Evade and Deceive? Citizen Responses to Surveillance

Kristine Eck*

Uppsala University

Sophia Hatz

Uppsala University

Charles Crabtree

Dartmouth College

Atsushi Tago

Waseda University

Friday 28th February, 2020

*Corresponding author: kristine.eck@pcr.uu.se. We would like to thank Christian Davenport, Chris Fariss, Espen Geelmuyden Rød, Rick Morgan, Yuri Zhukov. Feedback from participants at University of Michigan CPRD workshop, Uppsala University and Kobe University. All data files necessary to replicate the analysis presented in the article will be made available upon publication. R 3.5.1 was the statistical package used in this study. This study received IRB approval from Kobe University (nr. 29017) and was pre-registered at Evidence in Governance and Politics (EGAP). This research was funded by the Swedish Research Council (grant nr.2016-02632).

Abstract

How does state surveillance influence citizens' willingness to express political and social opinions? This article theorizes about different citizen responses to surveillance which fall on what we term the evasion-deception spectrum, including preference falsification, self-censorship, and opting-out. We present the results from an empirical exploration of these responses, drawing on an online survey experiment conducted in Japan. In our survey, we use a novel experimental stimulus to assess whether individuals engage in different forms of evasion and deception when plausibly under government surveillance. The study finds that citizens are substantially more likely to opt-out of sharing their opinions (by exiting a survey) when reminded of their government's capacity for monitoring. This occurs even when it implies a monetary cost (forfeiting payment for the survey) and even in a fully consolidated democracy, where freedoms of speech and opinion are legally codified. We conclude by discussing the implications of this finding for democratic deliberation and citizen-state relations.

Key words: state surveillance, preference falsification, self-censorship, Japan, survey experiments, repression

The development of the internet and other advanced communications technologies has been accompanied by increased opportunities for companies and states to monitor users' activities. As internet users the world over express anxiety over their online privacy, there is good reason to believe that online surveillance has important effects on users' willingness to express opinions and preferences.¹ Yet, we lack an understanding of the different ways in which individuals can respond to surveillance: do they respond by falsifying or by censoring their opinions? Do internet users take actions to avoid online monitoring? We also lack empirical evidence about the prevalence of these different behavioral responses. If people behave evasively or deceptively under conditions of surveillance, this itself creates challenges for observing the effects of surveillance.

This article studies the consequences of state surveillance of citizens' online activity in terms of its effects on citizens' willingness to express opinions on controversial political and social topics. We first develop a theoretical framework that conceptualises different kinds of citizen responses to state surveillance. The topic of preference falsification is well-established in the literature on citizen-state interactions. Scholars hold that citizens living under authoritarian regimes often falsify their preferences to avoid the potentially adverse repercussions associated with opposing the government (Kuran 1989, 1997). Self-censorship is a related topic which is typically discussed as a compliant response to censorship regimes (Cook and Heilmann 2013).

We bring these concepts together by more broadly considering the behaviors individuals use to avoid revealing their true preferences. In doing so, we introduce a continuum of evasion and deception, and distinguish three types of behavior citizens deploy when they wish to maintain their privacy. The first, preference falsification, involves outright deception

¹We define surveillance as 'any collection and processing of personal data, whether identifiable or not, for the purpose of influencing or managing those whose data have been garnered' (Lyon 2001). Unless otherwise specified, we are specifically interested in surveillance by a state (government) actor.

and is well-anchored in the existing research. The second, self-censorship, we define as the avoidance of stating an opinion and locate closer to the evasion end of the spectrum. Further, we introduce a third mode of response, opting-out, which involves both the avoidance of stating an opinion as well as taking actions to avoid surveillance itself. This response is most evasive, since it involves exiting an uncomfortable situation entirely; i.e. exiting the context in which surveillance occurs (Hirschman 1970).

To what extent do citizens engage in these deceptive and evasive tactics when they believe they might be under online government monitoring? This question poses research design and measurement challenges. For example, if certain types of individuals evade surveillance then we can only observe how surveillance effects the class of individuals who do not avoid surveillance. We devise an empirical strategy that addresses this type of challenge. Specifically, we design an online public opinion survey experiment that measures three types of responses to surveillance: preference falsification, self-censorship and opting-out. We incorporate an experimental component that manipulates the text of a website “cookie” to remind respondents that their personal data may be collected and their answers to questions could be monitored by the government. We compare the responses of survey experiment respondents randomly assigned to this surveillance treatment to those from respondents randomly assigned to a standard non-invasive website “cookie” in order to examine whether individuals falsify, self-censor or opt-out when plausibly under observation by the government.

Comparing survey experiment responses across participants who were reminded that they might be monitored and those who received no reminder, we do not find evidence of preference falsification or self-censorship. Survey respondents primed about government monitoring did not engage in deception by expressing different opinions (no preference falsification) nor did they avoid expressing opinions by selecting the response option “no-opinion” (no self-censorship). But we *do* find evidence of opting-out: being primed about government monitoring increased the probability that respondents exited the survey. Indeed, individuals who receive the government surveillance treatment are *8 percentage points* more likely to

attrit. Notably, opting-out occurred even though this action implied forfeiting payment for taking the survey. The results suggest that, in the context of an online survey, opting-out might be perceived as the optimal tactical response: opting-out allows an individual to avoid expressing their true opinions and also evades further data collection and monitoring.

This article contributes in a number of ways to understanding the relationship between surveillance and citizen interactions with authority. First, we further develop existing theory by making conceptual distinctions among different forms of evasion and deception: preference falsification, self-censorship, and opting out. In doing so, we identify behavioral alternatives that have received little attention in the literature. Second, we present an empirical exploration of these responses, drawing on a survey experiment designed to mimic conditions of online state surveillance. Finally, we extend the study of preferences falsification and self-censorship to democracies. While some empirical work has been done on these behaviors in anticipation of government sanctions, little has been conducted in the context of established, consolidated democracies. We focus on the democratic arena in order to explore the extent that evasion occurs even when citizens are not censored by the state and face no legal repercussions for holding political beliefs. To this end, we conduct our online survey experiment in Japan: a consolidated democracy where privacy is valued and protected as a civil right. Like most industrialized democracies, however, the Japanese government has recently passed regulation which expands the scope of the data it can collect on its populace. Japan is thus typical of many advanced democracies: on the one hand citizens enjoy high levels of civil rights, on the other hand, they are concerned about government surveillance.

That Japanese citizens opt-out of an online survey experiment when reminded of the government's capacity to collect personal data suggests that evasive responses occur even when freedoms of speech and opinion are legally codified. The findings of our study thus have implications for democratic deliberation and citizen-state interactions. State surveillance in democracies is problematic because it clashes with values like privacy. Democracy theorists and legal scholars stress that intellectual privacy – the freedom to read, to think, and to

communicate privately – is an essential predicate to democracy (Gavison 1980; Schwartz 1999). In line with our theoretical expectations, our study finds that when intellectual privacy is challenged by state surveillance, citizens engage in evasive behavior which involves avoiding sharing their political opinions. This in turn, undermines citizen participation in public fora for discussion of political issues. In this way, studying citizens’ reactions to perceived state surveillance provides a setting to think more carefully about the consequences of infringing empowerment rights in democracies.

