PROTESTS AS STRATEGIC GAMES: EXPERIMENTAL EVIDENCE FROM HONG KONG'S ANTIAUTHORITARIAN MOVEMENT*

DAVIDE CANTONI
DAVID Y. YANG
NOAM YUCHTMAN
Y. JANE ZHANG

Social scientists have long viewed the decision to protest as strategic, with an individual's participation a function of their beliefs about others' turnout. We conduct a framed field experiment that recalibrates individuals' beliefs about others' protest participation, in the context of Hong Kong's ongoing antiauthoritarian movement. We elicit subjects' planned participation in an upcoming protest and their prior beliefs about others' participation, in an incentivized manner. One day before the protest, we randomly provide a subset of subjects with truthful information about others' protest plans and elicit posterior beliefs about protest turnout, again in an incentivized manner. After the protest, we elicit subjects' actual participation. This allows us to identify the causal effects of positively and negatively updated beliefs about others' protest participation on subjects' own turnout. In contrast with the assumptions of many recent models of protest participation, we consistently find evidence of strategic substitutability. We provide guidance regarding plausible sources of strategic substitutability that can be incorporated into theoretical models of protests. *JEL* Codes: D74, D8, P0.

*Helpful and much appreciated suggestions, critiques, and encouragement were provided by Ned Augenblick, Doug Bernheim, Ernesto Dal Bó, Matthew Gentzkow, Peter Lorentzen, Muriel Niederle, Torsten Persson, and many seminar and conference participants. Moritz Leitner, Glen Ng, and Meggy Wan provided excellent research assistance. We thank Cathrin Mohr for providing data on East German protests. Cantoni acknowledges financial support from the LMUexcellent Junior Researcher Fund and the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant agreement no. 716837). The research described in this article was approved by the University of Munich IRB, protocol 2016-11; by the Stanford University Institutional Review Board, protocol 38481; by the University of California-Berkeley Committee for Protection of Human Subjects, protocol ID 2015-05-7571; and by the Hong Kong University of Science and Technology Human Participants Research Panel, submission 126. The experiment is registered on the AEA RCT registry with ID AEARCTR-0001423.

The Quarterly Journal of Economics (2019), 1–57. doi:10.1093/qje/qjz002.
Advance Access publication on January 21, 2019.

I. Introduction

Mass political movements have long demanded fundamental political rights. Citizens have taken to the streets from Tiananmen Square to Tahrir Square, from the women's suffrage movement to the civil rights movement, from the Velvet Revolution to Hong Kong's Umbrella Revolution. What drives individuals' decisions to participate in political protests such as these?

Strategic considerations have long been seen as crucial, with an individual's participation shaped by their beliefs about the participation of others. On the one hand, protests are a classic example of a political collective action problem: individuals have an incentive to free-ride on the costly participation of others, and may thus be less willing to turn out when they believe more others will do so, thus producing a game of strategic substitutes (Olson 1965; Tullock 1971; Palfrey and Rosenthal 1984). On the other hand, much recent theoretical work assumes strategic complementarity: this might arise because the cost of participation is anticipated to be lower when a protest is larger; because participatory utility is greater in a larger, more successful protest; or because one's utility under a postrevolution regime will be greater if one was a protest participant (see, for example, Bueno de Mesquita 2010; Edmond 2013; Passarelli and Tabellini 2017; Barberà and Jackson 2018). Indeed, in a recent review article, Gehlbach, Sonin, and Svolik (2016, 579) go so far as to affirm that strategic complementarity "*characterizes* mass protests"[1].

In this article, we identify the causal effect of beliefs about other individuals' protest turnout on one's own, conducting a framed field experiment with potential participants in an antiauthoritarian protest in Hong Kong. We study participation in a July 1 march, a yearly protest that represents an important component of Hong Kong's ongoing antiauthoritarian movement, epitomized by the recent Umbrella Revolution.[2] The July 1 march shares many essential characteristics with antiauthoritarian

1. See also Granovetter (1978), Kuran (1989, 1991, 1997), Chwe (2000), Fearon (2011), Kricheli, Livne, and Magaloni (2011), Egorov and Sonin (2018) for models of protests in which participation is a game of strategic complements. Some recent theoretical work allows for the possibility of strategic substitutability in the context of the same protest game: see Shadmehr and Bernhardt (2011) and Shadmehr (2018).

2. In Cantoni et al. (2016), we provide a complementary, descriptive study of the characteristics of supporters of Hong Kong's antiauthoritarian movement.

protests across time and space: participants aim to achieve policy concessions from an authoritarian regime by turning out in large numbers, facing the threat of government crackdown. In this context, we experimentally recalibrate individuals' beliefs about others' protest participation and study how these beliefs affect one's participation. We find consistent evidence of strategic substitutability in the decision to protest, challenging many recent models of protest participation that assume strategic complementarity.

Although much theoretical work has been done on the strategic element of the protest decision, empirical evidence on the causal effect of beliefs regarding others' protest turnout on one's own is extremely limited. Several recent articles have provided causal evidence on a bundle of "social" influences on protest participation: Enikolopov, Makarin, and Petrova (2016) present evidence that the diffusion of an online social network increased protest turnout in Russia; González (2016) provides evidence that peer participation in Chilean student protests increased one's own; and Manacorda and Tesei (2016) provide evidence that mobile phones' diffusion increased protest turnout in Africa. However, these analyses are unable to separately identify the effects of beliefs about the protest participation of others on one's own participation, instead estimating the combined effects of (i) learning about a protest's logistics (e.g., time and place), (ii) learning about the state of the world, and (iii) learning about others' protest turnout.[3]

Credibly testing for a causal effect of beliefs about others' turnout in the decision to protest against an authoritarian regime has been hindered by two empirical obstacles. First, antiauthoritarian political movements have typically been studied ex post (e.g., Kuran 1989, 1991, 1997; Opp and Gern 1993; Lohmann 1994). This not only generates selection issues—movements are generally studied after they have become large and successful—but also makes the prospective study of beliefs nearly impossible: belief elicitation would necessarily be retrospective and likely distorted by the realization of the political outcomes of interest.

Second, even when measured in real time, it is extremely difficult to exploit variation in beliefs to identify causal effects.

3. Other recent empirical work on the causes and consequences of mass political movements includes Madestam et al. (2013), Yanagizawa-Drott (2014), DellaVigna et al. (2014), and Acemoglu, Hassan, and Tahoun (2014).

Naturally occurring variation is very likely to be endogenous with respect to behaviors of interest. Experimental variation, for example, arising from an information treatment, runs into challenges from heterogeneous priors, which imply that the same information treatment can generate positive belief updating among one subset of the sample (i.e., those whose priors are below the information provided) and negative updating among another subset.[4] This means, for example, that even an effective intervention may produce average treatment effects on beliefs or behavior that spuriously appear to be null results. The average effects would simply reflect offsetting heterogeneous treatment effects of opposite signs. Thus, experimental interventions aimed at manipulating beliefs require carefully measured priors (and ideally posteriors as well) to determine exactly how the treatment affects particular individuals' beliefs, and through beliefs, behavior.

We overcome these obstacles as follows. First, we study participation within an ongoing, high-stakes political movement: Hong Kong's antiauthoritarian movement.[5] Because Hong Kong's democrats traditionally protest the rule of the Chinese Communist Party (CCP) each July 1, there exists a known protest about which we can elicit beliefs prospectively in real time. Second, using a three-part online experiment we conducted at the Hong Kong University of Science and Technology (HKUST), we are able to elicit the prior beliefs of more than 1,200 university students regarding the protest turnout of their classmates in the upcoming July 1 march (in an incentivized manner); we are then able to provide an information treatment to a random subset and elicit posterior beliefs (again in an incentivized manner); finally, we are able to elicit the students' own protest participation.

4. See Coffman, Featherstone, and Kessler (2015) for a theoretical analysis of the effects of information nudges.

5. The eventual success or failure of the movement is likely to have repercussions throughout "greater China" (and thus around the world) given concerns in Hong Kong, mainland China, and Taiwan over the increasingly authoritarian and nationalistic policies undertaken by the CCP. Our work contributes to a growing empirical literature on the political economy of the region: for example, Lorentzen (2013) highlights the central government's tolerance of certain types of protests; King, Pan, and Roberts (2013) study information control policies that aim at suppressing collective actions.

The goal of our experimental design is to isolate the causal effect of variation in beliefs regarding others' protest participation on one's own protest participation. To do so, we provide a random subset of individuals in our sample truthful information intended to shift beliefs regarding others' protest participation. A challenge we face is that such information must be provided prior to the protest itself—before we know the actual protest decisions of others. To solve this problem, one week before the protest, we collect information on individuals' planned protest turnout, on individuals' beliefs about others' planned turnout, and individuals' beliefs about others' future actual turnout at the protest. This allows us to provide truthful information regarding others' planned participation, plausibly affecting beliefs regarding others' actual protest participation.

A day before the protest, we provide a random subset of individuals in our sample truthful information about the planned participation of their classmates. We estimate the "first-stage" effect of information regarding others' planned participation on individuals' (posterior) beliefs regarding others' actual participation. Next we estimate the "reduced-form" effect of information regarding others' planned participation on individuals' own actual protest participation. Importantly, we split our analysis into two subsamples: those whose prior beliefs were below the true level of planned participation (whose beliefs regarding actual turnout, we expect, should be positively affected) and those whose prior beliefs were above the true level of planned participation (whose beliefs regarding actual turnout, we expect, should be negatively affected).[6]

Our findings consistently point to our sample of Hong Kong students viewing the strategic component of their protest decision as being a game of strategic substitutes. Among subjects whose prior beliefs regarding others' planned participation were below the truth, the experimental provision of information regarding

6. Note that in addition to providing evidence on balance between treatment and control groups in the full experimental sample, we present evidence of balance within each of these subsamples (see Section III). This sample split relies on our elicitation of beliefs about both others' planned participation and others' actual participation. The former gives us a measure of where priors stood relative to the experimental information we are able to provide to subjects prior to the protest itself (i.e., on planned participation). The latter gives us a measure of the priors we care about when examining belief updating in the first-stage analysis (i.e., changes in beliefs regarding actual participation).

the true level of other subjects' planned participation has a significant, positive effect on beliefs about actual participation in the protest, and a significant negative effect on subjects' own turnout. Among subjects whose prior beliefs regarding others' planned participation were above the truth, the experimental provision of information regarding the true level of other subjects' planned participation has a significant, negative effect on beliefs about actual participation in the protest, and a significant positive effect on subjects' own turnout.

We are able to address several concerns about our analysis. First, using list experiments, we provide evidence that our experimental subjects are willing to truthfully report on potentially sensitive political attitudes related to their participation in the July 1 protest; this helps assuage concerns regarding our reliance on a self-reported measure of protest turnout (see Section III.C).[7] Second, we can rule out a major threat to internal validity: the possibility that information about other subjects' turnout affected not only beliefs about others' protest participation but also beliefs about the "quality" of the political movement. Such a confounding "social learning" effect, however, would produce the appearance of strategic complementarity, not the strategic substitutability that we find. Third, we can address concerns regarding experimenter demand effects following a similar logic: typically, an experimenter's implied endorsement of an action (by indicating its popularity) would produce the appearance of strategic complementarity (see Section IV.D).

We find suggestive evidence of three sources of strategic substitutability in our context (see Section V). First, one aim of Hong Kong's protests is to get a sufficient number of people onto the street—this is a public good (as in Olson 1965; Tullock 1971; Palfrey and Rosenthal 1984) that could have tangible consequences and could serve as a signal of the movement's strength to the CCP and to other citizens. If subjects view attaining a threshold level of protest participation as producing a political public good, this will tend to produce strategic substitutability. Consistent with protest turnout being a public good, we find that

7. One might be particularly concerned about misreporting in survey data collected in authoritarian settings (see Reny 2016 for a discussion of challenges facing social scientists in China). Indeed, analyses of political behavior in real time are more common in settings that are already politically free (e.g., Gerber et al. 2011, 2017).

more prosocial subjects are more likely to protest, even conditional on ideology. Second, experimental subjects perceive a greater likelihood of a protest's success as protest size increases and a greater likelihood of government crackdown. If the latter dominates, the result can be strategic substitutability. Third, subjects may have social image concerns, and select into protest participation in part to signal their ideological "type" (as in Bénabou and Tirole 2011). If participation in a smaller protest sends a strong enough signal, individuals wanting to signal their antiauthoritarian ideology may differentially participate when they anticipate a protest will be small, thus producing a game of strategic substitutes. Consistent with this mechanism, we find that individuals who participate in the protests after learning that protests will be smaller than expected have ideologically more extreme friends than other protest participants.

Our results thus indicate that protests are not generically games of strategic complements—as assumed in much recent work. Models of protest participation must allow for the possibility of strategic substitutes. Yet it is important to emphasize that not all of the mechanisms we observe at work in Hong Kong will be present in all protests, and even if they are present, they will not always outweigh forces generating strategic complementarity emphasized in other work. As we discuss in Section VI, we believe that our findings reflect the fact that Hong Kong's protests are part of a long-running movement and that the Hong Kong government protects basic rights of association and expression. Strategic substitutability thus seems most likely to appear in protests that are part of larger movements and protests demanding rights from partially democratic regimes, while forces pushing toward strategic complementarity may dominate in one-shot protests that will end in the ousting of a dictator or the crushing of a movement.

In Section II we provide an overview of Hong Kong's ongoing democratic, antiauthoritarian movement and the July 1 march in particular. In Section III we describe our experimental design. In Section IV we present our main findings and discuss threats to internal validity. In Section V we discuss the theoretical implications of our findings and the characteristics of our setting that may generate strategic substitutability in the protest game. Finally, in Section VI, we discuss conditions under which strategic substitutability or complementarity are more likely, and offer concluding thoughts.

II. Hong Kong's Antiauthoritarian Movement

II.A. Political Context

Prior to 1997, Hong Kong was a British colony, with limited democratic political rights but strong protections of civil liberties and respect for the rule of law. On July 1, 1997, Hong Kong was returned to the People's Republic of China, to be ruled as a Special Administrative Region with its own quasi constitution—the Basic Law—and a promise from China that its institutions would be respected and maintained until 2047, under a policy known as "one country, two systems." The Basic Law left ambiguous several important details that have been bargained and battled over between the so-called pan-democracy and pro-Beijing camps.

The first ambiguity to generate mass political protests was regarding Article 23 of the Basic Law, which covered the legal regulation of speech and behavior that threatened the government. Under Beijing's encouragement, a law implementing provisions of Article 23—the National Security Bill—was proposed by the Hong Kong chief executive (the head of government) in September 2002 and was seen by many Hong Kong citizens as deeply threatening to their human rights and civil liberties.[8] The proposed legislation catalyzed a massive July 1 march (in 2003) in which an estimated half a million people protested. This expression of popular opposition led to the withdrawal of the bill, and no legislation on Article 23 has passed since.

More recently, political conflict has arisen from a second ambiguity in the Basic Law, regarding the method of selection of Hong Kong's chief executive. Article 45 of the Basic Law states the following: "The method for selecting the Chief Executive shall be specified in the light of the actual situation in the Hong Kong Special Administrative Region ... The ultimate aim is the selection of the Chief Executive by universal suffrage upon nomination by a broadly representative nominating committee in accordance with democratic procedures." While indicating an ultimate aim of universal suffrage, the Basic Law does not state when elections will be introduced, nor does it clarify the details of nomination. From Hong Kong's return to China until today, the chief executive has been selected by an election committee, rather

8. For a discussion of these concerns, see the University of Hong Kong's Human Rights Portal Page, "Research on Article 23," online at https://goo.gl/GdNcHY, accessed February 28, 2018.

than by universal suffrage; currently, the committee is composed of 1,200 members, and is widely seen as pro-Beijing.

