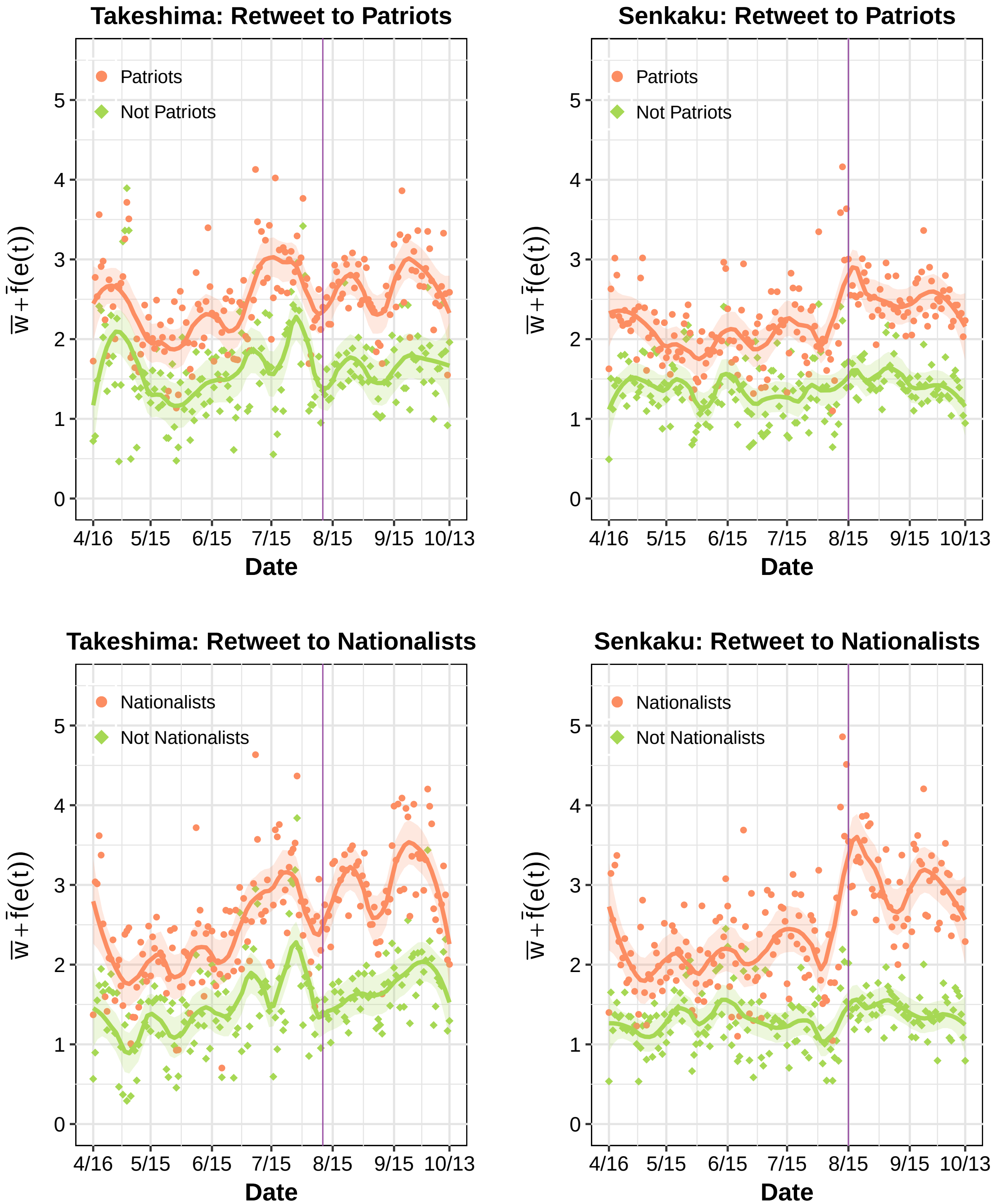

Figure 11: The Impact of Event (Time) and Personal Characteristics on Retweet to National Identity Holders

It should be noted that the result for Takeshima issue is rather mixed. We can still observe the pattern that the increase in the selection of identity holders is preceded by the threatening incident, and compare the entire periods before and after the incident, the average event impact scores for identity holders increase by 0.24 to 0.68 in the post-incident period compared to the pre-incident period (see Appendix B). On the other hand, the incident (August 10th) is not the only timing when there is a significant movement in the level of event impact scores. Scores also increased significantly during the period of both June to July and August to September. For Takeshima issue, it is unclear if threatening incident directly increases the reliance on identity holders.

6.2.2 On Links to Opinionated and Factual Domains

To evaluate H3b and H4, Figure 12 (for patriotic identity) and Figure 13 (for nationalistic identity) plot the daily average count of linked domains per user by each category[25]. First, it can be seen that links to any type of domains jump up after the incident. For Takeshima, there is a unique peak in frequency at the time of territorial issue incident, while for Senkaku, the territorial issue incident defined in this study is just one of the peaks. Timings of the peak seem to correspond with when the news on territorial issues is reported. For Takeshima, the news on the Lee Myung-bak's visit to the island on August 10th is the unique most significant news during the period. On the other hand, for Senkaku, news on the dispute are repeatedly aired during the period. The timing of the increase in links seems to correspond with the air of the news on the significant development of the issue. On the surface, this result is partially consistent with H3b, which predicts links to opinionated domains to increase but links to factual domains to decrease. However, similar to the discussion in the previous part, it is unclear whether the each rise is caused by individual's selective behavior, or by the increase in the pool of information after the incident. In contrast to the previous part, we have no information to control for the pool of information provided from each type of domains during the data collection period. Therefore, we pay more attention to the relative differences among domain categories and identity status.

First, the persistence of increase is different across domain categories. On the one side, the link frequency starts to decrease shortly after the increase. This short-term increase pattern is most prevalent for hard news. Curated news and soft news follow the similar but less prevalent patterns. On another side, the link frequency persists at high levels after the increase. Personal media shows this long-term increase pattern. After the increase in link frequency at the time of the territorial dispute incidents, the frequency does not decrease in the short span. Especially for Takeshima, the high link frequency persists for at least two months, until the end of data collection period. The above differences may have important implications in relations to the motivated reasoning. The accuracy motivation to access factual domains may decay shortly after the incident, but the directional motivation to access opinionated domains may persist for a longer time after the incident.

Second, comparing across the identity status, identity holders tend to have a steeper increase in link frequency than no identity holders. For example, taking the averages of entire periods before and after the incident, daily links to personal media for national identity holders on Takeshima issue increase by 0.32 to 0.35 in the post-incident period compared to the pre-incident period,

[25]The category "other" is excluded since there is no good way to interpret the result.

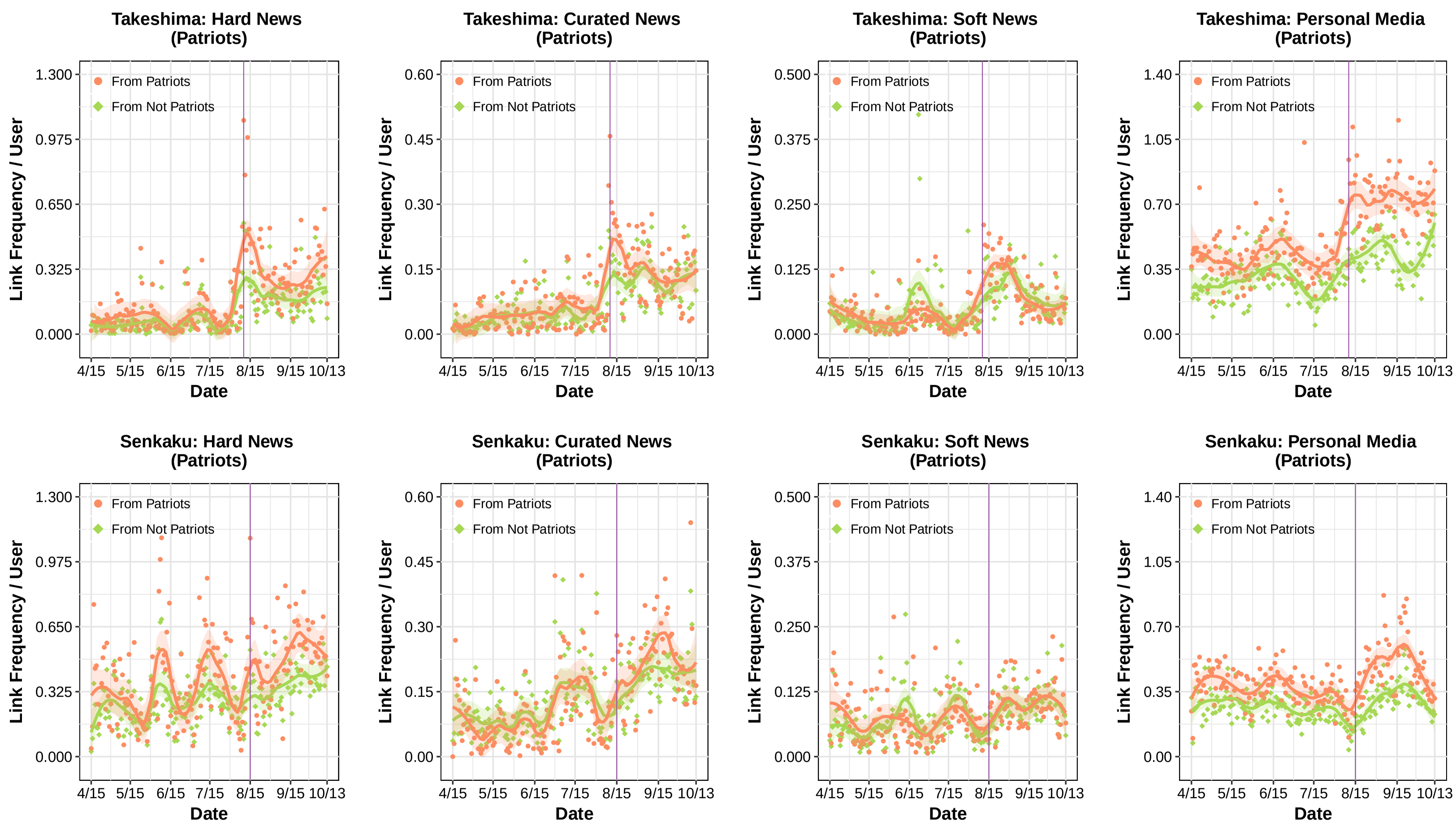

Figure 12: Average Daily Count of Web-Links per One User by Domain Categories (Patriots and Not Patriots)

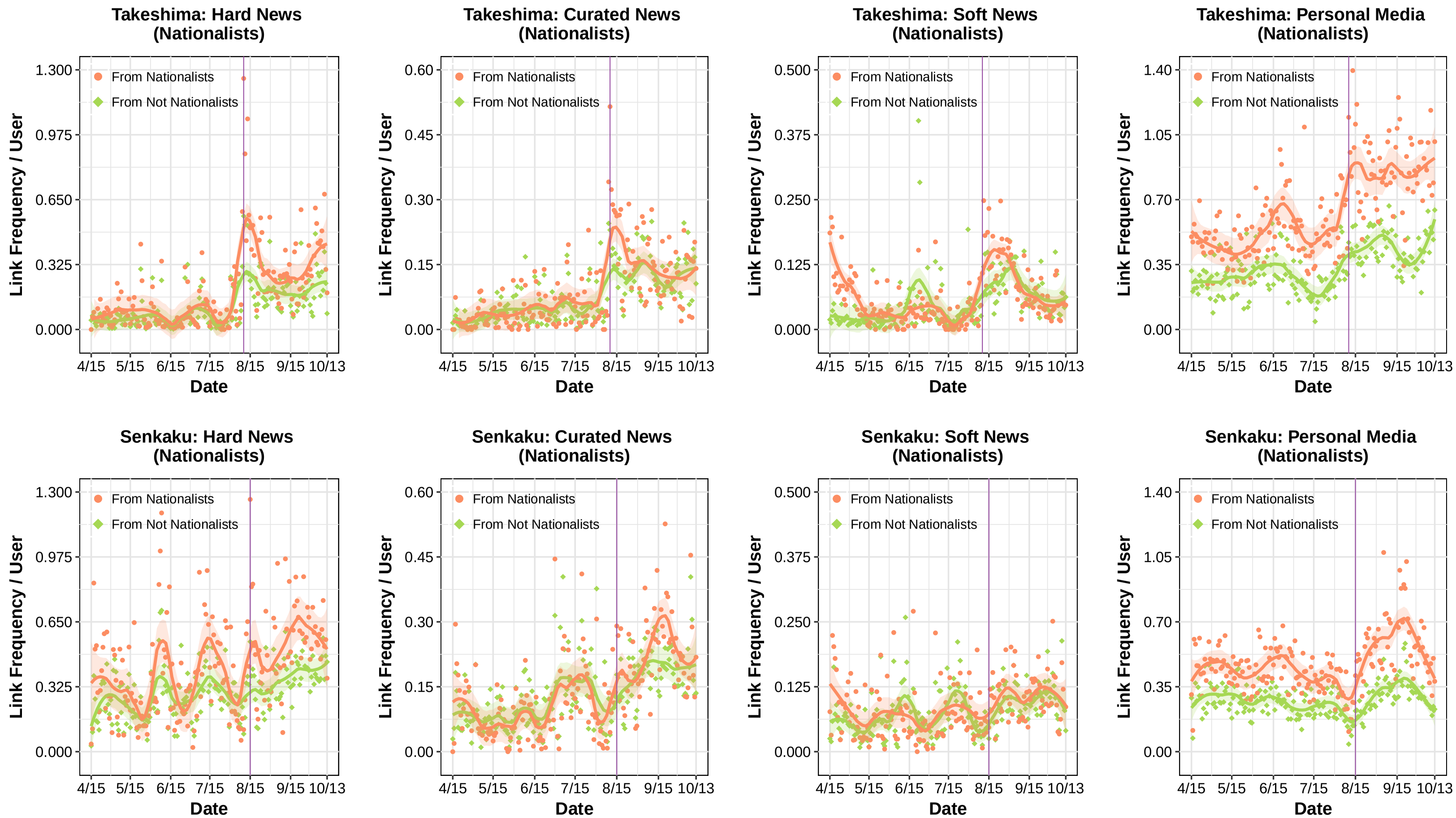

Figure 13: Average Daily Count of Web-Links per One User by Domain Categories (Nationalists and Not Nationalists)

while increases are limited to 0.15 to 0.16 for no identity holders (see Appendix B). This finding is consistent with H4, implies that identity holders are more sensitive to threats than no identity holders. Interestingly, among hard, curated and soft news, the difference is the largest among hard news and the smallest among soft news. After the incident, identity holders do have a stronger motivation to know about the facts than no identity holders. At the same time, the increase in personal media links after the incident is also much steeper for identity holders than for no identity holders. As already discussed, the increase in personal media links persists longer time than for hard news links, implying that directional motivation may prevail some period after the incident. All the above tendencies are slightly stronger for nationalistic identity than for patriotic identity.

In sum, we found that links to both factual and opinionated domains increase after the incident (partially consistent with H3b), while the direct cause of the increase is unclear. Comparing across domain categories, we find that increase in opinionated domain links (i.e., personal media) is more persistent than the increase in factual domain links (i.e., hard news). Lastly, identity holders record steeper increases in link frequency than no identity holders, implying that they are more sensitive to the threatening incidents (supports H4).

