Media Trust and Persuasion*

Shuhei Kitamura† Toshifumi Kuroda‡

July 8, 2021

Abstract

In this paper, we examine how media trust and persuasion are interlinked. We conducted a large-scale randomized field experiment in Japan in collaboration with the nation's most trusted media outlet in which the capacity for viewing its programs was increased. We find that our treatment increased the viewing time of its programs as well as support for government policies due to exposure to slanted information in favor of the government. We also find that the effect was larger for those who trusted the outlet more than other outlets and that their trust levels on the outlet were even increased by the treatment.

JEL code: K23, L82, D78

*This paper was previously titled as "Public Media Do Serve the State: A Field Experiment". We thank Tomomichi Amano, Amy Catalinac, Claudio Ferraz, Matthew Gentzkow, Masayuki Kudamatsu, Vardges Levonyan, John List, Emi Nakamura, Torsten Persson, David Strömberg, Kensuke Teshima, Yasutora Watanabe, and Teppei Yamamoto for their helpful comments and suggestions. We are also grateful to the seminar participants at GRIPS, Keio, Osaka, Osaka City, and Rochester and at the ES Virtual World Congress, the EEA-ESEM Annual Meetings at the University of Manchester, the JEA Spring Meeting at Musashi University, the Kansai Labor Economics Workshop, the Kwansei Gakuin IO Workshop, the Kyoto Summer Workshop on Applied Economics, the SIOE Conference at the Stockholm School of Economics, and the Workshop of Economic Analysis on Public Policy at National Chengchi University. All errors are our own.

†Associate Professor (Specially Appointed). CiDER, Osaka University, D74-1 Office for Industry-University Co-Creation (Bldg. D), 2-1 Yamadaoka, Suita, Osaka, 565-0871 Japan. Email: kitamura@cider.osaka-u.ac.jp. Phone: +81 6-6879-8543.

‡Associate Professor. Tokyo Keizai University. Email: kuroda@tku.ac.jp.

1 Introduction

Media trust and persuasion are interlinked. Individuals trust media that they prefer, while they obtain information from media outlets that they trust. In the United States, Republicans tend to trust right-leaning media outlets, while Democrats tend to trust left-leaning media outlets (Pennycook and Rand, 2019). Moreover, public media are most trusted media outlets in many countries (Newman et al., 2018).

Unpacking the interlinkage to further understand the influence of media on individuals is challenging but is important because slanted media outlets are known to have persuasive effect on voters (DellaVigna and Kaplan, 2007; Gerber et al., 2009; Martin and Yurukoglu, 2017; Durante et al., 2019) and political polarization has become a big social problem in countries such as the United States (Gentzkow, 2016).

In this paper, we study how media trust and persuasion are interlinked. We conducted a large-scale randomized experiment in Japan in collaboration with the nation's most trusted media outlet, or the national public TV company (henceforth Public TV).[1] In the experiment, which was conducted between November and December, 2016, we recruited 6,000 individuals and randomly increased the treatment group's capacity to view Public TV's programs by providing them with free subscriptions. Our main outcomes of interest are individuals' support for government policies and their subjective evaluations on the capability of the government.

We find that our treatment increased the subjects' viewing time of Public TV's programs as well as their support for certain policies, although we do not find that it changed their subjective evaluations on the capability of the government. Moreover, we find that the effect was larger for those who trusted Public TV more than private broadcasters, and that, more surprisingly, their trust levels were even increased by the treatment. We call this phenomenon as *endogenous persuasion*.[2] Individuals obtain information from a media outlet that they most trust and their trust levels on the outlet increase by doing so. Thus, trust levels are also endogenous and susceptibility to the persuasion depends on their trust levels on particular media outlets.

To further understand the information that the subjects were exposed during the experiment, we use a novel method to measure media slant. Specifically, we apply an unsupervised machine learning technique (Le and Mikolov, 2014) to measure the semantic similarity between the content

[1] The experiment is registered (AEARCTR-0005335).

[2] There is a related term called *self-persuasion* in Social Psychology. Compared to direct persuasion, *self-persuasion* is "indirect and entails placing people in situations where they are motivated to persuade themselves to change their own attitudes or behavior" (Aronson, 1999). Our is not necessarily self-persuasion because individuals are directly persuaded by media outlets but susceptibility to the persuasion depends on their trust levels on the outlets.

of TV programs and that of official statements made by domestic and foreign governments.[3] Using this similarity measure, we find that Public TV's information is closer to the domestic government's statements than to those of the foreign government, compared with the information provided by private media. Taken together, these results indicate that individuals especially those who trusted Public TV more than private broadcasters were more likely to be exposed to its slanted information in favor of the government during the experiment, and exposure to such information increased their trust levels on this outlet.

This study contributes to a large body of literature on media and information in social sciences, especially the recent literature on the role of trust in media consumption. In Economics, Gentzkow et al. (2020) theoretically shows that arbitrary small biases in processing noisy feedback can cause large ideological biases in trust in information sources as well as beliefs about facts. Our findings on endogenous persuasion is consistent with these results, although ours is an empirical study. In addition, in Psychology, Pennycook and Rand (2019) find that laypeople across the political spectrum rated mainstream sources more trustworthy than either hyperpartisan or fake news sources. Although they conclude that having algorithms up-rank content from trusted media outlets (i.e., mainstream sources) in social media platforms might reduce the spread of misinformation on social media, our results warn that up-ranking most trusted media outlets can exacerbate political polarization.[4]

This paper also relates to the theoretical literature documenting the role of independent media outlets. The existence of such media outlets should reduce the degree of media slant in the market by enabling consumers to cross-check information (Mullainathan and Shleifer, 2005; Gentzkow and Shapiro, 2006).[5] If public media's information is less slanted than that of private media for some reasons, they would play a crucial role in mitigating media slant. However, although Public TV is the sole public TV company in Japan, our study shows that its information is far from neutral, which raises a concern about the independence of the public media outlet.

[3] The technique has several advantages including no requirement for selecting words *ex ante*. Although previous studies often use predetermined word categories created by researchers (Tetlock, 2007; Qin et al., 2017, 2018), this step can be skipped in our method because the machine automatically computes the semantic similarity using a plain corpus, or a list of plain texts such as Google Books Corpora. In addition, compared with similar methods employed by previous studies, such as counting the frequency of words in texts (bag-of-words, bag-of-n-grams), our algorithm accounts for the ordering and semantics of words, as well as the contextual differences across documents show the performance of different supervised machine learning algorithms for measuring polarization (Ash and Labzina, 2019; Peterson and Spirling, 2018). Regarding the last point, roughly speaking, our algorithm employs a "fixed-effect" procedure in which the estimated parameter captures the contextual specificity of a document (or a TV program in our context) when maximizing the probability of observing a target word.

[4] For example, Republicans may still prefer Fox News, while Democrats may still prefer MSNBC. Due to endogenous persuasion, Republicans will trust Fox News more, and Democrats will trust MSNBC more.

[5] Strömberg (2015) also shows that competition leads to the under-provision of information compared with the social optimum, which public media or a Pigouvian subsidy can alleviate.

Third, this study adds a new evidence to the literature on media persuasion. Some studies find that a slanted TV channel has an impact on the vote share (DellaVigna and Kaplan, 2007; Martin and Yurukoglu, 2017; Durante et al., 2019) and political knowledge (Schroeder and Stone, 2015). Another study finds that slanted newspapers have no influence on the vote share, political knowledge, or political opinions, although reading such newspapers does increase support for Democratic candidates and raise voter turnout (Gerber et al., 2009). In our study, we show that public media in a democratic setting can also have a persuasive effect on their viewers.

In addition, it contributes to the literature on the sources of media slant. Media slant can be driven by either supply-side factors (e.g., owners, advertisers, and journalists) or demand-side factors (e.g., consumers).[6] Regarding the supply-side factors, media controlled by a state, party, or politician are slanted in China (Qin et al., 2018; Chen and Yang, 2019), Berlusconi's Italy (Durante and Knight, 2012), and Russia (Enikolopov et al., 2011). By contrast, Gentzkow and Shapiro (2010) also show that media slant in US newspapers is driven by consumer preferences rather than owner preferences.[7] These mixed findings in different countries suggest that the factors affecting media slant, either supply-side or demand-side factors, may depend on the degree of media independence and competition in the media market. Our study shows that public media outlets can be slanted as well, and we argue that supply-side factors are the main sources of the slant.

The literature on the impact of new information technology on political outcomes is also related to this work because our experiment involves online streaming. Previous studies show that voter turnout has fallen because of the diffusion of TV (Gentzkow, 2006) and the Internet (Falck et al., 2014; Campante et al., 2018; Gavazza et al., 2019). Using data from the United States, Prior (2005) finds that new information technology widened the knowledge gap and turnout between individuals who prefer news and those who prefer entertainment. By contrast, Boxell et al. (2017) find that a greater use of the Internet is not associated with political polarization in the United States.

Finally, our study contributes to the literature on the analysis of political language. Previous studies use "word scores" (Laver et al., 2003), think-tank citations (Groseclose and Milyo, 2005), and party phrases (Gentzkow and Shapiro, 2006; Jensen et al., 2012; Martin and Yurukoglu, 2017; Gentzkow et al., 2019). As mentioned above, our method has several advantages, including no requirement to select words of interest *ex ante*.

The rest of the paper is organized as follows. Sections 2 and 3 describe our experiment and

[6] Besley and Prat (2006) show that an incumbent is more likely to "buy" unfavorable information from a media outlet to avoid disclosing information to voters if it is less costly than the rents from holding office.

[7] Gentzkow et al. (2015) also find little evidence that the ruling party influences the partisan composition of the press.

data. Section 4 reports the results of the experiment. Section 5 uses machine learning to examine the mechanism behind these findings. Section 6 discusses the results and Section 7 concludes.

2 Experiment

An online randomized controlled trial was conducted in the greater Tokyo area in November and December 2016.[8] We recruited 6,000 Internet users using a mail survey.[9] The entire procedure was the following. First, we created 40 blocks according to age (20s, 30s, 40s, 50s, and 60s), sex (male and female), and TV viewing frequency (almost every day, at least once a week, at least once a month, and hardly ever or never). Then, we recruited individuals such that the percentage of each block in the sample was proportional to the population in each category in the greater Tokyo area.[10]

Recruitment was conducted by one of the largest Internet survey companies in Japan in terms of the number of registered individuals. The company sent a recruiting email to its registered individuals to explain the outline of the surveys and rewards.[11] The email clearly stated that the purpose was to survey the individual's daily TV viewing behavior including private programs, daily media access, and media evaluations, particularly TV. The email also stated that participants would keep a viewing record every day. To minimize selection bias, the respondents were not informed of the availability of free Internet streaming services at the time of recruitment. Participants were told about this benefit as a surprise just before the start of the experiment. Finally, we assigned 5,000 of the 6,000 individuals randomly to the treatment group and assigned the remainder to the control group.[12]

As the main treatment, we randomly provided the treatment group with free subscriptions to view Public TV's programs online.[13] In particular, treated individuals could view Public TV's

[8] Although some private broadcasters provided either free or paid Internet streaming services, Public TV did not provide such a service until a revision of the Broadcast Act in 2014, which allowed it to provide an Internet streaming service if it did not harm market competition. To test whether the streaming service met this criterion, Public TV conducted large-scale randomized experiments in 2015 and 2016; this study focuses on the 2016 experiment. Regarding the 2015 experiment, Kuroda et al. (2017) find no evidence that the streaming service harmed market competition. The 2016 experiment was significantly larger than the 2015 experiment.

[9] We determined the number to maximize the statistical power given the budget restrictions and government regulation.

[10] More specifically, we derived the numbers from Public TV's public opinion poll conducted in the greater Tokyo area in July 2016.

[11] The company did not disclose the details of the rewards to us, but they were based on its incentive criterion, according to the company's reply. Most importantly, there was a bonus for completing a viewing record (see the Data section).

[12] We used the `randomize` command in Stata 14. We allocated more than half of the individuals to the treatment group because of the second experiment, as described and analyzed in the Online Appendix.

[13] In the second experiment, we also sent program recommendations to randomly selected subgroups in the treatment

programs on web browsers and mobile phones. The services comprised live streaming and video on demand (VOD). The treatment thus increased the capacity for viewing Public TV's programs.

Although Public TV did not provide any live streaming service before the experiment began, it did offer a paid VOD service.[14] Some programs were not available during the experiment because of copyright issues.[15] In addition, because of government regulation, only programs that started at 7:00 am or after and ended by 11:00 pm were available for both services. By contrast, the VOD service was accessible at any time.

Figure 1 shows the timeline of the experiment. We started recruitment at the end of October 2016. The baseline survey was conducted at the same time. We then randomized the subjects. All participants started to keep a viewing record from November 14. The free online services became available for the treatment group on November 28. Finally, the endline survey was conducted after the end of the experiment on December 18.

3 Data

Our main data are each subject's viewing record.[16] Participants received an email from the survey company every day in which they were asked to mark the TV programs (Public TV's and private broadcasters' programs) that they viewed on that day. Although the data were self-reported, we alleviated potential measurement errors in several ways. First, to reduce the reporting cost, participants simply check-marked the names of the TV programs they viewed on the computer screen. Second, participants were incentivized for providing reports. That is, the bonus was given to individuals upon completion of the entire survey.[17] Third, we dropped unreliable samples as a robustness check. Finally, we employed the instrumental variable (IV) method.

In the baseline survey, we asked individuals' socioeconomic status, media usage, media evaluations, and political orientation. In the endline survey, we asked for their evaluations of policies and the government's capability in addition to media usage and media evaluations.

In addition, we collated unique data: TV scripts from all TV channels in the Tokyo region,

group on Fridays. The Online Appendix provides the analyses of this second treatment.

[14] The monthly subscription fee for the VOD service was 972 JPY (about 10 USD) or a minimum of 100 JPY (about 1 USD) per program.

[15] About 16% of total broadcasting time was not provided in these services.

[16] For treated individuals, the access log of Public TV's Internet streaming service was also available. We could not use the access log for the main analysis because there were no such data for the control group. See the Online Appendix for our analysis using these data for the subset of individuals.

[17] The survey company did not tell participants the exact amount or formula to avoid strategic behavior. Participants were also given three days to complete their reports.

including private broadcasters. We collected these data using a TV digital recorder.[18]

Since 1,524 individuals (1,280 in the treatment group and 244 in the control group) did not complete the endline survey, 23 people decided not to participate in the experiment, and an individual's viewing records were deemed to be unreliable,[19] our final sample became 4,452 (74.2%).[20]

We conducted balance checks between the treatment and control groups in the final sample using variables from the baseline survey (see the results in the Online Appendix). Most variables are balanced between the two groups, suggesting that the randomization was successful. Since some variables were not balanced,[21] we included them as the baseline control in the following analysis. Table 1 shows the summary statistics of the main variables.

4 Results

4.1 Effect on Total Viewing Time

The first analysis examines whether the treatment increases the viewing time of Public TV's programs, in particular Channel 1 (henceforth Public 1).[22] In particular, we run the following regression using ordinary least squares (OLS):

$$\Delta\text{Public_1}_i = \alpha + \beta\text{Treatment}_i + \mathbf{X_i}\gamma + \varepsilon_i, \tag{1}$$

for individual i, where $\Delta\text{Public_1}_i$ is the difference in the viewing time (in hours) of Public 1 before (November 14-27) and during (November 28-December 18) the experiment, Treatment_i is the treatment dummy, $\mathbf{X_i}$ is a vector of the control variables, and ε_i is the error term, which we assume is independently distributed across individuals.[23] Our main interest is the estimate of β. Using the difference in viewing time, we remove any systematic differences in the usual viewing pattern across individuals that we cannot perfectly control.

Figure 2 shows the average of the dependent variable for the treatment and control groups. It shows that the treatment group watched Public 1 approximately one hour longer than the control group during the experiment.

[18] The machine provided an API to extract program information including the titles, descriptions, and scripts. We used Python to obtain information from the recorder and saved them in CSV format.

[19] The individual reported that he/she viewed most TV programs in a day.

[20] The Online Appendix also provides an analysis for checking attrition bias.

[21] These variables were a dummy for supporting the Liberal Democratic Party, a minor party, and no party, a dummy for not reporting annual income and for reading the Sankei newspaper, and trust and accuracy measures.

[22] Public TV has two channels, 1 and 2. See below for further explanation about the second channel.

[23] Recall that there is a pre-treatment period in the viewing record from which we obtain information on pre-treatment viewing behavior.

Table 2 shows the average treatment effect on the viewing time of Public 1. The first column does not include the baseline control variables, while the next two columns do include them. The third column drops individuals who occasionally failed to report their viewing records for certain programs, although they were likely to view them, according to the access log. Overall, the treatment increased the viewing time of Public 1 by about one hour, corresponding to Figure 2. In the following analysis, we use the smallest sample to minimize measurement errors.

4.2 Effect by Week and on Other TV Channels

Table 3 shows the effect by week.[24] The effect appears in all three weeks of the experiment. Moreover, the size of the effects varies little across these weeks, although the effect in Week 2 is slightly larger than those in the other weeks. This finding shows that the treatment effect does not simply capture the effect of special daily programs.

Next, to understand the overall impact of our treatment, we investigate the effect on other TV channels. We are particularly interested in knowing whether the increase in the viewing time of Public 1 is explained by a decrease in the viewing time of another channel (Channels 2, 4, 6, 8, 10, and 12), holding total viewing time constant, or an increase in total viewing time.

Columns (1)–(7) in the top panel of Table 4 show the average treatment effects by channel. The dependent variables use the difference in viewing time as in equation (1), but for each channel. Column (1) replicates column (3) in Table 2 for comparison purposes. As shown in column (2), the treatment did not increase the viewing time of Channel 2 (henceforth Public 2), another channel broadcast by Public TV. Public 2 is slightly different from the other channels, as it focuses on educational programs. We also find that it was the least popular channel in our data.[25] These factors might explain why the treatment did not increase the viewing time of Public 2.

Columns (3)–(7) show the treatment effects for all the private broadcasters (which we call Private 4, 6, 8, 10, 12, respectively, hereafter). We find that the treatment effects for Privates 4 and 12 are negative, although the latter effect is relatively weak. By contrast, the last column of the top panel in Table 4 shows that the treatment did not affect the total viewing time of TV programs. Taken together, these results indicate that the increase in the viewing time of Public 1 is to some extent explained by the change in the viewing time of private channels, holding total viewing time constant. Using the difference in shares instead of the difference in hours for the

[24] Since there is no comparable weekly viewing time before the experiment, viewing time before the experiment is now included as an additional control variable in the regressions.

[25] The average viewing time of Public 2 during the experiment was only about one hour, while that of Public 1 was about 11 hours, as shown in Table 1.

dependent variables, the bottom panel of Table 4 shows a consistent picture; the treatment also increased the share of Public 1 and decreased the share of Private 4. The effect on Private 12 now comes out as insignificant.

4.3 Effect on Political Attitudes

Next, we investigate the effect of watching Public 1 on political attitudes. Two major political events occurred during the experiment and both of them were related to foreign policy. The first was President Vladimir Putin's visit to Japan on December 15 and 16 to attend a summit. One of the main topics of the meeting was the Northern Territories dispute, a territorial dispute between Japan and Russia over four islands (Etorofu, Kunashiri, Shikotan, and Habomai). The main interest of the Japanese public at the time was whether there would be any progress on that issue during the meeting (*The Economist*, 2016; Chugrov and Streltsov, 2017). The second major event was that then-Prime Minister Shinzō Abe announced that he would visit Pearl Harbor with then-President Barack Obama. Although the event actually occurred after the experiment (December 26–28), the first public announcement was made on December 5 (i.e., during the experiment).