Previous research

Scholars interested in citizen-state interaction have long focused on preference falsification, i.e. when citizens hide or disguise their true preferences. Broadly considered, preference falsification is a form of deception which people engage in on a regular basis as part of everyday life, usually for social and self-conceptual reasons. Our focus, however, is on preference falsification in the context of governance.

Preference falsification is widely discussed in the literature about authoritarian regimes. Scholars have noted that when citizens falsify their preferences, regimes do not realize the extent of underlying grievance and are therefore often surprised when waves of public dissent occur (Kuran 1989, 1997). Theories of informational cascades developed in the post-Soviet context suggest that, in at least some authoritarian contexts, citizens may hold private beliefs that differ from their public statements and in which they have a latent conflict with the state. These beliefs are kept private because authoritarian states typically circumscribe expressive freedoms² and individuals are risk-averse; they typically do not wish to incur the costs associated with contention against the regime. But once activated by the actions of first movers, latent grievances can be transformed into widespread social dissent, as was the case in Eastern Europe in 1989. The same idea has been applied to the Arab Spring:

²We use the term ‘expression’ to denote the right to share information or ideas freely through speech, writing, assembly, or any other form of communication.

countries with repressive laws that had previously successfully prevented challenges to the state experienced widespread protests and demonstrations (Hale 2013; Weyland 2012).

State censorship is therefore likely to induce preference falsification (Duvall and Stohl 1983). Surveillance is a tool that regimes use to enforce censorship regimes or monitor overt compliance with the state more generally. Most research and theorizing about surveillance and preference falsification has thus focused on authoritarian states (Jiang and Yang 2016; Kalinin 2016; King et al. 2013; Kuran 1997; Lorentzen 2014; Roberts 2018; Robinson and Tannenbergh 2018), since citizens who express dissenting opinions to the regime are more likely to pay costs in such states than in democratic states where freedom of expression and speech are subject to legal protections.

But this is a question that citizens in democratic states also grapple with, since some degree of ‘everyday state surveillance’ (Lyon 1994) is ubiquitous. Citizens are watched by video surveillance in public spaces, and states monitor internet traffic and other forms of communication to varying degrees. Although such monitoring in democracies is typically legally circumscribed to specific situations, sweeping anti-terror laws in most countries (Deibert 2003) have raised uncertainty amongst citizens about the scope of government surveillance. Legal scholars underscore the consequences of surveillance, arguing that it has a ‘chilling effect’, that is, citizens may be less willing to publicly express their (true) political views (Penney 2016). Thus far, however, only a few studies have considered the effects of surveillance in democracies. Considering the question of surveillance on preference falsification, Stoycheff (2016) argues that the effect of surveillance is conditional on social dynamics. She finds support for this claim from a small survey experiment (N=255) conducted on Facebook users, which found no effect of being reminded of government surveillance on expressing an opinion about a controversial subject unless the individual believed themselves to hold a minority opinion. Under such conditions, the treatment had a suppressive effect. In another study of online monitoring, Penney (2016) finds that after the June 2013 revelations of the US National Security Agency’s (NSA) widespread online surveillance there was an immedi-

ate and significant decrease in internet traffic to sites considered to be government privacy sensitive, such as Google, Bing, and Wikipedia pages.

Furthermore, there is little systematic empirical work which speaks to the broader array of behaviors individuals use to avoid revealing their true preferences. One reason for this is that the empirical study of phenomena such as preference falsification, dishonesty and social desirability bias poses many research design problems. Psychologists have noted that one of the great challenges of studying dishonesty is measurement precision and accuracy (Pierce and Balasubramanian 2015). In survey research and other fields that rely on respondent self-reporting the problem is discussed under the rubric of social desirability bias, or the disinclination of people to accurately report their true beliefs on sensitive topics (Fisher 1993). This work approaches the topic primarily from the point of view of measurement: scholars are not accurately measuring people’s opinions and attitudes and as a result the empirical foundations of our knowledge of sensitive political topics might be shaky. In recent years, scholars have designed experimental and quasi-experimental studies to circumvent these challenges. These studies confirm that political preference falsification does occur, but again, most of the evidence is from studies of authoritarian states (García-Ponce and Pasquale 2015; Jiang and Yang 2016; Kalinin 2016; Robinson and Tannenbergh 2018). In particular, Tannenbergh (2018) finds no evidence of preference falsification in a subset of African developing democracies with regards to questions about citizen-state relations. We build on this growing body of work, but focus on a broader array of behavioral outcomes, and do so in a fully consolidated democratic context.

Evade and deceive

We begin theoretically by casting the topic of preference falsification in broader terms, by considering a wider array of options available to citizens who do not wish to make their preferences public under conditions of state surveillance. Citizens can engage in evasion as well as deception, and these tactics can be active or passive.

A deception-evasion spectrum

We conceptualize three types of citizen responses to state surveillance which have to do with their expression of political opinions and preferences. These three types of responses can be located along a spectrum which ranges from deception to evasion. The first response takes as its departure point the literature on *preference falsification*. This type of response is the most deceptive of those which we consider: it involves stating outright mistruths. The second has also received scholarly attention, and is typically referred to as *self-censorship*. Self-censorship occurs when individuals preemptively avoid expressing an opinion, for example by remaining silent or stating that they have no opinion. Self-censorship is deceptive in the sense that it involves hiding or disguising true preferences, but also evasive in the sense that it avoids directly stating an opinion. The third response, which has been largely overlooked by scholars, is *opting-out*. This occurs when people exit a situation instead of expressing an opinion. We consider this response most evasive as it both avoids revealing an opinion and, in the context of surveillance, avoids further surveillance.

In theoretical terms, though, scholars have failed to distinguish between hedging (hiding and disguising preferences) and exiting; both have instead fallen within discussions of ‘self-censorship’ or ‘preference falsification’ cast more broadly. In doing so, scholars have failed to recognize that opting out is a conceptually unique tactic which has its own advantages and drawbacks. On the one hand, it insulates the individual from having to decide between committing outright deception or expressing a true preference and facing a possible confrontation for holding a dissenting view. It allows an individual to evade the risk of expressing a preference, just as in self-censorship. In contrast to self-censorship, however, opting-out also allows the individual to escape the situation in which preferences are being solicited and potentially monitored.