7 Analysis 2: Individual Roles in Society-Level Network

In this section, we focus on the network measure of *in-degrees* to assess the structure and individual roles in the society-level information communication network. By observing how much information communication is coming *into* an individual twitter user or web domain, we can evaluate how central/influential a user/domain is in the network. As a user/domain has more in-degree, more individuals in the network access or receive information from him, her or it. The first part evaluates the time-series movement in network-level measures of in-degree and its relevant statistics. The second part evaluates the time-series change in in-degrees of national identity holders and no identity holders. The third part evaluates the time-series change in in-degrees of opinionated (i.e., personal media) and factual (i.e., hard news) domains.

7.1 Threatening Event and Network-Level Change

To start with, as illustrated in Figure 4, we construct weighted directed network of information communication. To capture time-series change in the network structure, networks are generated from each day in the data (meaning that we have 182 networks for 182 days in our data). For each daily network, it is confirmed that in-degree distribution follows the power law. Figure 14 presents the example in-degree distribution from August 10th network on Takeshima issue. It indicates that the distribution has a long tail to the direction of high-degrees, meaning that the network includes some network hubs (particularly high-degree vertices).

In addition to average in-degree for the network, two additional measures are useful in assessing in-degree characteristics of the network. First, the standard deviation of in-degrees can capture how dispersed the in-degree distribution is. If average in-degree to stay constant, low standard

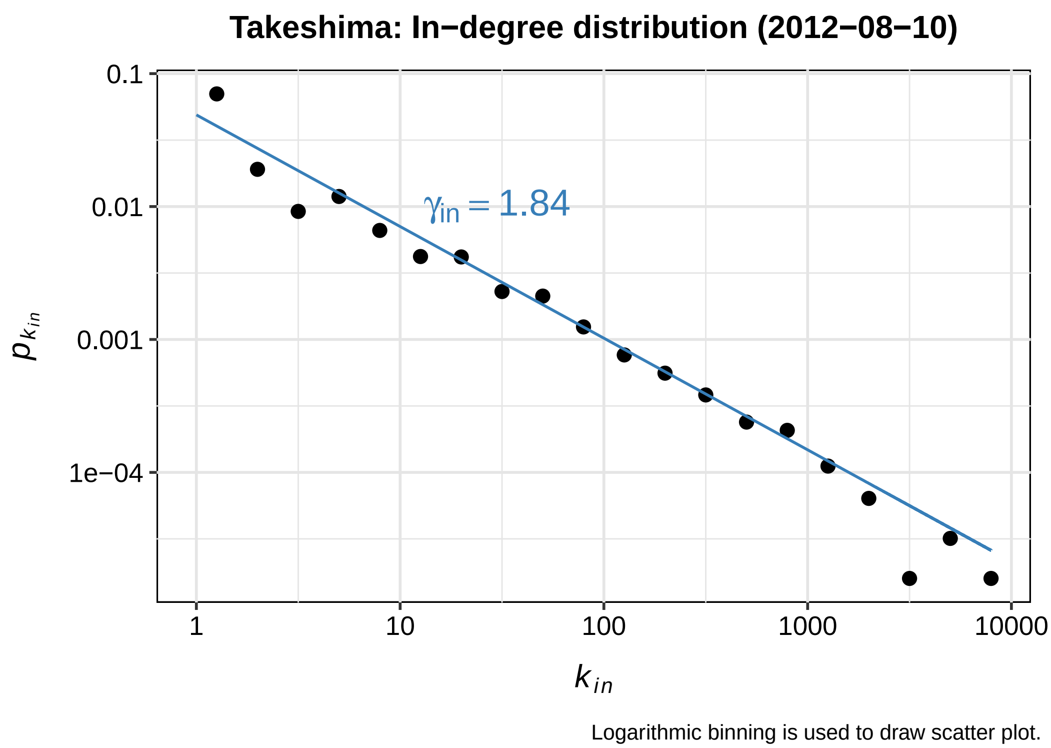

Figure 14: The In-degree Distributions of August 10th Network on Takeshima Issues

deviation implies that all users/domains tend to have equal-level of in-degrees, and high standard deviation implies that only some users/domains have high in-degree, while most of the others have low in-degree. Second is the measure called degree exponent. To make sense of this measure, the probability p_k that a vertex has exactly k edges ($k = 0, 1, 2, ...$) in a network following the power law can be formalized with an equation below (Barabási, 2016).

$$p_k = Ck^{-\gamma} \tag{7}$$

Here, C is constant and γ is called degree exponent. When γ becomes smaller, the tail of the distribution becomes longer. This means that there are few users or domains with extremely high in-degrees.

Figure 15 visualizes the daily time-series distribution of the average in-degree $\langle k_{in} \rangle$, the standard deviation $\sigma_{k_{in}}$, and the degree exponent of in-degree γ_{in}[26]. The first row shows the clear increase in the average in-degree after the territorial issue incident, both for Takeshima and Senkaku. Even some periods after the incident, the in-degrees stay at the higher level than the level before the incident. The second row indicates that the standard deviations of in-degree jump up at the moment of territorial issue incident. The score drops few days after the incident but stays at the slightly higher level than before the incident. The third row shows that the degree exponent drops after the incident. After the drop, it stays at the low level and does not return to the level before the incident.

[26]Parallel measures for out-degree are visualized in Appendix D

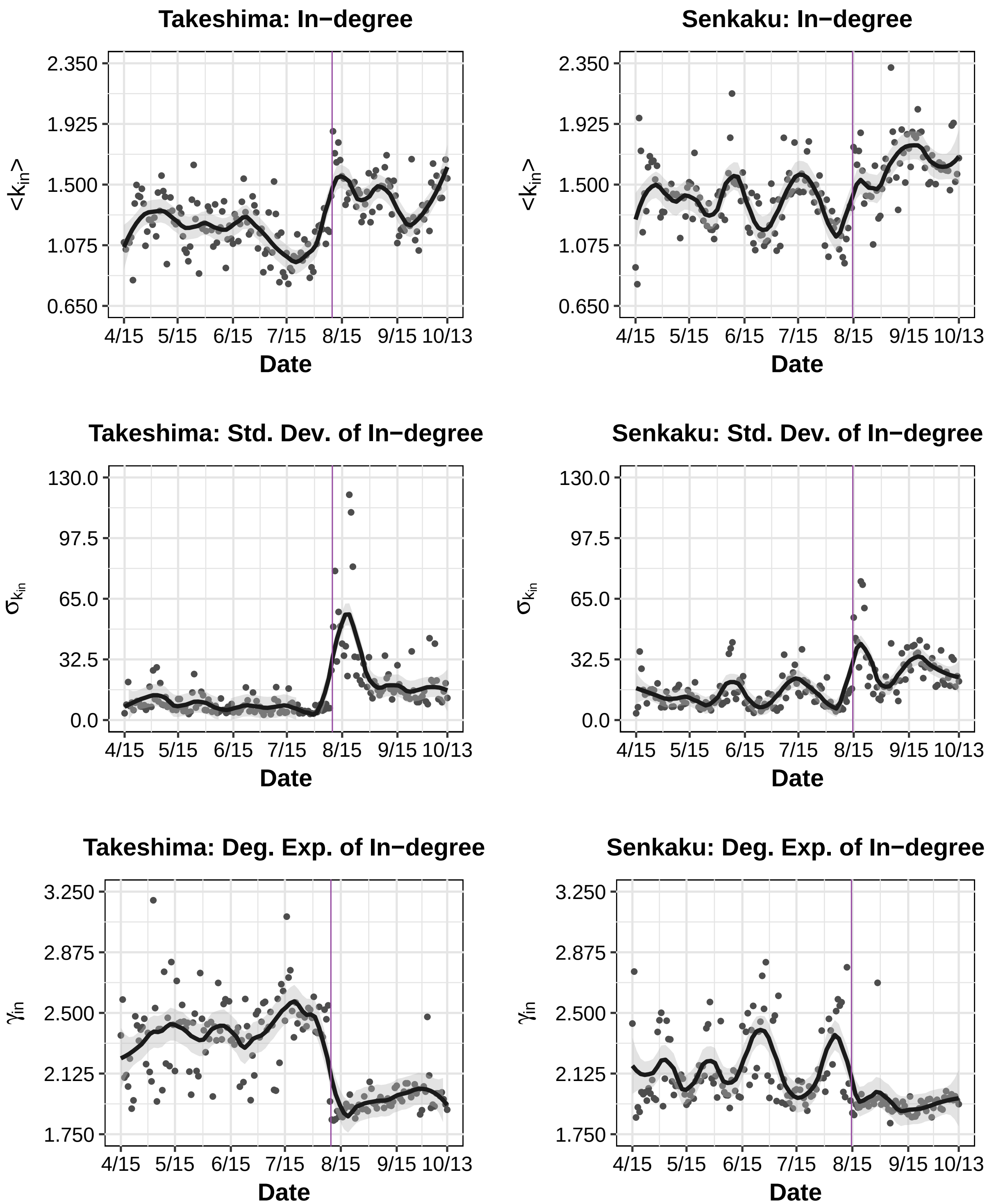

Figure 15: Time Series Distribution of In-degrees, Standard Deviation of In-degrees, and Degree Exponent of In-degrees

In summary, territorial issue incident coincides with the clear structural change in the in-degree characteristics of the network. Both in-degree and standard deviation of in-degree increase at the time of the incident, implying that information communication network becomes more active (dense) but more skewed (i.e., communication is concentrated around few in the network) after the incident. Degree exponent of in-degree drops after the incident, indicating that only a few users or domains gain the extremely high number of retweets or links after the incident. From the next part, we evaluate how these structural changes have differential impacts on identity holders and no identity holders, and opinionated domains and factual domains.

7.2 The Role of National Identity Holders

To evaluate H5a, we first compare the difference in the time-series *average* in-degree distribution of national identity holders and no identity holders. Figure 16 visualizes the result. From the first look, it can be seen that national identity holders have more in-degrees than no identity holders. For both Senkaku and Takeshima issue, national identity holders are retweeted larger amount of times than no identity holders. Then, comparing before and after the incident, in-degrees for identity holders jump up significantly after the incident. For example, average in-degree for patriots on Takeshima issue jumps up from the pre-incident level of slightly above one to over five on the incident day. In other words, patriotic users receive one retweet per day on average before the incident, but on the day of the incident (August 10th), they receive five retweets on average. The tendency is less clear for Senkaku issue, but it can be seen that identity holders tend to have higher in-degrees after the incident than before the incident. For both Takeshima and Senkaku issue, there is no apparent change in the in-degrees of no identity holders before and after the incident. The above assessments confirm H5a. In terms of average in-degree, national identity holders (but not no identity holders) do significantly gain influence in the network after the incident[27].

Another perspective to understand the influence of identity holders in the network is to look at the *total* in-degree distribution of national identity holders and no identity holders. Figure 17 presents the result. In contrast to the result in Figure 16, the total in-degree increases much more steeply after the incident for no identity holders than for identity holders. The combination of the previous result and this result implies that after the incident, even when identity holders are more likely to be retweeted on average, there are much more no identity holders *to be retweeted*. From this aspect, therefore, identity holders lost influence rather than gained influence after the incident[28].

In sum, the evidence give partial support to H5a. After the incident, identity holders, compared to no identity holders, tend to have more in-degrees on average, but fewer in-degrees in total. There are several caveats to the result. First, the average in-degree result does not necessarily imply that identity *cause* identity holders to gain influence on average. It might be the case that other characteristics of identity holders (e.g., the frequency of tweets, the number of followers or knowledge of the issue) make other users retweet them more after the incident. Second, the result does not necessarily show that identity holders gain or lose influence equally among everyone

[27]The mean comparisons of entire periods before and after the incident are provided in Appendix E.

[28]The mean comparisons of entire periods before and after the incident are provided in Appendix E.

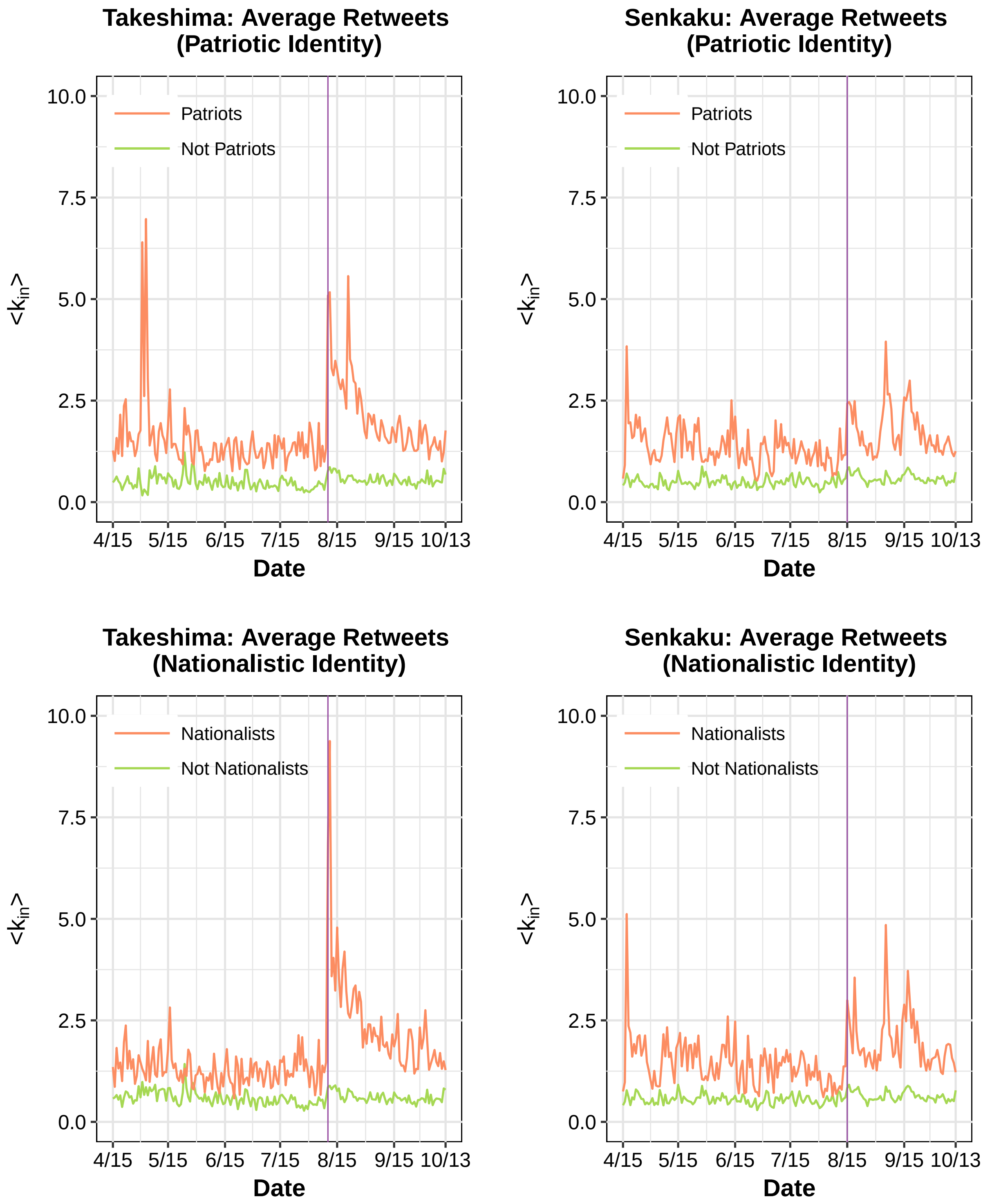

Figure 16: Time Series Distribution of Average In-degrees by Identity Holders and No Identity Holders