Figure 3 shows the popularity of each topic as measured by the frequency of Google searches in Japan according to Google Trends. It shows that "真珠湾" (Pearl Harbor) became more popular following the first public announcement by then-Prime Minister Abe and that "北方領土" (Northern Territories) became more popular during the summit. For comparison purposes, the search results for "交通事故" (traffic accident) were also presented, which showed no systematic pattern.

We included survey questions on these two major political events in the endline survey: How do you evaluate the government's measures against the Northern Territories dispute and how do you evaluate Prime Minister Abe's visit to Pearl Harbor with President Obama. Individuals rated each question on a scale of 0 (the lowest rating) to 10 (the highest rating). In addition, we asked individuals to evaluate the capability of the government to implement policies, with four possible answers from agree to disagree.

To check the effect of watching Public 1 on these outcomes, we run the following OLS regression:

$$\text{Evaluation}_i = \phi + \psi \Delta\text{Public_1}_i + \mathbf{X_i}\xi + v_i, \tag{2}$$

for individual i, where Evaluation_i is the political evaluation just described, $\Delta\text{Public_1}_i$ is the difference in the viewing time of Public 1 as in equation (1), $\mathbf{X_i}$ is a vector of the baseline control

variables, and ν_i is the error term, which we assume is independently distributed across individuals. We are interested in the estimate of $\boldsymbol{\psi}$.

Simply regressing model (2), the estimated $\boldsymbol{\psi}$ may be biased because $\Delta\text{Public_1}_i$ is likely to be endogenous. This bias can work in either direction. On the one hand, if individuals who view the channel longer tend to positively evaluate the government, the effect is likely to be overestimated. On the other hand, if individuals who view the channel longer tend to be skeptical about the government, the effect is likely to be underestimated. To isolate the causal effect, we thus use the treatment variable as the instrument for $\Delta\text{Public_1}_i$ in the IV framework.

Columns (1)–(3) in Table 5 show the results. The top panel shows the OLS estimates of equation (2). The results are not statistically significant for all the dependent variables. Next, the middle panel reports the IV (two-stage least squares, 2SLS) estimates in which we use the treatment variable as the instrument. Although watching Public 1 longer affects neither the evaluation of the government's capability nor the evaluation of Prime Minister Abe's visit to Pearl Harbor (columns (1) and (2), respectively), it has a statistically significant positive effect on the evaluation of the government's measures against the Northern Territories dispute (column (3)). That is, the treatment group is more likely to positively evaluate the government's effort against the territorial dispute than the control group. The first-stage F-statistic is slightly larger than the conventional level for the case of a single instrument. We also conduct weak instrument robust inference and report the resulting 95% confidence intervals and p-values. Column (3) shows that the confidence intervals include only positive values and that the effect remains significant.

The estimate (0.243) implies that a one standard deviation increase in the viewing time of Public 1 (8.848) increases the evaluation by 2.15, or approximately a half of the control mean. Hence, this effect is modestly large. Overall, the OLS estimates are underestimated, suggesting the second channel of bias as described above.[26]

Finally, the bottom panel shows the estimates of the reduced-form regressions. As expected, the intention-to-treat effect on the evaluation of the government's capability and Prime Minister Abe's visit to Pearl Harbor is null, while that on the evaluation of the Northern Territories dispute is statistically different from zero and positive.

Figure 4 plots the probability distribution of these three outcomes for the treatment and control groups. We regress the dependent variables on the baseline controls and the residuals are then plotted to draw these figures. As expected, the distributions for the two outcomes (the evaluation

[26] Another interpretation might be that the average treatment effect is smaller than the local average treatment effect. However, given that the independent variable is most likely to be endogenous, this is less likely to be the main reason.

of the government's capability and Prime Minister Abe's visit to Pearl Harbor) are the same. By contrast, the distributions for the Northern Territories dispute show that the evaluations shifted to the right for the treatment group compared with the control group. Moreover, the difference is observed over a wide range of domains, suggesting that the effect is not particularly driven by individuals who have certain prior opinions on that issue.

4.3.1 Effect of Not Viewing Private TV Programs

As shown earlier, Table 4 indicates some substitution between Public 1 and Private 4. In this section, we analyze whether the effect on political evaluations also comes from *not* viewing Private 4. Columns (4)–(6) in Table 5 use the difference in the viewing time of Private 4 instead of Public 1 as the independent variable. The top panel shows the OLS estimates, while the middle panel shows the IV (2SLS) estimates for which we use the treatment dummy as the instrument.

According to the table, the OLS estimates for the evaluation of the government's capability to implement policies and on Prime Minister Abe's visit to Pearl Harbor are positive and statistically significant; however, these estimates are most likely to be biased because of endogeneity. By contrast, the IV estimates are significant only for the evaluation of the government's measures against the Northern Territories dispute, although the effect is relatively weak. Moreover, the sign of all the IV estimates becomes negative. Because the F-statistic is smaller than the conventional level, weak instrument robust inference is also conducted. The corresponding confidence interval in column (3) indicates that the true effect is likely to be smaller than zero. Since we know from Table 4 that the treatment seemed to *decrease* the viewing time of Private 4, the negative sign indicates that the positive evaluation of the Northern Territories issue may also come from *not* viewing Private 4 to some degree. However, in the next section, we present evidence suggesting that the result of the positive evaluation is most likely due to viewing Public 1 more rather than viewing Private 4 less.

4.4 Heterogeneity by Trust Levels

Next, we examine the heterogeneous effect by exploiting different trust levels on various broadcasters. To do so, we first subtract trust levels on private broadcasters from trust levels on Public TV to create a variable for the net trust levels on Public TV. The variable take a positive value if an individual trusts Public TV more than private broadcasters, and *vice versa*.

Table 6 shows the results in which the interaction variables between the dummy variable made

from this variable and the treatment variable are included. Columns (1) and (2) show that the treatment effects are mostly driven by those who trusted Public TV more than private broadcasters. The results are consistent with observations that individuals tend to obtain information from media outlets that they trust (Gentzkow et al., 2020). Moreover, Column (3) shows an interesting result that (net) trust levels of those trust Public TV more than private broadcasters are increased by the treatment. As mentioned in the Introduction, we call this as *endogenous persuasion*.

5 Mechanism Behind the Findings

In this section, we use TV script data to investigate the mechanism behind these findings. In particular, we analyze information on the TV programs to which individuals were exposed during the experiment to understand why the treatment had a significant positive effect on the evaluation of certain policies and post-trust levels.

5.1 Closeness to Official Statements

Given that our main outcomes concern foreign policies, we hypothesize that treated individuals might have been exposed more to information that is more similar to that provided by the domestic government than by the foreign government.

To test this hypothesis, we apply an unsupervised machine learning algorithm to compute the semantic similarity between TV programs and the official statements made by domestic (Japan) and foreign governments (United States and Russia). Then, we calculate the sum of the similarity measure between each TV program and official document, weighted by the duration of each program, for each individual. In other words, this variable captures the average exposure to information provided by either government. Finally, exposure to the foreign government's information is subtracted from exposure to the domestic government's information to formulate a *relative* exposure variable.

Formally, for individual i and official statement $j \in \{\text{Japan, USA, Russia}\}$, we compute

$$\text{AvgExposure}_{ij} := \sum_{p} \{1[\text{View}_{ip}] \times \text{Duration}_{p} \times \text{Closeness}_{pj}\}, \tag{3}$$

where $1[\text{View}_{ip}]$ is an indicator variable that takes the value one if i viewed program p and zero otherwise, Duration_{p} denotes the duration of program p measured in hours, and Closeness_{pj} is the similarity between program p and the official statement made by government j, which takes a value

between -1 (very far) and 1 (very close). The higher the value of Closeness_{pj}, the more similar are program p and government j's official statement. We use machine learning to construct this closeness variable.[27]

Next, we take the difference between these measures for the domestic government and foreign government $k \in \{\text{USA, Russia}\}$:

$$\Delta\text{AvgExposure}_{ik} := \text{AvgExposure}_{i\text{Japan}} - \text{AvgExposure}_{ik}. \tag{4}$$

Intuitively, $\Delta\text{AvgExposure}_{ik}$ aims to capture the average exposure to the official statement made by the domestic government relative to the average exposure to the official statement made by the foreign government. This variable takes a higher positive value if individuals are exposed more to information more closely related to that provided by the domestic government than the foreign government.

5.2 Construction of the Closeness Variable

To construct the variable Closeness_{pj} in (3), we apply an unsupervised machine learning algorithm that learns fixed-length feature representations from pieces of texts (Le and Mikolov, 2014).

To obtain the semantics of words from unlabeled data, Firth (1957) initially proposed the distributional hypothesis, which assumes that words used in the same contexts tend to have similar meanings. In our case, the similarity of words is captured by the distributed semantics, which assumes that the meaning of a word is represented by fixed-length numerical vectors.[28]

We train a model such that it maximizes the probability of observing a word given a context (e.g., predict the fourth word from "the," "cat," and "sat").[29] Once the process is finished, we compute the similarity between each official statement and each TV program broadcast during the experiment using the semantic similarity between the words in these documents. We use the average of the word semantics in a document to compute the semantics of that document. Using this definition, two documents are semantically more similar if they contain more words that are semantically more similar.

Let $\mathscr{D} = \{D_1, ..., D_K\}$ be a set of documents (TV programs) used to train a machine and $\mathbf{X}$ be a vector of all the words in a training corpus, with length n. Each document D_k contains a sequence of words. Let $x_i \in \mathbf{X}$ be a target word and $\mathbf{c}_i^k \in \{0,1\}^{n \times 1}$ be the local context $x_{i-s}, ..., x_{i-1}, x_{i+1}, ...,$

[27] The procedure is described in the next section.
[28] The distributional hypothesis and distributed semantics are summarized in, for example, Eisenstein (2019).
[29] This is also called the paragraph vector with distributed memory model.

x_{i+s} of the target, where s is the maximum window size, given the sequence of words in document D_k. Each element c_{ij}^k of the vector takes the value one if word position j in $\mathbf{X}$ appears in the local context of x_i and zero otherwise.

Next, denote $\mathbf{P} \in \mathbb{R}^{h \times n}$ as the projection matrix from an input to a hidden space of length h and $\mathbf{P}^T \in \mathbb{R}^{n \times h}$ as the projection matrix from the hidden space to an output, where $\mathbf{p}_x$ is the column of $\mathbf{P}$ for word x.

Finally, let $\mathbf{D} \in \mathbb{R}^{h \times K}$ be a matrix in which each column represents the memory of the document and $\mathbf{d}^k$ be a column of $\mathbf{D}$, which represents a memory of document D_k.

Using a softmax function, the probability of observing the target word x_i in document D_k given the local context $\mathbf{c}_i^k$ and memory $\mathbf{d}^k$ is

$$P(x_i|\mathbf{c}_i^k, \mathbf{d}^k) = \frac{\exp(\mathbf{p}_{x_i}^T(\mathbf{P}\mathbf{c}_i^k + \mathbf{d}^k))}{\sum_{x \in \mathbf{X}} \exp(\mathbf{p}_x^T(\mathbf{P}\mathbf{c}_i^k + \mathbf{d}^k))}. \tag{5}$$

Then, the projection matrices are trained using stochastic gradient descent (where gradient is obtained via backpropagation) to minimize the loss function:

$$-\sum_{k=1}^{K}\sum_{i=1}^{I_k} \log P(x_i|\mathbf{c}_i^k, \mathbf{d}^k). \tag{6}$$

Given the trained $\mathbf{P}$, one can compute $\mathbf{d}^k$ by taking the average of the columns in $\mathbf{P}$ of the words included in document D_k.[30]

Suppose two documents each contain a single sentence: "I have a cat" and "I have a dog". Thus, $\dim \mathscr{D} = 2$ and $\mathbf{X} = (\text{dog}, \text{cat}, \text{a}, \text{have}, \text{I})^T$ without loss of generality. Let the window size and vector size be $s = 1$ and $h = 2$, respectively.

The sentence "I have a cat" has a sequence of words: "I," "have," "a," and "cat." The local context of "a," given the window size 1, is therefore "have" and "cat." Hence, c_3^1 is written as $(0,1,0,1,0)^T$. Similarly, the local context of "a" in the sentence "I have a dog" has $c_3^2 = (1,0,0,1,0)^T$. These local contexts of the "a"'s are projected onto a two-dimensional hidden space using matrix P and then onto an output using matrix P^T. These matrices are constantly updated in the training process so that "a" can be predicted with a high probability.

When the training is finished, words with similar semantics can be mapped to a similar location in the vector space. For tractability, one can typically evaluate the similarity between words x_i and x_j using cosine similarity $(\mathbf{q}_{x_i} \cdot \mathbf{q}_{x_j})/(\|\mathbf{q}_{x_i}\| \times \|\mathbf{q}_{x_j}\|)$.

[30] In practice, Le and Mikolov (2014) use hirerchical softmax to speed up the process. See Rong (2016).

Applying the same method to measure the semantic similarity between documents, we compute the semantic similarity between documents D_k and D_l using $(\mu_k \cdot \mu_l)/(\|\mu_k\| \times \|\mu_l\|)$, where μ_m is the average of the semantic vectors of the words contained in document D_m, treating duplicates as separate words. Therefore, we assume that the semantics of a document are represented by the average semantics of words in that document. We check this assumption when we set the parameter values.

As described in the Introduction, a major advantage of the algorithm is that it takes into account the ordering and semantics of words as well as the contextual difference across documents. In our case, a document means a TV program. That is, we allow for a word in a program to be semantically different from the same word in another program. Moreover, the algorithm does not require the *ex-ante* selection of words by researchers. Instead, it computes the semantic similarity between words using only a corpus.

The data used to train the machine (i.e., a corpus) are the scripts of TV programs broadcast after the experiment until August 28, 2018.[31] This choice of the corpus is based on the characteristics of our study; other studies may also use other corpora, such as Google Books Corpora, depending on their purpose. We use Python's `gensim` package for all the procedures.

Since Japanese texts, in contrast to western texts, are not separated into single words in the original data, we parse the scripts using a common Japanese parser called `MeCab`. We use only nouns, verbs, adjectives, and adverbs, so that our measures capture the similarity between texts in a meaningful way.[32] Moreover, we apply stemming for all words. The number of TV programs in the training data is 195,016, which contain 613,497 distinct words.

For the official statements, we use those made by the Japanese and US governments when Prime Minister Abe and President Obama visited Pearl Harbor[33] and those by the Japanese and Russian governments when Prime Minister Abe and President Putin met at the summit in Japan.[34] For

[31] The reason for not including the experiment period in the training data is to avoid any direct connection between the texts used for training and texts for measuring closeness. The reason for including only data until August 28, 2018 is that the latest version of the dictionary used in the parsing process (NEologd) contains data only until that date. Although the dictionary is constantly updated, we avoid using the running version to make the replication of our results possible.

[32] The following parts of speech are excluded: interjection, conjunction, particles, auxiliary verbs, prefix, and determiner.

[33] The official statements in Japanese are available on the Cabinet Office's website (http://www.kantei.go.jp/jp/97_abe/statement/2016/1227usa.html) and the US Embassy and Consulates in Japan's website (https://jp.usembassy.gov/ja/remarks-president-obama-and-prime-minister-abe-japan-pearl-harbor-ja/). The official statements in English are also available at the Cabinet Office's website (http://japan.kantei.go.jp/97_abe/statement/201612/1220678_11021.html) and the White House's website (https://obamawhitehouse.archives.gov/the-press-office/2016/12/28/remarks-president-obama-and-prime-minister-abe-japan-pearl-harbor).

[34] The official statements for both countries are only available in Japanese from the Cabinet Office's website

each case, the statements were about the same topic and were spoken to the public almost at the same time in the same place. Thus, the only difference between the official statements for each event is whether the statement was made by either the domestic government or the foreign government.

Before starting, the algorithm requires researchers to decide two parameters: window size (i.e., how many words are considered together in the moving window on texts, or s in the model) and vector size (i.e., how many dimensions are considered to capture the complex meanings of each word, or h in the model). To obtain the baseline numbers for these parameters, we compute the percentage of the closest programs to the official statements that can be categorized as either the United States or Russia topic to see how much we can successfully predict the relevant programs.

In particular, we first list the TV programs and order them from the most to the least close to each statement for each event. We then select several programs from the closest programs and check the percentage that can be categorized as either the United States or Russia topic.[35] Thus, a higher percentage means that we are more successful at predicting the relevant programs for each topic. We choose the top 7, 35, and 70 closest programs, which roughly correspond to 1%, 5%, and 10% of all TV programs in the data.

We start by using two window sizes (i.e., 5 and 15) for the moving window. Figure 5 plots a figure for a window size of 15 and Figure 6 plots a figure for a window size of 5. In both, the left figures use the closeness to the Japanese government's statement (top) and the US government's statement (bottom). By contrast, the right figures use the closeness to the Japanese government's statement (top) and the Russian government's statement (bottom).

These figures show that the machine with a window size of 15 is worse at predicting programs for the United States topic, particularly when we use the closeness to the official US government statement. In addition, Figure 6, which uses a window size of 5, shows that the percentage roughly decreases as the vector size increases for the top 7 programs if we use the closeness to the official Japanese government statement, while it increases for the top 7 programs if we use the closeness to the official US government statement, with the lowest hit when the vector size is 100.

Based on these analyses, we select a window size of 5 and a vector size of 200 for the main analysis. The Online Appendix provides the results using the other parameter values.

(https://www.kantei.go.jp/jp/97_abe/statement/2016/1216kaiken.html).

[35] See the Online Appendix for the procedure of categorizing programs into these topics.

5.3 Results

Table 7 shows the difference in the closeness between the programs and official statements without taking into account participants' viewing behavior (i.e., $\text{Duration}_p \times \text{Closeness}_{pj}$ in (3)). We provide the table to check for any difference in information from the supply side. We first take, for each program, the difference in the closeness to either government (domestic or foreign) and then compare the averages between the two types of broadcasters (Public 1 and private broadcasters). The top panel of the table also takes into account the duration of the program, while the bottom panel does not.

First, the table shows that all the differences are statistically significant. Moreover, they are all positive, indicating that Public 1's information is closer to the domestic government's than to that of the foreign government compared with private broadcasters' information. Interestingly, the difference between Japan and Russia is somewhat larger than that between Japan and the United States.

Next, we take into account participants' actual viewing behavior. Each point in Figure 7 corresponds to the individual's exposure to particular information, namely AvgExposure_{ij} in (3). Exposure to the domestic government's information is plotted against exposure to the foreign government's information. The red lines indicate the 45-degree lines.

Strikingly, the left panels show that the data points tend to be located above the 45-degree lines, especially for individuals who are away from the origin, indicating that they were more likely to be exposed to information closer to the domestic government's information than that of the foreign government by viewing Public 1's programs. This result becomes even starker when compared with the right panels, which plot the same figures, but for the private channels. This shows that the data points are more or less on the 45-degree lines even though there are five times more data points in the panels compared with the respective left panels. These results suggest that individuals were affected by information somewhat unique to Public 1 during the experiment.

To check whether the positive second-stage result on the political evaluations is explained by this informational difference, we next use the closeness measures in the regressions. The top panel of Table 8 shows the first-stage results in which we use the treatment dummy as the main independent variable and the similarity between Public 1 and official statements as the dependent variable. "Δ Japan–USA" means the difference in average closeness between the Japanese and US governments, while "Δ Japan–Russia" is the difference in average closeness between the Japanese and Russian governments.

Columns (1) and (4) include all TV programs broadcast by Public 1, columns (2) and (5) use only programs categorized as the topic related to Prime Minister Abe, and columns (3) and (6) use only programs categorized as the Russia topic. Overall, the table shows that the treatment group was more likely to receive information closer to that provided by the domestic government than by the foreign government. This finding may not be surprising given the slant in Public 1's information found before.

