Anti-corruption Campaigns and Popular Support for Authoritarian Governments: A Survey Experiment in China[*]

Yaoyao Dai[†]
New York University Abu Dhabi

ABSTRACT

Do anti-corruption campaigns increase popular support for authoritarian governments? Extending existing theoretical accounts of government performance, I argue that support depends in a conditional way on the perceived motivation and effectiveness of the anti-corruption campaign. I use a novel survey experiment in China to test my argument. As predicted, effective anti-corruption campaigns increase support for the government, but not if they are seen as a strategic move to purge political rivals. Due to possible preference falsification, anti-corruption campaigns perceived to be targeted at political rivals increase *reported* support for the government but not necessarily *actual* support. Overall, my results suggest that anti-corruption campaigns can help solve both intra-elite and mass-elite conflicts in authoritarian regimes. My paper contributes to the literature examining how authoritarian leaders use institutions and policies to consolidate their hold on power. It also has important implications for research on government support in both authoritarian and democratic settings.

[*]Note: I thank Jessamyn Abel, Molly Ariotti, Charles Crabtree, Kostanca Dhima, Matt Golder, Sona Golder, and Boliang Zhu, as well as audiences at the 2016 Annual Meeting of the American Political Science Association and the 2016 University of Pennsylvania Symposium on Contemporary China for their helpful comments on this paper. The data, codebook, and R code necessary to replicate the results and figures in this analysis will be made publicly available on my dataverse site and on my personal webpage at http://daiyaoyao.com on publication. The online survey experiment described in this paper received IRB approval (Study #00004159) at Pennsylvania State University.

[†]Postdoctoral Associate, Social Science Division, NYU Abu Dhabi (yaoyao.dai@nyu.edu).

1 Introduction

On June 11, 2015, Zhou Yongkang was sentenced to life in prison after being convicted of abusing power, accepting bribes, and revealing state secrets. In addition to being China's chief of domestic security, he had also served on China's highest decision-making body, the Standing Committee of the Politburo. The verdict, along with a video of his guilty plea, was announced on China Central Television's Daily News the same day. This court ruling came three years after his aides and close associates were first put under investigation (Xinhua News, 2015*a*; Wu, 2015; Jiang and Hanna, 2015). Although Zhou Yongkang is the highest-ranked person to be jailed for corruption in China's history, he is only one of numerous high-ranking officials, many of whom had once been considered untouchable, to be investigated and jailed in the anti-corruption campaign launched by President Xi Jinping in 2012. When announcing his anti-corruption campaign, Xi Jinping vowed to crack down not only on 'tigers' (high ranking officials), but also on 'flies' (low-level civil servants). In the first five years, over 350,000 officials were investigated and arrested on corruption-related charges.[1] There appears to be no end in sight for President Xi Jinping's anti-corruption campaign.

Anti-corruption campaigns, in which leading political figures espouse harsh campaign-style rhetoric against corruption and in which there are exhaustive reports on anti-corruption efforts and numerous officials are arrested, are common in authoritarian regimes (Wedeman, 2005). The anti-corruption campaign in China, while striking in its severity, is but one among many that have occurred in authoritarian regimes across the world. Andropov, Brezhnev, Khrushchev, and Gorbachev all launched anti-corruption campaigns when they were leader of the Soviet Union (Tarkowski, 1989; Anderson and Boettke, 1993; Duhamel, 2010). Current Russian President Vladamir Putin launched his own anti-corruption campaign at almost the same time as Xi Jinping in China (Krastev and Inozemtsev, 2013). More recently, Crown Prince Mohammed bin Salman launched an anti-corruption campaign in Saudi Arabia at the end of 2017 (Aleem, 2017). Similar anti-corruption campaigns have also occurred in other places such as Fiji (Larmour, 2012), Georgia (Darden, 2008), Hong Kong (World Bank, 1997), South Korea (Zhang and Lavena, 2015), and Vietnam (Meyers, 2018).

The last fifteen years have seen considerable research examining the institutions and policies authoritarian leaders adopt to consolidate their hold on power (Geddes, 2003; Lust-Okar, 2005;

[1]This number is based on information contained in the 2017 report of the Central Commission for Discipline Inspection (CCDI, 2017). Other sources put the number of punished party members at over one million (BBC, 2018).

Magaloni, 2006; Gandhi, 2008; Slater, 2010; Svolik, 2012; Blaydes, 2013; Manion, 2015). This research, though, has largely ignored the use of anti-corruption campaigns. This is somewhat surprising given the prevalence of these types of campaigns. What research there is tends to focus on *why* authoritarian leaders initiate anti-corruption campaigns. Some scholars argue that anti-corruption campaigns are introduced to keep corruption from reaching levels that might endanger economic performance (Manion, 2009; Wedeman, 2005) and build popular support by signaling that the government is responsive to citizen concerns (Gillespie and Okruhlik, 1991). Others claim that these campaigns are best understood as an efficient way for authoritarian leaders to eliminate political rivals and consolidate their hold on power (Zhu and Zhang, 2017; Jiang and Xu, 2015).

Rather than examine the causes of anti-corruption campaigns, I focus here on their consequences. In particular, I investigate whether and how anti-corruption campaigns increase popular support for authoritarian governments. Authoritarian leaders frequently claim they are launching anti-corruption campaigns in response to concerns raised by the people and to build popular support. On launching his anti-corruption campaign ahead of the 2015 presidential elections in Belarus, for example, President Alexander Lukashenko said that his government would become unacceptable to the people unless corruption was significantly curbed (Naviny.by, 2014). Similarly, in his January 2016 speech to the Discipline Inspection Committee in China, Xi Jinping stated that the primary goal of his anti-corruption campaign was to build public support for the government (The Beijing News, 2016). The media also often link anti-corruption campaigns with support for the government. For example, the BBC (2018) recently claimed that President Xi Jinping's anti-corruption campaign "has helped his popularity." But do anti-corruption campaigns actually increase popular support for authoritarian governments, and if so, how?

There is little systematic research addressing this question and the empirical evidence is largely mixed. On the one hand, Zhang and Lavena (2015) report that the anti-corruption campaign in South Korea during the 1970s proved popular, with 79.4% of respondents in a 1979 survey believing it had reduced corruption levels. On the other hand, Holmes (1993) argues that Gorbachev's anti-corruption campaign in the 1980s lowered citizen confidence in the government and accelerated the collapse of the Soviet Union. Similarly, Wang and Dickson (2018) find that support for the Chinese central government is negatively correlated with the number of anti-corruption cases in a province. Anecdotal evidence suggests that support for authoritarian governments may depend on the perceived motivation for the anti-corruption campaigns. For example, Morris and Klesner (2010) find that anti-corruption campaigns in Mexico had no impact on public support for the government as they were not considered a credible attempt to lower corruption. Along the same lines, Manion (2009) notes that citizens in Hong Kong were initially skeptical about the underlying motivation and credibility of the anti-corruption campaign that occurred during the 1970s.

In this article, I argue that whether anti-corruption campaigns increase popular support for authoritarian governments depends on the interaction between (i) citizen perceptions of the motivation behind the campaign and (ii) citizen perceptions of the campaign's effectiveness at reducing corruption. As the existing literature suggests, authoritarian leaders are motivated to implement an anti-corruption campaign either to purge their political rivals or because they wish to reduce levels of corruption and appear responsive to citizen preferences. Citizens generally like anti-corruption campaigns that are effective are reducing corruption. The motivation behind an anti-corruption campaign matters because it speaks to the likelihood that any reduction in corruption will persist into the future. Popular support for the government will be greatest when the government is perceived to be both motivated to lower corruption and effective at doing so. It follows that anti-corruption campaigns that are perceived to be effective at controlling corruption should always increase popular support for authoritarian governments, but less so if the campaign is perceived to be a vehicle for eliminating rivals. Conversely, it also follows that anti-corruption campaigns that are perceived to be targeting political rivals should always reduce popular support for the government, and more so if the campaign is perceived to be effective at reducing corruption.

This theoretical framework has two notable features. First, it recognizes that when citizens evaluate government performance, they take into account not only the outcomes produced by government policies but also the government's motivation in implementing those policies. Empirical studies that examine the determinants of government support often focus on economic performance (Lewis-Beck and Stegmaier, 2000; van der Brug, van der Eijk and Franklin, 2007; Nadeau, Lewis-Beck and Bélanger, 2013). To a large extent, these studies implicitly assume that economic outcomes are all that matter, not how or why those outcomes are produced. Second, my theoretical framework emphasizes the importance of citizen *perceptions* when it comes to evaluating authoritarian governments. In authoritarian regimes, particularly those that maintain a strict system of censorship, citizens do not typically have access to reliable information about policy outcomes or the motivation of the government. That perceptions may not be closely tied to objective reality in authoritarian regimes is important because it means that a researcher can manipulate the provision of information and realistically expect to shape citizen perceptions.

I test the implications of my theory using a novel online survey experiment related to the current anti-corruption campaign in China. In the experiment, I provide subjects with information to manipulate their perceptions about both the effectiveness of the anti-corruption campaign at reducing corruption and the motivation of the government in pursuing an anti-corruption campaign. In line with my theoretical expectations, the level of popular support for the Chinese government depends on the interaction between citizen perceptions of the motivation and effectiveness of the anti-corruption campaign. Popular support for the government increases when the anti-corruption campaign is perceived as effective at reducing corruption. As predicted, this positive effect declines and, indeed, disappears if citizens perceive that the campaign is primarily motivated by a desire to eliminate political rivals rather than reduce corruption. Contrary to my theoretical expectations, though, I find that support for the government *increases* when citizens perceive the campaign to be a vehicle for eliminating political rivals.

One potential explanation for why government support increases when the anti-corruption campaign is perceived as targeting political rivals has to do with preference falsification. Those citizens who believe that Xi Jinping is using the anti-corruption campaign to purge opponents may well be concerned with the repressive capacity of the state and thus exaggerate their support for the government. My survey experiment allows me to test this possibility as it includes questions that let me identify individuals who are likely to falsify their preferences based on their reported level of trust in the government. In line with the preference falsification story, I find that the perception that the anti-corruption campaign is targeted at political rivals *increases* support for the government among those identified as likely to falsify their preferences but *reduces* support for the government, as I originally predicted, among those identified as unlikely to falsify their preferences. These results are consistent with those found in other authoritarian regimes. For example, García-Ponce and Pasquale (2015) find that reported trust in the state in Zimbabwe is *higher* among citizens who have experienced repression than those who have not. These results also highlight the importance of distinguishing theoretically between *reported* levels of support for authoritarian governments and *actual* levels of support (Jiang and Yang, 2016).

2 Theory

Whether anti-corruption campaigns increase popular support for authoritarian governments depends jointly on whether citizens think the campaign is effective at reducing corruption *and* on the reasons why they think the government implemented the campaign.

