CAGE

Can Crises Affect Citizen Activism? Evidence from a Pandemic

CAGE working paper no. 693

November 2023

Farzana Afridi,
Ahana Basistha,
Amrita Dhillon,
Danila Serra

WARWICK
THE UNIVERSITY OF WARWICK

UKRI Economic and Social Research Council

Can Crises Affect Citizen Activism? Evidence from a Pandemic*

Farzana Afridi† Ahana Basistha‡ Amrita Dhillon§ Danila Serra¶

November 18, 2023

Abstract

We consider the largely unexpected shock caused by the second wave of the COVID-19 pandemic in India to assess whether major crises that impact the well-being of a large number of individuals can be catalysts for civic activism. Exploiting state-level variation in COVID-19 peaks and quasi-randomness in subjects' participation in an online survey fielded between March and July 2021, we elicit willingness to act against fraud and corruption in the provision of health services by supporting an NGO engaged in advocacy for health-sector regulation. By comparing responses of subjects surveyed before and after the COVID-19 peak in their state of residence, we find evidence of a large and significant increase in anti-corruption activism post peak. Our data suggest that this surge in activism can be attributed to heightened perceptions of corruption in the healthcare sector, increased awareness of individuals' own rights and entitlements, a greater willingness to take risks, and a positive shift in beliefs regarding others' willingness to fight corruption in the provision of healthcare services.

JEL codes: D73, D83, I15
Keywords: crisis, corruption, health sector, India, COVID-19

*We acknowledge financial support from the UK government, through DFID Accountability Initiative.
†Indian Statistical Institute (Delhi) and IZA (Bonn). Email: fafridi@isid.ac.in
‡Indian Statistical Institute (Delhi). Email: ahanabasistha@gmail.com.
§King's College London. Email: amrita.dhillon@kcl.ac.uk.
¶Texas A&M University. Email: dserra@tamu.edu.

1 Introduction

Direct lived experiences are often powerful motivators for belief formation and actions (Malmendier, 2021). By impacting large segments of the population, major crises caused by economic shocks, natural disasters, violent conflicts or health epidemics, can induce persisting changes in individual-level and society-level beliefs, expectations and behaviors.[1] The COVID-19 pandemic stands as an unparalleled global crisis, due to its profound impact on virtually every aspect of life worldwide. Similar to other catastrophic events, the pandemic also facilitated corruption in settings characterized by weak institutions, lack of transparency in the disbursement of emergency health funds, and little oversight over the use of such funds (Gallego et al., 2020; Rose-Ackerman, 2021; Vrushi and Kukutschka, 2021). Recent work has shown that the pandemic had lasting impacts on work preferences (Chen et al., 2023), political preferences (Baccini et al., 2021), confidence in the press and in governmental organizations (Brodeur et al., 2021).[2]

The second wave of the pandemic occurred in India largely unexpectedly in April 2021, and the number of daily cases peaked nationally on May 8, 2021. The number of people who died as a result is unprecedented in history (Jha et al., 2022). The grief and pain of losing a loved one touched nearly every family, regardless of socioeconomic status.[3] In this paper, we investigate whether the second wave of the COVID-19 pandemic, by amplifying difficulties in accessing medical care and by forcing large segments of the population to directly confront corruption within the healthcare system, increased individuals' willingness to engage in anti-corruption activism.

We exploit state-level variation in the timing of the second COVID-19 wave, between April and July 2021, and data from an online survey that we administered throughout India between February and July 2021. The survey aimed to capture experiences with the healthcare system, and individual willingness to engage in activism to curb fraud and corruption in hospitals. We implemented the study through Qualtrics, which manages large survey panels in India, in partnership with an NGO. Importantly, the data collection was

[1]Studies in psychology and neuroscience suggest that "emotional events often attain a privileged status in memory" (LaBar and Cabeza, 2006) and prior experiences influence how individuals learn (Sharpe et al., 2021; Spunt and Adolphs, 2017; Isen et al., 1978).

[2]For a review of the literature on the socioeconomic consequences of the COVID-19 pandemic see Brodeur et al. (2021).

[3]According to the World Health Organization dashboard, India had a cumulative 44,997,326 total cases as of 11 September 2023, second only to the United States, and the third highest death count. In particular, the second wave (April-July 2021) in India witnessed the highest surge in the world. India became the first country to report over 400,000 new cases in a single day on April 30, 2021 (`https://www.thehindu.com/news/national/coronavirus-india-becomes-first-country-in-the-world-to-report-over-400000-new-cases-on-april-30-2021/article61817889.ece`).

opened and closed multiple times by the Qualtrics team over the implementation period, to meet pre-determined quota and data quality checks. This generated quasi-random variation in the timing of individual participation in the survey, which allows us to compare, within a state, the preferences and decisions of individuals who participated in the survey before or after the COVID-19 peak.[4]

In the survey, not only did we elicit (unincentivized) willingness to "participate in a protest against corruption in the delivery of health services," but we also gave subjects the chance to take real action against fraud and corruption in the health sector. Specifically, at the end of the survey, we asked subjects whether they would be willing to sign a petition to the Ministry of Health to demand more accountability in the health sector, or make a donation to the NGO working to increase accountability and reduce corruption, or watch a 5 minute informational video on ways to take action against fraud in the health sector. We are therefore able to examine the impact of the health crisis caused by the second wave of the pandemic on both stated preferences and actual anti-corruption activism.

Data from our sample of nearly 900 Indian men, of which about 35 percent surveyed before the COVID-19 peak in their state of residence, and 65 percent surveyed after the peak, show evidence of a substantial and statistically significant increase in subjects' willingness to take action against healthcare corruption after being exposed to the second wave of the pandemic. We document an 11 percent increase in willingness to protest against healthcare corruption, a 30 percent increase in stated willingness to take action by signing a petition, making a donation to an NGO or watching an how-to-act video, and a 33 percent increase in actual engagement in one of these forms of activism.

Our analysis of mechanisms suggests that the increase in anti-corruption activism following the COVID-19 crisis likely stems from: (1) an increase in perceived corruption in the health sector, (2) greater awareness of own rights and entitlements regarding health provision, (3) increased risk tolerance and (4) a positive shift in individual beliefs about others' willingness to take action. We do not find any evidence of higher pro-sociality or lower tolerance of corruption resulting from the pandemic.

Our paper adds to the literature on how personal experiences of society-wide negative shocks, such as economic depressions and natural disasters, can affect individual beliefs, preferences and behaviors, both at the micro- and the macro-level. For instance, Malmendier and Nagel (2011) show that individuals' experiences of macro-economic shocks have long-term effects on risk attitudes and investment decisions, and Belloc et al. (2016) document that in the Middle Ages experiencing earthquakes facilitated the transition from autocratic to self-

[4]Our study design is sometimes referred to as an "unexpected event during study design" (e.g., see Muñoz et al., 2020, for a review).

governed institutions in Italy. A large literature on the microeconomics of violent conflicts provide evidence on the impact of war victimization on a large set of individual attitudes and preferences, including civic engagement and political participation (Bauer et al., 2016; Bellows and Miguel, 2009; Blattman, 2009). We add to these studies by examining the immediate impact of a negative shock, which heightened individual experiences of corruption in the health sector, on anti-corruption activism.

We also contribute to the literature on the impact of pandemics on individuals' preferences and attitudes. While there exists a large body of work on the effects of major health shocks on economic activity, health outcomes, and human capital development,[5] investigations of impacts on individual preferences and attitudes are scarce. Klemm and Mauro (2022) find that, in the US, the COVID-19 pandemic increased support for progressive tax reform, with the results being driven by individuals who experienced serious illness or job losses. Eichengreen et al. (2021) show that exposure to an epidemic over an individual's formative years (ages 18 to 25) reduces confidence in scientists, leading in lower vaccine take-up. Saka et al. (2022) find that past epidemic exposure during the formative years tends to lower individual confidence in both political institutions and the healthcare system.

The primary challenge in these studies lies in the identification of causal effects, due to the widespread and simultaneous diffusion of a pandemic across a population, and the non-random implementation of response strategies by national and local governments. We overcome these challenges by employing survey data collected as the second wave of the pandemic unexpectedly unfolded in India, and by leveraging time and space variation in exposure to the pandemic, together with quasi-randomness in the timing of survey participation. This allows us to identify how the health crisis, by amplifying individuals' personal encounters with corruption within the healthcare system, altered their preferences for anti-corruption activism and motivated subjects to participate in such efforts, thus becoming a catalyst for collective action.

2 Background and Context

2.1 India's Health Sector

The Indian healthcare system suffers from chronic under investment in key infrastructure and a high level of out-of-pocket expenditure, especially in poorer states (Garg and Karan, 2009; Das et al., 2016; Banerjee et al., 2008).[6] In fact, the healthcare system is characterized

[5]See Beach et al. (2022) for a recent review of this literature.

[6]According to the World Bank, India spent only USD 63.75 on health care per capita, versus a world average of USD 1,121.81 in 2019. Similarly, the share of out-of-pocket expenditure was roughly 55% of

by substantial state-level variation in healthcare spending as a proportion of budget, capacity of hospital beds (especially in Intensive Care Unit), and availability of health personnel and testing centers (Choutagunta et al., 2021). Moreover, the private sector has recently become a dominant player in healthcare provision, with minimum government regulation. This has led to nearly 70 percent of all health expenses occurring in the private sector, a substantially larger number than the average of 46 percent observed in low and middle countries.[7]

Capacity and budgetary constraints, insufficient regulation and lack of oversight all contribute to the proliferation of fraud and corruption in healthcare provision (Kumar, 2003).[8] Although corruption has been an enduring issue within the Indian healthcare sector, the COVID-19 pandemic significantly exacerbated its scope and impact. The absence of a regulatory framework posed a significant obstacle for states seeking to enforce accountability measures against corrupt individuals within healthcare facilities.[9] Corruption took place in the form of overcharging for COVID-related services, favoritism in service provision, and even the administration of fake vaccines for a fee.[10]

2.2 Pandemic Timeline

The first wave of the COVID-19 pandemic was marked by a strict nationwide lockdown with restrictions on domestic and international travel, stretching from 25 March 2020 to 31 May 2020, with the caseload peaking nationally around September 2020. In stark contrast, the second wave of the pandemic, which reached India during spring 2021, was marked by general unpreparedness and lack of coordinated efforts to keep the upsurge of COVID-19 cases under control. Moreover, in contrast with the first wave, the management of disease control and vaccinations was mostly shifted from the federal to the state governments (Press Information Bureau, 2021).

The 7 day moving average of the confirmed daily case-load surged from 65,144 on April 1 to about 392,331 by May 8, 2021, at its peak (World Health Organisation, 2022). According

current health expenditure in India, whereas the world average was about 18%.

[7]For more details, see data from https://data.worldbank.org/indicator/SH.XPD.PVTD.CH.ZS?end=2019&start=2019&view=map.