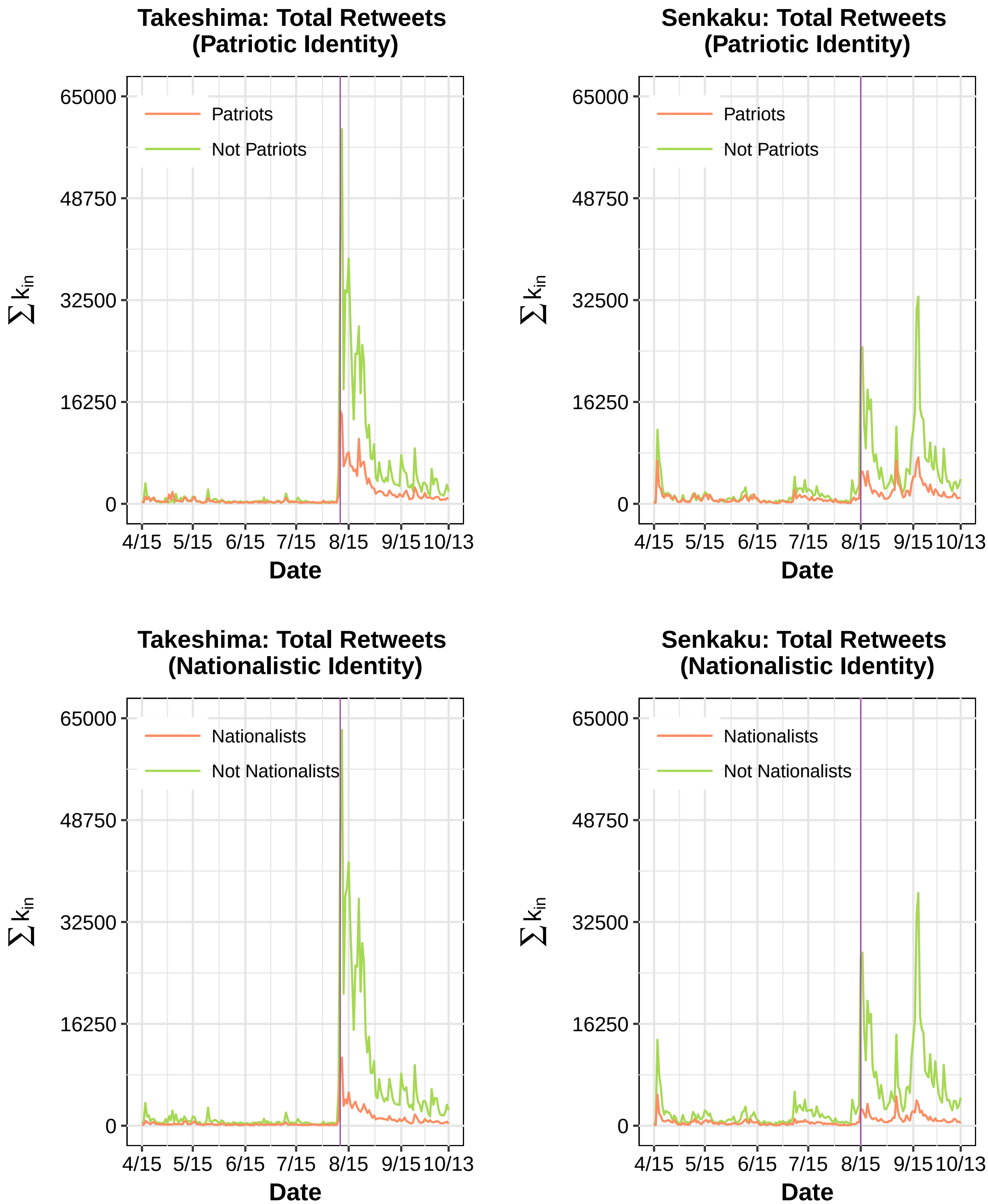

Figure 17: Time Series Distribution of Total In-degrees by Identity Holders and No Identity Holders

in the network. The analysis in section 6.2.1 indicates that (at least for Senkaku issue) national identity holders are more likely to retweet fellow national identity holders after the incident. It might be the case that identity holders gained influence more among the network of fellow identity holders, less among the network of no identity holders. Future studies should address those caveats to develop deeper understandings on the findings from this part.

7.3 The Role of Opinionated Link Domains

To assess H5b, we first compare the average in-degree for different category of domains. The previous analysis shows the clearest threat-related change among hard news domains (i.e., factual) and personal media domains (i.e., opinionated), thus we focus on those two categories of domains[29]. The upper half of Figure 18 visualizes the result. The general tendency shows that hard news domains tend to gain more average links than personal media domains after the territorial issue incident[30]. This finding contradicts with our expectation from H5b. On the other hand, considering the nature that hard news domains such as newspaper websites are often more accessible to a higher number of people than personal media domains, average in-degrees may not be comparable between two domains categories.

The look from total in-degrees is presented in the lower half of Figure 18. It shows that total in-degrees for hard news and personal media increase in different timings. For Takeshima issue, hard news domains gain total-indegrees immediately on the day of the incident, while personal media domains reach its peak in in-degree few days after the incident. For Senkaku issue, hard news domains similarly gain total in-degree the most on the day of territorial dispute incident. On the other hand, personal media domains gain the most in-degree on the next month of the territorial dispute incident (as we defined)[31]. Note that, at its peak, the total in-degrees of personal media reach the same or higher number than in-degrees of hard news. This result is consistent with the implication in section 6.2.2, that the threat impact on personal media domains is more persistent than the impact on hard news domains.

In sum, while the evidence from average in-degree is dis-confirming, the total in-degree result gives partial support to H5b. Even when factual domains (i.e., hard news) immediately gain total in-degree on the day of the incident, after few days to a month, opinionated domains (i.e., personal media) gain the same or higher in-degrees in total than factual domains in the network.

8 Discussion

The current project aims to understand the real-time social network dynamics of threatening events, national identity, and information communication. To achieve the purpose, we explore the patterns

[29] The tendencies for other two categories – soft news and curated news – generally fall in between hard news and personal media, while often more similar to the one for hard news.

[30] The mean comparisons of entire periods before and after the incident are provided in Appendix E.

[31] Another explanation can be that total in-degrees for hard news domains and personal media domains respond to different types of incidents.

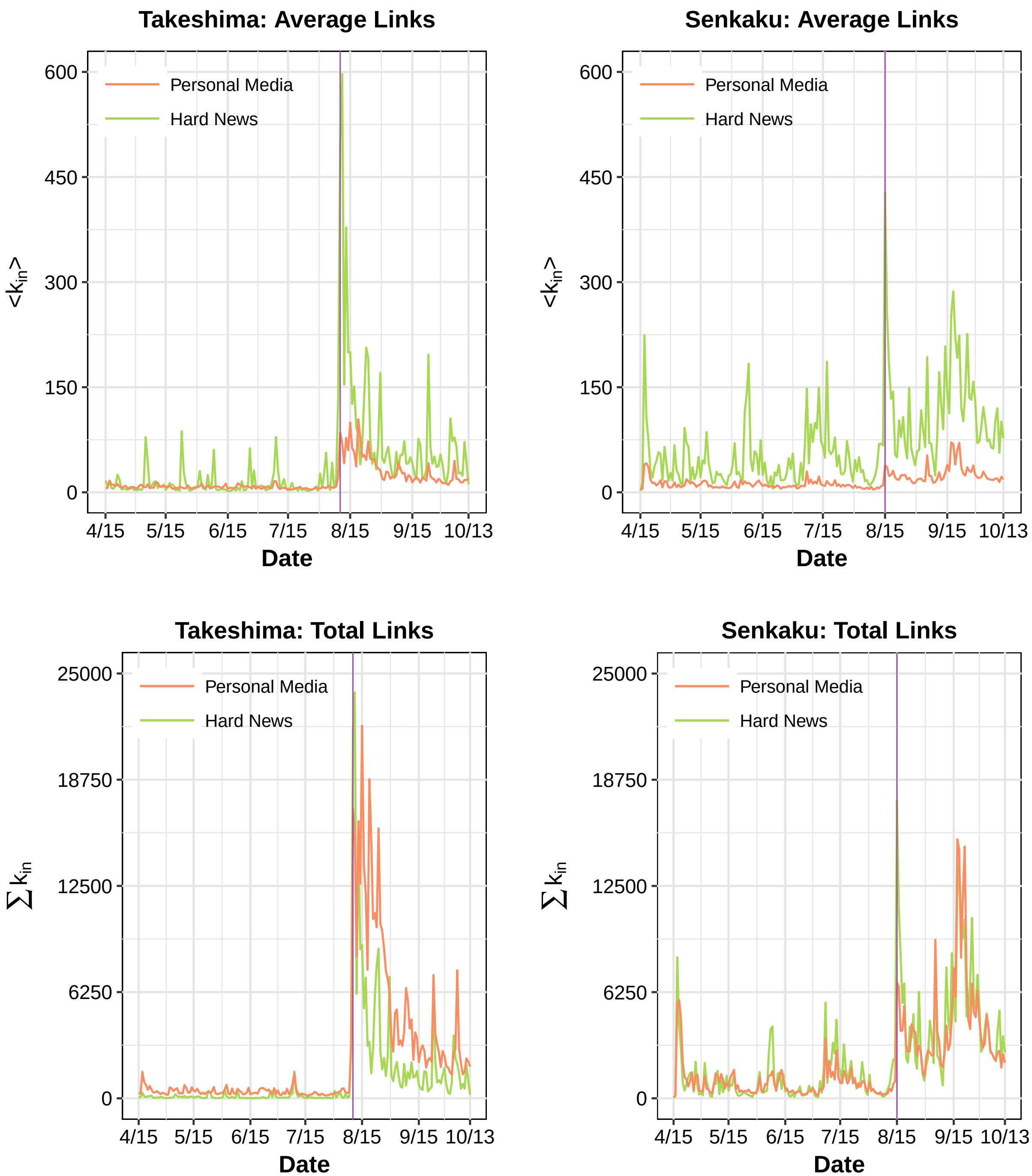

Figure 18: Time-Series Distribution of In-degrees by Opinionated and Factual Domains

in the Japanese twitter information communication network during the rise of territorial dispute and intergroup hostility between Japan and South Korea and China in 2012.

The findings in the first analysis indicate partial support for laboratory and survey based arguments on individual-level information communication behavior. To begin with, salient identity makes twitter users tweet more frequently (H1), more likely to retweet identity holders (H2a), and more frequently link both factual and opinionated domains (H2b). This result is consistent with our expectation that strong identity holder should look for other users with a strong identity, but partially dis-confirming the expectation that they link opinionated source to rationalize their pre-determined position rather than explore factual information. Second, territorial issue incidents involving China (i.e., Senkaku) lead identity holders (but less so for no identity holders) to retweet more to identity holders (H3a, H4). Given the heightened threat, individuals tend to access information aligned with their salient identity. For web-link domains, the incident is followed by more frequent links to both factual and opinionated domains, while there is a difference in the persistence of impact. The impact on opinionated personal media domains persists longer than the impact on factual hard news domains (H3b, H4).

The second analysis goes beyond the implication from lab and survey and assesses time-series change in societal-level network attributes. First, the occurrence of territorial issue incident coincides with the increase in average in-degrees (i.e., tendency that how much a user/domain is retweeted/linked by other users) for users/domains in the network. Second, average in-degree increases more sharply after the incident for national identity holders than for no identity holders, supporting H5a. The total in-degree, on the other hand, shows an opposite pattern supposedly because there are more no identity holders than identity holders to be retweeted. Third, comparing in-degrees for opinionated personal media domains and factual hard news domains, average in-degree increases more sharply after the incident for hard news domains than for personal media domains. Looking from total in-degree, link to opinionated domains and factual domains increase in different timings. The incident impact on total in-degree is immediate for factual domains, but late for opinionated domains. As a consequence, on the incident day, the total in-degree for hard news domain is higher than the personal media domain, but this relationship flips after few days to a month of the incident (H5b).

Two aspects of our findings are worth emphasizing. First, the problem of behavior aggregation is addressed in the comparison of average in-degree and total in-degree. Especially for the difference in in-degree for national identity holders and no identity holders, average in-degree comparison implies that national identity holders are more influential in the network after the incident, while total in-degree comparison implies the opposite. The finding suggests that making a conclusion based on either one of the findings lead to individualistic fallacy or ecological fallacy. Identity holders may become more influential after the incident in terms of average in-degree per individuals, but become less influential after the incident in terms of total in-degree for the group. Second, the dynamic assessment of dependent measures does give deeper insights into the impact of intergroup threat on information communication. For the analysis of web-links, we find that the increase in links after the incident is immediate but in short-lasting for factual domains, and late but long-lasting for opinionated domains. This finding suggests that if we just give a static assessment to one time-point (which is the case for most lab and survey based studies), we may make a misleading conclusion to the overall impact of threat on links to factual and opinionated domains.

This project makes contributions to the field of political behavior in at least three ways. First, the research design enables us to assess questions with a high level of external validity. In contrast to previous lab-experiment and survey based studies, we use field data from the real-world event and incorporate inter-dependency of individuals by embedding the network structure. In addition, the recent study by Bisbee and Larson (2017) shows that the structure of online network closely resembles that of an offline network. It suggests that observed behaviors in online twitter network have implications over behaviors in an offline social network. Second, the findings deepen substantive understandings of intergroup threat, group identity, and motivated reasoning. The narrow focus on individual-level and static data may generate the misleading assessment of theoretical expectations. Third, the method shows further possibilities to use tweets as a tool to analyze political behavior. We wish to demonstrate that machine learning and network analysis are effective ways to extract the rich potential of social media data.

While this paper makes a significant progress in understanding the real-world implications of intergroup threat, group identity and motivated reasoning, there are several caveats in interpreting the result. To start with, even when twitter sentiments are shown to be effective in predicting real-world political phenomena, it should be noted that the behavior of twitter users may not accurately represent the behavior of the general population. They tend to be younger than the general population, and in this study, we only cover users who posted at least one tweet relevant to territorial disputes. It is the best available data to answer our research questions, but we need to be cautious about generalizing the absolute values from our results.

Additionally, three points should be noted about our measurements. First, the measurement of territorial dispute incident is endogenous to data. Now, it is defined as the time of the most extreme increase in the tweet frequency. This procedure measures the threat as perception, not as an exogenous shock. Therefore, our analysis does not capture the causal effect of the exogenous threat. Rather, it captures the behavior and network of threatened individuals (regardless of there is an actual threat or not). Second, the identity of users is coded from user profiles, not tweet texts. Therefore, the coding is *static* and may miss some users who express identity in tweet texts at some point in the period. It is likely that the real identity is *dynamic*, thus the further study can explore the different types of measurement. For now, though, we argue that the current coding is conservative. It depicts users with very strong identity, who has a consistently salient identity. Third, similar to the second point, domain categories are not coded directly from the information contents. The code comes from the static domain address, which does not allow to capture dynamic transformation in the characteristics of their contents.

As the next step in this study, we have at least three major directions to develop the current analysis. First, we can apply machine learning to extract additional characteristics of twitter users. For one, we can analyze the tweet texts to extract positive negative sentiments. For another, we can analyze the retweet and web-link patterns of each user to find the group of users who have similar patterns of retweets or web-links. Second, we can explore the characteristics of the network more deeply. One way is the network community detection (Malliaros and Vazirgiannis, 2013). It is used to separate the network into sub groups to identify smaller communities within the network. After identifying community, we can assess the characteristics of users belong to each community. Lastly, we can assess dynamic rather than static characteristics of users and domains. Measuring and visualizing the dynamic characteristics of users in a dynamic network is a challenging task. The development of such method would contribute the development of network analysis and visualization.

Funding

This work was supported by the International Christian University (ICU) International Academic Exchange Endowment and the National Institute of Informatics (Japan), 2014 Open General Collaborative Research [grant number 47] (In Japanese: 国立情報学研究所 平成26年度共同研究 一般研究公募型［採択番号47］).