Before we continue, a proviso is in order. We have presented three closely related but conceptually distinct tactical responses, but we do not present expectations about *which* they will prefer. We expect such a calculation to be conditional not only on heterogeneous

personality traits, but also contextual conditions which influence individuals’ risk assessment. Citizens would presumably need to weigh expressive utility—intrinsic benefits gained from expressing opinions—and other benefits against the possible costs incurred. Given the paucity of research on this topic, our approach is to gain theoretical traction by conceptualizing different types of responses, and highlighting tactics like opting-out which have received little or no attention in previous research. Our focus is therefore to understand whether there is an effect of perceived online surveillance on people’s use of evasive and deceptive tactics; we consider this a necessary first step before we can begin to study tactical adjudication.

Authority

We consider the situation of state surveillance and ask how it impacts on individuals’ propensity to adopt evasive and deceptive tactics when asked about sensitive topics. As we have described, the majority of prior work on preference falsification and self-censorship in the context of citizen-state interactions has focused on situations of authoritarianism, modelling citizen responses in terms of resistance or compliance. False compliance and dissimulation are famously “weapons of the weak” (Scott 1985); they can be seen as everyday acts of resistance which serve important personal integrity ends (Scott 1990). Theorists of self-censorship, in contrast, typically articulate the phenomenon as a compliant response to censorship regimes.

Whether evasion and deception are acts of resistance or compliance, an underlying premise is that evasion and deception are reactions to state repression; the absence of a censor would imply that censorship is non-coercive (Cook and Heilmann 2013). We think this casts coercion in too narrow terms. In censorship regimes, a public authority deploys coercive tactics or the threat thereof to enforce explicit restrictions on expression. In self-censorship, there is neither censor nor prohibitions, but individuals preemptively abstain from expression that they deem to have the potential to bring them in contention with authority. The absence of a public authority tasked with censorship need not entail an absence

of perceived threat. Self-censorship can thus occur outside the context of censorship regimes, as can other forms of evasive and deceptive behavior. Outside authoritarian states, surveillance by authority figures in democratic states can affect individuals' propensity to adopt evasive and deceptive tactics. The act of observing can contain an *implicit* coercive element because of the power imbalance between the state and citizen; a risk-averse approach for the citizen is reticence. This effect should be heightened when individuals believe an authority figure to be observing their behavior, even if the authority figure has no right or incentive to sanction.

The focus on preference falsification and self-censorship in authoritarian states gives primacy to the possibility of negative externalities —sanctions —as a motivating factor. In a democratic setting where citizens typically are not punished by the state for their political beliefs, there is little reason to believe that there is any real risk for suffering costs under government surveillance. But individuals may also be driven by intrinsic motivations to react to online government surveillance; they may object in principal to perceived state intrusions into their privacy. This is one reason why privacy scholars insist that the right to privacy is fundamental to democratic practice, expressing concerns that self-censorship corrodes democratic values by diminishing accountability and inhibiting the sincerity on which democratic debate relies (Elster 1997; Festenstein 2018). While we cannot unpack the motivation driving the observable behaviors of interest—tactical evasion and deception under perceived government online surveillance—we posit that the possibility of intrinsic motivations can help to explain why we would expect citizens in democracies to use these tactics even when the possibility of extrinsic costs is negligible.

Hypotheses

From this discussion, we derive three hypotheses relating to our three main concepts: preference falsification, self-censorship, and opting-out. As indicated by the language used in the hypotheses, we will empirically test these expectations with an online survey experiment

described in the next section.

H1 (Preference falsification): Respondents answer differently to questions if they are exposed to information that the government might be able to observe their responses and other respondents.

H2 (Self-censorship): Respondents are more likely to conceal their opinions if they are exposed to information that the government might be able to observe their answers.

H3 (Opt-out): Respondents are more likely to leave a survey experiment if they think that the government might be able to observe their responses.

Empirical strategy

In contrast to prior work on topics related to state monitoring and censorship, we are not interested in autocracies or transitioning democracies, but rather full democracies with long histories of consolidated governance. For our empirical investigation, we choose Japan, a country with over 70 years of experience as a consolidated democracy.

Since Japan’s transition to a full democracy in 1946, the protection of privacy and freedom from censorship has been codified in its constitution and through special law. Yet, the Japanese government has also expanded and legally codified its monitoring of the populace. In particular, the recent passing of an ‘anti-conspiracy’ law in 2017 grants the Japanese government greater freedom in the use of surveillance. The law criminalizes discussion about committing a crime regardless of whether action is taken. To detect relevant discussions, the law allows the government widespread powers to surveil communications to search for crime-related conversations (Ogasawara 2017). The UN Special Rapporteur on the Right to Privacy has questioned the safeguards and remedies in Japanese law with regard to privacy and surveillance, stating that the ‘anti-conspiracy’ bill “may affect the exercise of the right to privacy as well as other fundamental freedoms given its potential broad application” (OHCHR 2017). Opponents to the law worry that “the Conspiracy Bill prevents people from talking about inconvenient and critical facts because of fear; fear of being criminalized

for communicating something unfavorable to the state” (Ogasawara 2017, 484).³ So although Japan offers high levels of democratic governance and civil rights, its public also has cause for concern about government monitoring of their online communication. Indeed, recent polls suggest that citizens in Japan are concerned about possible government surveillance; almost 60% of the population report being “very much” or “a great deal” worried about government wire-tapping or reading their mail or email (Inglehart et al. 2014).

In many ways, Japan is typical of the democratic context we are interested in. Privacy of personal information is valued in Japan at levels analogous to other industrialized nations (Adams, Murata, and Orito 2009; Murata et al. 2015). Yet, like the vast majority of industrialized countries in the post-9/11 era, Japan has enacted legislation to expand the capacity of the state to monitor Internet communication (Banisar and Davies 1999; Deibert 2003). In Japanese polls, as in cross-national surveys, citizens have expressed concern about privacy violations, believe surveillance laws are intrusive, and worry about providing personal information online (Banisar and Davies 1999; Eurobarometer 2008; Inglehart et al. 2014; Watson, Finn, and Barnard-Wills 2017).⁴ It is in the empirical case of Japan then, that we investigate our more general research question: the extent to which citizens respond to surveillance with deception and evasion with regards to expressing their political opinions.

³We provide additional background on privacy and government monitoring in Japan in online appendix E.

⁴See Watson, Finn, and Barnard-Wills (2017, 274-5) for a list of national and cross-national surveys on privacy, security and surveillance. Caution should be advised, however, when comparing levels of concern about surveillance across countries— for the very reason that individuals may falsify or censor their true preferences. As Tannenbergh (2018) argue, expressed opinions on sensitive topics may be systematically biased across countries as the perceived risk of repercussions varies across countries.

Research design

Studying preference falsification, self-censorship, and opting out requires researchers to observe individuals' expressed opinions both in the presence and absence of state surveillance. Even if observational data of this nature were obtainable, however, they are likely to be biased. Individuals who are surveilled are likely to differ systematically from those who are not, and these differences could explain differences in their willingness to express true preferences. If certain kinds of individuals evade surveillance, this in itself introduces strategic selection bias. We address these kinds of difficulties by designing an online experiment that randomly assigns individuals to surveillance (treatment) and non-surveillance (control) conditions. By virtue of the treatment being randomly assigned, we can attribute differences in the nature of responses to the effect of the surveillance treatment.