[8]See also the Transparency International Health Sector Corruption report here: https://www.transparency.org/en/our-priorities/health-and-corruptionTransparency International - Health Sector Corruption.

[9]For instance, the Clinical Establishments Act (2010), which provides for the registration and regulation of clinical establishments and prescribes minimum standards of facilities, had not been adopted at the time of the pandemic. For more press coverage of the limited implementation of this act, see: https://theprint.in/talk-point/fortis-regulate-charges-corporate-hospitals/17730/ and https://www.tribuneindia.com/news/punjab/7-yrs-on-clinical-act-hangs-fire-1100.

[10]For news coverage, see https://www.indiatoday.in/india/story/corruption-second-covid-pandemic-black-marketing-medicines-tiii-1799395-2021-05-06.

to news reports, in the absence of a nationwide lockdown, most states issued a complete or partial lockdown (Indian Express, 2021). Despite such localized lockdowns, each state showed similar increases in daily confirmed cases. Nevertheless, the different states reached their peak at different times, between April and July 2021. Specifically, ten states peaked before the national peak (May 8), nineteen states peaked on or after the national peak, but still in the month of May, one state peaked in June and one in July.[11]

3 Data and Methodololgy

3.1 Data

We conducted an online survey of adult Indian men implemented by Qualtrics between March and July 2021.[12] The study was framed as a general survey focused on understanding people's behavior and attitude during the pandemic, rather than health sector 'corruption' per se, in order to avoid priming the subjects.[13] The survey opened and closed on random dates between March and July 2021.[14] To identify the unanticipated COVID-19 wave and subsequent peaks in national and state-level daily caseloads, we utilize publicly available administrative data.[15] In Figure 1, we show the distribution of subjects according to the date of case peak experienced in their respective states of residence (in grey) and the timing of the survey (in red).[16] Depending on each respondent's random date of survey, we create a dummy variable that indicates whether the subject participated in the survey on or before the peak-date of daily COVID cases in his state of residence.[17]

[11] In the Online Appendix, we report the specific COVID-19 7-day moving average peak dates in each Indian state in Table A1.

[12] We recruited only men to avoid gender disparities arising from access to computer/mobile devices, healthcare and intra-household decision-making. Participation was conditional on monthly household income of INR 60,000 or less. Therefore, our average survey respondent is younger, more educated and belongs to wealthier households than the average Indian urban man.

[13] The data used in this paper is part of a larger study, which included the random allocation of subjects to different treatment conditions (Afridi et al., 2023). We employ only data from our baseline condition in this paper, and exploit unexpected variation in the timing of the second wave of COVID-19 and of the data collection. We pre-registered the larger experimental study on AsPredicted in March 2021.

[14] Qualtrics maintains a large pool of subjects who are representative of relatively younger, economically better-off and urban Indian population. Qualtrics sends a link to the potential respondents and waits for the sample to reach the target quota. Subjects in this survey did not get to see when the survey link would be on or for how long it would remain active. Typically, it took between 1 and 4 days to collect the target data. For detailed information on the dates the survey was open and closed by the Qualtrics team, see Table A2 in the Online Appendix.

[15] See: `https://www.covid19india.org/`.

[16] Throughout the paper, "peak" refers to the peak in daily new confirmed cases of COVID. Our main results are robust to using the peak in daily deaths instead.

[17] The national peak in daily cases was on May 8, 2021. The state-wise peaks in daily cases are provided in Table A1 in the Online Appendix.

Following a battery of initial questions on demographics and experience with healthcare, we elicited subjects' (hypothetical) agreement/disagreement with a statement related to their personal willingness to take part in a protest towards malfeasance/irregularities in the health sector during the pandemic. This constitutes our (hypothetical) outcome measure of willingness to protest.[18] Once subjects reached the end of the survey, we thanked them for their participation (following standard practice in online surveys) and then we invited them to "think about the problem of corruption and overcharging in Indian hospitals during the COVID-19 pandemic." We included a brief description of a nation-wide NGO that we partnered with - the All India Drug Action Network (A.I.D.A.N). A.I.D.A.N is a collective of medical and legal professionals that has been pressurizing local and federal government to better regulate health care in India, fostering transparency in pricing and providing redressal to patients who have been illegally overcharged. We then asked subjects whether they wanted to support A.I.D.A.N.'s activities or exit the survey. Subjects were randomly assigned to 4 different action treatment groups. Specifically, they could either sign a petition for improving regulation of the health sector, or make a monetary donation to the non-profit, or watch an informational video on A.I.D.A.N. activities and ways to get involved in the fight against corruption in the health sector.[19] The fourth action treatment presented subjects with all three actions and allowed them to choose among them or exit the survey.

We capture respondent's willingness to take action by pooling subjects' responses from all four action treatment groups in our analysis. This leads to two measure of anti-corruption activism: individual willingness to take action by selecting the activism option at the end of the survey rather than exiting, and a measure of actual activism, i.e., the decision to sign the petition to the Ministry of Health with a full name, or donating part of the survey earnings to the NGO, or watching the full 5 minute how-to-act video. As a result, our analysis is based on three outcome variables: (1) the hypothetical willingness to protest against corruption in the provision of health service, elicited through a survey question; (2) stated willingness to take action, as measured by the decision to not exit the survey and instead receive more information on one of the forms of activism presented to subjects, and (3) actual activism, measured by the decision to engage in one of the three forms of activism – signing the petition with a full name, donating a positive amount or watching the full 5-minutes of the how-to-act-video.[20]

[18] Willingness to engage in anti-corruption activism may be a sensitive issue for some individuals. Sensitivity bias in survey responses is attenuated in online surveys, as study participants remain anonymous to researchers.

[19] We developed each of these actions in partnership with A.I.D.A.N.

[20] For examples of online survey-based studies using a petition or a donation to measure individual preferences see for instance Alesina et al. (2018); Settele (2019); Bursztyn et al. (2020).

We consider the latter two outcomes (willingness to act and actual activism) as a *real-effort* behavioral measures since subjects had to incur a cost when they chose to engage in these actions instead of exiting the survey.[21] We therefore expect them to be less likely affected by social desirability bias and more likely to reflect individual true preferences regarding anti-corruption activism. While our "willingness to protest" measure is hypothetical and self-reported, it is still contextually important because, due to our study methodology (online survey), we could not include protesting in our real-effort elicitation of activism.

3.2 Empirical Methodology

3.2.1 Estimation

Our main estimating equation is the following:

$$Y_{ist} = \gamma + \beta_0 \; Post_{ist} + \beta_1 X_{ist} + \alpha_s + \varepsilon_{ist} \tag{1}$$

where $\mathrm{Y}_{\mathrm{ist}}$ is one of the three outcome variables defined above - willingness to protest, willingness to act, and the decision to engage in actual activism. $\mathrm{Post}_{\mathrm{ist}}$ is a dummy variable equal to 1 if the subject i's survey date t was after the 7-day moving average peak in daily COVID cases in his state of residence s, and 0 otherwise.[22] Hence, β_0 is our main coefficient of interest, capturing the effect of the second wave of pandemic on outcomes. $\mathrm{X}_{\mathrm{ist}}$ is a vector of individual characteristics, such as age, education, marital status, religion (1- Hindu; 0-otherwise), income (1 if household income is below INR 30K in previous month; 0 otherwise), assets (count of assets owned by a subject from a list of 10 common household assets), household composition, (i.e., having children, living with parents and living with an elderly person), mode of participation by phone and frequency of participation in online surveys. We also include state fixed effects α_s and an idiosyncratic error term $\varepsilon_{\mathrm{ist}}$. Throughout, we define the peak in terms of the 7-day moving average.[23] Standard errors are clustered at the state-month level to account for unobserved heterogeneity over time and space.

We supplement our analysis by estimating equation 1 at successive bandwidths around the timing of the shock, capturing the impact dissemination over time during the study period. Specifically, we estimate whether the impact of the crisis differs with the distance –

[21] By not exiting the survey subjects incurred the opportunity cost of time. By signing the petition they further incurred the cost of being identified by the Ministry of Health, and possibly receiving individual punishment. By making a donation, subjects incurred a pecuniary cost. By watching the 5-minute video, they incurred the opportunity cost of longer time spent on the survey.

[22] Bol et al. (2021) have also adopted a similar estimation strategy while measuring trust in government during COVID-19. For a recent review of studies using similar strategy, see Muñoz et al. (2020).

[23] Note that in the absence of centralized lock-downs, we can consider the stringency of localized lock-downs as a factor that varied at the state-level, and is absorbed by the state fixed effects.

measured in days – between the time of the survey and and the time of the state-level peak. We therefore define a model with a running variable, *Days*, ranging from -112 to 93, with 0 corresponding to the day of the peak, and 1 indicating the first day after the peak.[24]

We estimate the following equation:

$$Y_{ist} = \gamma + \beta_0 \; Post_{ist} + \beta_1 \; D_{ist} + \beta_2 \; D_{ist} \;\times\; Post_{ist} + \beta_1 X_{ist} + \alpha_s + \varepsilon_{ist} \tag{2}$$

where $\mathrm{D_{ist}}$ (i.e., the *Days* variable) denotes the difference in days between the interview date t and the peak date for subject i residing in state s. In this model, $Post_{ist}$ indicates the immediate impact of exposure to the peak, whereas the interaction term $D_{ist} \times Post_{ist}$ tests whether the impact became weaker or stronger over time.

Our secondary analysis aims to identify the primary mechanisms behind our main results. To this end, we employ equation 1 to examine the impact of the crisis on: i) beliefs about others' willingness to take action; ii) perceptions of corruption in the health sector; iii) awareness of own rights and entitlement in the health sector; iv) risk tolerance; v) pro-sociality and vi) tolerance of corruption. Given the number of outcome variables, we correct the p-values associated to individual hypotheses by employing the step-down multiple testing method developed by Romano and Wolf (2005).

Finally, we examine whether the impact of the crisis on activism is long-lasting. To do so, we augment our sample with additional survey data collected following the same procedure between August and November 2022, i.e., more than one year after our primary data collection. However, from this new wave of data we only have the self-reported willingness to protest as our measure of activism.

3.2.2 Identification Strategy and Robustness Checks

Since the experience of the pandemic or its timing is unlikely to be random, we utilize the quasi-random staggering of the time of the survey data collection within a state of residence, jointly with state-level variation in the timing of the pandemic, to assess the impact of the crisis on anti-corruption activism. Note that given our data, we cannot measure the change in willingness to act for a *given* individual; rather, we compare different individuals who were surveyed pre- and post-peak within the same state.

For causal estimation, we require that conditional on observables, assignment to pre or post periods is independent of the outcomes. We assume that this is the case given the quasi-randomness of the timing of the survey implementation by the Qualtrics team.

[24] For example, if the peak in subject's state of residence was on 9 May 2021 and he was interviewed on 10 May, then the Days variable would be equal to 1.