We can also check the heterogeneous effect by trust levels. Column (4) in Table 6 shows that individuals who trust Public TV more than private broadcasters are more likely to be exposed to slanted information.

Next, the top panel of Table 9 shows the OLS estimates in which the evaluation of the Northern Territories dispute are the dependent variables (i.e., the same dependent variables used in column (3) in Table 5) and the closeness measures are used as the independent variables. The bottom panel of the table shows the IV (2SLS) estimates in which the treatment variable is the instrument. The structure of the table is the same as Table 8; the dependent variable uses either all programs or a subset of them focusing on a particular topic depending on the column.

Although the effect is somewhat weak in some specifications, all the IV estimates show a positive sign, indicating that by being exposed to information closer to that provided by the domestic government, individuals are more likely to positively evaluate the government's measures against the Northern Territories dispute. Since the F-statistic in the first stage is smaller than the conventional level, we also conduct weak instrument robust inference and report the p-values and confidence intervals. Compared with standard confidence intervals, these intervals tend to be wider. However, they only include positive values, indicating that the true effect is more likely to be positive than negative or null.

In terms of the magnitude, using the estimate in column (1) (7.389), we find that a one standard deviation (0.681) increase in the independent variable increases the evaluation of the Northern Territories dispute by 5. Similarly, the estimate in column (4) (7.357) shows that a one standard deviation increase in the treatment variable increases the evaluation by 5. Therefore, the effect size is considerable.

Because of the (albeit weak) second-stage effect through Private 4 shown in the previous section, we also examine whether the effect on the political evaluation also comes from *not* being exposed to information closer to that provided by the foreign government. To check this possibility, we formulate closeness measures for Private 4 using the same procedure. The bottom panel of Table 8 shows the first-stage estimates of regressing these measures on the treatment dummy. The table

does not present clear evidence that the treatment induced individuals *not* exposed to information closer to that of the foreign government. Compared with the top panel in the same table, the estimates are also relatively small. This is not surprising given that individuals do not seem to be exposed to the slanted information through Private 4, as shown in Figure 7.

Taken together, the positive evaluation of the Northern Territories issue is more likely to be explained by viewers being exposed *more* to information closer to that provided by the domestic government rather than being exposed *less* to information closer to that provided by the foreign government.

Overall, we find that treated individuals positively evaluated certain foreign policies because they were exposed more to information closer to that provided by the domestic government during the experiment.

6 Discussion

Why in the first place is Public TV's information closer to that provided by the domestic government than that of the foreign government compared to private media information? As noted in the Introduction, the literature shows that media slant is driven by either demand-side factors or supply-side factors. Regarding the demand-side factors, since viewers pay receiving fees, regardless of their political orientation, it is hard to believe that consumers' political preferences (e.g., right vs. left) drive the slant in public media's information. The same argument holds for public media in other countries, such as Sweden and the United Kingdom, in which most adults pay similar fees or taxes.

Turning to the supply-side factors, since receiving fees are mandatory for anyone who has TV equipment, changing information in public media should not affect their revenue. Furthermore, as mentioned in the Background, their annual budgets and board members need the approval of the Diet and they do not receive any revenue from commercial advertising. These factors indicate that the most plausible explanation could be that public media are more susceptible to media capture by the government, if anything.

An alternative hypothesis could be that the public media's information might be more slanted toward the ruling party than the opposition party compared with that from the private media. We test this hypothesis in the Online Appendix, but find no results supporting it. However, we do find that public media information is slanted toward the ruling party. In other words, competition between political parties does not seem to explain our results.

Finally, although we find that the public media's information tends to be closer to the domestic government's information for both topics, only the result for the Northern Territories dispute in Table 5 is significant. A possible explanation might be that the summit only just ended when the endline survey was conducted, while Prime Minister Abe's visit to Pearl Harbor had not. The former event might thus have endorsed individuals' beliefs and hence influenced their survey answers.

7 Conclusion

In this paper, we study how media trust and persuasion are interlinked. We conducted a large-scale randomized experiment in Japan in collaboration with the nation's most trusted media outlet. In the experiment, we recruited 6,000 individuals and randomly increased the treatment group's capacity to view Public TV's programs by providing them with free subscriptions. We find that our treatment increased the subjects' viewing time of Public TV's programs as well as their support for certain policies. Interestingly, we find that the effect was larger for those who trusted Public TV more than private broadcasters, and that, more surprisingly, their trust levels were even increased by the treatment, which we call *endogenous persuasion* in the paper.

An important policy implication from our findings is that having more trustful media outlets in the market or directing individuals to such outlets is not enough or even necessary to mitigate political polarization because this can possibly exacerbate it. Rather, creating institutions that secure the independence of trustful public media may be needed to enable consumers to compare the information from slanted media sources with that from more neutral information sources (Mullainathan and Shleifer, 2005; Gentzkow and Shapiro, 2006).

Although we conducted a randomized experiment to obtain a generalizable causal effect, the external validity of our results should be tested in different settings.

References

Aronson, E. (1999). The power of self-persuasion. *American Psychologist 54*(11), 875–884.

Ash, E. and E. Labzina (2019). Does Politicized Media Distort Political Discourse? Evidence from U.S. Cable News Channels. *Mimeo*.

Besley, T. and A. Prat (2006). Handcuffs for the Grabbing Hand? Media Capture and Government Accountability. *American Economic Review 96*(3), 720–736.

Boxell, L., M. Gentzkow, and J. M. Shapiro (2017). Greater Internet use is not associated with faster growth in political polarization among US demographic groups. *Proceedings of the National Academy of Sciences 114*(40), 10612–10617.

Campante, F. R., R. Durante, and F. Sobbrio (2018). Politics 2.0: The Multifaceted Effect of Broadband Internet on Political Participation. *Journal of the European Economic Association 16*(4), 1094–1136.

Chen, Y. and D. Yang (2019). The Impact of Media Censorship: 1984 or Brave New World? *American Economic Review 109*(6), 2294–2332.

Chugrov, S. V. and D. V. Streltsov (2017). Interdependence of russo-japanese relations and mutual images of japan and russia. *Japanese Journal of Political Science 18*(1), 22–40.

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

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

Durante, R., P. Pinotti, and A. Tesei (2019). The Political Legacy of Entertainment TV. *American Economic Review 109*(7), 2497–2530.

Eisenstein, J. (2019). *Introduction to Natural Language Processing*. Cambridge, MA: MIT Press.

Enikolopov, R., M. Petrova, and E. Zhuravskaya (2011). Media and Political Persuasion: Evidence from Russia. *American Economic Review 101*(7), 3253–3285.

Falck, O., R. Gold, and S. Heblich (2014). E-lections: Voting Behavior and the Internet. *American Economic Review 104*(7), 2238–2265.

Firth, J. R. (1957). *Papers in linguistics, 1934-1951*. London: Oxford University Press.

Gavazza, A., M. Nardotto, and T. M. Valletti (2019). Internet and Politics: Evidence from U.K. Local Elections and Local Government Policies. *Review of Economic Studies 86*(5), 2092–2135.

Gentzkow, M. (2006). Television and Voter Turnout. *The Quarterly Journal of Economics 121*(3), 931–972.

Gentzkow, M. (2016). Polarization in 2016. *Mimeo*.

Gentzkow, M., N. Petek, J. M. Shapiro, and M. Sinkinson (2015). Do Newspapers Serve The State? Incumbent Party Influence on The US Press, 1869–1928. *Journal of the European Economic Association 13*(1), 29–61.

Gentzkow, M. and J. M. Shapiro (2006). Media Bias and Reputation. *Journal of Political Economy 114*(2), 280–316.

Gentzkow, M. and J. M. Shapiro (2010). What Drives Media Slant? Evidence from U.S. Daily Newspapers. *Econometrica 78*(1), 35–71.

Gentzkow, M., J. M. Shapiro, and M. Taddy (2019). Measuring Group Differences in High-Dimensional Choices: Method and Application to Congressional Speech. *Econometrica 87*(4), 1307–1340.

Gentzkow, M., M. B. Wong, and A. T. Zhang (2020). Ideological Bias and Trust in Information Sources. *Mimeo*.

Gerber, A. S., D. Karlan, and D. Bergan (2009). Does the Media Matter? A Field Experiment Measuring the Effect of Newspapers on Voting Behavior and Political Opinions. *American Economic Journal: Applied Economics 1*(2), 35–52.

Groseclose, T. and J. Milyo (2005). A Measure of Media Bias. *The Quarterly Journal of Economics 120*(4), 1191–1237.

Jensen, J., E. Kaplan, S. Naidu, and L. Wilse-Samson (2012). Political Polarization and the Dynamics of Political Language: Evidence from 130 Years of Partisan Speech. *Brookings Papers on Economic Activity*, 1–81.

Kuroda, T., D. Kido, and Y. Ichikawa (2017). Does the usage of TV streaming on the Internet bring about new functions to the public service broadcaster?: Evidence from a randomized field experiment. *Mimeo*.

Laver, M., K. Benoit, and J. Garry (2003). Extracting Policy Positions from Political Texts Using Words as Data. *American Political Science Review 97*(2), 311–331.

Le, Q. V. and T. Mikolov (2014). Distributed Representations of Sentences and Documents. *CoRR abs/1405.4*.

Martin, G. J. and A. Yurukoglu (2017). Bias in cable news: Persuasion and polarization. *American Economic Review 107*(9), 2565–2599.

Mullainathan, S. and A. Shleifer (2005). The Market for News. *American Economic Review 95*(4), 1031–1053.

Newman, N., R. Fletcher, A. Kalogeropoulos, D. A. L. Levy, and R. K. Nielsen (2018). *Reuters Institute Digital News Report 2018*. Oxford: Reuters Institute for the Study of Journalism.

Olea, J. L. M. and C. Pflueger (2013). A robust test for weak instruments. *Journal of Business & Economic Statistics 31*(3), 358–369.

Pennycook, G. and D. G. Rand (2019). Fighting misinformation on social media using crowdsourced judgments of news source quality. *Proceedings of the National Academy of Sciences 116*(7), 2521–2526.

Peterson, A. and A. Spirling (2018). Classification Accuracy as a Substantive Quantity of Interest: Measuring Polarization in Westminster Systems. *Political Analysis 26*(1), 120–128.

Prior, M. (2005). News vs. Entertainment: How Increasing Media Choice Widens Gaps in Political Knowledge and Turnout. *American Journal of Political Science 49*(3), 577–592.

Qin, B., D. Strömberg, and Y. Wu (2017). Why Does China Allow Freer Social Media? Protests versus Surveillance and Propaganda. *Journal of Economic Perspectives 31*(1), 117–140.

Qin, B., D. Strömberg, and Y. Wu (2018). Media bias in China. *American Economic Review 108*(9), 2442–2476.

Rong, X. (2016). word2vec Parameter Learning Explained. *Mimeo*.

Schroeder, E. and D. F. Stone (2015). Fox News and political knowledge. *Journal of Public Economics 126*, 52–63.

Strömberg, D. (2015). Media Coverage and Political Accountability: Theory and Evidence. In S. P. Anderson, J. Waldfogel, and D. Strömberg (Eds.), *Handbook of Media Economics*, Volume 1 of *Handbook of Media Economics*, pp. 595–622. Amsterdam: North Holland Publishing Co.

Tetlock, P. C. (2007). Giving Content to Investor Sentiment: The Role of Media in the Stock Market. *The Journal of Finance 62*(3), 1139–1168.

The Economist (2016). Two men in a tub. `https://www.economist.com/asia/2016/12/10/japans-prime-minister-plans-a-steamy-tete-a-tete-with-russias-president`.

Tables and Figures

Table 1: Summary statistics

	Mean	Std. Dev.	Min	Max	Obs.
Total view during the experiment (hour)					
Public_1	11.03	19.87	0.00	174.55	4452
Public_2	1.03	3.86	0.00	74.22	4452
Private_4	25.23	33.38	0.00	413.88	4452
Private_6	17.03	23.02	0.00	309.90	4452
Private_8	16.77	25.94	0.00	230.25	4452
Private_10	22.26	28.68	0.00	362.97	4452
Private_12	6.01	9.34	0.00	130.12	4452
Difference in total view					
Δ Public_1	0.94	8.69	−44.97	116.07	4452
Δ Public_2	0.24	2.14	−17.37	74.22	4452
Δ Private_4	8.16	15.13	−62.15	134.75	4452
Δ Private_6	5.18	11.27	−66.05	123.45	4452
Δ Private_8	5.79	11.76	−56.00	113.37	4452
Δ Private_10	7.45	13.16	−70.98	135.70	4452
Δ Private_12	1.24	4.97	−34.45	41.32	4452
Closeness					
Δ Japan-USA	0.37	0.66	−0.28	6.43	4452
Δ Japan-Russia	0.41	0.75	−0.34	6.26	4452
Treatment					
Treatment	0.83	0.37	0.00	1.00	4452
Ad.	0.67	0.47	0.00	1.00	4452
Ad. + info.	0.33	0.47	0.00	1.00	4452
Political evaluation					
Gen. eval.	2.33	0.88	1.00	4.00	4452
Pearl Harbor	6.85	2.54	0.00	10.00	4452
North. Terr.	4.51	2.55	0.00	10.00	4452

Table 2: The average treatment effect on watching Public 1

	Dependent variable: Δ Public_1		
	(1)	(2)	(3)
Treatment	0.815	0.887	1.082
	(0.329)**	(0.331)***	(0.336)***
Controls	No	Yes	Yes
Mean of control	0.27	0.27	0.27
R^2	0.00	0.01	0.01
Observations	4452	4452	4156

Note. Robust standard errors are in parentheses. The dependent variables are the difference in the viewing time of Public 1 before and during the experiment (in hours). "Treatment" is the treatment dummy. The control variables include a dummy for supporting the Liberal Democratic Party, a minor party, and no party, a dummy for not reporting annual income and for reading the Sankei newspaper, and trust and accuracy measures. Column (3) drops individuals who occasionally failed to report the viewing records of certain programs, although they were likely to have seen them according to the access log. *, **, and *** indicate $p < 0.10$, $p < 0.05$, and $p < 0.01$, respectively.

Table 3: The average treatment effect on watching Public 1, by week

	Dependent variable: Public_1		
	(1) Week 1	(2) Week 2	(3) Week 3
Treatment	0.300	0.420	0.321
	(0.125)**	(0.125)***	(0.125)**
Controls	Yes	Yes	Yes
Mean of control	3.72	3.36	3.47
R^2	0.80	0.78	0.77
Observations	4156	4156	4156

Note. Robust standard errors are in parentheses. The dependent variables are the weekly viewing time of Public 1 during the experiment (in hours). Week 1: November 28–December 4, Week 2: December 5–11, and Week 3: December 12–18. "Treatment" is the treatment dummy. The control variables include a dummy for supporting the Liberal Democratic Party, a minor party, and no party, a dummy for not reporting annual income and for reading the Sankei newspaper, trust and accuracy measures, and the viewing time of Public 1 before the experiment. *, **, and *** indicate $p < 0.10$, $p < 0.05$, and $p < 0.01$, respectively.

Table 4: The average treatment effect on watching TV programs, by channel

	Dependent variable: Δ TV view							
	(1) Public_1	(2) Public_2	(3) Private_4	(4) Private_6	(5) Private_8	(6) Private_10	(7) Private_12	(8) All
Level								
Treatment	1.082	−0.028	−1.538	−0.513	−0.237	−0.409	−0.355	−1.999
	(0.336)***	(0.077)	(0.624)**	(0.463)	(0.462)	(0.548)	(0.209)*	(1.473)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean of control	0.27	0.24	9.28	5.44	5.81	7.74	1.51	30.28
Mean of control (before)	10.28	0.71	17.96	11.16	10.24	15.06	5.04	70.44
Mean of control (during)	10.54	0.95	27.23	16.60	16.05	22.80	22.80	100.72
R^2	0.01	0.00	0.01	0.01	0.01	0.00	0.00	0.01
Observations	4156	4156	4156	4156	4156	4156	4156	4156
Share								
Treatment	0.014	−0.001	−0.010	−0.006	−0.001	0.003	0.001	
	(0.005)***	(0.001)	(0.005)**	(0.004)	(0.004)	(0.005)	(0.003)	
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	
Mean of control	-0.04	0.00	0.02	0.01	0.01	0.01	-0.01	
Mean of control (before)	0.18	0.01	0.24	0.15	0.14	0.21	0.08	
Mean of control (during)	0.14	0.01	0.26	0.16	0.15	0.22	0.07	
R^2	0.01	0.00	0.00	0.00	0.00	0.00	0.00	
Observations	3934	3934	3934	3934	3934	3934	3934	

Note. Robust standard errors are in parentheses. The dependent variables are the difference in the viewing time of TV programs before and during the experiment for each channel, except for column (8) for which the dependent variable uses the total viewing time of all channels (in hours). "Treatment" is the treatment dummy. The control variables include a dummy for supporting the Liberal Democratic Party, a minor party, and no party, a dummy for not reporting annual income and for reading the Sankei newspaper, and trust and accuracy measures. *, **, and *** indicate $p < 0.10$, $p < 0.05$, and $p < 0.01$, respectively.

Table 5: Effect of watching TV programs on political evaluations

	Dependent variable: Evaluation					
	Public_1			Private_4		
	(1) Gen. eval.	(2) Pearl Harbor	(3) North. Terr.	(4) Gen. eval.	(5) Pearl Harbor	(6) North. Terr.
OLS Estimates						
Δ TV view	0.002	0.006	−0.002	0.002	0.011	−0.000
	(0.001)	(0.004)	(0.004)	(0.001)**	(0.002)***	(0.002)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Mean of control	2.34	6.90	4.41	2.34	6.90	4.41
R^2	0.23	0.11	0.11	0.23	0.12	0.11
Observations	4156	4156	4156	4156	4156	4156
IV Estimates						
Δ TV view	0.042	0.044	0.243	−0.030	−0.031	−0.171
	(0.031)	(0.090)	(0.115)**	(0.023)	(0.065)	(0.092)*
Controls	Yes	Yes	Yes			
Mean of control	2.34	6.90	4.41	Yes	Yes	Yes
F-stat.	10.38	10.38	10.38	6.07	6.07	6.07
Effect. F-stat.	10.38	10.38	10.38	6.07	6.07	6.07
p-value (AR)	0.14	0.62	0.01	0.14	0.62	0.01
CI (AR)	[-.013, .144]	[-.153, .284]	[.074, .696]	[-.121, .010]	[-.285, .121]	[-.531,-.051]
CI (Wald)	[-.019, .103]	[-.133, .221]	[.017, .469]	[-.075, .016]	[-.158, .096]	[-.351, .009]
Observations	4156	4156	4156	4156	4156	4156
Reduced-form						
Treatment	0.046	0.048	0.263			
	(0.031)	(0.097)	(0.095)***			
Controls	Yes	Yes	Yes			
Mean of control	2.34	6.90	4.41			
R^2	0.23	0.11	0.12			
Observations	4156	4156	4156			

Note. Robust standard errors are in parentheses. The dependent variables are (a) the evaluation of the government's capability ("Do you evaluate that the current government has the capability to implement policies?" (4: Yes, 3: More or less, 2: Not quite, 1: No)) (columns (1) and (4)), (b) the evaluation of Prime Minister Abe's visit to Pearl Harbor with President Obama ("How do you evaluate Prime Minister Abe's visit to Pearl Harbor with President Obama?" (0 to 10 points)) (columns (2) and (5)), and (c) the evaluation of the government's measures against the Northern Territories dispute ("How do you evaluate the government's measures against the Northern Territories dispute?" (0 to 10 points)) (columns (3) and (6)).“Δ TV view" is the difference in the viewing time of either Public 1 (columns (1)–(3)) or Private 4 (columns (4)–(6)) before and during the experiment (in hours). The IV is the treatment dummy. The control variables include a dummy for supporting the Liberal Democratic Party, a minor party, and no party, a dummy for not reporting annual income and for reading the Sankei newspaper, and trust and accuracy measures. Effective F-stat. is the effective F-statistic of Olea and Pflueger (2013). We use Stata's `weakivtest` to compute the test statistic. Weak instrument robust inference (Anderson–Rubin (AR) test) is conducted and the corresponding p-values and confidence intervals are reported in rows labeled as "p-value (AR)" and "CI (AR)," respectively. We use Stata's command `rivtest` to compute the test statistic. *, **, and *** indicate $p < 0.10$, $p < 0.05$, and $p < 0.01$, respectively.