In 2014, the Twelfth National People's Congress proposed an election mode that would have allowed the citizens of Hong Kong a choice between two or three candidates, but these candidates would be selected by the same pro-Beijing committee as before.[9] In response to this limited expansion of democratic rights, a massive July 1 march was mobilized, with hundreds of thousands of citizens taking to the streets. Further escalation and a police crackdown precipitated the even larger-scale Umbrella Revolution, named for the ubiquitous umbrellas carried by participants. The Umbrella Revolution persisted for months, being slowly cleared out by police by the end of December 2014. Although the movement did not alter the policy proposed by Beijing, it did send a clear signal to the Hong Kong legislature (the "LegCo") that a circumscribed change in institutions was unacceptable to the people of Hong Kong. In June 2015, the LegCo struck down the Chinese proposal led by the opposition of the pan-democratic camp.

Since June 2015, the democratic movement in Hong Kong has both fragmented and radicalized. Recent encroachments on Hong Kong citizens' civil liberties, including the arrest of Hong Kong booksellers by the mainland Chinese government, have deepened some citizens' fear of the CCP and their sense of a Hong Kong identity very much distinct from—even opposed to—that of mainland China. The result is that Hong Kong citizens and political parties are now much more loudly calling for independence or self-determination. "Localist" violence has occasionally flared; new political parties, such as the student-led Demosistō, have formed and won seats in the 2016 LegCo election on platforms explicitly calling for self-determination.[10]

II.B. The July 1 Marches: Characteristics and Achievements

Marches on the anniversary of Hong Kong's handover to China, held each July 1, have been described as "the spirit of

9. Refer to https://goo.gl/0oyNmt, accessed February 28, 2018.

10. The legislators elected on a self-determination platform were since removed from office on various technicalities regarding their oath-taking, foreshadowing future conflict.

democratic struggle in Hong Kong."[11] The July 1 marches have played an important role in Hong Kong citizens' political engagement with the Chinese government and have achieved major policy changes and even constitutional concessions—particularly when large crowds of protesters were mobilized.[12] Each protest march, while part of a broader antiauthoritarian, democratic movement, is organized around a specific set of issues and policy aims. The first notable achievement came as a response to the CCP's September 2002 proposal for an antisubversion bill under Article 23, described above. The July 1, 2003, march included around 500,000 people—the largest political gathering in Hong Kong since the Chinese democracy movement of 1989. Not only was the proposed law withdrawn, the march eventually forced the resignation of multiple government officials, including the chief executive, Tung Chee-hwa.[13]

Another success followed the 2012 march, which included up to 400,000 people, and was part of a mobilization against a CCP proposal for a mandatory "moral and national curriculum" in Hong Kong schools. This proposal, too, was withdrawn shortly after the march. The 2014 march again saw hundreds of thousands of people demanding the popular nomination of chief executive candidates in the 2017 election. Although the march did not achieve citizen nomination of chief executive candidates, it did produce the massive Umbrella Revolution and led to the rejection of the CCP's proposal for partial democratic rights.

Like others before it, the July 1, 2016, march studied here was organized around important political aims: first, to denounce the perceived corruption of Beijing-backed Chief Executive C. Y. Leung; second, to mobilize support for democratic—especially the newly established localist—political parties in the run-up to the 2016 LegCo elections. Although the protest was smaller than some previous marches (turnout was under 100,000 participants), it is

11. "Sixteen Years of July 1st Marches: A Dynamic History of Hong Kong Citizens' Fight for Democracy," *Initium Media*, June 30, 2018, https://goo.gl/8bZDrf, accessed July 5, 2018.

12. A time series of turnout in July 1 marches can be seen in Online Appendix Figure A.1.

13. In an opinion piece tellingly titled "July 1st March Turnout Size Is Absolutely Important," former LegCo member Margaret Ng Ngoi-yee writes, "The turnout at the July 1st Marches is absolutely important. If not for 500,000 people taking to the street in 2003, Article 23 would have been legislated already." *Stand News*, June 29, 2018, https://goo.gl/vgP3WP, accessed July 5, 2018.

noteworthy that Leung chose not to run for reelection in 2017, despite being unconstrained by term limits; parties opposed to Beijing won 55% of the LegCo vote share, with localist political parties winning nearly 20% of the vote in the 2016 election.[14]

Some characteristics of Hong Kong's July 1 marches may appear idiosyncratic: they are regularly scheduled events, and they are largely tolerated by an authoritarian government. In fact, these characteristics appear in other contexts. First, regularly scheduled protests are utilized by many antiauthoritarian movements, from Russia's Strategy 31 movement demanding rights of assembly to the Monday demonstrations in Leipzig that precipitated the fall of the German Democratic Republic.[15] Second, authoritarian regimes are often surprisingly tolerant of protests, within limits. The Monday demonstrations in Leipzig were able to proceed in the late summer and autumn of 1989 despite the obvious possibility of crackdown.[16] In Russia, protesters recently organized rallies in support of opposition politician Alexei Navalny on Vladimir Putin's 65th birthday, in October 2017, and the *Financial Times* notes that in response to a protest of around 1,000 people in Moscow, "police largely left protesters alone."[17] Even in mainland China, the CCP tolerates particular protests (Lorentzen 2013). In each of these settings, there exists a threat of crackdown ex ante, and—even in Hong Kong—police do crack down when protests cross the line.

Thus, like other antiauthoritarian protests, Hong Kong's July 1 marches demand (and occasionally win) fundamental political rights—civil liberties and democratic institutions—from an authoritarian regime. Like other antiauthoritarian protests, turnout is important for success. The importance of protest size can be

14. Even using lower-end estimates of the protest size (30,000), as a percentage of the population, this would make the 2016 July 1 march around one-third of the size of the largest protest in U.S. history, the Women's March in 2017.

15. Strategy 31 is discussed by Luke Harding in "The Russian Protesters Who Won't Give Up," *The Guardian*, August 30, 2010, https://goo.gl/vfwZro, accessed December 9, 2017. Weeks of modestly sized, regularly scheduled protests prior to the massive events that led to the fall of the Berlin Wall can be seen in Online Appendix Figure A.2.

16. See Andrew Curry, "A Peaceful Revolution in Leipzig," *Spiegel Online*, October 9, 2009, https://goo.gl/iUakCp, accessed December 9, 2017.

17. Several dozen protesters were detained and then released in St. Petersburg, which saw a protest of more than 2,000 people. See Max Seddon and Henry Foy, "Anti-Putin Protests Mark Russian President's Birthday," *Financial Times*, October 7, 2017, https://goo.gl/4oWQzA, accessed December 9, 2017.

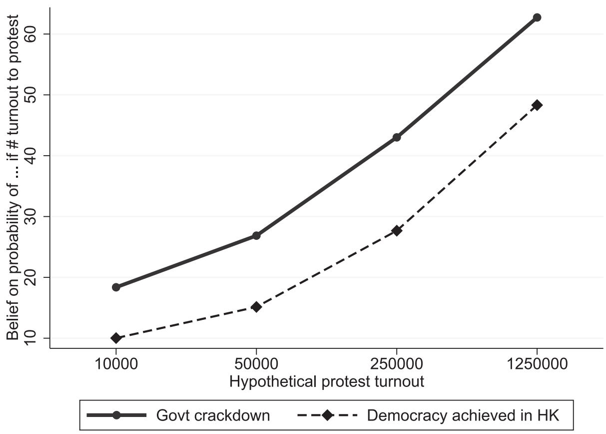


FIGURE I

Students' Beliefs Regarding the Benefits (the Chance of Achieving Democratic Institutions in Hong Kong) and Costs (the Chance of a Violent Government Crackdown) for Hypothetical Protests with Different Turnout Levels, Ranging from 10,000 to 1,250,000 Participants

seen in our survey data: subjects in our experiment believe there is a higher likelihood of protest success if a protest is larger (see Figure I). It can also be seen in the differences between July 1 march organizers' turnout estimates and the turnout estimates of the Hong Kong police. Organizers consistently exceed independent estimates of July 1 march size (and police estimates consistently fall below), with differences between the two reaching the tens or even hundreds of thousands (see Online Appendix Figure A.1).

Finally, like other antiauthoritarian protests, there is the potential for a high personal cost for turnout. Chinese authorities are deeply concerned about political instability in Hong Kong, at least in part because of potential spillovers into mainland China.[18] Thus, beyond the time cost and the experience of heat, humidity, and rain on a summer's day, the concern of the Chinese government implies the potential for high participation costs: the possibility of arrest and forceful police crackdowns using batons and tear gas—which have already occurred—and the potential

18. The Chinese government blocked Instagram—the last major uncensored social media platform available inside the Great Firewall—when the Umbrella Revolution broke out at the end of September 2014 (Hobbs and Roberts 2018).

for more violent suppression, particularly by the People's Liberation Army stationed in Hong Kong. A *New York Times* article describes the Umbrella Revolution in frightening terms: "On the first night, and for the next two weeks, rumors rippled through the [protesters'] camp. Protesters were fearful of a bloody crackdown, like what happened in Tiananmen Square." Interestingly, subjects in our experiment believe there is a higher likelihood of a government crackdown if a protest is larger (see Figure I). Our finding that beliefs about protest success and government crackdown both increase with protest size suggest that (much like in other antiauthoritarian protests) there are forces for strategic complementarity and for strategic substitutability in the July 1 march.[19]

III. Experimental Design

III.A. Design Overview

Our experiment was conducted online in three parts.[20] The goal of our experimental design is to isolate the causal effect of variation in beliefs regarding others' protest participation on one's own protest participation. To do so, we provide a random subset of individuals in our sample with truthful information intended to shift beliefs regarding others' protest participation. A

19. We discuss limitations on the generalizability of the Hong Kong context in the Conclusion (Section VI).

20. The experiment described here was preregistered with the AEA along with a second experiment, which varied persuasive messages regarding a democratic political party and examined the effects of these messages on contributions to that party as well as on political attitudes and beliefs. The "persuasion" experiment was completed in Part 1 of the study, and was cross-randomized with the intervention studied here; importantly, all of the data collection for the other experiment occurred prior to the experimental intervention studied here (which was implemented in Part 2 of the study). Reflecting the cross-randomization of the two experiments, in Online Appendix Table A.1, we show that the variables collected in Part 1 are generally balanced between the treatment and control groups in this study. In Section IV, we examine the impact of the unbalanced "Part 1" variables on our treatment effects, and find that they have almost no effect on our treatment effect estimates (results are reported in Online Appendix Table A.2). It is important to emphasize that the outcome variables considered in this article's analysis—posterior beliefs about the participation of others and subjects' own protest participation (collected in Part 3 of the study)—were the only outcome variables we collected following the experimental intervention we study here. We provide the full set of survey questions asked in Part 1 of the experiment (reformatted for brevity and organized thematically) in Online Appendix A.1. The full text of Parts 2 and 3 of the study are reproduced in Online Appendices A.2 and A.3, respectively.

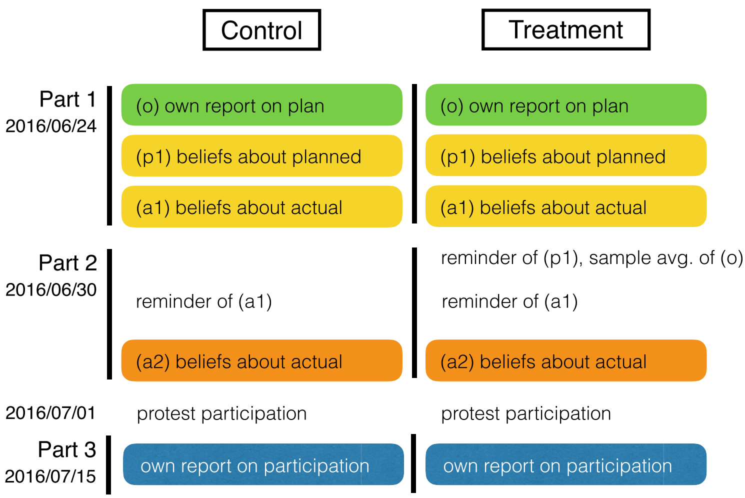
challenge we face is that such information must be provided prior to the protest itself—before we know the actual protest decisions of others.

We solve this problem by collecting information on individuals' beliefs about others' planned turnout, as well as beliefs about others' future actual turnout at the protest. These should be closely related, and crucially, we are able to elicit planned protest participation (as opposed to actual participation) prior to the protest. This allows us to provide truthful information regarding others' planned participation, plausibly affecting beliefs regarding others' actual protest participation.

We first estimate the "first-stage" effect of information regarding others' planned participation on individuals' (posterior) beliefs regarding others' actual participation. Next we estimate the "reduced-form" effect of information treatment regarding others' planned participation on individuals' own actual protest participation. Putting together the first stage and the reduced form, we can estimate the effect of a change in beliefs about others' participation on one's own using two-stage least squares.[21]

The broad outline of the design is as follows:

i. **Part 1:** On June 24, 2016, we elicited subjects' own planned participation in the upcoming July 1 march.



21. We discuss this two-stage estimate, particularly the implied exclusion restriction, in more detail below.

We also elicited subjects' beliefs regarding other subjects' planned protest participation. We refer to these as elicited priors regarding other subjects' planned participation. In the same survey, we elicited subjects' beliefs regarding other subjects' actual protest participation on July 1, 2016. We refer to these as elicited priors regarding other subjects' actual participation. Finally, we elicited subjects' beliefs regarding the total protest participation among all Hong Kong citizens on July 1, 2016. We refer to these as elicited priors regarding total actual turnout among all HK citizens.

ii. **Part 2:** On June 30, 2016, we provided a random subset of our experimental sample with a reminder of their prior beliefs regarding other subjects' planned participation, as well as information regarding the true level of planned protest participation in the experimental sample.[22] For both the information treatment group and the control group, we again elicited beliefs regarding other subjects' actual protest participation on July 1, 2016. We refer to these as elicited posteriors regarding other subjects' actual participation. Comparing posteriors between the treatment and control groups provides an estimate of the first-stage relationship. We also elicited subjects' beliefs regarding the total protest participation among all Hong Kong citizens on July 1, 2016. We refer to these as posteriors regarding total actual turnout among HK citizens.

iii. **Part 3:** On July 15, 2016, we elicited subjects' participation in the July 1 protest. Comparing participation rates between the treatment and control groups provides an estimate of the reduced-form relationship of interest. Self-reported July 1 protest participation is also the outcome in our two-stage estimates of the effects of beliefs regarding others' protest participation on one's own.

III.B. Experimental Sample

Our sample of experimental subjects is drawn from the population of students at HKUST. Studying a sample of students

22. We provide the reminder of subjects' priors to make the information treatment more salient, thus potentially increasing the power of our intervention. Of course, the reminder might serve as an anchor for subjects' responses, which could attenuate the treatment's effects.

to understand protest participation is ideal given students' importance in Hong Kong's democratic movement and in the localist political parties pursuing self-determination. In Part 1 of the study, we recruited participants on June 24, 2016, sending an email to the entire undergraduate population of HKUST.[23] We received 1,741 completed surveys, achieving a response rate of 19.1%. Among these, we focus on the 1,576 students who were either born in Hong Kong or moved there prior to high school (Hong Kong "natives"). Part 1 of the experiment elicited students' political preferences, beliefs, attitudes, and planned and past political protest behavior. Because protests occur every year on Handover Day, July 1, we asked a series of questions specifically eliciting planned participation in the upcoming July 1 protest, as well as (prior) beliefs about turnout at the protest.