Specifically, subjects could not select into either pre or post group, because the survey was opened and closed for several brief time spans, which was unpredictable from the respondents' viewpoint.[25] Therefore, we consider the assignments of individuals to a pre or a post group as good as random.[26] This does not necessarily mitigate concerns about unobserved differences and therefore needs further elaboration. In Table 1, we test and show that the pre and the post samples are similar in terms of observable characteristics, with only 2 out of 12 variables (marital status and being a Reserved Caste) being significantly different across groups. Additionally, as a robustness check, we repeat the main analysis after implementing entropy balancing (Hainmueller, 2012). This is a weighting process used to create balanced samples in observational studies with a binary treatment where the control group data can be re-weighted to match the co-variate moments in the treatment group. Column 4 of Table 1 confirms that after re-weighting, we see balance in all individual characteristics.[27]

We also check whether the quality of the data generated by the pre- and post-peak samples is consistent. Specifically, we examine response speed and subject comprehension. The former is the number of Qualtrics surveys completed per minute during the time the survey was open. The latter is the rate at which participants failed attention checks. In the Online Appendix, we show (Figure A3) that these two measures are not statistically different in the pre and post groups. As a further test of whether other simultaneous events may confound our results, for instance by affecting the selection of participants in the survey, we check whether the state-level unemployment rate differs pre- and post- peak. To this end, we use the nation-wide CMIE-CPHS data (CMIE, 2023) for 227,200 adult males residing in Indian states between March - July 2021. We rerun equation 1 with individual-level unemployment as the outcome variable. We find no evidence of significant changes in unemployment rates following the state-level COVID-19 peaks.[28] Finally, we provide several robustness checks in Section 4.1 to further assure that our estimation strategy is appropriate.

We note that our identification strategy does not assume that subjects interviewed before the peak did not expect the peak to occur. While this may be true in the states that were affected by the second COVID-19 wave early on, once the rapid increase of COVID-19 cases was covered by news outlets, it is possible that subjects surveyed before the peak in their

[25]From the back-end, we also ensured that a subject could not take the same survey more than once.

[26]Given that our survey period is fairly short, i.e., ranging from March 24 to July 26th, 2021, it is unlikely that other major systematic changes occurred in that time frame, besides the COVID-19 crisis.

[27]Propensity score methods are often used in observational studies to pre-process the data prior to the estimation of treatment effects under the assumption of selection on observables. In contrast to most other pre-processing methods, entropy balancing involves a re-weighting scheme that directly incorporates co-variate balance into the weight function that is applied to the sample units. An advantage of this method is that unlike the traditional coarsened exact matching, entropy balancing does not require huge data sets, and does not cause large portions of the sample to drop.

[28]See Table A4 in the Online Appendix.

state of residence expected the peak to occur soon. In our analysis of heterogeneous effects, we check whether the impact of the crisis on activism is larger for subjects in early-peak than those in late-peak states.

4 Results

We report the results from equation 1 in Panel A of Table 2. For each of our measures of activism (willingness to protest, willingness to take action, and decision to take action), we report estimates without (columns 1, 3, and 5) and with (columns 2, 4 and 6) entropy balancing. The estimates show that the willingness to protest against corruption increases by 7.5 percentage points after exposure to the COVID-19 peak. This translates into a 9 percent increase over the pre-peak mean of 83 percent. Similarly, the willingness to take action at the end of the survey increases by 7.5 percentage points after the pandemic peak, which corresponds to a 29 percentage increase over the pre-peak mean of 37 percent. Finally, columns 5 and 6 show that the likelihood to take an action increases by 8 percentage points, i.e., 35 percent over the pre-peak mean of 23 percent. The estimates are robust to entropy balancing, as shown in columns 2, 4 and 6, and to the Romano-Wolf correction for multiple hypothesis testing.[29]

In Panel B of Table 2, we estimate equation 2. For the hypothetical measure of willingness to protest neither the estimate of immediate effect nor exposure over time, is statistically significant (columns 1 and 2). However, we find that the immediate impact of exposure to peak on the real-effort willingness and decision to act is substantial. Columns 3 to 6 show that the immediate impact of the peak was an increase in willingness to act and in the likelihood to take action by 25 and 17 percentage points, respectively, although the latter impact loses statistical significance at conventional levels when implementing entropy balancing. The estimated coefficient of the interaction term *Days x Post* is statistically significant and negative, suggesting that the impact of the peak on activism decreases by between 0.04 and 0.07 percentage points for each day post peak. The positive coefficient of the non-interacted *Days* variable indicate that as the peak of COVID-19 became nearer, preferences for activism gradually increased.

[29]To graphically illustrate our results, we disaggregate the pre and post time periods as distance (in months) from the month containing the state peak. The time paths for our three measures of activism, shown in Figure A1 in the Online Appendix, confirm that our three measures of activism increased steadily in the 3 months following the state-level peak.

4.1 Robustness

As a robustness check, we vary the bandwidth around the state-level peak by ± 1 day(s). Recall that our overall time window is 125 day, i.e., from March 24 (when the data collection began) to July 26, 2021, when it ended. We start with $\pm$ 18 days around the peak, in order to ensure that we have enough observations to ensure balance in covariates between pre and post groups at every bandwidth.[30] The results, presented in Figure 2, show that the estimated impact of exposure to peak is consistently positive and stable over time, for all our outcomes. The observed stability also indicates that time trends or other events during our time frame of analysis are unlikely to be driving the estimated effects.

Our primary analysis employs state-month fixed effects. However, since the timing of the COVID-19 peak varied at the state level, as a robustness check, we estimate equation 1 by changing the level of clustering from state-month to state level. We use the wild cluster-bootstrap method (Roodman et al., 2019) to correct for the small number of clusters (35 states). Our estimates are robust to clustering the standard errors at the state level.[31]

4.2 Mechanisms

To examine what led to the increase in citizen activism following the health crisis, we conduct a secondary analysis where we assess the impact of state-level COVID-19 peaks on individual beliefs related to corruption in the health sector and others' willingness to act, as well as individual preferences such as risk tolerance and pro-sociality.[32]

The estimates reported in Table 3 show that the health crisis positively affected subjects' beliefs about others' willingness to take action against corruption in the health sector (column 1). We also see that the crisis increased both information about own rights concerning healthcare, and perception of corruption (columns 2 and 3). Our analysis of impacts on preferences shows that individuals' risk tolerance increased (column 4),[33] whereas their pro-sociality and tolerance of corruption were unaffected (columns 5 and 6). Our estimates are robust to correcting p-values for multiple hypothesis testing using the Romano and Wolf (2016)'s procedure.[34]

[30] Since we have small sample sizes for narrow bandwidths, we cannot adjust the standard errors by clustering at the state-month level.

[31] The results are reported in Table A5 in the Online Appendix.

[32] Details on how these survey-generated variables are created are provided in the Online Appendix.

[33] This is consistent with the body of literature providing evidence of an increase in risk tolerance following natural disasters. See for instance, the study by Islam et al. (2020) in Bangladesh. Regarding the COVID-19 pandemic specifically, Tsutsui and Tsutsui-Kimura (2022) shows that in Japan, people became more tolerant of risk after being exposed to COVID-19.

[34] We standardized our indexes, used in columns 2 to 6 of Table 3, around the pre-peak mean. Consequently, our estimates in columns 2 to 6 are expressed in standard deviations from the pre-peak mean.

4.3 Heterogeneous Impacts

To complement our main findings, we assess whether there are differential impacts of the health crisis on activism based on whether a state experienced the COVID-19 peak relatively early on or later in the second wave (relative to national peak in the full sample). It could be argued that the early peak states were more likely caught off-guard by the second wave of the pandemic, leading to greater casualties.[35] To graphically illustrate our findings, we plot the month-by-month time paths of our activism outcomes for the early and late peak sub-samples. Overall, in line with our expectations, we see stronger evidence of a positive impact of the crisis on activism in the early-peak than the late-peak states.[36]

We also ask if the estimated impacts differ by pre-existing state-level characteristics, such as corruption and health sector quality. To capture ex-ante levels of corruption at the state level, we use the Transparency International India (2019) report, which covers 20 Indian states. We create a dummy variable which is equal to 1 for *high corruption states* and 0 otherwise.[37] We also utilize data on hospitals from Kapoor et al. (2020), and data on state-level population for 2020 from Government of India (2019) to compute *hospital density*, i.e., the number of (public *and* private) hospitals per 100,000 people residing in a state. We find that the impact of the crisis on willingness to take action tends to be higher in states with high corruption and high hospital density, although actual activism is not significantly different across these dimensions.[38]

4.4 Longer-term Impacts

Does the increased willingness to take action against corruption persist over time? Between October and November 2022, about 14 months after the primary data collection, we conducted another online survey, following the same procedures as in March-July 2021. The one difference is that we only elicited the hypothetical measure of activism.

We refer to our additional sample of 849 survey participants as our *long-term sample* and the combination of the two samples as our *full sample*. We replicate the analysis conducted in Panel A of Table 2 for the long-term and the full samples, while also including 'Date of

[35] Table A1 in the online Appendix provides the full list of dates when state level peaks occurred.

[36] See Figure A2 in the Online Appendix.

[37] The high corruption states, as indicated in the report, are Punjab, Rajasthan, Uttar Pradesh, Bihar, Jharkhand, Karnataka, Telangana and Tamil Nadu. The states not included in the report are excluded from this analysis, reducing the number of observations to 848. For more details about the report, visit the website of Transparency International India.

[38] The estimates are reported in Table A6 in the Online Appendix. Note that the *high corruption* state indicator and the *hospital density* variable are captured at the state level and are time-invariant, hence these are perfectly collinear with state fixed effects, which we consequently drop from the specification.

Survey' fixed effects.[39] The estimates suggest that the positive impact of the health crisis on activism persisted over time. Specifically, in the long-term sample we still observe a 9 percentage points impact of the health crisis on the self-stated likelihood to join an anti-corruption protest.[40]

5 Conclusion

In this paper, we asked whether lived experience of malfeasance in the health sector, heightened in frequency and intensity by the unexpected surge of COVID-19 cases during the second wave of the pandemic in India, led to an increase in individuals' willingness to engage in anti-corruption activism. By exploiting state level variation in the occurrence of COVID-19 peaks, and quasi-randomness in the timing of subjects' participation in our survey, we find evidence of a significant increase in anti-corruption activism as a result of the health crisis and suggestive evidence that these effects persisted over a year later.

Our analysis shows that the crisis positively impacted individuals' information about their rights and entitlements, their risk tolerance, their perceptions of corruption in the health sector, and their beliefs about others' willingness to take action. This suggests that the intense negative experiences caused by the pandemic re-shaped and strengthened both individual and collective motivations to engage in anti-corruption activism.

[39]The longer time frame of analysis allows us to do that.

[40]The estimates are reported in Table A7 in the Online Appendix.

References

Afridi, F., A. Basistha, A. Dhillon, and D. Serra (2023). Activating change: The role of information and beliefs in social activism. IZA Discussion Papers 16358, Institute of Labor Economics (IZA).