Table 6: Heterogeneous effect by trust levels

	Dependent variable:			
	(1) Δ NHK Gen.	(2) North. Terr.	(3) Trust	(4) Closeness (Russia)
Treatment × Trust (Public more)	1.129	0.245	0.054	0.021
	(0.350)***	(0.097)**	(0.022)**	(0.011)*
Treatment × Trust (Private more)	0.041	0.661	0.270	0.019
	(0.346)	(0.497)	(0.169)	(0.013)
Trust	−0.850	−0.058	0.563	0.045
	(0.523)	(0.475)	(0.155)***	(0.017)***
Controls	Yes	Yes	Yes	Yes
Mean of control	0.267	4.405	0.220	0.126
H0: b[Treatment × Trust (High)] = b[Treatment × Trust (Low)] (p-value)	0.027	0.411	0.205	0.863
R^2	0.01	0.12	0.16	0.04
N	4156	4156	4156	4156

Note. Robust standard errors are in parentheses. The dependent variables are the difference in the viewing time of TV programs before and during the experiment for Public 1, difference in the closeness to the official statements made by the domestic and foreign governments, net trust levels on Public 1 relative to those on private channels, and the evaluation of the government's measures against the Northern Territories dispute, respectively. "Treatment × Trust (Public more)" is the interaction between the treatment variable and the net trust dummy which takes a value of one if an individual trusts Public TV more than private channels, and zero otherwise. "Treatment × Trust (Private more)" is the interaction between the treatment variable and one minus the net trust dummy. The control variables include a dummy for supporting the Liberal Democratic Party, a minor party, and no party, a dummy for not reporting annual income and for reading the Sankei newspaper, trust and accuracy measures, and the viewing time of Public 1 before the experiment. *, **, and *** indicate $p < 0.10$, $p < 0.05$, and $p < 0.01$, respectively.

Table 7: Difference in closeness

	Public_1 - Private TVs	p-value
Interacted with duration		
Japan–USA	0.225	0.025
Japan–Russia	1.133	0.000
Without interaction		
Japan–USA	0.015	0.000
Japan–Russia	0.031	0.000

Note. For each program, the difference in the closeness to either government (domestic or foreign) is computed. Then, their averages are compared for the two types of broadcasters (Public 1 and (the average of) private channels). The differences and associated p-values are reported. The top panel takes into account the duration of the program, while the bottom panel does not. We use machine learning to compute the closeness measures, where the window size is set to 5 and the vector size is set to 200. See the main text for more details.

Table 8: Effect on the difference in the closeness to the official statements

	Dependent variable: Average Exposure					
	Δ Japan–USA			Δ Japan–Russia		
	(1) All	(2) PM	(3) Russia	(4) All	(5) PM	(6) Russia
Public_1						
Treatment	0.034	0.022	0.011	0.037	0.025	0.016
	(0.014)**	(0.007)***	(0.005)**	(0.015)**	(0.009)***	(0.007)**
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Mean of control	0.35	0.13	0.08	0.40	0.18	0.13
R^2	0.74	0.66	0.60	0.77	0.65	0.63
Observations	4156	4156	4156	4156	4156	4156
Private_4						
Treatment	−0.005	−0.008	−0.007	0.008	−0.009	−0.007
	(0.013)	(0.008)	(0.005)	(0.016)	(0.009)	(0.006)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Mean of control	0.19	0.16	0.10	-0.10	0.14	0.10
R^2	0.30	0.50	0.44	0.03	0.46	0.42
Observations	4156	4156	4156	4156	4156	4156

Note. Robust standard errors are in parentheses. The dependent variables are the difference in the closeness to the official statements made by the domestic and foreign governments. We use machine learning to compute the closeness measures, where the window size is set to 5 and the vector size is set to 200. The top panel uses the programs on Public 1, while the bottom panel uses the programs on Private 4. The first three columns use the difference in average closeness to the official statements made by the Japanese and US governments during Prime Minister Abe's visit to Pearl Harbor with President Barack Obama. The next three columns use the difference in average closeness to the official statements made by the Japanese and Russian governments during President Putin's visit to Japan for the summit. Columns (1) and (4) include all programs. Columns (2) and (5) restrict the programs to those related to Prime Minister Abe, while columns (3) and (6) restrict them to those related to Russia. "Treatment" is the treatment dummy. The control variables include a dummy for supporting the Liberal Democratic Party, a minor party, and no party, a dummy for not reporting annual income and for reading the Sankei newspaper, trust and accuracy measures, and the viewing time of the respective channels before the experiment. *, **, and *** indicate $p<0.10$, $p<0.05$, and $p<0.01$, respectively.

Table 9: Effect on political evaluations of the difference in the closeness to the official statements

	Dependent variable: Evaluation (North. Terr.)					
	(1) All	(2) PM	(3) Russia	(4) All	(5) PM	(6) Russia
OLS Estimates						
Δ Japan-USA	0.291	0.375	0.794			
	(0.105)***	(0.213)*	(0.332)**			
Δ Japan-Russia				0.140	0.208	0.329
				(0.103)	(0.155)	(0.216)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Mean of control	4.41	4.41	4.41	4.41	4.41	4.41
R^2	0.12	0.12	0.12	0.12	0.12	0.12
Observations	4156	4156	4156	4156	4156	4156
IV Estimates						
Δ Japan-USA	7.871	12.384	24.994			
	(4.094)*	(5.715)**	(13.402)*			
Δ Japan-Russia				7.357	10.929	17.338
				(3.800)*	(5.544)**	(9.532)*
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Mean of control	4.41	4.41	4.41	4.41	4.41	4.41
F-stat.	6.29	10.21	5.69	6.21	6.78	5.18
Effect. F-stat.	6.29	10.21	5.69	6.21	6.78	5.18
p-value (AR)	0.00	0.00	0.00	0.00	0.00	0.00
CI (AR)	[2.522, 23.919]	[4.010, 34.787]	[7.482, 77.530]	[2.391, 22.255]	[3.246, 32.661]	[4.883, 54.702]
CI (Wald)	[-.152, 15.895]	[1.182, 23.585]	[-1.274, 51.262]	[-.092, 14.806]	[.063, 21.795]	[-1.344, 36.020]
Observations	4156	4156	4156	4156	4156	4156

Note. Robust standard errors are in parentheses. The dependent variables are the evaluation of the government's measures against the Northern Territories dispute. The main independent variables are the difference in the closeness to the official statements made by the domestic and foreign governments. We use machine learning to compute the closeness measures, where the window size is set to 5 and the vector size is set to 200. "Δ Japan–USA" is the difference in average closeness between TV programs and the official statements made by the Japanese and US governments during Prime Minister Abe's visit to Pearl Harbor with President Obama. "Δ Japan–Russia" is the difference in average closeness to the official statements made by the Japanese and Russian governments during President Putin's visit to Japan for the summit. Columns (1) and (4) include all programs. Columns (2) and (5) restrict the sample to programs related to Prime Minister Abe, while columns (3) and (6) restrict to programs related to Russia. The instrument is the treatment dummy. The control variables include a dummy for supporting the Liberal Democratic Party, a minor party, and no party, a dummy for not reporting annual income and for reading the Sankei newspaper, trust and accuracy measures, and the viewing time of Public 1 before the experiment. Effective F-stat. is the effective F-statistic of Olea and Pflueger (2013). We use Stata's `weakivtest` to compute the test statistic. Weak instrument robust inference (AR test) is conducted and the corresponding p-values and confidence intervals are reported in rows labeled as "p-value (AR)" and "CI (AR)," respectively. We use Stata's command `rivtest` to compute the test statistic. *, **, and *** indicate $p<0.10$, $p<0.05$, and $p<0.01$, respectively.

Figure 1: Timeline of the experiment

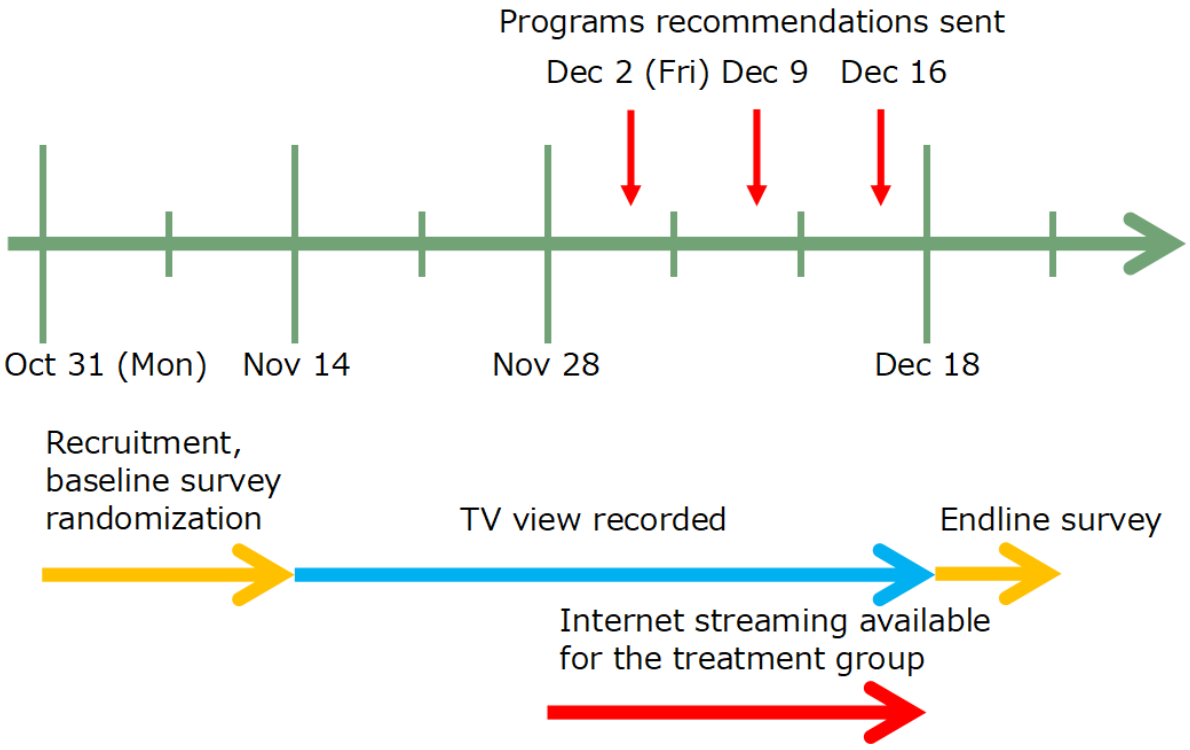
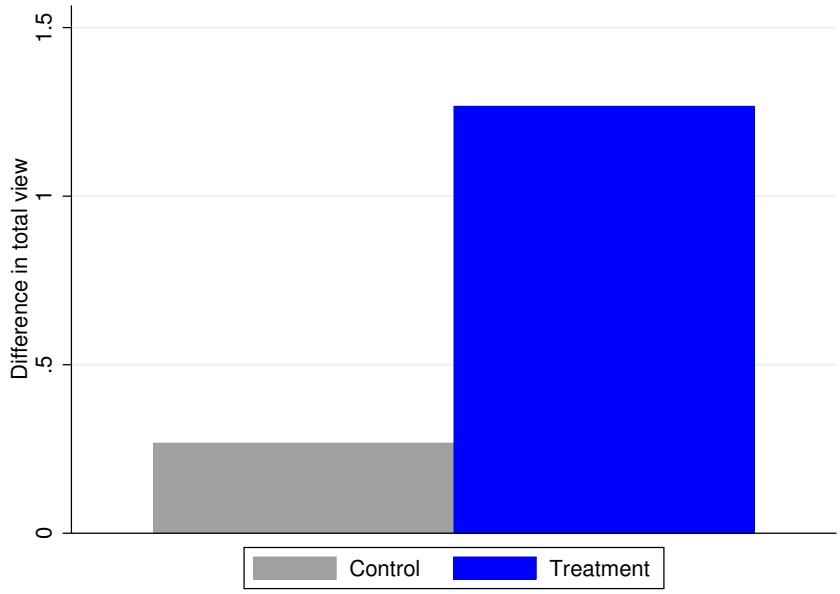
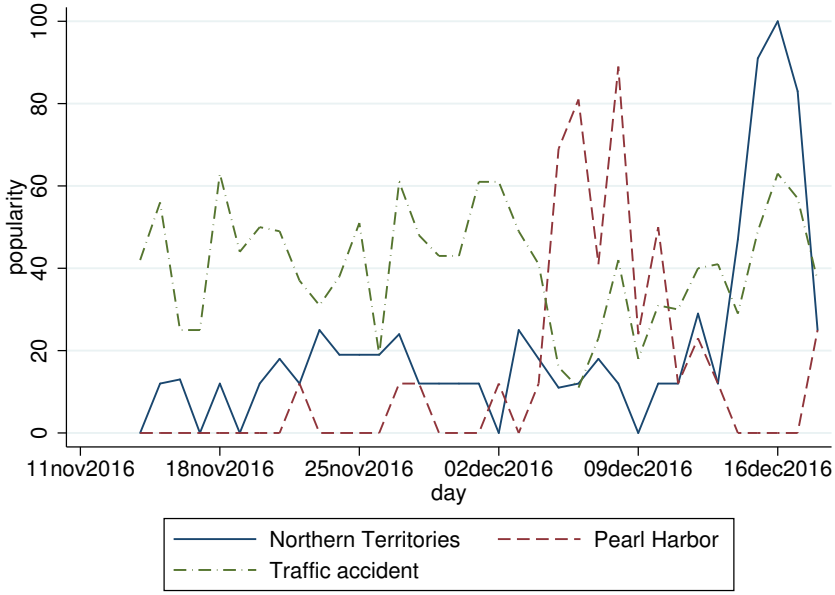


Figure 2: The average viewing time of Public 1, by treatment status



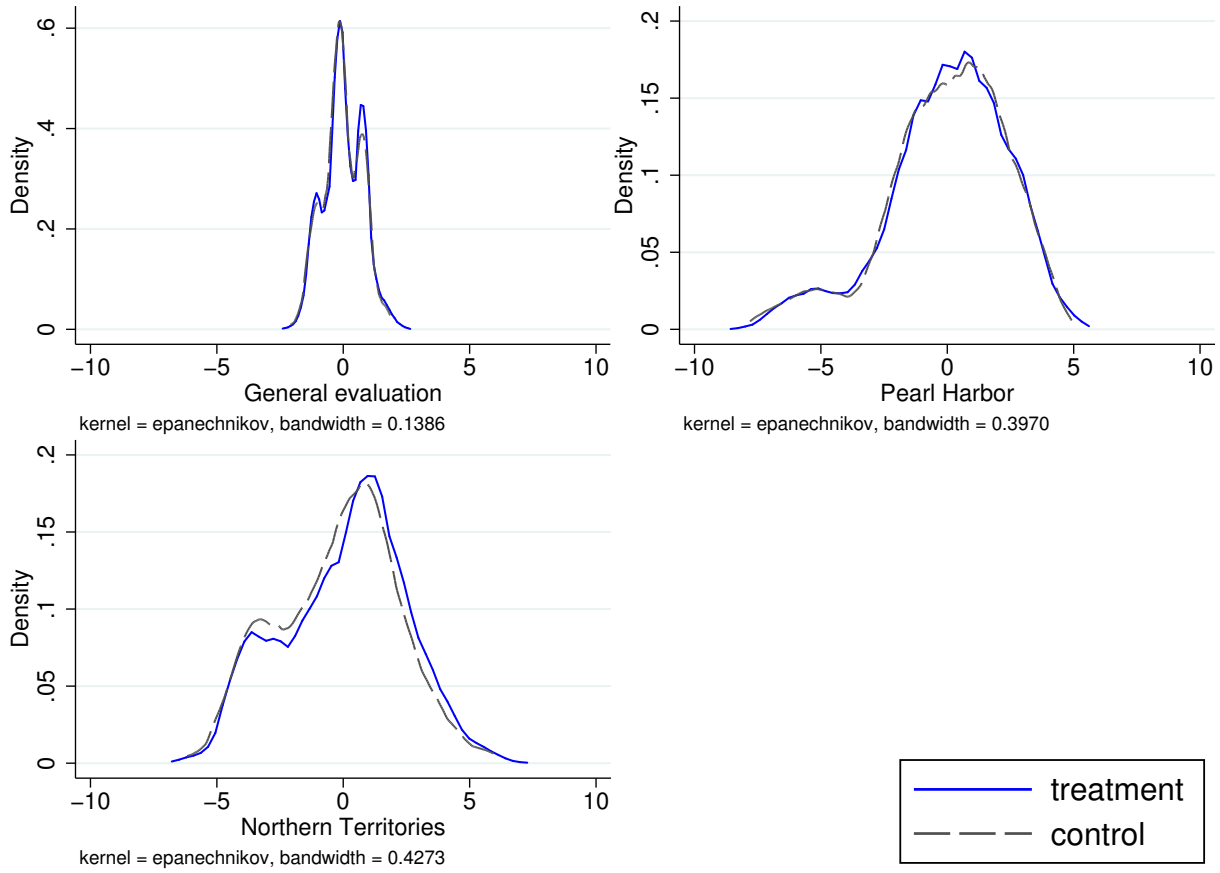
Note. Difference in the viewing time of Public 1 before and during the experiment are averaged for the treatment and control groups (in hours).

Figure 3: Popularity of each topic



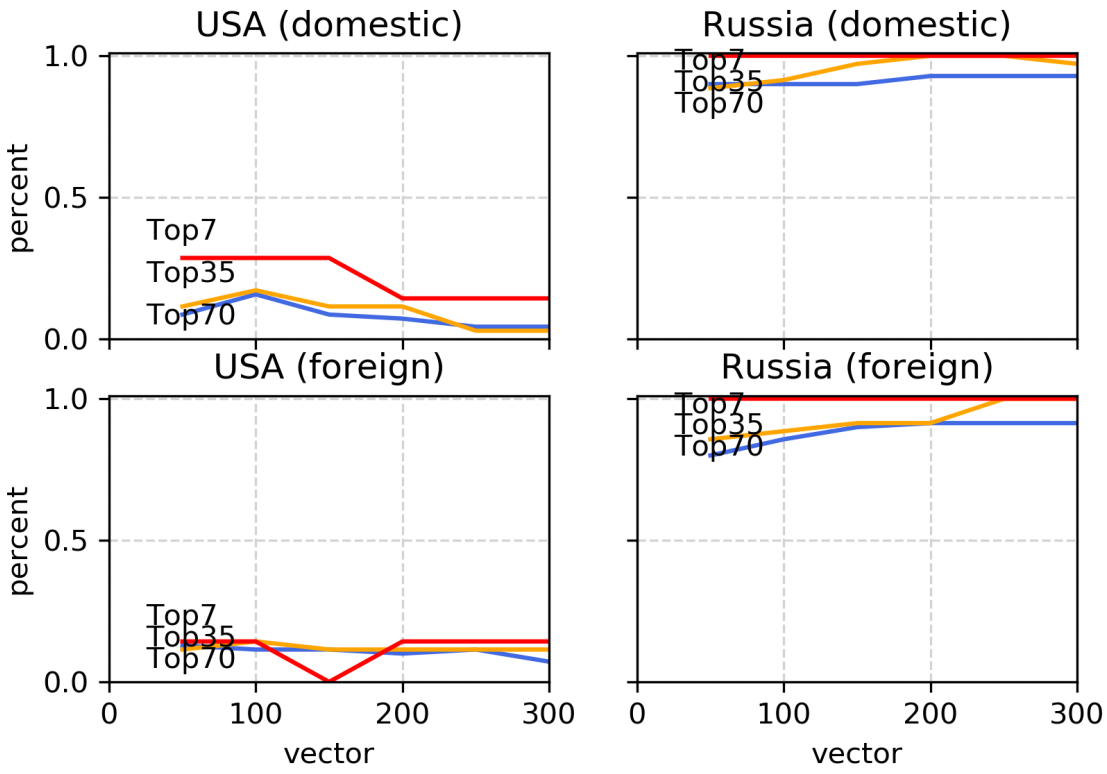
Note. Popularity is measured by the frequency of web searches using Google Trends. The location is set as Japan and the category is set as news. "北方領土" (Northern Territories) is used for Northern Territories, "真珠湾" (Pearl Harbor) is used for Pearl Harbor, and "交通事故" (traffic accident) is used for traffic accidents.