We paid students for their participation and provided additional payments as a function of their choices in incentivized games and in incentivized belief elicitations. On average, respondents received HK$205, approximately US$25, for completing this first survey. Our experimental intervention was conducted in Part 2 of the study, a very short online survey sent in an email on June 30, 2016, and completed by 1,303 Hong Kong native students. Along with the experimental intervention of interest, this second survey elicited (posterior) beliefs about turnout in the following day's protest. Students received a payment of HK$25 for completing the survey. Finally, in Part 3 of the study, we elicited students' participation in the July 1 protest of 2016 in a third online survey sent via email on July 15, 2016, and completed by 1,234 Hong Kong native students. Students who completed Part 3 of the study received an additional payment of HK$25. We present summary statistics for the observable characteristics of the experimental sample—those subjects who completed all three parts of the study—in Table I, columns (1) and (2).[24]

23. Our recruitment email informed students that we were researchers at HKUST, UC Berkeley, Stanford, and the University of Munich interested in understanding attitudes and preferences among college students in Hong Kong. The initial email did not explicitly mention our interest in political attitudes. All experimental materials were provided in English, the primary language of instruction at HKUST. Some bilingual support (i.e., materials provided in Chinese characters) was provided to clarify key terms.

24. In Online Appendix Table A.3 we present summary statistics for the 1,576 Hong Kong native students who completed Part 1 of the experiment, and for the 1,234 students in our experimental sample. One can see that the two groups

TABLE I

Summary Statistics and Balance Check

	Experimental sample			Prior belief on planned particip. below truth			Prior belief on planned particip. above truth		
Variables	Control mean (1)	Treatment mean (2)	*p*-value (3)	Control mean (4)	Treatment mean (5)	*p*-value (6)	Control mean (7)	Treatment mean (8)	*p*-value (9)
Male	0.554	0.535	.548	0.566	0.523	.233	0.522	0.565	.454
Birth year	1995.2	1995.2	.801	1995.4	1995.2	.259	1995.0	1995.2	.178
HK-oriented childhood env.	0.088	0.100	.827	0.069	0.089	.773	0.137	0.128	.932
Nonreligious	0.783	0.770	.595	0.809	0.771	.199	0.717	0.766	.317
HH economic & social status	0.029	−0.031	.324	0.019	−0.017	.623	0.055	−0.065	.279
Own projected economic status	0.001	−0.056	.348	0.037	−0.025	.400	−0.090	−0.130	.714
Planned to participate in protest	17.0	17.9	.688	13.9	14.2	.905	24.8	26.6	.714
Prior belief re: % others' planned particip.	15.7	15.8	.902	7.8	7.6	.501	36.0	35.4	.743
Prior belief re: % others' actual particip. (HKUST)	13.4	14.0	.470	6.0	5.9	.772	32.1	33.2	.546
Prior belief re: others' actual particip. (all HK)	158,243	153,665	.628	145,996	140,858	.635	189,458	183,875	.763
# of obs.	401	833		288	585		113	248	

Notes. Table tests for balance on observable characteristics (treatment versus control) for the full sample (columns (1)–(3)); the subsample of subjects whose prior beliefs regarding others' planned participation were below the true value of 17% (columns (4)–(6)); and the subsample of subjects whose prior beliefs regarding others' planned participation were above the true value of 17% (columns (7)–(9)). "HK-oriented childhood env." is a *z*-score index constructed from an indicator of whether the subject completed high school with English as the formal language of instruction (as opposed to Chinese), and the number of generations a subject's family has lived in Hong Kong. "HH economic & social status" is a *z*-score index constructed from self-reported total income earned by both parents (including sources of income such as dividends and rents), the number of real estate properties owned by a subject's parents/household in Hong Kong at the time of the survey, and whether a subject's father's and mother's highest educational attainment are above high school, respectively. "Own projected economic status" is a *z*-score index constructed from the projected median income of HKUST graduates in a subject's major/program, and self-reported expectations of relative income compared to classmates at HKUST at age 40. All *z*-score indices are weighted by the inverse covariance of the standardized outcomes, computed following Anderson (2008).

III.C. Elicitation of Plans, Beliefs, and Actual Protest Participation

1. Subjects' Planned Participation and Prior Beliefs. In Part 1 of the study, on June 24, 2016, we elicited subjects' own planned participation in the upcoming July 1, 2016, antiauthoritarian protest, asking the question that appears in Box I.

We next elicited subjects' beliefs regarding other subjects' planned protest participation (i.e., elicited priors regarding planned participation). This elicitation, like all other belief elicitation in this study, was conducted in an incentivized manner.[25] Specifically, we asked the question that appears in Box II.

In the same survey, we elicited subjects' beliefs regarding other subjects' actual protest participation on July 1, 2016 (i.e., elicited priors regarding actual participation). We asked the question that appears in Box III.

Responses in Part 1 indicated that 16.9% of all subjects (including nonnative Hong Kong students) planned to participate in the July 1 protest of 2016 (i.e., answered either "Yes" or "Not sure yet, but more likely than not" to the question regarding their plans for the upcoming July 1 protest). This is the number we use to provide "true" information regarding planned participation to experimental subjects (rounding to 17%). The experimental sample's average prior belief regarding planned protest participation was quite close to the truth, at 15.8%, but there was a great deal of variation around the truth (the standard deviation was also 15.8 percentage points).

---

are extremely similar. The experimental sample of students is also similar to—though not precisely representative of—the broader HKUST student body on the dimensions of school of enrollment (i.e., students' broad academic area), gender, and cohort (see Online Appendix Table A.4). Note that all of our findings are robust to reweighting our experimental sample to match the composition of the HKUST student body or to match the composition of the 1,576 Hong Kong native students who completed Part 1 of the experiment (see Section IV.D).

25. The survey and the incentives are necessarily coarse: we elicit respondents' beliefs on how many other subjects answered "Yes" or "Not sure yet, but more likely than not," not the entire distribution of beliefs. This corresponds to providing an incentive for subjects to report their belief regarding the modal outcome, rather than the mean (though these will correspond if subjects' distributions of beliefs are symmetric and single-peaked). We find that the distribution of prior beliefs regarding other students' planned participation has a mean and a mode very close to the true level of planned participation in the sample. This is consistent with the incentives we provided generating thoughtful, truthful responses.

BOX I: (O) OWN REPORT ON PLAN

Are you planning to participate in the July 1st march in 2016?

1 Yes
2 Not sure yet, but more likely than not
3 Not sure yet, but more unlikely than yes
4 No

BOX II: (P1) BELIEFS ABOUT PLANNED

Please guess what percentage of the participants from HKUST of this study plan to participate in the July 1st march in 2016 (answer either “Yes” or “Not sure yet, but more likely than not” to the above question on July 1st March in 2016).
If your guess is within 2 percentage points of the percent of students who actually answer either “Yes” or “Not sure yet, but more likely than not,” you will earn a bonus payment of HK$10.

BOX III: (A1) BELIEFS ABOUT ACTUAL (HKUST STUDENTS)

Please guess what percentage of the participants from HKUST of this study will participate in the July 1st march in 2016.
If your guess is within 2 percentage points of the percent of students who actually participate, you will earn a bonus payment of HK$10.

The sample’s average prior belief regarding others’ actual participation in the July 1 protest was 13.8%, slightly below the average prior belief about others’ planned participation. Prior beliefs about others’ planned and actual participation are strongly associated, as expected (the correlation is 0.83).[26]

Finally, we elicited subjects’ prior beliefs regarding the total actual turnout at the July 1 march. We asked the question that appears in Box IV.

The experimental sample’s average prior belief regarding total turnout in the July 1 protest was 155,153.

2. The Experimental Intervention and Posterior Beliefs. In Part 2 of the study, on June 30, 2016, we implemented the

26. In Online Appendix Figure A.3, we present the distributions of subjects’ prior beliefs regarding others’ planned participation and regarding others’ actual participation. One can see in the figure that the distribution of priors regarding actual participation is shifted slightly to the left of the distribution of beliefs regarding planned participation.

BOX IV: (A1) BELIEFS ABOUT ACTUAL (HK POPULATION)

How many people in total do you think will participate in the July 1st march in 2016?
If your guess is within 10% of what will be reported by the HKUPOP after the July 1st march in 2016, then you will earn a bonus payment of HK$10.
To give you a sense, according to HKUPOP's report, among the July 1st march that took place between 2003 and 2015:
The lowest attendance in a given year is: 17,000 (in 2008);
The highest attendance in a given year is: 462,000 (in 2003).

experimental intervention, randomly assigning two-thirds of subjects to the treatment group and one-third to the control group.[27] In Table I, columns (1)–(3), we present data on the background characteristics, economic status, protest plans, and prior beliefs of the treatment and control groups, and test for balance between them. One can see that the treatment and control groups are very similar on these margins. As discussed already, the impact of an information shock on beliefs, and thus behavior, should differ (having effects of opposite sign) depending on whether the information provided was above or below individuals' prior beliefs. We conduct much of our analysis separately examining individuals with priors above and below the information treatment, or pooling all subjects, but coding the treatment indicator as being equal to -1 for individuals with prior beliefs above the information treatment to make the treatment effect monotonic.[28] It is thus important to check for balance in the two subsamples of interest—subjects with priors above and below the true value of planned participation of

27. The decision to assign more individuals to the treatment group was made anticipating the possibility that some subjects may have ignored Part 2 of the study, and thus effectively ended up in the control condition. Under such a scenario, we could have examined protest behavior among individuals who were actually treated and among individuals who were assigned to the control condition or who did not complete the survey in Part 2 of the study. In practice, the vast majority of subjects completed all three parts of the study, so the additional individuals in the treatment group were not strictly necessary.

28. It is important to note that variation in individuals' prior beliefs was not experimentally induced. In Online Appendix Table A.5, we present predictors of individuals' own self-reported plans to participate in the protest as well as predictors of individuals' prior beliefs regarding other subjects' planned participation. We discuss the endogeneity of priors further in Section IV.

BOX V: REMINDER OF (P1), SAMPLE AVERAGE OF (O)

Recall that you guessed that *[Part 1 response]%* of HKUST survey participants would plan to attend the July 1 march.
Based on last week's survey, the true percentage of survey participants who plan to attend the July 1 march is 17%.

BOX VI: REMINDER OF (A1)

Remember that we offered you:
1—A HK$10 bonus payment for accurately guessing the percentage of HKUST survey participants who would actually attend this July 1 march;
2—An additional HK$10 bonus payment for accurately guessing the total number of Hong Kong citizens who would actually attend this July 1 march.
In last week's survey, you guessed that:
1—*[Part 1 response]%* of HKUST survey participants would attend this July 1 march;
2—A total of *[Part 1 response]* Hong Kong citizens would attend this July 1 march.

17%. One can see in Table I, columns (4)–(9), that treatment and control groups within each subsample are well balanced.[29]

Individuals in the treatment group—but not the control group—were reminded of their responses from Part 1 regarding other subjects' planned participation in the July 1 protest of 2016, and then told the actual level of other subjects' planned participation, as shown in Box V.

All subjects (treatment and control) were reminded of their responses from Part 1 regarding actual participation in the July 1 protest of 2016 (see Box VI).

All subjects were then given an opportunity to update their responses from Part 1 (see Box VII).

The experimental sample's average posterior belief regarding the percentage of other subjects who would actually participate

29. As noted already, the survey in Part 1 elicited a broad range of subject characteristics: political attitudes and beliefs; personality traits; and, preferences (among others). We present a comprehensive set of balance tests (for the entire experimental sample and for the two subsamples of interest) for 49 different variables in Online Appendix Table A.1, and we find statistically significant differences (at the 10% level) for 8 of the 49. In Section IV, we examine the impact of the unbalanced Part 1 variables on our treatment effects, and find that they have almost no effect on our treatment effect estimates (results are reported in Online Appendix Table A.2).

BOX VII: (A2) BELIEFS ABOUT ACTUAL (HKUST STUDENTS)

Perhaps since then your views have changed.
We now ask you again to provide guesses about actual attendance of the July 1 march.
Instead of your guesses in the previous survey, we will use today's guesses to determine your bonus payment.*

1 How many people in total do you think will participate in the July 1st march in 2016?
If your guess is within 10% of what will be reported by the HKUPOP after the July 1st march in 2016, then you will earn a bonus payment of HK$10.
To give you a sense, according to HKUPOP's report, among the July 1st march that took place between 2003 and 2015:
The lowest attendance in a given year is: 17,000 (in 2008);
The highest attendance in a given year is: 462,000 (in 2003).

2 Please guess what percentage of the participants from HKUST of this study will participate in the July 1st march in 2016?
If your guess is within 2 percentage points of the percent of students who actually participate, you will earn a bonus payment of HK$10.

*We chose to pay subjects based on their responses in Part 2 (rather than using responses in both Part 1 and Part 2) to minimize any income effects or strategic incentives (e.g., hedging). By reminding subjects of their responses in Part 1 and allowing them to hold the same posteriors as priors, subjects were still able to be paid based on their Part 1 responses.

in the July 1 protest was 14.5%; the average posterior belief regarding total actual turnout among Hong Kong citizens was 142,684. In fact, the July 1 protest of 2016 was smaller than subjects expected: the protest was attended by 3% of experimental subjects, and only 26,000 people overall.[30]

3. Measuring Protest Participation. In Part 3 of the study, on July 15, 2016, we elicited subjects' participation in the July 1 protest of 2016.[31] We asked subjects "Did you attend the July 1,

30. The smaller than expected protest was perhaps the result of an announcement on the morning of July 1 that the arrested and returned Hong Kong bookseller Lam Wing-kee would not lead the protest, as had been expected, out of fear for his safety. The Hong Kong Free Press headlined an article on June 28, "Organisers Expect 100,000 to Attend July 1 Democracy Rally Led by Bookseller and Ex-Prisoners of Conscience" (https://goo.gl/kgDYr1, accessed December 19, 2016). The same source headlined an article on July 1, "Returned HK Bookseller Will Not Lead July 1 Democracy March after 'Serious Threat' to Safety" (https://goo.gl/dHcR9b, accessed December 19, 2016).

31. We waited two weeks before collecting information on march attendance, as this gave us time to gauge whether subjects would be comfortable responding to a direct question about attendance. Because there had been no arrests in the

2016 March?" A response of "yes" to this question is our measure of individuals' protest participation.[32]

An important concern regarding our measure of protest participation is that experimental subjects may not report on their participation truthfully. This concern is particularly relevant in the context of an ongoing antiauthoritarian movement. However, there are several reasons to believe that self-reported protest turnout is a good measure in our context.[33] First, the particular protest that we study remained peaceful. While subjects faced a risk of government crackdown on the protest ex ante, there was no concern regarding legal sanctions on participants two weeks after the protest, when subjects' protest participation was elicited. Second, for fear of government sanction to produce measurement error, it would need to be the case that subjects were willing to take the risk of attending a (very public) protest, but unwilling to tell us in a private survey that they did so. Although this is possible (they may misperceive the observability of their protest choice and fear putting their behavior on the record), it strikes us as unlikely.

As a more direct test of our experimental subjects' willingness to truthfully respond to politically sensitive survey questions, in Part 1 of the study we elicited several key dimensions of political ideology that may be considered sensitive using "list experiments" (or the item count technique; Raghavarao and Federer 1979). The list experiment provides cover for the expression of possibly stigmatized attitudes at the individual level but allows the researcher to estimate the prevalence of these attitudes at the population level. We adopt a modified version of the standard list experiment (Coffman, Coffman, and Ericson 2017) in which we also directly elicit the potentially stigmatized attitudes from the control group. Thus, for each potentially sensitive political attitude, we are able

weeks after the march, we felt comfortable asking subjects directly about their participation.