Alesina, A., A. Miano, and S. Stantcheva (2018). Immigration and Redistribution. Technical report, National Bureau of Economic Research.

Baccini, L., A. Brodeur, and S. Weymouth (2021). The COVID-19 Pandemic and the 2020 US Presidential Election. Journal of Population Economics 34, 739–767.

Banerjee, A. V., E. Duflo, and R. Glennerster (2008). Putting a Band-Aid on a Corpse: Incentives for Nurses in the Indian Public Health Care System. Journal of the European Economic Association 6(2-3), 487–500.

Bauer, M., C. Blattman, J. Chytilová, J. Henrich, E. Miguel, and T. Mitts (2016). Can War Foster Cooperation? Journal of Economic Perspectives 30(3), 249–274.

Beach, B., K. Clay, and M. Saavedra (2022). The 1918 Influenza Pandemic and Its Lessons for COVID-19. Journal of Economic Literature 60(1), 41–84.

Belloc, M., F. Drago, and R. Galbiati (2016). Earthquakes, Religion, and Transition to Self-Government in Italian Cities. The Quarterly Journal of Economics 131(4), 1875–1926.

Bellows, J. and E. Miguel (2009). War and Local Collective Action in Sierra Leone. Journal of Public Economics 93(11-12), 1144–1157.

Blattman, C. (2009). From Violence to Voting: War and Political Participation in Uganda. American Political Science Review 103(2), 231–247.

Bol, D., M. Giani, A. Blais, and P. J. Loewen (2021). The Effect of COVID-19 Lockdowns on Political Support: Some Good News for Democracy? European Journal of Political Research 60(2), 497–505.

Brodeur, A., D. Gray, A. Islam, and S. Bhuiyan (2021). A Literature Review of the Economics of COVID-19. Journal of Economic Surveys 35(4), 1007–1044.

Brodeur, A., I. Grigoryeva, and L. Kattan (2021). Stay-at-Home Orders, Social Distancing, and Trust. Journal of Population Economics 34(4), 1321–1354.

Bursztyn, L., G. Egorov, and S. Fiorin (2020). From Extreme to Mainstream: The Erosion of Social Norms. American Economic Review 110(11), 3522–3548.

Chen, Y., P. Cortes, G. Kosar, J. Pan, and B. Zafar (2023). The Impact of COVID-19 on Workers' Expectations and Preferences for Remote Work. In AEA Papers and Proceedings, Volume 113, pp. 556–561. American Economic Association 2014 Broadway, Suite 305, Nashville, TN 37203.

Choutagunta, A., G. Manish, and S. Rajagopalan (2021). Battling COVID-19 with Dysfunctional Federalism: Lessons from India. Southern Economic Journal 87(4), 1267–1299.

CMIE (2023). Unemployment Rate in India. `https://unemploymentinindia.cmie.com/kommon/bin/sr.php?kall=wsttimeseries&index_code=050050000000&dtype=total`. Online; accessed 14 April 2023.

Das, J., A. Holla, A. Mohpal, and K. Muralidharan (2016). Quality and Accountability in Health Care Delivery: Audit-Study Evidence from Primary Care in India. American Economic Review 106(12), 3765–99.

Eichengreen, B., C. G. Aksoy, and O. Saka (2021). Revenge of the Experts: Will COVID-19 Renew or Diminish Public Trust in Science? Journal of Public Economics 193, 104343.

Falk, A., A. Becker, T. Dohmen, B. Enke, D. Huffman, and U. Sunde (2018). Global Evidence on Economic Preferences. The Quarterly Journal of Economics 133(4), 1645–1692.

Gallego, J. A., M. Prem, and J. F. Vargas (2020). Corruption in the times of Pandemia. Available at SSRN 3600572.

Garg, C. C. and A. K. Karan (2009). Reducing Out-of-pocket Expenditures to Reduce Poverty: A Disaggregated Analysis at Rural-urban and State Level in India. Health Policy and Planning 24(2), 116–128.

Government of India (2019). Rural Health Statistics. `https://hmis.nhp.gov.in/downloadfile?filepath=publications/Rural-Health-Statistics/RHS%202019-20.pdf`. Online; accessed 12 October 2022.

Hainmueller, J. (2012). Entropy Balancing for Causal Effects: A Multivariate Reweighting Method to Produce Balanced Samples in Observational Studies. Political Analysis 20(1), 25–46.

Indian Express (2021). Indian Express. `https://indianexpress.com/article/india/covid-19-second-wave-heres-a-list-of-states-that-have-imposed-lockdowns-7306634/`. Online; accessed 18 April 2022.

Isen, A. M., T. E. Shalker, M. Clark, and L. Karp (1978). Affect, Accessibility of Material in Memory, and Behavior: A Cognitive Loop? Journal of Personality and Social Psychology 36(1), 1.

Islam, A., C. M. Leister, M. Mahmud, and P. A. Raschky (2020). Natural Disaster and Risk-Sharing Behavior: Evidence from Rural Bangladesh. Journal of Risk and Uncertainty 61(1), 67–99.

Jha, P., Y. Deshmukh, C. Tumbe, W. Suraweera, A. Bhowmick, S. Sharma, P. Novosad, S. H. Fu, L. Newcombe, H. Gelband, et al. (2022). COVID Mortality in India: National Survey Data and Health Facility Deaths. Science, eabm5154.

Kapoor, G., S. Hauck, A. Sriram, J. Joshi, E. Schueller, I. Frost, R. Balasubramanian, R. Laxminarayan, and A. Nandi (2020). State-Wise Estimates of Current Hospital Beds, Intensive Care Unit (ICU) Beds and Ventilators in India: Are We Prepared for a Surge in COVID-19 Hospitalizations? MedRxiv.

Klemm, A. and P. Mauro (2022). Pandemic and Progressivity. International Tax and Public Finance 29(2), 505–535.

Kumar, S. (2003). Health Care Is among the Most Corrupt Services in India. BMJ: British Medical Journal 326(7379), 10.

LaBar, K. S. and R. Cabeza (2006). Cognitive Neuroscience of Emotional Memory. Nature Reviews Neuroscience 7(1), 54–64.

Malmendier, U. (2021). FBBVA Lecture 2020: Exposure, Experience, and Expertise: Why Personal Histories Matter in Economics. Journal of the European Economic Association 19(6), 2857–2894.

Malmendier, U. and S. Nagel (2011). Depression Babies: Do Macroeconomic Experiences Affect Risk Taking? The Quarterly Journal of Economics 126(1), 373–416.

Muñoz, J., A. Falcó-Gimeno, and E. Hernández (2020). Unexpected Event during Survey Design: Promise and Pitfalls for Causal Inference. Political Analysis 28(2), 186–206.

Oppenheimer, D. M., T. Meyvis, and N. Davidenko (2009). Instructional Manipulation Checks: Detecting Satisficing to Increase Statistical Power. Journal of Experimental Social Psychology 45(4), 867–872.

Press Information Bureau (2021). PIB. `https://pib.gov.in/PressReleasePage.aspx?PRID=1710190`. Online; accessed 16 April 2022.

Romano, J. P. and M. Wolf (2005). Stepwise Multiple Testing as Formalized Data Snooping. Econometrica 73(4), 1237–1282.

Romano, J. P. and M. Wolf (2016). Efficient Computation of Adjusted p-Values for Resampling-Based Stepdown Multiple Testing. Statistics & Probability Letters 113, 38–40.

Roodman, D., M. Ø. Nielsen, J. G. MacKinnon, and M. D. Webb (2019). Fast and Wild: Bootstrap Inference in Stata Using Boottest. The Stata Journal 19(1), 4–60.

Rose-Ackerman, S. (2021). Corruption and COVID-19. EUNOMÍA. Revista en Cultura de la Legalidad (20), 16–36.

Saka, O., B. Eichengreen, and C. Aksoy (2022). The Political Scar of Epidemics. The Economic Journal.

Settele, S. (2019). How Do Beliefs About the Gender Wage Gap Affect the Demand for Public Policy? Available at SSRN 3382325.

Sharpe, M. J., H. M. Batchelor, L. E. Mueller, M. P. Gardner, and G. Schoenbaum (2021). Past Experience Shapes the Neural Circuits Recruited for Future Learning. Nature Neuroscience 24(3), 391–400.

Spunt, R. P. and R. Adolphs (2017). A New Look at Domain Specificity: Insights from Social Neuroscience. Nature Reviews Neuroscience 18(9), 559–567.

Transparency International India (2019). India Corruption Survey. https://transparencyindia.org/wp-content/uploads/2019/11/India-Corruption-Survey-2019.pdf. Online; accessed 25 January 2022.

Tsutsui, Y. and I. Tsutsui-Kimura (2022). How Does Risk Preference Change under the Stress of COVID-19? Evidence from Japan. Journal of Risk and Uncertainty, 1–22.

Vrushi, J. and R. Kukutschka (2021). Why Fighting Corruption Matters In Times Of COVID-19. In Corruption Perceptions Index 2020: Research Analysis. Transparency International.

World Health Organisation (2022). World Health Organisation. https://www.who.int/news/item/05-05-2022-14.9-million-excess-deaths-were-associated-with-the-covid-19-pandemic-in-2020-and-2021. Online; accessed 19 April 2022.

6 Tables and Figures

Table 1: Descriptive Statistics and Balance Tests

	Pre-2nd Wave (1)	Post-2nd Wave (2)	Difference (3)=(1)-(2)	After Entropy Balancing (4)=(1)-(2)
Age 45+	0.149	0.144	0.005	-0.000
			[0.873]	[0.999]
Married	0.411	0.511	-0.100**	0.000
			[0.024]	[0.996]
Has Children	0.356	0.375	-0.019	0.000
			[0.636]	[0.998]
Lives with Parents	0.272	0.275	-0.003	0.000
			[0.934]	[1.000]
Reserved Caste	0.466	0.628	-0.162***	0.000
			[0.000]	[0.993]
Hindu	0.709	0.772	-0.064	0.000
			[0.219]	[0.994]
College Educated	0.819	0.779	0.039	0.000
			[0.200]	[0.992]
Monthly Income <INR 30K	0.434	0.469	-0.035	0.000
			[0.421]	[0.997]
Asset Index	6.197	5.997	0.201	0.003
			[0.305]	[0.990]
Frequent Survey Participants	0.738	0.781	-0.043	0.000
			[0.120]	[0.991]
Mobile Survey	0.680	0.642	0.038	0.000
			[0.359]	[0.994]
Lives with Elderly	0.592	0.576	0.017	0.000
			[0.712]	[0.996]
N	309	589		

Notes: "Pre-2nd wave" ("Post-2nd wave")indicates that the subject participated in the survey before (after) the COVID-19 peak in his state of residence. "Age 45+" is a dummy equal to 1 for subjects aged 45 and above, 0 otherwise; "Reserved Caste" is a dummy indicating SC (Schedule Caste), ST (Scheduled Tribe) and other back classes (OBC) subjects, who are socio-economically deprived individuals in India, 0 otherwise. "Income < INR 30K" indicates subjects with monthly household income below INR 30K in the previous month; "Asset Index" indicates a count of assets owned by a subject from a list of ten household assets; "Frequent Survey Participant" is equal to 1 for individuals who state to participate in Qualtrics surveys at least once a week, 0 otherwise; "Mobile Survey" is equal to 1 for subjects who participated in the survey using a mobile phone; "Lives with elderly" is equal to 1 if the subject lives with a household member who is older than 60, 0 otherwise; "Has children" is 1 if subject has children, 0 otherwise; "Lives with parents" is 1 if subject lives with parents, 0 otherwise. Column (3) reports the difference between the characteristic of the subjects surveyed pre- and post-state peak. Column (4) reports differences in individual characteristics after adjusting for entropy re-weighting (Hainmueller, 2012). Standard errors are clustered at state-month level. p-values reported in square brackets. * $p < .10$, ** $p < .05$, *** $p < .01$.

