

NBER WORKING PAPER SERIES

TRUST IN STATE AND NON-STATE ACTORS:  
EVIDENCE FROM DISPUTE RESOLUTION IN PAKISTAN

Daron Acemoglu  
Ali Cheema  
Asim I. Khwaja  
James A. Robinson

Working Paper 24611  
<http://www.nber.org/papers/w24611>

NATIONAL BUREAU OF ECONOMIC RESEARCH  
1050 Massachusetts Avenue  
Cambridge, MA 02138  
May 2018

We thank Niharika Singh, Christina Brown, Amanda Rudin, and Landin Smith for outstanding research assistance in Cambridge and Nadia Hasham, Neha Zaigham, Ahmed Raza, Zain Chaudhry, and Kamran Niazi for outstanding research assistance in Lahore. This paper was funded through support from the Abdul Latif Jameel Poverty Action Lab's Governance Initiative, as well as the Roy and Lila Ash Center for Democratic Governance and Innovation, the Weatherhead Center for International Affairs, and the South Asia Institute at Harvard University.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2018 by Daron Acemoglu, Ali Cheema, Asim I. Khwaja, and James A. Robinson. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Trust in State and Non-State Actors: Evidence from Dispute Resolution in Pakistan  
Daron Acemoglu, Ali Cheema, Asim I. Khwaja, and James A. Robinson  
NBER Working Paper No. 24611  
May 2018  
JEL No. C91,C93,D02,D73,D74,D83,K40,O17,P16

### **ABSTRACT**

Lack of trust in state institutions, often due to poor service provision, is a pervasive problem in many developing countries. If this increases reliance on non-state actors for crucial services, the resulting self-reinforcing cycle can further weaken the state. This paper examines whether such a cycle can be disrupted. We focus on dispute resolution in rural Punjab, Pakistan. We find that providing information about reduced delays in state courts leads to citizens reporting higher willingness to use state courts and to greater fund allocations to the state in two lab-in-the-field games designed to measure trust in state and non-state actors in a high-stakes setting. More interestingly, we find indirect effects on non-state actors. After receiving state positive information, respondents report lower likelihood of using non-state institutions and reduce funds allocated to them in field games. Furthermore, we find similar direct and indirect effects on a battery of questions concerning people's beliefs about these actors, including a decreased allegiance to the non-state actor. We rationalize these results with a model of motivated reasoning whereby reduced usage of non-state institutions makes people less likely to hold positive views about them. These results indicate that, despite substantial distrust of the state in Pakistan, credible new information can change beliefs and behavior. The feedback loop between state ineffectiveness and the legitimacy of non-state actors may be reversible.

Daron Acemoglu  
Department of Economics, E52-446  
MIT  
77 Massachusetts Avenue  
Cambridge, MA 02139  
and CIFAR  
and also NBER  
daron@mit.edu

Ali Cheema  
Department of Economics  
Lahore University of Management Sciences  
and Institute of Development Studies  
cheema@lums.edu.pk

Asim I. Khwaja  
Harvard Kennedy School  
Harvard University  
79 JFK Street  
Cambridge, MA 02138  
and NBER  
asim\_ijaz\_khwaja@harvard.edu

James A. Robinson  
University of Chicago  
Harris School of Public Policy  
1155 East 60th Street  
Chicago, Illinois 60637  
and NBER  
jamesrobinson@uchicago.edu

# 1 Introduction

Endemic lack of state capacity is one of the most challenging problems facing many less developed societies around the world.<sup>1</sup> Though the ineffectiveness of state institutions has complex historical and contemporary causes ranging from institutional deficiencies to corruption and lack of adequate resources, it becomes exacerbated as it undermines trust in state institutions and what is sometimes referred to as the state’s “legitimacy”. This problem is further intensified as powerful non-state actors step in to fill the void, providing competing services such as protection, conflict resolution, or even broader public goods (see, for example, Clark (2004), Harmsen (2010), and Roy (2013) for case studies in the context of various Middle Eastern countries). One might conjecture that the shifting balance between state and non-state actors creates a feedback process — perhaps even a vicious cycle — where state weakness leads to greater trust in competing non-state actors, further reducing trust in the state and the state’s capacity to perform its functions.

Although this interplay between state and non-state actors is plausible, there is little direct evidence that the strength of non-state actors derives from the weakness of state institutions. Similarly, there has been no investigation, to the best of our knowledge, of whether the pervasive lack of trust in the state in many societies can be redressed. In this paper, we investigate these issues by using “lab-in-the-field” experiments in rural Punjab, Pakistan. Our study aims to shed light on two related questions. First, after documenting the pervasive belief among Pakistanis about the ineffectiveness of state courts, we study whether providing objective media reported information about improved service delivery (as evidenced by reductions in the number of pending cases in these courts) can change people’s beliefs and behavior. In particular, we ask if this information makes them more willing to use, interact with, and trust these state institutions. Second, more pertinent to the issue of negative feedback between state and non-state institutions, we investigate whether such positive information about state courts makes our subjects less willing to interact with competing non-state actors, and whether it also changes their beliefs about these non-state actors — even though this information has no direct relevance to the non-state actors’ effectiveness.

Pakistan is an ideal setting for such an investigation because of the well-recognized weakness of state institutions and the associated low levels of trust in the state (Jackson et al. (2014), Cheema et al. (2017)), as well as the critical role that various non-state actors have come to play in parts of the country, especially in dispute resolution (see Chaudhary (1999), Gayer (2014), Shinwari (2015), Siddique (2013), Siddique (2015)). Ineffectiveness of state courts in Pakistan is one of the key dimensions of state weakness that has both spawned general discontent and generated support for non-state actors such as the Taliban. For example, in 2009 Taliban militants took control of parts

---

<sup>1</sup>A growing literature in political science, sociology, and economics emphasizes the central role of state capacity for economic development (see, inter alia, Johnson (1982), Amsden (1992), Wade (1990)), while weak and ineffective (“low capacity”) states are often argued to cause poverty, instability and even civil war (e.g., Migdal (1988), Herbst (2014), Centeno (2002), Acemoglu et al. (2011), Besley and Persson (2009)).

of Pakistan’s frontier province of Khyber Pakhtunkhwa (KP) and instituted parallel justice and administrative systems based on Sharia and funded through taxes they imposed on the population (Rana (2009), Rehman et al. (2014)).<sup>2</sup> Whilst the Taliban and other associated extremist Islamist groups are the most visible threat to state institutions in Pakistan, the involvement of non-state actors in the provision of dispute resolution and public services is much more pervasive, and our study focuses on the role of *panchayats* in dispute resolution.<sup>3</sup> Panchayats, comprised of groups of village elders and other influential locals, are a mainstay of conflict resolution in rural Pakistan and the main alternative to state courts. Notably, panchayats base their rulings on cultural norms and do not typically follow laws promulgated by the Pakistani state.<sup>4</sup>

We start with a simple conceptual framework that motivates our overall approach and experimental design. Our model considers an individual who is interacting with two forums, state and non-state, under incomplete information about their quality (effectiveness, corruption, ease of access, etc.). This interaction has two dimensions. The first is a contribution, which could be voluntary or correspond to less evasion of tax obligations, which is used by the forum to deliver services. The unknown quality of the forum affects the benefits obtained from these services. The second is a relationship-specific effort, broadly corresponding to the notion of an individual developing an allegiance to one (or both) of these forums. We show that providing positive (media-reported) information about the quality of state institutions increases contributions to and allegiance towards the state forum. Perhaps more interestingly, we further show that when individuals engage in “motivated reasoning” and upgrade their beliefs concerning the state forum,<sup>5</sup> they will also downgrade their beliefs concerning the non-state forum, and, consequently, reduce their contributions to and allegiance towards the non-state actor. This, in turn, exacerbates their shift away from non-state institutions in response to positive information about state institutions.

---

<sup>2</sup>After the army retook control of the province, the President of Pakistan established Sharia courts as part of the settlement to end the conflict. This was an acknowledgment of the discontent surrounding the state courts that had fueled the Taliban’s rise (Walsh (2009), Siddique (2013)).

<sup>3</sup>A study directly focusing on the Taliban would have been unsafe in the current Pakistani context. It was also not possible to obtain official permission for research in areas undergoing violent insurgency.

<sup>4</sup>It is important to emphasize that, in contrast to panchayats in India, which are part of the local government structure, the panchayat system in Pakistan is entirely outside of the control of the state and often competes with it (Chaudhary (1999), Shinwari (2015), Siddique (2015)). Like India, Pakistan’s local governments have also introduced provisions to enact local mediation bodies called *Musalihat-i-Anjuman*; however, these bodies have not been institutionalized because local governments have been periodically disbanded (Shinwari (2015), Cheema et al. (2010)).

<sup>5</sup>Motivated reasoning refers to the possibility that individuals manipulate their own beliefs either for direct benefit or for strategic purposes. The theory of motivated reasoning in psychology goes back at least to Festinger’s (1962) theory of cognitive dissonance and to Bem (1967). See also Kunda (1990), Gilovich and Ross (2015), Epley and Gilovich (2016), Edwards and Smith (1996), Jost et al. (2003) and Kahan (2013) for more recent discussions, and Trivers (2011) and Von Hippel and Trivers (2011) for an approach emphasizing the benefits of motivated reasoning from an evolutionary viewpoint. In economics, the first application of motivated reasoning might be Akerlof and Dickens’s (1982) use of ideas from cognitive dissonance in occupation choice. For more recent contributions in economics, see Loewenstein (1987), Rabin (1994), Carrillo and Mariotti (2001), Caplin and Leahy (2001), Bénabou and Tirole (2002), Bénabou and Tirole (2004), Bénabou and Tirole (2016), Brunnermeier and Parker (2005), Van den Steen (2004), and Ortoleva and Snowberg (2015).

In our empirical analysis, we use two approaches to capture behavior and attitudes towards state and non-state institutions. First, we collect survey information on expected usage and assessment of state courts and non-state dispute resolution forums like panchayats. Second, we design relatively high-stakes lab games meant to both address concerns arising from using self-reported data and to elucidate different aspects of behavior toward these forums. Our subjects can earn as much as 550 PKR (approximately 5.3 USD) in these games, equivalent to one and a half times the average daily earnings in our sample. The first game, which we call the fund dictator game, is a version of the well-known dictator game in experimental economics. It gives our subjects a choice between allocating a certain pot of money between themselves and a fund that would help others access state courts. They then make the same decision for a fund to improve access to panchayats using a separate pot of money. Thus, this game is meant to measure how much they are willing to contribute to these two forums. Our second game, referred to as the investment game, is related to the trust game in experimental economics. It asks our subjects to invest money in a complainant’s case that is being addressed by either the state or the non-state actor, with the understanding that this “investment” will be repaid depending on the effectiveness of the relevant actor. This second game is therefore aimed at measuring the subject’s trust and belief in the effectiveness of the two competing forums. The game is framed in terms that evoke familiar co-investments that rural respondents often engage in.

Our within-subject experimental design relies on first capturing baseline behavior and beliefs and subsequently providing respondents with information about improved court efficiency. We then measure how this information changes their game behavior and responses about expected usage, beliefs, trust, and allegiance. At baseline, people report significantly lower expected usage and also choose somewhat lower levels of allocations to the state in both games. Once people receive the informational treatment, we see a notable improvement in expected usage of state courts — indicating that the information we provided is indeed believed. We also estimate fairly large *direct effects* on their allocations to the state in both the fund dictator and investment games. These direct impacts are not driven by “social experimenter effects” whereby our subjects change their behavior because they think this is what we would like them to do. We verify this by providing a randomly-selected sample with a statement that does not contain any information about improved performance of state courts, but a clearly stated opinion favorable to the state. There are no significant changes in the allocations in the two games following this “social experimenter treatment,” and there is a much smaller effect on self-reported expected usage. We then explicitly “net out” any such social experimenter effects and still find large and robust effects of the main informational treatment. We also confirm the robustness of these direct effects using an extensive set of checks that address concerns regarding empirical specification, outliers, and respondent understanding. Overall, it appears that despite the deep-rooted mistrust of the state, credible new information can make people change

their views and behavior. There is nothing hardwired about the lack of trust in the state in rural Pakistan.

Our second set of results is more striking. We find large and robust *indirect effects*. Following the positive information about state courts, self-reported expected usage of panchayats declines sharply, and our subjects choose significantly lower allocations towards panchayats in both the fund dictator and investment games. Again, these results hold after netting-out any social experimenter effects, and are robust to a range of data and specification checks.

We then include additional respondents in our study to further investigate the mechanisms through which these direct and indirect effects are operating. This analysis examines the impact of our state positive informational treatment on two specific dimensions of their beliefs about the effectiveness of state and non-state forums — “service effectiveness,” which concerns the quality of the service, and “enforcement effectiveness,” which measures how well the judgment in the relevant forum will be enforced. In addition, we also ask questions on ease of access to the relevant forum, trust in the forum, and “allegiance,” meant to capture subjects’ willingness to exert effort towards building a relationship with the relevant forum. This additional analysis allows us to first verify that the informational treatment improves our subjects’ beliefs about state effectiveness, access, and trust. Interestingly, we find no evidence of increased effort on fostering allegiance towards the state, perhaps suggesting that the respondents do not consider this to be as feasible. We then turn to indirect effects, and find that beliefs regarding effectiveness and trust concerning the non-state forum deteriorate following the positive information about state courts. We also find evidence that efforts toward establishing allegiance to the non-state actor decreases following the positive state information provision. These overall results are useful in two related ways for our discussion of potential mechanisms. First, in terms of our conceptual framework, they provide evidence of motivated reasoning, whereby our subjects, once they expect to interact less with non-state actors, feel they no longer need to entertain as positive views of them. Second, they can be interpreted as evidence of the feedback mechanisms mentioned above — whereby positive views about non-state actors are fed by negative beliefs regarding the effectiveness of state institutions, and vice versa.

Though our main results focus on an informational treatment that provides positive information about state courts, in our pilots we also tried the three remaining combinations. We provided (objective, media-reported) information about less successful dimensions of the state court’s performance as well as negative and positive information about panchayats. Despite the smaller sample sizes in these cases, we once again find direct and indirect effects consistent with the feedback mechanisms linking beliefs about the state and non-state actors. This suggests that the feedback between perceptions of state and non-state forums holds more generally than our baseline focusing on the impact of positive new information about state courts.

Our paper contributes to a number of literatures. While there is an extensive literature on the

implications of low state capacity in the development process and a similarly large literature on the origins of state capacity, there is little work about how state and non-state institutions interact and compete. There has also been only limited emphasis on the role of trust and cooperation of citizens on the effective functioning of state institutions. This is suggested in Peter Evans' notion of "embedded autonomy" (Evans 2012), and in a few works in political science (e.g., Mishler and Rose (2001), Rothstein and Stolle (2008), Zmerli and Newton (2008)). Furthermore, it has been modeled in the context of "consensually strong states," defined as states that derive authority from citizens who have the capability to rein them back (Acemoglu 2005). Recent work by Dell et al. (2015) argues that the greater capacity of the north Vietnamese state (relative to areas in the south that were under the historical influence of the Khmer Empire) is related to the cooperation of villagers. As mentioned above, the rise of extremist religious organizations such as Hamas, the Muslim Brotherhood, and various Salafist groups has been linked to the weakness of the state in the qualitative literature on Middle Eastern politics (e.g., Clark (2004), Kepel (2009), Roy (2013)). However, to the best of our knowledge, it has not been developed systematically nor empirically investigated.

Our approach is related to and builds on several different strands in the experimental economics literature as well. There is a growing literature using experimental methods to measure trust, beliefs, and norms in various different settings (see, for example, Berg et al. (1995), Burks et al. (2003), Dufwenberg and Gneezy (2000), Bohnet and Zeckhauser (2004), Camerer and Fehr (2004), Ashraf et al. (2006), Johnson and Mislin (2011)). Most of this literature does not focus on attitudes towards state institutions, the notable exception being Cox et al. (2009). Another branch of the literature related to our work investigates various dimensions of extremism. For example, Bullock et al. (2011) and Blair et al. (2013) look at support for militant groups in Pakistan, while Delavande and Zafar (2012) and Bursztyn et al. (2016) focus on anti-American attitudes. There is also a small literature using lab-in-the-field games in the context of development economics (see the survey by Cardenas and Carpenter (2008)).

Finally, some works in the sociology and social psychology literatures are also related to our paper. The emphasis on political trust and its crucial role in the functioning of institutions goes back at least to Coleman (1990). Anderson (2010), Sullivan and Transue (1999) and Schoon and Cheng (2011) emphasize the role of individual experiences in shaping political trust, while a number of other works develop similar ideas in the context of organizations (e.g., Kramer (1999)).

The remainder of this paper is organized as follows. Section 2 presents our conceptual framework. Section 3 describes the context of dispute resolution in Pakistan. Section 4 provides the details of our experimental design and empirical strategy. Section 5 presents our main empirical results, and section 6 concludes. The Appendix contains additional robustness checks, treatment texts, and selected question texts.