Figure 4: Distributions of political outcomes: Treatment versus control



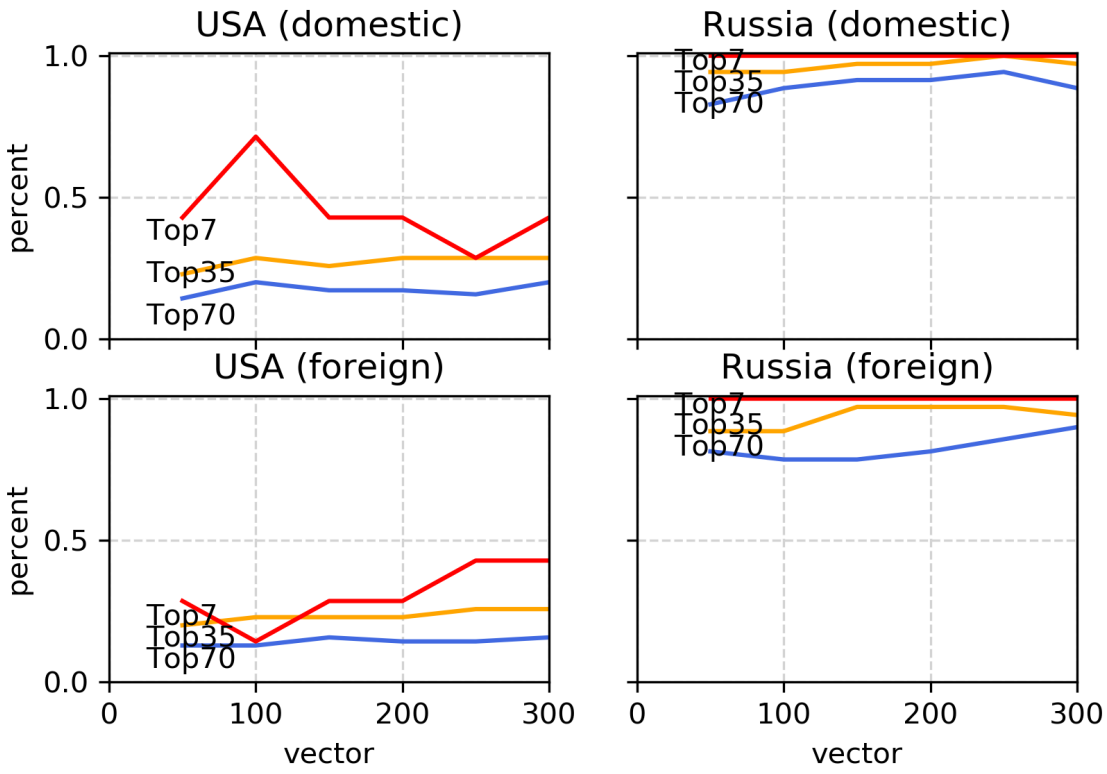
Note. The kernel density plots of the residuals from the regressions of the dependent variables on the control variables. The dependent variable for the top-left figure is the evaluation of the government's capability ("Do you evaluate that the current government has the capability to implement policies?" (4: Yes, 3: More or less, 2: Not quite, 1: No)). The dependent variable for the top-right figure is the evaluation of Prime Minister Abe's visit to Pearl Harbor with President Obama ("How do you evaluate Prime Minister Abe's visit to Pearl Harbor with President Obama?" (0 to 10 points)). The dependent variable for the bottom-left figure is the evaluation of the government's measures against the Northern Territories dispute ("How do you evaluate the government's measures against the Northern Territories dispute?" (0 to 10 points)). The control variables include a dummy for supporting the Liberal Democratic Party, a minor party, and no party, a dummy for not reporting annual income and for reading the Sankei newspaper, and trust and accuracy measures.

Figure 5: Predictability of the machine learning, United States and Russia topics (window=15)



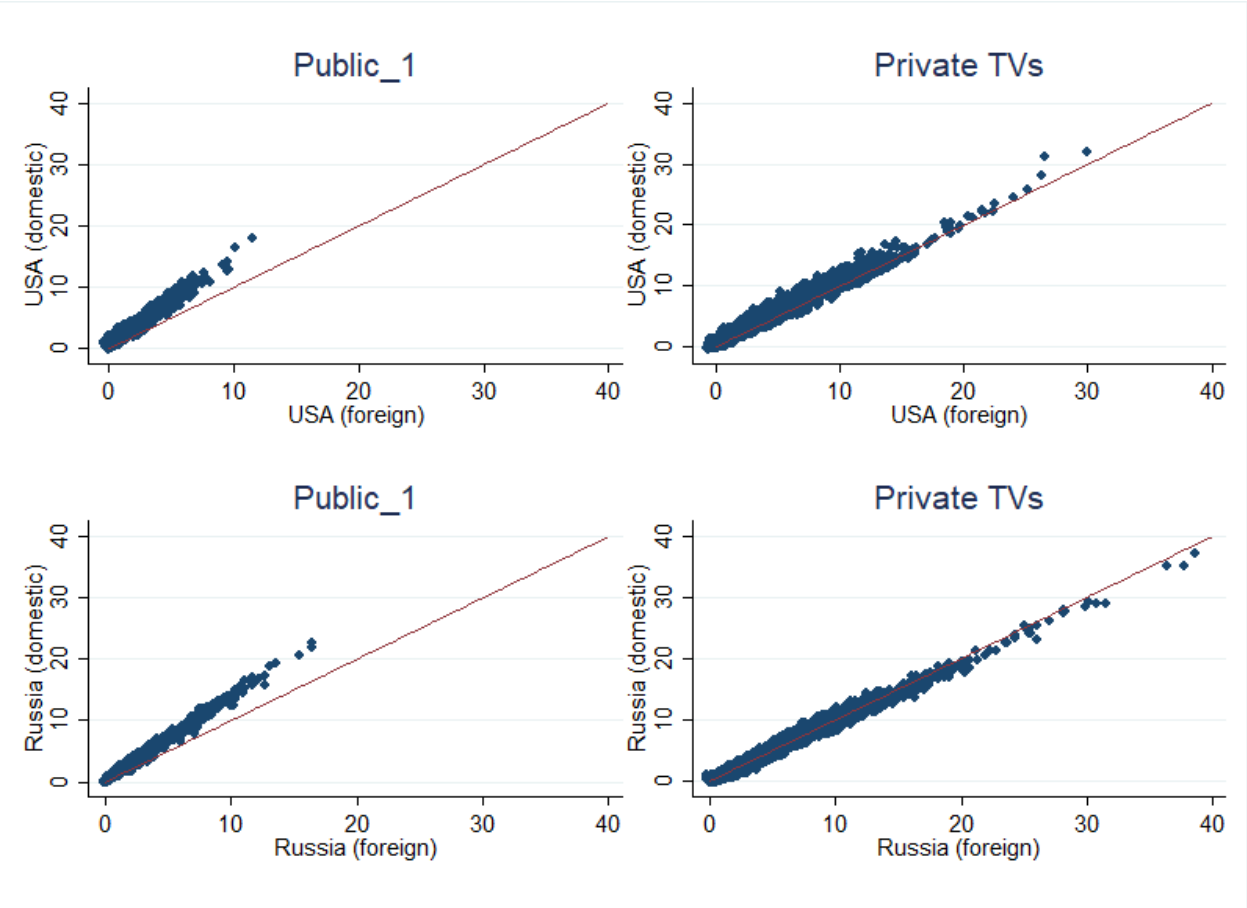
Note. The figures plot the percentage of TV programs categorized as either the United States topic or the Russia topic for those that are the closest to the official statements (top 7, top 35, and top 70 closest programs): the left figures use the closeness to the Japanese government's statement (top) and the US government's statement (bottom) during Prime Minister Abe's visit to Pearl Harbor with President Obama; the right figures use the Japanese government's statement (top) and the Russian government's statement (bottom) during President Putin's visit to Japan for the summit. We use machine learning to compute the closeness measures, where the window size is set to 15. See the main text for more details.

Figure 6: Predictability of the machine learning, United States and Russia topics (window=5)



Note. The figures plot the percentage of TV programs categorized as either the United States topic or the Russia topic for those that are the closest to the official statements (top 7, top 35, and top 70 closest programs): the left figures use the closeness to the Japanese government's statement (top) and the US government's statement (bottom) during Prime Minister Abe's visit to Pearl Harbor with President Obama; the right figures use the Japanese government's statement (top) and the Russian government's statement (bottom) during President Putin's visit to Japan for the summit. We use machine learning to compute the closeness measures, where the window size is set to 5. See the main text for more details.

Figure 7: Average closeness to the official statements interacted with individuals' viewing time



Note. These figures plot the closeness to the official statements interacted with individuals' viewing time (AvgExposure in (3)). Each data point represents an individual. The figure shows the results for Public 1 and all private channels. In the top-left panel, the closeness between Public 1's programs and the Japanese government's statement (interacted with duration) (y-axis) is plotted against the closeness between its programs and the US government's statement (interacted with duration) (x-axis), both made when Prime Minister Abe visited Pearl Harbor with President Obama. The top-right panel shows the same, except that closeness is measured for the private channels. Since all five private channels are plotted on the same figure, the number of data points in this panel is five times larger than that in the top-left panel. The bottom panels show the same, except that closeness is measured for the official statements made by the Japanese and Russian governments at the time of the Japan–Russia summit. We use machine learning to compute the closeness measures, where the window size is set to 5, and the vector size is set to 200. See the main text for more details.

Online Appendix for *Media Trust and Persuasion*

Shuhei Kitamura* Toshifumi Kuroda†

July 8, 2021

Abstract: This Online Appendix includes the tables, figures, and additional analyses excluded from the study because of space constraints. The sections are organized as follows. Section A.1 explains the background information. Section A.2 shows the instructions for the experiment. Section A.3 presents the balance tables. Section A.4 describes the selection criteria for each topic. Section A.5 examines the effect of the program recommendations. Section A.6 discusses an attrition issue. Finally, Section A.7 provides examples of the closest words to official statements and examines different cases for formulating the closeness measures.

A.1 Background

This section provides background information for this study.

A.1.1 Market Structure

The television market in Japan has a dual structure that consists of one nationwide public media outlet (henceforth Public TV) and several regional private broadcasters. Public TV has two channels, 1 and 2. In addition, local private channels in each region broadcast the programs of five networks (whose headquarters are all in Tokyo) and/or their own programs. No private broadcaster is allowed to have more than one channel in a region. In addition, the number of private broadcasters differs across Japan. The greater Tokyo area, for example, has five major private channels: 4, 6, 8, 10, and 12. The last entry or exit of private broadcasters in this area occurred in 1995.

Public TV's market share varies by region, as the number of private broadcasters varies (from two to six). In the greater Tokyo area in 2016, Channel 1 had a market share of 18.4% compared

*Osaka University. Email: kitamura@cider.osaka-u.ac.jp.

†Tokyo Keizai University. Email: kuroda@tku.ac.jp.

with Channels 4 (22.5%), 6 (16.0%), 8 (15.5%), 10 (19.8%), and 12 (7.8%) according to Dentsu Media Innovation Lab (2018).

A.1.2 Source of Revenue and Ownership

Public TV is mostly (99% of annual income) funded by the receiving fees paid by households and enterprises that have TV equipment, regardless of whether they actually view the outlet's programs or not. It also receives a relatively small (0.5% of annual income in 2016) government subsidy for providing international and party election broadcasts. The Broadcast Act prohibits any revenue from advertising and states that the Diet must approve the outlet's annual budget.

The 12 members of the Board of Governors, which decides the outlet's management policy and operations, are appointed by the Prime Minister and approved by the Diet.[1] An appointment sometimes comes to public attention. For example, one appointment in 2013 gained media attention because some members were too closely related to Prime Minister Abe (*Nikkei*, 2013). Historically, all members have been Japanese nationals.

By contrast, the main source of revenue for private broadcasters is advertising (82.4% of their annual income in 2016) according to Japan Commercial Broadcasters Association (2017). Foreign nationals can buy shares in a private broadcaster, although the Broadcast Act regulates that foreign ownership with voting rights should not exceed 20%.[2]

A.1.3 Program Contents

Regarding the editing of broadcasting programs, the Broadcast Act rules that the broadcaster "shall not harm public safety or good morals," "shall be politically fair," "shall not distort the facts," and "shall clarify the points at issue from as many angles as possible where there are conflicting opinions concerning an issue" (Ministry of Internal Affairs and Communications, 2010). Although the same rule applies to all broadcasters, the program contents differ for public and private outlets. For example, Krauss (2000) finds that the frequency at which journalists and newscasters express their personal opinions about a news item in private broadcasting programs is twice that in the outlet's programs. In addition, when the newscasters comment on a particular news story, they tend to provide objective and obvious information rather than their personal opinions. According

[1] They are typically business managers, academics, and celebrities. The term of appointment is three years, although it can be renewed.

[2] If we include foreign ownership without voting rights, the share of foreign shareholding can exceed 20%. However, in practice, it does not deviate too much from this threshold.

to an online survey conducted by Newman et al. (2018), Public TV is the most trusted news outlet in Japan.[3] The style of its news programs might be one of the major reasons for this trust.

A.2 Instructions for the Experiment

Figure A.1 shows the original instructions sent by the survey company at the time of recruitment. Figure A.2 presents their translation in English.

A.3 Balance Table for the Main Experiment

Table A.1 shows the balance table of the variables of individual characteristics, party support, media use, media trust, and evaluations of public TV programs. It shows the average value for the treatment and control groups as well as the p-value of the difference. Most of the variables are balanced and those variables that are not balanced are included as the baseline control in the analysis in the study.

A.4 Selection Criteria for Each Topic

The TV programs in our data are categorized into three topics: (a) Prime Minister Shinzō Abe (henceforth PM topic), (b) the relationship between the United States and Japan (henceforth USA topic), and (c) the relationship between Russia and Japan (henceforth Russia topic). This section describes the selection criteria for each topic. We intentionally set relatively general selection criteria to avoid the risk of excluding relevant programs. After selecting the programs, we checked their content to avoid including irrelevant programs.

For (a), we first selected those programs that contain words related to Prime Minister Shinzō Abe. For a program to be categorized as the PM topic, the program should contain "安倍" (Abe) and at least one of the following words: "首相" (Prime Minister), "総理" (Prime Minister), "総裁" (party president), "内閣" (cabinet), "政権" (administration), "外交" (diplomacy), "日本" (Japan), "晋三" (Shinzō), "安倍氏" (Mr. Abe), or "安倍さん" (Mr. Abe) These words tend to co-occur in a sentence by forming a compound word such as "安倍首相" (Prime Minister Abe).[4]

[3] We also find that it is more trusted than private broadcasters in our sample.

[4] Including relatively general compound words such as "安倍氏" (Mr. Abe) and "安倍さん" (Mr. Abe) in the list aims to select programs that contain "Mr. Abe." We found only one program containing "Mr. Abe" that was not the Prime Minister (Abe is a relatively common last name) and excluded that program from the PM topic.

For (b), we first split the words into three themes: America, Pearl Harbor Visit, and World War II. Prime Minister Abe and President Obama visited Pearl Harbor in December 2016. For a program to be categorized as the USA topic, it should contain at least one word from both (i) *and* (ii), from both (i) *and* (iii), or from (iv) according to the following list:[5]

(i) America: "アメリカ" (America), "米国" (America), "米大統領" (American President), "オバマ" (Obama), "日米" (Japan–America).

(ii) World War II: "第二次大戦" (World War II), "第二次世界大戦" (World War II), "太平洋戦争" (The Pacific War), "日本軍" (The Japanese Army), "開戦" (outbreak of war), "終戦" (end of war), "敗戦" (losing a war), "７５年" (75 years).

(iii) Pearl Harbor Visit: "追悼" (mourning), "慰霊" (memorial), "式典" (ceremony), "戦没者" (war dead), "献花" (offering of flowers), "ハワイ" (Hawaii), "オアフ島" (Oahu Island), "ホノルル" (Honolulu), "フォード島" (Ford Island).

(iv) Pearl Harbor Visit: "真珠湾" (Pearl Harbor), "パールハーバー" (Pearl Harbor), "パール・ハーバー" (Pearl Harbor), "アリゾナメモリアル" (Arizona Memorial), "アリゾナ・メモリアル" (Arizona Memorial), "アリゾナ記念館" (Arizona Memorial), "戦艦アリゾナ" (Battleship Arizona), "戦艦ミズーリ" (Battleship Missouri).

Words related to the Pearl Harbor Visit were split into two groups: (iii) a set of words that need an additional word to be categorized as the Pearl Harbor Visit (e.g., "式典" (ceremony) may just mean a ceremony unrelated to the Pearl Harbor Visit) and (iv) a set of words directly related to Pearl Harbor.

For (c), we first split the words into three themes; Russia, the Japan–Russia Summit, and the Northern Territories issue. The Summit took place in Nagato, a city in the Yamaguchi Prefecture, and in Tokyo in December 2016. President Putin landed at the Yamaguchi Ube Airport on December 15. For a program to be categorized as the Russia topic, it should contain at least one word from both (i) *and* (ii), from both (i) *and* (iii), or from (iv) according to the following list:

(i) Russia: "ロシア" (Russia), "露西亜" (Russia), "ソ連" (Soviet), "ソビエト" (Soviet), "露大統領" (Russian President), "プーチン" (Putin), "日ロ" (Japan–Russia), "日露" (Japan–Russia).

[5] We used Google Translate for the translation. If a word was incorrectly translated, we corrected it manually.

(ii) Japan–Russia Summit: “会談” (meeting), “交渉” (negotiation), “首脳” (leaders), “来日” (visiting Japan), “長門” (Nagato), “宇部” (Ube), “山口” (Yamaguchi).

(iii) Northern Territories dispute: “領土問題” (territorial dispute), “四島” (four islands), “４島” (four islands), “二島” (two islands), “２島” (two islands), “返還” (return).

(iv) Northern Territories dispute: “北方領土” (Northern Territories), “北方四島” (Four Northern Islands), “北方４島” (Four Northern Islands), “南クリル” (South Kuril Islands), “色丹” (Shikotan), “択捉” (Etorofu), “歯舞” (Habomai), “平和条約” (Peace Treaty), “日ソ共同宣言” (The Japan–Soviet Joint Declaration), “５６年宣言” (The Declaration of 1956), “５６年の宣言” (The Declaration of 1956).