Our study is comprised of four parts. First, respondents are provided with an informed consent statement that asks them to participate in a survey experiment on social and political attitudes conducted by University researchers. They must consent to continue. Second, they are asked four demographic questions about their sex, education, age, income. They are also asked about their interest in politics and attention to politics.⁵ Third, they receive our experimental surveillance treatment, which we describe in more detail below. Fourth, respondents are asked their opinion on eight topics across different sectors (political, social, economic and environmental). These topic areas were chosen for their potential controversiality, in order to maximize the possibility that individuals might want to hide their preferences. We also chose these divisive issues on the basis of newspaper coverage, case expertise, and discussions with citizens; they were also pre-tested and revised in a pilot sur-

⁵We collect this information for two reasons: to potentially examine heterogeneous treatment effects, and to ensure that exiting the survey experiment will entail a cost since some (small) amount of time has already been spent on it and the payment will be forfeited.

vey.⁶ See Figure A-1 in the online appendix for an illustration of the flow of the survey experiment.

The central feature of our survey experiment is our experimental treatment. It takes the form of a website cookie message that our survey experiment respondents encounter. The vast majority of websites collect and store user information through cookies (also called HTTP cookies, web cookies, Internet cookies, or browser cookies) which are small pieces of data sent from a website and stored on the user’s computer while the user is browsing. In some countries, the law requires that websites which track user data through cookies must inform users of this, typically doing so with a small pop-up window or banner.

In our control (non-surveillance) condition we use fairly standard language to notify respondents that the website uses cookies:

The next set of questions ask about your views on some controversial political issues.

While we ensure your answers are kept confidential, we would like to remind you that this
website uses cookies.

The government surveillance treatment combines this cookie warning with a reminder about the possibility of government monitoring:

The next set of questions ask about your views on some controversial political issues.

While we ensure your answers are kept confidential, we would like to remind you that this
website uses cookies.

When you answer this survey, you send pieces of data to us, including your IP address, your geographic location, your survey answers, and other information. The Government of Japan reserves the right to monitor activity on this website and can request this information from
us.

One possible concern with a comparison of the government surveillance and non-surveillance (control) cookie messages is that we not be identifying the effect of government surveillance

⁶The pilot study was conducted in January 2018 on a sample of 1646 respondents recruited on Yahoo Crowdsourcing Japan, a platform analogous to M-Turk.

but surveillance more generally. Our control condition reminds respondents that the website hosting the survey experiment collects cookies. In the message which we designed to mimic government surveillance, we remind respondents of the specific kinds of personal data that are collected with cookies and that the Japanese government could request this information. The government surveillance condition differs from the control condition in two ways then: (1) it details the specific kinds of personal data that are collected, and (2) it reminds respondents that the government can collect this data. Another way of thinking about this issue is that we use a ‘double-barreled treatment’ (Kertzer and Brutger 2016) and cannot identify whether (1) or (2) explains the results we observe.

To tease out a government monitoring effect from a more general monitoring effect, we also included a third kind of cookie message—which we call a company surveillance treatment—in our survey. This cookie message contained the same information regarding the specific kinds of personal data collected with cookies, but stated that the website (a corporate actor) uses this information to improve services:

The next set of questions ask about your views on some controversial political issues.

While we ensure your answers are kept confidential, we would like to remind you that this website uses cookies.

When you answer this survey, you send pieces of data to us, including your IP address, your geographic location, your survey answers, and other information. We use this information to help us deliver our services.

In all three conditions (control, government surveillance and company surveillance), we prime the respondents to think that they will be asked about their opinions on controversial political issues in order to focus their attention on the implications of the cookie text. Respondents were assigned to one of the three conditions with equal probability.

The next important feature of our survey experiment design regards how we measure different forms of deception and evasion. In order to capture the three modes of citizen responses we’ve conceptualised –preference falsification, self-censorship, and opting-out– we

designed three versions of the survey. All conditions (control, government surveillance and company surveillance) are present in the three versions of the online survey. Respondents were assigned to one of the three survey experiment versions with equal probability, and the three versions were conducted simultaneously within the same survey experiment project, thus ensuring that respondents could not participate more than once and thereby ensuring against spillover (Trochim and Donnelly 2008).

Survey version 1 (V1) measures **preference falsification**. In this version, respondents were presented with the randomly assigned cookie message and then proceeded to the opinion questions. For each opinion question, we are then able to look at the difference in expressed preferences across treatment and control groups. If there is a difference, we can conclude that the surveillance treatment has an effect on respondents’ propensity to express their true preferences. In real life, people are rarely forced to express an opinion (Hayes, Glynn, and Shanahan 2005), but this version of the survey experiment is designed to create such a forced option: respondents are told that they must answer all questions in order to complete the survey experiment and receive payment. Survey V1 was thus designed to prevent evasion and force a statement of preference.

Survey version 2 (V2) measures **self-censorship**. This version is identical to V1, except that we added the additional answer option ‘no opinion’ to all survey experiment questions (see Table A-1 in the online appendix for a clarification of how V1 and V2 differ). By including a ‘no opinion’ option, respondents are no longer forced to express their preferences; they are given an option to evade expressing their true beliefs on a question-by-question basis. While respondents must answer all questions to receive payment, nothing prevents them from answering ‘no opinion’ across all of the eight opinion questions. We measure self-censorship as the extent of ‘no opinion’; if there is a difference in the extent of ‘no opinion’ comparing treatment and control groups we can conclude that the treatment has an effect on respondents tendency to self-censor.

Survey version 3 (V3) measures **opting-out**. In this version, respondents are asked

during the cookie condition whether they consent to continue. Respondents who agreed to the cookie condition continued to the opinion questions, while those who disagreed with the policy exited the survey experiment at that point, ostensibly without payment.⁷ To accomplish this, the text below was added to the treatment and control cookie messages in V3.

Please agree to our use of cookies to continue with the survey. To exit the survey (without payment) please click "I do not agree".

As in V1, there is no ‘no opinion’ option in the subsequent opinion questions in V3. Instead, attrition becomes a meaningful metric in V3. To assess whether surveillance induces opting-out, we can compare the propensity of exiting the survey experiment across treatment and control groups.

Data

The survey experiment was fielded in Japan in June 2018.⁸ We conduct the survey experiment online for four reasons. First, a primary domain for democratic expression and debate about political and social issues today is online, making the internet an important forum for deliberation and participation (Farrell 2012; Gibson and Cantijoch 2013; Oser, Hooghe, and Marien 2013; Schlozman, Verba, and Brady 2010). Second, a rash of news coverage of events related to online surveillance—from revelations about NSA monitoring to corporate hacks—has raised global awareness and concern about online privacy (Stoycheff 2016). Third, the online setting allows us to create treatments that would be impractical in face-to-face settings. Finally, researchers have shown that web-based data collection is more likely than face-to-face interviews to elicit truthful responses to sensitive questions in democratic countries (Heerwegh 2009; Holbrook and Krosnick 2009).