Table 2: Impact of COVID-19 State-Level Exposure on Activism

	Willing to Protest		Willing to Act		Took Action	
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A						
Post	0.075***	0.075***	0.107**	0.116***	0.073**	0.080***
	(0.025)	(0.021)	(0.042)	(0.032)	(0.035)	(0.023)
Romano-Wolf p-value		[0.004]		[0.004]		[0.004]
Panel B						
Post	0.017	0.040	0.211***	0.248**	0.185***	0.171**
	(0.058)	(0.056)	(0.075)	(0.107)	(0.067)	(0.076)
Days x Post	0.001	0.000	-0.006***	-0.007**	-0.006***	-0.004**
	(0.001)	(0.001)	(0.002)	(0.003)	(0.002)	(0.002)
Days	0.000	0.000	0.009***	0.008***	0.007***	0.005***
	(0.002)	(0.001)	(0.002)	(0.003)	(0.002)	(0.002)
Observations	898	898	898	898	898	898
Pre-peak Mean Y	0.832	0.832	0.372	0.372	0.233	0.233
State FE	no	yes	no	yes	no	yes
Controls	no	yes	no	yes	no	yes
Entropy Balancing	no	no	no	no	no	no

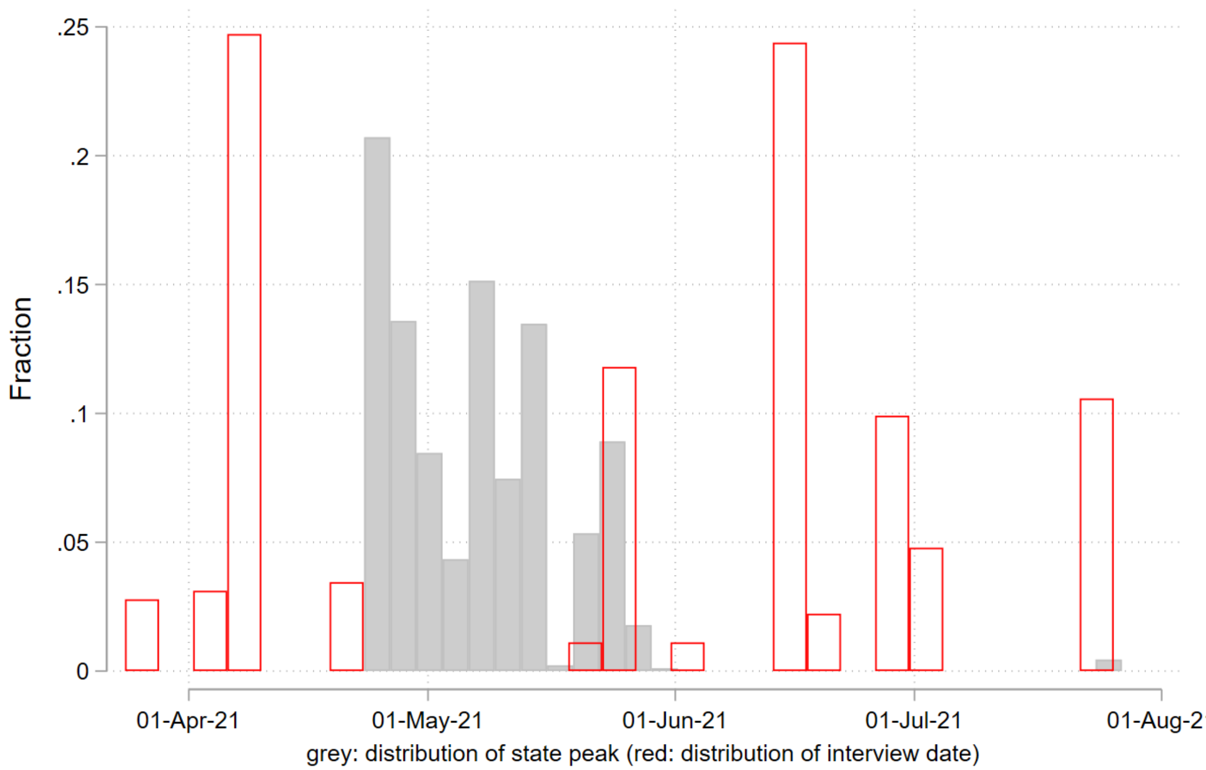
Notes: In both Panels A and B, "Post" is a dummy equal to 1 if the subject was surveyed after the peak in daily COVID-19 cases in his state of residence, and equal to 0 if he was surveyed before the peak. In Panel B, we also control for "Days", which is the difference between the date of the survey and the date of the state-level COVID-19 peak, and its interaction with the 'Post' indicator. The dependent variable in columns (1)-(2) is a 0-1 dummy which is equal to 1 if the respondent states to be willing to participate in a protest against corruption in the provision of health services. The dependent variable in columns (3)-(4) is a 0-1 dummy, which is equal to 1 if the respondent is willing to either sign a petition, or donate, or watch a 5 minute video on how to fight fraud and corruption in the health sector. The dependent variable in columns (5)-(6) is a 0-1 dummy that is equal to 1 if the respondent actually signed a petition to the Health Ministry (with his full name), or donated a positive amount to an NGO fighting fraud and demanding transparency in the health sector, or watched a 5 minute video on to help the fight against fraud and corruption in the health sector. Controls include age, marital status, caste, religion, education, income, asset, indicator for frequent repeated participation in Qualtrics surveys, dummy variable for mode of survey participation (phone or computer), indicators for household composition. We report multiple hypothesis corrected p-values using the Romano-Wolf procedure (Romano and Wolf, 2016) in square brackets (across 3 outcomes reported in Table 2 and 6 outcomes reported in Table 3). Results unchanged after adjusting for entropy re-weighting (Hainmueller, 2012), as shown in Table A3 (columns (1) - (3)). Standard errors clustered at the state-month level reported in parentheses. * $p < .10$, ** $p < .05$, *** $p < .01$

Table 3: Mechanisms

	Belief about Others	Information (Rights)	Corruption Perception	Risk Tolerance	Pro-sociality	Corruption Tolerance
	(1)	(2)	(3)	(4)	(5)	(6)
Post	5.947***	0.203***	0.222**	0.152**	-0.079	0.026
	(1.469)	(0.076)	(0.094)	(0.065)	(0.054)	(0.076)
Romano-Wolf p-value	[0.002]	[0.010]	[0.024]	[0.034]	[0.178]	[0.629]
Observations	898	898	898	898	898	898
Pre-peak Mean Y	57.961	0.000	0.000	0.000	0.000	0.000
State FE	yes	yes	yes	yes	yes	yes
Controls	yes	yes	yes	yes	yes	yes
Entropy Balancing	no	no	no	no	no	no

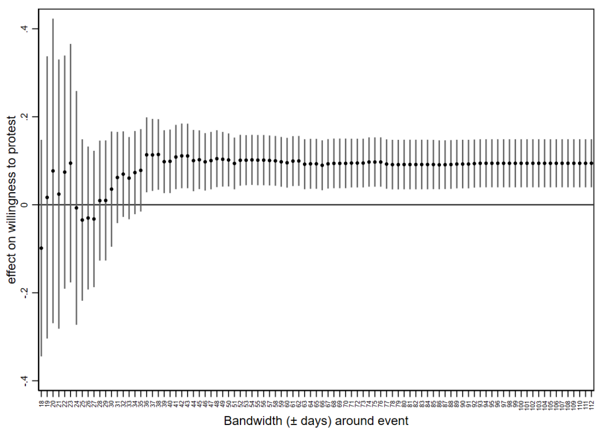
Notes: "Post" is a dummy equal to 1 if the subject was surveyed after the peak in daily COVID-19 cases in his state of residence, and equal to 0 if he was surveyed before the peak. In column (1), the dependent variable - "Belief about Others" - is the percentage of previous survey participants that the respondent believes had answered "yes" to the statement "I am willing to raise my voice and participate in a protest against corruption in the provision of health service". The belief elicitation was incentivized. In column (2) the dependent variable is an index of the subject's degree of information about his own rights and entitlement regarding the provision of health services. In column (3), the dependent variable is an index of individual perceptions of corruption in the health sector. In column 4, the dependent variable - Risk Tolerance - indicates willingness to take risk on a scale from 0 to 10 (0 indicates completely unwilling, and 10 indicates very willing to take risks). In column (5), the dependent variable is a Pro-sociality index, which aggregates answers to questions eliciting trust, altruism and retaliatory preferences (i.e., tendency to punish people who treat you unfairly). The tolerance index - in column (6) - aggregates the respondents' general attitude towards corruption, which combined two questions- (1) the extent to which they think it's justified to pay bribe, or avoid fare or allow doctors to overcharge, and (2) how many people in their community would expect them to complain if they were overcharged or asked to pay a bribe by a doctor. Information on how the indexes are constructed from survey questions is provided in the Online Appendix B. The dependent variables in columns (2) - (6) are standardized around the mean in the pre-peak period, therefore their estimated coefficients are expressed in standard deviations from this mean. Controls include age, marital status, caste, religion, education, income, asset, indicator for frequent repeated participation in Qualtrics surveys, dummy variable for mode of survey participation (phone or computer), indicators for household composition. We report multiple hypothesis corrected p-values using the Romano-Wolf procedure (Romano and Wolf, 2016) in square brackets (across 3 outcomes reported in Table 2 and 6 outcomes reported in Table 3). Results unchanged after adjusting for entropy re-weighting (Hainmueller, 2012), as shown in Table A3 (columns (1) - (3)). Standard errors clustered at the state-month level reported in parentheses. * $p < .10$, ** $p < .05$, *** $p < .01$

Figure 1: Survey Date and State COVID-19 Peaks

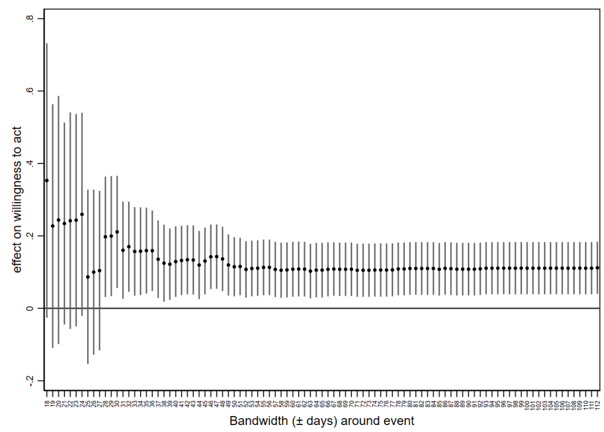


Notes: The red bins in the figure represent the distribution of survey dates. Individuals were surveyed only once, either before or after the COVID-19 peak in their state of residence, during the second wave of the pandemic in India. The grey bins depict the distribution of state-level COVID-19 case peaks.