## 2 Conceptual Framework

In this section, we briefly discuss a simple conceptual framework that motivates our empirical work and helps with the interpretation of the results.

### 2.1 Setup

We consider a setting in which two actors, one state and one non-state, offer services to an individual. Crucially, the individual is uncertain about the quality and effectiveness of these two actors. For simplicity, we collapse this uncertainty into two binary variables representing the qualities of the two actors,  $\theta_S \in \{0, 1\}$  and  $\theta_N \in \{0, 1\}$ . We denote the prior beliefs about these two quality variables by  $\pi_S^0 = \mathbb{E}[\theta_S | \text{Prior information}]$  and  $\pi_N^0 = \mathbb{E}[\theta_N | \text{Prior information}]$ . Our experimental treatment will be to provide (some of) the subjects with relevant information (signal  $\sigma_S$ ) about the quality of the state actor.<sup>6</sup> We denote the beliefs of the individual after he receives this information by  $\pi_S = \mathbb{E}[\theta_S | \text{Prior information}, \sigma_S]$  and  $\pi_N = \mathbb{E}[\theta_N | \text{Prior information}, \sigma_S]$ .

We also assume each individual has the following (expected) utility function:

$$c + a_S \pi_S g_S + a_N \pi_N g_N - \Gamma_S(a_S) - \Gamma_N(a_N), \quad (1)$$

where  $c$  denotes consumption, and  $g_S$  and  $g_N$  are services provided respectively by the state and the non-state actor. These terms are multiplied by  $\pi_S$  and  $\pi_N$  to capture the fact that the quality of the relevant actor matters for utility (or effectiveness of services). In our model, this is the reason why the individual cares about the quality of the state and non-state actors. In mapping these terms to our setting, we interpret them as incorporating not just what the individual himself is receiving but also the services to the entire community that the individual may care about (this is relevant in the context of our fund dictator game where our subjects make contributions for other people's benefit). Relatedly, in the context of the investment game these terms may be interpreted as additional income resulting from the individual's investment in the court case of another plaintiff.

Utility from the services provided by the state and the non-state actor depends on  $a_S$  and  $a_N$  as well. These variables capture both the usage of the relevant forum by the individual and the actor-specific efforts that he undertakes. For instance, the individual needs to spend time with the relevant forum to understand how it operates, and, in the case of the panchayat, he may need to invest time and effort in order to stand in the good grace of the elders who will make the ruling. These variables might additionally capture psychological benefits related to the allegiance of the individual towards the state or the non-state actor (e.g., whether the individual feels comfortable interacting with the relevant actor, or whether he feels that he is receiving services from a forum

---

<sup>6</sup>As we discuss below, in our pilot experiments, we also gave information about the non-state actor, but here we focus on our main treatment and sample for which the information provided concerned the effectiveness of the state actor.



to which he does not belong). The last two terms in the utility function are the costs of these actor-specific efforts, which are assumed to be convex and differentiable.

An important aspect of the utility function in (1) is its quasi-linearity: because utility from consumption is linear, there are no income effects. This is appropriate in our setting for at least two reasons. First, our experimental design clarifies that our subjects will be paid for only one of the games they are playing, so we expect only limited across-game effects working through the marginal utility of consumption. In particular, conditional on one of the games being chosen for payment, allocations in the other games have no impact on the subject's take-home amount and thus should not generate any income effects on their choices. Second, one of our games, the investment game, focuses on investing resources with potential returns, so income effects should not be present in this context. These justifications notwithstanding, below we explain further how the results would be different if instead we introduced income effects.

The individual starts with income  $y$ , and has to decide how much to contribute to the state and the non-state actor, denoted by  $T_S$  and  $T_N$ . In reality, this includes taxes (which individuals can try to evade by taking costly actions) and voluntary contributions. In our games, individuals have an explicit decision to make about these contributions. We assume that these contributions affect the amount of services that the individual receives via the following relationships

$$g_S = h_S(T_S) \text{ and } g_N = h_N(T_N),$$

where  $h_S(\cdot)$  and  $h_N(\cdot)$  are continuously differentiable, increasing and concave functions.

The budget constraint of the individual takes the form

$$c + T_N + T_S \leq y. \tag{2}$$

Putting all these together, and taking his beliefs as given, the individual's maximization problem is

$$U(\pi_S, \pi_N) = \max_{\{c, T_i, a_i\}_{i \in \{S, N\}}} c + a_S \pi_S h_S(T_S) + a_N \pi_N h_N(T_N) - \Gamma_S(a_S) - \Gamma_N(a_N) \tag{3}$$

subject to (2). For future reference, this equation also defines  $U(\pi_S, \pi_N)$  as the expected utility level of the individual as a function of his beliefs.

Before we analyze the solution to this maximization problem and how it changes with information about the effectiveness of the state actor, we discuss how individuals will update their beliefs once they are given such information. We start with the benchmark Bayesian updating model.

## 2.2 Bayesian Updating

Consider how the provision of signal  $\sigma_S$  affects the individual's beliefs under Bayesian updating. Following this signal, the Bayesian posterior about the state actor, denoted by  $\pi_S^B$ , can be computed as

$$\pi_S^B = \frac{\Pr[\sigma_S|\theta_S = 1]\pi_S^0}{\Pr[\sigma_S|\theta_S = 1]\pi_S^0 + \Pr[\sigma_S|\theta_S = 0](1 - \pi_S^0)}, \quad (4)$$

where  $\Pr[\sigma_S|\theta_S]$  is the probability of receiving signal  $\sigma_S$  conditional on the value of  $\theta_S$ . Since the signal in our experimental treatment is designed to be good news about the quality of the state actor, we assume that  $\Pr[\sigma_S|\theta_S = 1] > \Pr[\sigma_S|\theta_S = 0]$ . The Bayesian updating formula, (4), then yields

$$\pi_S^B > \pi_S^0.$$

What about  $\pi_N$ ? In our experimental setup, the signal is designed to be only directly relevant about the quality of the state actor, with no information about the quality of the non-state actor. Mathematically, this would imply  $\Pr[\sigma_S|\theta_N = 1] = \Pr[\sigma_S|\theta_N = 0]$ . If this is indeed the case, the Bayesian posterior,  $\pi_N^B$ , about the non-state actor satisfies

$$\pi_N^B = \frac{\Pr[\sigma_S|\theta_N = 1]\pi_N^0}{\Pr[\sigma_S|\theta_N = 1]\pi_N^0 + \Pr[\sigma_S|\theta_N = 0](1 - \pi_N^0)} = \pi_N^0, \quad (5)$$

where the second equality follows immediately from the fact that  $\Pr[\sigma_S|\theta_N = 1] = \Pr[\sigma_S|\theta_N = 0]$ .

## 2.3 Basic Implications

We now study the implications of the signal  $\sigma_S$  on beliefs and behavior.

**Proposition 1** *Suppose that the individual is Bayesian. Then, the provision of signal  $\sigma_S$  (which is good news for  $\theta_S$  and not directly relevant for  $\theta_N$ ) weakly increases  $\pi_S$ ,  $T_S$ , and  $a_S$ , and has no effect on  $\pi_N$ ,  $T_N$ , and  $a_N$ .*

**Proof.** Recall that in this case,  $\pi_S^B > \pi_S^0$  and  $\pi_N^B = \pi_N^0$  as a result of Bayesian updating as in (5). Clearly, the budget constraint (2) will bind. Substituting from this constraint for  $c$ , (3) can be written as

$$\max_{\{T_i, a_i\}_{i \in \{S, N\}}} y - T_S - T_N + a_S \pi_S h_S(T_S) + a_N \pi_N h_N(T_N) - \Gamma_S(a_S) - \Gamma_N(a_N).$$

Now it is straightforward to see that this problem is supermodular in  $(T_S, a_S, \pi_S)$  and also separable between this vector and  $(T_N, a_N, \pi_N)$ . This implies that the optimal values of the control variables for the individual can be written as  $T_S^*(\pi_S)$ ,  $a_S^*(\pi_S)$ ,  $T_N^*(\pi_N)$ , and  $a_N^*(\pi_N)$ , where each one of these functions is (weakly) increasing. Since  $\pi_S$  increases and  $\pi_N$  remains constant, the claims in the

proposition follow immediately. In particular,  $T_S$  and  $a_S$  (weakly) increase and  $T_N$  and  $a_N$  do not change. Moreover, if  $T_S$  and  $a_S$  are interior, the changes are strict. ■

To provide further intuition, note first that the signal  $\sigma_S$  naturally increases  $\pi_S$ . These improved beliefs about the quality of state institutions make the state a better forum for the individual, who is then induced to increase  $T_S$  (unless this was at the corner solution of zero). This in turn increases the value of using the state forum and raising the relationship-specific effort (or allegiance) towards it,  $a_S$ . In this Bayesian baseline, given the nature of the signal  $\sigma_S$ , there is no impact on the beliefs concerning the non-state actor,  $\pi_N$ . In addition, given the quasi-linear preferences and the separable nature of the costs of allegiance towards the state and the non-state actors, contributions and allegiance to the non-state actor are unaffected.

Proposition 1 provides some intuitive basic implications concerning the direct effects of new information about the quality of state institutions, which will be borne out by our empirical results. But as emphasized in the Introduction, a critical part of our focus is on indirect effects, which are not in line with the predictions of this proposition. Anticipating what is to come, in our empirical work we find powerful negative indirect effects on beliefs concerning the non-state actor ( $\pi_N$ ), allocations in our two games (corresponding to  $T_N$ ), and on measures of allegiance ( $a_N$ ). Why might this be? One obvious possibility is that various assumptions underlying Proposition 1 do not hold in practice. We discuss four related sets of issues:

1. If utility from consumption were given by a strictly concave utility function  $u(c)$ , then increased contributions to the state would raise the marginal utility of consumption, encouraging a decline in contributions to the non-state actor,  $T_N$ , and via this channel a reduction in  $a_N$ . However, there would continue to be no impact on  $\pi_N$ . Moreover, as noted above, this mechanism should not apply to the allocations in the investment game because the amount allocated represents an investment made by the respondent based on what he believes is the expected effectiveness of the relevant actor.
2. If, instead of a separable cost function for the relationship-specific efforts, we had a joint, nonseparable cost function  $\Gamma(a_S, a_N)$ , where the two types of efforts are substitutes, increased effort towards the state forum would lead to a reduction in relationship-specific effort towards the non-state forum and consequently in  $T_N$ . But there would again be no impact on beliefs on the effectiveness of panchayats,  $\pi_N$ .
3. There may be greater substitution between using state courts and panchayats than posited in our model, and if so, greater use of state courts could directly translate into lower use of panchayats. Though this may affect our question regarding perceived usage of the non-state actor, it should not impact game allocations or beliefs concerning the non-state actor,  $\pi_N$ .

4. Even though the information we provided was not directly informative about  $\theta_N$ , it may have been perceived this way by our subjects. This might be the case, for example, if they believed that  $\theta_S$  and  $\theta_N$  were negatively correlated. If there were such an effect on  $\pi_N$ , this would naturally lead to lower  $T_N$  and  $a_N$ . Though it is important to note this possibility, our fieldwork revealed no apparent reasons to suspect that this was important in practice. During our pilots, we ensured that the information treatment was worded carefully to prevent such a misunderstanding. Subsequent interviews with respondents indicated that they understood that the two forums were separate and that they did not equate improvements in one with changes in the other.

Instead of these possibilities, we argue that a model of “motivated reasoning,” whereby individuals manipulate their own beliefs to be in line with their actions provides a better match to the ideas discussed in the Introduction and to our setting, and also crucially implies a change in beliefs concerning the non-state actor — a pattern we document in the data. We next present such a model.

## 2.4 Motivated Reasoning

We now present a simple variation on the utility function in (1) that incorporates motivated reasoning. In our setting, motivated reasoning is relevant in part because conditional on choosing a particular actor as their service provider (and having invested in relationship-specific effort for the actor), individuals may have greater reason to convince themselves that this actor is providing high-quality services (and perhaps that it is honest, well-meaning, and effective). More formally, under motivated reasoning, individuals choose not only their behavior but also their beliefs, and will do so to make their beliefs more congruent with their behavior. A simple way of modeling motivated reasoning is therefore to allow a simultaneous choice over behavior and beliefs with a penalty for deviations of these beliefs from the Bayesian benchmark as in the next optimization problem:

$$\max_{\{c, T_i, a_i, \pi_i\}_{i \in \{S, N\}}} \quad c + a_S \pi_S h_S(T_S) + a_N \pi_N h_N(T_N) - \Gamma_S(a_S) - \Gamma_N(a_N) - d[U(\pi_S, \pi_N) - U(\pi_S^B, \pi_N^B)], \quad (6)$$

again subject to (2). Note that compared to (3) there are now two important differences. First, there is an explicit choice over beliefs,  $\pi_S$  and  $\pi_N$ . Because these beliefs multiply  $a_S h_S(T_S)$  and  $a_N h_N(T_N)$ , this choice creates a force towards beliefs that are more favorable towards the actor that the individual is using and investing in. Second, the second line introduces a penalty for the deviation of these beliefs from their Bayesian counterparts,  $\pi_S^B$  and  $\pi_N^B$ . In particular, here  $d[\cdot]$  is a convex function that is increasing when its argument is positive and decreasing when it is negative, and we also assume that it is differentiable with  $d'[0] = 0$ . This penalty is in terms of the difference between the (actual) utility level  $U(\pi_S, \pi_N)$  as defined in (3) when the individual’s

beliefs are  $\pi_S$  and  $\pi_N$ , and his utility under Bayesian beliefs is  $U(\pi_S^B, \pi_N^B)$ , which is the maximum utility he can achieve. Convexity implies that the penalty for further deviations is greater when  $U(\pi_S, \pi_N)$  is further apart from  $U(\pi_S^B, \pi_N^B)$ .<sup>7</sup> Overall, this objective function captures in a simple way the trade-off between the direct utility of congruence between beliefs and actions versus the cost of manipulating beliefs in terms of departures from the best ex-post decisions (which would give utility  $U(\pi_S^B, \pi_N^B)$ ).

The following proposition describes the results from our simple conceptual framework in the presence of motivated reasoning.

**Proposition 2** *Suppose that the individual engages in motivated reasoning. Then the provision of information  $\sigma_S$  (which is good news for  $\theta_S$  and not directly relevant for  $\theta_N$ ), (weakly) increases  $\pi_S$ ,  $T_S$ , and  $a_S$ , and (weakly) decreases  $\pi_N$ ,  $T_N$ , and  $a_N$ .*

**Proof.** Consider the same maximization problem as in the proof of Proposition 1, which defines

$$U(\pi_S, \pi_N) = y - T_S^*(\pi_S) - T_N^*(\pi_N) - a_S^*(\pi_S)\pi_S h_S(T_S^*(\pi_S)) \\ - a_N^*(\pi_N)\pi_N h_N(T_N^*(\pi_N)) - \Gamma_S(a_S^*(\pi_S)) - \Gamma_N(a_N^*(\pi_N)).$$

Now substituting this into (6), we obtain an objective function that is supermodular in  $(T_S, a_S, \pi_S, \pi_N^B, -T_N, -a_N, -\pi_N)$ . Consequently, the signal  $\sigma_S$  that increases  $\pi_S^B$  (and does not affect  $\pi_N^B$ ) will lead to a (weak) increase in  $T_S, a_S$ , and  $\pi_S$  and a (weak) decrease in  $T_N, a_N$ , and  $\pi_N$ , establishing the desired result. All changes are strict when these variables are interior. ■

As before, the positive information  $\sigma_S$  about the state actor increases the Bayesian benchmark beliefs about this actor,  $\pi_S^B$ . Because of the convex penalty term,  $\sigma_S$  also induces an increase in the beliefs that the individual holds about this actor,  $\pi_S$ . Given these improved beliefs, the individual raises his contribution to the state actor,  $T_S$  (and his relationship-specific effort  $a_S$ ). In turn, these higher levels of  $T_S$  and  $a_S$  induce a further increase in beliefs concerning the quality of the state actor,  $\pi_S$ , amplifying these qualitative effects.

In addition, and crucially for our purposes, a higher  $\pi_S$  for given  $\pi_S^B$  and  $\pi_N^B$  now increases the deviation between  $U(\pi_S, \pi_N)$  and  $U(\pi_S^B, \pi_N^B)$ , because the deviation of actual payoff from its maximum value (realized under Bayesian beliefs) increases. From the convexity of the penalty function  $d[\cdot]$ , this increases the marginal cost of motivated reasoning, which also applies to beliefs concerning the other actor. Consequently,  $\pi_S$  and  $\pi_N$  become substitutes, and any change that induces an individual to increase  $\pi_S$  also triggers a decline in  $\pi_N$ , which in turn leads to lower contributions to the non-state actor,  $T_N$ , and lower relationship-specific effort,  $a_N$ . Intuitively, the convex penalty for the deviation of beliefs from their Bayesian counterparts implies that incentives

<sup>7</sup>Without this penalty term, the individual would have an incentive to choose the highest possible values of  $\pi_S$  and  $\pi_N$ .

to improve beliefs about the state also create incentives to make beliefs about the non-state actor more realistic (and less positive).<sup>8</sup>

Therefore, differently from Proposition 1, Proposition 2 introduces a natural reason for why beliefs about the non-state actor will deteriorate in response to positive news about the state actor, and consequently, contributions and allegiance to this actor will also decline. In our empirical work, we investigate the implications summarized in Propositions 1 and 2.