32. We also ask those who attended the march a small number of follow-up questions. First, we asked them to indicate which of 28 groups' crowds they joined at the protest (we also gave them the option of "Others"). Next, we asked for general impressions of the protest in an open-ended manner, subject to a 300-word maximum. Finally, we asked about the number of their friends who attended the protest. Because these questions were only asked of individuals who attended the march, we do not consider them outcome variables in this analysis.

33. This discussion of Hong Kong students' willingness to report their political attitudes and behavior truthfully closely follows Cantoni et al. (2016).

TABLE II

Item Count Experiments and Willingness to Respond to Direct Questions

Attitudes	"Yes" in direct question	Δ When cover is provided
Unfavorable view of CCP	0.923	0.020
		[0.059]
Consider self Hong Kongese	0.878	– 0.063
		[0.051]
Support for HK independence	0.465	0.057
		[0.057]
Support violence in pursuit of HK's political rights	0.216	0.173***
		[0.050]

Notes. The left column presents the fraction of 790 Hong Kong local students who expressed the corresponding attitude in response to a direct question. The right column presents the difference between that fraction and the fraction estimated to support the attitude using the item count technique (list experiment). The 790 students asked the direct questions also represent the control group for the list experiment; the remaining 786 Hong Kong local students represent the treatment group. Assignment to "direct question" and "list experiment" conditions was random among individuals who completed Part 1 of the study. *: Significant at 10%; **: 5%; ***: 1%. This table also appears in Cantoni et al. (2016).

to compare the study sample's estimated adherence to that attitude when cover is provided (based on our list experiment) with the expressed adherence to that attitude in response to a direct question (from the control group). We can then test whether there exists a significant distortion in the expression of a particular attitude in response to a direct question (e.g., due to stigma or fear).

In Table II, we present the fraction of our sample expressing support for Hong Kong independence; who consider themselves Hong Kongese rather than Chinese; who have a favorable view of the ruling CCP; and who support the use of violence in pursuit of Hong Kong's political rights (these estimates are based on subjects completing Part 1 of the study). In the left column, we simply present the population estimate of adherence to a political attitude based on direct questions. In the right column, we show the difference between the estimate based on direct questions and the estimate based on the list experiment. One can see that for three of the political attitudes, there is no significant effect of providing respondents with cover for expressing their views: this is true even for self-reported support for Hong Kong's independence—a much more extreme political position than simply attending a July 1 march.

Observing that the provision of cover by the list experiment does not affect estimated support for several sensitive attitudes, one might have been concerned that the "cover" provided by the list experiment was insufficient to elicit truthful reporting. However, Table II does show one significant difference between direct elicitation and list experiments: many students in our sample support the use of violence to achieve Hong Kong's political rights but are afraid to say so when directly asked. Finding a significant gap between direct questions and the list experiment on this dimension suggests that subjects do value the cover provided by the list experiment when it is needed—but it is not needed in response to political questions within the range of nonviolent opposition to the CCP. Because participation in the July 1 march falls within this nonviolent range, we are confident that subjects are willing to respond truthfully in response to direct questions about their participation.

IV. Main Results

IV.A. The First Stage: Effects on Posterior Beliefs

We begin by presenting the effects of the information treatment on individuals' beliefs regarding actual participation in the July 1 march—the first stage. Our focus is on posterior beliefs regarding the percentage of other experimental subjects who would actually participate in the July 1 march, rather than on beliefs regarding the total number of participants in the Hong Kong population. The former beliefs are more directly linked to the information provided regarding subjects' planned participation. We also present some evidence on posterior beliefs regarding total turnout among HK citizens.

The effect of the information treatment—reminding treatment group subjects of their prior beliefs regarding other subjects' planned participation and informing them that 17% of experimental subjects planned to attend the protest—can be seen in the distributions of beliefs regarding subjects' actual participation, presented in Figure II. Given that the information we provided to subjects was above the prior beliefs of some (regarding planned participation) and below the prior beliefs of others, if subjects believed that the information provided was truthful and updated their priors regarding actual participation in the direction of the new information, one would expect to see a more compressed distribution of posteriors in the treatment group than in the control group. Indeed, this is precisely what one observes

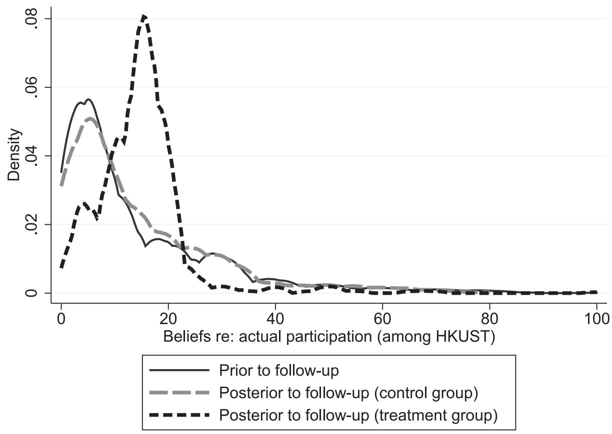


FIGURE II

Distribution of Prior Beliefs and Posterior Beliefs Regarding the Actual Protest Participation of HKUST Survey Participants

in Figure II: one can see in the figure that the treatment group's posteriors are distributed much more tightly between 10% and 20%.[34]

We more closely examine the anticipated heterogeneous effects of the information treatment depending on subjects' prior beliefs. In Figure III, we present a binned scatter plot of the change in beliefs (posteriors minus priors) regarding other subjects' actual participation against subjects' priors regarding other subjects' planned participation. In the left panel, one can see that, as predicted, subjects in the treatment group with priors regarding planned participation below (above) the information provided consistently updated their beliefs regarding other subjects' actual participation positively (negatively). Subjects in the treatment group with priors more distant from the information provided updated their beliefs more than those with priors closer to the information provided. Individuals in the control group with lower priors tended to update their beliefs positively, and vice versa, but the changes in beliefs are tiny compared to those observed in the treatment group (see the right panel of Figure III).[35]

34. A Kolmogorov-Smirnov test of equality of posterior distributions between the treatment and control groups strongly rejects the null ($p < .001$).

35. The updating of beliefs among the control group may result from newly acquired information from outside the study, or from information spilling over from

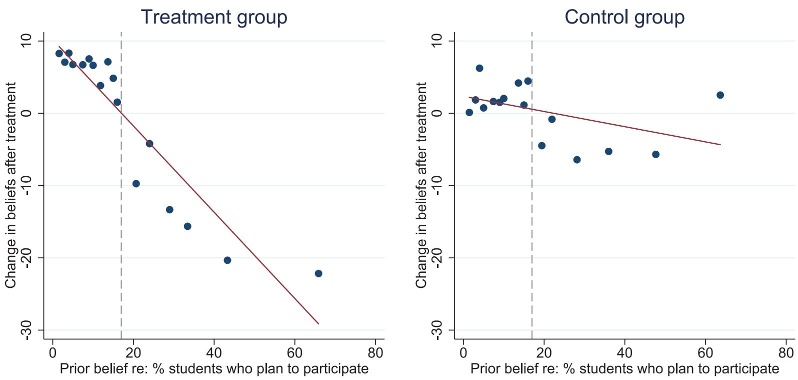


FIGURE III

Changes in Beliefs (Posteriors Minus Priors) Regarding the Actual Protest Participation of HKUST Survey Participants, Presented Separately by Subjects' Treatment Status

Observing that belief updating in the treatment group systematically differs in sign between subjects with priors above and below the information provided, we present in Figure IV the prior and posterior beliefs regarding other subjects' actual participation for the treatment and control groups, split by priors regarding planned participation above and below the information treatment. Recall that treatment and control groups are generally balanced on observable characteristics within each of these subsets (see Table I, columns (4)–(9)). One can see in the figure that among individuals with priors regarding other subjects' planned participation below (above) the true level, there is a significantly greater increase (decrease) in posteriors among the treatment group than among the control group.

We next estimate regression models predicting posterior beliefs regarding other subjects' actual participation as a function of treatment status, controlling for individuals' prior beliefs regarding others' actual participation.[36] To begin, we pool all subjects, but we code the treatment variable as being equal to

the treatment group; such information spillovers would tend to bias estimated effects (in both the first stage and the reduced form) toward 0.

36. In Online Appendix Table A.6 we present an alternative specification, in which the outcome variable is a subject's change in beliefs (posteriors minus priors) regarding other subjects' actual participation. We focus on belief levels in the main text as this specification is less restrictive on the coefficient on prior beliefs and

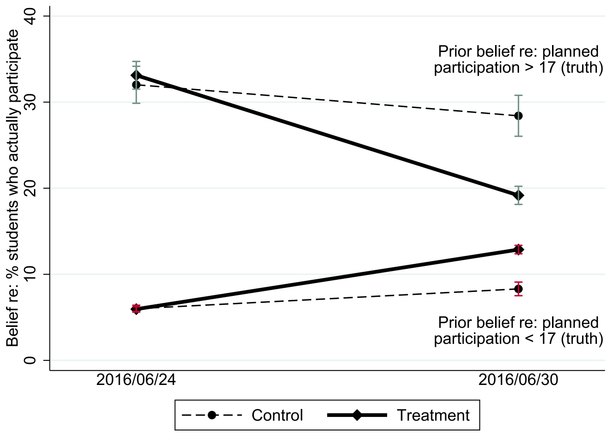


FIGURE IV

Treatment Effect on Beliefs ("First Stage")

Graph shows prior (measured June 24, 2016) and posterior (measured June 30, 2016) beliefs regarding the actual protest participation of HKUST survey participants, split according to subjects' treatment status and according to prior beliefs about other subjects' planned participation. Subsamples of subjects are divided according to whether beliefs regarding the planned protest participation of HKUST survey participants were above or below the true level of 17%.

−1 for individuals whose prior beliefs regarding the planned participation of other subjects were above the truth, in order to make the treatment effect monotonic. The coding of the treatment variable reflects our strong priors, as well as the evidence presented in Figure III, that individuals in the treatment group update their beliefs about the actual participation of others in opposite directions depending on whether their prior beliefs about the planned participation of others were above or below the information provided. One can see in Table III, Panel A, column (1), that the experimental treatment statistically significantly moved beliefs, by just below 6 percentage points.[37]

because models of protest participation typically focus on belief levels, rather than belief changes, as the drivers of protest participation.

37. We explore alternative "switching points" for the treatment variable (i.e., different levels of priors above which we code treatment equal to −1) in Online Appendix Figure A.4. Specifically, we estimate the specification presented in Table III, Panel A, column (1), and plot the point estimate (and 95% confidence interval) for integer switching points from 0 to 100. One can see in the left panel

TABLE III
TREATMENT EFFECT ON POSTERIOR BELIEFS

	Posterior belief on participation among HKUST students (%)			Posterior belief on total participation among HK population (#)		
	All subjects	Prior below truth	Prior above truth	All subjects	Prior below truth	Prior above truth
Sample	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Baseline						
Treatment	5.891***			11,423.9*		
(direction adj.)	[0.567]			[5,988.6]		
Treatment		4.457***	– 9.459***		13,198.0**	– 7,013.5
		[0.545]	[1.413]		[6,541.3]	[13,108.6]
Panel B: With controls						
Treatment	6.119***			14,133.0**		
(direction adj.)	[0.572]			[6,150.0]		
Treatment		4.638***	– 10.120***		15,086.2**	– 10,936.5
		[0.532]	[1.452]		[6,686.1]	[13,747.8]
Panel C: Trimmed						
Treatment	5.144***			13,051.2**		
(direction adj.)	[0.555]			[6,226.9]		
Treatment		4.279***	– 7.568***		15,911.0**	– 5,050.6
		[0.568]	[1.371]		[6,540.9]	[14,994.8]
Observations	1,234	873	361	1,234	873	361
DV mean (control grp.)	14.04	8.44	28.30	139,878	128,084	169,940
DV std. dev. (control grp.)	14.10	8.46	15.54	155,482	147,528	171,162
DV mean (all)	14.50	11.40	22.02	142,684	134,454	162,586
DV std. dev. (all)	10.83	7.99	12.92	142,685	139,385	148,689

Notes. Table shows first-stage effects: the effects of the experimental treatment on subjects' posterior beliefs regarding others' actual protest participation, conditional on subjects' prior beliefs regarding others' actual participation. Columns (1)–(3) show effects on posterior beliefs regarding other experimental subjects' actual participation, and columns (4)–(6) show effects on posterior beliefs regarding the total turnout at the protest by all Hong Kong citizens. In columns (1) and (4), all subjects are pooled and treatment is coded as -1 for subjects whose priors regarding others' planned participation were above the true value of 17%; pooled regressions also control for an indicator of whether subjects' priors regarding others' planned participation were above true value and its interaction with subjects' prior beliefs regarding others' actual participation. Columns (2) and (5) ((3) and (6)) show effects on posterior beliefs for the subsample of subjects whose priors regarding others' planned participation were below (above) the true value of 17%. Panel A is estimated without any additional controls beyond the corresponding levels of prior beliefs; Panel B replicates the analysis in Panel A but adds controls for subjects' background characteristics and economic status (gender, year of birth, a z-score index measuring whether subjects were raised in a Hong Kong–oriented environment, whether subjects were raised in a religious household, the economic status of the household, and own projected economic status); Panel C estimates the baseline specification of Panel A excluding those individuals in the experimental sample with the 5% lowest and the 5% highest prior beliefs regarding other subjects' planned participation. Number of observations refers to that in the baseline specification. *: Significant at 10%; **: 5%; ***: 1%.

Next we split the experimental sample into groups with priors regarding planned participation above and below the true level of planned participation of 17%. In Table III, Panel A, columns (2) and (3), one can see statistically significant belief updating in each subsample. Among individuals with prior beliefs regarding planned participation of other subjects below (above) the truth, treatment increases (decreases) beliefs regarding other subjects' actual turnout by around 4.5 (9.5) percentage points. A test of equality of the coefficients estimated in these columns is rejected with $p < .001$.

of Online Appendix Figure A.4 that we find the strongest first-stage effects when the coding of the treatment indicator switches at 17%.

We explore the robustness of the first-stage estimates along two margins in Table III. First, we examine whether controlling for subject characteristics affects the estimated treatment effects. One can see in Table III, Panel B, columns (1)–(3), that controlling for subjects' background characteristics and economic status (variables presented in Table I) does not meaningfully affect the estimated treatment effects in either the pooled regression or the split sample regressions. As a second robustness exercise, we consider the possibility that our results are strongly influenced by individuals with extreme priors; we thus drop from our sample the 5% of subjects with the lowest prior beliefs and the 5% of subjects with the highest prior beliefs regarding the planned participation of others. One can see in Table III, Panel C, columns (1)–(3), that dropping individuals with extreme priors does not greatly affect our results (though the coefficient estimated from the sample with priors above 17% is somewhat attenuated).

An additional important consideration is whether our statistical inferences based on traditional standard errors are sound.[38] As an alternative, we randomly assign (fictional) treatment status (in the same two-thirds treatment, one-third control ratio used in the actual experiment) and estimate first-stage treatment effects 10,000 times each for the subsample of subjects with prior beliefs regarding others' planned participation below 17% and the subsample with prior beliefs regarding others' planned participation above 17%. We can compare the t-statistics from the estimated treatment effects from the fictional treatment assignments to the t-statistics from the actual treatment assignment (the actual estimates are those in Table III, Panel A, columns (2)–(3)). We find that our p-values using randomization inference, based on two-sided tests, are very similar to those using standard inference (see Figure V).