2.1 Effectiveness

In a broad sense, the existing literature argues that incumbent government support is tied to government performance — incumbent support is expected to be high when the government performs well. In the retrospective voting literature, for example, the support for the incumbent government is often tied to economic outcomes, which are, at least partially, attributed to the incumbent's performance in office (Lewis-Beck and Stegmaier, 2000). While most studies of popular government support have focused on rich democracies, there is growing evidence that performance also matters for government support in less developed and more authoritarian settings. Wright and Stein (2010), for example, find that economic outcomes influence individual-level support for leaders in nineteen non-democracies. Rosenfeld (2018) finds that economic growth and perceptions of economic performance play an important role in how individuals evaluate regional governments in Russia. There is also evidence that economic performance affects public support for the government in China (Yang and Tang, 2010; Lewis-Beck, Tang and Martini, 2014).

Although research on incumbent government support, in both democratic and authoritarian settings, has typically focused on economic outcomes, support for governments should depend on their performance in any policy area that citizens care about. One such area is corruption. Corruption is a worldwide concern (Transparency International, 2017; Anderson and Tverdova, 2003). For example, a 2014 survey by the Pew Research Center in 34 countries across the Middle East, Asia, Latin America, Africa, and Eastern Europe finds that corruption was considered the second largest problem facing people, after crime. On average, 76% of people claimed that corruption was a 'very big problem' (Fisman and Golden, 2017, 147). A more recent survey in China finds that corrupt officials are viewed as the biggest challenge to the country, more problematic than even pollution and economic inequality (Wike and Stokes, 2016). All of this suggests that anti-corruption campaigns that are effective at reducing corruption should increase popular support for the government.

Support for incumbent governments depends on how citizens *perceive* government performance. There is debate in the existing literature as to how accurately citizen perceptions of government performance align with actual performance. Some economic voting scholars, for example, argue that citizen perceptions of the economy are distorted by things like partisan bias (Evans and Andersen, 2006; Anderson, 2007). Other scholars, though, argue that citizen perceptions are closely tied to economic reality and that objective conditions drive voter behavior (Lewis-Beck, Nadeau and Elias, 2008; Nadeau, Lewis-Beck and Bélanger, 2013; Lewis-Beck, Martini and Kiewiet, 2013).

The connection between perceptions of government performance and actual government performance is likely to be weaker in authoritarian countries as citizens often lack reliable information about objective conditions due to government censorship (Hollyer, Rosendorff and Vreeland, 2015) and data manipulation (Magee and Doces, 2015; Wallace, 2016). Citizens in authoritarian countries are also often suspicious of official statistics (Rosenfeld, 2018). When it comes to corruption, citizens do not always have much direct experience with corruption (Zhu, Lu and Shi, 2013) and studies find that citizen (and expert) perceptions rarely reflect actual levels of corruption (Olken, 2009; Treisman, 2007). From a theoretical perspective, the key point here is that we would expect support for the incumbent government to increase if citizens *think* anti-corruption campaigns are effective at reducing corruption irrespective of whether corruption levels actually decline.

2.2 Motivation

Research on government support typically assumes that citizens only care about the effectiveness of government policies. Citizens are expected to look favorably on the government when, say, unemployment is low, inflation is low, or economic growth is high (Powell and Whitten, 1993). They are not thought to care about how or why these outcomes are produced. In what follows, though, I argue that citizens care not only about the effectiveness of government policies but also about the motivation behind those policies.

One area in which the motivation of the government is sometimes raised is political business cycles. In a political business cycle, the government actively manipulates the economy to engineer a short-term economic high to coincide with an election (Nordhaus, 1975; MacRae, 1977). Traditionally, citizens are assumed to assess governments based on the economic outcomes they produce at election time, and are not expected to take account of either why those outcomes are produced or the future consequences of the government's economic policies. These assumptions have been challenged by rational expectations scholars who argue that citizens should understand the motivation behind the government's expansionary pre-election economic policies as well as the expected post-election costs associated with these policies (Persson and Tabellini, 1990; Rogoff and Sibert, 1988; Cukierman and Meltzer, 1986). In effect, citizens should look unfavorably on governments they think are trying to manipulate the economy solely for their own political gain.

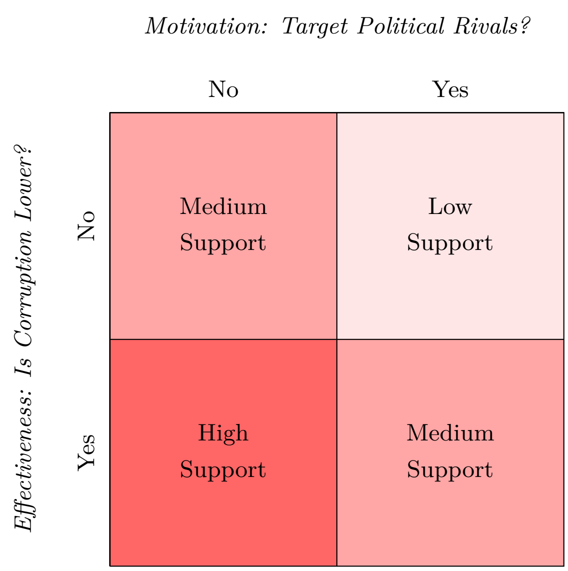
As with political business cycles, we might expect citizens to care about why anti-corruption campaigns are introduced when evaluating the government. The existing literature suggests that governments introduce anti-corruption campaigns either because they want to reduce corruption and appear responsive to citizen concerns or because they want to purge political rivals and consolidate their hold on power. Some scholars, for example, argue that anti-corruption campaigns are introduced to keep corruption levels below the point at which they might endanger economic performance and encourage mass rebellion (Wedeman, 2005; Manion, 2009). However, others argue that they are introduced as a relatively uncontroversial way for authoritarian leaders to efficiently purge political rivals (Zhu and Zhang, 2017; Jiang and Xu, 2015). Eliminating rivals on the grounds that they are corrupt is likely to generate significantly less public opposition than simply removing them for political or ideological reasons. Anti-corruption campaigns that are introduced to eliminate political rivals are an abuse of power for private gain and, hence, a form of corruption in itself. As a result, citizens are much more likely to look favorably on a government if they think it has introduced an anti-corruption campaign to reduce corruption than if they think it is simply using the campaign to strengthen its grip on power.

2.3 The Interaction between Effectiveness and Motivation

Perceptions about the effectiveness of anti-corruption campaigns should interact with perceptions about the motivation behind these campaigns to determine popular support for the government. This is because the motivation for introducing an anti-corruption campaign speaks to the likelihood that any observed reduction in corruption will persist into the future. Corruption may well decline in the short term irrespective of why the government introduces an anti-corruption campaign. Officials who are uncertain as to exactly why an anti-corruption campaign has been introduced are likely to respond initially by engaging in less corruption (Wedeman, 2005). If these officials realize over time that the campaign is driven by a desire to eliminate political rivals rather than reduce corruption, perhaps because they see who is being targeted, then they are likely to return to their previously corrupt ways. Citizens who perceive that the government is primarily motivated by a desire to eliminate rivals should understand this dynamic and expect that any immediate reduction in corruption will be relatively short lived. Indeed, anti-corruption campaigns that are targeted at political rivals may even lead to higher levels of corruption in the future as officials (and citizens) learn that the government is uninterested in cracking down on actual corruption.

The specific conditionality underlying my theory is graphically illustrated in Figure 1. Popular support for the incumbent government will be high when citizens perceive that the government is motivated to reduce corruption *and* is effective at doing so (bottom left quadrant). This is because citizens in this scenario value the reduction in corruption and expect it to persist into the future. In contrast, popular support for the government will be low when citizens perceive that the government is motivated by a desire to target political rivals *and* is ineffective at lowering corruption (top right quadrant). This is because citizens in this scenario dislike both the government's

Figure 1: Anti-Corruption Campaigns and Popular Support for Authoritarian Governments



Note: *Effectiveness* captures whether the government's anti-corruption campaign is perceived as having lowered corruption. *Motivation* captures whether the government's anti-corruption campaign is perceived as primarily targeting political rivals as opposed to reducing corruption. Cell entries denote the predicted level of popular support for the government under the four different scenarios.

abuse of power in using the anti-corruption campaign to eliminate political rivals and the fact that corruption has not declined. Popular support for the government will be moderately high when citizens perceive that the government is effective at reducing corruption but motivated by a desire to eliminate political rivals (bottom right quadrant). This is because citizens in this scenario value the reduction in corruption but don't expect the reduction in corruption to persist into the future. Popular support for the government will also be moderately high when citizens perceive that the government is motivated to reduce corruption but is ineffective at doing so (top left quadrant). In this scenario, the citizens value the fact that the government is at least trying to lower corruption, but would prefer that it was more effective at doing so.

Two conditional hypotheses can be drawn from this interactive theoretical framework (Berry, Golder and Milton, 2012). According to the *Effectiveness Hypothesis*, individuals will evaluate the government more favorably if they perceive that the anti-corruption campaign is effective at reducing corruption. This positive effect should be smaller, and may disappear entirely, if individuals believe that the anti-corruption campaign is primarily targeted at eliminating political rivals as opposed to reducing corruption. This is because any reduction in corruption levels in this scenario will be considered only a side-effect of the anti-corruption campaign and unlikely to persist into the future. According to the *Rivals Hypothesis*, individuals will always evaluate the government less favorably if they perceive that the anti-corruption campaign is primarily targeted at eliminating political rivals. The inherent symmetry of interactions means that this negative effect will be larger if individuals believe that the anti-corruption campaign has been effective at reducing corruption. This is because the government's reason for introducing the anti-corruption campaign matters more when the campaign appears to be effective at reducing corruption than when this is not the case.

> **Effectiveness Hypothesis**: Individuals will evaluate the government more favorably if they believe the anti-corruption campaign is effective at reducing corruption. This positive effect will be smaller, and may even disappear, if they believe the campaign is primarily targeted at eliminating political rivals as opposed to reducing corruption.

> **Rivals Hypothesis**: Individuals will evaluate the government less favorably if they believe the anti-corruption campaign is primarily targeted at eliminating political rivals as opposed to reducing corruption. This negative effect will be larger if they believe the campaign is effective at reducing corruption.

3 Empirics

I test my hypotheses using a novel survey experiment related to the current anti-corruption campaign in China. The survey experiment was conducted in March 2016.

3.1 Experimental Research Design

The respondents in my survey experiment come from a popular Chinese crowd-sourcing website, similar to Amazon's Mechanical Turk in the United States, that recruits and compensates agents for performing tasks.[2] Respondents were directed to an external Qualtrics survey where the experiment actually took place.[3] The experiment adopts a fully crossed 2 by 2 factorial design that explicitly matches the interactive theoretical framework illustrated in Figure 1. In the experiment, I provide respondents with information about the effectiveness and motivation of the anti-corruption campaign. The first factor has to do with the effectiveness of the campaign and takes on two values: either the anti-corruption campaign has been effective at reducing corruption or it has not been effective. The second factor has to do with the motivation of the campaign and also takes on two values: either the anti-corruption campaign is targeted at eliminating political rivals or it is not. The 2 by 2 factorial design results in four basic treatments:

1. *Effective & Rivals (ER)* Treatment: The anti-corruption campaign has been *effective* at reducing corruption and has been *targeted at political rivals*.