### 3 Background and Context

#### 3.1 Dispute Resolution, Courts, and *Panchayats* in Pakistan

Dispute resolution is one of the most important services demanded by Pakistani citizens and one of the Pakistani state’s core responsibilities.<sup>9</sup> Disputes are a particularly common occurrence in our setting, Pakistani Punjab, as manifested by high litigation rates — about three times as high as the rates of litigation in Indian parts of colonial Punjab.<sup>10</sup> Consistent with this, in our survey one in every five households reports that they have used the court system in the last three months. This is especially noteworthy given, as we note below, that courts are not the preferred choice. Recent studies show that a majority of cases that end up in courts involve disputes around land, property, inheritance, and contract (Siddique (2013)),<sup>11</sup> which suggests that many disputes arise due to weak enforcement of property rights or contracts. Becoming embroiled in a dispute imposes significant costs. Estimates suggest that cases take many years to resolve and involve sizable financial costs (Chemin (2009), Siddique (2016)).

Popular media is also full of accounts of the miscarriage of justice and the high cost of accessing court-based justice, especially for the poor and those in rural areas (see Javed (2017), Nekokara (2016), Shinwari (2015)). As previously mentioned, the rallying slogan used by the Taliban insurgency in KP province was the provision of cheap and swift justice (Khan et al., 2000). The third most popular political party in the country, the Pakistan Tehreek-e-Insaaf, or Movement for Justice Party, has made access to justice the pillar of its platform.

Dispute resolution therefore offers an ideal setting for our study. We further narrow our focus to

---

<sup>8</sup>Note, in particular, that this formulation implies that the individual will always hold more positive beliefs about the actors he is interacting with, so the downgrading of his beliefs about the non-state actor will make them closer to their Bayesian counterpart.

<sup>9</sup>In *Sh. Liaquat Hussain v. Federation of Pakistan* (PLD 1999 SC 504), the Supreme Court of Pakistan held that “(t)he right of access to justice to all is enshrined in the Constitution, which is also found in the doctrine of ‘due process of law’. Such a right includes the right to be treated according to law; the right to have a fair and proper trial and the right to have an impartial Court or Tribunal.”

<sup>10</sup>The partition of British India split the former province of colonial Punjab into the Punjab province in Pakistan and the states of Punjab, Haryana and Himachal Pradesh in India. The litigation rates in these Indian states ranged between 5.3 and 9.2 per 1000 persons between 2005 and 2010 (Eisenberg et al. (2013)) compared to 17 per 1000 in Pakistani Punjab during the same period (authors’ estimates are based on the Lahore High Court Annual Reports).

<sup>11</sup>In a survey of litigants in Lahore, Siddique (2013) finds that approximately 57.5% of court cases involved land, property and inheritance disputes, 18% involved marital or guardianship cases, and around 8% were contract disputes.

rural areas, where there is not only a salient demand for dispute resolution but also clearly defined state and non-state actors that provide competing dispute resolution services.

On the state’s side, the judicial system operated by the Pakistani state consists of courts backed by the police.<sup>12</sup> It is an adversarial and retributive judicial system that is divided into courts of first instance (both civil and criminal) and appellate courts, which have the power to review the decisions of the lower courts. The state primarily works through three key actors — police, judges, and lawyers. The police are responsible for the maintenance of law and for the administration of criminal justice, making them the typical first point of contact for citizens in criminal matters. Judges, with special legal training, adjudicate on the basis of codified procedures and consistent application of state law. Lawyers are meant to assist the courts in reaching just decisions.

Non-state actors have historically run parallel dispute resolution forums in rural areas of Pakistan that are distinct from the state judicial system.<sup>13</sup> These non-state forums are (typically) ad hoc local councils of respectable villagers called *panchayats*,<sup>14</sup> and are usually given the authority to resolve disputes on behalf of residents of the community (Chaudhary (1999), Soomro and Chandio (2013), Ayaz and Fleschenberg (2009)). 80% of villages in our sample report the presence of such a system in their community, dealing with a wide array of cases including theft, robbery, family feuds, small social complaints, and land disputes. Dispute resolution in panchayats relies on local customary law and practices rather than on Pakistan’s legal code. They are run exclusively by influential villagers, such as leaders of clans and tribes, family elders, and powerful members of the community, who have neither training in legal matters nor any practice in following the law. Members of panchayats are invariably male and rarely include non-Muslim minorities. While the state justice system is relatively punitive, panchayat decisions tend to be restorative. They use a combination of mediation, compromise, and penalties, including social ostracism, boycott, and sometimes even physical retaliation. Enforcement of panchayat decisions is typically underpinned by the threat of sanctions by the community or powerful members of the community (Shinwari (2015), Siddique (2015), Chaudhary (1999)).

---

<sup>12</sup>In colonial India, officers of the executive (as opposed to the judicial) branch were invested with specific judicial powers under the criminal procedure and penal codes. Whereas this system has continued in post-independence India, it was abolished in Pakistan as a result of the Devolution Reforms of 2001 to achieve separation of powers between the judiciary and the executive (Article 175(3) of the Constitution of Pakistan). A consequence has been a significant expansion of the remit of the courts in Pakistan’s justice system.

<sup>13</sup>Hoebel (1965) observed more than three decades ago that, “the legal system of Pakistan does not constitute a neatly integrated whole; it is made up of an undetermined multiplicity of subsystems. Deeply embedded in the village and tribal areas of Pakistan is a vast array of local folk systems of law varying from village to village” (Chaudhary (1999)). Even today non-state actors play a critical role in dispute resolution in much of Pakistan (Chaudhary (1999), Siddique (2015)).

<sup>14</sup>In other areas of Pakistan, panchayats are also called *kath*, *paryah*, *faislo*, or *jirga* (Chaudhary (1999), Shinwari (2015)). The panchayat system is not new to the Indian subcontinent, and it remains fairly prevalent in both India and Pakistan. There are references to it in the Sanskrit epic of 8th and 9th centuries B.C.E., the *Mahabharata*, and it also appears to have continued through the period of Muslim rule. This is in contrast to the state judicial system, which is a product of British colonial rule (Siddique, 2015).

This discussion clarifies that the panchayats, which ignore the formal law and compete with the state, clearly correspond to a non-state actor. Since independence in 1947, Pakistan’s state judicial system has indeed been highly suspicious of such parallel non-state forums, viewing them as antithetical to the state system. This is in sharp contrast to India, which has tried to incorporate panchayats into the formal state apparatus. In fact, the report of the Pakistan Government’s Law Reform Commission 1967-70 argued that “it will be a retrograde step to revert to the primitive method of administration of justice by taking our disputes to a group of ordinary laymen ignorant of modern complexities of life and not conversant with legal concepts and procedures” (Chaudhary (1999)). In 2004, the Sindh High Court banned trials under the non-state system and declared these forums illegal (Cawasjee (2004), Brohi (2016)). Similarly, the Supreme Court of Pakistan has made a series of rulings during the past decade decreeing many panchayat decisions to be unconstitutional (Brohi (2016)). It has specifically targeted panchayat decisions that sanction direct vengeance for murders and forced marriages of young girls as punishment for crimes committed by their male relatives (Shinwari (2015), Siddique (2015)).

### **3.2 Access Challenges and Mistrust towards the State**

Since our primary information treatment is to provide positive information about the state dispute resolution actor (courts), it is important to understand what the baseline views towards courts are, as well as whether citizens are well informed about service changes (improvements).

What is noteworthy at the outset is that citizens generally have a poor view of the state judicial system, with courts generally rated as unreliable and unfair. Siddique (2013) finds that three quarters of respondents in a survey of litigants were dissatisfied or deeply dissatisfied with the pace at which their case was proceeding and about the same proportion could not predict when a verdict would arrive. Respondents report how “For 20 years have I been waiting for justice. Judges and lawyers ensure that case does not come to a conclusion,” “(M)y family has withered away while pursuing this matter,” and they bemoan how “(T)his legal system is a complete failure.” These assessments are in sharp contrast to views on panchayats which, despite their tendency to go against the writ of the state and their complete failure to incorporate disenfranchised members of society such as women and low-income groups, are preferred over the police and state courts. This is a pattern that is clearly confirmed in our data. Respondents in our baseline surveys report that their likelihood of using a court to settle a dispute is only 38%, while the likelihood of using a panchayat is 67%. This preference for panchayats appears to be a reflection of the difficulties in accessing and using state courts and a general mistrust towards them.

As suggested by the above testimonials, a common refrain about state courts is about endemic delays. Over 80% of respondents in Siddique’s (2013) sample felt that there were significant delays when going through courts, while 27% of litigants had their case stuck in the court system for more



than five years. The majority did not know when a verdict was expected. As noted previously, a major factor in the rise of the Taliban was their promise of speedy dispute resolution; decisions would be made (almost) on the spot in the Sharia courts. This desire for faster decisions is unsurprising as Siddique (2016) and Chemin (2009) report that cases in state courts take on average two or three years. Shinwari (2015) also finds that the lack of speedy justice is one of the biggest complaints made against the courts by over three quarters of the respondents in his nationally representative survey. Consistent with this, Chemin (2009) estimated that more than 1.2 million cases were pending in 2001, and recent estimates suggest that this number may have climbed up to 2 million (Siddiqi (2016)).

Another recurrent theme in complaints is cost. Since resolving disputes through state courts is a lengthy process, an individual needs to consider the loss of daily wages, cost of transportation, and the legal fees necessary to enter and remain in the system.<sup>15</sup> While panchayats are located within most villages, courts are fewer in number and are usually situated in the main district city.<sup>16</sup> With a bare-bones infrastructure, meager public transport, and additional costs of traveling long distances to and from courts, the idea of approaching the state for assistance does not even occur to many rural residents. For villagers living in more remote areas, access to the local district courts is difficult, if not impossible. In contrast, panchayats offer quick resolution by gathering the disputing parties directly in the village. Their proximity allows each party to bring their supporters to the meetings with limited expenses. In a culture where showing strength in numbers is pervasive, the opportunity to be escorted by your well-wishers to a dispute is favored to trudging to a distant court alone. This access issue is further compounded by the fact that news of reforms in the state system, such as reductions in the number of pending cases, may not have reached many villagers at all. Thus, despite being dominated by powerful individuals known to make rulings that depend on political, social, religious, and cast affiliations, panchayats are often preferred over state courts due to citizens having a better understanding of how the system works in their respective communities (Shinwari, 2015).

Mistrust towards state courts further exacerbates these tendencies. This mistrust is partly rooted in the systemic lack of access to legal knowledge and effective legal counsel. According to Siddique (2015), the likelihood of receiving incorrect legal assistance from poorly trained lawyers is higher for uneducated, low-income households. In addition, most non-profit organizations working to provide

---

<sup>15</sup>Based on interviews with lawyers at the Sessions Courts, we found that different types of cases vary in length and cost. For example, an inheritance case could last more than two years on average with anywhere between 5,000 PKR and 200,000 PKR in costs. The resolution of business related cases may be faster, though even those can take upwards of six months with potentially significant costs for the parties (Siddique (2013)). Shinwari (2015) finds that low income households and women particularly face high costs of accessing formal justice institutions, in part, due to the process being time consuming, the high lawyer fees, and the long distances to court.

<sup>16</sup>Unlike India, which has taken a number of measures to strengthen dispute resolution bodies located in villages and working with state institutions (e.g., Zainulbhai (2016)), Pakistan has failed to institutionalize formal mediation at the village level even though they have been enacted under its local government laws. This means that village litigants have to travel to district headquarter towns and cities to access dispute resolution services from the state.

legal assistance to marginalized communities and low-income groups do not have the resources to reach individuals in rural areas. This is true across the country, including in the population sampled in this study where respondents quoted their inability to speak with anyone who could provide them with credible legal advice. In cases where respondents approached state courts, their typical means of legal counsel was from other laymen who had experience in the system themselves but did not possess a legal education or complete information on how to navigate the judicial system. The cost of acquiring legal counsel thus contributes both to mistrust and to the practical difficulties of accessing state courts.

For those who are able to surpass the hurdle of accessing legal assistance — either in the form of aid, advice, or practical representation — a common refrain is the paucity of justice once one enters the system. Our own survey confirms the high level of corruption that our sample population felt exists within both the police and the court system. Respondents were quick to add that justice can only be found through deep pockets, and that only people (unlike themselves) who are capable of lining the pockets of those who matter are able to obtain a verdict in their cases.<sup>17</sup> It is, therefore, not surprising that nearly 90% of our respondents who had accessed the police or the state justice system in the three months preceding our survey felt that the police cannot be trusted, and that another 65.7% thought that the courts are not trustworthy. These misgivings are shared with the community as anecdotes and warnings that inform the views of those members who have yet to face any type of dispute. The experiences of their neighbors or relatives appear to make them think twice about reaching out to state courts.

With these pervasive difficulties of access, persistent delays, and perceived unfairness of state courts, rural Pakistanis' views toward the state judicial system for dispute resolution is understandable. Because disputes are very common in Pakistani Punjab, many of our subjects still use state courts, but their often negative experiences reinforce their poor perceptions of the state. While ultimately these attitudes can only change if there are significant improvements in the performance of state courts on a multitude of fronts, the remoteness of state courts implies that rural Pakistanis may plausibly be unaware of recent changes. Therefore, credible information provision concerning recent developments regarding improved access and effectiveness for courts might still have an impact on perceptions, as we will see.

## 4 Experimental Design and Estimation

Understanding the determinants of individual decisions to access state or non-state actors for dispute resolution is challenging given how many factors and conditions can impact such behavior. In order to overcome these challenges, we utilize a within-subject experimental study design whereby we

---

<sup>17</sup>Siddique (2013) finds that 47% of the respondents in Lahore felt that the laws are either biased against them or unjust, and 34% of those surveyed were dissatisfied with the competence of the judicial officer.

examine how the behavior of individuals changes in response to information they receive. We expose individuals to information on state effectiveness and then see how their views and behavior change towards state versus non-state actors. While it would have been preferable to observe changes in terms of the forums people use in a dispute that they are actually experiencing, identifying populations that are experiencing a dispute at an early enough stage that their actions may be impacted by information they receive is nontrivial. While we are exploring such an approach in subsequent work, in this paper we adopt a more straightforward strategy. We draw from a random sample of (rural) individuals and conduct initial surveys to capture their baseline beliefs and behavior. We then (experimentally) expose these individuals to different types of information and repeat these surveys to see what impact this information has on their behavior and beliefs. To help address concerns that survey responses may not reflect real behavior, our main focus is on experimental games with real stakes that reveal respondents' behavior towards both non-state and state actors. In this section we detail the information treatments, the data and sampling, experimental protocols, and estimation strategy.

#### 4.1 Information Treatments

We are interested in whether perceptions of state (relative) effectiveness can change attitudes and behavior towards state and non-state actors. Given the generally poor views of state actors and the likelihood that the average citizen may not be fully informed, our primary information treatment is therefore one that provides true, positive (favorable) evidence on the performance of the formal justice system. We refer to this as the “**state positive treatment (SP)**.” Based on our field discussions where a variety of information primes were discussed, and given that many of our respondents felt that they would not get effective justice because of the pervasive delays in the court system, we chose to focus on information that signals a change in these delays. This is also desirable because, regardless of the legitimacy or beliefs concerning judicial biases, a reduction in delays is likely to be widely attractive. This treatment provides respondents the following information about a district (Multan) in a nearby region to our study sample district (Sargodha): *“The legal system and judges have formed a new judicial policy. This policy was introduced in Multan and has resolved 6000 pending cases in two months. For this reason, Multan’s number of pending cases has decreased by 20%. This policy has now been implemented in Sargodha and it is estimated that most pending cases could potentially be resolved within a year.”*

Our initial design included other variations, in particular, state negative, non-state positive and non-state negative treatments. However, our pilots revealed that the impact of these different treatments was fairly symmetric, so for the sake of statistical power we decided to only scale up the state positive version. We should note that while the state negative treatment also primed on decision-making delays, the non-state positive (negative) primes included information about the

(in)effectiveness of the decisions made by the panchayats as well. This suggests that the informational impacts we observed are not just about changes in delays, but that respondents react analogously when informed about more (less) effective resolution of disputes as well. The Appendices provide the details of these treatments and the results.