The words related to the Northern Territories dispute were split into two groups: (iii) a set of words that needed an additional word to be categorized as the Northern Territories dispute (e.g., “四島” (four islands) without “北方” (Northern) could just mean four islands unrelated to the Northern Territories) and (iv) a set of words directly related to the Northern Territories dispute.

A.5 Effect of Program Recommendations

In addition to the main experiment, we sent program recommendations every Friday to examine whether they would alter the individual’s response. This section examines the effect of this second experiment.

We used two sub-experiments to evaluate the role of program recommendations. In the first sub-experiment, the treatment group in the main experiment was randomly divided into two subgroups: a group receiving program recommendations for programs broadcast by Public Channel 1 (henceforth Public 1) (treatment) and a group not receiving them (control). This sub-experiment aimed to capture the effect of program recommendations.

In the second sub-experiment, we divided the sample of this treatment group into two subgroups: a group receiving the program recommendations with additional information on the content of programs (treatment) and a group receiving them without such information (control). The motivation behind this second sub-experiment was to examine the role of relatively rich information on program recommendations.

Figure A.3 shows examples of program recommendations. The recommendation in the top panel was used for the control group in the second sub-experiment, while the recommendation in

the bottom panel was used for the treatment group in the second sub-experiment. It shows that the bottom panel includes an additional row, which briefly explains the content of the programs.

A.5.1 Balance tables

Table A.2 shows the balance table for the first sub-experiment, which is similar to Table A.1, but using only the sample in the treatment group in the main experiment. It shows that most variables are balanced. The unbalanced variables are included as control variables in the following analysis.

By contrast, Table A.3 is the balance table for the second sub-experiment. Again, most of the variables are balanced. We include unbalanced variables as control variables in the following analysis.

A.5.2 Results

Table A.4 shows the effect of program recommendations on watching Public 1; columns (1)–(3) show the results for the first sub-experiment, while columns (4)–(6) show the results for the second sub-experiment. Similar to Table 2 in the study, the first columns for each experiment (columns (1) and (4)) do not include control variables, the second columns (columns (2) and (5)) include them, and the last columns (columns (3) and (6)) drop individuals who occasionally failed to report the viewing records of certain programs, although they were likely to have watched them according to the access log. Overall, the insignificant results in all the columns indicate that the effect of program recommendations are not statistically different from zero.

Because we have access logs for individuals in the treatment group in the main experiment, we also conduct a similar exercise using the log data.

Table A.5 shows the results. The control variables are the same as those in Table A.4. The first two columns examine the effect of program recommendations and recommendations with additional information, respectively on the viewing time of Public 1 using the log data. As expected, the effect is not significantly different from zero.

Columns (3)–(6) in Table A.5 use the dummy variable, which takes the value one if the individual watches Public 1 for at least one minute and zero otherwise, as the dependent variables. The results in columns (3) and (4) indicate that program recommendations increased the likelihood of watching Public 1, according to both the linear probability model (column (3)) and the probit model (column (4)). By contrast, no effect was detected for providing additional information on program advertisements.

Taken together, the effect of program recommendations seems to be null or limited in our sample.

A.6 Attrition

As mentioned in the Data section in the study, 1,524 individuals (1,280 in the treatment group and 244 in the control group) did not complete the endline survey. This section discusses the possibility of attrition bias.

To check the possibility of attrition on the observables, we regress a dummy variable, which takes the value one if an individual does not complete the endline survey, and zero otherwise on the baseline variables in Table A.1. Table A.6 shows the results. First, we find that most variables are statistically insignificant at the 10% level. Most importantly, systematic attrition did not occur with respect to the viewing behavior of Public 1. Second, we find that sex, age, Internet use, and trust in web-based information are relatively strong predictors of attrition. This finding indicates that younger women who use the Internet more often on holidays than on weekdays and who are less likely to trust information from websites (which excludes information whose original source is the TV or a newspaper) tended not to complete the endline survey.

Next, we apply the weighted least squares method (Fitzgerald et al., 1998). Specifically, we first regress the attrition dummy on Internet use variables as well as sex, age, and trust in web-based information using the probit model. Then, we run the weighted least squares method using the inverse of the predicted probability as the weight.

Table A.7 is the same as Table 2 in the study, except that it shows the weighted least squares estimates. Although the estimates are all statistically significant, they are about 40 minutes longer. Thus, if attrition is made on the observables, we can at least say that the main estimates in the study serve as the lower bound of the true effect. Since we cannot check the credibility of this assumption, we use the original unweighted estimates in the main analysis.

A.7 Closeness Measures

A.7.1 Example of Words

This section shows an example of the closest words to official statements used in the main analysis. Table A.8 shows the top 10 words that appear in the corpus at least 10 times and that are the closest to the official statement made by the Japanese and US governments. First, most words are highly

related to Pearl Harbor, military, and war, which validates the use of our closeness measures. However, a few words do not necessarily seem to be related to the topic. This is particularly the case when closeness is measured between the content of private broadcasters' programs and Japanese official statements. This might be because the Japanese government's statement contains more abstract and poetic expressions. Finally, there are some differences between Public 1 and private broadcasters. For example, "パールハーバー" (Pearl Harbor) appears only for private broadcasting.

Similarly, Table A.9 shows the top 10 closest words to the official statement made by the Japanese and Russian governments when President Putin visited Japan for the summit. First, similar to the previous table, most words are highly related to the Northern Territories dispute and the Japan–Russia relationship. The number of seemingly unrelated words is also fewer in this case. Lastly, there are some differences between Public 1 and private broadcasting. For example, "経済連携協定" (Economic Partnership Agreement) appears only for Public 1. A possible explanation could be that the government might have wanted to stress the achievement of the summit (i.e., an agreement with Russia on economic cooperation).

A.7.2 Alternative Vector Size

The analysis in the study uses a vector size of 200. This section analyzes an alternative vector size. Table A.10 shows the difference in average closeness to official statements, similar to Table 6 in the study but with 100, instead of 200, vectors. As before, Public 1's information is closer to the domestic government than to the foreign government compared with private broadcasters' information. Moreover, the difference between Japan and Russia is larger than the difference between Japan and the United States.

Table A.11 shows the first-stage results using these alternative closeness measures. The results are similar to before: treated individuals were more exposed to information provided by the domestic government.

Table A.12 shows the effect on political evaluations using the alternative vector size. Again, the results are qualitatively similar to the results in the study.

A.7.3 Excluding Low Frequency Words

Since there are many low frequency words in the training data, this section examines the case in which these words are dropped from the data. Figure A.4 plots the percentage of TV programs

categorized as either the USA or the Russia topic for those closest to official statements, similar to Figure 6 in the study, but using only words that appear at least 50 times in the corpus. The figures appear to be similar to each other. The machine is better at predicting programs categorized as the Russia topic than programs categorized as the USA topic, which we also observe in Figure 6 in the study.

Table A.13 shows the difference in average closeness to official statements, similar to Table 6 in the study but dropping low frequency words. The difference is still positive and statistically significant.

Table A.14 shows the first-stage results and Table A.15 shows the effect on political evaluations using these alternative measures. Overall, the results are similar to the ones in the study.

It is hard to say which method performs better given these results. Moreover, the definition of "low frequency" can be ad hoc. Therefore, we decided to use the full sample in the main analysis.

A.7.4 Closeness to Party Speeches

In the study, we measure the semantic similarity between TV programs and the official statements made by domestic and foreign governments and find that the positive evaluation of foreign policies is caused by the feature that public media's information is slanted toward the domestic government rather than the foreign government.

Another hypothesis could be that public media's information might be slanted toward the ruling party rather than the opposition party and that treated individuals might have been exposed to such information. The difference between this hypothesis and the hypothesis in the study is that the former reflects the competition between political parties within a country, while the latter is related to national identity. The latter hypothesis does not necessarily rule out the former, however. Further, it might be difficult to distinguish between these hypotheses if information in favor of the domestic government is also emphasized by the ruling party.

To check whether political competition explains our findings, we measure the semantic similarity between TV programs and party speeches made by the major ruling party (Liberal Democratic Party, henceforth LDP) and the major opposition party (Democratic Party, henceforth DP) in this section. To make the results more comparable, we focus on party speeches on the relationship with Russia in the Foreign Affairs Committee meetings in the 192nd Diet (September 26–December 14, 2016).[6] We searched the minutes of the meetings that contain "ロシア" (Russia) using the

[6] Recall that the experiment took place between November 14 and December 18, 2016.

database of the minutes of the Diet.[7] The database has three meetings, namely those held on October 14, October 26, and November 4, from which we obtained the speeches made by the party members of the LDP and DP. Finally, the procedure for measuring semantic similarity is the same as described in the study.

Table A.16 shows the supply-side results. This table is similar to Table 6 in the study but uses the semantic similarity to the party speeches instead of official statements. The table shows that all values are positive and statistically significant, meaning that Public 1's information is closer to the major ruling party than the major opposition party compared with private channels' information.

Next, we construct the average exposure variable as in the study using the alternative closeness variable and regress it on the treatment dummy to check whether the treated group was exposed more to information favoring the ruling party than the opposition party compared with the control group. Interestingly, Table A.17 shows no significant result, which contrasts with the top panel in Table 7 in the study.

Thus, although we do find that public media's information seems to be slanted toward the ruling party compared with private media's information, this cannot explain the increase in the positive evaluation of foreign policies found in the study. Therefore, the positive evaluation of foreign policies seems to be driven by information related to national identity (domestic vs. foreign) rather than political competition (ruling vs. opposition).

References

Dentsu Media Innovation Lab (2018). *Jyouhou Media Hakusyo*. Tokyo: Diamond Inc.

Fitzgerald, J., P. Gottschalk, and R. Moffitt (1998). An Analysis of Sample Attrition in Panel Data: The Michigan Panel Study of Income Dynamics. *The Journal of Human Resources 33*(2), 251–299.

Japan Commercial Broadcasters Association (2017). *Nihon Minkan Housou Nenkan*. Tokyo: Kōken Publishing.

Krauss, E. S. (2000). *Broadcasting Politics in Japan: NHK and Television News*. Ithaca: Cornell University Press.

Ministry of Internal Affairs and Communications (2010). *The Broadcast Act*. `http://www.soumu.go.jp/main_sosiki/joho_tsusin/eng/Resources/laws/pdf/090204_5.pdf`.

[7] http://kokkai.ndl.go.jp/.

Newman, N., R. Fletcher, A. Kalogeropoulos, D. A. L. Levy, and R. K. Nielsen (2018). *Reuters Institute Digital News Report 2018*. Oxford: Reuters Institute for the Study of Journalism.

Nikkei (2013). NHK Keiei Iin Kokkai Doui. `https://www.nikkei.com/article/DGXNASDD080D5_Y3A101C1TJ0000/`.

Olea, J. L. M. and C. Pflueger (2013). A robust test for weak instruments. *Journal of Business & Economic Statistics 31*(3), 358–369.

Tables and Figures

Table A.1: Balance table for the main experiment.

	Control	Treatment	p-value	Observations
Individual characteristics				
Age	47.754	47.303	0.384	4452
	(0.481)	(0.213)		
Female	0.466	0.478	0.554	4452
	(0.018)	(0.008)		
Education	14.752	14.744	0.922	4452
	(0.069)	(0.031)		
Annual income	6.275	6.105	0.333	3926
	(0.170)	(0.072)		
Annual income not reported	0.092	0.124	0.013	4452
	(0.011)	(0.005)		
Origin (prefecture)	12.307	12.310	0.963	4452
	(0.057)	(0.027)		
Occupation	2.851	2.798	0.504	4452
	(0.073)	(0.033)		
Family type	2.777	2.734	0.393	4452
	(0.046)	(0.021)		
Family size	2.839	2.771	0.171	4452
	(0.047)	(0.020)		
Public awareness	2.413	2.395	0.590	4452
	(0.031)	(0.013)		
Policy evaluation	2.307	2.276	0.352	4452
	(0.030)	(0.013)		
Party support				
Liberal-Democrat	0.288	0.243	0.009	4452
	(0.017)	(0.007)		
Democrat	0.049	0.049	0.994	4452
	(0.008)	(0.004)		
Komei	0.019	0.019	0.991	4452
	(0.005)	(0.002)		
Communist	0.031	0.027	0.592	4452
	(0.006)	(0.003)		
Ishin	0.023	0.023	0.946	4452
	(0.005)	(0.002)		
Social-Democrat	0.008	0.005	0.239	4452
	(0.003)	(0.001)		
Other	0.007	0.015	0.067	4452
	(0.003)	(0.002)		
No party	0.490	0.526	0.075	4452
	(0.018)	(0.008)		
No answer	0.086	0.094	0.504	4452
	(0.010)	(0.005)		

Note. Standard errors are in parentheses.

Balance table for the main experiment (cont.).

	Control	Treatment	p-value	Observations
TV view freq. (realtime)				
Public_1	4.643	4.540	0.320	4452
	(0.094)	(0.042)		
Public_2	2.984	2.986	0.985	4452
	(0.077)	(0.035)		
Private_4	5.495	5.421	0.367	4452
	(0.074)	(0.034)		
Private_6	5.337	5.249	0.278	4452
	(0.073)	(0.034)		
Private_8	5.113	5.092	0.805	4452
	(0.076)	(0.035)		
Private_10	5.359	5.333	0.751	4452
	(0.074)	(0.034)		
Private_12	4.883	4.846	0.658	4452
	(0.075)	(0.034)		
TV view freq. (recorded)				
Public_1	2.615	2.657	0.615	4452
	(0.077)	(0.034)		
Public_2	2.013	2.114	0.148	4452
	(0.061)	(0.029)		
Private_4	3.323	3.442	0.177	4452
	(0.080)	(0.036)		
Private_6	3.345	3.354	0.924	4452
	(0.080)	(0.036)		
Private_8	3.216	3.273	0.513	4452
	(0.079)	(0.036)		
Private_10	3.352	3.405	0.546	4452
	(0.080)	(0.036)		
Private_12	3.076	3.154	0.358	4452
	(0.077)	(0.035)		
Internet use				
Weekdays (duration)	4.574	4.507	0.403	4434
	(0.072)	(0.033)		
Holidays (duration)	4.893	4.913	0.814	4429
	(0.076)	(0.035)		
On mobile phones (freq.)	4.081	4.058	0.804	4452
	(0.083)	(0.037)		

Note. Standard errors are in parentheses.

Balance table for the main experiment (cont.).

	Control	Treatment	p-value	Observations
Reading newspaper				
Yomiuri	0.248	0.245	0.868	4452
	(0.016)	(0.007)		
Asahi	0.205	0.203	0.926	4452
	(0.015)	(0.007)		
Mainichi	0.040	0.037	0.712	4452
	(0.007)	(0.003)		
Sankei	0.049	0.036	0.093	4452
	(0.008)	(0.003)		
Nikkei	0.151	0.145	0.660	4452
	(0.013)	(0.006)		
Sport	0.037	0.031	0.365	4452
	(0.007)	(0.003)		
Local	0.052	0.054	0.778	4452
	(0.008)	(0.004)		
Other	0.009	0.011	0.665	4452
	(0.003)	(0.002)		
Not reading	0.384	0.380	0.860	4452
	(0.018)	(0.008)		
Media trust				
Public TV	3.048	2.984	0.030	4452
	(0.026)	(0.012)		
Private channels	2.829	2.796	0.227	4452
	(0.025)	(0.011)		
Newspaper	2.956	2.908	0.081	4452
	(0.025)	(0.011)		
Publication	2.580	2.556	0.370	4452
	(0.025)	(0.011)		
Internet	2.469	2.424	0.114	4452
	(0.026)	(0.012)		
Evaluation on Public TV programs				
Accuracy	2.822	2.749	0.041	4452
	(0.032)	(0.015)		
Fairness	2.493	2.447	0.188	4452
	(0.031)	(0.014)		
Observations	753	3699		4452

Note. Standard errors are in parentheses.

Table A.2: Balance table for the second treatment (ad.).

	Control	Treatment	p-value	Observations
Individual characteristics				
Age	47.525	47.248	0.605	3699
	(0.476)	(0.238)		
Female	0.475	0.479	0.877	3699
	(0.018)	(0.009)		
Education	14.817	14.726	0.245	3699
	(0.069)	(0.035)		
Annual income	6.102	6.106	0.983	3242
	(0.161)	(0.081)		
Annual income not reported	0.115	0.126	0.420	3699
	(0.012)	(0.006)		
Origin (prefecture)	12.337	12.303	0.609	3699
	(0.058)	(0.030)		
Occupation	2.758	2.808	0.550	3699
	(0.073)	(0.037)		
Family type	2.765	2.726	0.455	3699
	(0.047)	(0.023)		
Family size	2.805	2.763	0.415	3699
	(0.046)	(0.023)		
Public awareness	2.377	2.400	0.503	3699
	(0.030)	(0.015)		
Policy evaluation	2.235	2.286	0.127	3699
	(0.030)	(0.015)		
Party support				
Liberal-Democrat	0.219	0.249	0.088	3699
	(0.015)	(0.008)		
Democrat	0.056	0.048	0.342	3699
	(0.009)	(0.004)		
Komei	0.015	0.020	0.418	3699
	(0.004)	(0.003)		
Communist	0.034	0.025	0.185	3699
	(0.007)	(0.003)		
Ishin	0.018	0.024	0.293	3699
	(0.005)	(0.003)		
Social-Democrat	0.001	0.005	0.149	3699
	(0.001)	(0.001)		
Other	0.014	0.016	0.715	3699
	(0.004)	(0.002)		
No party	0.545	0.521	0.237	3699
	(0.018)	(0.009)		
No answer	0.098	0.093	0.658	3699
	(0.011)	(0.005)		

Note. Standard errors are in parentheses.

Balance table for the second treatment (ad.) (cont.).

	Control	Treatment	p-value	Observations
TV view freq. (realtime)				
Public_1	4.493	4.552	0.580	3699
	(0.095)	(0.047)		
Public_2	2.959	2.992	0.707	3699
	(0.078)	(0.039)		
Private_4	5.437	5.417	0.815	3699
	(0.076)	(0.038)		
Private_6	5.249	5.249	0.996	3699
	(0.076)	(0.037)		
Private_8	5.117	5.086	0.717	3699
	(0.079)	(0.039)		
Private_10	5.343	5.330	0.879	3699
	(0.077)	(0.038)		
Private_12	4.825	4.852	0.757	3699
	(0.079)	(0.038)		
TV view freq. (recorded)				
Public_1	2.668	2.654	0.869	3699
	(0.077)	(0.038)		
Public_2	2.066	2.125	0.406	3699
	(0.062)	(0.032)		
Private_4	3.436	3.444	0.930	3699
	(0.082)	(0.041)		
Private_6	3.320	3.362	0.640	3699
	(0.081)	(0.040)		
Private_8	3.279	3.272	0.941	3699
	(0.080)	(0.040)		
Private_10	3.407	3.404	0.977	3699
	(0.081)	(0.040)		
Private_12	3.143	3.157	0.881	3699
	(0.080)	(0.039)		
Internet use				
Weekdays (duration)	4.447	4.522	0.363	3687
	(0.071)	(0.037)		
Holidays (duration)	4.833	4.932	0.257	3684
	(0.074)	(0.040)		
On mobile phones (freq.)	4.060	4.058	0.982	3699
	(0.085)	(0.042)		

Note. Standard errors are in parentheses.

Balance table for the second treatment (ad.) (cont.).