2. *Effective & Not Rivals (ENR)* Treatment: The anti-corruption campaign has been *effective* at reducing corruption and has *not been targeted* at political rivals.

[2] A common concern with convenience samples like this one is that they may be unrepresentative of the broader population of interest, potentially leading to experiments with low external validity (McDermott, 2011). Recent research, though, has shown that this concern is often misplaced. Several studies, for example, have found that Mechanical Turk (MTurk) samples are more diverse and representative of the overall population than in-person convenience samples (Paolacci, Chandler and Ipeirotis, 2011; Berinsky, Huber and Lenz, 2012). When it comes to many important demographic characteristics (Berinsky, Huber and Lenz, 2012), urban-rural and occupational status (Huff and Tingley, 2015), and political ideology (Clifford, Jewell and Waggoner, 2015), MTurk samples have also been found to be similar to nationally-representative samples. More importantly, several studies in the areas of political science, law, and psychology have shown that estimates of average treatment effects from MTurk samples are statistically and substantively similar to estimates based on nationally-representative samples (Berinsky, Huber and Lenz, 2012; Clifford, Jewell and Waggoner, 2015; Firth, Hoffman and Wilkinson-Ryan, 2017). The main reason for this seems to be treatment homogeneity (Coppock, Leeper and Mullinix, 2018). Almost all of this research has focused on crowd-sourcing websites like MTurk in the United States. The only study to date that has examined crowd-sourcing websites in China finds that they tend to be representative of the Chinese online population and is optimistic about the external validity of experiments using these samples (Li, Shi and Zhu, 2018).

[3] To prevent repeat participation by the same respondent, each unique account at the recruiting platform and each unique IP address were allowed to participate only once in the experiment.

3. *Not Effective & Rivals (NER)* Treatment: The anti-corruption campaign has *not been effective* at reducing corruption and has been *targeted* at political rivals.

4. *Not Effective & Not-Rivals (NENR)* Treatment: The anti-corruption campaign has *not been effective* at reducing corruption and has *not been targeted* at political rivals.

The goal of these treatments is to change respondent perceptions of the anti-corruption campaign. For perceptions to change, though, two things are necessary. First, respondents must have some uncertainty as to the effectiveness of, and motivation for, the anti-corruption campaign. If respondents have strong priors about the anti-corruption campaign, then the information treatments are likely to have little effect. Official reports about corruption cases in China are often censored and media coverage of the anti-corruption campaign is always positive, frequently stating that the campaign has been successful at reducing corruption and is not being used to eliminate rivals, (Xinhua News, 2015*b*).[4] Most citizens, though, are aware of the censorship and pro-government bias in the media, and are thus uncertain about the true success of, and motivation for, the anti-corruption campaign. While many citizens seek out news from their social networks, evidence suggests that citizens are also skeptical of this 'grapevine news' (Truex, 2014; Zhu, Lu and Shi, 2013).[5]

Second, respondents must find the information treatments credible (Acharya, Blackwell and Sen, 2018). To maximize the credibility of the treatments, respondents are told that the information is based on research from Yale University and the University of Chicago. Why? Survey experiments in China find that the source for a piece of information can influence how credible respondents find that information (Stockmann, 2013; Huang, 2015*a*). Studies have found that citizens are aware of a pro-regime bias in state-controlled media but that they are even more skeptical of foreign media (Truex, 2014). The pro-regime bias and overwhelmingly positive coverage of the anti-corruption campaign in China means that the Chinese media are not a suitable information source as they would not be considered a credible source for all of the information treatments. In particular,

[4]That reports about corruption cases are censored became evident in 2013 when censorship instructions regarding the coverage of General Xu Caihou's death when he was under investigation for corruption were leaked. In the leaked instructions, authorities requested that all media "use only Xinhua wire copy. Do not make the story headline news on the homepage. Close comment sections" (China Digital Times, 2015).

[5]Since the implementation of my experiment in 2016, it has become increasingly clear to many Chinese citizens that President Xi Jinping introduced his anti-corruption campaign to eliminate political rivals and consolidate his hold on power. Indeed, it is now widely accepted that the anti-corruption campaign laid the necessary groundwork for the March 2018 removal of presidential term limits that allows Xi Jinping to remain president for life. These term limits had been introduced by Deng Xiaoping in the 1980s to prevent the rise of another Mao Zedong.

the treatments claiming that the anti-corruption campaign has failed to reduce corruption and has been targeted at political rivals would likely be considered fake. Chinese citizens appear to have more trust in university research. In my survey, I asked respondents to indicate their level of trust in different groups of people. University professors were considered the second most trusted group after relatives.[6] Given the sensitive nature of the anti-corruption campaign, though, it is unclear whether Chinese academics would be allowed to report the information found in all four of my treatments, particularly the negative ones. This is why I use foreign universities for the information source. I chose Yale and Chicago because these universities are likely to be recognized as real and respectable by many Chinese citizens and because my four information treatments are based on the actual results found in two studies conducted at these universities.[7]

To be clear, the goal of providing an information source for the treatments in my experiment is not to evaluate 'source effects' (the source is constant across all treatments), but rather to maximize the likelihood that respondents will find the information credible and change their perceptions of the anti-corruption campaign. It is important to note that to the extent that respondents fail to find the information treatments credible or have strong prior beliefs about the anti-corruption campaign, I will not find the effects predicted by my theory. As an example of one of my information treatments, here is the exact wording for the *Effective & Not Rivals (ENR)* treatment:

> Researchers at the University of Chicago and Yale University have spent the last three years studying the anti-corruption campaign started after the 18^{th} Party Congress in China. They have collected and analyzed data from numerous sources, including the promotion track and network of officials, the experience and opinion of both experts and ordinary citizens, as well as economic and trade data. This research shows that since 2013, the anti-corruption campaign in China **has been effective** at reducing the level of corruption and it **has not been targeted** at political rivals.

Since my theory addresses how perceptions of anti-corruption campaigns *change* the level of popular government support, my dependent variable, *Change in Support*, captures the change in a respondent's level of support for the government since the start of the anti-corruption campaign. It

[6] 87% and 80% of respondents reported at least some trust in their relatives and university professors. Only 65% indicated at least some trust in the next most trusted group — central government officials.

[7] At Yale, Qian and Wen (2015) find that the anti-corruption campaign in China has been effective at reducing corruption that is easily detected by the public but ineffective at reducing corruption that is hard to observe. At Chicago, Jiang and Xu (2015) find that the anti-corruption campaign has centralized power in a small inner circle and that political rivals have been targeted. At the same time, they also find that personal ties to top leaders has not protected officials from corruption charges.

is measured on a $1-7$ scale, with 1 meaning that support for the government has greatly decreased and 7 meaning that it has greatly increased.[8] The precise survey question is:

> Since the 18^{th} Party Congress in 2012, our government has initiated a series of anti-corruption enforcements. How has your support for the government changed since the 18^{th} Congress in 2012?

Respondents are asked this question immediately after receiving one of the information treatments.

Experiments with a factorial design are intended to test conditional claims. As such, the results from these experiments are appropriately evaluated with an interaction model (Brambor, Clark and Golder, 2006; Berry, Golder and Milton, 2012). I use the following interaction specification to test the *Effectiveness Hypothesis* and the *Rivals Hypothesis*,

$$\text{Change in Support} = \beta_0 + \beta_1\text{Effective} + \beta_2\text{Rivals} + \beta_3\text{Effective} \times \text{Rivals} + \epsilon. \quad (1)$$

Effective is a dichotomous variable that equals 1 if the respondent receives information that the anti-corruption campaign is effective at reducing corruption and 0 otherwise. *Rivals* is a dichotomous variable that equals 1 if the respondent receives information that the anti-corruption campaign is targeted at political rivals and 0 otherwise. The interaction term, *Effective*×*Rivals* is included to test the conditionality of my hypotheses.

The marginal effect of *Effective* is $\beta_1 + \beta_3 Rivals$. According to the *Effectiveness Hypothesis*, β_1 should be positive as support for the government is expected to increase when the respondent believes the campaign has been effective at reducing corruption *and* is not targeted at eliminating political rivals. This positive effect is expected to be smaller and may even disappear if the respondent believes the campaign is targeted at political rivals. Thus, β_3 should be negative and $\beta_1 + \beta_3$ should be greater than or equal to 0. The marginal effect of *Rivals* is $\beta_2 + \beta_3 Effective$. According to the *Rivals Hypothesis*, β_2 should be negative as support for the government is expected to decrease when the respondent believes the campaign is targeted at eliminating rivals *and* is not effective at reducing corruption. This negative effect is expected to be larger if the respondent believes the campaign is effective at reducing corruption. Thus, both β_3 and $\beta_2 + \beta_3$ should be negative.

[8] An advantage of focusing on *change* in support for the government in an authoritarian regime like China is that respondents may be more willing to indicate that their support has declined as opposed to that their support is low. Rather than simply asking respondents whether their support for the government has declined, remained the same, or increased, the seven-point scale allows respondents to indicate that their support has declined a little, potentially also increasing the chances that a respondent would be comfortable reporting a decline in support for the government.

3.2 The Sample and Potential Threats to Internal Validity

Although the four information treatments were randomly assigned, there are two potential threats to the random assignment assumption: (i) non-random attrition and (ii) non-random ignorance of the treatment. Consider non-random attrition first. The respondents in the survey experiment volunteered to participate in the study and could therefore quit the survey at any point. The random treatment assumption holds if respondents quit at random. Bias is introduced, though, if some of the treatments cause respondents to quit the study early (Zhou and Fishbach, 2005). This was not the case here. Of the 609 individuals who began the survey experiment, only nine (1.5%) dropped out *after* receiving their treatment but before registering their change in support for the government. The fact that these nine respondents received different information treatments suggests that treatment assignment and attrition were unrelated.[9] 41 (6.7%) respondents dropped out *before* receiving their treatment, meaning that 559 respondents completed the survey.

Although the four treatments were randomly assigned, it is possible that the respondents in the different treatment groups could still differ demographically. Table 1 provides information on the balance of demographic information across the four treatment groups. The *Not Effective & Not Rivals (NENR)* treatment group is used as the baseline for conducting difference-in-means tests. As Table 1 indicates, the four treatment groups are mostly balanced demographically. Among the seven demographic characteristics evaluated in the survey, difference-in-means tests indicate that there is a slight imbalance in *Income* and *Student* between the group receiving the *Not Effective & Rivals (NER)* treatment and the group receiving the *Not Effective & Not Rivals (NENR)* treatment ($p < 0.1$). As a result, I control for *Income* and *Student* in some of the upcoming empirical analyses.