Experimental subjects who updated their beliefs regarding other subjects' turnout at the protest also may have updated their beliefs regarding the turnout of Hong Kong citizens more generally. We thus next examine the effect of the treatment on subjects' beliefs regarding protest turnout among the entire Hong Kong population. We replicate the specifications in Table III, columns (1)–(3), but using as our outcome the posterior beliefs regarding the total turnout in the July 1 protest of 2016 (and

38. See Deaton and Cartwright (2018) for a discussion of challenges to statistical inference in randomized controlled trials.

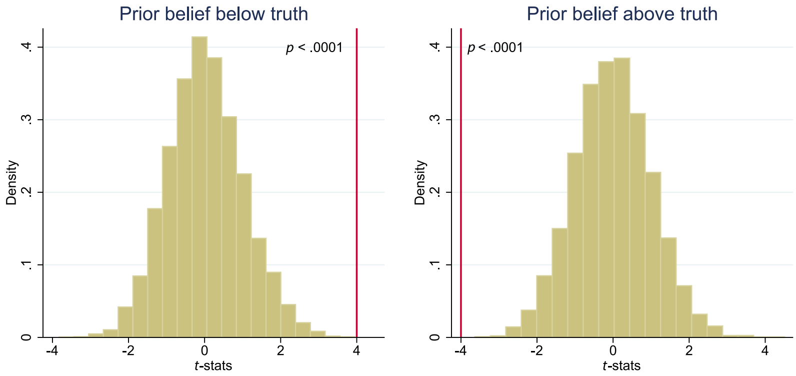


FIGURE V

Distribution of t-Statistics from Estimating the Treatment Effect on Posterior Beliefs (First Stage) Based on 10,000 Random Assignments of Treatment Status (Two-Thirds Treatment, One-Third Control) to Study Participants

Subsamples of subjects are divided according to whether beliefs regarding the planned protest participation of HKUST survey participants were above or below the true level of 17%. Red vertical lines indicate the t-statistics from the actual treatment assignment (drawn at 4 and -4, rather than their true values, for ease of visualization), with indicated p-values from two-sided tests.

controlling for prior beliefs regarding the total turnout at the protest). In Table III, columns (4)–(6), one can see suggestive evidence that the treatment affected beliefs regarding total protest size in the same direction as it affected beliefs regarding other subjects' turnout.

IV.B. The Reduced Form: Effects on Protest Turnout

We now turn to examining the effects of the information treatment on individuals' own protest participation. As we did in the analysis of the first stage, we split the experimental sample into two groups: first, subjects whose prior beliefs regarding other subjects' planned turnout were below the truth; and second, subjects whose prior beliefs regarding other subjects' planned turnout were above the truth. In the previous section we saw that in the former group, the treatment increased beliefs regarding other subjects' turnout, while in the latter group, the treatment reduced beliefs regarding other subjects' turnout.

In Figure VI, we present turnout levels among subjects in the treatment and control groups in the two subsamples split according to priors. One can see in the figure that in the subsample whose priors were below the truth, the information treatment caused a

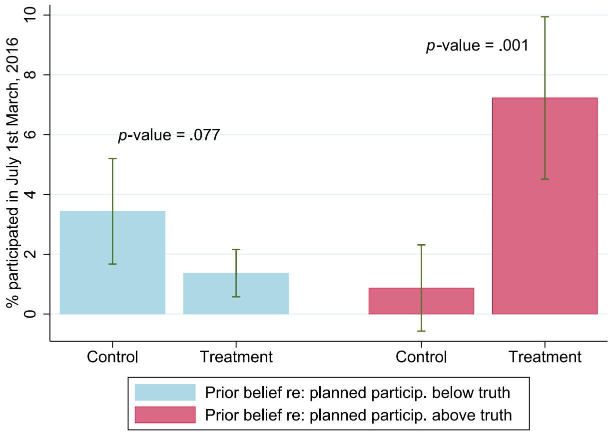


FIGURE VI

Treatment Effect on Protest Participation ("Reduced Form")

Graph shows self-reported participation in the July 1 protest of 2016, split according to subjects' treatment status and according to prior beliefs about other subjects' planned participation. Subsamples of subjects are divided according to whether beliefs regarding the planned protest participation of HKUST survey participants were above or below the true level of 17%.

statistically significant fall in turnout; in the subsample whose priors were above the truth, the information treatment caused a statistically significant increase in turnout. In other words, we find evidence that the protest decision is a negative function of beliefs regarding the turnout of others. It is worth emphasizing that this relationship is found in independent tests on two distinct subsamples: both the subsample with prior beliefs below the information provided and the subsample with prior beliefs above. The protest game in this setting is one of strategic substitutes.

We next turn to regression analysis of the reduced-form relationship between treatment and protest participation. As in the first-stage analysis, we begin by pooling all subjects, coding the treatment variable as being equal to −1 for individuals whose prior beliefs regarding the planned participation of other subjects were above the truth, in order to make the treatment effect monotonic. In Table IV, Panel A, column (1), one can see that the treatment causes a statistically significant 2.7 percentage point change in turnout in the opposite direction of the change in beliefs.[39] In

39. As was done in the first-stage analysis, we explore alternative switching points for the treatment variable (i.e., different levels of priors above which we code

TABLE IV
TREATMENT EFFECT ON PROTEST PARTICIPATION

	Participated in 2016 July 1 march		
Sample	All subjects (1)	Prior below truth (2)	Prior above truth (3)
Panel A: Baseline			
Treatment (direction adj.)	−2.713*** [0.765]		
Treatment		−2.105* [1.182]	6.373*** [1.873]
Panel B: With controls			
Treatment (direction adj.)	−2.151** [0.913]		
Treatment		−2.066* [1.222]	7.106*** [2.066]
Panel C: Trimmed			
Treatment (direction adj.)	−2.467*** [0.797]		
Treatment		−2.230* [1.250]	5.711*** [2.042]
Observations	1,234	873	361
DV mean (control grp.)	2.743	3.472	0.885
DV std. dev. (control grp.)	16.35	18.34	9.41
DV mean (all)	2.998	2.062	5.263
DV std. dev. (all)	17.06	14.22	22.36

Notes. Table shows reduced-form estimates, reporting the effects of the experimental treatment on subjects' own protest participation. In columns (1) and (4), all subjects are pooled and treatment is coded as −1 for subjects whose priors regarding others' planned participation was above the true value of 17%. Columns (2)–(3) show effects on posterior beliefs for the subsample of subjects whose priors regarding others' planned participation were below (above) the true value of 17%. Panel A is estimated without any controls. Panel B replicates the analysis in Panel A but adds controls for subjects' background characteristics and economic status (gender, year of birth, a z-score index measuring whether subjects were raised in a Hong Kong–oriented environment, whether subjects were raised in a religious household, the economic status of the household, and own projected economic status), and for subjects' prior beliefs regarding HKUST students' planned participation in the July 1 march. Panel C estimates the baseline specification of Panel A, but excluding those individuals in the experimental sample with the 5% lowest and the 5% highest prior beliefs regarding other subjects' planned participation. Number of observations refers to that in the baseline specification. *: Significant at 10%; **: 5%; ***: 1%.

Table IV, Panel A, columns (2) and (3), we split the experimental sample into groups with priors regarding planned participation above and below the true level of planned participation of 17%,

treatment equal to −1) in Online Appendix Figure A.4. Specifically, we estimate the specification presented in Table IV, Panel A, column (1), and plot the point estimate (and 95% confidence interval) for integer switching points from 0 to 100. One can see in the right panel of the figure that we find the strongest reduced-form effects when the coding of treatment switches in the range between 15% and 30%.

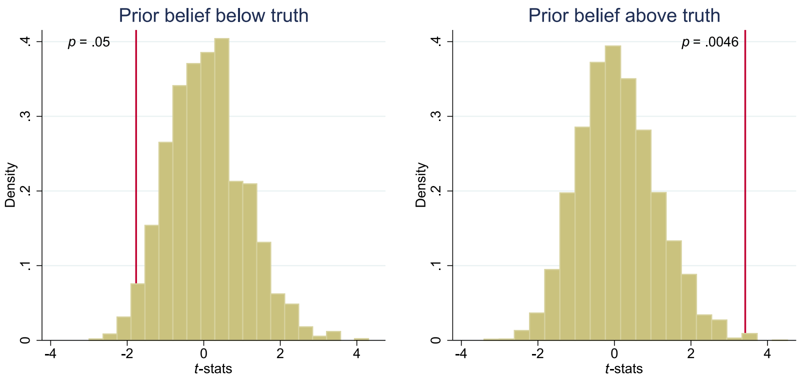


FIGURE VII

Distribution of t-Statistics from Estimating the Reduced-Form Effect of Treatment on Protest Participation Based on 10,000 Random Assignments of Treatment Status (Two-Thirds Treatment, One-Third Control) to Study Participants

Subsamples of subjects are divided according to whether beliefs regarding the planned protest participation of HKUST survey participants were above or below the true level of 17%. Red vertical lines indicate the t-statistics from the actual treatment assignment, with indicated p-values from two-sided tests.

and find significant effects in each subsample, matching the results shown in Figure VI (a test of equality of the coefficients estimated in Table IV, Panel A, columns (2) and (3) is rejected with $p < .001$).

As in the first-stage analysis, we address concerns about statistical inferences based on traditional standard errors in our study of the reduced-form effects. We again use randomization inference as an alternative, randomly assigning (fictional) treatment status and estimating reduced-form treatment effects 10,000 times each for the subsample of subjects with prior beliefs regarding others' planned participation below 17% and the subsample with prior beliefs regarding others' planned participation above 17%. We then compare the t-statistics from the estimated treatment effects from the fictional treatment assignments to the t-statistics from the actual treatment assignment (the actual estimates are those in Table IV, Panel A, columns (2) and (3)). We again find that our p-values using randomization inference, based on two-sided tests, are very similar to those using standard inference (see Figure VII).

The estimated reduced-form treatment effects in Table IV, Panel A, are not only statistically significant but also indicate a substantively significant effect of the information treatment on political behavior. One can see this by comparing the variation in protest participation explained by the treatment to the variation explained by the rich set of additional individual covariates we collected in Part 1 of the experiment. In Online Appendix Table A.7, one can see that the treatment actually has greater explanatory power than economic preferences, personality traits, cognitive ability, background characteristics, or economic status.

The magnitude of the treatment effect can also be benchmarked against estimated political mobilization effects in the existing literature in economics and political science. We follow DellaVigna and Gentzkow (2010) in calculating the persuasion rate implied by our treatment: the fraction of individuals who participate in the protest when treated but who would not have turned out to protest in the absence of the treatment. Among the subsample of individuals with priors above the information we provided—whom the treatment moved in the direction of greater participation—we find a persuasion rate of 6.43%. DellaVigna and Gentzkow (2010) calculate persuasion rates for seven interventions that stimulate turnout, from get-out-the-vote campaigns to media content; our treatment, though it did not directly encourage protest turnout, has effects that exceed most of the impersonal persuasive messages, for example, TV content or mass mailings encouraging voter turnout.[40]

In Table IV, Panels B and C, we conduct robustness exercises analogous to those considered in the first-stage analysis. One can see that neither including controls for subjects' prior beliefs regarding other subjects' planned participation, subjects' background characteristics, and economic status (Panel B) nor trimming the 5% of subjects with the lowest and the 5% of subjects with the highest prior beliefs regarding planned participation (Panel C), greatly affects the estimated reduced-form treatment effects.

40. The information treatment is not as effective as more personal interventions (e.g., door-to-door canvassing). For comparison, the findings in Gerber and Green (2000) imply persuasion rates of 1% for impersonal get-out-the-vote mailings and 15.6% for door-to-door visits. The magnitude of the treatment effect is sensible given the nature of our interactions with the students in our sample: our repeated interactions were more personal than a mass mailing or mass media campaign, but were less personal than a face-to-face interaction.

We explore two more dimensions of robustness in Online Appendix Table A.8. First, we consider sample restrictions depending on subjects' own plans to participate in the July 1 march. One might believe that individuals who reported, without any uncertainty, a plan not to attend the march (in Part 1 of the experiment) would be unlikely to have had their protest decision affected on the margin by the experimental treatment. Consistent with this expectation, dropping this group from our analysis results in larger estimated treatment effects than in the baseline specification. In contrast, one might expect individuals who reported, without any uncertainty, a plan to attend the march would most likely be closer to the margin—either confirmed in their plans or deterred from them by the information treatment. Consistent with this logic, when we drop subjects who planned to attend the march, we find attenuated (though still statistically significant) treatment effects. A second exercise considers an alternative construction of the outcome variable: an indicator of a change from plans in Part 1 of the experiment (no certain plan to attend) to behavior in Part 3 (actual attendance at the march). We find qualitatively similar results when we consider "changed plans" as the outcome variable.

We next examine heterogeneity in the reduced-form treatment effect associated with subjects' priors regarding the planned participation of others. In Figure III one could see that the largest first-stage effects were produced among individuals with prior beliefs far from the information provided by the experimental treatment. We assess whether the largest reduced-form effects are also seen among those subjects whose prior beliefs were far from the true level of 17%. To do so, we regress protest participation on the interaction between a treatment group dummy variable and 5-percentage-point bins of priors regarding other subjects' planned participation (as well as lower-order terms). In Figure VIII, one can see the estimated coefficients on the interaction terms and their 95% confidence intervals. The effect of the treatment on protest turnout was, indeed, greatest among individuals whose priors were furthest from the information provided. These findings provide reassuring evidence of consistency between the first-stage effects and the reduced form.[41]

41. A less parametric analysis of heterogeneity in the treatment effect associated with subjects' priors—locally weighted regression estimates of the treatment effect across subjects' priors—can be seen in Online Appendix Figure A.5. The patterns are broadly similar to those in Figure VIII.

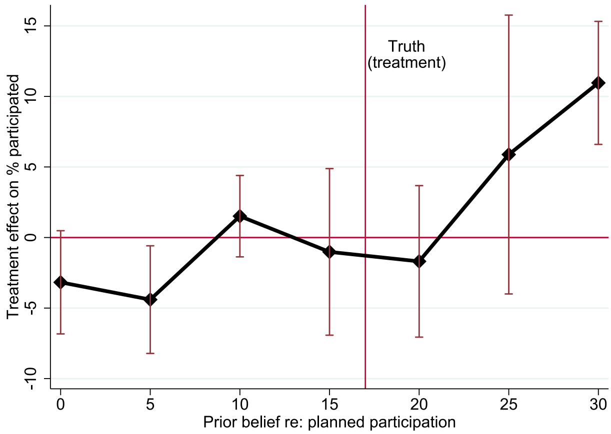


FIGURE VIII

Treatment Effect on Protest Participation (Reduced Form), Disaggregated by Prior Beliefs Regarding the Planned Participation of HKUST Survey Participants (Bins Are Five Percentage Points Wide)

Of course, the variation across subjects in their prior beliefs regarding others' planned participation is not experimentally induced, and is correlated with other variables that also shape the decision to protest (recall that predictors of prior beliefs can be seen in Online Appendix Table A.5). We next explore the extent to which variation in individual characteristics—rather than variation in prior beliefs—is likely to be behind the heterogeneous treatment effects observed in Figures III and VIII. We begin, in Table V, Panel A, column (1), by estimating a linear model of heterogeneous first-stage treatment effects associated with subjects' prior beliefs regarding other subjects' planned participation (roughly, fitting a line through Figure III). As expected, we estimate a significant, positive y-intercept and a significant, negative coefficient on the interaction between the treatment dummy and prior beliefs. In Table V, Panel B, column (1), we estimate a linear model of heterogeneous reduced-form treatment effects associated with subjects' prior beliefs regarding other subjects' planned participation (roughly, fitting a line through Figure VIII). As expected, we estimate a significant, negative y-intercept and a significant, positive coefficient on the interaction between the treatment dummy and prior beliefs.