⁷We ultimately paid all respondents, regardless of whether they exited prior to completion.

⁸This study received IRB approval from (redacted) and was pre-registered at Evidence in Governance and Politics (EGAP).

We contracted Qualtrics Research Services to recruit in sample, which in turn subcontracted to local Japanese survey companies. Respondents received remuneration of about 50 cents to US\$1.30 for participation. The final sample size is a total of 4514 respondents. Details about sampling, treatment assignment and respondent demographics are provided in online appendix A.

Results

We first examine responses to the eight opinion questions across the three different randomly assigned versions of the survey, which were designed to pick up the three forms of citizen responses we theorized: preference falsification (V1); self-censorship (V2), and opting out (V3). For this, we recode each of the opinion questions to binary indicators, with ‘1’ indicating a specific preference. Figure 1 shows the proportions of respondents who expressed the specified preferences, within each of the three survey experiment versions.

Figure 1 about here

Because respondents were randomly assigned to one of the three survey experiment versions, we can attribute differences in the propensity of expressing a certain preference to differences in the survey experiment versions. It is important to note, however, that in V3 (‘opt-out’), these are the preferences among the respondents who agreed to continue after the cookie message; we cannot observe from the preferences of those who opted-out of V3. Since the respondents who stayed in V3 are likely to differ systematically from the respondents in V1 and V2, we cannot compare preferences across V3 and the other survey experiment versions. What we can see in Figure 1 is that, generally, lower proportions of respondents assigned to V2 (who had the ‘no-opinion’ answer option) expressed the preferences we coded, compared to respondents assigned to V1 (who did not have a ‘no-opinion’ option). This suggests that self-censorship occurs, but it does not tell us whether self-censorship is due to the surveillance treatments we incorporated. For this, we analyse the impact of our surveillance treatments in each survey experiment version.

Preference falsification

To test if monitoring leads to preference falsification (Hypothesis 1), we assess whether respondents' expressed preferences on the eight opinion questions differ across the surveillance and non-surveillance conditions. Given the dichotomous nature of our outcome variables, and in line with our pre-registration plan, we estimate the average treatment effect of our experimental interventions with linear probability models.⁹ We present the results of these analyses in Figure 2.

Figure 2 about here

Counter to H1, we see no evidence of preference falsification across any of the opinion questions under the either surveillance treatment (government or company). More specifically, we cannot reject the null that the government surveillance treatment had no effect on respondent answers. Importantly, the 95 confidence intervals cover only trivial values, ruling out the possibility that the treatments had a meaningful effect.

⁹We use dichotomous outcome measures and report the results of linear probability models to make interpretation of our results easier. We obtain substantively similar results if we use logit or probit models instead. When using the untransformed versions of our outcomes, we use ordinary least squares regression models. All models include the following mean-centered covariates collected before the experimental treatments: age, education, gender, income, sex, attention to politics, and satisfaction with politics. As we describe in our pre-registration plan, we use HC2 robust standard errors to make inferences. We obtain substantively similar results if we use classic standard errors instead. Given plausible assumptions about the sizes of treatment effects, statistical power simulations demonstrate that our probability of rejecting a false null hypothesis at the 0.05 significance level exceeds 80 percent for our study.

Self-censorship

Survey version 2 was designed to detect evidence of self-censorship. To test if monitoring leads to self-censorship (Hypothesis 2), we assess whether respondents who were exposed to different cookie messages differ in the extent to which they select ‘no opinion’ in V2. The results are shown in Figure 3.

Figure 3 about here

In contrast to our expectations (H2), we find no evidence that government or company surveillance increases the tendency of respondents to censor their preferences in any of the opinion questions.¹⁰ Once again we see here that the confidence intervals cover only a range of trivial values, providing evidence that our treatment had a negligible effect.

Opting-out

Finally, we assess whether monitoring induces respondents to opt-out of the survey experiment by comparing treatment and control groups within survey experiment version 3. In V3, the cookie message included the option to “disagree” with the use of cookies and exit the survey experiment (ostensibly without payment). Our outcome measure in this survey experiment version is whether respondents opt-ed out of the survey experiment (by clicking “disagree”) or stayed in (by clicking “agree”).

We see substantial variation in this outcome measure: 21% of the people who received the generic control cookie message left the survey, while 29% of the people who received the government surveillance treatment message left the survey experiment and 24% of those

¹⁰To be precise: we see no evidence of self-censorship in response to the company surveillance treatment across the opinion questions, with the exception of the question about the Japan Self-Defense Forces (SDF). We account for this multiple comparison issue by applying a Bonferroni correction to the p-value associated with the coefficient for the company treatment. Once we do this, we cannot reject the null hypothesis that the company treatment did not lead to self-censorship for the SDF outcome ($p > 0.05$).

who received the company surveillance treatment left. To examine whether our surveillance cues caused people to opt-out, we estimate linear probability models. We present the results from these models in Figure 4.

Figure 4 about here.

The figure shows that, in line with H3, respondents who received the government surveillance treatment were significantly more likely to decline consent and exit the survey, compared to respondents assigned to the control.¹¹ The effect of the treatment is substantively large, indicating that individuals who received the treatment were about 8 percentage points more likely to attrit.¹²

One concern might be, though, that survey respondents exit because after the government condition because of a breakdown in trust between respondents and researchers. We can think of a key reason why this might not be the case. This is because we do not find that the company surveillance treatment causes respondents to exit at a higher rate. If our experimental prompt causes respondents to feel they were lied to in the informed consent statement at the beginning of the survey, they should exit the survey after the company surveillance treatment, just as they did after the government surveillance treatment. In fact,

¹¹One concern with this finding might be that it is a statistical fluke, the result of repeatedly testing many hypotheses with the same data. To address the multiple comparison issue, we apply a Bonferroni correction to the p-value associated with the coefficient for the government treatment. After this adjustment, we can still reject the null hypothesis that the treatment had no effect ($p < 0.05$).

¹²The overlapping confidence intervals in Figure 4 are not necessarily evidence that the difference between point estimates is not statistically significant (Schenker and Gentleman 2001). To assess whether the difference between the estimated effects of the government and company surveillance treatments is statistically significant, we re-estimate our model with the company surveillance condition as the reference category. We find that the government surveillance treatment is still statistically significant ($p < 0.10$).

we might expect a greater break down in respondent-researcher trust after the company treatment, because the language in this cookie message suggests that the researchers are choosing to monitor respondents, instead of passively having to comply with a government directive. That we find the opposite suggests that a breakdown in trust is not the primary driver of respondent exit.