References

Barabási, Albert-László. 2016. *Network Science*. Cambridge, UK: Cambridge University Press.

Bermingham, Adam and Alan F. Smeaton. 2011. On Using Twitter to Monitor Political Sentiment and Predict Election Results. In *Sentiment Analysis where AI meets Psychology (SAAIP) Workshop at the International Joint Conference for Natural Language Processing (IJCNLP),13th November 2011*. Chiang Mai, Thailand: .

Bisbee, James and Jennifer M. Larson. 2017. "Testing Social Science Network Theories with Online Network Data: An Evaluation of External Validity." *American Political Science Review* 111(3):502–521.

Bizman, Aharon and Yoel Yinon. 2001. "Intergroup and Interpersonal Threats as Determinants of Prejudice: The Moderating Role of In-Group Identification." *Basic and Applied Social Psychology* 23(3):191–196.

Bollen, Johan, Huina Mao and Xiaojun Zeng. 2011. "Twitter Mood Predicts the Stock Market." *Journal of Computational Science* 2(1):pp. 1–8.

Breiman, Leo. 2001. "Random Forests." *Machine Learning* 45(1):pp. 5–32.

Brylka, Asteria, Tuuli Anna Mähönen and Inga Jasinskaja-Lahti. 2015. "National Identification and Attitudes towards Russian Immigrants in Finland: Investigating the Role of Perceived Threats and Gains." *Scandinavian Journal of Psychology* 56(6):670–677.

Ceron, Andrea, Luigi Curini and Stefano M. Iacus. 2015. "Using Sentiment Analysis to Monitor Electoral Campaigns: Method Matters–Evidence From the United States and Italy." *Social Science Computer Review* 33(1):pp. 3–20.

Ceron, Andrea, Luigi Curini, Stefano M Iacus and Giuseppe Porro. 2014. "Every Tweet Counts? How Sentiment Analysis of Social Media Can Improve Our Knowledge of Citizens'Political Preferences with an Application to Italy and France." *New Media & Society* 16(2):pp. 340–358.

Charles-Toussaint, Gifflene C. and H. Michael Crowson. 2010. "Prejudice against International Students: The Role of Threat Perceptions and Authoritarian Dispositions in U.S. Students." *The Journal of Psychology* 144(5):413–428. PMID: 20806848.

Cleveland, William S. 1979. "Robust Locally Weighted Regression and Smoothing Scatterplots." *Journal of the American Statistical Association* 74(368):829–836.

Corkalo, Dinka and Zeljka Kamenov. 2003. "National Identity and Social Distance: Does In-group Loyalty Lead to Outgroup Hostility?" *Review of Psychology* 10(2):85–94.

Cutler, Adele and John R. Stevens. 2006. [23] Random Forests for Microarrays. In *DNA Microarrays, Part B: Databases and Statistics*, ed. Alan Kimmel and Brian Oliver. Vol. 411 of *Methods in Enzymology* Academic Press pp. 422–432.

Deans, Phil. 2000. "Contending Nationalisms and the Diaoyutai/Senkaku Dispute." *Security Dialogue* 31(1):119–131.

Druckman, Daniel. 1994. "Nationalism, Patriotism, and Group Loyalty: A Social Psychological Perspective." *Mershon International Studies Review* 38(1):pp. 43–68.

Druckman, James N., Jordan Fein and Thomas J. Leeper. 2012. "A Source of Bias in Public Opinion Stability." *The American Political Science Review* 106(2):pp. 430–454.

Fischle, M. 2000. "Mass Response to the Lewinsky Scandal: Motivated Reasoning or Bayesian Updating?" *Political Psychology* 21(1):pp. 135–159.

Games, Paul A and John F Howell. 1976. "Pairwise Multiple Comparison Procedures with Unequal N's and/or Variances: a Monte Carlo Study." *Journal of Educational Statistics* 1(2):113–125.

Gayo-Avello, Daniel. 2012. "No, You Cannot Predict Elections with Twitter." *Internet Computing, IEEE* 16(6):pp. 91–94.

Gibler, Douglas M., Marc L. Hutchison and Steven V. Miller. 2012. "Individual Identity Attachments and International Conflict: The Importance of Territorial Threat." *Comparative Political Studies* .

Harrison, Neil and Nicola Peacock. 2010. "Cultural Distance, Mindfulness and Passive Xenophobia: Using Integrated Threat Theory to Explore Home Higher Education Students' Perspectives on 'Internationalisation at Home'." *British Educational Research Journal* 36(6):877–902.

Hayes, Andrew F. and Klaus Krippendorff. 2007. "Answering the Call for a Standard Reliability Measure for Coding Data." *Communication Methods and Measures* 1(1):pp. 77–89.

Huddy, Leonie and Nadia Khatib. 2007. "American Patriotism, National Identity, and Political Involvement." *American Journal of Political Science* 51(1):pp. 63–77.

Hutchison, Marc L. and Douglas M. Gibler. 2007. "Political Tolerance and Territorial Threat: A Cross-National Study." *The Journal of Politics* 69(1):128–142.

Jin, Mingzhe and Masakatsu Murakami. 2007. "Authorship Identification Using Random Forests." *Proceedings of the Institute of Statistical Mathematics* 55(2):pp. 255–268.

Karasawa, Minoru. 2002. "Patriotism, Nationalism, and Internationalism Among Japanese Citizens: An Etic-Emic Approach." *Political Psychology* 23(4):pp. 645–666.

Killick, Rebecca and Idris Eckley. 2014. "Changepoint: An R package for Changepoint Analysis." *Journal of Statistical Software* 58(3):1–19.

Kim, Sung-youn, Charles S. Taber and Milton Lodge. 2010. "A Computational Model of the Citizen as Motivated Reasoner: Modeling the Dynamics of the 2000 Presidential Election." *Political Behavior* 32(1):1–28.

Kunda, Ziva. 1990. "The Case for Motivated Reasoning." *Psychological Bulletin* 108(3):pp. 480–498.

Kwak, Haewoon, Changhyun Lee, Hosung Park and Sue Moon. 2010. What is Twitter, a Social Network or a News Media? In *Proceedings of the 19th International Conference on World Wide Web*. WWW '10 New York, NY, USA: ACM pp. 591–600.

Li, Qiong and Marilynn B. Brewer. 2004. "What Does It Mean to Be an American? Patriotism, Nationalism, and American Identity after 9/11." *Political Psychology* 25(5):pp. 727–739.

Lodge, Milton and Charles S. Taber. 2013. *The Rationalizing Voter*. New York, NY: Cambridge University Press.

Malliaros, Fragkiskos D. and Michalis Vazirgiannis. 2013. "Clustering and Community Detection in Directed Networks: A Survey." *Physics Reports* 533(4):95 – 142. Clustering and Community Detection in Directed Networks: A Survey.

Mccormack, Gavan. 2000. "Nationalism and Identity in Post-Cold War Japan." *Pacifica Review: Peace, Security & Global Change* 12(3):247–263.

Ministry of Internal Affairs and Communications. 2012. *Joho Tsushin Medhia no Riyo Jikan to Joho Kodo ni Kansuru Chosa Hokokusho [The Report of Survey on Use Time of Information Communcation Media and Information Behavior]*.

Morrison, Kimberly Rios and Oscar Ybarra. 2008. "The Effects of Realistic Threat and Group Identification on Social Dominance Orientation." *Journal of Experimental Social Psychology* 44(1):pp. 156–163.

Murthy, Dhiraj. 2015. "Twitter and Elections: Are Tweets, Predictive, Reactive, or a Form of Buzz?" *Information, Communication & Society* 18(7):pp. 816–831.

Riek, Blake M., Eric W. Mania and Samuel L. Gaertner. 2006. "Intergroup Threat and Outgroup Attitudes: A Meta-Analytic Review." *Personality and Social Psychology Review* 10(4):pp. 336–353.

Rose, Caroline. 2000. "'Patriotism is Not Taboo': Nationalism in China and Japan and Implications for Sino-Japanese Relations." *Japan Forum* 12(2):pp. 169–181.

Rozman, Gilbert. 2013. "Narrowing the Gap between China and Japan: Three Dimensions of National Identity and the Korean Factor." *Japanese Journal of Political Science* 14(1):31?49.

Sasada, Hironori. 2006. "Youth and Nationalism in Japan." *SAIS Review of International Affairs* 26(2):pp. 109–122.

Stephan, Walter G and Cookie White Stephan. 2000. An Integrated Threat Theory of Prejudice. In *Reducing Prejudice and Discrimination*, ed. Stuart Oskamp. Lawrence Erlbaum Associates pp. 23–45.

Stephan, Walter G., Kurt A. Boniecki, Oscar Ybarra, Ann Bettencourt, Kelly S. Ervin, Linda A.

Jackson, Penny S. McNatt and C. Lausanne Renfro. 2002. “The Role of Threats in the Racial Attitudes of Blacks and Whites.” *Personality and Social Psychology Bulletin* 28(9):1242–1254.

Stephan, Walter G. and Marisa D. Mealy. 2011. Intergroup Threat Theory. In *The Encyclopedia of Peace Psychology*, ed. Daniel J. Christie. Blackwell Publishing Ltd.

Stephan, Walter G., Oscar Ybarra, Carmen Martnez Martnez, Joseph Schwarzwald and Michal Tur-Kaspa. 1998. “Prejudice toward Immigrants to Spain and Israel: An Integrated Threat Theory Analysis.” *Journal of Cross-Cultural Psychology* 29(4):pp. 559–576.

Stephan, Walter G., Rolando Diaz-Loving and Anne Duran. 2000. “Integrated Threat Theory and Intercultural Attitudes.” *Journal of Cross-Cultural Psychology* 31(2):240–249.

Suh, B., Lichan Hong, P. Pirolli and Ed H. Chi. 2010. Want to be Retweeted? Large Scale Analytics on Factors Impacting Retweet in Twitter Network. In *2010 IEEE Second International Conference on Social Computing (SocialCom)*. pp. 177–184.

Suzuki, Shogo. 2015. “The Rise of the Chinese ‘Other’ in Japan’s Construction of Identity: Is China a Focal Point of Japanese Nationalism?” *The Pacific Review* 28(1):95–116.

Suzuki, Takafumi. 2009. “Extracting Speaker-specific Functional Expressions from Political Speeches Using Random Forests in Order to Investigate Speakers’ Political Styles.” *Journal of the American Society for Information Science and Technology* 60(8):pp. 1596–1606.

Taber, Charles S. and Milton Lodge. 2006A. “Motivated Skepticism in the Evaluation of Political Beliefs.” *American Journal of Political Science* 50(3):755–769.

Tajfel, Henri. 1974. “Social Identity and Intergroup Behaviour.” *Information (International Social Science Council)* 13(2):65–93.

Tajfel, Henri and John Turner. 1979. An Integrative Theory of Intergroup Conflict. In *The Social Psychology of Intergroup Relations*. Brooks Cole Publishing.

Tanaka, Seiki. 2015. “The Microfoundations of Territorial Disputes: Evidence from a Survey Experiment in Japan.” *Conflict Management and Peace Science* .

Tumasjan, Andranik, Timm O. Sprenger, Philipp G. Sandner and Isabell M. Welpe. 2011. “Election Forecasts With Twitter: How 140 Characters Reflect the Political Landscape.” *Social Science Computer Review* 29(4):pp. 402–418.

Tversky, Amos and Daniel Kahneman. 1974. “Judgment Under Uncertainty: Heuristics and Biases.” *Science* 185(4157):pp. 1124–1131.

Velasco González, Karina, Maykel Verkuyten, Jeroen Weesie and Edwin Poppe. 2008. “Prejudice towards Muslims in The Netherlands: Testing Integrated Threat Theory.” *British Journal of Social Psychology* 47(4):667–685.

Verkuyten, Maykel. 2009. “Support for Multiculturalism and Minority Rights: The Role of National Identification and Out-group Threat.” *Social Justice Research* 22(1):31–52.

Yogeeswaran, Kumar and Nilanjana Dasgupta. 2014. "Conceptions of National Identity in a Globalised World: Antecedents and Consequences." *European Review of Social Psychology* 25(1):189–227.

Appendix

A The List of Domain Coding

Domains denoted as "SUPPRESSED" contain some personal information/names. Complete lists are available upon request from the authors.

Table A.1: Coding of Domain Characteristics (Ranking 1-100)

Rank	Domain	Count	Category	Rank	Domain	Count	Category
1	sankei.jp.msn.com	241686	Hard News	51	SUPPRESSED	6846	Personal Media
2	news.yahoo.co.jp	170416	Curated News	52	www.news-postseven.com	6635	Soft News
3	t.co	136126	Others	53	SUPPRESSED	6631	Personal Media
4	news.nicovideo.jp	92943	Soft News	54	jp.reuters.com	6015	Hard News
5	www.yomiuri.co.jp	63946	Hard News	55	www.fnn-news.com	5942	Hard News
6	blog.livedoor.jp	61567	Personal Media	56	www.tokyo-np.co.jp	5833	Hard News
7	twitter.com	60485	Personal Media	57	p.twipple.jp	5747	Personal Media
8	www.youtube.com	59490	Personal Media	58	matome.naver.jp	5659	Personal Media
9	www3.nhk.or.jp	48251	Hard News	59	SUPPRESSED	5582	Personal Media
10	www.47news.jp	45575	Hard News	60	zasshi.news.yahoo.co.jp	5543	Curated News
11	www.nikkei.com	40514	Hard News	61	SUPPRESSED	5310	Personal Media
12	www.asahi.com	38730	Hard News	62	www.recordchina.co.jp	4993	Others
13	www.jiji.com	37703	Hard News	63	www.excite.co.jp	4960	Curated News
14	SUPPRESSED	37479	Personal Media	64	SUPPRESSED	4861	Personal Media
15	www.amazon.co.jp	34501	Others	65	headlines.yahoo.co.jp	4858	Curated News
16	SUPPRESSED	34024	Personal Media	66	news.nifty.com	4572	Curated News
17	news.google.com	33357	Curated News	67	news.tbs.co.jp	4572	Hard News
18	mainichi.jp	29761	Hard News	68	blogs.yahoo.co.jp	4557	Personal Media
19	news.livedoor.com	28330	Curated News	69	SUPPRESSED	4554	Personal Media
20	SUPPRESSED	25134	Personal Media	70	www.rakuten.co.jp	4482	Others
21	SUPPRESSED	23125	Others	71	www.nikkansports.com	4465	Soft News
22	SUPPRESSED	21252	Personal Media	72	business.nikkeibp.co.jp	4464	Hard News
23	on-msn.com	20810	Curated News	73	jp.wsj.com	4255	Hard News
24	twitpic.com	17587	Personal Media	74	SUPPRESSED	4142	Personal Media
25	error.fc2.com	16907	Personal Media	75	SUPPRESSED	4043	Personal Media
26	www.zakzak.co.jp	16769	Soft News	76	news.google.co.jp	3964	Curated News
27	SUPPRESSED	16256	Personal Media	77	SUPPRESSED	3960	Personal Media
28	rd.yahoo.co.jp	15999	Personal Media	78	SUPPRESSED	3625	Personal Media
29	www.searchina.net	15374	Soft News	79	www.sponichi.co.jp	3620	Soft News
30	www.j-cast.com	15268	Soft News	80	www.google.com	3538	Others
31	SUPPRESSED	14476	Personal Media	81	www.chijihon.metro.tokyo.jp	3431	Personal Media
32	SUPPRESSED	13695	Personal Media	82	is.gd	3430	Others
33	SUPPRESSED	12502	Personal Media	83	news.infoseek.co.jp	3299	Curated News
34	www.nicovideo.jp	12384	Personal Media	84	SUPPRESSED	3196	Curated News
35	blog.goo.ne.jp	12164	Personal Media	85	SUPPRESSED	3119	Others
36	togetter.com	11789	Personal Media	86	SUPPRESSED	3036	Personal Media
37	SUPPRESSED	11626	Others	87	get-lives.info	2994	Curated News
38	SUPPRESSED	10684	Personal Media	88	www.uniqlo.com	2863	Others
39	blogos.com	10462	Soft News	89	getnews.jp	2855	Soft News
40	www.facebook.com	10093	Personal Media	90	prize-survey.com	2797	Others
41	news.goo.ne.jp	10034	Curated News	91	nicoimage.com	2792	Personal Media
42	SUPPRESSED	9898	Personal Media	92	petitions.whitehouse.gov	2748	Others
43	SUPPRESSED	9567	Personal Media	93	www.hugedomains.com	2734	Others
44	news.biglobe.ne.jp	8560	Curated News	94	www.tv-asahi.co.jp	2726	Hard News
45	SUPPRESSED	8147	Personal Media	95	www.news24.jp	2679	Hard News
46	live.nicovideo.jp	7505	Personal Media	96	www.okinawatimes.co.jp	2645	Hard News
47	www.iza.ne.jp	7496	Soft News	97	SUPPRESSED	2604	Personal Media
48	SUPPRESSED	7262	Personal Media	98	survey-winner.net	2586	Others
49	SUPPRESSED	7090	Personal Media	99	SUPPRESSED	2563	Personal Media
50	t.c	7034	Others	100	the-liberty.com	2544	Personal Media