Now consider non-random ignorance of the treatment. Some respondents may be inattentive and fail to read the information treatment carefully. If this is the case, the respondents do not actually 'receive' the treatment assigned to them. No threat to the internal validity of the experiment occurs if the respondents who miss the information treatment do so at random. However, if there are some factors or characteristics common to this group of respondents that make them less attentive *and* cause them to exhibit different levels of regime support than others, then the

[9] To be specific, two respondents received the ENR treatment, three received the ER treatment, three received the NENR treatment, and one received the NER treatment.

Table 1: Demographic Balance across Treatment Groups

	Group Means				
	NENR	*ENR*	*ER*	*NER*	Sample
Female	0.45 (0.50)	0.42 (0.50)	0.41 (0.50)	0.48 (0.50)	0.44 (0.50)
Age	27.61 (6.10)	27.65 (7.07)	28.06 (7.64)	28.08 (6.50)	27.85 (6.84)
Education	6.26 (1.04)	6.09 (1.34)	6.27 (0.98)	6.36 (1.15)	6.25 (1.13)
Urban	0.51 (0.50)	0.46 (0.50)	0.55 (0.50)	0.49 (0.50)	0.50 (0.50)
Income	1.92 (1.12)	1.88 (1.30)	1.94 (1.26)	2.21* (1.42)	1.99 (1.28)
Government Employee	0.11 (0.31)	0.12 (0.33)	0.13 (0.33)	0.12 (0.33)	0.12 (0.33)
Student	0.13 (0.33)	0.19 (0.40)	0.13 (0.34)	0.21* (0.41)	0.16 (0.37)
N	142	131	144	142	559

$^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$ (two-tailed)

Note: Table 1 indicates the means for different demographic variables across the treatment groups and the sample as a whole; standard deviations are shown in parentheses. *NENR* refers to the *Not Effective & Not Rivals* treatment, *ENR* refers to the *Effective & Not Rivals* treatment, *ER* refers to the *Effective & Rivals* treatment, and *NER* refers to the *Not Effective and Rivals* treatment. *NENR* is treated as the baseline group for conducting difference-in-means tests. Welch's *t*-test, which allows for unequal variances, was used for the difference-in-means tests. All demographic information was gathered prior to the information treatments.

results will be biased. To distinguish between those who did not 'receive' a treatment and those who plausibly did, and to be able to test whether these two groups are systematically different, I included a validation check question (Kane and Barabas, 2019).[10] Of the 559 respondents who completed the survey, 383 (68.5%) passed the validation check and 176 (31.5%) failed it. As Table 2 indicates, there is little difference in the demographic characteristics of those who passed the validation check and those who failed the validation check. Difference-in-means tests indicate that those who passed the validation check are slightly older (28.35) than those who failed it (27.11). As a result, I control for *Age* in some of the upcoming empirical analyses. As Figure 2 indicates, there is also little difference in the geographic distribution of respondents in the full sample (panel

[10]Specifically, respondents were asked to identify "Which of the following statements describes the study conducted by the researchers at the University of Chicago and Yale University?" Only one of the provided statements referred to the anti-corruption campaign. Respondents who failed to recognize that the study mentioned in the information treatment had to do with the anti-corruption campaign were coded as having failed the validation check.

Table 2: Validation Check: Pass and Fail

	Group Means		
	Passed Validation	Failed Validation	Difference (t-statistic)
Female	0.42 (0.49)	0.47(0.50)	-0.05 (-1.20)
Age	28.30 (7.45)	26.86 (5.11)	1.44*** (2.64)
Education	6.26 (1.12)	6.21 (1.17)	0.05 (0.60)
Urban	0.52 (0.50)	0.45 (0.50)	0.07 (1.12)
Income	2.04 (1.30)	1.88 (1.23)	0.16 (1.16)
Government Employee	0.11 (0.33)	0.13 (0.33)	-0.02 (0.73)
Student	0.16 (0.37)	0.18 (0.38)	-0.02 (-0.60)
N	383	176	

$^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$ (two-tailed)

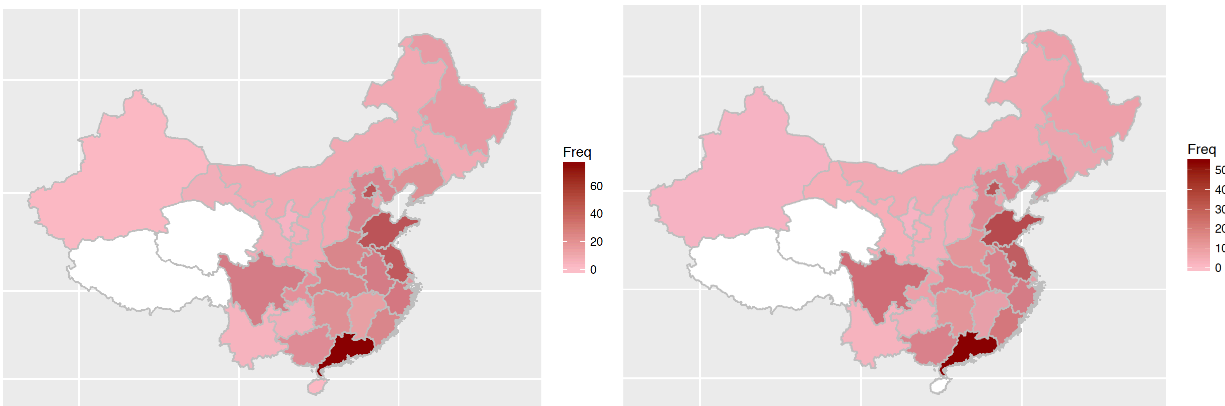
Note: The first two columns in Table 2 indicate the means for different demographic variables depending on whether the respondent passed the validation check (*Passed Validation*) or failed the validation check (*Failed Validation*); standard deviations are shown in parentheses. The third column indicates the differences in these two means and the t-statistics associated with these differences. Welch's t-test, which allows for unequal variances, was used for the difference-in-means tests. All demographic information was gathered prior to the information treatments.

a) and that of respondents who passed the validation check (panel b). In order to address any remaining concern with bias produced by the non-random ignorance of the treatments, I present results in my upcoming empirical analyses from models that include only those respondents who passed the validation check as well as from models that include the full sample.[11]

[11] One concern with looking at results from models that include only those respondents who passed the validation check has to do with possible post-treatment bias (Montgomery, Nyhan and Torres, 2018). Post-treatment bias arises if the treatment a respondent receives affects the probability that she passes the validation check. Results shown in Online Appendix A indicate that none of the information treatments have a statistically significant effect on whether a respondent passes the validation check. Indeed, the estimated effect of the treatments is substantively close to zero. Aranow, Baron and Pinson (2015, 4) suggests that a problem might still remain if "the types of subjects who fail [the validation check] under one treatment [are] not …the same as those who fail under a different treatment." Tests indicate that there is no statistical difference, at least with respect to demographic characteristics, between those who fail the validation check under the different treatments. Of course, this does not mean that differences might not still exist with respect to unobservables. This is why I present results from models that include the full sample of respondents as well as models that include only those respondents who passed the validation check. The coefficients from the two types of models are not significantly different.

Figure 2: Geographic Distribution of Respondents by Province

(a) Full Sample | (b) Passed Validation



Note: Figure 3 shows the geographic distribution of respondents across provinces. Panel (a) refers to the full sample, whereas panel (b) refers only to those respondents who passed the validation check. With the exception of Tibet and Qinghai, respondents from each of the 22 provinces and the four directly controlled municipalities (Beijing, Tianjin, Shanghai, Chongqing) participated in the survey experiment.

3.3 Results

As an initial test of my hypotheses, consider Figure 3, which shows how *Change in Support* varies across the different treatment groups. The shaded cells report the mean level of *Change in Support* in each of the four treatment groups; standard deviations are shown in parentheses.[12] The *Difference in Means* row indicates how *Change in Support* differs (across rows) depending on whether the respondents perceive the anti-corruption campaign to be effective or not at reducing corruption; *t*-statistics from difference-in-means tests are shown in parentheses. The *Difference in Means* column indicates how *Change in Support* differs (across columns) depending on whether the respondents perceive the anti-corruption campaign to be targeted at political rivals or not.

There is strong support for the *Effectiveness Hypothesis.* As predicted, perceiving the anti-corruption campaign to be effective at reducing corruption significantly increases the level of support for the central government when the campaign is not targeted at political rivals. This is indicated by the positive and statistically significant difference reported in the first cell of the *Difference in Means* row. This positive effect declines, and indeed disappears, when the anti-corruption campaign is targeted at political rivals. This is indicated by the substantively small and statistically

[12]Recall that *Change in Support* is measured on a $1-7$ scale, with 1 meaning that the support for the central government has greatly decreased and 7 meaning that support for the central government has greatly increased.

Figure 3: An Initial Test of the Hypotheses

Effectiveness: Is Corruption Lower?	*Motivation: Target Political Rivals?* No	*Motivation: Target Political Rivals?* Yes	**Difference in Means** (Rivals − Non-Rivals)
No	4.92 (1.28)	5.33 (1.37)	**0.41**** (**1.94**)
Yes	5.26 (1.09)	5.32 (1.15)	**0.06** (**0.34**)
Difference in Means (Effective − Not Effective)	**0.34*** (**1.94**)	**−0.01** (**0.08**)	

Note: The shaded cells report the mean level of *Change in Support* in the four treatment groups based on those respondents who passed the validation check; standard deviations are included in parentheses. *Change in Support* is measured on a $1-7$ scale, with 1 meaning that support for the government has greatly decreased and 7 meaning that support for the government has greatly increased. The *Difference in Means* row indicates how *Change in Support* differs depending on whether the anti-corruption campaign is perceived as effective or not at reducing corruption. The *Difference in Means* column indicates how *Change in Support* differs depending on whether the anti-corruption campaign is perceived to be targeted at political rivals or not; t-statistics from difference-in-means tests are shown in parentheses. Welch's t-test, which allows for unequal variances, was used for the difference-in-means tests.

insignificant difference reported in the second cell of the *Difference in Means* row.

In contrast, there is no support for the *Rivals Hypothesis*. Contrary to expectations, perceiving the anti-corruption campaign to be targeted at eliminating political rivals always increases support for the central government. This is indicated by the positive differences reported in the *Difference in Means* column. Indeed, this positive effect is substantively large and statistically significant when the anti-corruption campaign is not effective at reducing levels of corruption.