	Control	Treatment	p-value	Observations
Reading newspaper				
Yomiuri	0.240	0.247	0.724	3699
	(0.016)	(0.008)		
Asahi	0.209	0.202	0.653	3699
	(0.015)	(0.007)		
Mainichi	0.045	0.035	0.198	3699
	(0.008)	(0.003)		
Sankei	0.046	0.034	0.098	3699
	(0.008)	(0.003)		
Nikkei	0.139	0.147	0.617	3699
	(0.013)	(0.006)		
Sport	0.034	0.030	0.560	3699
	(0.007)	(0.003)		
Local	0.059	0.053	0.557	3699
	(0.009)	(0.004)		
Other	0.014	0.010	0.457	3699
	(0.004)	(0.002)		
Not reading	0.380	0.381	0.971	3699
	(0.018)	(0.009)		
Media trust				
Public TV	2.970	2.988	0.555	3699
	(0.027)	(0.013)		
Private channels	2.775	2.801	0.332	3699
	(0.026)	(0.012)		
Newspaper	2.889	2.912	0.421	3699
	(0.026)	(0.013)		
Publication	2.512	2.566	0.058	3699
	(0.026)	(0.013)		
Internet	2.404	2.429	0.396	3699
	(0.026)	(0.013)		
Evaluation on Public TV programs				
Accuracy	2.708	2.759	0.171	3699
	(0.033)	(0.017)		
Fairness	2.414	2.455	0.255	3699
	(0.032)	(0.016)		
Observations	732	2967		3699

Note. Standard errors are in parentheses.

Table A.3: Balance table for the second treatment (ad. + info.).

	Control	Treatment	p-value	Observations
Individual characteristics				
Age	47.010	47.486	0.317	2967
	(0.335)	(0.337)		
Female	0.479	0.478	0.960	2967
	(0.013)	(0.013)		
Education	14.724	14.728	0.954	2967
	(0.049)	(0.049)		
Annual income	6.189	6.022	0.303	2594
	(0.113)	(0.116)		
Annual income not reported	0.119	0.132	0.265	2967
	(0.008)	(0.009)		
Origin (prefecture)	12.280	12.325	0.455	2967
	(0.044)	(0.041)		
Occupation	2.792	2.823	0.672	2967
	(0.052)	(0.052)		
Family type	2.721	2.731	0.829	2967
	(0.033)	(0.033)		
Family size	2.768	2.758	0.819	2967
	(0.032)	(0.032)		
Public awareness	2.390	2.410	0.513	2967
	(0.021)	(0.021)		
Policy evaluation	2.270	2.303	0.280	2967
	(0.021)	(0.021)		
Party support				
Liberal-Democrat	0.247	0.250	0.857	2967
	(0.011)	(0.011)		
Democrat	0.047	0.048	0.954	2967
	(0.006)	(0.006)		
Komei	0.022	0.017	0.416	2967
	(0.004)	(0.003)		
Communist	0.026	0.025	0.891	2967
	(0.004)	(0.004)		
Ishin	0.022	0.027	0.350	2967
	(0.004)	(0.004)		
Social-Democrat	0.007	0.003	0.130	2967
	(0.002)	(0.002)		
Other	0.017	0.014	0.542	2967
	(0.003)	(0.003)		
No party	0.515	0.527	0.524	2967
	(0.013)	(0.013)		
No answer	0.097	0.089	0.424	2967
	(0.008)	(0.007)		

Note. Standard errors are in parentheses.

Balance table for the second treatment (ad. + info.) (cont.).

	Control	Treatment	p-value	Observations
TV view freq. (realtime)				
Public_1	4.534	4.570	0.711	2967
	(0.067)	(0.067)		
Public_2	2.997	2.987	0.898	2967
	(0.056)	(0.056)		
Private_4	5.423	5.412	0.880	2967
	(0.053)	(0.053)		
Private_6	5.251	5.247	0.966	2967
	(0.053)	(0.053)		
Private_8	5.128	5.043	0.275	2967
	(0.055)	(0.056)		
Private_10	5.361	5.299	0.413	2967
	(0.052)	(0.054)		
Private_12	4.843	4.861	0.810	2967
	(0.053)	(0.054)		
TV view freq. (recorded)				
Public_1	2.626	2.682	0.460	2967
	(0.054)	(0.054)		
Public_2	2.082	2.169	0.178	2967
	(0.045)	(0.046)		
Private_4	3.384	3.504	0.141	2967
	(0.057)	(0.058)		
Private_6	3.337	3.387	0.538	2967
	(0.056)	(0.057)		
Private_8	3.243	3.301	0.473	2967
	(0.056)	(0.057)		
Private_10	3.324	3.484	0.048	2967
	(0.057)	(0.057)		
Private_12	3.098	3.215	0.135	2967
	(0.055)	(0.056)		
Internet use				
Weekdays (duration)	4.541	4.504	0.615	2956
	(0.053)	(0.052)		
Holidays (duration)	4.916	4.949	0.679	2954
	(0.056)	(0.056)		
On mobile phones (freq.)	4.046	4.070	0.774	2967
	(0.059)	(0.059)		

Note. Standard errors are in parentheses.

Balance table for the second treatment (ad. + info.) (cont.).

	Control	Treatment	p-value	Observations
Reading newspaper				
Yomiuri	0.243	0.251	0.601	2967
	(0.011)	(0.011)		
Asahi	0.191	0.213	0.136	2967
	(0.010)	(0.011)		
Mainichi	0.036	0.034	0.823	2967
	(0.005)	(0.005)		
Sankei	0.032	0.035	0.702	2967
	(0.005)	(0.005)		
Nikkei	0.143	0.150	0.605	2967
	(0.009)	(0.009)		
Sport	0.023	0.037	0.025	2967
	(0.004)	(0.005)		
Local	0.054	0.052	0.846	2967
	(0.006)	(0.006)		
Other	0.010	0.011	0.867	2967
	(0.003)	(0.003)		
Not reading	0.392	0.369	0.203	2967
	(0.013)	(0.013)		
Media trust				
Public TV	2.973	3.003	0.271	2967
	(0.019)	(0.019)		
Private channels	2.811	2.792	0.446	2967
	(0.017)	(0.017)		
Newspaper	2.918	2.907	0.682	2967
	(0.018)	(0.018)		
Publication	2.570	2.563	0.791	2967
	(0.018)	(0.018)		
Internet	2.434	2.424	0.676	2967
	(0.018)	(0.018)		
Evaluation on Public TV programs				
Accuracy	2.735	2.782	0.156	2967
	(0.024)	(0.023)		
Fairness	2.431	2.479	0.137	2967
	(0.023)	(0.023)		
Observations	1480	1487		2967

Note. Standard errors are in parentheses.

Table A.4: The average treatment effect of advertisements on watching Public 1.

	Dependent variable: Δ Public_1					
	(1)	(2)	(3)	(4)	(5)	(6)
Ad.	0.396	0.407	0.501			
	(0.364)	(0.363)	(0.379)			
Ad. + info.				0.127	0.123	0.027
				(0.323)	(0.323)	(0.346)
Controls	No	Yes	Yes	No	Yes	Yes
Mean of control	0.76	0.76	0.88	1.10	1.10	1.35
R^2	0.00	0.00	0.00	0.00	0.00	0.00
Observations	3699	3699	3403	2967	2967	2704

Note. Robust standard errors are in parentheses. The dependent variables are the difference in the viewing time of Public 1 before and during the experiment (in hours). "Ad." and "Ad. + info." are the treatment dummies for the first and second sub-experiments, respectively. See the text for more details. The control variables for columns (1)–(3) are a dummy for supporting the LDP and for reading the Sankei newspaper. The control variables for columns (4)–(6) are the frequency of watching Channel 10 and a dummy for reading sports newspapers. Columns (3) and (6) drop individuals who occasionally failed to report the viewing records of certain programs, although they were likely to have seen them according to the access log. *, **, and *** indicate $p < 0.10$, $p < 0.05$, and $p < 0.01$, respectively.

Table A.5: The average treatment effect of advertisements on watching Public 1 using log data.

	Dependent variable: TV view					
	Minutes		Dummy			
	(1) OLS	(2) OLS	(3) OLS	(4) Probit	(5) OLS	(6) Probit
Ad.	−0.200		0.031	0.440		
	(0.812)		(0.007)***	(0.122)***		
Ad. + info.		0.488			−0.009	−0.114
		(0.524)			(0.008)	(0.085)
Controls	Yes	No	Yes	Yes	No	No
Mean of control	0.93	0.55	0.02	0.04	0.02	0.04
R^2	0.02	0.02	0.07		0.07	
Number of obs.	3403	2704	3403	3403	2704	2704

Note. Robust standard errors are in parentheses. The dependent variables for columns (1) and (2) are the viewing time of Public 1 (in minutes) and for columns (3)–(6) are the dummy, which takes the value one if an individual watched at least a minute of Public 1 during the experiment. *Ad.* and *Ad. + info.* are the treatment dummies for the first and second sub-experiments, respectively. See the text for more details. The control variables for columns (1), (3), and (4) are a dummy for supporting the LDP and for reading the Sankei newspaper. The control variables for columns (2), (5), and (6) are the frequency of watching Channel 10 and a dummy for reading sports newspapers. *, **, and *** indicate p <0.10, p <0.05, and p <0.01, respectively.

Table A.6: Probit estimates of the attrition dummy on the covariates.

		Coef.	S.E.
Individual characteristics	Age	−0.016	(0.002)***
	Female	0.199	(0.044)***
	Education	−0.014	(0.011)
	Annual income	−0.006	(0.006)
	Origin (Prefecture)	−0.014	(0.012)
	Occupation	0.022	(0.011)*
	Family type	0.003	(0.021)
	Family size	−0.004	(0.022)
	Public awareness	0.084	(0.025)***
	Policy evaluation	−0.037	(0.029)
Party support	Liberal-Democrat	0.304	(0.185)
	Democrat	0.118	(0.204)
	Komei	0.381	(0.224)*
	Communist	0.243	(0.213)
	Ishin	0.015	(0.228)
	Social-Democrat	−0.275	(0.388)
	No party	0.059	(0.184)
	No answer	0.243	(0.192)
TV view freq. (realtime)	Public_1	−0.011	(0.012)
	Public_2	0.024	(0.013)*
	Private_4	−0.014	(0.021)
	Private_6	0.044	(0.025)*
	Private_8	0.004	(0.019)
	Private_10	−0.009	(0.022)
	Private_12	−0.012	(0.017)
TV view freq. (recorded)	Public_1	−0.006	(0.014)
	Public_2	−0.018	(0.017)
	Private_4	−0.012	(0.022)
	Private_6	−0.019	(0.026)
	Private_8	−0.001	(0.024)
	Private_10	0.014	(0.022)
	Private_12	0.023	(0.017)
Internet use	Weekdays (duration)	−0.069	(0.016)***
	Holidays (duration)	0.053	(0.015)***
	On mobile phones (freq.)	0.053	(0.010)***
Reading newspaper	Yomiuri	−0.044	(0.065)
	Asahi	−0.017	(0.067)
	Mainichi	−0.065	(0.112)
	Sankei	−0.033	(0.108)
	Nikkei	0.047	(0.068)
	Sport	0.058	(0.114)
	Local	0.162	(0.095)*
	Other	−0.197	(0.206)
	Not reading	0.081	(0.071)
Media trust	Public TV	0.048	(0.042)
	Private channels	−0.068	(0.046)
	Newspaper	0.068	(0.044)
	Publication	−0.026	(0.040)
	Internet	−0.084	(0.035)**
Evaluation on Public TV programs	Accuracy	0.001	(0.033)
	Fairness	−0.021	(0.034)
	Constant	−0.031	(0.330)
	Observations	4983	

Note. Standard errors are in parentheses. A dummy for not reporting annual income and for supporting another party were dropped because of collinearity. *, **, and *** indicate $p < 0.10$, $p < 0.05$, and $p < 0.01$, respectively.

Table A.7: The average treatment effect on watching Public 1 (weighted least squares).

	Dependent variable: Δ Public_1		
	(1)	(2)	(3)
Treatment	1.397	1.459	1.459
	(0.452)***	(0.451)***	(0.451)***
Controls	No	Yes	Yes
Mean of control	0.27	0.27	0.27
R^2	0.00	0.01	0.01
Observations	4132	4132	4132

Note. Robust standard errors are in parentheses. The dependent variables are the difference in the viewing time of Public 1 before and during the experiment (in hours). "Treatment" is the treatment dummy. The control variables are a dummy for supporting the LDP, a minor party, and no party, a dummy for not reporting annual income and for reading the Sankei newspaper, and trust and accuracy measures. Column (3) drops individuals who occasionally failed to report the viewing records of certain programs, although they were likely to have seen them according to the access log. The weight is the inverse of the predicted probability from regressing the attrition dummy on the Internet use variables, sex, age, and trust in web-based information using the probit model. *, **, and *** indicate $p < 0.10$, $p < 0.05$, and $p < 0.01$, respectively.

Table A.8: Closeness and frequency of the 10 closest words to the Japanese and US governments' official statements.

Japanese government			US government		
Word	Count	Closeness	Word	Count	Closeness
Public Channel 1					
戦没者 (war dead)	12	0.5712	旧日本軍 (Imperial Japanese Army)	42	0.5804
真珠湾 (Pearl Harbor)	98	0.5664	真珠湾攻撃 (Attack on Pearl Harbor)	72	0.5739
真珠湾攻撃 (Attack on Pearl Harbor)	72	0.5630	最高司令官 (HQ Commander)	11	0.5587
サンプソン (Sampson)	39	0.5475	戦没者 (war dead)	12	0.5429
旧日本軍 (Imperial Japanese Army)	42	0.5470	真珠湾 (Pearl Harbor)	98	0.5287
最高司令官 (HQ Commander)	11	0.5369	出撃 (sally)	20	0.5239
哀悼 (condolence)	13	0.5313	戦艦 (warship)	37	0.5139
はなめがね (pince-nez)	15	0.5298	侵攻 (invasion)	13	0.5101
国家評議会議長 (Chairman of State Council)	27	0.5273	日本軍 (Japanese Army)	14	0.5098
佐世保高専 (Sasebo Kōsen)	32	0.5267	軍人 (soldier)	21	0.5076
Private channels					
パールハーバー (Pearl Harbor)	20	0.5916	旧日本軍 (Imperial Japanese Army)	34	0.5804
アリゾナ記念館 (Arizona Memorial Hall)	16	0.5904	真珠湾攻撃 (Attack on Pearl Harbor)	82	0.5739
惨禍 (calamity)	13	0.5663	パールハーバー (Pearl Harbor)	20	0.5469
ゴルフ日本シリーズ (Golf Nippon Series)	18	0.5630	アリゾナ記念館 (Arizona Memorial Hall)	16	0.5429
戦没者 (war dead)	13	0.5517	惨禍 (calamity)	13	0.5372
真珠湾 (Pearl Harbor)	246	0.5476	沿岸警備隊 (Coast Guard)	11	0.5287
真珠湾攻撃 (Attack on Pearl Harbor)	82	0.5475	艦船 (warship)	17	0.5164
ゲレロ (Guerrero)	11	0.5472	戦没者 (war dead)	13	0.5139
ハバロフスク (Khabarovsk)	16	0.5463	海兵隊 (Marines)	59	0.5101
マオリ (Maori)	35	0.5451	真珠湾 (Pearl Harbor)	246	0.5098

Table A.9: Closeness and frequency of the 10 closest words to the Japanese and Russian governments' official statements.

Japanese government			Russian government		
Word	Count	Closeness	Word	Count	Closeness
Public Channel 1					
日ロ (Japan-Russia)	232	0.6430	日ロ (Japan-Russia)	232	0.5733
北方領土問題 (Northern Territories dispute)	119	0.6130	四島 (four islands)	21	0.5610
北方四島 (Northern four islands)	132	0.6070	北方四島 (Northern four islands)	132	0.5469
平和条約 (Peace Treaty)	272	0.5996	締結 (conclusion (of agreements))	156	0.5459
経済連携協定 (Economic Partnership Agreement)	18	0.5958	北方領土問題 (Northern Territories dispute)	119	0.5430
領土問題 (territorial dispute)	100	0.5869	経済連携協定 (Economic Partnership Agreement)	18	0.5386
締結 (conclusion (of agreements))	156	0.5722	加盟国 (member states)	33	0.5370
日米同盟 (Japan-US alliance)	68	0.5694	北方領土 (Northern Territories)	339	0.5329
四島 (four islands)	21	0.5694	同盟国 (allies)	28	0.5309
北方領土 (Northern Territories)	339	0.5641	国際機関 (international organizations)	12	0.5301
Private channels					
二国間 (between two countries)	11	0.6439	二国間 (between two countries)	11	0.5905
日ロ (Japan-Russia)	198	0.6430	両国間 (between both countries)	14	0.5828
北方領土問題 (Northern Territories dispute)	368	0.6130	日ロ (Japan-Russia)	198	0.5733
北方四島 (Northern four islands)	187	0.6070	四島 (four islands)	100	0.5610
安保条約 (Security Treaty)	12	0.5996	北方四島 (Northern four islands)	187	0.5469
平和条約 (Peace Treaty)	560	0.5958	安保条約 (Security Treaty)	12	0.5459
両国間 (between both countries)	14	0.5869	クリミア半島 (Crimean Peninsula)	14	0.5430
領土問題 (territorial dispute)	468	0.5739	締結 (conclusion (of agreements))	336	0.5386
締結 (conclusion (of agreements))	336	0.5722	北方領土問題 (Northern Territories dispute)	368	0.5370
日露 (Japan-Russia)	192	0.5694	日露 (Japan-Russia)	192	0.5329

Table A.10: Difference in closeness (100 vectors).

	Public_1–Private TVs	p-value
Interacted with duration		
Japan–USA	0.812	0.000
Japan–Russia	2.151	0.000
Without interaction		
Japan–USA	0.033	0.000
Japan–Russia	0.049	0.000

Note. For each program, the difference in closeness to either government (domestic or foreign) is computed. Then, their averages are compared between the two types of broadcasters (Public 1 or (the average of) private channels). The difference and associated p-value are reported. The top panel takes into account the duration of the program, while the bottom panel does not. We use machine learning to compute the closeness measures, where the window size is set to 5 and the vector size is set to 100. See the main text for more details.

Table A.11: Effect on the difference in closeness to the official statements (100 vectors).

	Dependent variable: Average Exposure					
	Δ Japan–USA			Δ Japan–Russia		
	(1) All	(2) PM	(3) Russia	(4) All	(5) PM	(6) Russia
Treatment	0.042	0.026	0.013	0.054	0.041	0.025
	(0.018)**	(0.009)***	(0.006)**	(0.021)**	(0.015)***	(0.010)**
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Mean of control	0.47	0.17	0.10	0.45	0.28	0.19
R^2	0.75	0.67	0.60	0.69	0.65	0.62
Observations	4156	4156	4156	4156	4156	4156

Note. Robust standard errors are in parentheses. The dependent variables are the difference in closeness to the official statements made by the domestic and foreign governments. We use machine learning to compute the closeness measures, where the window size is set to 5 and the vector size is set to 100. The first three columns use the difference in average closeness to the official statements made by the Japanese and US governments during Prime Minister Abe's visit to Pearl Harbor with former President Obama. The next three columns use the difference in average closeness to the official statements made by the Japanese and Russian governments during President Vladimir Putin's visit to Japan for the summit. Columns (1) and (4) include all programs. Columns (2) and (5) restrict the programs to those related to Prime Minister Abe, while columns (3) and (6) restrict them to those related to Russia. "Treatment" is the treatment dummy. The control variables are a dummy for supporting the LDP, a minor party, and no party, a dummy for not reporting annual income and for reading the Sankei newspaper, trust and accuracy measures, and the viewing time of Public 1 before the experiment. *, **, and *** indicate $p < 0.10$, $p < 0.05$, and $p < 0.01$, respectively.

Table A.12: Effect on political evaluations of the difference in closeness to the official statements (100 vectors).