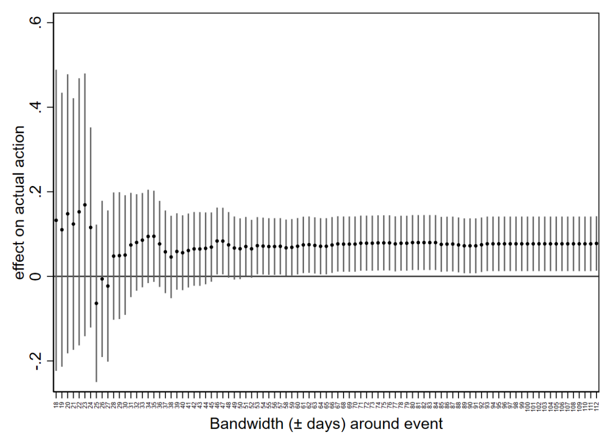
Figure 2: Effect of COVID-19 Exposure Over Multiple Bandwidths



(a) Effect on Willingness to Protest



(b) Effect on Willingness to Act



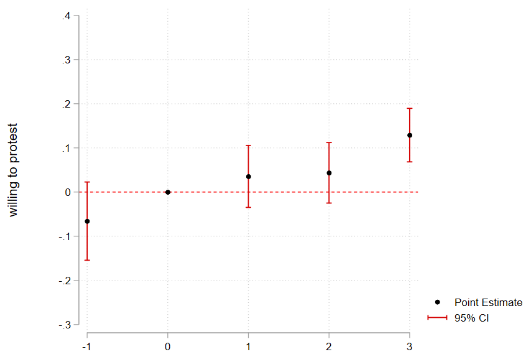
(c) Effect on Actual Action

Notes: The figure plots the coefficients of 'Post' from equation 1, including controls, state fixed effects and entropy balancing, for different time bandwidths. 'Post' is a dummy equal to 1 if the subject participated in the survey after the COVID-19 peak in the state of residence, 0 otherwise. The point estimates are denoted by black dots and the 95% confidence interval appears in gray. The initial bandwidth is ±18 days around the COVID-19 peak of each state, to ensure co-variate balancing. Starting from 18, the bandwidth is then increased progressively by 1 day, up to 112 days pre- and post-state peak. The small sample size makes it unfeasible to employ bandwidths lower than ±18 days. Controls include age, marital status, caste, religion, education, income, asset, indicator for frequent repeated participation in Qualtrics surveys, dummy for mode of survey participation, indicators of household composition, such as, a dummy indicating if the subject has children, another dummy if he lives with parents, and a third dummy if he lives with an elderly (older than 60 years).

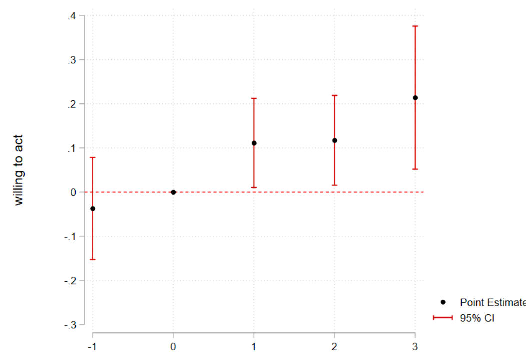
Appendix

A Additional Analysis

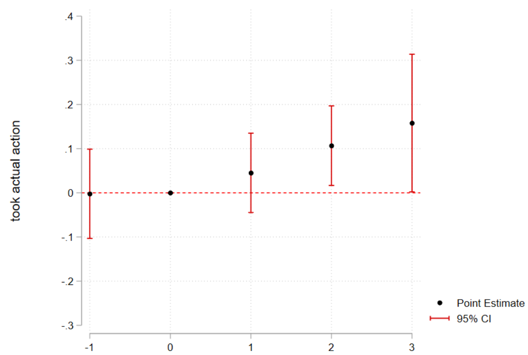
Figure A1: Activism by Month of Survey



(a) Willingness to Protest



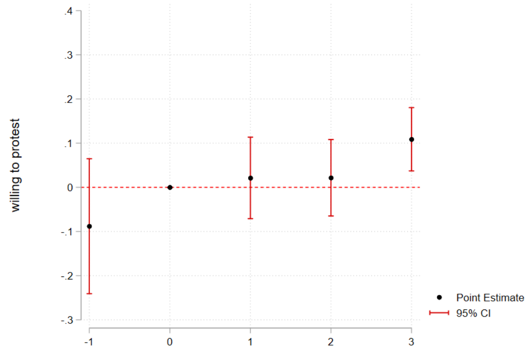
(b) Willingness to Act



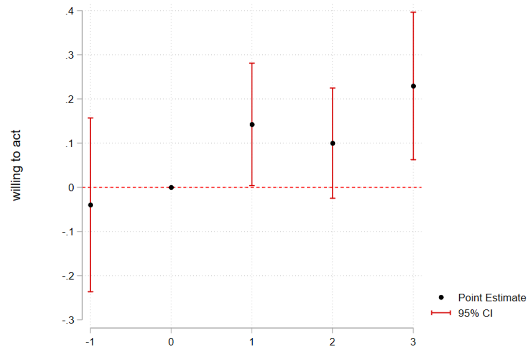
(c) Took Action

Notes: Figure (a), (b), and (c) show, respectively, time paths for subjects' willingness to protest, their willingness to engage in anti-corruption activism, and their likelihood of taking an actual anti-corruption action for the full sample over 5 months. The point estimates are denoted by black dots and 95% confidence interval are displayed in red. The numbers on the horizontal axis indicate the distance from the state level peak, in months. "0" on the horizontal axis indicates the month when the peak in daily cases was reached in a given state. Total count of subjects=898. Controls include age, marital status, caste, religion, education, income, asset, indicator for frequent repeated participation in Qualtrics surveys, dummy for mode of survey participation, indicators of household composition, such as, a dummy indicating if the subject has children, another dummy if he lives with parents, and a third dummy if he lives with an elderly (older than 60 years). We include state fixed effects in all specifications.

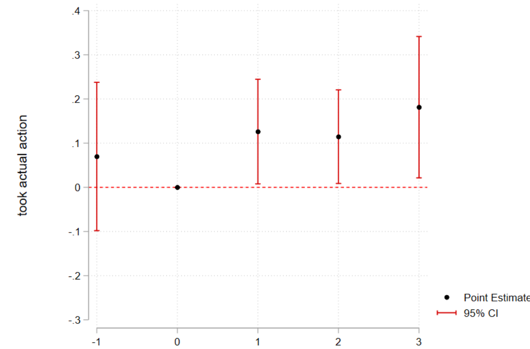
Figure A2: Activism by Month of Survey for Early Peak and Late Peak States



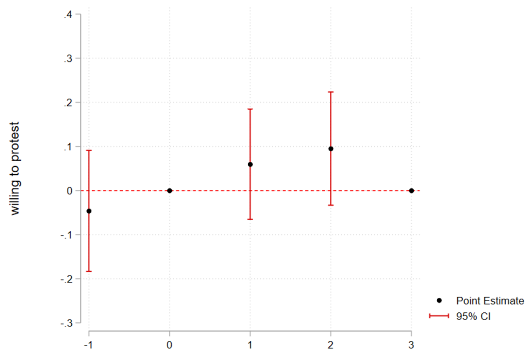
(a) Willingness to Protest (Early Peak)



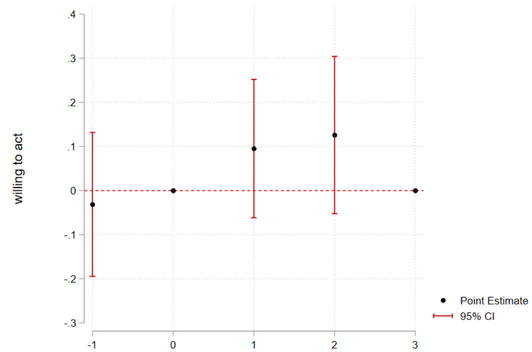
(b) Willingness to Act (Early Peak)



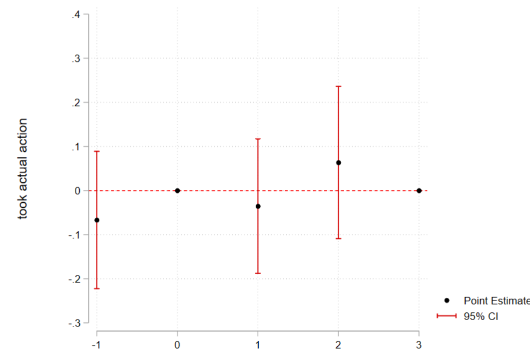
(c) Took Action (Early Peak)



(d) Willingness to Protest (Late Peak)



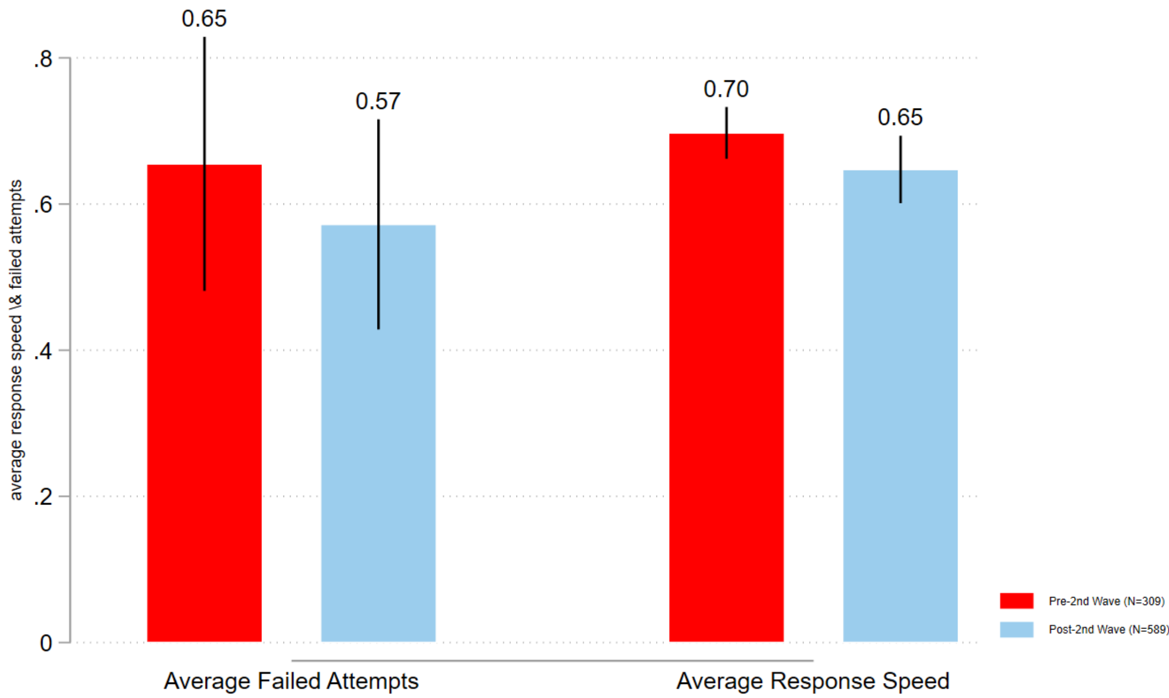
(e) Willingness to Act (Late Peak)