While the difference-in-means tests reported in Figure 3 provide an initial test of my hypotheses, they do not provide a direct test of the purported interaction between the perceived effectiveness and motivation of the anti-corruption campaign, nor do they control for the unbal-

Table 3: Perceptions of the Anti-corruption Campaign and Popular Support for the Government

Dependent Variable: *Change in Support* $(1-7)$				
	Passed Validation		Full Sample	
	Model 1	Model 2	Model 3	Model 4
Effective	0.510*	0.550**	0.274	0.262
	(0.267)	(0.268)	(0.215)	(0.218)
Rivals	0.705***	0.796***	0.606***	0.624***
	(0.264)	(0.271)	(0.218)	(0.224)
Effective × Rivals	−0.530	−0.645*	-0.523*	-0.551*
	(0.371)	(0.380)	(0.305)	(0.312)
Age		0.041***		0.041***
		(0.015)		(0.014)
Income		−0.080		0.003
		(0.080)		(0.067)
Student		0.012		0.179
		(0.275)		(0.229)
Observations	383	380	559	545
AIC	1188.806	1181.304	1774.616	1725.94
Log-likelihood	-585.403	-578.652	-878.308	-850.967

$^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$ (two-tailed)

Note: The dependent variable, *Change in Support*, captures the change in the respondent's level of support for the government since the start of the anti-corruption campaign in China in 2012. It is measured on a $1-7$ scale, with 1 meaning that support for the government has greatly decreased and 7 meaning that support for the government has greatly increased. Estimates are based on ordered logit models. The cutpoints from these models, which tests reveal are statistically different, are not shown. Standard errors are shown in parentheses. *Passed Validation* include only those respondents who passed the validation check, whereas *Full Sample* includes all respondents.

anced demographic variables. To better test my hypotheses, I estimate an ordered logit model.[13] The results from four slightly different specifications are shown in Table 3. The first two models are for the sample that includes only those who passed the validation check, whereas the last two models are for the full sample. Models 2 and 4 include the demographic controls mentioned earlier

[13]This involves feeding the interactive specification shown in Eq. (1) through the link function for the ordered logit model. Results are robust to simply estimating Eq. (1) by ordinary least squares.

(*Age*, *Income*, *Student*) to account for possible confounders that might influence a respondent's chance of receiving a particular treatment and her support for the government.

The results in Table 3 again provide strong support for the *Effectiveness Hypothesis*. As predicted, the coefficients on *Effective* are positive in all four models and are statistically significant for those models where I include only those respondents who passed the validation check. This indicates that respondents who think the anti-corruption campaign is effective at reducing corruption and is not targeted at political rivals have increased their support for the government. Also as predicted, the coefficient on the interaction term, *Effective* $\times$ *Rivals*, is negative in all four models and statistically significant in three of them. This indicates that the positive effect of thinking the anti-corruption campaign is effective at reducing corruption declines if the respondent also believes the campaign is targeted at political rivals. In other words, perceptions about the effectiveness and motivation of the anti-corruption campaign matter in a conditional way for determining the government's level of popular support.[14] These results are substantively meaningful. An anti-corruption campaign that is perceived to be effective at reducing corruption but not targeted at rivals increases the probability that a respondent will report an increase in her support for the government by 16% [1%, 33%]; two-tailed 95% confidence intervals are shown in square brackets.[15] A similarly effective anti-corruption campaign that is perceived to be targeted at rivals decreases the probability that a respondent will report an increase in her support for the government by a substantively small and statistically insignificant 2% [−10%, 8%]. In sum, perceiving that the anti-corruption campaign is effective at reducing corruption increases government support but only when respondents think the government is not strategically using the campaign to target its political rivals.

The results in Table 3 provide no support for the *Rivals Hypothesis*. Recall that the *Rivals Hypothesis* predicts that individuals will reduce their support for the government if they think the anti-corruption campaign is targeted at political rivals. Contrary to this prediction, the coefficient on *Rivals* is positive and statistically significant in all four of the models in Table 3. This indicates

[14]That my information treatments have a significant effect on popular support for the government indicates that they were considered credible by the respondents. Importantly, the fact that my information treatments could change respondent support for the government also indicates that Chinese citizens did not hold strong prior beliefs in 2016 about the effectiveness (and motivation) of the anti-corruption campaign.

[15]The percentage change in predicted probability is calculated using the results from Model 2 in Table 3 and for the scenario where the respondent is not a student and has mean age and income. Increased support for the government aggregates the three categories in the dependent variable where respondents indicate that their support has increased a little, increased, or greatly increased.

that respondents who think the anti-corruption campaign is targeted at political rivals *increase* their support for the government even when they think the campaign is ineffective at reducing corruption. Although the coefficient on the interaction term is negative, the effect of *Rivals* remains positive when the respondent thinks the anti-corruption campaign is effective at reducing corruption. Substantively, the percentage change in predicted probability that a respondent reports increased support for the government when she thinks the campaign is targeted at rivals is 22% [7%, 39%] for the case where the campaign is perceived as ineffective at reducing corruption and 3% [−7%, 14%] for the case where the campaign is perceived as effective at reducing corruption.

3.4 Preference Falsification and Treatment Heterogeneity

The results with respect to the *Rivals Hypothesis* are, at first glance, surprising, as they suggest that support for authoritarian governments increases when respondents think governments are using anti-corruption campaigns to target their political rivals and consolidate their hold on power.[16] One potential explanation has to do with preference falsification (Kuran, 1991) and the sensitive nature of government support in China. Fearing punishment from societal or state actors, citizens often falsify their preferences and provide socially desirable or government-preferred answers to sensitive survey items (Tourangeau and Yan, 2007). The fact that the respondents in my experiment are randomly assigned to the different information treatments means that any social desirability bias should be similar across all of the treatment groups and, thus, that the estimated treatment effects should be unbiased. However, the two treatments indicating that the anti-corruption campaign has been used to eliminate political rivals may signal or reveal the regime's willingness to repress. As a result, those respondents who receive these treatments may be more likely to falsify their preferences and report an increase (or greater increase) in support for the government. If this is correct, it would explain the positive effect of *Rivals* on support for the government in Table 3.

[16] One might think that Chinese citizens simply like a strong government that is able to eliminate its rivals and so better promote stability and efficiency. However, the results from my experiment are not consistent with a *strong government story*. First, it is not the case that respondents always increase their support for the government when they perceive that the anti-corruption campaign is targeted at rivals. As Figure 3 indicates, the *Rivals* treatment only has a positive and statistically significant effect on government support when the anti-corruption campaign is perceived to be ineffective at reducing corruption. Second, the *strong government* story would predict a positive interaction between the *Rivals* treatment and the *Effectiveness* treatment. In other words, the *Rivals* treatment should have a larger positive effect when the anti-corruption campaign is effective as opposed to ineffective at reducing corruption. As the results in Table 3 indicate, though, the coefficients on the interaction terms are negative in all four models. The bottom line is that there is little support for a *strong government story*.

Support for this line of reasoning comes from several other studies that have found a positive relationship between repression, preference falsification, and support for authoritarian governments. García-Ponce and Pasquale (2015), for example, find that reported trust in the government in Zimbabwe is higher among citizens who have experienced state repression. Similarly, Jiang and Yang (2016) find that reported support for the Chinese government increased among Shanghai citizens after a political purge. In line with the preference falsification story, Jiang and Yang (2016) find that while *reported* support for the government increased, *actual* support decreased.

My experiment also provides support for the *preference falsification story*. Prior to receiving the treatment, respondents are asked a series of questions, including one about trust in the central government. The trust question allows me to identify, at least in a rudimentary way, those respondents who are less likely to falsify their attitudes towards the government. Trust in the government is a sensitive topic in China. Given this, those who express high trust in the government could be honest supporters of the government or they could be lying. However, those expressing low trust in the government are very likely expressing their true feelings. According to the *preference falsification story*, at least some of the 'high trust' respondents who receive a *Rivals* treatment (either *ER* or *NER*) should report increased government support. However, 'low trust' respondents who receive these treatments should, in line with the original *Rivals Hypothesis*, report reduced government support. In effect, there should be treatment effect heterogeneity with respect to the two *Rivals* treatments across high and low trust respondents. This is exactly what I find.

In Table 4, I report results from four ordered logit models with the same basic specification as before. The first two models distinguish between high trust and low trust respondents, while the third presents results from a fully interactive specification that allows us to see whether the effects of the treatments are significantly different across low trust and high trust respondents (Brambor, Clark and Golder, 2006). While the first three models include only those respondents who passed the validation check, the fourth model includes the full sample. *Low Trust* is a dichotomous variable that equals 1 if the respondent reported no trust, very low trust, or low trust in the central government, and 0 otherwise. Of the 383 respondents who passed the validation check, 99 (25.8%) reported low trust in the government and 284 (74.1%) reported high trust. The results are highly supportive of the *preference falsification story*.

Table 4: Treatment Effect Heterogeneity and Trust in the Central Government

Dependent Variable: *Change in Support*				
	Passed Validation			Full Sample
	High Trust	Low Trust	Interaction	Interaction
Effective	0.639*	0.195	0.613*	0.379
	(0.327)	(0.496)	(0.322)	(0.265)
Rivals	0.997***	−0.237	0.958***	0.903***
	(0.320)	(0.498)	(0.315)	(0.262)
Effective × Rivals	−0.670	−0.399	−0.646	-0.561
	(0.438)	(0.734)	(0.433)	(0.363)
Low Trust			−0.867**	-0.726**
			(0.407)	(0.332)
Effective × Low Trust			−0.386	−0.190
			(0.604)	(0.476)
Rivals × Low Trust			−1.206**	−1.288**
			(0.602)	(0.491)
Effective × Rivals × Low Trust			0.195	−0.218
			(0.869)	(0.690)
Observations	284	99	383	557
AIC	809.9402	339.1274	1142.334	1688.532
Log-likelihood	−395.97	−160.56	−558.167	−831.266

$^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$ (two-tailed)

Note: The dependent variable, *Change in Support*, captures the change in the respondent's level of support for the government since the start of the anti-corruption campaign in China in 2012 and is measured on a $1 - 7$ scale. Estimates are based on ordered logit models. The cutpoints from these models, which tests reveal are statistically different, are not shown. *Age*, *Income*, and *Student* are controlled for in all four models but are not shown in Table 4. *High Trust* and *Low Trust* refer to respondents who reported high or low trust in the central government.