Table A.2: Coding of Domain Characteristics (Ranking 101-200)

Rank	Domain	Count	Category	Rank	Domain	Count	Category
101	fc2.in	2530	Personal Media	151	pbs.twimg.com	1594	Personal Media
102	SUPPRESSED	2523	Personal Media	152	SUPPRESSED	1585	Others
103	SUPPRESSED	2491	Personal Media	153	news.ameba.jp	1574	Curated News
104	www.chosunonline.com	2488	Others	154	ptic.jp	1558	Personal Media
105	SUPPRESSED	2468	Personal Media	155	detail.chiebukuro.yahoo.co.jp	1541	Personal Media
106	kinbricksnow.com	2395	Soft News	156	url.os7.biz	1534	Others
107	SUPPRESSED	2383	Personal Media	157	SUPPRESSED	1504	Personal Media
108	jp.WSJ.com	2381	Hard News	158	SUPPRESSED	1503	Personal Media
109	www.mofa.go.jp	2378	Others	159	SUPPRESSED	1495	Personal Media
110	SUPPRESSED	2266	Personal Media	160	SUPPRESSED	1492	Personal Media
111	www.newsweekjapan.jp	2254	Hard News	161	SUPPRESSED	1486	Personal Media
112	london2012.nikkansports.com	2233	Soft News	162	SUPPRESSED	1479	Personal Media
113	www.ustream.tv	2213	Personal Media	163	SUPPRESSED	1474	Personal Media
114	www.whitehouse.gov	2159	Others	164	www.nikkeibp.co.jp	1468	Hard News
115	account.nicovideo.jp	2137	Personal Media	165	www.twitlonger.com	1458	Others
116	wedge.ismedia.jp	2132	Hard News	166	SUPPRESSED	1447	Personal Media
117	fb.me	2118	Personal Media	167	pt.afl.rakuten.co.jp	1416	Others
118	ryukyushimpo.jp	2067	Hard News	168	SUPPRESSED	1391	Personal Media
119	SUPPRESSED	2024	Personal Media	169	SUPPRESSED	1388	Personal Media
120	SUPPRESSED	1999	Personal Media	170	football-station.net	1379	Curated News
121	news.guideme.jp	1991	Curated News	171	SUPPRESSED	1354	Personal Media
122	agora-web.jp	1990	Soft News	172	sp.recordchina.co.jp	1334	Others
123	www.chugoku-np.co.jp	1979	Hard News	173	www.afpbb.com	1334	Hard News
124	SUPPRESSED	1971	Personal Media	174	SUPPRESSED	1331	Personal Media
125	news.tv-asahi.co.jp	1944	Hard News	175	www.sanspo.com	1330	Soft News
126	jbpress.ismedia.jp	1939	Soft News	176	SUPPRESSED	1326	Personal Media
127	www.pref.shimane.lg.jp	1921	Others	177	www.nishinippon.co.jp	1325	Hard News
128	SUPPRESSED	1906	Personal Media	178	www.kantei.go.jp	1322	Others
129	SUPPRESSED	1905	Personal Media	179	SUPPRESSED	1318	Personal Media
130	www.bloomberg.co.jp	1902	Hard News	180	SUPPRESSED	1294	Personal Media
131	www.yaeyama-nippo.com	1897	Hard News	181	SUPPRESSED	1294	Personal Media
132	d.hatena.ne.jp	1881	Personal Media	182	SUPPRESSED	1286	Personal Media
133	dd.hokkaido-np.co.jp	1863	Hard News	183	SUPPRESSED	1277	Personal Media
134	socialnews.rakuten.co.jp	1863	Curated News	184	j.plustar.jp	1269	Others
135	SUPPRESSED	1858	Personal Media	185	2naname.com	1262	Others
136	imageshack.com	1857	Personal Media	186	SUPPRESSED	1253	Personal Media
137	urx.nu	1832	Others	187	SUPPRESSED	1253	Personal Media
138	SUPPRESSED	1817	Personal Media	188	jp.sputniknews.com	1250	Others
139	ja.wikipedia.org	1789	Others	189	SUPPRESSED	1245	Personal Media
140	www.jcp.or.jp	1761	Personal Media	190	SUPPRESSED	1240	Personal Media
141	SUPPRESSED	1746	Personal Media	191	SUPPRESSED	1240	Personal Media
142	ja.m.wikipedia.org	1714	Others	192	www.chunichi.co.jp	1224	Hard News
143	SUPPRESSED	1706	Personal Media	193	micro.asahi.com	1194	Hard News
144	feeds.feedburner.com	1680	Others	194	SUPPRESSED	1188	Personal Media
145	www.sankeibiz.jp	1673	Hard News	195	itunes.apple.com	1184	Others
146	SUPPRESSED	1666	Personal Media	196	www.livedoor.com	1180	Others
147	SUPPRESSED	1655	Personal Media	197	hochi.yomiuri.co.jp	1179	Soft News
148	SUPPRESSED	1653	Personal Media	198	SUPPRESSED	1162	Others
149	SUPPRESSED	1637	Others	199	seiji.yahoo.co.jp	1151	Curated News
150	SUPPRESSED	1628	Personal Media	200	po.st	1144	Others

Table A.3: Coding of Domain Characteristics (Ranking 201-300)

Rank	Domain	Count	Category	Rank	Domain	Count	Category
201	www.google.co.jp	1141	Others	251	SUPPRESSED	768	Personal Media
202	diamond.jp	1141	Hard News	252	SUPPRESSED	760	Personal Media
203	SUPPRESSED	1128	Personal Media	253	ictmax.org	749	Curated News
204	www.dailymotion.com	1123	Personal Media	254	SUPPRESSED	749	Personal Media
205	ceron.jp	1122	Personal Media	255	SUPPRESSED	733	Personal Media
206	gendai.net	1119	Soft News	256	SUPPRESSED	729	Personal Media
207	rocketnews24.com	1113	Soft News	257	www.takeshimakyoto.jp	715	Others
208	tweetbuzz.jp	1107	Personal Media	258	SUPPRESSED	713	Soft News
209	SUPPRESSED	1103	Personal Media	259	SUPPRESSED	708	Personal Media
210	news.mixi.jp	1090	Curated News	260	SUPPRESSED	707	Personal Media
211	realtime.wsj.com	1090	Hard News	261	plaza.rakuten.co.jp	685	Others
212	gendai.ismedia.jp	1082	Hard News	262	www.mod.go.jp	678	Others
213	SUPPRESSED	1078	Soft News	263	www.cnn.co.jp	677	Hard News
214	SUPPRESSED	1075	Personal Media	264	SUPPRESSED	673	Personal Media
215	SUPPRESSED	1065	Personal Media	265	SUPPRESSED	671	Personal Media
216	www.letscorp.net	1059	Others	266	SUPPRESSED	671	Personal Media
217	SUPPRESSED	1058	Personal Media	267	SUPPRESSED	669	Personal Media
218	tr.twipple.jp	1054	Personal Media	268	taipeiquote.com	661	Curated News
219	www.yahoo.co.jp	1033	Others	269	www.jimin.jp	661	Personal Media
220	SUPPRESSED	1007	Personal Media	270	SUPPRESSED	656	Personal Media
221	seiga.nicovideo.jp	1004	Personal Media	271	SUPPRESSED	644	Personal Media
222	SUPPRESSED	1001	Personal Media	272	SUPPRESSED	644	Personal Media
223	digital.asahi.com	1001	Hard News	273	SUPPRESSED	632	Personal Media
224	nureinu.net	997	Others	274	logmemo.org	629	Personal Media
225	SUPPRESSED	972	Personal Media	275	jump.cx	628	Others
226	SUPPRESSED	972	Others	276	SUPPRESSED	626	Personal Media
227	SUPPRESSED	966	Personal Media	277	japanese.donga.com	622	Others
228	SUPPRESSED	957	Personal Media	278	SUPPRESSED	619	Personal Media
229	SUPPRESSED	949	Others	279	twitcasting.tv	614	Personal Media
230	SUPPRESSED	943	Personal Media	280	SUPPRESSED	612	Personal Media
231	SUPPRESSED	933	Personal Media	281	SUPPRESSED	612	Personal Media
232	SUPPRESSED	921	Personal Media	282	www.e-gov.go.jp	611	Others
233	SUPPRESSED	921	Personal Media	283	rabitsokuhou.2chblog.jp	591	Personal Media
234	www.tokyo-sports.co.jp	919	Soft News	284	news-social.com	591	Curated News
235	SUPPRESSED	911	Personal Media	285	www.metro.tokyo.jp	588	Others
236	SUPPRESSED	905	Personal Media	286	photozou.jp	588	Personal Media
237	SUPPRESSED	902	Personal Media	287	london.yahoo.co.jp	588	Curated News
238	SUPPRESSED	894	Personal Media	288	SUPPRESSED	585	Personal Media
239	SUPPRESSED	894	Personal Media	289	v.qq.com	585	Personal Media
240	SUPPRESSED	892	Personal Media	290	president.jp	584	Hard News
241	SUPPRESSED	883	Personal Media	291	SUPPRESSED	583	Personal Media
242	SUPPRESSED	879	Personal Media	292	SUPPRESSED	578	Personal Media
243	www.nhk.or.jp	873	Hard News	293	inewsjp.com	578	Curated News
244	www.googmail.info	862	Others	294	rentalserver.fc2.com	576	Personal Media
245	SUPPRESSED	858	Personal Media	295	sportsnavi.yahoo.co.jp	576	Soft News
246	SUPPRESSED	856	Personal Media	296	soccerunderground.com	575	Soft News
247	i.jiji.jp	854	Hard News	297	SUPPRESSED	574	Personal Media
248	SUPPRESSED	837	Personal Media	298	SUPPRESSED	568	Personal Media
249	SUPPRESSED	804	Personal Media	299	SUPPRESSED	567	Personal Media
250	www.viglink.com:80	795	Others	300	SUPPRESSED	565	Personal Media

Table A.4: Coding of Domain Characteristics (Ranking 301-400)

Rank	Domain	Count	Category	Rank	Domain	Count	Category
301	SUPPRESSED	562	Personal Media	351	SUPPRESSED	421	Personal Media
302	SUPPRESSED	552	Personal Media	352	SUPPRESSED	412	Personal Media
303	market.android.com	547	Others	353	twitenna.jp	410	Others
304	SUPPRESSED	545	Personal Media	354	SUPPRESSED	410	Personal Media
305	SUPPRESSED	544	Personal Media	355	SUPPRESSED	405	Personal Media
306	www.msn.com	536	Others	356	www.oita-press.co.jp	400	Hard News
307	wpb.shueisha.co.jp	531	Soft News	357	news.merumo.ne.jp	400	Soft News
308	www.cyzowoman.com	525	Soft News	358	www.instagram.com	398	Personal Media
309	p.twimg.com	522	Personal Media	359	SUPPRESSED	397	Personal Media
310	SUPPRESSED	521	Personal Media	360	news.so-net.ne.jp	394	Curated News
311	SUPPRESSED	519	Personal Media	361	SUPPRESSED	394	Personal Media
312	tinyurl.com	518	Others	362	plus.google.com	393	Personal Media
313	SUPPRESSED	515	Personal Media	363	SUPPRESSED	393	Personal Media
314	SUPPRESSED	508	Personal Media	364	SUPPRESSED	393	Personal Media
315	SUPPRESSED	507	Personal Media	365	SUPPRESSED	391	Personal Media
316	goo.gl	506	Others	366	SUPPRESSED	389	Personal Media
317	SUPPRESSED	501	Personal Media	367	SUPPRESSED	382	Personal Media
318	SUPPRESSED	500	Personal Media	368	www.fukuishimbun.co.jp	382	Hard News
319	SUPPRESSED	488	Personal Media	369	kyushu.yomiuri.co.jp	378	Hard News
320	www.mesotw.com	488	Others	370	SUPPRESSED	369	Personal Media
321	www.cyzo.com	487	Soft News	371	SUPPRESSED	367	Personal Media
322	SUPPRESSED	485	Personal Media	372	shikakukoushin.jimdo.com	366	Others
323	www.hokkoku.co.jp	485	Hard News	373	SANSPO.COM	365	Soft News
324	SUPPRESSED	484	Personal Media	374	SUPPRESSED	364	Personal Media
325	news.kstyle.com	481	Soft News	375	SUPPRESSED	362	Personal Media
326	news.ifeng.com	480	Others	376	SUPPRESSED	359	Personal Media
327	ww11.gigamode.net	479	Others	377	SUPPRESSED	359	Personal Media
328	www.geocities.jp	473	Personal Media	378	SUPPRESSED	353	Personal Media
329	SUPPRESSED	468	Personal Media	379	SUPPRESSED	352	Personal Media
330	ow.ly	468	Others	380	www.daily.co.jp	352	Soft News
331	SUPPRESSED	468	Personal Media	381	SUPPRESSED	347	Personal Media
332	okwave.jp	466	Personal Media	382	SUPPRESSED	347	Personal Media
333	SUPPRESSED	463	Personal Media	383	SUPPRESSED	345	Personal Media
334	linkenquete.web.fc2.com	462	Others	384	SUPPRESSED	344	Personal Media
335	SUPPRESSED	461	Personal Media	385	www.yukawanet.com	343	Soft News
336	m.youtube.com	461	Personal Media	386	SUPPRESSED	342	Personal Media
337	SUPPRESSED	459	Personal Media	387	SUPPRESSED	341	Personal Media
338	SUPPRESSED	459	Hard News	388	SUPPRESSED	341	Personal Media
339	SUPPRESSED	458	Personal Media	389	SUPPRESSED	337	Personal Media
340	SUPPRESSED	444	Personal Media	390	www.yyc.co.jp	337	Personal Media
341	SUPPRESSED	442	Personal Media	391	SUPPRESSED	336	Personal Media
342	SUPPRESSED	441	Personal Media	392	SUPPRESSED	333	Personal Media
343	vippers.jp	441	Personal Media	393	photo.sankei.jp.msn.com	332	Hard News
344	SUPPRESSED	440	Personal Media	394	SUPPRESSED	330	Personal Media
345	2ii.jp	439	Others	395	SUPPRESSED	326	Personal Media
346	SUPPRESSED	431	Personal Media	396	SUPPRESSED	325	Personal Media
347	SUPPRESSED	430	Personal Media	397	SUPPRESSED	325	Personal Media
348	SUPPRESSED	423	Personal Media	398	www.cinematoday.jp	325	Soft News
349	SUPPRESSED	423	Personal Media	399	www.causes.com	324	Personal Media
350	SUPPRESSED	421	Personal Media	400	SUPPRESSED	323	Personal Media