To properly assess this effect size, though, we need to place it in context. On the one hand, it is striking that a prime about government surveillance would elicit such a high exit rate for Japanese respondents.¹³ This is because Japan is a fully developed democracy and boasts many legal protections for freedom of speech. As a consolidated democracy, we might expect that the effect of the government surveillance treatment will be weaker in comparison to less consolidated democracies or autocracies. On the other hand, one might be concerned that our effect size is so large, in part, because of the language we use in our treatment texts. The possible concern here might be that the language is either too direct or artificial. We think that a very promising direction for future work would be to construct more subtle or unobtrusive treatments.

Discussion

We theorized that there are a range of tactical responses to state surveillance: citizens may engage in deception by falsifying their true preferences, may evade and deceive by censoring their true preferences, or they may evade surveillance entirely by opting-out of situations in which they cannot express their preferences privately. Our hypotheses did not attempt to adjudicate between these different evasive tactics under consideration. Our findings suggest,

¹³This is particularly the case because our control condition might already prime respondents to think about surveillance. If so, then we should be less likely to observe an effect of our government surveillance treatment on respondent behavior. In this light, we might consider the strong, substantive effect we find as a lower bound on the effect we might find with a different control condition that does not reference surveillance.

however, that in at least some situations opting-out may be the first line of defense against surveillance. This is an important observation since opting out is the tactic which has received least theoretical attention. The field has focused on more overt forms of deception and self-censorship but our findings indicate that individuals may adopt more subtle, less identifiable steps to preserve their integrity.

Opting-out may be an optimal tactical response for several reasons. First, it is a pragmatic approach to the discomfort of being watched (Marx 2003). Kuran (1997, p.5) notes that a defining feature of preference falsification is the discomfort produced as a consequence of lying; this discomfort is avoided through exit. It is also simple means of avoiding a direct confrontation with authority, especially in situations in which it is unclear in which direction one needs to align one's preferences. Perhaps most importantly, however, the strategy of opting-out allows an individual to escape an uncomfortable or threatening context entirely. In the context of an internet survey, opting out offers the advantage of avoiding further monitoring: exiting the survey experiment means no more data will be collected via cookies.

While opting-out has clear advantages to respondents, it also has clear disadvantages, at least in the context of our survey. First, opting-out means respondents forfeit payment for completing the survey. This may make opt-ing out puzzling since rational theories of utility maximization would lead us to expect that preference falsification or self-censorship to be the most logical responses in this setting. By completing the survey experiment with false responses or 'no opinion', the respondent receives the economic reward and cannot pay a cost for dissembling since there is no way to validate whether individual responses represent true preferences or not. Especially in democratic contexts, where preference falsification and self-censorship may not be met with reprisals from a state, opting-out is relatively more costly. Yet what we find is the opposite: respondents receiving the government surveillance treatment prefer to forfeit their private benefit. From rationalist point of view, this makes sense if we consider the benefit of avoiding exposure to further surveillance.

A second disadvantage of opting-out is that it requires survey experiment respondents

to forfeit their opportunity to express their opinions. This is qualitatively different from both preference falsification and self-censorship: in these tactics, citizens disguise or hide their true preferences but stay in the forum for political deliberation. In opt-ing out, the respondent withdraws from political deliberation entirely. The strategy of exiting thus has implications for the resilience of democratic participation under conditions of government surveillance. In our survey, we find that when a tradeoff presents itself between freedom of expression and protecting online privacy, individuals choose to protect their privacy. The broader implication is that while a majority of citizens may feel safe expressing opinions and engaging in debate on online platforms even if they believe they might be monitored, at least some percentage will abstain entirely.

Conclusion

In this article, we conceptualized three points on an evasion-deception continuum: preference falsification, self-censorship, and opting-out. We theorized that if reminded of the possibility of online surveillance by the government, citizens would be more likely to engage in these tactics. We examined this argument in the context of an established democracy where civil liberties are well-entrenched. To test our argument, we developed a novel survey experiment which we conducted in Japan, in which we manipulated the text of a website cookie to remind respondents of the possibility of online surveillance.

We found that respondents who were reminded of the possibility of government surveillance were more likely to exit the survey experiment (opt-out) compared to respondents who received a typical cookie message. This occurred even while it implied respondents forfeit both payment for the survey experiment and the opportunity to express their opinions, and even in a context where freedom of expression and intellectual privacy is legally codified. We found no evidence of preference falsification or self-censorship: survey experiment respondents did not provide false answers nor did they select ‘no-opinion’ on controversial topics when reminded of government monitoring. We also failed to find evidence of deceptive or

evasive behaviours when reminding survey experiment respondents of monitoring by a corporate actor, suggesting that opting-out is a response specific to situations of government surveillance (see Table B-1 in the online appendix for a summary of the study findings).

There are several implications of these findings. First, they suggest that online state surveillance has a non-negligible effect on citizens and that this may translate into a reluctance amongst part of the populace to engage in online discussions, even in democracies. In our study, we found that surveillance can create trade-offs between fundamental democratic values such as freedom of expression and intellectual privacy: we find that individuals prefer to forfeit an opportunity to express their political opinions in order to protect their online privacy. This in turn, underscores the importance of perceptions of privacy and security for accurate measurement in survey research. Second, the fact that we found this result in an established democracy suggests that the effect may be even more profound in semi-democracies and autocracies, and that engaging in various forms of deception and evasion may be everyday forms of resistance in these settings. Third, the findings indicate the utility of considering a wider array of evasive behaviors than has been previously considered in the literature, which has typically truncated its examination of this spectrum to the most overt forms.

The problem of strategic selection bias poses an important challenge to the study of this topic; certain kinds of individuals self-select into surveillance (are more likely to be targeted by monitoring programmes) and out of surveillance (are more likely to avoid surveillance). Our solution was a survey experiment, where conditions mimicking online surveillance could be randomly assigned. This gave us leverage on this topic, but it also came with limitations. Our approach is reductive in that it focuses narrowly on one form of expressed behavior: willingness to state an opinion in an online survey. We imagine that evasion and deception could take very different forms in other contexts. On other online platforms, such social media applications and websites, evasion and deception could occur via encryption, VPNs, by clearing cookies and by going offline. In face-to-face surveys or live discussions, individuals

may try to hide their true attitudes, give misinformation or exit the conversation. In the same vein, a government's collection of data online is only one form of surveillance; many other modes exist. Our focus took as its departure point an interest in strategies to evade online surveillance, which we hope it will inspire further work on the growing field of research on resistance to surveillance more generally (Fernandez and Huey 2009; Lyon 2001; Martin, Van Brakel, and Bernhard 2009; Marx 2003; Ullrich and Knopp 2018).