Consistent with the *preference falsification story*, the coefficient on *Rivals* is positive and statistically significant for the high trust respondents. This indicates that high trust respondents who think the anti-corruption campaign is targeted at political rivals report increased support for the government when they think the campaign is ineffective at reducing corruption. Although the coefficient on *Effective* × *Rivals* is negative (as predicted), the effect of *Rivals* remains positive when high trust respondents think the anti-corruption campaign is effective at reducing corruption. In effect, high trust respondents always report increased support for the government when they get a *Rivals* treatment. The results for the low trust respondents are quite different and, importantly, are consistent with the original *Rivals Hypothesis*. As predicted, the coefficients on *Rivals* and *Effective* × *Rivals* are both negative.[17] This indicates that low trust respondents, who are less likely to falsify their preferences, report *reduced* trust for the government when they believe the anti-corruption campaign is being used to target the government's political rivals. As expected, this difference in how high trust and low trust respondents respond to the *Rivals* treatments is statistically significant. This is indicated by the negative and statistically significant coefficients on *Rivals* × *Low Trust* in the fully interactive *Passed Validation* and *Full Sample* models. While I have argued that the *Rivals* treatments might induce preference falsification by signalling the repressive capacity of the state, the *Effective* treatments should have no such effect. As a result, we should not see a difference across high and low trust respondents with respect to the *Effective* treatments. In line with this reasoning, the coefficients on the two interaction terms that include *Effective* and *Low Trust* in the *Passed Validation* and *Full Sample* models are both statistically insignificant.[18]

Taken together, the results in Table 4 are consistent with the claim in the *Rivals Hypothesis* that individuals reduce their *true* support for the government when they perceive that the anti-corruption campaign is motivated by a desire to eliminate political rivals. Things are slightly different, though, when it comes to *reported* government support. Those who think the anti-corruption campaign is targeted at rivals are particularly aware of the government's repressive capacity and, as a result, many of them report increased, not decreased, support for the government, as this is the government-preferred response. This suggests that authoritarian governments can boost their *true* support by implementing an anti-corruption campaign that is effective at reducing corruption. Alternatively, or in addition, they can increase their *reported* support by implementing an anti-corruption campaign that targets political rivals and signals their repressive capacity. The

[17]That these coefficients are not statistically significant is not entirely surprising given that there are only 99 respondents, divided across four treatment groups, in the 'low trust' sample.

[18]I have used reported trust in the central government to distinguish between those respondents who are unlikely to have falsified their preferences (low trust) and those respondents who might have falsified their preferences (high trust). Consistent with this, the results in the *Passed Validation* and *Full Sample* models indicate that low trust respondents always report less support for the government than high trust respondents in all four treatment conditions. This is indicated by the negative coefficients on *Low Trust*, *Effective* × *Low Trust*, and *Rivals* × *Low Trust*.

benefit of this second strategy is that authoritarian governments not only get to eliminate political rivals, helping to solve intra-elite conflicts, they also get increased *reported* support, helping to solve mass-elite conflicts (Svolik, 2012). That true popular support does not increase with this second strategy is not necessarily problematic as the simple act of increasing reported support in the government is enough to make the coordination and collective mobilization of opposition groups more difficult (Kuran, 1991; Huang, 2015*b*; Hollyer, Rosendorff and Vreeland, 2015).

Conclusion

Although there is increasing research on the institutions and policies that authoritarian leaders employ to stabilize their regimes, little attention has been paid to the use of anti-corruption campaigns. This is despite the fact that these types of campaigns are common in authoritarian regimes around the world. To the extent that scholars have examined anti-corruption campaigns, they have focused on their causes – why do authoritarian leaders implement these campaigns? I take a different tack and instead address the consequences of anti-corruption campaigns. In particular, I examine whether anti-corruption campaigns increase the popular support for authoritarian governments. Authoritarian leaders frequently claim that their campaigns are designed to increase popular support by responding to citizen concerns. There is little systematic evidence, though, as to whether and why anti-corruption campaigns increase popular support for authoritarian governments.

In this paper, I argue that much depends on the perceived effectiveness and motivation of the anti-corruption campaign. Anti-corruption campaigns that are considered effective at reducing corruption should increase support for authoritarian governments but less so, or not at all, if the campaign is seen as a tool to eliminate political rivals. Empirical support for this comes from my experiment examining the anti-corruption campaign in China. Relatedly, anti-corruption campaigns that are seen as targeting political rivals should decrease popular support for authoritarian governments, and increasingly so if these campaigns are considered effective at reducing corruption. The results from my experiment in China paint a more nuanced picture here. While there is some evidence that the 'true' level of government support declines when the anti-corruption campaign is seen as politically motivated, 'reported' levels of government support actually increase. This is because anti-corruption campaigns that target political rivals signal the repressive capacity of the state, leading many citizens to falsify their true opinions and report *increased* government support. This increase in reported support for the government is likely to make it harder for opposition groups to coordinate and mobilize. Thus, by launching anti-corruption campaigns to eliminate political rivals, authoritarian leaders not only reduce the likelihood of intra-elite conflict, they also reduce the likelihood of mass-elite conflict. As such, anti-corruption campaigns can be a particularly effective strategy for promoting authoritarian survival. Among other things, my analysis highlights the importance of distinguishing, both theoretically and empirically, between levels of actual and reported support for the government in studies of authoritarian politics.

In addition to advancing our understanding of anti-corruption campaigns and authoritarian survival, my paper has broader implications for the study of government support in democracies and authoritarian regimes more generally. Existing studies of incumbent support typically assume that citizens care only about the outcomes produced by government policies. In many situations, though, citizens are likely to also care about the motivation behind a government's policies. There are at least two reasons for this. First, citizens may care intrinsically about whether the government is implementing a particular policy to promote its own goals or those of society more generally. Second, the government's motivation can speak to the likelihood that any policy outcome, such as lower corruption or better economic performance, will persist into the future. While rational expectations scholars have suggested that government motivation should be important for citizens when evaluating economic outcomes in the context of political business cycles, empirical research has yet to take this into account. My analysis is the first to empirically find that government effectiveness and motivation interact to determine incumbent government support.

References

Acharya, Avidit, Matthew Blackwell and Maya Sen. 2018. "Analyzing causal mechanisms in survey experiments." *Political Analysis* 26(4):357–378.

Aleem, Zeeshan. 2017. "Saudi Arabia's anti-corruption purge is all about life after oil." *Vox*, November 29.

Anderson, Christopher. 2007. "The end of economic voting? Contingency dilemmas and the limits of democratic accountability." *Annual Review of Political Science* 10:271–296.

Anderson, Christopher J. and Yuliya V. Tverdova. 2003. "Corruption, political allegiances, and attitudes toward government in contemporary democracies." *American Journal of Political Science* 47(1):91–109.

Anderson, Gary M. and Peter J. Boettke. 1993. "Perestroika and public choice: The economics of autocratic succession in a rent-seeking society." *Public Choice* 75(2):101–118.

Aranow, Peter M., Jonathon Baron and Lauren Pinson. 2015. "A note on dropping experimental subjects who fail a manipulation check." Unpublished manuscript, Yale University.

BBC. 2018. "China's Xi allowed to remain 'president for life' as term limits removed." March 11.

Berinsky, Adam J., Gregory A. Huber and Gabriel S. Lenz. 2012. "Evaluating online labor markets for experimental research: Amazon.com's Mechanical Turk." *Political Analysis* 20(3):351–368.

Berry, William D., Matt Golder and Daniel Milton. 2012. "Improving tests of theories positing interaction." *Journal of Politics* 74(3):653–671.

Blaydes, Lisa. 2013. *Elections and Distributive Politics in Mubarak's Egypt, 2011.* New York: Cambridge University Press.

Brambor, Thomas, William Roberts Clark and Matt Golder. 2006. "Understanding interaction models: Improving empirical analyses." *Political Analysis* 14(1):63–82.

CCDI. 2017. "王岐山在十八届中纪委三次全会上的工作报告." October 20.

China Digital Times. 2015. "Minitrue: Downplay death of Xu Caihou." March 16.

Clifford, Scott, Ryan M. Jewell and Philip D. Waggoner. 2015. "Are samples drawn from Mechanical Turk valid for research on political ideology." *Research and Politics* 4(2).

Coppock, Alexander, Thomes J. Leeper and Kevin J. Mullinix. 2018. "Generalizability of heterogeneous treatment effect estimates." Unpublished manuscript, Yale University.

Cukierman, Alex and Allen Meltzer. 1986. "A positive theory of discretionary policy, the cost of a democratic government, and the benefit of a constitution." *Economic Inquiry* 24:367–388.

Darden, Keith. 2008. "The integrity of corrupt states: Graft as an informal state institution." *Politics & Society* 36(1):35–59.

Duhamel, Luc. 2010. *The KGB Campaign against Corruption in Moscow, 1982-1987.* Pittsburgh: University of Pittsburgh Press.

Evans, Geoffrey and Robert Andersen. 2006. "The political conditioning of economic perceptions." *Journal of Politics* 68(1):194–207.

Firth, Kristin, David A. Hoffman and Tess Wilkinson-Ryan. 2017. "Law and psychology grows up, goes online and replicates." Public Law and Legal Theory Research Paper Series # 17-38, University of Pennsylvania Law School.

Fisman, Ray and Miriam A. Golden. 2017. *Corruption: What Everyone Needs to Know*. New York: Oxford University Press.

Gandhi, Jennifer. 2008. *Political Institutions under Dictatorship*. New York: Cambridge University Press.

García-Ponce, Omar and Benjamin Pasquale. 2015. "How political repression shapes attitudes toward the state: Evidence from Zimbabwe." Unpublished manuscript, New York University.

Geddes, Barbara. 2003. *Paradigms and Sand Castles: Theory Building and Research Design in Comparative Politics*. New York: Cambridge University Press.

Gillespie, Kate and Gwenn Okruhlik. 1991. "The political dimensions of corruption cleanups: A framework for analysis." *Comparative Politics* 24(1):77–95.

Hollyer, James R., B. Peter Rosendorff and James Raymond Vreeland. 2015. "Transparency, protest, and autocratic instability." *American Political Science Review* 109(04):764–784.

Holmes, Leslie. 1993. *The End of Communist Power: Anti-Corruption Campaigns and Legitimation Crisis (Europe and the International Order)*. New York: Oxford University Press.

Huang, Haifeng. 2015*a*. "International knowledge and domestic evaluations in a changing society: The case of China." *American Political Science Review* 109(3):613–634.

Huang, Haifeng. 2015*b*. "Propaganda as signaling." *Comparative Politics* 47(4):419–444.

Huff, Connor and Dustin Tingley. 2015. "Who are these people? Evaluating the demographic characteristics and political preferences of MTurk survey respondents." *Research and Politics* 2(3):351–368.

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.

Jiang, Junyan and Yan Xu. 2015. "Popularity and power: The political logic of anticorruption in authoritarian regimes." Unpublished manuscript, University of Chicago.

Jiang, Steven and Jason Hanna. 2015. "China's ex-security chief Zhou Yongkang sentenced to life for bribery." *CNN*, June 11.

Kane, John V. and Jason Barabas. 2019. "No harm in checking: Using factual manipulation checks to assess attentiveness in experiments." *American Journal of Political Science* 63(1):239–249.

Krastev, Ivan and Vladislav Inozemtsev. 2013. "Putin's self-destruction — Russia's new anti-corruption campaign will sink the regime." *Foreign Affairs*, June 9.

Kuran, Timur. 1991. "Now out of never: The element of surprise in the east European revolution of 1989." *World Politics* 44:7–48.

Larmour, Peter. 2012. Populist anti-corruption and military coup: The clean up campaign in Fiji 2006-07. In *Governments, NGOs and Anti-Corruption: The New Integrity Warriors*, ed. Luís De Sousa, Barry Hindess and Peter Larmour. New York: Routledge.

Lewis-Beck, Michael, Nicholas F. Martini and D. Roderick Kiewiet. 2013. "The nature of economic perceptions in mass publics." *Electoral Studies* 32:524–528.