TABLE V
HETEROGENEOUS TREATMENT EFFECTS BY PRIOR BELIEFS

	(1)	(2)	(3)
Panel A: Treatment effect on posterior beliefs (first stage)			
Treatment	7.417***	3.635	7.206***
	[0.827]	[3.314]	[2.311]
Treatment × prior beliefs	– 0.430***		
	[0.059]		
Treatment × predicted prior beliefs		– 0.204	– 0.419***
		[0.211]	[0.153]
Treatment × residual prior beliefs			– 0.441***
			[0.062]
Panel B: Treatment effect on protest participation (reduced form)			
Treatment	– 4.112***	– 10.970**	– 11.140**
	[1.444]	[4.822]	[4.679]
Treatment × prior beliefs	0.284***		
	[0.068]		
Treatment × predicted prior beliefs		0.691**	0.701**
		[0.341]	[0.331]
Treatment × residual prior beliefs			0.246***
			[0.068]
Observations	1,234	1,176	1,176
1st stage DV mean (control grp.)	14.04	14.10	14.10
1st stage DV std. dev. (control grp.)	14.10	14.24	14.24
1st stage DV mean (all)	14.50	14.50	14.50
1st stage DV std. dev. (all)	10.83	10.85	10.85
2nd stage DV mean (control grp.)	2.743	2.880	2.880
2nd stage DV std. dev. (control grp.)	16.35	16.75	16.75
2nd stage DV mean (all)	2.998	2.976	2.976
2nd stage DV std. dev. (all)	17.06	17.00	17.00

Notes. Table shows first-stage (Panel A) and reduced-form (Panel B) estimates, reporting the effects of the experimental treatment on subjects' posterior beliefs about others' participation and on subjects' own protest participation. In column (1), the model allows treatment effects to vary with subjects' prior beliefs regarding the planned protest participation of other experimental subjects. In column (2), the model allows treatment effects to vary with subjects' predicted prior beliefs based on observables (the predicted levels of priors regarding others' planned participation are estimated using a linear model including all 47 factors listed in Online Appendix Table A.5). In column (3), the model allows treatment effects to vary with subjects' residual prior beliefs, the component of beliefs not predicted by observables, as well as with subjects' predicted priors. All regressions include the relevant lower-order term for prior beliefs. *: Significant at 10%; **: 5%; ***: 1%.

We exploit our rich information about subjects to predict their priors using a linear model including all 47 factors listed in Online Appendix Table A.5. These predicted priors represent "problematic" variation in beliefs: heterogeneity in treatment effects associated with predicted priors may be driven by subjects' observable characteristics rather than by beliefs. Having calculated predicted priors for each individual, we are left with their residual priors—the component of prior beliefs unexplained by the large set of observable characteristics.[42] In Table V, Panels A and B, column (2), we estimate linear models of heterogeneous first-stage and reduced-form treatment effects associated with subjects' predicted prior beliefs; these estimates are of the same sign as but somewhat less precise than estimates using actual priors. In Table V, Panels A and B, column (3), we estimate linear models of heterogeneous first-stage and reduced-form treatment effects associated with both subjects' predicted prior beliefs and residual priors. One can see that while there exists statistically significant heterogeneity in the treatment effect associated with the "problematic" variation in predicted priors, there also exists statistically significant heterogeneity in the treatment effect associated with just the residual priors. This provides reassurance that the heterogeneous treatment effects we observe are indeed driven by variation in beliefs.

IV.C. Two-Stage Estimates: The Effects of Beliefs on Turnout

Thus far we have shown that providing information regarding the true level of planned protest turnout among our experimental sample caused (i) beliefs regarding actual turnout to change; and (ii) subjects' own turnout to change, with beliefs and turnout moving in opposite directions. We combine the two effects—first stage and reduced form—in a two-stage analysis that allows us to estimate the effect of beliefs regarding others' turnout on one's own turnout. It is worth emphasizing that one should not interpret the two-stage estimates too literally: we have already shown that the information treatment affected beliefs regarding the turnout of other experimental subjects and Hong Kong citizens more generally, so we cannot estimate the "pure" causal effect of a 1 percentage point change in beliefs regarding other subjects' turnout.[43]

42. Note that it is not the case that heterogeneity in treatment effects associated with predicted priors is necessarily driven by observables, rather than by beliefs, but simply that this cannot be ruled out. Note, too, that heterogeneity in treatment effects associated with residual priors is not necessarily driven by beliefs alone, as it might be driven by unobserved characteristics. Our rich data collection makes the latter less of a concern but does not eliminate it.

43. Another important reference population whose turnout might affect subjects' own is subjects' close friends. Although we do not observe prior and posterior beliefs regarding close friends' turnout, it is worth noting that we do not expect these beliefs to have been affected by our experimental intervention: students likely knew well whether their close friends would attend the July 1 march

With this caveat in mind, we still believe this exercise is instructive. Specifically, we predict individuals' protest turnout using their posterior beliefs regarding other experimental subjects' actual participation, instrumenting for the latter with the treatment dummy (when examining the pooled sample, we instrument for posteriors setting the treatment equal to -1 for individuals with prior beliefs regarding the planned participation of others greater than 17%).

We begin by estimating the two-stage model using the pooled sample, controlling only for subjects' prior beliefs regarding other subjects' actual turnout. In Table VI, Panel A, column (1), one can see that we estimate just over a 0.5 percentage point decrease in one's own turnout for every 1 percentage point increase in one's posterior beliefs regarding the turnout of others. In Table VI, Panel A, columns (2)–(3), we split the experimental sample into groups with priors regarding planned participation above and below the true level of planned participation of 17%, and find significant effects in each subsample. It is worth noting that although we had found larger first-stage and reduced-form treatment effects in the subsample with priors above 17% (see Tables III and IV), in Table VI, Panel A, columns (2) and (3), we find very similar estimated effects of beliefs about others' turnout on one's own turnout in the two subsamples—around a 0.5 percentage point effect (a test of equality of the coefficients estimated in Table VI, Panel A, columns (2) and (3), fails to reject the null, with $p = .585$).

In Table VI, Panels B and C, we conduct robustness exercises analogous to those considered in the first-stage and reduced-form analyses above. One can see that neither including controls for subjects' background characteristics and economic status (Panel B), nor trimming the 5% of subjects with the lowest and the 5% of subjects with the highest prior beliefs regarding planned participation (Panel C), greatly affects the estimated effects of

regardless of their treatment status. We explore heterogeneity in the first-stage and reduced-form treatment effects depending on whether subjects reported in Part 1 the expectation that a close friend would participate. We find that first-stage treatment effects are not significantly different depending on close friends' attendance, but the reduced-form treatment effects are larger among individuals' who expected a close friend to attend (see Online Appendix Table A.9). This reflects the fact that individuals on the margin of attendance (and so more responsive to the treatment) were differentially those who expected their friends to attend the July 1 march as well.

TABLE VI
TWO-STAGE ESTIMATES OF PROTEST PARTICIPATION

	Participated in 2016 July 1 march		
Sample	All subjects (1)	Prior below truth (2)	Prior above truth (3)
Panel A: Baseline			
Posterior belief	– 0.553***	– 0.468**	– 0.654**
	[0.177]	[0.236]	[0.264]
Panel B: With controls			
Posterior belief	– 0.554***	– 0.447*	– 0.660***
	[0.175]	[0.231]	[0.254]
Panel C: Trimmed			
Posterior belief	– 0.607***	– 0.514**	– 0.754**
	[0.212]	[0.261]	[0.354]
Observations	1,234	873	361
1st stage DV mean (control grp.)	14.04	8.44	28.30
1st stage DV std. dev. (control grp.)	14.10	8.46	15.54
1st stage DV mean (all)	14.50	11.40	22.02
1st stage DV std. dev. (all)	10.83	7.99	12.92
2nd stage DV mean (control grp.)	2.743	3.472	0.885
2nd stage DV std. dev. (control grp.)	16.35	18.34	9.41
2nd stage DV mean (all)	2.998	2.062	5.263
2nd stage DV std. dev. (all)	17.06	14.22	22.36

Notes. Table shows two-stage estimates of the effects of beliefs about other subjects' turnout in the protest on one's own turnout. The first stage estimates the effects of the experimental treatment on subjects' posterior beliefs regarding other subjects' actual protest participation, conditional on the corresponding prior beliefs. The second stage exploits variation in beliefs regarding other subjects' participation induced by the experimental treatment to estimate the effect of beliefs about others' protest turnout on one's own turnout. In column (1), all subjects are pooled and treatment is coded as −1 for subjects whose priors regarding others' planned participation were above the true value of 17%; pooled regression also controls for an indicator of whether subjects' priors regarding others' planned participation were above true value and its interaction with subjects' prior beliefs regarding others' actual participation. Columns (2)–(3) show estimates for the subsample of subjects whose priors regarding others' planned participation were below (above) the true value of 17%. Panel A is estimated without any controls beyond the corresponding prior beliefs. Panel B replicates the analysis in Panel A but adds controls for subjects' background characteristics and economic status (gender, year of birth, a z-score index measuring whether subjects were raised in a Hong Kong–oriented environment, whether subjects were raised in a religious household, the economic status of the household, and own projected economic status). Panel C estimates the baseline specification of Panel A, but excluding those individuals in the experimental sample with the 5% lowest and the 5% highest prior beliefs regarding other subjects' planned participation. Number of observations refers to that in the baseline specification. *: Significant at 10%; **: 5%; ***: 1%.

beliefs on protest turnout. Finally, in Online Appendix Table A.10, we replicate Table VI, but consider belief changes, rather than levels, as the endogenous regressor of interest (for which we use the treatment dummy as an instrument). We find similar results using this alternative specification.

IV.D. Internal Validity

We address several concerns regarding the internal validity of our estimates. First, as noted, we found that our treatment and control groups were unbalanced at the 5% level on 6 of 47 subject characteristics elicited in Part 1 (summary statistics for these variables and t-tests of equality of means between treatment and control groups are presented in Online Appendix Table A.1). To determine whether observable differences between treatment and control groups might drive our results, we estimate our baseline first-stage and reduced-form specifications (shown in Table III, Panel A, column (1) and Table IV, Panel A, column (1), respectively; replicated in Online Appendix Table A.2, column (1)), but controlling for the interaction between each unbalanced characteristic and the treatment dummy. One can see in Online Appendix Table A.2, columns (2)–(9), that controlling for these interactions leaves our treatment effect estimates practically unaffected.

A second concern is that selective attrition between Part 1 of our experiment and Part 3 might affect our estimates (though our finding of insignificant differences between the Part 1 sample's characteristics and the Part 3 sample's, in Online Appendix Table A.3, is reassuring). To address concerns about selective attrition, we reweight the sample of individuals who completed Part 3 to match the sample of individuals who completed Part 1 (but not necessarily the rest) of the study. We match the Part 1 sample on subjects' gender, birth year, childhood environment, religiosity, household income, and economic status. In Online Appendix Table A.11, one can see that our estimated first-stage and reduced-form treatment effects are largely unaffected by this reweighting exercise.[44]

Another possibility is that information regarding other subjects' protest plans not only affected beliefs about others' participation but also affected beliefs about the political movement

44. In Online Appendix Table A.11, we replicate our main results, but weighting observations to match the HKUST student population. We again find that our results are essentially unchanged. It is important to note that our reweighting exercises can only provide evidence of similar observable characteristics between our experimental sample and the reference samples: the HKUST student population or the sample of individuals who completed Part 1 of the study. We cannot rule out unobservable differences between these groups, though the similarity of observable characteristics is reassuring.

itself—as if the protest were a consumption good with uncertain quality. However, in this case, the standard social learning logic would suggest that positive updating of beliefs regarding the number of other subjects joining a protest should lead students to update positively about protest "quality." This would produce the appearance of strategic complementarity, not the strategic substitutability that we observe.

Standard experimenter demand effects would likely produce the appearance of strategic complementarity—when the experimenter tells a subject that a behavior is more prevalent than the subject expected, the subject seems likely to assume that the experimenter is trying to induce that behavior, and thus be more likely to report it. Further arguing against experimenter demand effects (or other unintended consequences of the experimental treatment) driving our findings, we also find evidence consistent with strategic substitutability in the protest decision even in the absence of any experimental treatment. Simply examining naturally occurring variation in beliefs among our control group subjects, one sees a negative relationship between posterior beliefs regarding the actual participation of others and subjects' own participation in the July 1 march (see Online Appendix Figure A.6).[45]

V. Discussion

Our experiment estimates the causal effect of beliefs about others' protest participation on one's own, exploiting variation in beliefs driven by a specific level of planned turnout, from a single protest, in a particular setting. One naturally wonders how to generalize from such an estimate. In Section II.B, we discussed several important dimensions along which Hong Kong's July 1 marches share similarities with other antiauthoritarian protests:

45. We also observe a negative relationship between subjects' self-reported past protest participation and their beliefs about other subjects' support for the antiauthoritarian movement, in responses recorded in Part 1 (see Online Appendix Figure A.7). It is important to emphasize that, while suggestive, the correlations observed in Online Appendix Figures A.6 and A.7 must be interpreted with caution. Among other concerns, the levels of posterior beliefs in the control group are quite strongly correlated with their accuracy, and subjects with more accurate beliefs may turn out more than others for reasons other than strategic considerations. Past protest participation may shape beliefs about others' ideology, in addition to being shaped by them.

as in other protests, the Hong Kong marches aim to achieve concrete political changes by mobilizing large numbers of individuals, with these individuals facing a potential government crackdown. Thus, on its face, there is reason to expect that protest behavior in the July 1 march will have some degree of external validity to other antiauthoritarian protests.

Yet it must be acknowledged that even in the same context we study, even in the specific protest we study, it is possible that at a different level of protest participation the causal effect of beliefs about others' turnout on one's own could differ—indeed, it may even flip signs. Given the possible heterogeneity in the treatment effect we estimate, an important question is to what extent does our single empirical parameter estimate bound the set of theoretical models of protest participation?

Although our estimate does not pin down a single model of protest participation, we argue that it does provide important bounds on the set of models consistent with our estimate: much recent theoretical work on protest participation assumes only the possibility of strategic complementarity in the protest decision. This assumption has become so widespread that we treat it as a "benchmark" model in what follows. Our findings reject this benchmark model. Although we cannot claim that a specific alternative model produces our findings of strategic substitutability, we provide evidence consistent with three plausible mechanisms that could generate strategic substitutability in our setting and could easily be incorporated into the benchmark model to allow for the strategic substitutability that we find.

V.A. The Benchmark Model

The recent political economy literature has typically modeled protest participation as a global game (or similar), with the stage game featuring strategic complementarity, and by assumption ruling out the possibility of strategic substitutes. Here we present a simple, but general, version of the "benchmark" model of the stage game and relate it to several recent articles that emphasize different underlying sources of strategic complementarity.