One potential drawback of the within-subject design is that respondents may change their views and behaviors after the informational treatment for other reasons. Most saliently, they may feel obligated to do so given what the experimenter has just said, despite no real change in their view. A direct way to deal with this “social experimenter effect” is by using a treatment that directly primes it. In order to do so we provided the following **social experimenter treatment (SET)** statement from the surveyors to a randomly selected group of respondents (again, after the baseline surveys and games): *“So I’ve been thinking about the current state of affairs and how the state’s been dealing with everything and while I don’t really know how great a job state institutions are doing, in my personal opinion, I really like the state system.”* We then repeated the surveys and games after this treatment. Using this sample, we can “net out” any potential social experimenter effects. We also directly use a pure cross-subject design as an alternative strategy as described below.<sup>18</sup>

## 4.2 Sample and Data

The survey respondents are rural male household heads between the ages of 20 and 64 years and spanning four districts in Punjab (Pakistan’s most populous state) and over 400 households. The results of the pilot indicated that subjects from this gender and age bracket are most likely to be decision-makers within a household and are best suited to understand and relate to the contextual framework of the survey and the behavioral experiments. Out of the total number of pilot surveys, close to 100 were conducted with women. However, we found that women did not have as much direct experience with the state or non-state forums as men had, making them less suitable as respondents for our purposes. Moreover, no obvious patterns of heterogeneity emerged, and this, combined with statistical sampling/power and logistical constraints, led us to limit our sample to males. After pilots within urban, peri-urban, and rural settings, we also decided to conduct the experiments in rural areas only, because engagement with non-state actors was more easily identifiable within this setting. While urban areas did have non-state actors, the specific actor varied considerably. In contrast, in rural areas, the panchayat was invariably identified as the primary non-state actor. Moreover, based on the 1998 Population Census, a majority (68.7%) of Pakistan’s population still resides in rural villages (Pakistan Bureau of Statistics). Our primary sample consists of randomly

---

<sup>18</sup>While both the state positive and social experimenter treatments are randomly assigned across individuals, the fraction assigned to either treatment varies across six sampling strata, because after the first few surveys we recalibrated the sample sizes of these two treatments. To avoid any concerns related to “p-hacking”, sample sizes were adjusted using information only on standard errors of outcomes of interest, and not based on estimated effect sizes, p-values or t-statistics. Our within-subject design is unaffected by this recalibration, and in any specification that involves cross-subject comparisons we always include strata fixed effects to capture any baseline differences across strata.

drawn rural households in one district (Sargodha) of Punjab. The district is fairly representative of Punjab and was drawn according to a pre-existing sampling frame (the 2007 Sargodha Village and Household Survey, SVHS) from a primary survey designed to be representative at the district level and conducted by Cheema and Naseer (2013). We carried out 2,986 surveys of male household heads from 37 villages drawn from the previous study sample.

In order to measure potential heterogeneous treatment effects by caste, we stratified our sample population (in more than half of the villages) by identifying neighborhoods with high, low, and middle tier households through a pre-existing definition of caste from SVHS.<sup>19</sup> While the strata size varied somewhat across the villages, our final sample consists of roughly equal shares of all three castes (35% high-caste, 34% middle-caste, and 31% low-caste respondents). Survey participants were recruited through door-to-door household visits in a pre-survey location round. The identified and available households were then randomly assigned to primary and replacement lists to meet a pre-specified target for each village. We carried out 45 surveys in each village (or in each settlement if a village consisted of more than one settlement). This number is based on the number and speed of enumerators employed, coupled with our decision not to return to a village or settlement for a second day of surveying, in order to minimize the risk that information about how our games are conducted would spread in the village, thereby contaminating our sample.

Table 1 presents demographic summary statistics for the 2,986 respondents in our sample, as well as the baseline values for our main outcome variables.<sup>20</sup> Recall that our primary sample includes only men, and from Panel A we see that they have an average age of 37 and 72% have some formal education (though less than 10% have a high school equivalent degree). While 28% own land and average monthly household income is 17,007 PKR (approximately \$165), there is also considerable heterogeneity in land ownership and income. Though there is reasonable usage of state services, as mentioned previously, our subjects have a significantly higher expected usage of the non-state actor (panchayat). Interestingly, despite this difference in expected usage, baseline allocations to the state actor in the two games are only a little lower than those to the non-state actor.

### 4.3 Experimental Procedures and Games

Our primary protocol is to approach a randomly selected household and seek their consent to have a discussion around issues regarding dispute resolution. We introduce ourselves as researchers interested in understanding the different forums for dispute resolution in their community, solicit their views, and play some experimental games. Participants are informed that they will receive a token 50 PKR payment for agreeing to participate, plus they have the potential of earning significant

---

<sup>19</sup>Caste, or quom, as it is referred to in Cheema and Naseer (2013), is defined as a social group based on patrilineal descent. Castes are further distinguished into high or low depending on colonial assignment of ownership of land.

<sup>20</sup>Observation counts vary because of non-response, especially for the income questions, and over-time changes in the survey.

payoffs from the games.

After receiving consent, respondents are asked a series of questions regarding dispute resolution and their views on the effectiveness of both state courts and panchayats. The survey instrument was designed to understand the actions of respondents with regard to effectiveness of state and non-state actors. The survey includes a question on the expected usage of the state and non-state forums, which we use throughout the paper. In later stages of our fieldwork we also included questions on individual beliefs regarding the effectiveness of the relevant forum as well as general trust and allegiance towards the forum.<sup>21</sup> In addition to information on our subjects' own expected usage of state and non-state forums, we also gathered information about their expectations of others' usage (all of these variables are on a scale from 0 to 10).<sup>22</sup> By comparing responses to these questions at baseline and post-treatment, we can measure the change in a subject's own expected usage and their perceptions of others' usage of state and non-state forums resulting from our informational and social experimenter treatments.

In addition to the baseline surveys, respondents play two different games — the fund dictator and the investment game — before and after the information treatment. The two games are designed in order to capture different aspects of citizens' views of state and non-state actors. The fund dictator game is meant to measure a change in beliefs and related behavior concerning how beneficial state forums are to the general population of the subject's locality, while the investment game is aimed at measuring a change in beliefs and related behavior concerning how effective the two forums are in resolving disputes.

The fund dictator game is set up along the lines of a standard dictator game, where we seek to understand respondents' proclivity to assist those in their communities in accessing state or non-state dispute resolution forums. We do so by asking individuals to contribute to two potential funds in their home area; one (state fund) to assist those individuals in the community who prefer to go through the state system in order to resolve disputes, and the other (non-state fund) for those who would rather go to the non-state alternative. Respondents can confidentially make a choice to give all, some, or none of a specified amount to the fund in question based on their beliefs about the relevant actor's benefit to citizens and the respondent's level of trust in the chosen actor. The surveyor also explains that such a fund is anticipated to be set up by the research organization so the

---

<sup>21</sup>There was initially a concern that including a detailed set of questions regarding effectiveness of state and non-state actors at baseline could have its own priming effect and confound our interpretation of state positive and social experimenter informational treatments. We included these richer set of questions (pre- and post- the information provision) in subsequent samples in order to shed further light on the mechanisms at play. Moreover, given our budget and power calculations, we could only provide the state positive informational treatment and not the social experimenter treatment to the sample where we asked these additional belief questions.

<sup>22</sup>In particular, we ask about self engagement: "*How likely are you to use the state or panchayat, 0 meaning you would not go to the forum if you had a dispute and 10 meaning you would definitely go to the forum if you had a dispute?*" Regarding perceptions of other villagers' engagement, we ask: "*If 10 people have different types of dispute (of the kind listed above), how many of them are likely to go to the courts to resolve their disputes, 0 meaning no one would go to courts and 10 meaning, all 10 people will go to the courts?*"

money that the respondent allocates to that fund will actually be donated.<sup>23</sup> We therefore expect the respondent to allocate more money to the fund he believes will be more useful and/or to the one he feels more altruistic towards. In order to avoid any mechanical spillover effects, participants receive two separate endowments of 250 PKR, which they can allocate to a specific fund or keep for themselves. Thus, in the baseline play, for example, an individual may decide to donate 100 PKR to the state fund and keep 150 PKR for himself out of the first endowment, while donating 150 PKR to the non-state fund and keeping 100 PKR for himself out of the second endowment.<sup>24</sup>

The investment game, while less standard, is designed to specifically have the participants focus on which forum is most effective in getting (the appropriate) results, and subsequently, a better return on their own investments. Respondents are told to consider two hypothetical members of their community, each of whom is experiencing a dispute, but one member has chosen to take his case through state institutions and the other to the panchayat. Both members are rightfully owed remunerations from a defendant, and the respondent is given a chance to invest an amount  $X$  out of 250 PKR in the plaintiff’s case in return for a share of the remuneration. They are told the share they receive will be  $\eta X$ , where  $\eta \in [0, 2]$  measures the effectiveness of the forum. If a forum is completely ineffective, then  $\eta = 0$  and the plaintiff will receive no remunerations; if a forum is fully effective, then  $\eta = 2$  and the plaintiff will receive full remuneration. The respondents are not given a specific value for  $\eta$ , but they are informed that it has been calibrated for each forum to reflect reality.<sup>25</sup> Hence, the more effective the respondent believes a dispute resolution forum to be, the more he will invest in that forum. While the specifics of the game may have been somewhat unusual, the context is not, since villagers are familiar with investing in each other’s projects or lending money to help each other out where the return/likelihood of repayment is contingent on the success of the endeavor.<sup>26</sup>

---

<sup>23</sup>We have a total of 157,000 PKR allocated to the non-state fund and 145,200 PKR for the state fund. We are working towards setting up these funds in the context of our follow-up study focusing on individuals currently experiencing disputes.

<sup>24</sup>These contributions will clearly depend on a respondent’s overall degree of altruism, but our within-subject design removes this dependence.

<sup>25</sup>After piloting different options, we chose to keep the return on investment from both the state and the non-state forum consistent at a factor of one, ensuring the respondent would receive a payment (the exact amount he allocates in one of the four rounds in each game, i.e., 250 PKR) whilst remaining within our budget. We could have varied the return on invested allocation between the two forums but since respondents do not discover the return on investment until both experiments are completed and only find out the return on investment on one forum depending on which round is randomly chosen for payment, it was deemed simpler to not vary the return.

<sup>26</sup>We arrived at the game designs described above through multiple iterations of pilots with alternative framings. The goal was to ensure that these games were both well understood and tailored to the context so that they would appear familiar to our respondents, especially since they were unlikely to have ever experienced such lab-in-field games before. For example, our initial framing of the fund dictator game involved the respondents allocating money into a “voucher” which would be used to help access the relevant forum. However, interviews revealed several issues with this framing and highlighted the non-applicability of a voucher system to the panchayat. Similarly, our initial variant of the investment game had the respondents take on the role of the state/non-state actor and allowed them to decide how they believed each actor would allocate remunerations; however, follow-up interviews revealed the role-playing aspect of this game confused many respondents. We additionally piloted a version where the respondents were given two envelopes in which they were told the payout from each forum had already been placed, and they were then asked

Once the basic design of the games had been finalized, we ran additional pilots to ensure that the language and format details were all easily comprehended by our respondents. Our original survey was created in Urdu (the national language), but initial piloting revealed that the nuances of the games were best understood in Punjabi (the local vernacular), prompting us to present the information in Punjabi. Similarly, we varied the sequence in which respondents played the games to see if their understanding differed depending on which game was played first. We found no such effects, but decided to keep the order randomization to account for any potential level or treatment effects induced by order. The game payoff amounts were also piloted in order to arrive at an amount that was large enough to create credible stakes without being too costly. Finally, we ensured that the wording was such that respondents' game allocations followed their own beliefs rather than other potentially salient factors, such as what they may have thought the surveyor/research team wanted. For a few respondents we directly asked about their thought process in making their allocations and found little evidence of any such concerns.

We also took several steps to ensure high quality responses in the games. Our first strategy was to give each respondent 50 PKR at the start as a participation fee, building credibility with respect to our intention to pay out their winnings in cash. Respondents then played both the fund and investment game three times. First, they played a practice round of each, followed by a discussion to ensure they had understood the game procedures. A customized board was used as a visual tool during the explanations, which also provided sub-divisions between the baseline and post-treatment allocations for each experiment. The benefit of such a tool is three-fold: i) it provides a visual aid for respondents, ii) it creates a sense of privacy (each section has a cover that hides the allocations of the respondent from the enumerator), and iii) it provides demarcation between the two baseline rounds and the two post-treatment rounds for each experiment. Each section is further subdivided to depict the respondent's allocations for state/self and non-state/self. The board is shown in Figure A1 in Appendix A. After the enumerator was sure the respondent had understood the game, a baseline round was played. Next, the respondents were provided with one of our information treatments. Finally, they were asked to (confidentially) review their baseline allocations and play an end-line game. For both games participants were paid for only one of the four allocation decisions (baseline state, baseline non-state, end-line state, or end-line non-state allocation) made under each game type. For the fund dictator game, the payoff was simply one of the four amounts the participant selected to keep for himself. In the investment game, the respondent received (for one of the four choices) both the money allocated to himself and a return on the investment made in the given dispute resolution forum, that is,  $(250 - X) + \eta X$ .

---

how much they would be willing to pay in order to receive each envelope. Again, this was found to be confusing, and the more natural and familiar framing of investing in a community member's case was chosen instead.



## 5 Results

In this section, we present our main results. As detailed in the previous section, our main treatment is to provide our subjects with (objective, media-reported) information about the improved performance of state courts. We then measure how this affects their (self-reported) expected usage of courts, and their behavior in the two games described above; the fund dictator game, which is designed to measure their willingness to contribute to the state or non-state institutions, and the investment game, which is designed to measure their assessment of the effectiveness of these two institutions. We first present our results concerning the effects of our informational treatment on beliefs and behavior towards state courts. We then turn to spillover effects toward non-state institutions, and proceed to present a range of robustness results. We conclude with an investigation of potential channels by looking at the effects of these treatments on beliefs and allegiance to the different actors.

### 5.1 Direct Effects

Our primary informational treatment, described above, provides positive information about the effectiveness of state courts. Recall from Table 1 that most individuals start with a low perception of the state relative to non-state (panchayat) institutions. We therefore expect this information to potentially change people’s expectations of future usage and behavior in our games.

Our first results are from the within-subject design already described in the previous section and are presented in Panel A of Table 2. We estimate

$$Y_{it} = \alpha + \beta \text{Post}_t + \delta_i + \varepsilon_{it}, \tag{7}$$

where  $Y_{it}$  is our outcome variable of interest (expected usage or contributions in the two games) and  $\text{Post}_t$  is a dummy designating the second stage of our experiment, after the subject has been given the informational treatment. The parameter  $\beta$  is our coefficient of interest, and measures the within-subject effect of the informational treatment — how much a given person changes their reported perceptions and behavior following the new information. In addition,  $\varepsilon_{it}$  is an error term capturing all omitted influences, and  $\delta_i$  is a person fixed effect. The inclusion of these fixed effects has no impact on the estimates of  $\beta$ , but affects the standard errors.

Panel A of Table 2 reports estimates of equation (7) for our three main outcome variables, expected usage, allocation to the state in the dictator fund game, and allocation to the state in the investment game. We see uniformly positive and precisely estimated effects on the three outcome variables. The table also includes the estimate of the constant,  $\alpha$ , which corresponds to the pre-treatment average. For example, for expected usage, shown in column 1, the informational treatment increases the (self-reported) expected usage variable by about 20% — by 0.8 (s.e. = 0.05) starting

from a base of about 4.1. This closes 27% of the pre-treatment gap between expected usage of state and non-state institutions.

In contrast to the expected usage variable, our two other main variables are not based on self-reports and come from our subjects' behavior in relatively high-stakes experiments. The picture they paint is very similar to the expected usage variable.

Column 2 shows a large impact on the amount allocated to state courts in the fund dictator game. Starting from a base of 104.8, this allocation increases, on average, by 15.4 (s.e. = 1.30). This corresponds to an average post-treatment allocation of 116 (equivalent to 1.10 dollars) or roughly 23% of a day's wages in our sample villages. Put differently, this is about a 15% increase from the pre-treatment base, and is larger than the gap between pre-treatment contributions to the state and the non-state actors. The pattern for the investment game (column 3) is similar; in this case we see an increase of 14.6 (s.e. = 1.34) from a base of 115.4, which again more than closes the gap between the pre-treatment allocations to the state and the non-state institutions. Figure 1 visually summarizes the magnitude of these direct effects.

Panel B of the table turns to our investigation of the social experimenter effect. One may be concerned that our estimates in Panel A reflect a desire on the part of our subjects to act in a way that they think the experimenter would like to see. Since the informational treatment is providing positive news about the state, the subjects may infer that we would like them to become more positive towards the state and act in that way. We use our social experimenter treatment to gauge the extent of these effects. As described in the previous section, we provide a statement that only contains an opinion about state courts, without any clear, objective information. The experimenter effect should be, if anything, stronger in this case because of the stated opinion, but in the absence of objective information there should be no, or only very limited, updating of beliefs (there can be some updating, for example, if the respondent believes that the experimenter's opinion was in turn informed by some relevant metric).