(f) Took Action (Late Peak)

Notes: We divide Indian states in two groups: those who peaked early, i.e., before the national peak of May 8th (7 day M.A.), and those who peaked late. Figures (a), (b) and (c) show time paths for subjects' willingness to protest against fraud and corruption in the health sector, willingness to engage in activism against such corruption, and actual activism for the early-peak sample, respectively, over 5 months. Figures (c), (d) and (e) show the time paths for the same outcome variables when restricting the sample to the late-peak Indian states. The point estimates are denoted by black dots and 95% confidence interval are displayed in red. The numbers on the horizontal axis indicate the distance from the state level peak, in months. "0" on the horizontal axis indicates the month when the peak in daily cases was reached in a given state. Controls include age, marital status, caste, religion, education, income, asset, indicator for frequent repeated participation in Qualtrics surveys, dummy for mode of survey participation, indicators of household composition, such as, a dummy indicating if the subject has children, another dummy if he lives with parents, and a third dummy if he lives with an elderly (older than 60 years). We include state fixed effects in all specifications.

Figure A3: Data Quality Checks Before & After COVID Case-peak



Notes: "Failed attempt" is a continuous variable indicating the number of times a subject attempted the attention manipulation check question embedded in the survey before answering it correctly. The response speed is the number of surveys completed per minute to obtain the quota set by Qualtrics. Specifically, each time the Qualtrics team fielded the survey, they had a target quota that they needed to meet before temporarily closing the survey for data quality checks.

Table A1: COVID-19 Peak Dates by State

State	state-peak
NCT Of Delhi	23-Apr-21
Maharashtra	24-Apr-21
Dadra Nagar Haveli Daman & Diu	26-Apr-21
Uttar Pradesh	27-Apr-21
Chhattisgarh	28-Apr-21
Jharkhand	28-Apr-21
Madhya Pradesh	29-Apr-21
Gujarat	30-Apr-21
Telangana	1-May-21
Bihar	6-May-21
Rajasthan	8-May-21
National Peak on 8-May-21	
Chandigarh	9-May-21
Haryana	9-May-21
Jammu And Kashmir	9-May-21
Karnataka	9-May-21
Goa	11-May-21
Uttarakhand	11-May-21
Kerala	12-May-21
Punjab	12-May-21
Himachal Pradesh	13-May-21
West Bengal	15-May-21
Nagaland	18-May-21
Andhra Pradesh	20-May-21
Assam	22-May-21
Lakshadweep	25-May-21
Meghalaya	25-May-21
Tamil Nadu	25-May-21
Tripura	25-May-21
Orissa	26-May-21
Sikkim	1-Jun-21
Manipur	27-Jul-21

Notes: The table reports the date at which each Indian state experienced the highest number (peak) of confirmed COVID-19 (in terms of 7 day MA) cases during the second wave of the pandemic. Data source: covid19india.org.

Table A2: Survey Dates

Survey dates	N. of Subjects	Cumulative
24-Mar-21	17	17
25-Mar-21	8	25
2-Apr-21	28	53
7-Apr-21	222	275
21-Apr-21	20	295
22-Apr-21	11	306
National peak on 8-May-21		
19-May-21	10	316
26-May-21	106	422
1-Jun-21	9	431
3-Jun-21	1	432
15-Jun-21	132	564
16-Jun-21	82	646
17-Jun-21	5	651
18-Jun-21	20	671
30-Jun-21	89	760
1-Jul-21	43	803
26-Jul-21	95	898
Total	898	

Notes: We report the number of subjects who participated in the survey by date of the survey, before and after the national (7-day moving average) COVID-19 peak. The last column reports the cumulative number of survey respondents.

Table A3: Impact of Second-wave on Willingness to Protest, Willingness to Act, and Actual Action, and Mechanisms

	Willing to Protest (1)	Willing to Act (2)	Took Action (3)	Belief about Others (4)	Information (Rights) (5)	Corruption Perception (6)	Risk Tolerance (7)	Pro-sociality (8)	Corruption Tolerance (9)
Post	0.094***	0.112***	0.078***	6.503***	0.287***	0.238***	0.223***	-0.053	0.073
	(0.021)	(0.030)	(0.022)	(1.347)	(0.070)	(0.086)	(0.065)	(0.052)	(0.068)
Romano-Wolf p-value	[0.002]	[0.004]	[0.008]	[0.002]	[0.002]	[0.022]	[0.010]	[0.417]	[0.417]
Observations	898	898	898	898	898	898	898	898	898
Pre-peak Mean Y	0.832	0.372	0.233	57.961	0.000	-0.000	-0.000	0.000	-0.000
State FE	yes	yes	yes	yes	yes	yes	yes	yes	yes
Controls	yes	yes	yes	yes	yes	yes	yes	yes	yes
Entropy Balancing	yes	yes	yes	yes	yes	yes	yes	yes	yes

Notes: "Post" is a dummy equal to 1 if the subject was surveyed after the peak in daily COVID-19 cases in his state of residence, and equal to 0 if he was surveyed before the peak. The dependent variable in column (1) is a 0-1 dummy which is equal to 1 if the respondent states to be willing to participate in a protest against corruption in the provision of health services. The dependent variable in column (2) is a 0-1 dummy, which is equal to 1 if the respondent is willing to either sign a petition, or donate, or watch a 5 minute video on how to fight fraud and corruption in the health sector. The dependent variable in column (3) is a 0-1 dummy that is equal to 1 if the respondent actually signed a petition to the Health Ministry (with his full name), or donated a positive amount to an NGO fighting fraud and demanding transparency in the health sector, or watched a 5 minute video on to help the fight against fraud and corruption in the health sector. In column (4), the dependent variable - "Belief about Others" - is the percentage of previous survey participants that the respondent believes had answered "yes" to the statement "I am willing to raise my voice and participate in a protest against corruption in the provision of health service". The belief elicitation was incentivized. In column (5) the dependent variable is an index of the subject's degree of information about his own rights and entitlement regarding the provision of health services. In column (6), the dependent variable is an index of individual perceptions of corruption in the health sector. In column (7), the dependent variable - Risk Tolerance - indicates willingness to take risk on a scale from 0 to 10 (0 indicates completely unwilling, and 10 indicates very willing to take risks). In column (8), the dependent variable is a Pro-sociality index, which aggregates answers to questions eliciting trust, altruism and retaliatory preferences (i.e., tendency to punish people who treat you unfairly). The tolerance index - in column (9) - aggregates the respondents' general attitude towards corruption, which combined two questions- firstly, the extent to which they think it's justified to pay bribe, or avoid fare or allow doctors to overcharge, and secondly how many people in their community would expect them to complain if they were overcharged or asked to pay a bribe by a doctor. Information on how the indexes are constructed from survey questions is provided in the Online Appendix B. The dependent variables in columns (5) - (9) are standardized around the mean in the pre-peak period, therefore their estimated coefficients are expressed in standard deviations from this mean. Control variables include age, marital status, caste, religion, education, income, asset, indicator for frequent repeated participation in Qualtrics surveys, dummy for mode of survey participation, indicators of household composition, such as, a dummy variable for mode of survey participation (phone or computer), indicators for household composition. We report multiple hypothesis corrected p-values using the Romano-Wolf procedure (Romano and Wolf, 2016) in square brackets (across all 9 outcomes reported in columns (1) - (9)). "Entropy Balancing" indicates that the observations are weighted by the control variables so that the first and second moments of the distribution of co-variates among "Post" subjects mimic those of the "Pre" subjects (Hainmueller, 2012). Standard errors clustered at the state-month level reported in parentheses. * $p < .10$, ** $p < .05$, *** $p < .01$

Table A4: Impact of Second-wave on Probability of Employment

	Employed	
	(1)	(2)
Post	0.002	-0.002
	(0.004)	(0.005)
Observations	227,200	227,200
Pre-peak Mean Y	0.657	0.657
State FE	yes	yes
Controls	no	yes

Notes: This regression analysis uses data from the nation-wide Consumer Pyramids Household Survey (CPHS) of the Centre for Monitoring Indian Economy (CMIE), in which each household is surveyed three times a year. We use data for the sample of adult male Indians interviewed between the months of March-July 2021. "Employed" is a 0-1 dummy variable equal to 1 if the subject worked on the day of the survey or the day prior, or was not working but returning to work in the near future. "Post" is a dummy equal to 1 if the subject was surveyed in a month after the month containing the peak in daily COVID-19 cases in his state of residence, and equal to 0 if he was surveyed before the month containing the peak. Controls include indicators for marital status, religion, caste category, literacy status, and age (in years). Standard errors clustered at the state-month level reported in parentheses. * $p < .10$, ** $p < .05$, *** $p < .01$

Table A5: Robustness: Clustering Standard Errors at the State Level

	Willing to Protest		Willing to Act		Took Action	
	(1)	(2)	(3)	(4)	(5)	(6)
Post	0.075***	0.094***	0.116***	0.112***	0.080**	0.078***
	[0.003]	[0.002]	[0.003]	[0.004]	[0.007]	[0.011]
Observations	898	898	898	898	898	898
State FE	yes	yes	yes	yes	yes	yes
Controls	yes	yes	yes	yes	yes	yes
Entropy Balancing	no	yes	no	yes	no	yes