Table A.5: Coding of Domain Characteristics (Ranking 401-500)

Rank	Domain	Count	Category	Rank	Domain	Count	Category
401	SUPPRESSED	321	Personal Media	451	SUPPRESSED	253	Personal Media
402	SUPPRESSED	320	Personal Media	452	SUPPRESSED	253	Personal Media
403	SUPPRESSED	319	Personal Media	453	cro.st	250	Others
404	SUPPRESSED	315	Personal Media	454	getgold.jp	247	Soft News
405	SUPPRESSED	309	Personal Media	455	SUPPRESSED	246	Personal Media
406	SUPPRESSED	309	Personal Media	456	SUPPRESSED	244	Soft News
407	SUPPRESSED	309	Personal Media	457	mixi.jp	243	Personal Media
408	SUPPRESSED	308	Personal Media	458	SUPPRESSED	243	Personal Media
409	2crd.com	307	Others	459	SUPPRESSED	243	Personal Media
410	www.kahoku.co.jp	306	Hard News	460	SUPPRESSED	242	Personal Media
411	www.kobe-np.co.jp	304	Hard News	461	SUPPRESSED	242	Personal Media
412	SUPPRESSED	304	Personal Media	462	www3.to	241	Others
413	SUPPRESSED	302	Personal Media	463	buzztter.com	241	Personal Media
414	SUPPRESSED	296	Personal Media	464	00fenrir.com	240	Others
415	SUPPRESSED	296	Personal Media	465	www.ch-sakura.jp	239	Soft News
416	www.itmedia.co.jp	294	Soft News	466	SUPPRESSED	237	Personal Media
417	www.chibanippo.co.jp	293	Hard News	467	SUPPRESSED	237	Personal Media
418	archive.gohoo.org	292	Others	468	SUPPRESSED	236	Personal Media
419	SUPPRESSED	292	Soft News	469	www.tv-tokyo.co.jp	236	Hard News
420	www.shikoku-np.co.jp	291	Hard News	470	SUPPRESSED	236	Personal Media
421	news-log.jp	289	Soft News	471	SUPPRESSED	234	Personal Media
422	imgnews.naver.net	286	Personal Media	472	SUPPRESSED	232	Others
423	www.followjp.com	285	Personal Media	473	SUPPRESSED	231	Personal Media
424	www.fiat-auto.co.jp	284	Others	474	SUPPRESSED	231	Personal Media
425	SUPPRESSED	283	Personal Media	475	gurumi-ch.com	231	Others
426	SUPPRESSED	282	Personal Media	476	news.mynavi.jp	230	Curated News
427	bit.ly	281	Others	477	SUPPRESSED	230	Personal Media
428	SUPPRESSED	277	Personal Media	478	SUPPRESSED	230	Personal Media
429	nikkansports.com	277	Soft News	479	SUPPRESSED	228	Personal Media
430	woman.infoseek.co.jp	273	Soft News	480	SUPPRESSED	227	Personal Media
431	SUPPRESSED	273	Personal Media	481	mns.jp	226	Others
432	SUPPRESSED	272	Personal Media	482	www.nytimes.com	226	Hard News
433	www.wa-dan.com	270	Hard News	483	SUPPRESSED	225	Others
434	SUPPRESSED	269	Personal Media	484	SUPPRESSED	222	Curated News
435	SUPPRESSED	269	Personal Media	485	SUPPRESSED	222	Hard News
436	SUPPRESSED	267	Personal Media	486	SUPPRESSED	222	Personal Media
437	SUPPRESSED	267	Personal Media	487	SUPPRESSED	221	Personal Media
438	www.y-mainichi.co.jp	265	Hard News	488	SUPPRESSED	220	Personal Media
439	via.me	265	Personal Media	489	SUPPRESSED	220	Personal Media
440	SUPPRESSED	263	Personal Media	490	SUPPRESSED	219	Personal Media
441	SUPPRESSED	262	Personal Media	491	SUPPRESSED	219	Personal Media
442	SUPPRESSED	262	Personal Media	492	SUPPRESSED	218	Personal Media
443	www.cybozu.net	262	Curated News	493	SUPPRESSED	217	Personal Media
444	SUPPRESSED	261	Personal Media	494	SUPPRESSED	217	Personal Media
445	research-panel.jp	260	Others	495	twishort.com	216	Others
446	www.tachiagare.jp	259	Personal Media	496	jcc.jp	216	Curated News
447	play.google.com	257	Others	497	logtters.jp	216	Personal Media
448	SUPPRESSED	255	Personal Media	498	video.fc2.com	215	Personal Media
449	SUPPRESSED	254	Personal Media	499	box600.bluehost.com	214	Personal Media
450	SUPPRESSED	253	Personal Media	500	www.lib.city.minato.tokyo.jp	214	Others

B Mean Comparisons of Pre- and Post-Incident Information Behavior

Table B.1: Pre- and Post-Incident Mean Comparisons of the Result in Figure 9 (Means are Calculated from the Entire Period of Before/After the Incident)

ID Type	Issue	ID Stat.	Mean (Before)	SE	Mean (After)	SE	Mean (After-Before)	SE
Nationalism	Takeshima	No ID	0.280	0.010	0.227	0.007	-0.053	0.012
Nationalism	Takeshima	Have ID	0.443	0.008	0.395	0.007	-0.048	0.011
Patriotism	Takeshima	No ID	0.432	0.014	0.328	0.009	-0.104	0.016
Patriotism	Takeshima	Have ID	0.635	0.009	0.526	0.008	-0.108	0.013
Nationalism	Senkaku	No ID	0.249	0.008	0.179	0.006	-0.070	0.010
Nationalism	Senkaku	Have ID	0.410	0.008	0.363	0.007	-0.047	0.010
Patriotism	Senkaku	No ID	0.364	0.010	0.284	0.008	-0.080	0.013
Patriotism	Senkaku	Have ID	0.551	0.008	0.472	0.008	-0.079	0.011

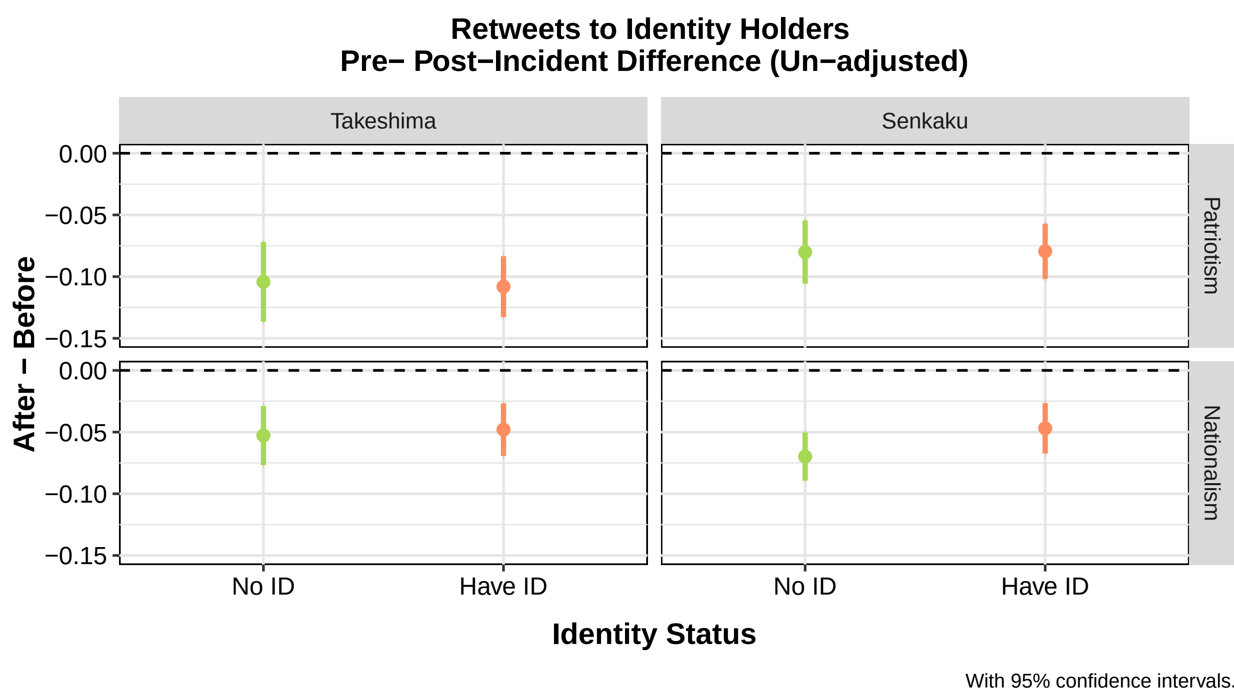

Figure B.1: Mean Differences of the Result in Figure 9 (Means are Calculated from the Entire Period of Before/After the Incident)

Table B.2: Pre- and Post-Incident Mean Comparisons of the Result in Figure 11 (Means are Calculated from the Entire Period of Before/After the Incident)

ID Type	Issue	ID Stat.	Mean (Before)	SE (Before)	Mean (After)	SE (After)	Mean (After-Before)	SE (After-Before)
Nationalism	Takeshima	No ID	1.461	0.058	1.721	0.050	0.260	0.077
Nationalism	Takeshima	Have ID	2.344	0.065	3.021	0.069	0.677	0.095
Patriotism	Takeshima	No ID	1.622	0.062	1.642	0.047	0.020	0.078
Patriotism	Takeshima	Have ID	2.409	0.056	2.652	0.053	0.243	0.077
Nationalism	Senkaku	No ID	1.287	0.036	1.422	0.034	0.135	0.049
Nationalism	Senkaku	Have ID	2.194	0.059	3.002	0.061	0.808	0.085
Patriotism	Senkaku	No ID	1.360	0.038	1.463	0.033	0.104	0.050
Patriotism	Senkaku	Have ID	2.097	0.046	2.498	0.036	0.401	0.058

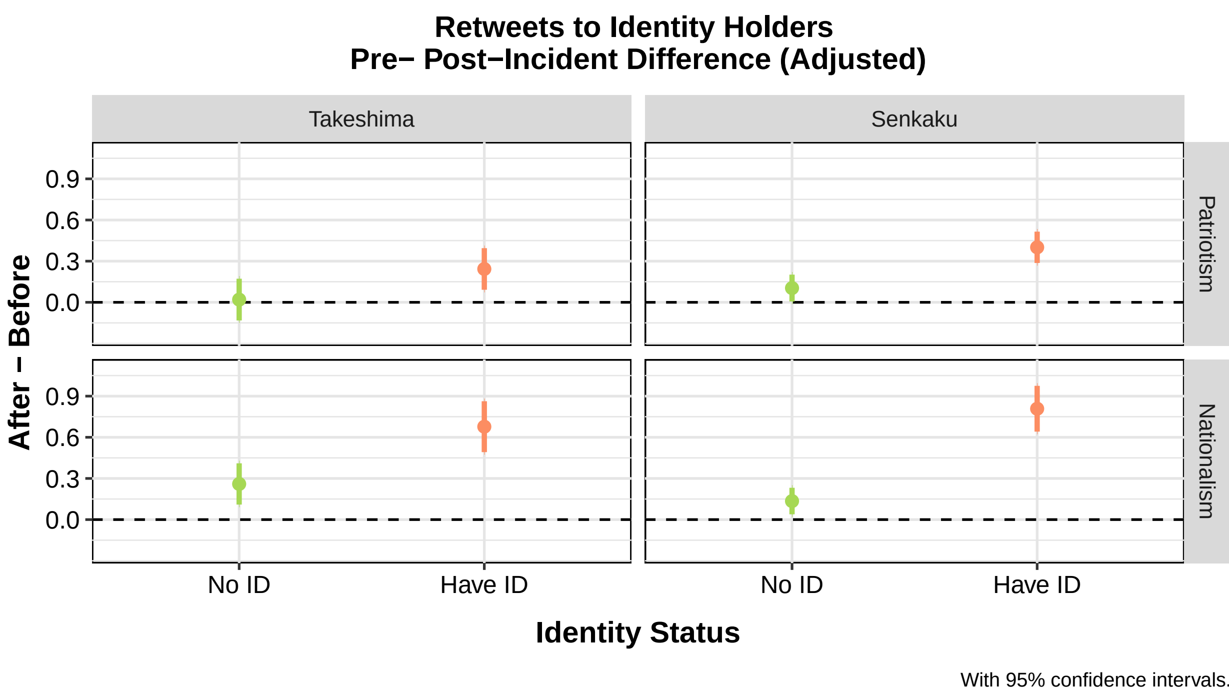

Figure B.2: Mean Differences of the Result in Figure 11 (Means are Calculated from the Entire Period of Before/After the Incident)

Table B.3: Pre- and Post-Incident Mean Comparisons of the Result in Figure 12 (Means are Calculated from the Entire Period of Before/After the Incident)

Domain Cat.	Issue	ID Stat.	Mean	SE	Mean	SE	Mean	SE
			(Before)		(After)		(After-Before)	
Hard News	Takeshima	No ID	0.061	0.007	0.212	0.014	0.151	0.016
Hard News	Takeshima	Have ID	0.091	0.009	0.329	0.024	0.237	0.025
Curated News	Takeshima	No ID	0.043	0.004	0.129	0.007	0.085	0.008
Curated News	Takeshima	Have ID	0.048	0.005	0.155	0.010	0.106	0.011
Soft News	Takeshima	No ID	0.038	0.005	0.081	0.005	0.043	0.007
Soft News	Takeshima	Have ID	0.034	0.003	0.088	0.006	0.054	0.007
Personal Media	Takeshima	No ID	0.288	0.010	0.436	0.014	0.148	0.017
Personal Media	Takeshima	Have ID	0.423	0.012	0.741	0.017	0.318	0.021
Hard News	Senkaku	No ID	0.265	0.012	0.364	0.014	0.098	0.018
Hard News	Senkaku	Have ID	0.337	0.019	0.508	0.024	0.171	0.031
Curated News	Senkaku	No ID	0.108	0.007	0.185	0.008	0.077	0.010
Curated News	Senkaku	Have ID	0.099	0.007	0.219	0.011	0.120	0.014
Soft News	Senkaku	No ID	0.066	0.004	0.098	0.005	0.031	0.007
Soft News	Senkaku	Have ID	0.070	0.004	0.101	0.005	0.031	0.007
Personal Media	Senkaku	No ID	0.264	0.007	0.306	0.011	0.042	0.013
Personal Media	Senkaku	Have ID	0.363	0.009	0.485	0.019	0.122	0.021