We see in Panel B of Table 2 that the responses of our main outcome variables following the social experimenter treatment are very different than in Panel A: there is a small response for expected usage, which increases by 0.28 (s.e. = 0.05), that is, by about a third of our estimate in Panel A. When we turn to the higher stakes allocation decisions in the two games, there are no statistically significant effects from the social experimenter treatment.

This question is more formally investigated in Panel C, where we pool our informational treatment and social experimenter treatment samples, and estimate the differential effect of the informational treatment relative to the social experimenter treatment. Table A1 in Appendix A shows that individual characteristics and baseline responses are broadly balanced between our state positive and social experimenter treatment samples. In particular, out of the 25 variables on individual characteristics and baseline responses available for our entire sample, three of them (perceptions of

others’ expected usage of state and non-state forums as well as the allocation made to the state in the fund game) show differences that are significant at 10%, which is approximately what we would expect due to sampling variation. Therefore, we estimate the following model:

$$Y_{it} = \alpha + \gamma \text{Post}_i + \beta \text{InfoTreat}_i \cdot \text{Post}_t + \delta_i + \text{Post}_t \cdot S_i + \varepsilon_{it}, \quad (8)$$

where the parameter of interest, again denoted by  $\beta$ , is now the coefficient on the dummy for the second stage of the experiment,  $\text{Post}_i$ , interacted with the group that receives the informational treatment, for whom the dummy  $\text{InfoTreat}_i$  takes the value 1. We also include a full set of interactions between  $\text{Post}_t$  and strata fixed effects, denoted by  $S_i$  for the strata to which individual  $i$  belongs, to account for the strata-level differences in sampling probabilities for the two treatments. The main effect of the second stage, the coefficient on the dummy  $\text{Post}_i$  denoted by  $\gamma$ , is estimated from our subjects that have been given the social experimenter treatment.

The results from the estimation of equation (8), reported in Panel C, are consistent with what can be seen by comparing Panels A and B: there is a large effect of the informational treatment even once the social experimenter effect is netted out. Figure 2, analogously to Figure 1, depicts these estimates netting out the social experimenter effects, which can be seen to have broadly the same magnitudes as the baseline estimates in Figure 1.

Overall, we conclude from Table 2 that there are robust and fairly large effects of our informational treatment, which provide accurate but favorable information about the effectiveness of state courts, on self-reported expected usage and on allocations in high-stakes experimental games. These results suggest that, despite the widespread and deep-rooted inefficiencies of state institutions in Pakistan, citizens are willing to update their views and behavior in response to (credible) information about improved service quality. We next turn to the spillover effects of this informational treatment on behavior concerning the non-state actor.

## 5.2 Indirect Effects

As highlighted in our theoretical discussion, we may also expect the informational treatment to impact our subjects’ interactions with the relevant non-state actor (panchayat), especially under motivated reasoning as in Proposition 2. This issue is investigated in Table 3, which has an identical structure to Table 2, except that the dependent variable is (self-reported) expected usage and the allocations in the two games for the non-state actor (panchayat). The informational treatment continues to be the same — providing positive information about the effectiveness of *state courts* — so that our focus is now the indirect or spillover effects on the *non-state actor*.

The results concerning these indirect effects are very clear-cut. We find uniformly negative effects for the non-state actor.

In Panel A of Table 3, we start with our within-subject design, summarized in equation (7), for the sample that has received the informational treatment. The estimates in the three columns are directly comparable to the estimates in Panel A of Table 2. In column 3 we see a negative effect,  $-0.55$  (s.e. =  $0.04$ ). This is about 30% smaller than the positive estimate in Table 2, but still very precisely estimated and significant. If we add up the direct and the indirect effects, we see that the simple informational treatment closes 53% of the gap between self-reported expected usage of state and non-state institutions in the baseline.

In columns 2 and 3 we see similar effects for the allocations in the fund dictator and investment games. The estimates are again very precise and statistically significant,  $-10.4$  (s.e. =  $1.30$ ) in column 2, and  $-12.13$  (s.e. =  $1.35$ ) in column 3. The effect in column 2 is about 30% smaller than the direct positive effect in Table 2, while the estimate in column 3 is about 15% smaller. Figure 1 also presents these indirect effects and highlights that their magnitudes, though somewhat smaller, are comparable to the direct effects.

In Panel B, we estimate the social experimenter effects in the same fashion as in Panel B of Table 2, and find similar results. There is a negative impact on self-reported expected usage, but this is much smaller than the impact of the informational treatment in Panel A. The effects on the allocations in the two games are small and imprecisely estimated, insignificant for the fund dictator game and marginally significant at 10% for the investment game.

When we net out the social experiment effects in Panel C using the specification in (8), we continue to find significant negative effects from the informational treatment on the non-state actor, though we lose some precision for the game allocations. These estimates are also plotted in Figure 2.

Overall, our results show substantial negative spillovers on non-state institutions once individuals update their beliefs about the effectiveness of state institutions. These results suggest that information about the effectiveness of state institutions may not only convince people to engage more with these institutions, but also to disengage with non-state actors providing substitute services. We next investigate the robustness of these results before turning to a more detailed examination of beliefs and trust that help shed light on the channels of behavioral impact we see.

### 5.3 Robustness

In this subsection, we report several robustness checks on the results presented so far.

Our main specifications utilize a within-subject design, whereby the effects of interest are estimated by using information on how our subjects change their expectations and behavior after being provided with relevant information. Although our experiments were designed for within-subject comparisons, our data still enable a cross-subject strategy (albeit with lower statistical precision) where we simply compare the post information expected usage and allocations in the two games

between subjects who received the information treatment and the social experimenter treatment.<sup>27</sup> These results are presented in Table 4. The first panel of this table focuses on direct effects, while the second panel is for indirect effects. The results are quite similar to those presented in Tables 2 and 3, even if less precise. For example, we see a direct effect of 0.82 (s.e.= 0.19) for expected usage, of 19.71 (s.e. = 3.93) for allocation in the fund dictator game, and 21.31 (s.e. = 3.91) in the investment game. The indirect effects reported in Panel B are similar as well, though again less precisely estimated and thus statistically significant at 1% only for expected usage.

Finally on the same issue, Appendix Table A2 presents results from a specification where we run cross-subject regressions but control for the baseline (pre-information treatment) values of the dependent variable in order to capture cross-subject baseline differences. The estimates in Table A2 show that the coefficients are now strongly significant and very close to those in Tables 2 and 3, highlighting the importance of accounting for the (substantial) cross-subject heterogeneity at baseline.

Overall, although designed for within-subject comparisons, we find it reassuring that the data generated from our experiment yield very similar results with these across-subject specifications.

Table 5 investigates whether subjects correctly forecast how others’ expected usage and allocation choices in the two games change in response to the same information. Columns 1-3 are analogues of the specifications in Panel A of Tables 2 and 3, while columns 4-6 are analogues of the specifications in Panel C (which nets out the social experimenter effect). The results are in line with those in Tables 2 and 3, and show that our subjects not only respond to the positive information about the effectiveness of state institutions by reporting greater expected usage and changing their behavior in the two games, but also predict, correctly and with considerable precision, that others will do likewise, and with very similar magnitudes. We find these patterns to be a reassuring reality check on our main results.

We have also carried out a series of additional robustness checks to ensure that our results are not unduly affected by individuals who may not have fully understood the games, by outliers, or by those who are at the “boundary” (meaning that before the informational treatment they were allocating the minimum or the maximum amount to one of the two actors). Even though we carefully explained the procedures and details of each game to our subjects at each stage, we also wanted to ensure that any remaining imperfect understanding of the game would not influence our results. For this reason, we collected information on the enumerators’ and the respondents’ own assessments of how well they understood the game. In Appendix Table A3, we report estimates after removing any respondents who received a score below 5 (out of 10) from either their own reports or from the enumerators. The results are very similar to our baseline estimates. In Appendix Table A4, we

---

<sup>27</sup>As noted previously, given sampling probabilities of the two treatments varied by strata, we also include a full set of strata fixed effects in this specification.

show estimates from specifications that down-weight outliers (following the procedure of Li, 1985), which are also very similar to our baseline results. In Appendix Table A5, we report estimates after excluding subjects that have contributions in any of the games that are at the boundary (i.e. maximum or minimum value). The results are again very similar to the baseline results. We also randomized the order in which the respondents played different games and answered different questions, and the results in Appendix Table A6 show that there are no robust order effects.

In our pilots, in addition to our baseline informational treatment that provides positive information about the effectiveness of state institutions, we experimented with treatments that provide (accurate) negative information about the effectiveness of state institutions as well as negative or positive information about the effectiveness of non-state actors. Though naturally under-powered, the results in these smaller pilot samples are in line with our expectation. In particular, with the “state negative treatment,” we find statistically significant negative direct effects for usage and investment game (reported in Panel A of Appendix Table A7 - the fund game was not played at the time this pilot was conducted), and a positive and statistically significant indirect effect for expected usage, but imprecise estimates for allocations in the investment game (Panel B). With the “non-state negative treatment,” Appendix Table A8 shows negative and statistically significant direct effects for the non-state actor, and positive and generally statistically significant positive indirect effects for state institutions. With the “non-state positive treatment,” for which we have the smallest sample, the effects have the expected signs (presented in Appendix Table A9), but they are mostly imprecise. Overall, we find it reassuring that, even with the smaller samples, we are able to detect effects consistent with our theoretical expectations, and also comforting that when the information is negative, these effects indeed show negative direct and positive indirect effects.

Our final robustness check investigates whether (positive or negative) “social multiplier” effects, related to expectations of our subjects concerning others’ information or behavior, complicate the interpretation of our results. To check for such social multiplier effects, we emphasized to a random subset of our subjects that others in their village were also given the same information. Using this variation, we estimated social multiplier effects (interacting our  $Post_i$  variable with a dummy for having received this reminder). As Appendix Table A10 shows, we find no evidence of any such effects.

Overall, our main results appear quite robust to a range of variations and are not unduly affected by outliers or respondents that may not have fully understood the context or the details of the game.

#### **5.4 Potential Channels: Effects on Beliefs, Trust and Allegiance**

We next turn to an examination of the potential channels through which the behavioral changes we have documented may be operating. As explained above, it is instructive to look at the effects of our informational treatment on beliefs about the effectiveness of both state and non-state institutions,

trust towards these actors, and allegiance to these actors (corresponding to relationship-specific efforts). Changes in beliefs concerning the effectiveness of state institutions serve primarily as a validity check on the informational treatment. However, recall from Proposition 2 that in the presence of motivated reasoning we should observe changes in beliefs concerning the effectiveness of non-state institutions as well as changes in allegiance.

We investigate these issues in Table 6. Panel A looks at the direct effects on state courts, while Panel B looks at the indirect effects on panchayats. Each column is for one of five variables. Our first two variables are “service effectiveness,” which proxies our subjects’ beliefs about how effectively a forum provides services to an ordinary person and “enforcement effectiveness,” which corresponds to how effectively a forum is able to enforce its verdicts. We then consider our question on access which measures how easy it is for an ordinary person to access the forum. Our fourth variable, trust, provides a general measure of trust in the relevant forum. Our final variable, “allegiance,” measures how much effort the individual expects to put in in order to build a relationship with the relevant forum, and thus corresponds to our theoretical allegiance variable. Our *ex ante* expectations were that we should find direct positive effects for all five of these measures for state courts, and potentially negative indirect effects reflecting motivated reasoning for at least some of these variables for the panchayats.<sup>28</sup>

The results are broadly in line with these expectations. Panel A of Table 6 shows that for the first four variables we find sizable and generally precisely estimated positive direct effects. For example, we estimate a direct effect of 0.52 (s.e. = 0.06) for service effectiveness, 0.26 (s.e. = 0.04) for enforcement effectiveness, 0.41 (s.e. = 0.06) for access, and 0.58 (s.e. = 0.05) for trust. These effects are sizable, typically about 10% of the baseline values (as shown by the constants in the table). We do not find any direct effects on allegiance, a pattern to which we return below.

More interestingly, and in line with the predictions of Proposition 2, we also find fairly precisely estimated, though quantitatively smaller, indirect effects on beliefs, trust, and allegiance concerning non-state institutions. For example, in column 1 of Panel B, the estimate for the indirect effect on service effectiveness for panchayats is negative and significant, -0.07 (s.e. = 0.03). In column 2, for

---

<sup>28</sup>As noted in footnote 21, these belief, trust, and allegiance questions were asked at both baseline and after the informational treatment only at later stages of our fieldwork and therefore only applied to about a third of our sample. Our budget, coupled with our concerns of not having a large enough sample for these questions under the state positive informational treatment, meant that we could not provide the social experimenter treatment to a subsample of our subjects answering these questions. Since our initial sample results do not show any sizable social experimenter effects, we believe that the impacts for these belief, trust, and allegiance questions are also very unlikely to be due to social experimenter effects. Moreover, while we did not ask these detailed questions in the initial sample, we did ask two questions about the relative effectiveness of state and non-state actors in delivering justice and enforcing verdicts at the end of our survey. Since we have the social experimenter treatment for this sample, we can directly test whether social experimenter effects could have a similar impact on these beliefs. Appendix Table A11 presents the results using a cross-subject design (since we have answers to these questions only at the end of the survey) and confirms there are significant effects of the state positive treatment relative to the social experimenter treatment. This lends further support to our decision to exclude the social experimenter treatment in our later-stage sample for budgetary and logistical reasons.

service effectiveness, the estimate is -0.09 (s.e. = 0.03). There are similarly statistically significant and quite precise negative effects on overall trust. The only exception to this pattern is in column 3, for access, where we have a positive and significant estimate. It is not entirely surprising that there should be no negative indirect effect on this variable since it is unlikely that the information about reduced delays in state courts should make our subjects think that accessing panchayats becomes harder. Why might we be estimating a positive (rather than zero) effect? One possibility is that our subjects expect others to switch away from panchayats to state courts and thus also expect that accessing panchayats would get even easier, though we are not certain that this is the right explanation for the pattern we are seeing, and we are not able to verify its applicability.

We also have a sizable effect on allegiance in column 5, -0.19 (s.e. = 0.04). This indirect estimate might be somewhat surprising, since Panel A does not show a direct impact on allegiance. The lack of a direct effect might be due to sample variability, though this seems somewhat unlikely since we can comfortably reject positive direct effects of the same magnitude as the indirect effects. An alternative explanation for this pattern is that our subjects do not perceive the need to invest much effort to build a relationship with state institutions (beyond understanding how some basic aspects of the courts operate), but such relationship-specific investments are important for traditional institutions. Then improved beliefs about the effectiveness of state institutions do not trigger a major increase in efforts to build a relationship with state courts, but still lead to significant declines in efforts targeted at traditional, non-state actors.<sup>29</sup> That said, our subjects reported that they view such investments to be important not just for non-state actors, but also for state courts. Yet this might reflect the need to understand how the courts work, and not translate into changes in relationship-specific efforts in response to new information.

As a summary, Figure 3 depicts these direct and indirect estimates on beliefs and allegiance as well.

We also checked the robustness of the results reported in Table 6. Appendix Tables A12 and A13 show that these results are robust to dropping subjects who may not have fully understood the context or the details of some of the games and to dropping observations on the boundaries.

Overall, the results in Table 6 suggest that the mechanisms we conjectured previously are operating and influencing individual beliefs, trust, and allegiance for state institutions, but even more importantly, consistent with the type of motivated reasoning highlighted in Proposition 2, there are changes in these variables for non-state institutions as well. We find this pattern not just reassuring regarding some of the key mechanisms in our conceptual framework, but also highly suggestive regarding the bigger picture questions we started with. Namely, the negative indirect effects on beliefs imply that a significant cause of citizens turning to non-state actors for services and protection may

---

<sup>29</sup>This is not inconsistent with our theoretical framework; if  $\Gamma_S(\cdot)$  is (nearly) vertical, there would not be a response of allegiance towards the state, but there would continue to be an allegiance response for the non-state actor.



be the perceived ineffectiveness and corruption of state institutions, and this may in turn feed into more positive beliefs concerning the effectiveness and capabilities of non-state actors.

## 5.5 Heterogeneous Effects

We investigated the heterogeneity of the direct and indirect effects we have reported so far by various individual and village level characteristics. Though these often do predict baseline allocations or expected usage, we do not find much evidence of heterogeneous impact of the informational treatment. Specifically, we find no evidence of any heterogeneity by individual caste or other individual characteristics, nor is there any evidence that the effects are heterogeneous by village characteristics.

We also investigated whether the responses of our subjects are different depending on their initial beliefs by estimating the following generalization of (7):

$$Y_{it} = \alpha + \beta \text{Post}_i + \pi^{\text{state}} \text{InitialBelief}_i^{\text{state}} \cdot \text{Post}_t + \pi^{\text{non-state}} \text{InitialBelief}_i^{\text{non-state}} \cdot \text{Post}_t + \delta_i + \varepsilon_{it}, \quad (9)$$

where  $\text{InitialBelief}_i^{\text{state}}$  and  $\text{InitialBelief}_i^{\text{non-state}}$  measure individual  $i$ 's initial beliefs about the state and the non-state actors. This equation, therefore, allows the effects of the informational treatment to be different by the baseline beliefs of individuals concerning the state and/or the non-state actor. We estimated this equation using baseline expected usage to proxy for initial beliefs.