Lewis-Beck, Michael S. and Mary Stegmaier. 2000. "Economic determinants of electoral outcomes." *Annual Review of Political Science* 3:183–219.

Lewis-Beck, Michael S., Richard Nadeau and Angelo Elias. 2008. "Economics, party, and the vote: Causality issues and panel data." *American Journal of Political Science* 52(1):84–95.

Lewis-Beck, Michael S., Wenfang Tang and Nicholas F. Martini. 2014. "A Chinese popularity function: Sources of government support." *Political Research Quarterly* 67(1):16–25.

Li, Xiaojun, Weiyi Shi and Boliang Zhu. 2018. "The face of internet recruitment: Evaluating the labor markets of online crowdsourcing platforms in China." *Research and Politics* 5(1):1–8.

Lust-Okar, Ellen. 2005. *Structuring Conflict in the Arab World: Incumbents, Opponents, and Institutions*. New York: Cambridge University Press.

MacRae, Duncan. 1977. "A political model of the business cycle." *Journal of Political Economy* 95:239–263.

Magaloni, Beatriz. 2006. *Voting for Autocracy: Hegemonic Party Survival and its Demise in Mexico*. New York: Cambridge University Press.

Magee, Christopher S. P. and John A. Doces. 2015. "Reconsidering regime type and growth: Lies, dictatorships, and statistics." *International Studies Quarterly* 59(2):223–237.

Manion, Melanie. 2009. *Corruption by Design: Building Clean Government in Mainland China and Hong Kong*. Cambridge: Harvard University Press.

Manion, Melanie. 2015. *Information for Autocrats: Representation in Chinese Local Congresses*. New York: Cambridge University Press.

McDermott, Rose. 2011. "New directions for experimental work in international relations." *International Studies Quarterly* 55(2):503–520.

Meyers, Jessica. 2018. "Vietnam's corruption crackdown is targeting the powerful. Has it gone too far?" *Los Angeles Times*, January 16.

Montgomery, Jacob M., Brendan Nyhan and Michelle Torres. 2018. "How conditioning on post-treatment variables can ruin your experiment and what to do about it." *American Journal of Political Science* 62(3):760–775.

Morris, Stephen D. and Joseph L. Klesner. 2010. "Corruption and trust: Theoretical considerations and evidence from Mexico." *Comparative Political Studies* 43(10):1258–1285.

Nadeau, Richard, Michael S. Lewis-Beck and Éric Bélanger. 2013. "Economics and elections revisited." *Comparative Political Studies* 46(5):551–573.

Naviny.by. 2014. "Lukashenka on the fight against corruption: Let them shoot me better or hang me than I will retreat here." July 31.

Nordhaus, William D. 1975. "The political business cycle." *Review of Economic Studies* 42:169–190.

Olken, Benjamin A. 2009. "Corruption perceptions vs. corruption reality." *Journal of Public Economics* 93(7):950–964.

Paolacci, Gabriele, Jesse Chandler and Panagiotis G. Ipeirotis. 2011. "Running experiments on Amazon Mechanical Turk." *Judgment and Decision Making* 5(5):411–419.

Persson, Torsten and Guido Tabellini. 1990. *Macroeconomic Policy, Credibility, and Politics*. London: Harwood Academic Press.

Powell, G. Bingham and Guy D. Whitten. 1993. "A cross-national analysis of economic voting: Taking account of the political context." *American Journal of Political Science* 37(2):391–414.

Qian, Nancy and Jaya Wen. 2015. "The impact of Xi Jinping's anti-corruption campaign on luxury imports in China." Unpublished manuscript, Yale University.

Rogoff, Kenneth and Anne Sibert. 1988. "Equilibrium political business cycles." *Review of Economic Studies* 55:1–16.

Rosenfeld, Bryn. 2018. "The popularity costs of economic crisis under electoral authoritarianism: Evidence from Russia." *American Journal of Political Science* 62(2):382–397.

Slater, Dan. 2010. *Ordering Power: Contentious Politics and Authoritarian Leviathans in Southeast Asia*. New York: Cambridge University Press.

Stockmann, Daniela. 2013. *Media Commercialization and Authoritarian Rule in China*. New York: Cambridge University Press.

Svolik, Milan W. 2012. *The Politics of Authoritarian Rule*. New York: Cambridge University Press.

Tarkowski, Jacek. 1989. "Old and new patterns of corruption in Poland and the USSR." *Telos* 1989(80):51–62.

The Beijing News. 2016. "习近平反腐八大新提法：加强对国家机器的监督." January 12.

Tourangeau, Roger and Ting Yan. 2007. "Sensitive questions in surveys." *Psychological Bulletin* 133(5):859.

Transparency International. 2017. "Corruption Perceptions Index 2016.".

Treisman, Daniel. 2007. "What have we learned about the causes of corruption from ten years of cross-national empirical research?" *Annual Review of Political Science* 10:211–244.

Truex, Rory. 2014. "Who believes the People's Daily? Bias and credibility in authoritarian media." Unpublished manuscript, Princeton University.

van der Brug, Wouter, Cees van der Eijk and Mark Franklin. 2007. *The Economy and the Vote: Economic Conditions and Elections in Fifteen Countries*. New York: Cambridge University Press.

Wallace, Jeremy L. 2016. "Juking the stats? Authoritarian information problems in China." *British Journal of Political Science* 46(1):11–29.

Wang, Yuhua and Bruce Dickson. 2018. "How corruption investigations undermine regime support: Evidence from China." Unpublished manuscript, Harvard University.

Wedeman, Andrew. 2005. "Anticorruption campaigns and the intensification of corruption in China." *Journal of Contemporary China* 14(42):93–116.

Wike, Richard and Bruce Stokes. 2016. "Chinese public sees more powerful role in world, names U.S. as top threat." *Pew Research Center*, October 5.

World Bank. 1997. *World Development Report.* Washington, DC: Oxford University Press.

Wright, Joseph and Elizabeth Stein. 2010. "Economic outcomes and support for authoritarian incumbents: Results from survey data." Paper presented at the Annual Meeting of the American Political Science Association, Washington D.C.

Wu, Yuwen. 2015. "Profile: China's fallen security chief Zhou Yongkang." *BBC Chinese*, October 12.

Xinhua News. 2015*a*. "周永康一审被判无期徒刑." June 11.

Xinhua News. 2015*b*. "[新华访谈] 习近平访美谈反腐的三个 "第一次"." September 25.

Yang, Qing and Wenfang Tang. 2010. "Exploring the sources of institutional trust in China: Culture, mobilization, or performance?" *Asian Politics & Policy* 2(3):415–436.

Zhang, Yahong and Cecilia Lavena. 2015. *Government Anti-corruption Strategies: A Cross-cultural Perspective.* New York: CRC Press.

Zhou, Haotian and Ayelet Fishbach. 2005. "The pitfall of experimenting on the web: How unaffected selection attrition leads to surprising (yet false) research conclusions." *Journal of Personality and Social Psychology* 111(4):493–504.

Zhu, Jiangnan and Dong Zhang. 2017. "Weapons of the powerful: Authoritarian elite competition and politicized anticorruption in China." *Comparative Political Studies* 50(9):1186–1220.

Zhu, Jiangnan, Jie Lu and Tianjian Shi. 2013. "When grapevine news meets mass media: Different information sources and popular perceptions of government corruption in mainland China." *Comparative Political Studies* 46(8):920–946.

Appendix A: Validation Check and Possible Post-Treatment Bias

In the main text, I raise the issue of non-random ignorance of treatment. Some respondents may be inattentive or fail to read the information treatment provided in the survey experiment carefully. If there are some factors or characteristics that are common to this group of respondents that make them less attentive *and* cause them to exhibit different levels of regime support than others, then my results will be biased. One way to ameliorate this concern is to focus on only those respondents who pass a validation check. My validation check question asks respondents to identify the subject of the study mentioned in their information treatment. Note that the validation check question did not ask what the results of the study were, just what the study was about. Those respondents who were unable to identify that the study concerned the anti-corruption campaign were coded as having failed the validation check. Given that those respondents who failed the validation check could not recall the subject matter of the study, it is almost certainly the case that they did not receive their specific information treatment regarding the results of the study.

One concern, though, with looking at results from models that include only those respondents who passed the validation check has to do with post-treatment bias (Montgomery, Nyhan and Torres, 2018). Post-treatment bias arises if the treatment a respondent receives affects the probability that she passes the validation check. As Table 5 indicates, there is no statistically significant difference between the proportion of respondents who passed and failed the validation check for any of the information treatments. Not only are the differences for each treatment statistically insignificant, but they are also substantively close to zero. This result is confirmed when I estimate a logit model where the dependent variable is 1 if the respondent passes the validation check and 0 otherwise. As Table 6 indicates, the coefficients on the four dichotomous variables representing the different treatments are statistically insignificant and close to zero.

Aranow, Baron and Pinson (2015, 4) suggest that a problem might still remain if "the types of subjects who fail [the validation check] under one treatment [are] not …the same as those who fail under a different treatment." This does not appear to the case here, at least with respect to observable demographic characteristics. In Table 7, I present the results from a series of logit models, one for each treatment group, where I examine how respondent demographics affect the

Table 5: Validation Check: Pass and Fail

	Group Proportions		
	Passed Validation	Failed Validation	Difference (t-statistic)
NENR Treatment	0.26 (0.44)	0.24 (0.43)	0.02 (0.57)
ENR Treatment	0.22 (0.42)	0.26 (0.44)	−0.04 (−1.00)
ER Treatment	0.26 (0.44)	0.26 (0.44)	0.00 (−0.14)
NER Treatment	0.26 (0.44)	0.24 (0.43)	0.02 (0.57)
N	383	176	

$^{*}p < 0.05$; $^{**}p < 0.01$ (two-tailed)

Note: *NENR* refers to the *Not Effective & Not Rivals* treatment, *ENR* refers to the *Effective & Not Rivals* treatment, *ER* refers to the *Effective & Rivals* treatment, and *NER* refers to the *Not Effective and Rivals* treatment.

Table 6: Information Treatment and Validation Check

	Dependent Variable: Passed Validation Check (1, 0)
ENR Treatment	−0.129
	(0.272)
ER Treatment	−0.125
	(0.264)
NER Treatment	0.095
	(0.278)
NENR Treatment	−0.660
	(0.770)
Demographics	Yes
Observations	529
Log Likelihood	−317.538

$^{*}p < 0.05$; $^{**}p < 0.01$ (two-tailed)

Note: Table 6 presents results from a logit model where the dependent variable is 1 if the respondent passes the validation check and 0 otherwise. Standard errors are shown in parentheses. The model also includes the same seven demographic variables as seen in Table 2.