	Dependent variable: Evaluation (North. Terr.)					
	(1) All	(2) PM	(3) Russia	(4) All	(5) PM	(6) Russia
OLS Estimates						
Δ Japan-USA	0.178	0.264	0.691			
	(0.080)**	(0.166)	(0.265)***			
Δ Japan-Russia				0.046	0.124	0.274
				(0.072)	(0.099)	(0.142)*
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Mean of control	4.41	4.41	4.41	4.41	4.41	4.41
R^2	0.12	0.12	0.12	0.12	0.12	0.12
Observations	4156	4156	4156	4156	4156	4156
IV Estimates						
Δ Japan-USA	6.403	10.234	20.459			
	(3.537)*	(4.887)**	(11.118)*			
Δ Japan-Russia				4.943	6.636	10.973
				(2.541)*	(3.280)**	(5.890)*
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Mean of control	4.41	4.41	4.41	4.41	4.41	4.41
F-stat.	5.18	8.70	5.35	6.44	7.55	5.58
Effect. F-stat.	5.18	8.70	5.35	6.44	7.55	5.58
p-value (AR)	0.00	0.00	0.00	0.00	0.00	0.00
CI (AR)	[1.782, 20.267]	[3.074, 29.391]	[5.931, 64.043]	[1.421, 14.906]	[2.091, 19.491]	[3.277, 34.060]
CI (Wald)	[-.529, 13.335]	[.655, 19.812]	[-1.333, 42.251]	[-.038, 9.924]	[.208, 13.064]	[-.571, 22.517]
Observations	4156	4156	4156	4156	4156	4156

Note. Robust standard errors are in parentheses. The dependent variables are the evaluation of the government's measures against the Northern Territories dispute. The main independent variables are the difference in closeness to the official statements made by the domestic and foreign governments. We use machine learning to measure closeness, where the window size is set to 5 and the vector size is set to 100. "Δ Japan–USA" is the difference in average closeness between TV programs and the official statements made by the Japanese and US governments during Prime Minister Abe's visit to Pearl Harbor with former President Obama. "Δ Japan–Russia" is the difference in average closeness to the official statements made by the Japanese and Russian governments during President Putin's visit to Japan for the summit. Columns (1) and (4) include all programs. Columns (2) and (5) restrict the programs to those related to Prime Minister Abe, while columns (3) and (6) restrict them to those related to Russia. The instrument is the treatment dummy. The control variables are a dummy for supporting the LDP, a minor party, and no party, a dummy for not reporting annual income and for reading the Sankei newspaper, trust and accuracy measures, and the viewing time of Public 1 before the experiment. Effective F-stat. is the effective F-statistic of Olea and Pflueger (2013). We use Stata's `weakivtest` to compute the test statistic. Weak instrument robust inference (Anderson-Rubin (AR) test) is conducted and the corresponding p-value and confidence interval are reported in the rows labeled "p-value (AR)" and "CI (AR)," respectively. We use Stata's command `rivtest` to compute the test statistic. *, **, and *** indicate $p<0.10$, $p<0.05$, and $p<0.01$, respectively.

Table A.13: Difference in closeness (drop low frequency words).

	Public_1–Private TVs	p-value
Interacted with duration		
Japan–USA	0.339	0.001
Japan–Russia	1.431	0.000
Without interaction		
Japan–USA	0.015	0.000
Japan–Russia	0.037	0.000

Note. For each program, the difference in closeness to either government (domestic or foreign) is computed. Then, their averages are compared between the two types of broadcasters (Public 1 or (the average of) private channels). The difference and associated p-value are reported. The top panel takes into account the duration of the program, while the bottom panel does not. We use machine learning to compute the closeness measures, where the window size is set to 5 and the vector size is set to 200. Only words with a minimum 50 counts are used in the data training process. See the main text for more details.

Table A.14: Effect on the difference in closeness to the official statements (drop low frequency words).

	Dependent variable: Average Exposure					
	Δ Japan–USA			Δ Japan–Russia		
	(1) All	(2) PM	(3) Russia	(4) All	(5) PM	(6) Russia
Treatment	0.031	0.019	0.009	0.035	0.025	0.014
	(0.010)***	(0.007)***	(0.004)**	(0.014)**	(0.010)**	(0.007)**
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Mean of control	0.25	0.11	0.05	0.36	0.18	0.12
R^2	0.72	0.58	0.54	0.75	0.64	0.63
Observations	4156	4156	4156	4156	4156	4156

Note. Robust standard errors are in parentheses. The dependent variables are the difference in closeness to the official statements made by the domestic and foreign governments. We use machine learning to compute the closeness measures, where the window size is set to 5 and the vector size is set to 200. Only words with a minimum 50 counts are used in the data training process. The first three columns use the difference in average closeness to the official statements made by the Japanese and US governments during Prime Minister Abe's visit to Pearl Harbor with former President Obama. The next three columns use the difference in average closeness to the official statements made by the Japanese and Russian governments during President Vladimir Putin's visit to Japan for the summit. Columns (1) and (4) include all programs. Columns (2) and (5) restrict the programs to those related to PM, while columns (3) and (6) restrict them to those related to Russia. "Treatment" is the treatment dummy. The control variables are a dummy for supporting the LDP, a minor party, and no party, a dummy for not reporting annual income and for reading the Sankei newspaper, trust and accuracy measures, and the viewing time of Public 1 before the experiment. *, **, and *** indicate $p < 0.10$, $p < 0.05$, and $p < 0.01$, respectively.

Table A.15: Effect on political evaluations of the difference in closeness to the official statements (drop low frequency words).

	Dependent variable: Evaluation (North. Terr.)					
	(1) All	(2) PM	(3) Russia	(4) All	(5) PM	(6) Russia
OLS Estimates						
Δ Japan-USA	0.419	0.544	1.380			
	(0.136)***	(0.217)**	(0.395)***			
Δ Japan-Russia				0.191	0.225	0.390
				(0.101)*	(0.146)	(0.219)*
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Mean of control	4.41	4.41	4.41	4.41	4.41	4.41
R^2	0.12	0.12	0.12	0.12	0.12	0.12
Observations	4156	4156	4156	4156	4156	4156
IV Estimates						
Δ Japan-USA	8.612	13.830	29.458			
	(4.061)**	(6.581)**	(15.092)*			
Δ Japan-Russia				7.598	10.930	19.275
				(3.944)*	(5.718)*	(11.248)*
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Mean of control	4.41	4.41	4.41	4.41	4.41	4.41
F-stat.	9.04	8.63	6.35	6.01	6.13	4.33
Effect. F-stat.	9.04	8.63	6.35	6.01	6.13	4.33
p-value (AR)	0.00	0.00	0.00	0.00	0.00	0.00
CI (AR)	[2.662, 24.530]	[4.189, 39.625]	[8.543, 88.617]	[2.444, 23.060]	[3.458, 33.345]	[5.469, 63.365]
CI (Wald)	[.652, 16.571]	[.932, 26.727]	[-.122, 59.038]	[-.133, 15.329]	[-.277, 22.137]	[-2.770, 41.320]
Observations	4156	4156	4156	4156	4156	4156

Note. Robust standard errors are in parentheses. The dependent variables are the evaluation of the government's measures against the Northern Territories dispute. The main independent variables are the difference in closeness to the official statements made by the domestic and foreign governments. We use machine learning to compute the closeness measures, where the window size is set to 5 and the vector size is set to 200. Only words with a minimum of 50 counts are used in the data training process. "Δ Japan–USA" is the difference in average closeness between TV programs and the official statements made by the Japanese and US governments during Prime Minister Abe's visit to Pearl Harbor with former President Obama. "Δ Japan–Russia" is the difference in average closeness to the official statements made by the Japanese and Russian governments during President Putin's visit to Japan for the summit. Columns (1) and (4) include all programs. Columns (2) and (5) restrict the programs to those related to PM, while columns (3) and (6) restrict to those related to Russia. The instrument is the treatment dummy. The control variables are a dummy for supporting the LDP, a minor party, and no party, a dummy for not reporting annual income and for reading the Sankei newspaper, trust and accuracy measures, and the viewing time of Public 1 before the experiment. Effective F-stat. is the effective F-statistic of Olea and Pflueger (2013). We use Stata's `weakivtest` to compute the test statistic. Weak instrument robust inference (AR test) is conducted and the corresponding p-value and confidence interval are reported in the rows labeled "p-value (AR)" and "CI (AR)," respectively. We use Stata's command `rivtest` to compute the test statistic. *, **, and *** indicate $p<0.10$, $p<0.05$, and $p<0.01$, respectively.

Table A.16: Difference in closeness (party speeches).

	Public_1–Private TVs	p-value
Interacted with duration		
Ruling–Opposition	0.932	0.000
Ruling–Opposition (drop low freq.)	1.179	0.000
Ruling–Opposition (100 vectors)	1.689	0.000
Without interaction		
Ruling–Opposition	0.011	0.000
Ruling–Opposition (drop low freq.)	0.010	0.000
Ruling–Opposition (100 vectors)	0.029	0.000

Note. For each program, the difference in closeness to either party (LDP (ruling party) or DP (opposition party)) is computed. Then, their averages are compared between the two types of broadcasters (Public 1 or (the average of) private channels). The difference and associated p-value are reported. The top panel takes into account the duration of the program, while the bottom panel does not. We use machine learning to compute the closeness measures, where the window size is set to 5. For each panel, the first row uses 200 vectors, the second row drops low frequency words, and the third row uses 100 vectors. See the main text for more details.

Table A.17: Effect on the difference in closeness to party speeches.

	Dependent variable: Average Exposure		
	(1) All	(2) PM	(3) Russia
Treatment	−0.006	0.000	0.005
	(0.011)	(0.004)	(0.004)
Controls	Yes	Yes	Yes
Mean of control	-0.02	0.03	0.04
R^2	0.05	0.18	0.40
Observations	4156	4156	4156

Note. Robust standard errors are in parentheses. The dependent variables are the difference in closeness to the party speeches made by the ruling and opposition parties. We use machine learning to compute the closeness measures, where the window size is set to 5 and the vector size is set to 200. Column (1) includes all programs. Column (2) restricts the programs to those related to PM, while column (3) restricts them to those related to Russia. "Treatment" is the treatment dummy. The control variables are a dummy for supporting the LDP, a minor party, and no party, a dummy for not reporting annual income and for reading the Sankei newspaper, trust and accuracy measures, and the viewing time of Public 1 before the experiment. *, **, and *** indicate $p < 0.10$, $p < 0.05$, and $p < 0.01$, respectively.

Figure A.1: Original instructions

アンケートにご協力いただきありがとうございます。このアンケートでは、後日実施する【メディア利用に関する日記式調査】にご協力いただける方を募集しています。

【調査概要】

本調査は、XX（以下、X）が、消費者のテレビの視聴動向や、メディア利用動向を把握するために実施します。
調査期間内に同じ方に複数の調査をさせていただき、日々のテレビの視聴動向を調査することで、視聴者のメディアの利用シーンやニーズを検討いたします。
調査期間は約1ヶ月（11月14日から12月19日まで）です。

本調査の構成は、以下の通りです。

○毎日のテレビ番組の視聴動向調査：
毎日のテレビ番組については、見た番組を番組表の中でチェックします。
毎日ですが、簡単なアンケートです。

○メディア利用に関する調査：
日々のメディア利用や、メディアに対する考え方、Xに対する考え方について伺います。
調査開始時と終了時の2回行います。

毎日アンケートすべてに回答するのは大変だと思いますので、毎日のテレビ番組の視聴動向調査とメディア利用に関する調査はアンケートの回答に3日間の余裕がございます。

【データの取扱い】

アンケートの回答内容は、個人情報（特定の個人を識別する情報）を排除した形で、Xに提供いたします。
本調査結果は個人を特定しない統計的な利用の範囲で活用し、Z等で公表します。
本調査でXが取得した情報を、本調査の目的以外に利用したり、第三者へ提供することは行いません。

【ポイント】

アンケートは毎日届きます。
ポイントは後日まとめて付与させていただきます。
また約1ヶ月間の回答した回数に応じて、後日ボーナスポイントを差し上げます。

【備考】

- 【メディア利用に関する日記式調査】への参加依頼のご連絡は、調査のオペレーションを担当するYからさせていただきます。
- 本調査に参加してくださる方は、本調査の条件について、同意をいただいたものとさせていただきます。
- 応募いただいた中から抽選で、11月13日（日）までに次のステップのご連絡をさせていただきます。
  ご連絡がなかった場合には、何卒ご了承下さい。

Note. XX and X are the name of Public_1 and its abbreviation, respectively. Y is the name of the survey company. Z is the name of Public_1's website.

Figure A.2: Instructions translated into English

Thank you for participating in the survey.
In this survey, we recruit people who can participate in a Record-Based Survey on Media Usage, which will be implemented later.

[Survey outline]

The surveys are implemented in order for XX (hereafter, X) to understand consumers' TV viewing and media usage habits.
During the survey period, we will conduct multiple surveys, and examine viewers' media usage and needs by investigating their daily TV viewing habits.
The survey period will last for approximately one month (from November 14 to December 19).

The composition of the surveys is as follows:

◯ Surveys of daily TV viewing habits:
Checkmark TV program(s) you saw in the TV guide every day.
Even though the survey is conducted daily, it is easy to answer.

◯ Media usage surveys:
Your daily media usage and your thoughts on media and X.
The survey will be conducted at the beginning and at the end of the survey period.

As it may be difficult to answer the questionnaire every day, you can take three days to answer either type of survey.

[Data handling]

Your answers will be given to X after removing personal information from the questionnaires (i.e., information that identifies a particular individual).
The results of the surveys will only be used for statistical purposes without identifying any individual, and will be published in, e.g., Z.
The information obtained by X will not be used for any purpose other than that for which the survey was intended. No information contained in the survey will be given to third parties.

[Reward points]

You will receive a questionnaire every day.
Reward points will be given later.
In addition, bonus points will be given according to the number of answers you give.

[Remarks]

・A participation request for a Record-Based Survey on Media Usage will be sent by Y, which is responsible for survey operations.
・By participating in this survey, we assume that you have agreed to the terms and conditions of the survey.
・Some of the applicants will be chosen at random and informed about the next step of the survey. We kindly ask for your understanding if you are not chosen as a candidate.

Note. XX and X are the name of Public_1 and its abbreviation, respectively. Y is the name of the survey company. Z is the name of Public_1's website.

Figure A.3: Examples of program recommendations

今週末のおすすめの番組

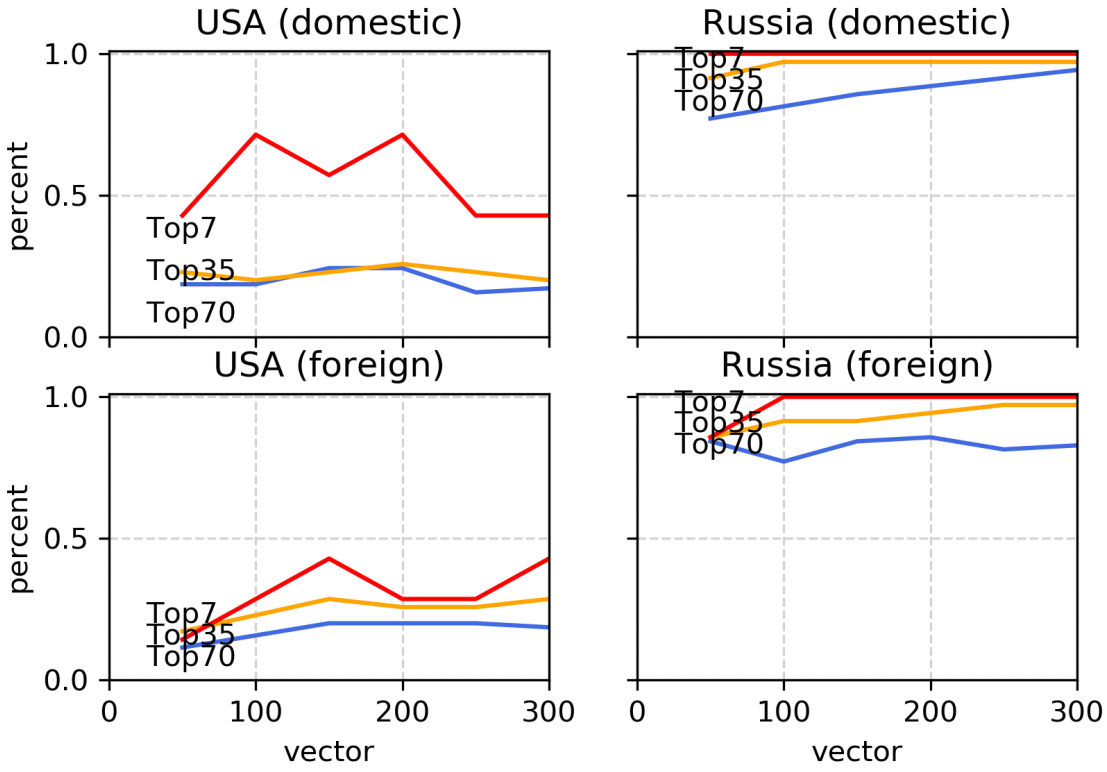
番組名	ＮＨＫスペシャル 戦艦武蔵の最期 ～映像解析　知られざる“真実”～	ＴＶシンポジウム 複雑化する世界の中で ～今求められる人道支援路とは～	ＮＨＫニュース７
放送日 時間	12月4日　日曜日 午後9時～9時50分 （総合）	12月3日　土曜日 午後2時～3時 （Ｅテレ）	12月3日(土)･4日(日) 午後7時～7時30分 （総合）

今週末のおすすめの番組

番組名	ＮＨＫスペシャル 戦艦武蔵の最期 ～映像解析　知られざる“真実”～	ＴＶシンポジウム 複雑化する世界の中で ～今求められる人道支援路とは～	ＮＨＫニュース７
放送日 時間	12月4日　日曜日 午後9時～9時50分 （総合）	12月3日　土曜日 午後2時～3時 （Ｅテレ）	12月3日(土)･4日(日) 午後7時～7時30分 （総合）
内容	日本の最高機密として極秘に建造された戦艦「武蔵」。太平洋戦争末期、壮絶な最期を遂げたとされるが、資料が少なく、真相は謎に包まれてきた。去年３月、フィリピン沖１２００ｍの深海で武蔵が発見され大ニュースとなった。ＮＨＫは膨大な未公開映像とデータを入手。最新の映像解析技術で知られざる武蔵の“真実”が浮かび上がった。驚くべき内部構造や凄まじい攻撃力。定説を覆す“意外な姿”とは。戦艦武蔵の謎に迫る。	世界各地で増加する武力紛争による難民、そして地震や洪水による被災者の増加など支援を必要な人は年々増えています。今、必要な人道支援について議論します。	夜７時、これさえ見れば１日が分かる。今日の日本・世界の今を、あなたのもとへ 【キャスター】高瀬耕造，【サブキャスター】橋本奈穂子，【気象キャスター】菊池真以

Note. The top panel does not include the description of the program (“control” in the second sub-experiment), while the bottom panel includes it (“treatment” in the second sub-experiment). See the main text for more details.

Figure A.4: Predictability of the machine, USA, and Russia topics (drop low frequency words).



Note. The figures plot the percentage of TV programs categorized as either the USA topic or the Russia topic for those that are the closest to official statements (top 7, top 35, and top 70 closest programs): the left figures use closeness to the Japanese government's statement (top) and the US government's statement (bottom), respectively during Prime Minister Abe's visit to Pearl Harbor with former President Obama; the right figures use the Japanese government's statement (top) and the Russian government's statement (bottom), respectively during President Putin's visit to Japan for the summit. We use machine learning to compute the closeness measures, where the window size is set to 5. Only words with a minimum of 50 counts are used in the data training process. See the main text for more details.