The results are reported in Appendix Table A14, and overall show some interesting heterogeneities: higher (more favorable) initial beliefs for state courts often translate into smaller positive effects on expected usage and the allocations in the two games following our state-positive informational treatment. We did not detect any heterogeneity for the direct effects from initial beliefs concerning the non-state actor. Conversely, we find some evidence that the negative indirect effects for the non-state actor are more pronounced for those with higher (more favorable) initial beliefs concerning this actor, and do not depend on initial beliefs about state courts.

Though interesting, it is worth noting that these patterns are not as robust or precisely estimated as our main results, nor are they directly consistent with a simple Bayesian updating story. In such a story, initial belief differences concerning the state and the non-state actors should have opposite effects, and there should be a non-monotonic relationship between initial beliefs and responsiveness (as those who are very sure about either the state or the non-state actors should not respond much to new information). The results in Appendix Table A14 do not show this pattern of opposite effects from initial beliefs about the state and the non-state actors, and we did not find any systematic evidence of non-monotonic effects. Whether this specific pattern of heterogeneities conditional on baseline beliefs is consistent with motivated reasoning will in turn depend on the details of the model of motivated reasoning and in particular on the degree to which distortions of beliefs are

persistent.

## 6 Conclusion and Future Directions

Lack of trust in state institutions, which discourages citizens from seeking services from and interacting with these institutions, is a pervasive problem in many developing countries. Though this lack of trust is often well grounded in evidence of weakness, ineffectiveness, and corruption of state institutions, it also exacerbates these problems by encouraging citizens to disengage with the state. It also drives them to seek protection, conflict resolution, and public services from non-state actors, which further undermines trust in the state. This feedback process might even be argued to create a vicious circle of weakness and illegitimacy. Despite the ubiquity of these issues in many developing societies, there is relatively little evidence about whether this lack of trust is real, what its implications are, and how the interplay between state and non-state actors contributes to it.

Our paper provides evidence on this set of issues. We focus on rural Punjab in Pakistan, where lack of trust in state institutions is widespread. Our results show fairly large and very robust *direct effects* from the informational treatment: after receiving favorable information about state courts, our subjects become more likely to use state courts and more willing to allocate funds in the experiments toward state institutions. These results indicate that, despite the deep-rooted mistrust of the Pakistani state, accurate, credible information can change people’s beliefs and behavior.

Interestingly, we also estimate large and robust *indirect effects* on the relevant non-state actor (panchayats): after the same informational treatment, our subjects report that they are less likely to use non-state institutions for dispute resolution and allocate fewer resources to panchayats in our two games.

Our survey questions enable us to dig deeper into the mechanisms through which these direct and indirect effects are working. We find that the informational treatment does indeed improve people’s beliefs about state courts. More importantly, it makes them more pessimistic about the effectiveness of non-state institutions as well, and reduces both their trust in and allegiance to these competitors of the state.

We argue that our evidence can be best understood from the viewpoint of a “motivated reasoning” model, where reduced usage of these non-state institutions makes people less likely to hold positive views about them. In terms of the larger motivation of our project, these patterns, and especially the indirect effects on panchayats, provide some evidence for one aspect of the negative feedback cycle — whereby positive views of non-state institutions are fed by the ineffectiveness of state institutions and vice versa — but also suggest that this negative feedback can be reversed if and when state institutions become more effective and credible information about their performance is provided to citizens.

We view our paper as a first step in a systematic investigation of the interplay between state and non-state institutions and citizens' beliefs about and trust in these institutions. In ongoing work, we are pursuing this agenda further by focusing on populations that have recently experienced a dispute and are at the initial stages of considering which (state or non-state) forum to utilize. In addition to providing such individuals with analogous informational treatments about the state's recent improvements, we are exploring ways to better provide state services to them. The goal is to measure not only self-reported and game based outcomes but also to follow these individuals over time and examine which forum they choose to engage with and in what manner. This will allow us to examine whether the results documented in this paper continue to operate over extended periods of time by impacting decisions people make when dealing with the actual disputes they are facing.

## References

- Acemoglu, D. (2005). Politics and economics in weak and strong states. *Journal of Monetary Economics*, 52(7):1199–1226.
- Acemoglu, D., Ticchi, D., and Vindigni, A. (2011). Emergence and persistence of inefficient states. *Journal of the European Economic Association*, 9(2):177–208.
- Akerlof, G. A. and Dickens’s, W. T. (1982). The economic consequences of cognitive dissonance. *American Economic Review*, 72(3):307–319.
- Amsden, A. H. (1992). *Asia’s Next Giant: South Korea and Late Industrialization*. Oxford University Press on Demand.
- Anderson, M. R. (2010). Community psychology, political efficacy, and trust. *Political Psychology*, 31(1):59–84.
- Ashraf, N., Bohnet, I., and Piankov, N. (2006). Decomposing trust and trustworthiness. *Experimental Economics*, 9(3):193–208.
- Ayaz, A. and Fleschenberg, A. (2009). *The Gender Face of Asian Politics*. Oxford University Press, USA.
- Bem, D. J. (1967). Self-perception: An alternative interpretation of cognitive dissonance phenomena. *Psychological Review*, 74(3):183.
- Bénabou, R. and Tirole, J. (2002). Self-confidence and personal motivation. *The Quarterly Journal of Economics*, 117(3):871–915.
- Bénabou, R. and Tirole, J. (2004). Willpower and personal rules. *Journal of Political Economy*, 112(4):848–886.
- Bénabou, R. and Tirole, J. (2016). Bonus culture: Competitive pay, screening, and multitasking. *Journal of Political Economy*, 124(2):305–370.
- Berg, J., Dickhaut, J., and McCabe, K. (1995). Trust, reciprocity, and social history. *Games and Economic Behavior*, 10(1):122–142.
- Besley, T. and Persson, T. (2009). The origins of state capacity: Property rights, taxation, and politics. *American Economic Review*, 99(4):1218–44.
- Blair, G., C., C. F., Malhotra, N., and Shapiro, J. N. (2013). Poverty and support for militant politics: Evidence from Pakistan. *American Journal of Political Science*, 57(1):30–48.

- Bohnet, I. and Zeckhauser, R. (2004). Trust, risk and betrayal. *Journal of Economic Behavior & Organization*, 55(4):467–484.
- Brohi, N. (2016). Women, violence and jirgas. Retrieved from: <http://af.org.pk/gep/images/publications/Research>
- Brunnermeier, M. K. and Parker, J. A. (2005). Optimal expectations. *American Economic Review*, 95(4):1092–1118.
- Bullock, W., Imai, K., and Shapiro, J. N. (2011). Statistical analysis of endorsement experiments: Measuring support for militant groups in Pakistan. *Political Analysis*, 19(4):363–384.
- Burks, S. V., Carpenter, J. P., and Verhoogen, E. (2003). Playing both roles in the trust game. *Journal of Economic Behavior & Organization*, 51(2):195–216.
- Bursztyjn, L., Callen, M., Ferman, B., Gulzar, S., Hasanain, A., and Yuchtman, N. (2016). Political identity: Experimental evidence on anti-Americanism in Pakistan. *Unpublished manuscript, Univ. Chicago, Chicago, IL Article Locations: Article Location Article Location Article Location Article Location*.
- Camerer, C. F. and Fehr, E. (2004). Measuring social norms and preferences using experimental games: A guide for social scientists. *Foundations of Human Sociality: Economic Experiments and Ethnographic Evidence from Fifteen Small-scale Societies*, 97:55–95.
- Caplin, A. and Leahy, J. (2001). Psychological expected utility theory and anticipatory feelings. *The Quarterly Journal of Economics*, 116(1):55–79.
- Cardenas, J. C. and Carpenter, J. (2008). Behavioural development economics: Lessons from field labs in the developing world. *The Journal of Development Studies*, 44(3):311–338.
- Carrillo, J. D. and Mariotti, T. (2001). Electoral competition and politician turnover. *European Economic Review*, 45(1):1–25.
- Centeno, M. A. (2002). *Blood and Debt: War and Statemaking in Latin America*. Princeton: Princeton University Press.
- Chaudhary, M. A. (1999). *Justice in Practice: Legal Ethnography of a Pakistani Punjabi Village*. Oxford University Press, USA.
- Cheema, A., Hameed, Z., and Shapiro, J. (2017). Victimization, citizen engagement, and policing in Lahore. *Institute of Development and Economic Alternatives, Policy Report, Lahore, Pakistan*.

- Cheema, A., Khan, A. Q., and Myerson, R. (2010). Breaking the countercyclical pattern of local democracy in Pakistan. *Unpublished manuscript. Available from <http://home.uchicago.edu/rmyerson/research/pakdemoc.pdf>.*
- Cheema, A. and Naseer, M. F. (2013). Historical inequality and intergenerational educational mobility: The dynamics of change in rural Punjab. *The Lahore Journal of Economics*, 18(special edition):211.
- Chemin, M. (2009). The impact of the judiciary on entrepreneurship: Evaluation of Pakistan's "access to justice programme". *Journal of Public Economics*, 93:114–125.
- Clark, J. A. (2004). *Islam, Charity, and Activism: Middle-class Networks and Social Welfare in Egypt, Jordan, and Yemen*. Indiana University Press.
- Coleman, J. (1990). Foundations of social theory. *Cambridge, MA: Belknap*.
- Cowasjee, A. (2004). A first step. Dawn Newspaper. Retrieved from: <https://www.dawn.com/news/1072875>.
- Cox, J. C., Ostrom, E., Walker, J. M., Castillo, A. J., Coleman, E., Holahan, R., Schoon, M., and Steed, B. (2009). Trust in private and common property experiments. *Southern Economic Journal*, pages 957–975.
- Delavande, A. and Zafar, B. (2012). How deeply held are anti-American attitudes among Pakistani youth? evidence using experimental variation in information.
- Dell, M., Lane, N., and Querubin, P. (2015). State capacity, local governance, and economic development in Vietnam. *NBER Working Paper*, pages 1–40.
- Dufwenberg, M. and Gneezy, U. (2000). Measuring beliefs in an experimental lost wallet game. *Games and Economic Behavior*, 30(2):163–182.
- Edwards, K. and Smith, E. E. (1996). A disconfirmation bias in the evaluation of arguments. *Journal of Personality and Social Psychology*, 71(1):5.
- Eisenberg, T., Kalantry, S., and Robinson, N. (2013). Litigation as a measure of well-being. *DePaul Law Review*, 62(2).
- Epley, N. and Gilovich, T. (2016). The mechanics of motivated reasoning. *Journal of Economic Perspectives*, 30(3):133–40.
- Evans, P. B. (2012). *Embedded Autonomy: States and Industrial Transformation*. Princeton University Press.

- Festinger's, L. (1962). *A Theory of Cognitive Dissonance*, volume 2. Stanford University Press.
- Gayer, L. (2014). *Karachi: Ordered Disorder and the Struggle for the City*. Oxford University Press (UK).
- Gilovich, T. and Ross, L. (2015). *The Wisest One in the Room*. Free Press.
- Harmsen, E. (2010). *Islam, Civil Society and Social Work: Muslim Voluntary Welfare Associations in Jordan between Patronage and Empowerment*. Amsterdam University Press.
- Herbst, J. (2014). *States and Power in Africa: Comparative Lessons in Authority and Control*. Princeton University Press.
- Hoebel, E. A. (1965). Part two ethnographic approaches: Fundamental cultural postulates and judicial lawmaking in Pakistan. *American Anthropologist*, 67(6):43–56.
- Jackson, J., Asif, M., Bradford, B., and Zakria Zakar, M. (2014). Corruption and police legitimacy in Lahore, Pakistan. *British Journal of Criminology*, 54(6):1067–1088.
- Javed, A. (2017). Justice for rich, jail for poor. <https://nation.com.pk/07-May-2017/justice-for-rich-jail-for-poor>. Accessed: 2018-04-10.
- Johnson, C. (1982). *MITI and the Japanese Miracle: The Growth of Industrial Policy: 1925-1975*. Stanford University Press.
- Johnson, N. D. and Mislin, A. A. (2011). Trust games: A meta-analysis. *Journal of Economic Psychology*, 32(5):865–889.
- Jost, J. T., Glaser, J., Kruglanski, A. W., and Sulloway, F. J. (2003). Political conservatism as motivated social cognition. *Psychological bulletin*, 129(3):339.
- Kahan, D. M. (2013). Ideology, motivated reasoning, and cognitive reflection: An experimental study. *Judgement and Decision Making*, 8(4):407–424.
- Kepel, G. (2009). *Beyond Terror and Martyrdom*. Harvard University Press.
- Khan, M., Kapoor, S., and Cooraswamy, R. (2000). Domestic violence against women and girls. *Innocenti Digest*, 6:1–30.
- Kramer, R. M. (1999). Trust and distrust in organizations: Emerging perspectives, enduring questions. *Annual Review of Psychology*, 50(1):569–598.
- Kunda, Z. (1990). The case for motivated reasoning. *Psychological Bulletin*, 108(3):480.

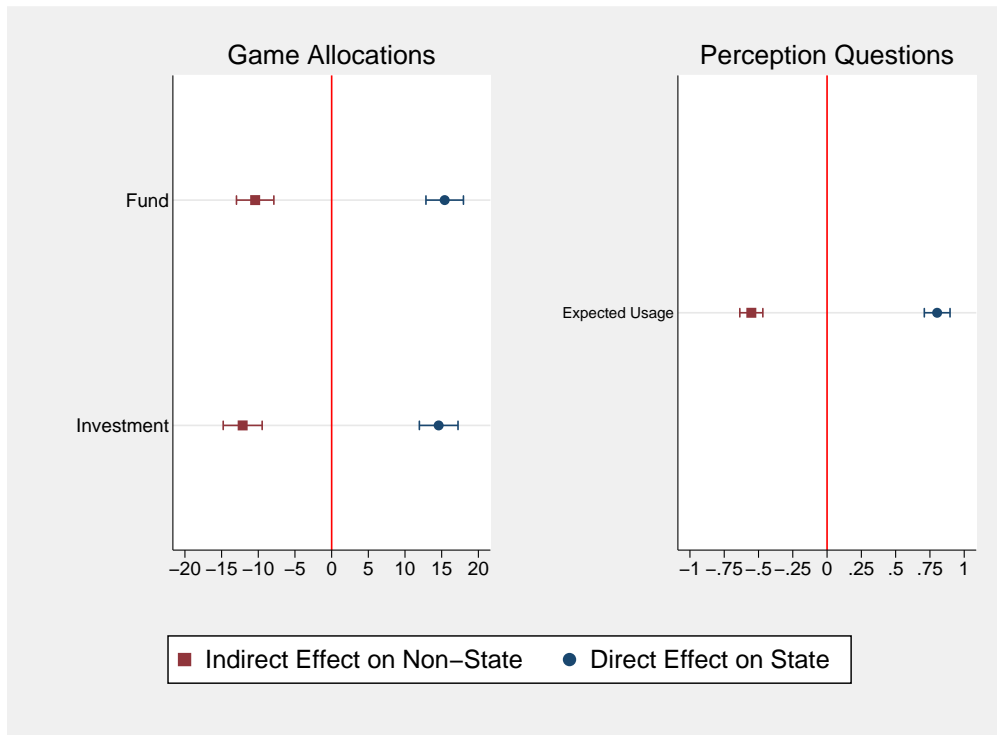
- Loewenstein, G. (1987). Anticipation and the valuation of delayed consumption. *The Economic Journal*, 97(387):666–684.
- Migdal, J. S. (1988). *Strong Societies and Weak States: State-society Relations and State Capabilities in the Third World*. Princeton University Press.
- Mishler, W. and Rose, R. (2001). What are the origins of political trust? testing institutional and cultural theories in post-communist societies. *Comparative Political Studies*, 34(1):30–62.
- Nekokara, M. A. (2016). Access to justice and legal aid. <https://www.dawn.com/news/1301948>. Accessed: 2018-04-10.
- Ortoleva, P. and Snowberg, E. (2015). Overconfidence in political behavior. *American Economic Review*, 105(2):504–35.
- Rabin, M. (1994). Cognitive dissonance and social change. *Journal of Economic Behavior & Organization*, 23(2):177–194.
- Rana, M. A. (2009). Taliban insurgency in Pakistan: A counterinsurgency perspective. *Conflict and Peace Studies*, 2(2):12.
- Rehman, Z. U., Haider, I., and Zahid, F. M. (2014). Swat: an unquiet calm. Retrieved from: <https://www.dawn.com/news/1133198>.
- Rothstein, B. and Stolle, D. (2008). The state and social capital: An institutional theory of generalized trust. *Comparative politics*, 40(4):441–459.
- Roy, S. (2013). *Hamas and Civil Society in Gaza: Engaging the Islamist Social Sector*. Princeton University Press.
- Schoon, E. and Cheng, H. (2011). Determinants of political trust: A lifetime learning model. *Developmental Psychology*, 47(3):619.
- Shinwari, N. A. (2015). Understanding the informal justice system: Opportunities and possibilities for legal pluralism in Pakistan.
- Siddiqi, F. (2016). Is speedy justice possible? <https://www.dawn.com/news/1255917>. Accessed: 2018-04-10.
- Siddique, O. (2013). *Pakistan's Experience with Formal Law: an Alien Justice*. Cambridge University Press.
- Siddique, O. (2015). Legal advisory and representational services in south Punjab: Baseline study of availability and priority needs.