Table 7: Demographics, Information Treatment, and Validation Check

	Dependent Variable: Passed Validation Check (1, 0)			
	ENR Treatment	ER Treatment	NER Treatment	NENR Treatment
Female	−0.426	0.065	−0.042	−0.362
	(0.415)	(0.388)	(0.418)	(0.392)
Age	0.030	0.031	0.022	0.099*
	(0.038)	(0.030)	(0.043)	(0.049)
Education	0.259	0.126	0.019	−0.185
	(0.169)	(0.192)	(0.196)	(0.203)
Student	0.891	−0.518	0.220	0.498
	(0.631)	(0.578)	(0.588)	(0.633)
Income	0.210	−0.062	0.199	0.041
	(0.184)	(0.176)	(0.190)	(0.199)
Urban	−0.455	0.287	0.162	0.068
	(0.462)	(0.387)	(0.448)	(0.406)
Government Employee	−0.261	0.082	−0.190	−0.620
	(0.597)	(0.579)	(0.663)	(0.608)
Observations	123	139	128	139
Log Likelihood	−73.865	−84.833	−72.849	−79.538

$^{*}p < 0.05$; $^{**}p < 0.01$ (two-tailed)

Note: Table 7 presents results from a series of logit models, one for each treatment group, where the dependent variable is 1 if the respondent passed the validation check and 0 otherwise. Standard errors are shown in parentheses.

probability of passing the validation check. Results from the fully interactive model in Table 8 indicate that the coefficients associated with each of the demographic variables in Table 7 are not statistically different across each of the four information treatments; none of the coefficients on the interaction terms in Table 8 are statistically significant.[1] Of course, the respondents who failed the

[1]The positive and statistically significant coefficient on *Age* for the *NENR* treatment in Table 7 indicates only that the effect of *Age* is significantly different from zero for the *NENR* treatment, not that it is significantly different from the effect of *Age* in the other treatments. As previously indicated, the insignificant coefficients on the interaction terms associated with *Age* from the fully interactive model in Table 8 show that there is no statistically significant difference in the effect of *Age* across the different treatments.

Table 8: Demographics, Information Treatment, and Validation Check: Interactive Specification

	Dependent Variable: Passed Validation Check (1, 0)
Female	−0.362 (0.392)
Age	0.099* (0.049)
Education	−0.185 (0.203)
Student	0.498 (0.633)
Income	0.041 (0.199)
Urban	0.068 (0.406)
Government Employee	−0.620 (0.608)
ENR Treatment	−1.191 (2.267)
ER Treatment	−0.316 (2.238)
NER Treatment	0.370 (2.420)
Female × ENR Treatment	−0.064 (0.571)
Female × ER Treatment	0.427 (0.551)
Female × NER Treatment	0.320 (0.573)
Age × ENR Treatment	−0.069 (0.062)
Age × ER Treatment	−0.069 (0.057)
Age × NER Treatment	−0.078 (0.065)
Education × ENR Treatment	0.444 (0.265)
Education × ER Treatment	0.311 (0.280)
Education × NER Treatment	0.204 (0.282)
Student × ENR Treatment	0.393 (0.893)
Student × ER Treatment	−1.016 (0.857)
Student × NER Treatment	−0.277 (0.864)
Income × ENR Treatment	0.169 (0.271)
Income × ER Treatment	−0.103 (0.266)
Income × NER Treatment	0.158 (0.275)
Urban × ENR Treatment	−0.523 (0.615)
Urban × ER Treatment	0.219 (0.560)
Urban × NER Treatment	0.094 (0.604)
Government Employee × ENR Treatment	0.359 (0.852)
Government Employee × ER Treatment	0.702 (0.840)
Government Employee × NER Treatment	0.430 (0.900)
Observations	529
Log Likelihood	−313.089

$^{*}p < 0.05$; $^{**}p < 0.01$ (two-tailed)

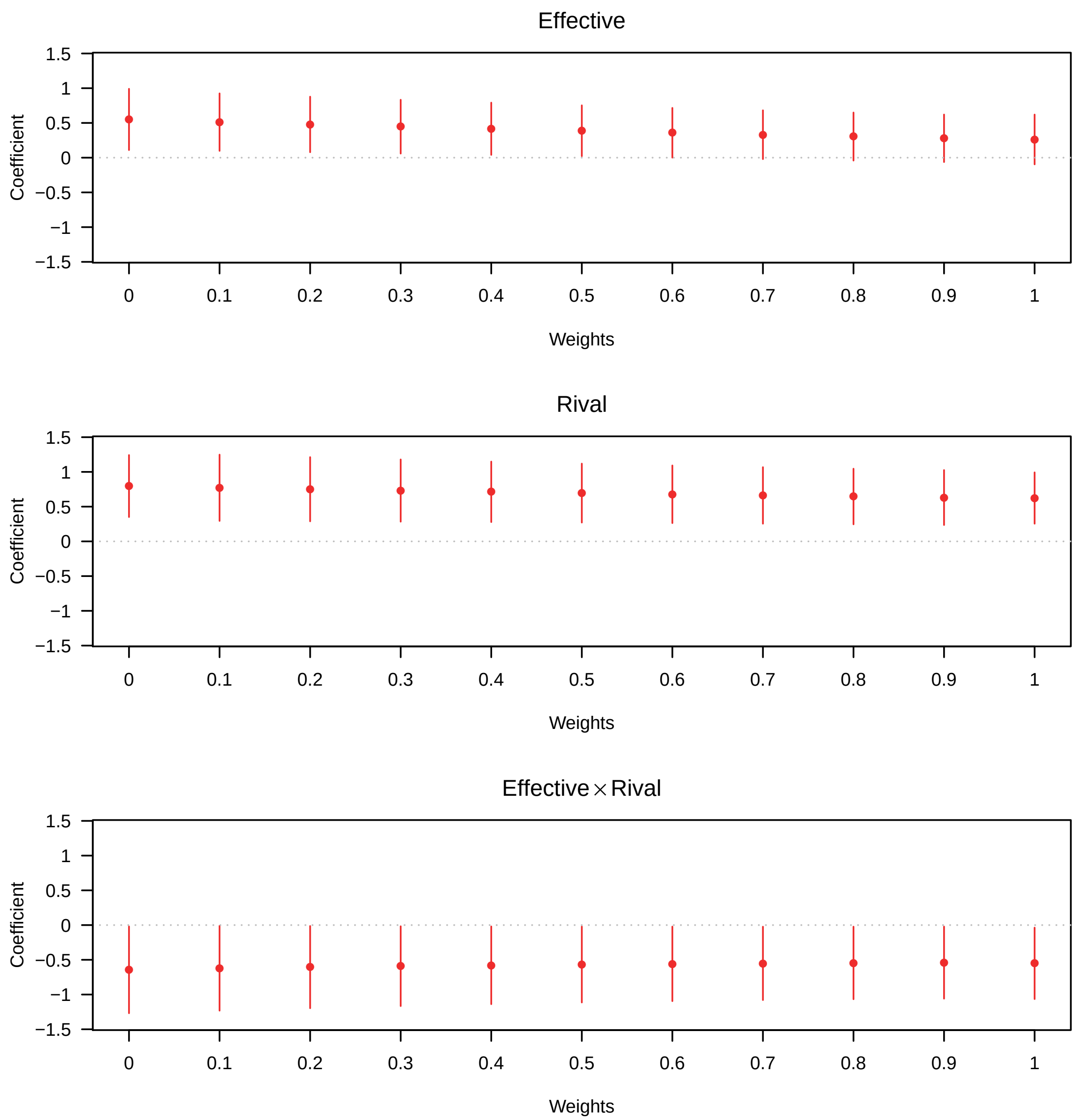
Note: Table 8 presents results from a logit model, where the dependent variable is 1 if the respondent passed the validation check and 0 otherwise. Standard errors are shown in parentheses.

validation check may differ across treatment groups for unobservable reasons. This is why I always present results in the main text from both models that include only those respondents who passed the validation check *and* models that include the full sample.

Another way to examine whether post-treatment bias is a problem is with the help of fractional pooling (Bartels, 1996). Fractional pooling involves using different weights for those respondents who failed the validation check. The *Passed Validation* model essentially weights these respondents as 0 and the *Full Sample* model weights them as 1. In Figure 4, I show how the coefficients on *Effective*, *Rivals*, and *Effective* × *Rivals* change when I include all respondents but alter the weight placed on those respondents who failed the validation check from 0 to 1. Results are for the model specification where I include *Age*, *Income*, and *Student* as controls.[2] To reiterate, the coefficient and confidence interval on the left of each panel ($Weight = 0$) are identical to those that I previously reported from the *Passed Validation* model in Table 3 where only those respondents (380) who passed the validation check are included. Similarly the coefficient and confidence interval on the right of each panel ($Weight = 1$) are identical to those that I previously reported from the *Full Sample* model in Table 3 where all 545 respondents are included and 166 of the sample failed the validation check. The coefficients and confidence intervals in between these two end points ($0 < Weight < 1$) come from models where I include those who failed the validation check but alter the weight that I give these observations in the model. One thing to note here is that the sign of the coefficients in each panel remains the same across all values of *Weight*. More important, though, is that there is no substantive or statistically significant difference in the magnitude of the coefficients as I alter *Weight* from 0 to 1. All of this suggests that we should not be too concerned with potential post-treatment bias in the *Passed Validation* models.

[2] To clarify, when *Weight* in Figure 4 is equal to 1, each of the 545 respondents is treated equally and so each receives a weight of $1/545 = 0.0018$. Of these 545 respondents, 166 (30.5%) failed the validation check. When *Weight* in Figure 4 is equal to 0.5, those respondents who failed the validation check account for $30.5\%/2 = 15.25\%$ of the sample. As a result, those respondents who failed the validation check receive a weight of $0.1525/166 = 0.0009$ and those respondents who passed the validation check receive a weight of $0.8475/380 = 0.0022$. When *Weight* in Figure 4 is equal to 0, those respondents who failed the validation check receive a weight of 0 and those respondents who passed the validation check receive a weight of $1/380 = 0.0026$. This same process, with differing weights, is used to estimate all of the points in Figure 4.

Figure 4: Models with Different Weights on Respondents Who Failed the Validation Check



Note: Figure 4 shows how the coefficients on *Effective*, *Rival*, and *Effective* × *Rival* change when I place different weights on those respondents who failed the validation check. The coefficients come from models where I control for *Age*, *Income*, and *Student*. The solid red lines represent two-tailed 90% confidence intervals. When the respondents who failed the validation check receive a weight of 0, we obtain the same results as we did earlier with the *Passed Validation* model in Table 3. When the respondents who failed the validation check receive a weight of 1, we obtain the same results as we did earlier with the *Full Sample* model in Table 3.

Online Appendix: References

Aranow, Peter M., Jonathon Baron and Lauren Pinson. 2015. "A note on dropping experimental subjects who fail a manipulation check." Unpublished manuscript, Yale University.

Bartels, Larry M. 1996. "Pooling disparate observations." *American Journal of Political Science* 40(3):905–942.

Montgomery, Jacob M., Brendan Nyhan and Michelle Torres. 2018. "How conditioning on post-treatment variables can ruin your experiment and what to do about it." *American Journal of Political Science* 62(3):760–775.