Our findings also suggest that the field would benefit from more nuanced theorizing about strategies of evasion and deception, in order to distinguish various tactics and the conditions under which they are employed. In considering how citizens adjudicate between tactics, we suspect that both psychological and rational mechanisms are activated, and that these are heavily conditioned by contextual factors. Democratic and autocratic contexts create different incentive structures in terms of the likelihood of reprisals for expressing controversial opinions as well as the intrinsic benefits of exercising voice. There is much more to be done in terms of understanding citizens' relationship to state surveillance. How are costs calculated and how on what criteria do individuals optimize when estimating whether to acquiesce or resist? How do state institutions and practices impact on that calculus? Our survey experiment departed from an atomistic individual; respondents did not know each other and could not coordinate responses. But in other situations, individuals are embedded in social networks which may help them to mitigate the effects of surveillance or develop collective solutions (see Earl and Kimport (2010) and Tufekci (2017) on internet activism).

While we do not see any reason that the results of our study should be particular to Japan, it is worth keeping in mind the ways in which Japan differs from other consolidated democracies. First, Japan's experience with autocracy is within recent memory. Although only a few of our respondents would likely be able to remember this period, its legacy is likely to have after-effects. García-Ponce and Pasquale (2015), for example, infer from a study of Zimbabwe that exposure to violent repression leads citizens to report inflated levels of trust in state institutions. Jiang and Yang (2016) find that a political purge in Shanghai increased

the propensity of respondents in a subsequent opinion poll to express support for the Chinese regime. In line with these studies, it is possible that a history of political repression affects how Japanese citizens express political opinions, and that we would not find similar results in consolidated democracies which are farther removed from their illiberal pasts. That said, we do not wish to overstate Japan's uniqueness; it is clear that citizens living in diverse political contexts are aware of the possibility of government surveillance and that many find it cause for concern.

We were intrigued by the powerful effect of our government surveillance treatment on opting-out. Our demographic covariates were too broad to capture what drives these individuals to react as they do. We wonder what traits distinguish these individuals from those who chose to participate in the study. In other contexts, are these the people who would be first movers in resistance movements? Or is their exit from the survey experiment an indication of a powerful risk aversion?

Finally, we were puzzled by the fact that when given the opportunity to opt-out in one version of our survey, the government surveillance treatment prompted many respondents to do so. But when we did not provide an offer to opt-out in the other two versions of our survey, the government treatment did not induce respondents to simply close the web browser. The implications of this discrepancy are worrisome. Why do individuals seem to need permission to opt-out? Why be obedient to our artificial, non-enforceable survey experiment authority? We hope that our findings and these remaining puzzles will inspire researchers interested in understanding issues of resistance and conformity in the context of citizen-state relations, particularly given the increasingly prevalence of state monitoring.

References

- Adams, Andrew A, Kiyoshi Murata, and Yohko Orito. 2009. "The Japanese Sense of Information Privacy". *AI & Society* 24 (4): 327–341.
- Banisar, David, and Simon Davies. 1999. "Global Trends in Privacy Protection: An International Survey of Privacy, Data Protection, and Surveillance Laws and Developments". *J. Marshall J. Computer & Information Law* 18 (1).
- Cook, Philip, and Conrad Heilmann. 2013. "Two Types of Self-censorship: Public and Private". *Political Studies* 61 (1): 178–196.
- Deibert, Ronald J. 2003. "Black Code: Censorship, Surveillance, and the Militarisation of Cyberspace". *Millennium: Journal of International Studies* 32 (3): 501–530.
- Duvall, Raymond D., and Michael Stohl. 1983. "Governance by Terror". In *The Politics of Terrorism*, ed. by Michael Stohl, 179–219. New York: Marcel Dekker.
- Earl, Jennifer, and Katrina Kimport. 2010. "Changing the World One Webpage at a Time: Conceptualizing and Explaining Internet Activism". *Mobilization* 15 (4): 425–446.
- Elster, Jon. 1997. "The Market and the Forum: Three Varieties of Political Theory". In *Deliberative Democracy*, ed. by James Bohman and William Rehg, 471–498. Cambridge, MA: MIT Press.
- Eurobarometer. 2008. *Data Protection in the European Union: Citizens' Perceptions*. Flash Eurobarometer Series No. 225.
- Farrell, Henry. 2012. "The Consequences of the Internet for Politics". *Annual Review of Political Science* 15:35–52.
- Fernandez, Luis A, and Laura Huey. 2009. "Is Resistance Futile? Thoughts on Resisting Surveillance". *Surveillance & Society* 6 (3): 199–202.
- Festenstein, Matthew. 2018. "Self-censorship for Democrats". *European Journal of Political Theory* 17 (3): 324–342.

- Fisher, Robert J. 1993. "Social Desirability Bias and the Validity of Indirect Questioning". *Journal of consumer research* 20 (2): 303–315.
- García-Ponce, Omar, and Benjamin Pasquale. 2015. "How Political Repression Shapes Attitudes toward the State: Evidence from Zimbabwe". Unpublished manuscript.
- Gavison, Ruth. 1980. "Privacy and the Limits of Law". *The Yale Law Journal* 89 (3): 421–471.
- Gibson, Rachel, and Marta Cantijoch. 2013. "Conceptualizing and Measuring Participation in the Age of the Internet: Is online Political Engagement Really Different to Offline?" *The Journal of Politics* 75 (3): 701–716.
- Hale, Henry E. 2013. "Regime Change Cascades: What We Have Learned from the 1848 Revolutions to the 2011 Arab Uprisings". *Annual Review of Political Science* 16:331–353.
- Hayes, Andrew F, Carroll J Glynn, and James Shanahan. 2005. "Willingness to Self-censor: A Construct and Measurement Tool for Public Opinion Research". *International Journal of Public Opinion Research* 17 (3): 298–323.
- Heerwegh, Dirk. 2009. "Mode Differences Between Face-to-face and Web Surveys: An Experimental Investigation of Data Quality and Social Desirability Effects". *International Journal of Public Opinion Research* 21 (1): 111–121.
- Hirschman, Albert O. 1970. *Exit, Voice, and Loyalty: Responses to Decline in Firms, Organizations, and States*. Harvard University Press.
- Holbrook, Allyson L, and Jon A Krosnick. 2009. "Social Desirability Bias in Voter Turnout Reports: Tests Using the Item Count Technique". *Public Opinion Quarterly* 74 (1): 37–67.
- Inglehart, R., et al. 2014. *World Values Survey: Round Six - Country-Pooled Datafile Version*.
- Jiang, Junyan, and Dali L Yang. 2016. "Lying or Believing? Measuring Preference Falsification from a Political Purge in China". *Comparative Political Studies* 49 (5): 600–634.