In the benchmark model, individual *i*'s utility from protest participation can be written as follows:

$$U_i = \mathbb{1}_{P_i=1}(V_i(P_{-i}, S(P_{-i})) - C_i(P_{-i}, S(P_{-i}))),$$

where $\mathbb{1}_{P_i=1}$ is an indicator that individual i participated in the protest and the utility from participating (U_i) is a function of the benefits from participation (V_i) and the costs (C_i). Costs and benefits, in turn, are assumed to depend directly on the participation of other individuals (the level of which is denoted P_{-i}), as well as on the success of a protest (an indicator denoted S), which itself may depend on the participation of others.

The benchmark model first follows the recent literature in assuming that protests are more likely to succeed when they are larger and that individuals derive differential utility from participating in protests that succeed (or that are more likely to succeed). Bueno de Mesquita (2010, 449), for example, assumes that "there is some portion ... of the payoffs from regime change that can only be accessed by those who participate ... Substantively, this could be because those who actively participate in revolution gain privileged status after regime change occurs or because there are expressive benefits to having participated in a victorious uprising." Edmond (2013, 1425) writes simply that "the more citizens participate in [a protest], the more likely it is that the regime is overthrown and so the more likely it is that any individual also participates." Passarelli and Tabellini (2017, 910) write that individuals may derive greater utility from the "feeling of contributing to a more meaningful event with a greater chance of success." Barberà and Jackson (2018, 7–8) assume that individuals attain utility "from knowing or being able to say that they participated in a revolution that was successful: from having been one of those who stormed the Bastille, participated in the Salt March, or protests in Tunisia, etc."

Formally, the benchmark model assumes:

$$\frac{\partial S}{\partial P_{-i}} > 0 \qquad \text{and} \qquad \frac{\partial V}{\partial S} > 0.$$

It is worth noting that the crucial assumption that a protest is perceived to be more likely to succeed when more people turn out (i.e., $\frac{\partial S}{\partial P_{-i}} > 0$) receives empirical support in our experimental context. We elicit experimental subjects' beliefs about the likelihood that a Hong Kong prodemocracy protest would succeed as a function of protest turnout and find that these are strongly increasing (see Figure I).

The benchmark model also assumes that the benefits from participating in a protest are larger when it is attended by more people—even if the increased participation does not affect the protest's success. This may be, for example, because participation in a larger protest increases the expressive utility a participant attains. In a model of citizen unrest in response to perceived unfair government policy, Passarelli and Tabellini (2017, 910) write that “the psychological benefit of a public display of anger is stronger if the emotion is more widely shared.”

Thus, whether because of direct effects or because of indirect effects working through protest success, the benchmark model assumes that the benefits of protest attendance increase with the attendance of others—thus generating strategic complementarity. Formally:

$$\frac{\mathrm{d}\,V}{\mathrm{d}\,P_{-i}} > 0.$$

The benchmark model assumes similar effects of others’ turnout arising from the costs side. The cost of protest attendance may be lower when protests are successful—perhaps because in a successful protest the regime concedes, rather than cracks down. If, as assumed above, protest success is an increasing function of turnout, then greater turnout will lower protest costs, alongside increasing benefits. The cost of attendance may also be falling in protest size independent of success if costs per person are lower in a larger crowd. For example, Passarelli and Tabellini (2017, 910) write that “[strategic] complementarity could also be on the cost side: the probability of being arrested is smaller in a larger crowd.”

Thus, the benchmark model assumes:

$$\frac{\partial C}{\partial S} < 0.$$

This assumption, along with the above assumption that $\frac{\partial S}{\partial P_{-i}} > 0$, as well as direct reductions of costs arising from larger turnout, imply another force generating strategic complementarity:

$$\frac{\mathrm{d}\,C}{\mathrm{d}\,P_{-i}} < 0.$$

In sum, the benchmark model—synthesizing several recent articles—allows for different underlying sources of strategic complementarity.[46] All of these generate a common prediction: if an individual believes that the turnout of others will be greater, she is unambiguously more likely to turn out to a protest herself.

V.B. Sources of Strategic Substitutability in the Protest Decision

Our findings reject the benchmark model of protest turnout that assumes only the possibility of strategic complementarity in the decision to protest. It is important to emphasize that we do not claim that protests are always games of strategic substitutes, but that models of protest turnout must allow for the possibility of strategic substitutability. We propose adding to the benchmark model's stage game utility function a term that allows for strategic substitutability: specifically, a term whose first derivative with respect to protest size is negative, at least over some range.

While we cannot conclusively identify specific mechanisms that produce the pattern of strategic substitutability we observe, we are able to provide suggestive evidence of three plausible drivers: first, that protests produce a public good from attaining a threshold level of participation; second, that an individual's expected costs of protest turnout increase with protest size; and third, that an individual's benefits from signaling her antiauthoritarian type decrease with protest size.

1. Threshold Public Goods as a Function of Turnout. As described above, the benchmark model assumed $\frac{\mathrm{d}V}{\mathrm{d}P_{-i}} > 0$, arising from several sources of differential utility experienced when participating in larger, more successful protests. However, there may be benefits from one's own protest turnout that over some range are decreasing in the turnout of others. A classic argument in the political economy literature (Olson 1965; Tullock 1971; Palfrey and Rosenthal 1984) is that protest participation produces a political public good. This good might be tangible: attaining a sufficient turnout level might send the regime a signal that directly affects policy (for example, in the 2017 march, significant turnout might have persuaded Leung not to run for a second term as chief executive). The public good produced by a protest might also be symbolic: sufficient turnout might signal the existence of a "critical mass" of protesters, which sustains the political

46. This literature (Bueno de Mesquita 2010; Edmond 2013; Barberà and Jackson 2018; Passarelli and Tabellini 2017) is discussed in more detail in Online Appendix B.

movement; note that this signal is not inconsequential, as it shapes subsequent protest turnout in the political movement.[47]

We find evidence consistent with protest turnout being a public good in our setting. We construct an index of subjects' prosociality (altruism and reciprocity) using their responses in Part 1 of the study. We find that this prosociality index is significantly correlated with protest turnout unconditionally, as well as conditional on treatment status, posterior beliefs about others' turnout, and on subjects' own antiauthoritarian ideology (see Online Appendix Table A.12). This is consistent with a theoretical and empirical literature linking prosociality to public goods contributions (see, e.g., Croson 2007).

We can incorporate the production of a public good—either an intermediate, tangible public good (something meaningful but short of democratization) or a symbolic public good—from sufficient protest turnout into our model by adding a term, $\phi(P_{-i})$, to the benefits function:

$$V_i = V_i(S(P_{-i}), \phi(P_{-i})).$$

In this case, the benefits from turnout are a function of (i) the success of the protest (i.e., achieving democratization) $S(P_{-i})$, which is monotonically increasing in P_{-i} and will tend to produce strategic complementarity; and (ii) the production of an intermediate tangible public good or a symbolic public good $\phi(P_{-i})$, which could be increasing in P_{-i} over some range (i.e., below the level at which the public good is completely produced), but decreasing in P_{-i} around the point at which individuals believe their turnout is no longer needed to produce the public good.

Over some range, particularly in a protest in which the production of the public good (e.g., signaling the movement's strength) is more relevant than the achievement of the ultimate success of the movement (e.g., achieving democratization), the production of a public good from protest participation may generate strategic substitutability.[48] Thus, in this case, $\frac{\mathrm{d}V}{\mathrm{d}P_{-i}}$ may be positive or negative, thus allowing for strategic complementarity or substitutability.

47. Barberà and Jackson (2018) present a model in which protest turnout in one period affects beliefs about eventual movement success, and thus turnout in subsequent periods.

48. Note that a public goods game may not be a game of strategic substitutes in all settings: subjects may play a strategy of conditional cooperation, which

2. Increasing Costs as Turnout Rises. The benchmark model assumed that an individual's expected cost of protest attendance was falling with the expected turnout of others. The logic was that the costs of a government crackdown would be diffused across more people, thus reducing an individual's own cost. However, this logic ignores the possibility that the likelihood of a government crackdown is increasing in protest size, at least over some range. Consistent with this proposed mechanism behind strategic substitutability in our setting, one can see in Figure I that our experimental subjects believe that a government crackdown is an increasing function of protest size.

This source of strategic substitutability can easily be incorporated into the benchmark model. Suppose the cost of turnout is a function of (i) the cost borne by a protester conditional on crackdown, $c_i(P_{-i})$; (ii) the success of the protest, $S(P_{-i})$; and (iii) the probability of crackdown, $\pi(P_{-i})$. Then, the cost function would be:

$$C_i = C_i(\pi(P_{-i}) \times c_i(P_{-i}), S(P_{-i})),$$

with $\frac{\partial c_i}{\partial P_{-i}} < 0$ (following the logic of the benchmark model), $\frac{\partial S}{\partial P_{-i}} > 0$ (again, following the benchmark model), and $\frac{\partial \pi}{\partial P_{-i}} > 0$. In this case, $\frac{\mathrm{d}C}{\mathrm{d}P_{-i}}$ may be positive or negative, allowing for strategic complementarity or substitutability.

3. Decreasing Signaling Value of Participation as Turnout Rises. We finally propose an additional element of the benefits of protest participation, which may be greater when participating in smaller protests: if one cares about signaling one's antiauthoritarian type, the value of that signal, $\theta_i(P_{-i})$, may be decreasing in protest size, at least over some range.[49]

generates strategic complementarity (see, e.g., Fischbacher, Gächter, and Fehr 2001).

49. We follow the logic of Bénabou and Tirole (2011), who present a model in which an individual trades off honor from taking a costly, meritorious action against stigma from not taking that action. When fewer individuals undertake the costly action, it increases the honor associated with taking the action (because the signal of one's type conditional on taking the action becomes more positive), but also reduces the stigma associated with not taking the action (because the signal of one's type conditional on not taking the action becomes less negative). Under a small set of assumptions, the honor associated with taking a less common action increases more than the stigma from not taking the action falls when the base rate

We find several pieces of suggestive evidence consistent with individuals' protest participation being shaped by their political identities (i.e., types), and the signaling of those identities. First, we find that students' antiauthoritarian identities (self-reported and revealed in incentivized lab experiments) are strongly associated with their protest participation plans (see Online Appendix Table A.5). In a subsequent survey of participants in the July 1, 2017 march in which we directly asked protest participants why they attended the march, 45% of participants responded, "Being politically active is an important component of my identity."[50]

If subjects selected into smaller protests in response to the perceived signaling value, one would expect individuals with very antiauthoritarian friends—to whom signaling one's type would be especially valuable—to have differentially attended the July 1 march after updating their beliefs about total protest size downward. Analogously, one would expect that individuals who attend a protest specifically to signal their type to their antiauthoritarian friends might select out of attendance after updating their beliefs about total protest size upward. Consistent with this prediction, we find that treatment group protest participants with high priors regarding other subjects' planned participation—a group who turned out to protest after updating their beliefs regarding protest turnout downward—have more antiauthoritarian friends than protest participants in the control group (see Online Appendix Figure A.8). We find analogous results among individuals selecting out of protests when updating their beliefs about turnout upward. This suggests that subjects who had the strongest social-image reasons to signal their antiauthoritarian identities differentially selected into protests they believed would

of taking the action is rare. The logic is as follows: when a meritorious action is taken by a small minority in the tail of a distribution, the conditional mean moving further into the tail changes more in the minority tail of the distribution than in the majority. This model produces strategic substitutes in taking the meritorious action when starting from a low base, but strategic complements when starting from a high base. Enikolopov et al. (2018) present a dynamic model in which social-image considerations motivate persistent protest participation over time.

50. The participants in the 2017 survey included some individuals surveyed as part of our experiment conducted in 2016 as well as additional respondents. Among the 2017 survey sample, we directly elicited all 59 protest participants' reasons for attending the 2017 march. Individuals' identity was the second most frequent response, with 27 individuals reporting it. The most frequent response was, "I believed the march would produce political change," reported by 28 individuals.

be smaller and selected out of protests they believed would be larger.

We can incorporate the signaling value of protest attendance into a benefits function alongside a threshold public good. Suppose the benefits from turnout are a function of (i) the success of the protest, $S(P_{-i})$, discussed above; (ii) the production of a public good, $\phi(P_{-i})$, also discussed above; and (iii) the value of the signal of one's type, $\theta_i(P_{-i})$. Then, the benefits function would be:

$$V_i = V_i(S(P_{-i}), \phi(P_{-i}), \theta_i(P_{-i})),$$

with $\frac{\partial V}{\partial S} > 0$ (as in the benchmark model), $\frac{\partial \phi}{\partial P_{-i}} < 0$ (over some range), and $\frac{\partial \theta}{\partial P_{-i}} < 0$ (when P_{-i} is small enough). In this case, $\frac{\mathrm{d}\, V}{\mathrm{d}\, P_{-i}}$ may be positive or negative, again allowing for strategic complementarity or substitutability.

Thus, while only suggestive, the patterns in our data are consistent with plausible mechanisms underlying strategic substitutability in our setting. Importantly, these mechanisms can easily be incorporated into the benchmark model to allow for either strategic complementarity or substitutability in the protest decision, making it consistent with our main findings.[51]

VI. Conclusion

According to the human rights organization Freedom House, as of 2016, 26% of the world's population—nearly two billion people—live in states classified as "not free." Recent protests from the Arab Spring to Russia or Venezuela provide reminders that citizens of unfree states today, as in the past, continually rise up

51. There exist other potential mechanisms that could generate strategic substitutability for which we do not find evidence in our setting (though they may be present in other protests). Shadmehr and Bernhardt (2011) propose a model in which moderates update their beliefs about the gains from protest negatively if they believe other agents' thresholds for turnout are too low. This makes moderates select out of protesting when they update their beliefs positively—the opposite of what we find in Online Appendix Figure A.8. Myatt (2017) proposes a model in which protest voters may moderate their expression if they believe that extremists are too popular. Again, this model would suggest that moderates select out of protests when they update their beliefs positively, in contrast to our findings. Finally, Shadmehr (2018) proposes protester perceptions of pivotality as potential drivers of strategic substitutability; we did not ask subjects about these perceptions.

and demand political rights.[52] Given the prevalence of authoritarian regimes, it is important to understand individuals' decisions to participate in such protest movements.[53]

We conduct an experiment to study one dimension of individuals' protest decision: how one's turnout is affected by beliefs about the turnout of others. We find consistent evidence indicating that Hong Kong students considering participation in the July 1 protest of 2016 viewed the strategic element of their decision as a game of strategic substitutes. Individuals in our sample who were induced by the experiment to positively update their beliefs about others' turnout became less likely to participate themselves; individuals who were induced by the experiment to negatively update their beliefs about others' turnout became more likely to participate themselves. Although our finding of strategic substitutability in the protest decision is estimated from a particular context, the result provides guidance regarding models of protest turnout: specifically, our findings reject many recent models that assume only the possibility of strategic complementarity in the protest decision.

While shedding light on individuals' decisions to participate in an antiauthoritarian protest, our experiment raises the question: how general is our finding of strategic substitutability likely to be? It is important to emphasize that the mechanisms we observe at work in Hong Kong will not all be present in all protests, and even if they are present, they will not always outweigh forces generating strategic complementarity emphasized in other work.

Among the mechanisms we highlight, we believe that the existence of a political public good produced from attaining a threshold level of participation is likely to be particularly important in

52. Data come from Freedom House's (2016) "Freedom in the World" report, https://freedomhouse.org/sites/default/files/FH_FITW_Report_2016.pdf, accessed August 5, 2016.