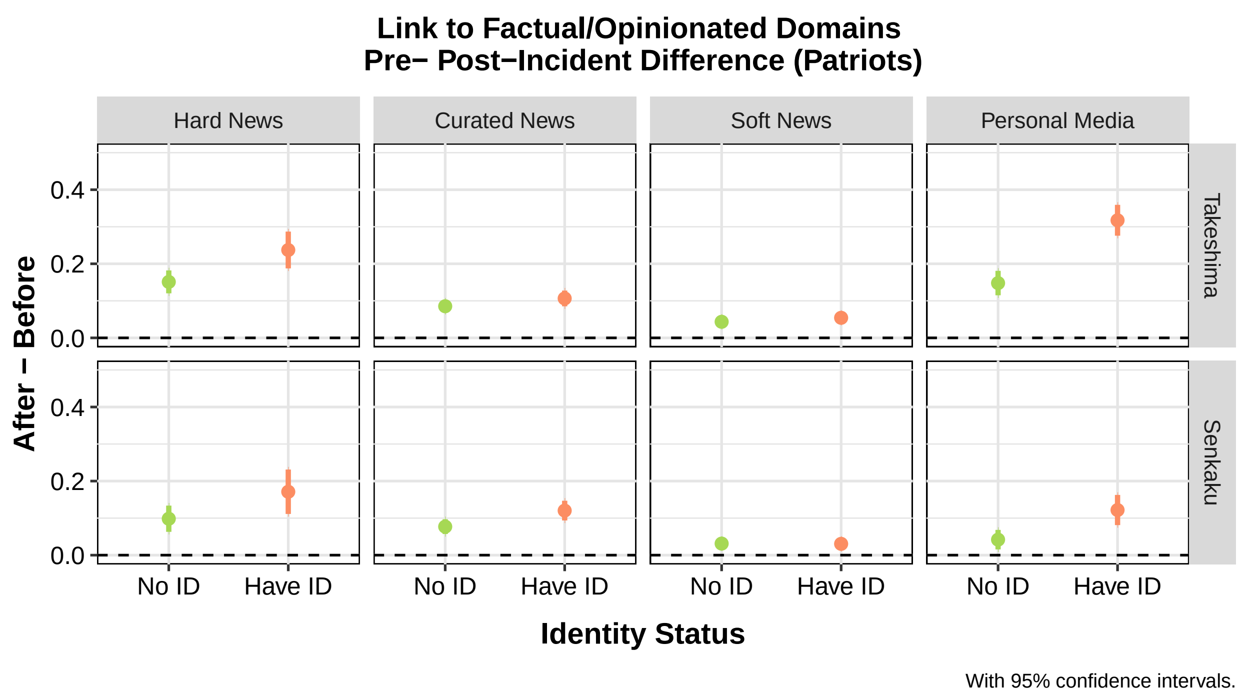

Figure B.3: Mean Differences of the Result in Figure 12 (Means are Calculated from the Entire Period of Before/After the Incident)

Table B.4: Pre- and Post-Incident Mean Comparisons of the Result in Figure 13 (Means are Calculated from the Entire Period of Before/After the Incident)

Domain Cat.	Issue	ID Stat.	Mean (Before)	SE (Before)	Mean (After)	SE (After)	Mean (After-Before)	SE (After-Before)
Hard News	Takeshima	No ID	0.064	0.007	0.217	0.014	0.153	0.016
Hard News	Takeshima	Have ID	0.089	0.009	0.353	0.027	0.264	0.028
Curated News	Takeshima	No ID	0.044	0.004	0.130	0.007	0.086	0.008
Curated News	Takeshima	Have ID	0.047	0.005	0.157	0.011	0.110	0.012
Soft News	Takeshima	No ID	0.035	0.005	0.081	0.005	0.046	0.007
Soft News	Takeshima	Have ID	0.045	0.004	0.093	0.007	0.047	0.008
Personal Media	Takeshima	No ID	0.281	0.010	0.441	0.013	0.161	0.016
Personal Media	Takeshima	Have ID	0.512	0.014	0.869	0.024	0.357	0.028
Hard News	Senkaku	No ID	0.269	0.012	0.369	0.014	0.100	0.018
Hard News	Senkaku	Have ID	0.353	0.021	0.555	0.028	0.202	0.035
Curated News	Senkaku	No ID	0.107	0.007	0.187	0.008	0.080	0.010
Curated News	Senkaku	Have ID	0.100	0.007	0.225	0.012	0.124	0.014
Soft News	Senkaku	No ID	0.067	0.004	0.097	0.005	0.031	0.006
Soft News	Senkaku	Have ID	0.073	0.005	0.109	0.006	0.036	0.008
Personal Media	Senkaku	No ID	0.263	0.007	0.310	0.011	0.047	0.013
Personal Media	Senkaku	Have ID	0.424	0.010	0.577	0.022	0.153	0.024

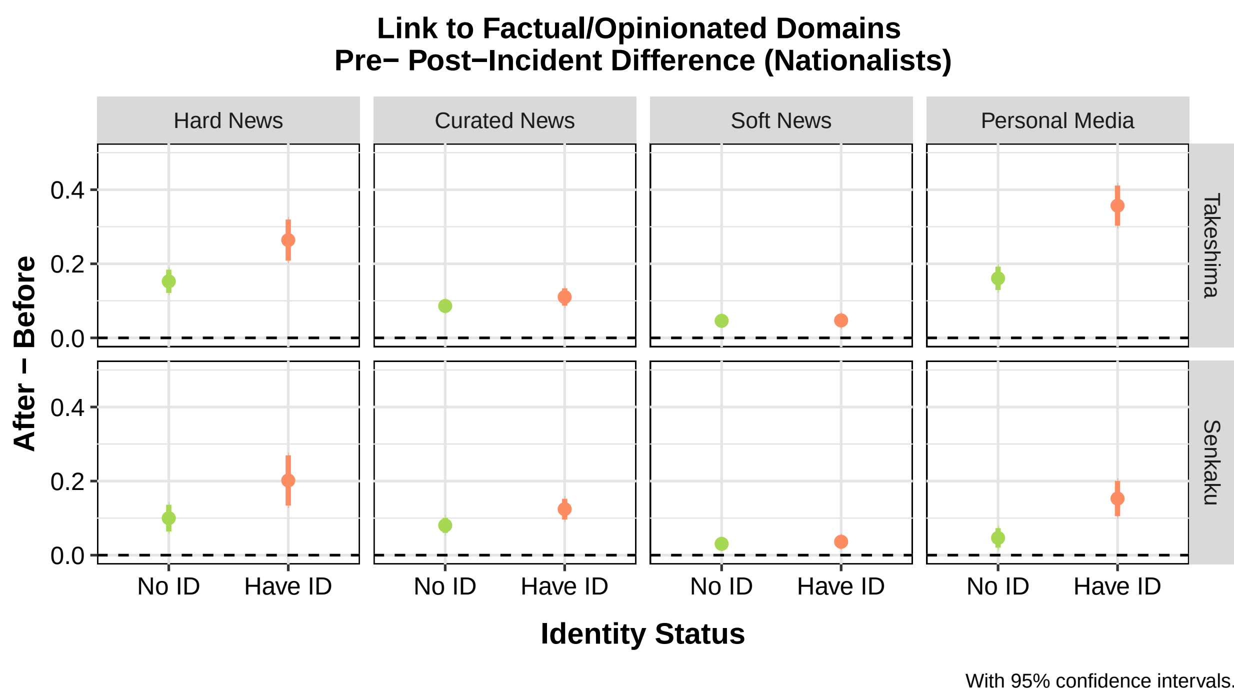

Figure B.4: Mean Differences of the Result in Figure 13 (Means are Calculated from the Entire Period of Before/After the Incident)

C Difference in Differences in Event Impact on Retweets

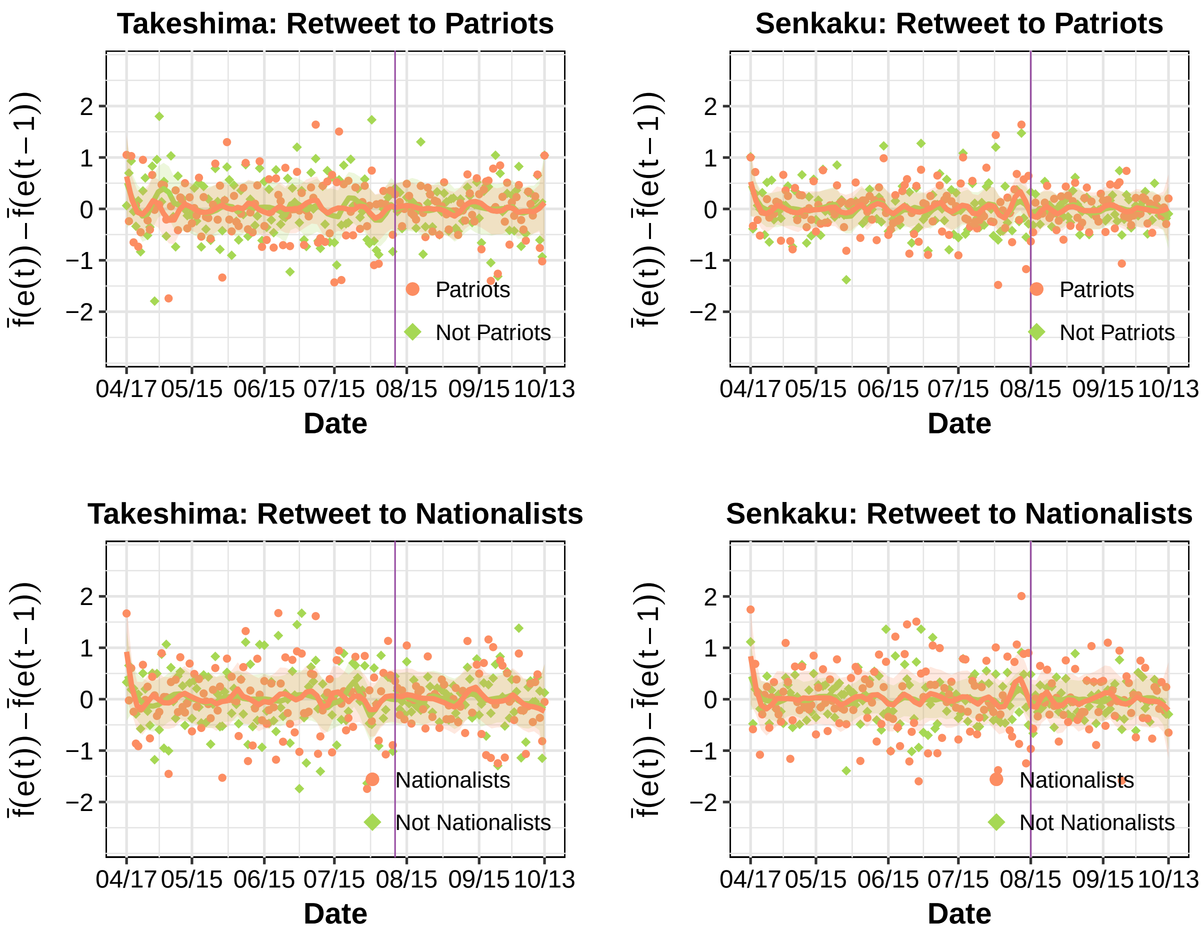

Figure C.1: Difference in Differences in Event Impact on Retweets (k=1)

In addition to the analysis in Figure 11, from Equation 6, we can represent difference in differences in the event impact to the users in t and $t-k$.

$$\bar{f}\left(e\left(t\right)\right)-\bar{f}\left(e\left(t-k\right)\right)=E\left[\bar{r}\left(t\right)\right]/R(t,\tau)-E\left[\bar{r}\left(t-k\right)\right]/R(t-k,\tau) \tag{8}$$

This way, we can eliminate $\bar{w}$ from the equation and focus on the pure event impact. Figure C.1 and Figure C.2 present this result with $k=1$ and $k=7$. While $k=1$ difference do not show the clear difference by identity status, $k=7$ do show the similar tendency as in Figure 11: Following territorial issue incidents in Senkaku, identity holders strengthen the tendency to retweet identity holders more than no identity holders. This implies that the event impacts on identity holders are larger than no identity holder, because the impacts on identity holders are more persistent. There is a tendency that as k becomes larger, difference in differences in event impact becomes larger.

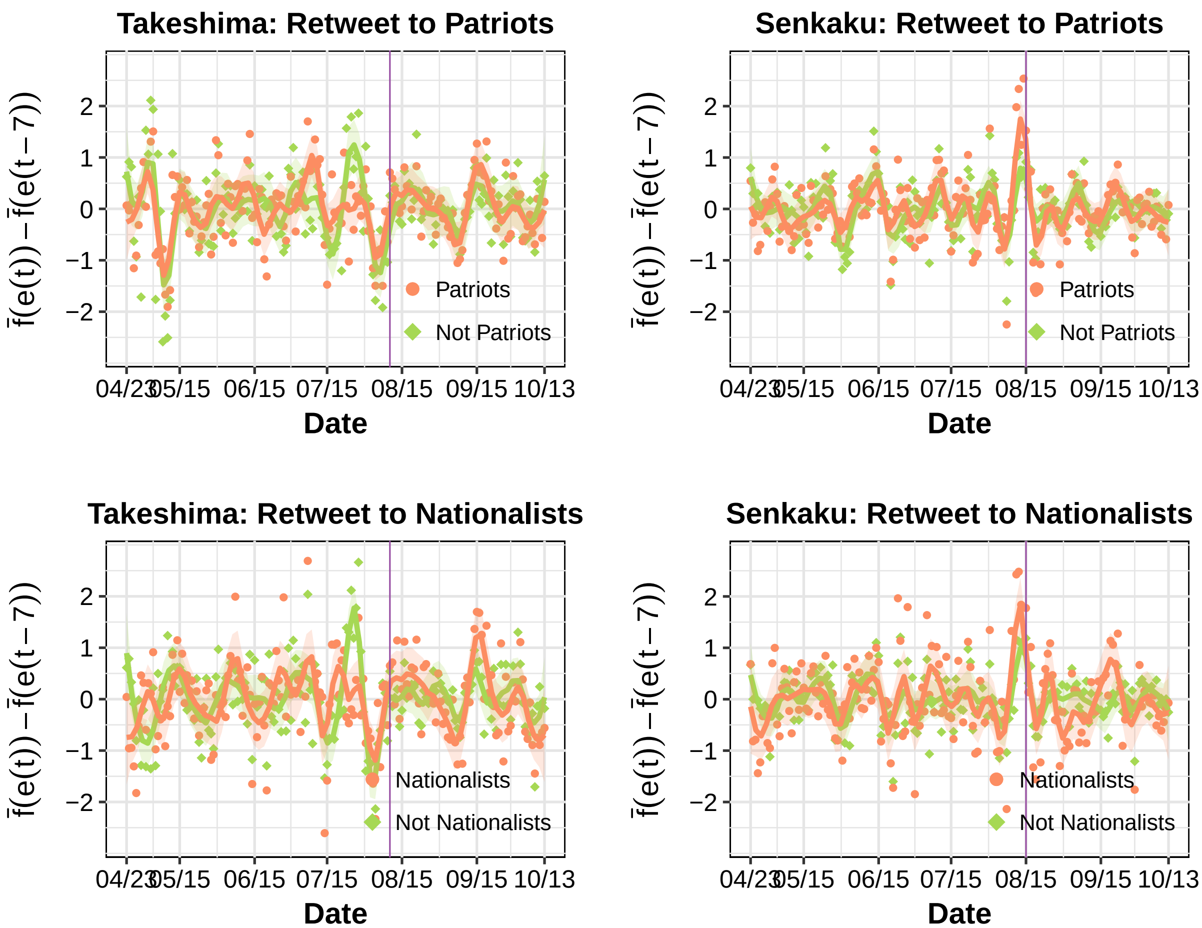

Figure C.2: Difference in Differences in Event Impact on Retweets (k=7)