- Siddique, O. (2016). Caseflow management in courts of Punjab: Framework, practices and reform measures. European Union Punjab Access to Justice Project. Accessed: 2018-04-10.
- Soomro, A. N. and Chandio, A. A. (2013). Challenges to good governance: A case study of issues in local government system of Pakistan. *Journal of African & Asian Local Government Studies*, 2(4).
- Sullivan, J. L. and Transue, J. E. (1999). The psychological underpinnings of democracy: A selective review of research on political tolerance, interpersonal trust, and social capital. *Annual Review of Psychology*, 50(1):625–650.
- Trivers, R. (2011). *The Folly of Fools: The Logic of Deceit and Self-deception in Human Life*. Basic Books (AZ).
- Van den Steen, E. (2004). Rational overoptimism (and other biases). *American Economic Review*, 94(4):1141–1151.
- Von Hippel, W. and Trivers, R. (2011). Reflections on self-deception. *Behavioral and Brain Sciences*, 34(1):41–56.
- Wade, R. (1990). *Governing the Market: Economic Theory and the Role of Government in East Asian Industrialization*. Princeton University Press.
- Walsh, D. (2009). Pakistan bows to demand for sharia law in taliban-controlled Swat Valley. *The Guardian*, 14.
- Zainulbhai, T. (2016). Justice for all: Improving the lok adalat system in India. *Fordham International Law Journal*, 35.
- Zmerli, S. and Newton, K. (2008). Social trust and attitudes toward democracy. *Public opinion quarterly*, 72(4):706–724.

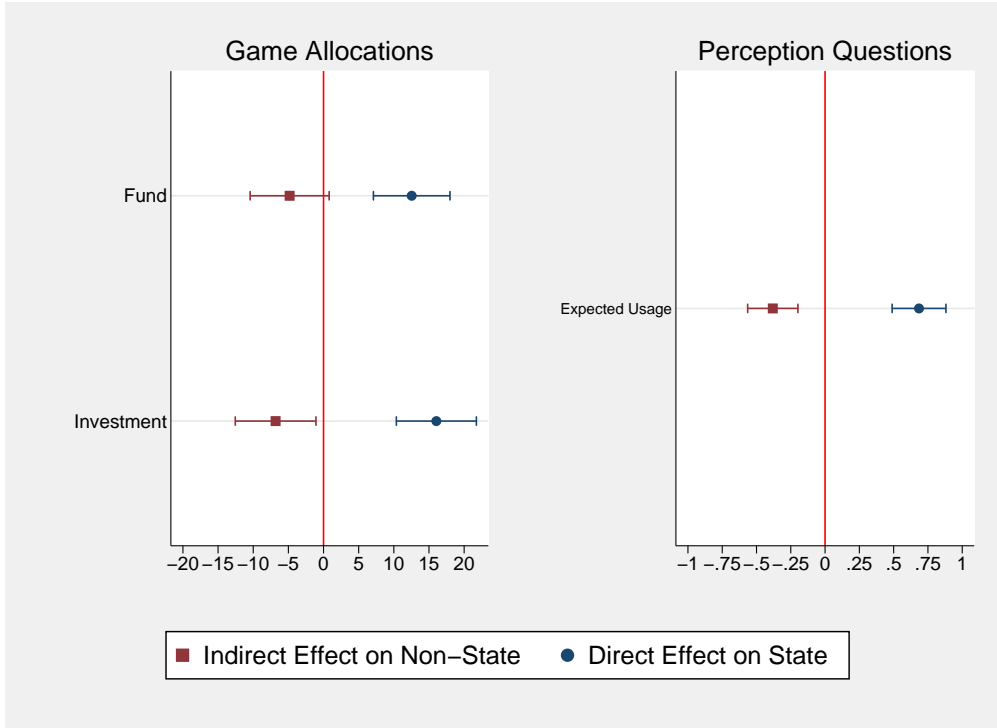
# Main Figures and Tables

Figure 1: Effect of Treatments on Outcomes



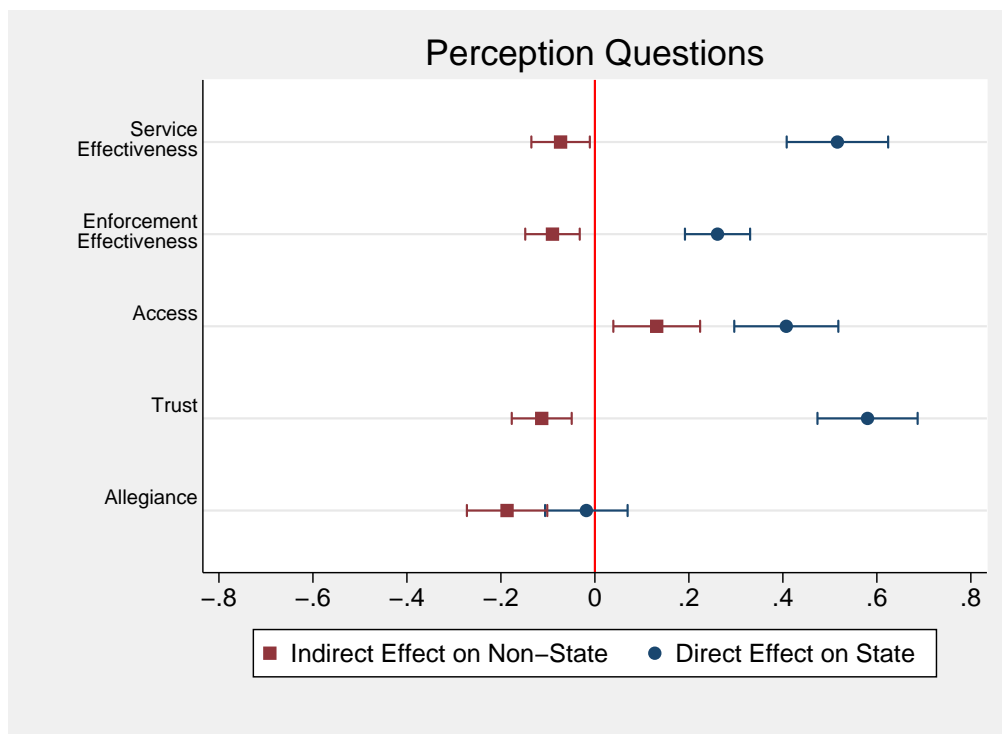
*Notes:* This figure reports the effect of the state-positive treatment on our main outcomes for both forums (analogous to Panels A and C in Tables 2 and 3). Outcomes are regressed on a post indicator and individual fixed effects. Effect sizes shown are the coefficient from the post indicator. Effects on non-state outcomes are shown in red squares, and effects on state outcomes are shown in blue squares. Expected Usage asks respondents how likely (0-10) they are to use each forum, Fund Allocation is the amount the respondent allocated to the forum in the Fund Game, and Investment Allocation is the amount allocated to the forum in the Investment Game. 95% confidence interval shown alongside coefficient.

Figure 2: Effect of Treatments on Outcomes (Net Social-Experimenter)



*Notes:* This figure reports the effect of the state-positive treatment on our main outcomes for both forums, net any social experimenter effect (analogous to Panel C in Tables 2 and 3). Outcomes are regressed on a post indicator, a state-positive treatment indicator, their interaction, a set of individual fixed effects, and a set of strata fixed effects interacted with the post indicator. Effect sizes shown are the coefficients from the interaction term. Effects on non-state outcomes are shown in red squares, and effects on state outcomes are shown in blue squares. Expected Usage asks respondents how likely (0-10) they are to use each forum, Fund Allocation is the amount the respondent allocated to the forum in the Fund Game, and Investment Allocation is the amount allocated to the forum in the Investment Game. 95% confidence interval shown alongside coefficient.

Figure 3: **Effect of Treatments on Forum Perception Outcomes**



*Notes:* This figure reports the effect of the state-positive treatment on the outcomes in Table 6 for both forums. Outcomes are regressed on a post indicator and individual fixed effects. Effect sizes shown are the coefficient from the post indicator. Effects on non-state outcomes are shown in red squares, and effects on state outcomes are shown in blue squares. Expected Usage asks respondents how likely they are to use each forum. Service Effectiveness measures how effective they believe the forum to be at providing service to the ordinary man, Enforcement Effectiveness measures how effective respondents believe the forum to be at enforcing its verdicts, Access measures how difficult respondents believe it is for any ordinary man to access the forum, Trust measures how much respondents trust forum officials, and Allegiance measures how much effort the respondent would make in order to strengthen their connection to the forum. All perception questions are on a scale of 0-10. 95% confidence interval shown alongside coefficient.

Table 1: **Summary of Key Variables**

	Mean	Std. Dev.	N
<b>Demographic Variables:</b>			
Age	37.57	12.38	2984
Income	17001.42	68015.62	1629
Formal Education	0.72	0.45	2944
Asset Index	-0.01	0.80	2151
Landlord	0.28	0.45	2986
<b>Service Usage:</b>			
Gov. Health Center	0.39	0.49	2159
Gov. Education	0.44	0.50	2159
Police	0.20	0.40	2159
Courts	0.21	0.41	2159
<b>Civic Values:</b>			
Helped Neighbor	0.50	0.50	2156
Resolved Dispute	0.40	0.49	2155
Civic Donation	0.06	0.24	2155
NGO Member	0.06	0.23	2153
<b>State Outcomes:</b>			
<b>Self:</b>			
Expected Usage	3.81	3.76	2757
Invest Allocation	111.57	79.63	2872
Fund Allocation	103.39	78.32	2870
<b>Others:</b>			
Other Expected Usage	4.97	2.96	2973
Other Fund	112.66	65.67	2426
Other Invest	119.35	67.31	2421
<b>Non-State Outcomes:</b>			
<b>Self:</b>			
Expected Usage	6.76	3.65	2757
Invest Allocation	116.51	80.94	2872
Fund Allocation	105.68	80.79	2870
<b>Others:</b>			
Other Expected Usage	6.05	3.02	2972
Other Fund	113.78	70.90	2428
Other Invest	120.30	70.75	2425

*Notes:* This table presents summary of our respondents' demographics, as well as baseline summary statistics for our primary outcome variables. Income given as household monthly income. Landlord is an indicator for if the respondent owns land. Asset index calculated using factor analysis on number of chairs, beds, motorcycles, bicycles, buffaloes, goats, radios, televisions, and heaters the respondent's household owned. Service usage questions are indicators for if the respondent had used the given service in the three months prior to surveying. Similarly, civic values questions ask respondents if they have, in the three months prior to surveying, helped their neighbor (e.g. harvesting, building, or home repairs), helped resolve a community dispute, donated (money or time) to a non-governmental social welfare/civic organization, or held membership in an NGO. Observation counts change due to some non-responsiveness and changes in survey instrument over time.

Table 2: **Direct Effect on State Actor**

	(1) Expected Usage	(2) Fund Allocation	(3) Invest Allocation
<b>Panel A: Effect of State Positive Treatment</b>			
Post Treatment	0.80*** (0.05)	15.41*** (1.30)	14.59*** (1.34)
Constant	4.06*** (0.03)	104.82*** (0.92)	115.39*** (0.95)
<b>Panel B: Effect of Social Experimenter Treatment</b>			
Post Treatment	0.28*** (0.05)	1.72 (1.66)	2.23 (1.81)
Constant	3.23*** (0.04)	100.32*** (1.18)	103.26*** (1.28)
<b>Panel C: Netting Out Social Experimenter Effect</b>			
Post $\times$ State-Positive	0.69*** (0.10)	12.54*** (2.78)	16.05*** (2.90)
<b>Observations:</b>			
Panel A:	3812	3918	3938
Panel B:	1702	1822	1806
Panel C:	5514	5740	5744

*Notes:* This table examines the (direct) effect of the state-positive treatment on our main outcomes concerning the state forum. All regressions include two observations per respondent, a pre-treatment observation and a post-treatment observation. Outcomes are regressed on a post-treatment indicator. Panel A reports results for those respondents who received the state-positive treatment, and Panel B reports the same outcomes for respondents who received the social experimenter treatment. Panel C uses both set of respondents to net out the social experimenter effect from the state positive effect. Expected Usage asks respondents how likely (0-10) they are to use each forum, Fund Allocation is the amount the respondent allocated to the forum in the Fund Game, and Investment Allocation is the amount allocated to the forum in the Investment Game. Observation counts vary slightly due to response rate differences and small changes in survey questions. Regressions in all panels include individual fixed effects. Panel C also includes strata fixed effects interacted with a post indicator. This ensures the cross-treatment comparisons in Panel C are not affected by any strata-level differences that may arise due to the varying sampling probabilities (across strata) for the state positive and social experimenter treatments. Standard errors reported in parentheses. \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

Table 3: **Indirect Effect on Non-state Actor**

	(1) Expected Usage	(2) Fund Allocation	(3) Invest Allocation
<b>Panel A: Effect of State Positive Treatment</b>			
Post Treatment	-0.55*** (0.04)	-10.42*** (1.30)	-12.13*** (1.35)
Constant	6.64*** (0.03)	103.99*** (0.92)	116.68*** (0.96)
<b>Panel B: Effect of Social Experimenter Treatment</b>			
Post Treatment	-0.24*** (0.06)	-2.84 (1.87)	-3.23* (1.82)
Constant	7.01*** (0.04)	109.31*** (1.32)	116.15*** (1.29)
<b>Panel C: Netting Out Social Experimenter Effect</b>			
Post × State-Positive	-0.38*** (0.09)	-4.82* (2.87)	-6.81** (2.93)
<b>Observations:</b>			
Panel A:	3810	3918	3938
Panel B:	1700	1822	1806
Panel C:	5510	5740	5744

*Notes:* This table examines the (indirect) effect of the state-positive treatment on our main outcomes concerning the non-state forum. All regressions include two observations per respondent, a pre-treatment observation and a post-treatment observation. Outcomes are regressed on a post-treatment indicator. Panel A reports results for those respondents who received the state-positive treatment, and Panel B reports the same outcomes for respondents who received the social experimenter treatment. Panel C uses both set of respondents to net out the social experimenter effect from the state positive effect. Expected Usage asks respondents how likely (0-10) they are to use each forum, Fund Allocation is the amount the respondent allocated to the forum in the Fund Game, and Investment Allocation is the amount allocated to the forum in the Investment Game. Observation counts vary slightly due to response rate differences and small changes in survey questions. Regressions in all panels include individual fixed effects. Panel C also includes strata fixed effects interacted with a post indicator. This ensures the cross-treatment comparisons in Panel C are not affected by any strata-level differences that may arise due to the varying sampling probabilities (across strata) for the state positive and social experimenter treatments. Standard errors reported in parentheses. \* p<0.10, \*\* p<0.05, \*\*\* p<0.01

Table 4: **Effect of State-Positive Treatment (Across Subject)**

	(1) Expected Usage	(2) Fund Allocation	(3) Invest Allocation
<b>Panel A: Direct Effect on State Institutions</b>			
SP Dummy	0.82*** (0.19)	19.71*** (3.93)	21.31*** (3.91)
<b>Panel B: Indirect Effect on Non-State</b>			
SP Dummy	-0.65*** (0.19)	-3.41 (3.88)	-3.10 (3.94)
Observations	2755	2870	2872

*Notes:* This table examines the effect of the state-positive treatment using a pure cross-subject design. The sample is therefore limited to only post-treatment observations, and a regression is run of outcomes on a state-positive treatment indicator. The comparison category are respondents who received the social-experimenter treatment, so the coefficient on the state-positive indicator shows the effect of SP-treatment beyond that which can be explained by the SE effect. Panel A reports results for outcomes concerning the State Forum, and Panel B reports the same outcomes for the non-state Forum. Expected Usage asks respondents how likely (0-10) they are to use each forum, Fund Allocation is the amount the respondent allocated to the forum in the Fund Game, and Investment Allocation is the amount allocated to the forum in the Investment Game. Observation counts vary slightly due to response rate differences and small changes in survey questions. All regressions include strata fixed effects. This ensures the cross-treatment comparisons are not affected by any strata-level differences that may arise due to the varying sampling probabilities (across strata) for the state positive and social experimenter treatments. Standard errors reported in parentheses. \* p<0.10, \*\* p<0.05, \*\*\* p<0.01