53. Research on authoritarian regimes, movements opposed to them, and the consequences of constraints on rulers typically considers aggregate behavior, rather than individual behavior, as we do. A large literature has studied the consequences of political constraints for economic growth (e.g., DeLong and Shleifer 1993; Przeworski and Limongi 1993; Przeworski et al. 2000; Gerring et al. 2005; Rodrik and Wacziarg 2005; Persson and Tabellini 2006, 2008; Papaioannou and Siourounis 2008; Acemoglu and Robinson 2012; Bates, Fayad, and Hoeffler 2012; Meyersson 2016; Acemoglu et al. 2015). Relatedly, a growing theoretical and empirical literature studies the extension of the franchise (e.g., Acemoglu and Robinson 2000, 2006; Lizzeri and Persico 2004; Llavador and Oxoby 2005; Aidt and Franck 2012, 2015).

long-lasting political movements, for example, the women's suffrage movements in the early twentieth century United States and Britain; the civil rights movement in the 1960s United States; and political movements in unconsolidated democracies, like Russia's ongoing anticorruption protests and protests in Venezuela. In these settings, getting people on the street may in itself be an aim of the protest: achieving target protest sizes could directly achieve policy concessions, and also signal the strength of a movement to potential future participants. Strategic substitutability is most likely during stages of a political movement when the public goods component of a protest looms large relative to a movement's ultimate success. This does not make the protest inconsequential: tangible achievements are possible even in the absence of complete success; movement survival and growth in the future may depend on turnout in the present. It is important to emphasize, however, that the dynamic signaling element of this mechanism will likely not be relevant in one-shot mass events, like the Arab Spring protests, which will either topple a regime or be crushed. One-shot protests are more likely to exhibit strategic complementarity.

The second force for strategic substitutability in our setting is subjects' perception of an increase in the probability of a government crackdown for larger protests. This force seems most likely to exist in protests demanding political rights from regimes that are not totalitarian (again, the women's suffrage movements, the civil rights movement, and protests in partial democracies). Many regimes, even authoritarian ones, allow some freedom of expression and association; thus, small protests are likely to be tolerated. These regimes may not tolerate destabilizing large events, and thus may be perceived as being more likely to crack down on larger protests. This mechanism producing strategic substitutability is unlikely to be present in settings where protests are patently illegal, such as the Soviet Union and much of Eastern Europe prior to 1989, North Korea, or in much of the Arab world. In these cases, a government crackdown is essentially guaranteed, so the probability of bearing a private cost conditional on a crackdown becomes the only consideration. In such settings, larger protests are likely to be perceived as lower cost, thus pushing toward strategic complementarity.

Finally, the Bénabou and Tirole (2011) logic of greater individual signaling value of attending a smaller protest seems likely to be quite general: political radicals will be able to more strongly signal their type when a protest is small. It is important to note, however, that the magnitude of this effect may differ substantially across types of protesters and across protests. Again, this force toward substitutability may be especially pronounced in ongoing movements, in which participation becomes part of one's identity and may be less relevant in one-shot events aimed at toppling a dictator.

Future work thus should shed more light on the mechanisms that underlie the strategic interactions in the protest game, ideally exploring their heterogeneity across varying political settings. It should also engage in the empirical analysis of protests that are linked dynamically as part of larger political movements. In our setting, a large July 1 march not only affects policy immediately but also shapes the beliefs and behavior of citizens in the antiauthoritarian movement more broadly. This means there are dynamic considerations among protesters, learning among protest participants and nonparticipants, and potentially important cross-protest spillovers. Given that many protests are not solitary events, understanding these dynamic processes is an important area for future research.

Ludwig-Maximilians-Universität Munich, Centre for Economic Policy Research, and CESifo
Harvard University and Abdul Latif Jameel Poverty Action Lab
London School of Economics, National Bureau of Economic Research, and CESifo
Hong Kong University of Science and Technology

Supplementary Material

An Online Appendix for this article can be found at *The Quarterly Journal of Economics* online. Data and code replicating tables and figures in this article can be found in Cantoni et al. (2019), in the Harvard Dataverse, doi:10.7910/DVN/GVOMUR.

References

Acemoglu, Daron, Tarek A. Hassan, and Ahmed Tahoun, "The Power of the Street: Evidence From Egypt's Arab Spring," NBER Working Paper No. 20665, 2014.

Acemoglu, Daron, Suresh Naidu, Pascual Restrepo, and James A. Robinson, "Democracy Does Cause Growth," MIT Working Paper, 2015, http://economics.mit.edu/files/11227.

Acemoglu, Daron, and James A. Robinson, "Why Did the West Extend the Franchise? Democracy, Inequality, and Growth in Historical Perspective," *Quarterly Journal of Economics*, 115 (2000), 1167–1199.

———, *Economic Origins of Dictatorship and Democracy* (Cambridge: Cambridge University Press, 2006).
———, *Why Nations Fail: The Origins of Power, Prosperity, and Poverty* (New York: Crown Business, 2012).
Aidt, Toke S., and Raphaël Franck, "How to Get the Snowball Rolling and Extend the Franchise: Voting on the Great Reform Act of 1832," *Public Choice*, 155 (2012), 229–250.
———, "Democratization under the Threat of Revolution: Evidence from the Great Reform Act of 1832," *Econometrica*, 83 (2015), 505–547.
Anderson, Michael L., "Multiple Inference and Gender Differences in the Effects of Early Intervention: A Reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects," *Journal of the American Statistical Association*, 103 (2008), 1481–1495.
Barberà, Salvador, and Matthew O. Jackson, "A Model of Protests, Revolution, and Information," Working Paper, Stanford University, 2018, https://ssrn.com/abstract=2732864.
Bates, Robert H., Ghada Fayad, and Anke Hoeffler, "The State of Democracy in Sub-Saharan Africa," *International Area Studies Review*, 15 (2012), 323–338
Bénabou, Roland, and Jean Tirole, "Laws and Norms," NBER Working Paper No. 17579, 2011.
Bueno de Mesquita, Ethan, "Regime Change and Revolutionary Entrepreneurs," *American Political Science Review*, 104 (2010), 446–466.
Cantoni, Davide, David Y. Yang, Noam Yuchtman, and Jane Zhang, "The Fundamental Determinants of Anti-Authoritarianism," Working Paper, UC Berkeley, 2016, http://faculty.haas.berkeley.edu/yuchtman/Noam_Yuchtman_files/antiauthoritarian_draft.pdf.
———, "Replication Data for: 'Protests as Strategic Games: Experimental Evidence from Hong Kong's Antiauthoritarian Movement'," (2019), Harvard Dataverse, doi: 10.7910/DVN/GVOMUR.
Chwe, Michael Suk-Young, "Communication and Coordination in Social Networks," *Review of Economic Studies*, 67 (2000), 1–16.
Coffman, Katherine B., Lucas C. Coffman, and Keith M. Marzilli Ericson, "The Size of the LGBT Population and the Magnitude of Anti-Gay Sentiment Are Substantially Underestimated," *Management Science*, 63 (2017), 3168–3186.
Coffman, Lucas C., Clayton R. Featherstone, and Judd B. Kessler, "A Model of Information Nudges," Working Paper, Ohio State University, 2015, https://site.stanford.edu/sites/default/files/a_model_of_information_nudges_20151124.pdf.
Croson, Rachel, "Theories of Commitment, Altruism and Reciprocity: Evidence from Linear Public Goods Games," *Economic Inquiry*, 45 (2007), 199–216.
Deaton, Angus, and Nancy Cartwright, "Understanding and Misunderstanding Randomized Controlled Trials," *Social Science & Medicine*, 210 (2018), 2–21.
DellaVigna, Stefano, Ruben Enikolopov, Vera Mironova, Maria Petrova, and Ekaterina Zhuravskaya, "Cross-Border Media and Nationalism: Evidence from Serbian Radio in Croatia," *American Economic Journal: Applied Economics*, 6 (2014), 103–132.
DellaVigna, Stefano, and Matthew Gentzkow, "Persuasion: Empirical Evidence," *Annual Review of Economics*, 2 (2010), 643–669.
DeLong, J. Bradford, and Andrei Shleifer, "Princes and Merchants: European City Growth before the Industrial Revolution," *Journal of Law and Economics*, 36 (1993), 671–702.
Edmond, Chris, "Information Manipulation, Coordination, and Regime Change," *Review of Economic Studies*, 80 (2013), 1422–1458.
Egorov, Georgy, and Konstantin Sonin, "Elections in Non-Democracies," Working Paper, University of Chicago, 2018.
Enikolopov, Ruben, Alexey Makarin, and Maria Petrova, "Social Media and Protest Participation: Evidence from Russia," Working Paper, Universitat Pompeu Fabra, 2016, https://ssrn.com/abstract=2696236.

Enikolopov, Ruben, Alexey Makarin, Maria Petrova, and Leonid Polishchuk, "Social Image, Networks, and Protest Participation," Working Paper, Universitat Pompeu Fabra, 2018.

Fearon, James D., "Self-Enforcing Democracy," *Quarterly Journal of Economics*, 126 (2011), 1661–1708.

Fischbacher, Urs, Simon Gächter, and Ernst Fehr, "Are People Conditionally Cooperative? Evidence from a Public Goods Experiment," *Economics Letters*, 71 (2001), 397–404.

Gehlbach, Scott, Konstantin Sonin, and Milan W. Svolik, "Formal Models of Nondemocratic Politics," *Annual Review of Political Science*, 19 (2016), 565–584.

Gerber, Alan S., and Donald P. Green, "The Effects of Canvassing, Telephone Calls, and Direct Mail on Voter Turnout: A Field Experiment," *American Political Science Review*, 94 (2000), 653–663.

Gerber, Alan, Mitchell Hoffman, John Morgan, and Collin Raymond. "One in a Million: Field Experiments on Perceived Closeness of the Election and Voter Turnout," Working Paper, University of Toronto, 2017.

Gerber, Alan S., Gregory A. Huber, David Doherty, Conor M. Dowling, Connor Raso, and Shang E. Ha, "Personality Traits and Participation in Political Processes," *Journal of Politics*, 73 (2011), 692–706.

Gerring, John, Philip Bond, William Barndt, and Carola Moreno, "Democracy and Growth: A Historical Perspective," *World Politics*, 57 (2005), 323–364.

González, Felipe, "Collective Action in Networks: Evidence from the Chilean Student Movement," Working Paper, UC-Berkeley, 2016.

Granovetter, Mark, "Threshold Models of Collective Behavior," *American Journal of Sociology*, 83 (1978), 489–515.

Hobbs, William, and Margaret E. Roberts, "How Sudden Censorship Can Increase Access to Information," *American Political Science Review*, 112 (2018), 621–636.

King, Gary, Jennifer Pan, and Margaret E. Roberts, "How Censorship in China Allows Government Criticism but Silences Collective Expression," *American Political Science Review*, 107 (2013), 326–343.

Kricheli, Ruth, Yair Livne, and Beatriz Magaloni, "Taking to the Streets: Theory and Evidence on Protests under Authoritarianism," APSA Annual Meeting Paper, 2011.

Kuran, Timur, "Sparks and Prairie Fires: A Theory of Unanticipated Political Revolution," *Public Choice*, 61 (1989), 41–74.

———, "The East European Revolution of 1989: Is it Surprising that We Were Surprised?," *American Economic Review*, 81 (1991), 121–125.

———, *Private Truths, Public Lies: The Social Consequences of Preference Falsification* (Cambridge, MA: Harvard University Press, 1997).

Lizzeri, Alessandro, and Nicola Persico, "Why Did the Elites Extend the Suffrage? Democracy and the Scope of Government, with an Application to Britain's 'Age of Reform'," *Quarterly Journal of Economics*, 119 (2004), 707–765.

Llavador, Humberto, and Robert J. Oxoby, "Partisan Competition, Growth, and the Franchise," *Quarterly Journal of Economics*, 120 (2005), 1155–1189.

Lohmann, Susanne, "The Dynamics of Informational Cascades: The Monday Demonstrations in Leipzig, East Germany, 1989–91," *World Politics*, 47 (1994), 42–101.

Lorentzen, Peter, "Regularizing Rioting: Permitting Public Protest in an Authoritarian Regime," *Quarterly Journal of Political Science*, 8 (2013), 127–158.

Madestam, Andreas, Daniel Shoag, Stan Veuger, and David Yanagizawa-Drott, "Do Political Protests Matter? Evidence from the Tea Party Movement," *Quarterly Journal of Economics*, 128 (2013), 1633–1685.

Manacorda, Marco, and Andrea Tesei, "Liberation Technology: Mobile Phones and Political Mobilization in Africa," Working Paper, Queen Mary University of London, 2016.

Meyersson, Erik, "Political Man on Horseback: Military Coups and Development." Working Paper, SITE, Stockholm School of Economics, 2016, https://erikmeyersson.files.wordpress.com/2015/04/coups_meyersson_1504.pdf.
Myatt, David P., "A Theory of Protest Voting," *The Economic Journal*, 127 (2017), 1527–1567.
Olson, Mancur. *The Logic of Collective Action: Public Goods and the Theory of Groups* (New York: Schocken Books, 1965).
Opp, Karl-Dieter, and Christiane Gern, "Dissident Groups, Personal Networks, and Spontaneous Cooperation: The East German Revolution of 1989," *American Sociological Review*, 58 (1993), 659–680.
Palfrey, Thomas, and Howard Rosenthal, "Participation and the Provision of Discrete Public Goods: A Strategic Analysis," *Journal of Public Economics*, 24 (1984), 171–193.
Papaioannou, Elias, and Gregorios Siourounis, "Democratisation and Growth," *Economic Journal*, 118 (2008), 1520–1551.
Passarelli, Francesco, and Guido Tabellini, "Emotions and Political Unrest," *Journal of Political Economy*, 125 (2017), 903–946.
Persson, Torsten, and Guido Tabellini, "Democracy and Development: The Devil in the Details," *American Economic Review*, 96 (2006), 319–324.
———, "The Growth Effect of Democracy: Is it Heterogenous and How Can It Be Estimated?," In *Institutions and Economic Performance,* Elhanan Helpman, ed. (Cambridge, MA: Harvard University Press, 2008).
Przeworski, Adam, Michael E. Alvarez, José Antonio Cheibub, and Fernando Limongi. *Democracy and Development: Political Institutions and Well-Being in the World, 1950–1990* (Cambridge: Cambridge University Press, 2000).
Przeworski, Adam, and Fernando Limongi, "Political Regimes and Economic Growth," *Journal of Economic Perspectives*, 7 (1993), 51–69.
Raghavarao, Damaraju, and Walter T. Federer, "Block Total Response as an Alternative to the Randomized Response Method in Surveys," *Journal of the Royal Statistical Society, Series B*, 41 (1979), 40–45.
Reny, Marie-Eve, "Authoritarianism as a Research Constraint: Political Scientists in China," *Social Science Quarterly*, 97 (2016), 909–922.
Rodrik, Dani, and Romain Wacziarg, "Do Democratic Transitions Produce Bad Economic Outcomes?," *American Economic Review*, 95 (2005), 50–55.
Shadmehr, Mehdi, "Tullock's Paradox, Pivotal Revolutionaries and the Strength of Weak States," Working Paper, University of Calgary, 2018.
Shadmehr, Mehdi, and Dan Bernhardt, "Collective Action with Uncertain Payoffs: Coordination, Public Signals and Punishment Dilemmas," *American Political Science Review*, 105 (2011), 829–851.
Tullock, Gordon, "The Paradox of Revolution," *Public Choice*, 11 (1971), 89–99.
Yanagizawa-Drott, David, "Propaganda and Conflict: Evidence from the Rwandan Genocide," *Quarterly Journal of Economics*, 129 (2014), 1947–1994.