D Time-Series Distribution of Out-Degree Relevant Characteristics

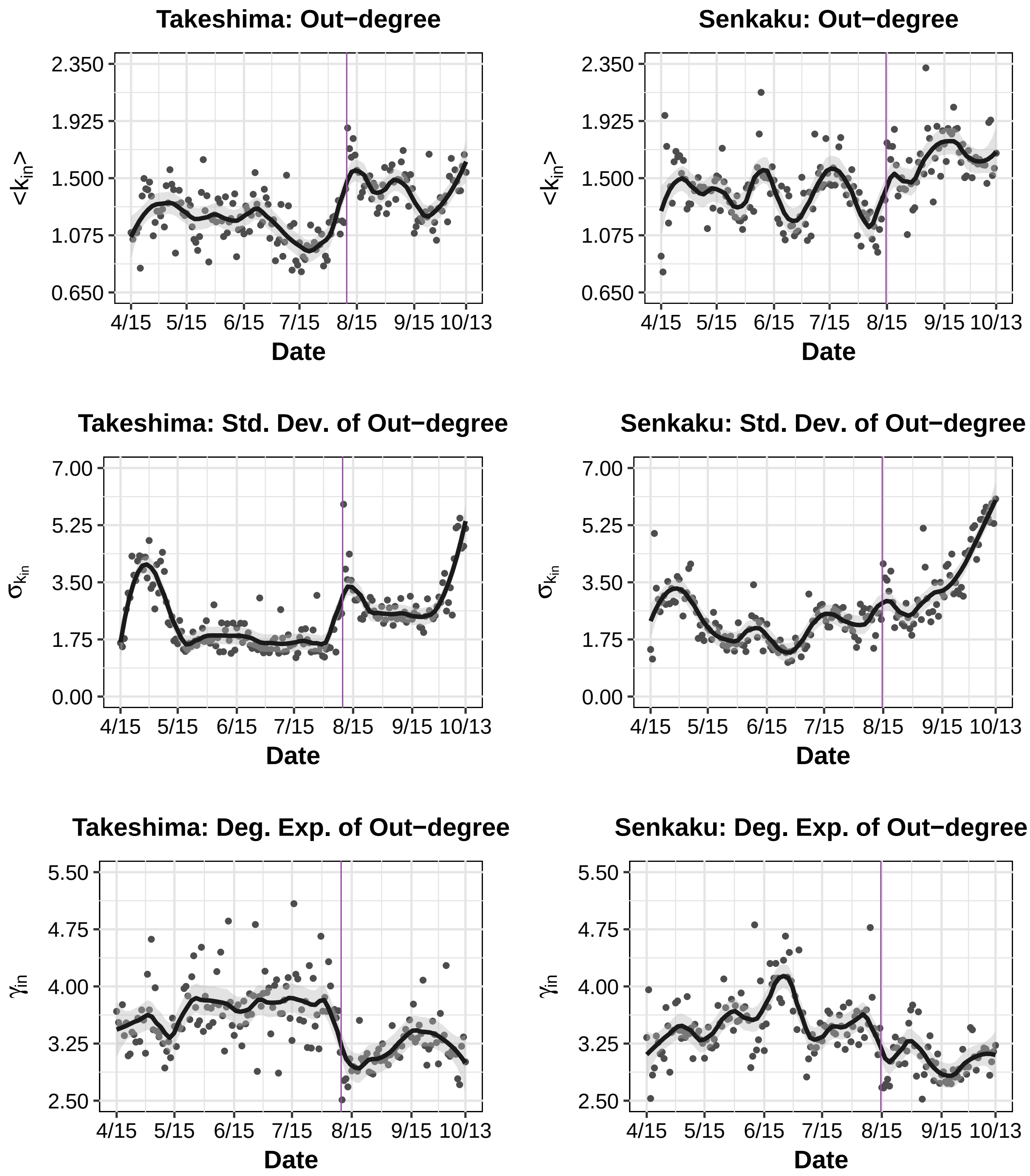

Figure D.1: Time Series Distribution of Out-Degrees, Standard Deviation of Out-Degrees, and Degree Exponent of Out-Degrees

E Mean Comparisons of Pre- and Post-Incident In-degrees

Table E.1: Pre- and Post-Incident Mean Comparisons of the Result in Figure 16 (Means are Calculated from the Entire Period of Before/After the Incident)

Unit	ID Type	Issue	ID Stat.	Mean (Before)	SE (Before)	Mean (After)	SE (After)	Mean (After-Before)	SE (After-Before)
Average	Nationalism	Takeshima	No ID	0.588	0.016	0.609	0.016	0.021	0.023
Average	Nationalism	Takeshima	Have ID	1.287	0.035	2.423	0.164	1.136	0.168
Average	Patriotism	Takeshima	No ID	0.479	0.016	0.562	0.015	0.083	0.022
Average	Patriotism	Takeshima	Have ID	1.477	0.075	2.119	0.119	0.642	0.141
Average	Nationalism	Senkaku	No ID	0.551	0.011	0.638	0.016	0.088	0.019
Average	Nationalism	Senkaku	Have ID	1.416	0.051	1.976	0.090	0.560	0.104
Average	Patriotism	Senkaku	No ID	0.489	0.011	0.588	0.015	0.099	0.018
Average	Patriotism	Senkaku	Have ID	1.336	0.042	1.764	0.075	0.428	0.086

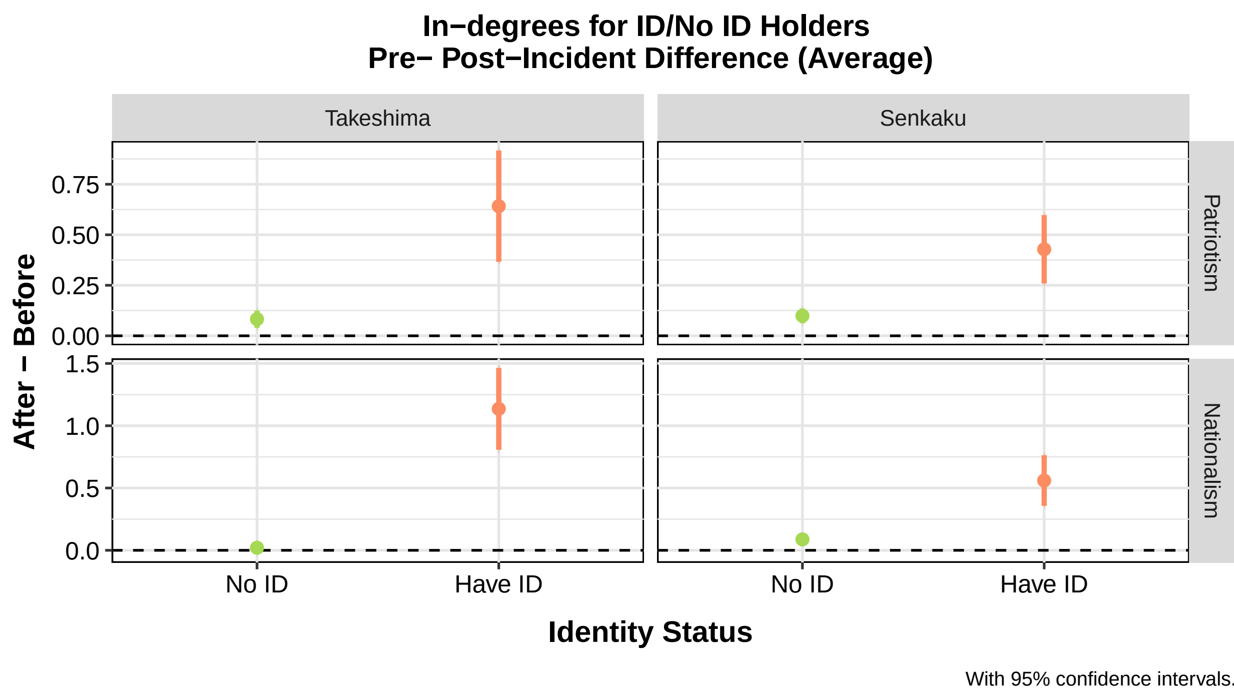

Figure E.1: Mean Differences of the Result in Figure 16 (Means are Calculated from the Entire Period of Before/After the Incident)

Table E.2: Pre- and Post-Incident Mean Comparisons of the Result in Figure 17 (Means are Calculated from the Entire Period of Before/After the Incident)

Unit	ID Type	Issue	ID Stat.	Mean (Before)	SE (Before)	Mean (After)	SE (After)	Mean (After-Before)	SE (After-Before)
Average	Nationalism	Takeshima	No ID	0.588	0.016	0.609	0.016	0.021	0.023
Average	Nationalism	Takeshima	Have ID	1.287	0.035	2.423	0.164	1.136	0.168
Average	Patriotism	Takeshima	No ID	0.479	0.016	0.562	0.015	0.083	0.022
Average	Patriotism	Takeshima	Have ID	1.477	0.075	2.119	0.119	0.642	0.141
Average	Nationalism	Senkaku	No ID	0.551	0.011	0.638	0.016	0.088	0.019
Average	Nationalism	Senkaku	Have ID	1.416	0.051	1.976	0.090	0.560	0.104
Average	Patriotism	Senkaku	No ID	0.489	0.011	0.588	0.015	0.099	0.018
Average	Patriotism	Senkaku	Have ID	1.336	0.042	1.764	0.075	0.428	0.086

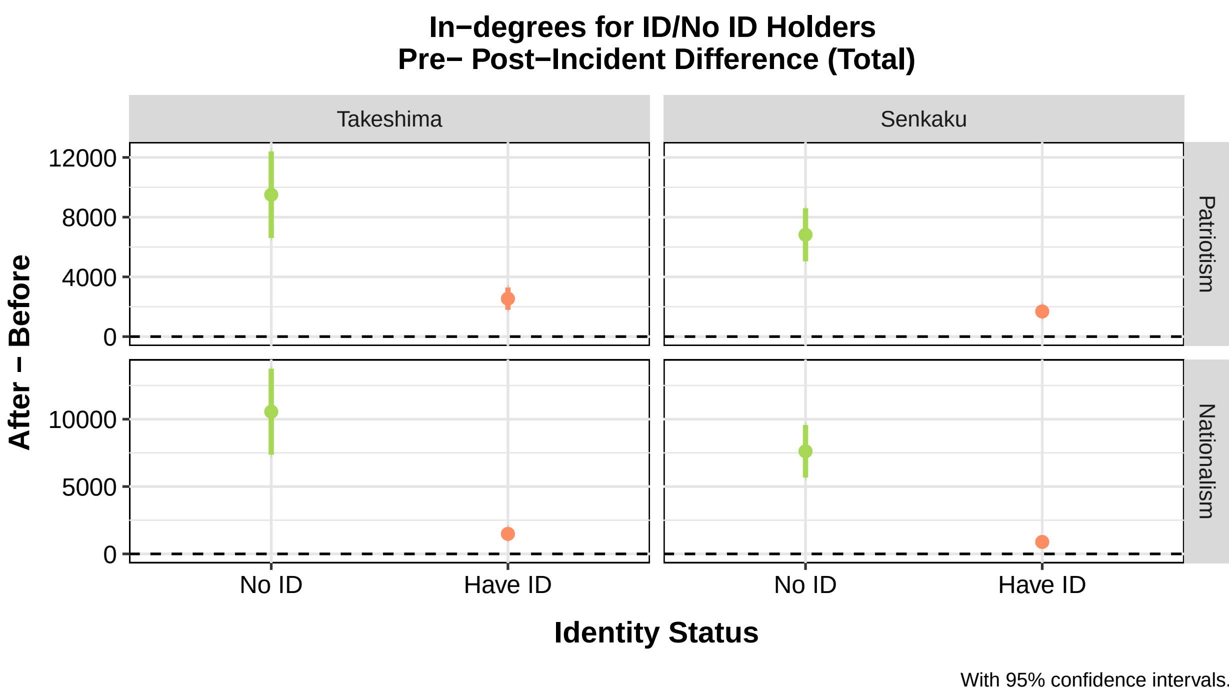

Figure E.2: Mean Differences of the Result in Figure 17 (Means are Calculated from the Entire Period of Before/After the Incident)

Table E.3: Pre- and Post-Incident Mean Comparisons of the Result in Figure 18 (Means are Calculated from the Entire Period of Before/After the Incident)

Unit	Domain Cat.	Issue	Mean	SE	Mean	SE	Mean	SE
			(Before)		(After)		(After-Before)	
Average	Hard News	Takeshima	13.424	1.824	87.520	13.005	74.095	13.132
Average	Personal Media	Takeshima	7.709	0.266	33.299	2.780	25.590	2.793
Total	Hard News	Takeshima	130.359	32.431	3186.677	558.338	3056.318	559.279
Total	Personal Media	Takeshima	421.128	29.277	6073.723	625.511	5652.595	626.196
Average	Hard News	Senkaku	45.530	3.446	119.830	9.371	74.300	9.984
Average	Personal Media	Senkaku	10.974	0.552	26.772	1.748	15.798	1.833
Total	Hard News	Senkaku	1057.475	112.380	5075.433	441.457	4017.958	455.536
Total	Personal Media	Senkaku	912.926	81.444	4547.800	404.359	3634.874	412.480

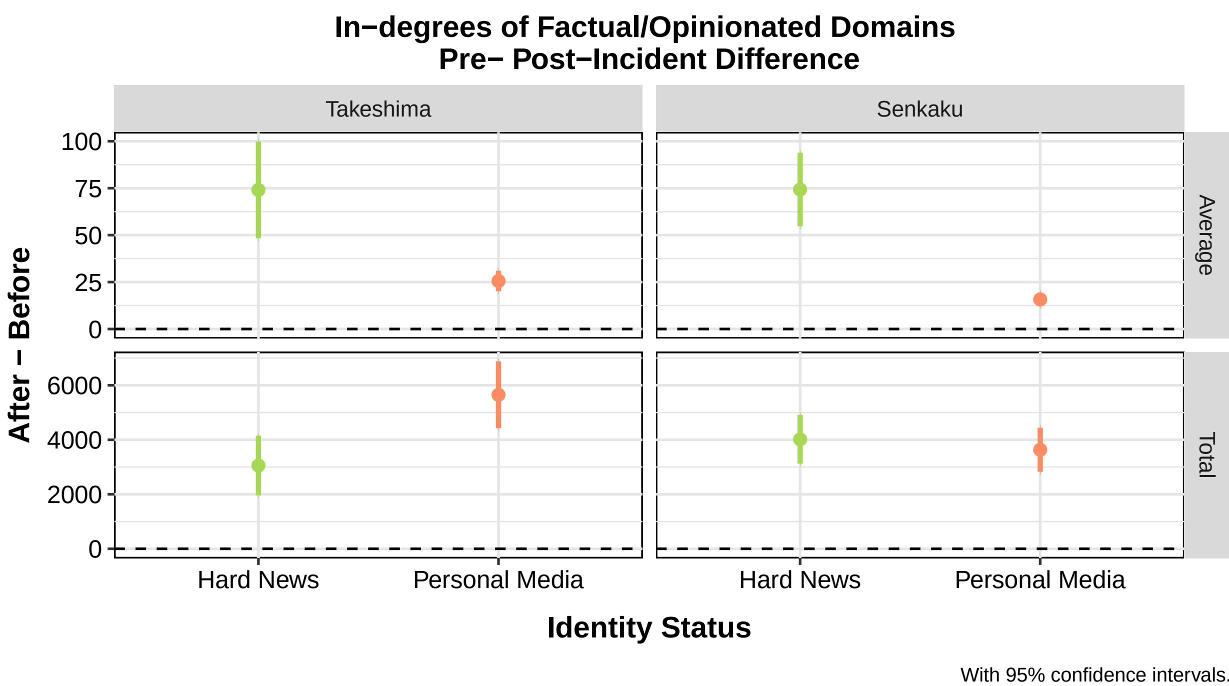

Figure E.3: Mean Differences of the Result in Figure 18 (Means are Calculated from the Entire Period of Before/After the Incident)