Notes: "Post" is a dummy equal to 1 if the subject was surveyed after the peak in daily COVID-19 cases in his state of residence, and equal to 0 if he was surveyed before the peak. The dependent variable in columns (1)-(2) is a dummy which equals 1 if the respondent states to be willing to participate in a protest against corruption in the provision of health services. The dependent variable in columns (3)-(4) is a dummy which equals 1 if the respondent is willing to either sign a petition, or donate, or watch a 5 minute video on how to fight fraud and corruption in the health sector. The dependent variable in column (5)-(6) is a dummy that equals 1 if the respondent actually signed a petition to the Health Ministry (with his full name), or donated a positive amount to our collaborator NGO fighting fraud and demanding transparency in the health sector, or watched a 5 minute video on to help the fight against fraud and corruption in the health sector. Controls include age, marital status, caste, religion, education, income, asset, indicator for frequent repeated participation in Qualtrics surveys, dummy variable for mode of survey participation (phone or computer), indicators for household composition."Entropy Balancing" indicates that the observations are weighted by the control variables using entropy balancing; the weighting is such that the first and second moments of the distribution of covariates among "Post" subjects mimic those of the "Pre" subjects (Hainmueller, 2012). Standard errors are clustered at the state level, and wild-cluster bootstrapping is used to correct for the small number of clusters (N=31). p-values reported below coefficients in square brackets: * $p < .10$, ** $p < .05$, *** $p < .01$

Table A6: Heterogeneity by High Corruption Level, and by Hospital Density in Indian States

	Willing to Protest (1)	Willing to Act (2)	Took Action (3)
Panel A			
Post x *high corruption state*	0.039	0.137**	-0.019
	(0.044)	(0.066)	(0.048)
Post	0.074**	0.058	0.093***
	(0.028)	(0.037)	(0.028)
Observations	848	848	848
Panel B			
Post x *hospital density*	0.003	0.012**	0.006
	(0.005)	(0.006)	(0.004)
Post	0.079**	0.039	0.041
	(0.033)	(0.059)	(0.039)
Observations	898	898	898
State FE	yes	yes	yes
Controls	yes	yes	yes
Entropy Balancing	yes	yes	yes

Notes: In both Panels A and B, "Post" is a dummy equal to 1 if the subject was surveyed after the peak in daily COVID-19 cases in his state of residence, and equal to 0 if he was surveyed before the peak. In Panel A, *high corruption state* is a dummy equal to 1 for Indian state with high corruption levels, according to the India Corruption Survey Report by Transparency International India (2019), and 0 otherwise. Missing observations in Panel A correspond to 9 states that were not included in the 2019 survey. In Panel B, *Hospital density* indicates the number of hospitals per 100,000 population in a state. Data on state level population is taken from Government of India (2019), whereas data on the number of hospitals (public and private) is taken from Kapoor et al. (2020). Controls include age, marital status, caste, religion, education, income, asset, indicator for frequent repeated participation in Qualtrics surveys, dummy variable for mode of survey participation (phone or computer), indicators for household composition. "Entropy Balancing" indicates that the observations are weighted by the control variables using entropy balancing; the weighting is such that the first and second moments of the distribution of covariates among "Post" subjects mimic those of the "Pre" subjects (Hainmueller, 2012). Standard errors clustered at the state-month level reported in parentheses. * $p < .10$, ** $p < .05$, *** $p < .01$

Table A7: Long Term Impact of COVID-19 on Willingness to Protest

	Long-term Sample			Full Sample		
	(1)	(2)	(3)	(4)	(5)	(6)
Post	0.136***	0.091***	0.049	0.111***	0.094***	0.572***
	(0.024)	(0.022)	(0.025)	(0.077)	(0.019)	(0.107)
Observations	1158	1158	1158	1747	1747	1747
Pre-peak Mean Y	0.832	0.832	0.832	0.832	0.832	0.832
Controls	no	yes	yes	no	yes	yes
State FE	no	yes	yes	no	yes	yes
Date of Survey FE	no	no	yes	no	no	yes
Entropy Balancing	no	yes	yes	no	yes	yes

Notes: The dependent variable is equal to 1 if the survey participant stated his willingness to protest to fight corruption in the provision of health services. The *Long-term Sample* refers to subjects who participated in the survey more than a year after the initial study, i.e., in October and November 2022. The *Full Sample* refers to all survey participants, i.e., those surveyed between March and July 2021, and those surveyed in October and November 2022. In the *Long-term Sample*, in columns (1) - (3), "Post" is a 0-1 dummy that equals 1 if the subject participated in the survey in October or November 2022, over a year after the COVID-19 state-level peak; it is equal to 0, if the subject participated in the survey before the peak. In the *Full Sample*, in columns (4) - (6), "Post" is a 0-1 dummy that equals 1 if the subject participated in the survey after the state-level COVID-19 peak in either 2021 & 2022, and 0 if he participated before the peak. The longer time span allows us to include Date of Survey Fixed Effects in the most comprehensive specifications, in columns (3) and (6). Controls include age, marital status, caste, religion, education, income, indicator for frequent repeated participation in Qualtrics surveys, dummy variable for mode of survey participation (phone or computer), an indicator for household composition. "Entropy Balancing" indicates that the observations are weighted by the control variables using entropy balancing; the weighting is such that the first and second moments of the distribution of covariates among "Post" subjects mimic those of the "Pre" subjects (Hainmueller, 2012). Standard errors clustered at the state-month level reported in parentheses. * $p < .10$, ** $p < .05$, *** $p < .01$

B Data Generation Procedures

B.1 Sampling

In order to measure whether the subjects are paying attention to the survey and understand the instructions of the questionnaire, we employ a variety of checks and screener questions within the survey.

- The first screener question is a simple one to catch subjects who paid the least attention. Following the suggestions of Oppenheimer et al. (2009), we include the following question: "People are very busy these days and many do not have time to follow what goes on in the government. Some do pay attention to politics but do not read questions carefully. To show that you have read this much, please ignore the question below and just select the option C from the four choices below. That's right, just select the option C from the four choices below.

  How interested are you in information about what's going on in government and politics? (answer choices: option A/ option B/ option C/ option D)"

  Subjects who failed to pick option C are considered as 'inattentive'. We don't outright disqualify these subjects from continuing the survey, but they are not included in the final analysis sample.

- We then place three training questions prior to the belief questions that were real-effort, to capture the subjects' comprehension of how much they're going to earn from the real-effort questions. Using the set of training questions, we measure the number of failed attempts for each subject to grasp their prospective earnings.

- Finally, we include a descriptive question; "Some people who are asked to pay bribes do not complain about it. Why do you think this is the case? Please type your response in the text box below." We consider as inattentive the subjects who enter nonsensical text when answering this question.

Overall, we find that these three indicators of attention are highly correlated. Inattentive subjects are also more likely to have a much higher number of failed attempts in the training questions, and are more likely to leave a gibberish answer in the descriptive question. We do not find the proportion of inattentive subjects to vary significantly between treatment groups. Hence, from the main analysis sample, we decide to exclude them. This brings our subject pool to 898.

B.2 Procedure for Standardization and Index Construction

We constructed indices for corruption experience and individual preferences. These are the average of the relevant standardized variables, as listed in below. The procedure is as follows-

- Individual variables are coded such that the positive direction always corresponded with "higher" outcome for all sub-components of the aggregate index, 0 otherwise.
- Each variable is normalized by subtracting the overall control (pre-2nd wave) mean and dividing by the control group standard deviation. The index is then generated by averaging over relevant components.
- The final index is then re-scaled such that the control mean is 0 and the standard deviation is 1.

B.2.1 Corruption Perception

The corruption perception index aggregates the following survey questions.

- "Please consider all the contact you or members of your household had with health workers in clinics or hospitals since April 2020 till date. How many times did you have to pay extra money to obtain a medical service? (never/1/2/.../10/more than 10 times)."[41]
- "In your opinion, has the level of corruption in the health sector during the COVID-19 pandemic - (increased a lot/ increased somewhat/ stayed the same/ decreased somewhat/ decreased a lot)"[42]?
- "According to your experience, the current level of corruption in the health sector is - (not a problem at all/ a small problem/ a moderate problem/ a major problem)"[43].

B.2.2 Information (Rights)

Subjects' information on rights and entitlements are captured through this index, which aggregates the following survey questions.

- "Do you know what is the rate you have to pay per day for an ICU bed at your local hospital?"[44]

[41] response coded into a continuous variable.
[42] response coded into a continuous variable with higher value indicating increase in corruption.
[43] response coded into a continuous variable with higher value indicating bigger problem.
[44] response coded into a dummy=0 if subject answered with 'don't know', 1 otherwise.

- "Do you think you or a member of your household were illegally overcharged by the healthcare professionals for the hospital stay? - (does not apply / don't know or can't say/ no/ yes)"[45]

B.2.3 Corruption Tolerance

The corruption tolerance index aggregates the following survey questions.

- "Please tell us for each of the following actions whether you think it can never be justified, always be justified or something in between using a scale of 1 to 10 below (1 denotes never justifiable, and 10 denotes always justifiable)."[46]
  - avoiding fare on a public transport
  - doctors overcharging for a hospital bed during COVID-19 pandemic
  - someone accepting a bribe in course of their duties.
- "How many people in your community do you think expects you to complain if you are overcharged or asked to pay a bribe by a doctor? (nobody/ a few people/ many people/ most people/ everybody)."[47]

B.2.4 Preferences

'Risk' is a self-assessed measure of risk preference. Similarly, the pro-sociality index is generated by combining self-assessment indices of trust, retaliation and altruism. These variables are measured following Falk et al. (2018):

- The *risk index* is computed using response to "Please tell us, in general, how willing or unwilling are you to take risks, using a scale of 0 to 10 below (0 indicates completely unwilling, and 10 indicates very willing to take risks.) (answer choices: completely unwilling 0/ 1/ ..../very willing 10)"
- *Trust* is computed using response to "Please tell us whether the following statement describes you as a person: you assume that people only have the best intentions, using a scale of 0 to 10 below (0 indicates that the statement does not describe you at all, and 10 indicates that the statement describes you perfectly). (doesn't describe you at all 0/1/ .../ describes you perfectly 10)."

[45] response coded into a dummy=1 if subject answered with a 'yes.

[46] responses coded into a continuous variable.

[47] response coded into a dummy=1 if subject answered with 'nobody'.

- *Retaliatory behavior* is based on response to
  - "Please tell us whether, if you are treated very unjustly, you will take revenge at the first opportunity, even if there is a cost to do so, using a scale of 0 to 10 below (0 indicates you are completely unwilling to take revenge, 10 indicates you are very willing to take revenge)."
  - "Please tell us how willing you are to punish someone who treats you unfairly, even if there may be costs for you, using a scale of 0 to 10 below (0 indicates you are completely unwilling to do so, 10 indicates you are very willing to do so)."
  - "Please tell us how willing you are to punish someone who treats others unfairly, even if there may be costs for you, using a scale of 0 to 10 below (0 indicates you are completely unwilling to do so, 10 indicates you are very willing to do so)."
- *Altruism* is measured by response to "Please tell us how willing you are to give to good causes without expecting anything in return, using a scale of 0 to 10 below (0 indicates you are completely unwilling to give, 10 indicates you are very willing to give) (answer choices: completely unwilling to give 0/ 1/ ..../ very willing to give 10)."

The trust, altruism and reverse-coded retaliation measures are combined to create the pro-sociality index using the same process described above.