- Kalinin, Kirill. 2016. "The Social Desirability Bias in Autocrat's Electoral Ratings: Evidence from the 2012 Russian Presidential Elections". *Journal of Elections, Public Opinion and Parties* 26 (2): 191–211.
- Kertzer, Joshua D, and Ryan Brutger. 2016. "Decomposing Audience Costs: Bringing the Audience Back into Audience Cost Theory". *American Journal of Political Science* 60 (1): 234–249.
- King, Gary, et al. 2013. "How Censorship in China Allows Government Criticism but Silences Collective Expression". *American Political Science Review* 107 (2): 1–18.
- Kuran, Timur. 1989. "Sparks and Prairie Fires: A Theory of Unanticipated Political Revolution". *Public Choice* 61 (1): 41–74.
- Kuran, Timur. 1997. *Private Truths, Public lies: The Social Consequences of Preference Falsification*. Cambridge, MA: Harvard University Press.
- Lorentzen, Peter. 2014. "China's Strategic Censorship". *American Journal of Political Science* 58 (2): 402–414.
- Lyon, David. 1994. *The Electronic Eye: The Rise of Surveillance Society*. Minneapolis: University of Minnesota Press.
- Lyon, David. 2001. *Surveillance Society: Monitoring Everyday Life*. Buckingham: Open University Press.
- Martin, Aaron K, Rosamunde E Van Brakel, and Daniel J Bernhard. 2009. "Understanding Resistance to Digital surveillance: Towards a Multi-disciplinary, Multi-actor Framework". *Surveillance & Society* 6 (3): 213–232.
- Marx, Gary T. 2003. "A Tack in the Shoe: Neutralizing and Resisting the New Surveillance". *Journal of Social Issues* 59 (2): 369–390.

- Murata, Kiyoshi, et al. 2015. “So What if the State is Monitoring Us? Snowden’s Revelations Have Little Social Impact in Japan”. *ACM SIGCAS Computers and Society* 45 (3): 361–368.
- Ogasawara, Midori. 2017. “Surveillance at the Roots of Rveryday Interactions: Japan’s Conspiracy Bill and its Totalitarian Effects”. *Surveillance & Society* 15 (3/4): 477–485.
- OHCHR. 2017. “UN Office of the High Commissioner on Human Rights, Mandate on the Special Rapporteur on the Right to Privacy”. OL JPN 3/2017.
- Oser, Jennifer, Marc Hooghe, and Sofie Marien. 2013. “Is Online Participation Distinct from Offline Participation? A Latent Class Analysis of Participation Types and their Stratification”. *Political Research Quarterly* 66 (1): 91–101.
- Penney, Jonathon W. 2016. “Chilling Effects: Online Surveillance and Wikipedia Use”. *Berkeley Technology Law Journal* 31:117.
- Pierce, Lamar, and Parasuram Balasubramanian. 2015. “Behavioral Field Evidence on Psychological and Social Factors in Dishonesty and Misconduct”. *Current Opinion in Psychology* 6:70–76.
- Roberts, Margaret E. 2018. *Censored: Distraction and Diversion Inside Chinas Great Firewall*. Princeton: Princeton University Press.
- Robinson, Darrel, and Marcus Tannenbergh. 2018. “Self-censorship in Authoritarian states: Response Bias in Measures of Popular Support in China”. V-Dem Working Paper, nr. 2018-66.
- Schenker, Nathaniel, and Jane F. Gentleman. 2001. “On Judging the Significance of Differences by Examining the Overlap Between Confidence Intervals”. *The American Statistician* 55 (3): 182–186.
- Schlozman, Kay Lehman, Sidney Verba, and Henry E Brady. 2010. “Weapon of the Strong? Participatory Inequality and the Internet”. *Perspectives on Politics* 8 (2): 487–509.

- Schwartz, Paul M. 1999. "Privacy and Democracy in Cyberspace". *Vand. L. Rev.* 52:1609–1702.
- Scott, James C. 1985. *Weapons of the Weak: Everyday Forms of Peasant Resistance*. New Haven: Yale university Press.
- Scott, James C. 1990. *Domination and the Arts of Resistance: Hidden Transcripts*. New Haven: Yale University Press.
- Stoycheff, Elizabeth. 2016. "Under Surveillance: Examining Facebook's Spiral of Silence Effects in the Wake of NSA Internet Monitoring". *Journalism & Mass Communication Quarterly* 93 (2): 296–311.
- Tannenberg, Marcus. 2018. "The Autocratic Trust Bias: Politically Sensitive Survey Items and Self-censorship". V-Dem Working Paper, nr.2017-49.
- Trochim, William M.K., and James P. Donnelly. 2008. *Research Methods Knowledge Base*. 3rd. Mason, OH: Atomic Dog.
- Tufekci, Zeynep. 2017. *Twitter and Tear Gas: The Power and Fragility of Networked Protest*. New Haven: Yale University Press.
- Ullrich, Peter, and Philipp Knopp. 2018. "Protesters' Reactions to Video Surveillance of Demonstrations: Counter-Moves, Security Cultures, and the Spiral of Surveillance and Counter-Surveillance". *Surveillance & Society* 16 (2): 183–202.
- Watson, Hayley, Rachel L Finn, and David Barnard-Wills. 2017. "A Gap in the Market: The Conceptualisation of Surveillance, Security, Privacy and Trust in Public Opinion Surveys". *Surveillance & Society* 15 (2): 269–285.
- Weyland, Kurt. 2012. "The Arab Spring: Why the Surprising Similarities with the Revolutionary Wave of 1848?" *Perspectives on Politics* 10 (4): 917–934.

Figures

Figure 1: *Preferences by survey experiment version*

Bar heights represent the proportion of respondents who expressed the specified preference in response to an opinion question. Error bars are 95% confidence intervals. $n = 4514$.

Figure 2: *Effect of surveillance treatments on preference falsification (V1)*

Using survey experiment V1, the dependent variable is the probability of expressing a given political preference. Plotted points indicate estimated coefficients (average treatment effects) and solid lines denote 95 percent confidence intervals calculated from HC2 robust standard errors. Estimated via linear probability models. Regressions include the following mean-centered covariates: age, education, gender, income, sex, attention to politics, and satisfaction with politics. $n = 1448$.

Figure 3: *Effect of surveillance treatments on self-censorship (V2)*

Using survey experiment V2, the dependent variable is the probability of selecting “no opinion” in response to a political preference question. Plotted points indicate estimated coefficients (average treatment effects) and solid lines denote 95 percent confidence intervals calculated from HC2 robust standard errors.

Estimated via linear probability models. Regressions include the following mean-centered covariates: age, education, gender, income, sex, attention to politics, and satisfaction with politics. $n = 1562$.

Figure 4: *Effect of surveillance treatments on opting-out (V3)*

Using survey experiment V3, the dependent variable is the probability of disagreeing to continue (exiting the survey) after a surveillance treatment. Plotted points indicate estimated coefficients (average treatment effects) and solid lines denote 95 percent confidence intervals calculated from HC2 robust standard errors. Estimated via linear probability models. Regressions include the following mean-centered covariates: age, education, gender, income, sex, attention to politics, and satisfaction with politics. $n = 1504$.