Table 5: **Beliefs about Others**

	State-Positive			Netting Social Experimenter		
	(1) Expected Usage	(2) Fund Allocation	(3) Invest Allocation	(4) Expected Usage	(5) Fund Allocation	(6) Invest Allocation
<b>Panel A: Direct Effect on State Institution</b>						
Post Treatment	0.66*** (0.04)	13.60*** (1.09)	12.65*** (1.10)			
Post × SP				0.59*** (0.08)	8.22*** (2.68)	13.20*** (2.64)
Constant	5.12*** (0.03)	113.12*** (0.77)	122.04*** (0.78)			
<b>Panel B: Indirect Effect on Non-State</b>						
Post Treatment	-0.46*** (0.04)	-7.73*** (1.13)	-10.33*** (1.07)			
Post × SP				-1.32** (0.59)	-7.73*** (2.77)	-10.05*** (2.64)
Constant	6.01*** (0.03)	112.46*** (0.80)	119.88*** (0.76)			
Observations	4024	3500	3496	5944	4856	4850
Respondents	2012	1750	1748	2972	2428	2425

*Notes:* This table examines the effect of the state-positive treatment on alternative outcomes concerning our respondents' beliefs about others. Columns (1)-(3) report results of regression analogous to those from Panel A of Table 2 (regression on post-treatment indicator only), while Columns (4)-(6) report results from regressions analogous to those from Panel C of Table 2 (netting out the social-experimenter effect through interacting post-treatment and SP-treatment dummies). All regressions include two observations per respondent, a pre-treatment observation and a post-treatment observation. Panel A reports results for outcomes concerning the state forum, and Panel B reports the same outcomes for the non-state forum. Expected Usage asks respondents how likely (0-10) they believe others are to use a forum, Fund Allocation is the amount the respondent believes others would allocate to the forum in the Fund Game, and Investment Allocation is the amount the respondent believes others would allocate in the Investment Game. Observation counts vary slightly due to response rate differences and small changes in survey questions (Columns (4) to (6) have more observations because they include the sample with the social experimenter treatment as well). All regressions include individual fixed effects. Regressions in columns (4)-(6) also includes strata fixed effects interacted with a post indicator. This ensures the cross-treatment comparisons in these columns are not affected by any strata-level differences that may arise due to the varying sampling probabilities (across strata) for the state positive and social experimenter treatments. Standard errors reported in parentheses. \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

Table 6: **Forum Perception Questions**

	(1) Service Effectiveness	(2) Enforcement Effectiveness	(3) Access	(4) Trust	(5) Allegiance
<b>Panel A: Direct Effect on State Institution</b>					
Post Treatment	0.52*** (0.06)	0.26*** (0.04)	0.41*** (0.06)	0.58*** (0.05)	-0.02 (0.04)
Constant	4.12*** (0.04)	6.73*** (0.03)	3.06*** (0.04)	5.16*** (0.04)	7.33*** (0.03)
<b>Panel B: Indirect Effect on Non-State</b>					
Post Treatment	-0.07** (0.03)	-0.09*** (0.03)	0.13*** (0.05)	-0.11*** (0.03)	-0.19*** (0.04)
Constant	5.37*** (0.02)	6.23*** (0.02)	7.74*** (0.03)	6.51*** (0.02)	7.13*** (0.03)
Observations	1644	1640	1642	1644	1648
Respondents	822	820	821	822	824

*Notes:* This table examines the effect of the state-positive treatment on alternative outcomes concerning respondents' beliefs about different aspects of each forum. All regressions include two observations per respondent, a pre-treatment observation and a post-treatment observation. Outcomes are regressed on a post-treatment indicator. Panel A reports results for outcomes concerning the state forum, and Panel B reports the same outcomes for the non-state forum. All outcomes are measured on a scale of 0-10. Service Effectiveness measures how effective they believe the forum to be at providing service to the ordinary man, Enforcement Effectiveness measures how effective respondents believe the forum to be at enforcing it verdicts, Access measures how difficult respondents believe it is for any ordinary man to access the forum, Trust measures how much respondents trust forum officials, and Allegiance measures how much effort the respondent would make in order to strengthen their connection to the forum. Observation counts are lower than in other tables due to these questions being added to the survey during later rounds. All regressions include individual fixed effects. Standard errors reported in parentheses. \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

# Appendix A: Additional Figures and Tables

Figure A1: Boards Used for Field Games



(a) Board with game allocation slots closed.



(b) Board with game allocation slots open.

Table A1: **Baseline Balance on Outcomes and Covariates**

	State-Positive		Social Experimenter		Balance		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Mean	Obs	Mean	Obs	Difference	p-value	Total Obs
<b>Demographics:</b>							
Age	37.84	2022	36.99	962	0.85	0.16	2984
Income	16023.06	916	18258.34	713	-2235.28	0.55	1629
Formal Education	0.73	1981	0.71	963	0.02	0.44	2944
Asset Index	-0.03	1195	0.02	956	-0.04	0.26	2151
Landlord	0.29	2023	0.26	963	0.03	0.13	2986
<b>Service Usage:</b>							
Gov. Health Center	0.39	1197	0.39	962	0.01	0.79	2159
Gov. Education	0.44	1197	0.44	962	-0	0.99	2159
Police	0.21	1197	0.20	962	0.01	0.74	2159
Courts	0.20	1197	0.22	962	-0.02	0.28	2159
<b>Civic Values:</b>							
Helped Neighbor	0.51	1195	0.49	961	0.01	0.58	2156
Resolved Dispute	0.39	1195	0.41	960	-0.02	0.36	2155
Civic Donation	0.06	1197	0.07	958	-0.01	0.41	2155
NGO Member	0.06	1195	0.05	958	0.01	0.44	2153
<b>State Outcomes:</b>							
<b>Self:</b>							
Expected Usage	3.85	1906	3.71	851	0.14	0.48	2757
Invest Allocation	113.23	1969	107.97	903	5.26	0.19	2872
Fund Allocation	105.67	1959	98.50	911	7.17	0.07	2870
<b>Others:</b>							
Other Expected Usage	5.08	2014	4.72	959	0.36	0.01	2973
Other Fund	113.66	1748	110.10	678	3.56	0.35	2426
Other Invest	120.98	1746	115.14	675	5.84	0.14	2421
<b>Non-State Outcomes:</b>							
<b>Self:</b>							
Expected Usage	6.68	1906	6.94	851	-0.26	0.16	2757
Invest Allocation	117.68	1969	113.97	903	3.71	0.36	2872
Fund Allocation	106.13	1959	104.72	911	1.41	0.73	2870
<b>Others:</b>							
Other Expected Usage	5.95	2012	6.26	960	-0.31	0.03	2972
Other Fund	114.09	1750	112.98	678	1.11	0.79	2428
Other Invest	119.83	1748	121.50	677	-1.66	0.69	2425

*Notes:* This table presents balance checks on the respondent-level treatment randomization between the state positive and social experimenter. Strata fixed effects are included due to treatment assignment probability varying by strata. Income given as household monthly income. Landlord is an indicator for if the respondent owns land. Asset index calculated using factor analysis on number of chairs, beds, motorcycles, bicycles, buffaloes, goats, radios, televisions, and heaters the respondent's household owned. Service usage questions are indicators for if the respondent had used the given service in the three months prior to surveying. Similarly, civic values questions ask respondents if they have, in the three months prior to surveying, helped their neighbor (e.g. harvesting, building, or home repairs), helped resolve a community dispute, donated (money or time) to a non-governmental social welfare/civic organization, or held membership in an NGO. Columns (1)-(2) report the mean and observations for the respondents who received the state-positive treatment, Columns (3)-(4) report the mean and observations for the respondents who received the Social-Experimenter Treatment, Column (5) shows the difference between the means of these two groups, Column (6) shows the p-value for a t-test measuring if that difference is different from zero, and Column (7) shows the total number of observations.

Table A2: Controlling for Baseline Values

	(1) Expected Usage	(2) Fund Allocation	(3) Invest Allocation
<b>Panel A: Direct Effect on State Institutions</b>			
State-Positive	0.70*** (0.10)	14.33*** (2.60)	17.53*** (2.68)
Baseline	0.86*** (0.01)	0.75*** (0.01)	0.72*** (0.01)
Constant	0.67*** (0.08)	27.13*** (2.36)	30.17*** (2.49)
<b>Panel B: Indirect Effect on Non-State</b>			
State-Positive	-0.41*** (0.09)	-4.41* (2.61)	-5.76** (2.69)
Baseline	0.89*** (0.01)	0.71*** (0.01)	0.71*** (0.01)
Constant	0.59*** (0.10)	25.72*** (2.39)	27.87*** (2.54)
Observations	2755	2870	2872
Respondents	1378	1435	1436

*Notes:* This table examines the effect of the state-positive treatment on our main outcomes independent of any potentially confounding social-experimenter effect. The sample is limited to only post-treatment observations, and a regression is run of outcomes on a state-positive treatment indicator and the subject's baseline response. The comparison category are respondents who received the social-experimenter treatment, so the coefficient on the state-positive indicator shows the effect of SP-treatment beyond that which can be explained by the SE effect. Panel A reports results for outcomes concerning the State Forum, and Panel B reports the same outcomes for the non-state forum. Expected Usage asks respondents how likely (0-10) they are to use each forum, Fund Allocation is the amount the respondent allocated to the forum in the Fund Game, and Investment Allocation is the amount allocated to the forum in the Investment Game. Observation counts vary slightly due to response rate differences and small changes in survey questions. All regressions include strata fixed effects. Standard errors reported in parentheses. \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.011$

Table A3: **Removing Observations with Poor Understanding**

	(1) Expected Usage	(2) Fund Allocation	(3) Invest Allocation
<b>Panel A: Direct Effect on State Institutions</b>			
Post Treatment	0.80*** (0.05)	15.27*** (1.33)	14.72*** (1.38)
Constant	4.13*** (0.03)	105.74*** (0.94)	116.00*** (0.97)
<b>Panel B: Indirect Effect on Non-State</b>			
Post Treatment	-0.55*** (0.04)	-10.05*** (1.33)	-12.04*** (1.40)
Constant	6.64*** (0.03)	103.06*** (0.94)	116.73*** (0.99)
Observations	3630	3716	3752
Respondents	1815	1858	1876

*Notes:* This table examines the effect of the state-positive treatment on our main outcomes, dropping observations that could potentially be low-quality. After playing each game we had both the surveyor and respondent rank the respondent's understanding of the game on a scale of 1-10. These regressions drop respondents who received below a 5 at any point. All regressions include two observations per respondent, a pre-treatment observation and a post-treatment observation. Outcomes are regressed on a post-treatment indicator. Panel A reports results for outcomes concerning the state forum, and Panel B reports the same outcomes for the non-state Forum. Expected Usage asks respondents how likely (0-10) they are to use each forum, Fund Allocation is the amount the respondent allocated to the forum in the Fund Game, and Investment Allocation is the amount allocated to the forum in the Investment Game. Observation counts vary slightly due to response rate differences across the different LHS variables Sample is restricted to those who received the state positive treatment and never received below a 5 for understanding each game. All regressions include individual fixed effects. Standard errors reported in parentheses. \* p<0.10, \*\* p<0.05, \*\*\* p<0.01

Table A4: **Robust to Outliers**

	(1) Expected Usage	(2) Fund Allocation	(3) Invest Allocation
<b>Panel A: Direct Effect on State Institution</b>			
Post Treatment	0.85*** (0.13)	17.00*** (2.75)	15.92*** (2.83)
Constant	4.01*** (0.09)	101.30*** (1.96)	112.83*** (2.02)
<b>Panel B: Indirect Effect on Non-State</b>			
Post Treatment	-0.64*** (0.13)	-9.78*** (2.71)	-12.85*** (2.73)
Constant	6.87*** (0.09)	99.21*** (1.93)	114.91*** (1.94)
Observations	3811	3981	3991
Respondents	1906	1991	1996

*Notes:* This table examines whether the effect of the state-positive treatment on our main outcomes is robust to outliers using Stata's `rreg` command. All regressions include two observations per respondent, a pre-treatment observation and a post-treatment observation. Outcomes are regressed on a post-treatment indicator. Panel A reports results for outcomes concerning the State Forum, and Panel B reports the same outcomes for the Non-State Forum. Expected Usage asks respondents how likely (0-10) they are to use each forum, Fund Allocation is the amount the respondent allocated to the forum in the Fund Game, and Investment Allocation is the amount allocated to the forum in the Investment Game. Observation counts vary slightly due to response rate differences and small changes in survey questions. \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

Stata's `rreg` command first eliminates gross outliers by performing an initial screening based on Cook's distance  $> 1$ . Then, an iteration process begins in which weights are calculated based on absolute residuals (both Huber iterations and biweight iterations are used, as suggested by Li(1985)). The iterating stops when the maximum change between the weights from one iteration to the next is below tolerance. With Stata defaults, robust regression is about 95 percent as efficient as OLS (Hamilton, 1991). Source: <http://www.ats.ucla.edu/stat/stata/dae/rreg.htm>.

Table A5: **Dropping Observations on the Boundary**

	(1) Expected Usage	(2) Fund Allocation	(3) Invest Allocation
<b>Panel A: Direct Effect on State Institutions</b>			
Post Treatment	0.72*** (0.05)	13.72*** (1.20)	14.56*** (1.22)
Constant	4.78*** (0.03)	104.64*** (0.85)	110.81*** (0.86)
<b>Panel B: Indirect Effect on Non-State</b>			
Post Treatment	-0.49*** (0.05)	-8.91*** (1.19)	-10.78*** (1.15)
Constant	5.49*** (0.03)	102.84*** (0.84)	111.96*** (0.82)
Observations	1816	2902	2934
Respondents	908	1451	1467

*Notes:* This table re-examines the effect of the state-positive treatment on our main outcomes, dropping observations that we're potentially constrained from changing their initial outcome due to the bounded nature of our outcomes. These regressions omit observations where the respondent allocated the maximum (e.g. 250 PKR in either game) or minimum (e.g. 0 PKR in either game) to the relevant forum in the baseline. All regressions include two observations per respondent, a pre-treatment observation and a post-treatment observation. Outcomes are regressed on a post-treatment indicator. Panel A reports results for outcomes concerning the state forum, and Panel B reports the same outcomes for the non-state forum. Expected Usage asks respondents how likely (0-10) they are to use each forum, Fund Allocation is the amount the respondent allocated to the forum in the Fund Game, and Investment Allocation is the amount allocated to the forum in the Investment Game. Observation counts vary slightly due to response rate differences and small changes in survey questions. All regressions include individual fixed effects. Standard errors reported in parentheses. \* p<0.10, \*\* p<0.05, \*\*\* p<0.01



Table A6: **Effect of Forum Allocation Order**

	(1) Expected Usage	(2) Fund Allocation	(3) Invest Allocation
<b>Panel A: Direct Effect on State Institution</b>			
Pre-First $\times$ Post	0.02 (0.12)	5.85* (3.50)	-1.88 (3.51)
Post-First $\times$ Post	-0.01 (0.12)	-0.36 (3.50)	-0.84 (3.51)
Post Treatment	0.63*** (0.11)	14.16*** (3.16)	15.79*** (3.17)
Constant	4.77*** (0.04)	98.78*** (1.23)	118.99*** (1.24)
<b>Panel B: Indirect Effect on Non-State</b>			
Pre-First $\times$ Post	-0.04 (0.11)	-2.98 (3.37)	-5.58 (3.74)
Post-First $\times$ Post	0.02 (0.11)	2.42 (3.37)	3.91 (3.74)
Post Treatment	-0.44*** (0.10)	-12.53*** (3.05)	-11.80*** (3.38)
Constant	6.63*** (0.04)	98.28*** (1.19)	116.46*** (1.32)
Observations	1648	1650	1650
Respondents	824	825	825

*Notes:* This table examines whether the State-Positive treatment has a differential impact depending on if respondents allocated to the court first. For a subset of our respondents we randomize the order in which they allocated to each forum, both for each game and pre/post treatment. All regressions include two observations per respondent, a pre-treatment observation and a post-treatment observation. Outcomes are regressed on a post-treatment indicator, a Pre-Court first indicator (in the pre-information provision games, the first game played was with the state actor), and a Post-Court first indicator (in the post-information provision games, the first game played was with the state actor). The coefficient on their interaction represents the effect of allocating to the court first vs. to the panchayat first. Panel A reports results for outcomes concerning the state forum, and Panel B reports the same outcomes for the non-state forum. Expected Usage asks respondents how likely (0-10) they are to use each forum, Fund Allocation is the amount the respondent allocated to the forum in the Fund Game, and Investment Allocation is the amount allocated to the forum in the Investment Game. All regressions include individual fixed effects. Standard errors reported in parentheses. \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

Table A7: **Effect of State Negative Treatment (Pilot)**

	(1) Other Usage	(2) Invest Allocation
<b>Panel A: Direct Effect on State Institution</b>		
Post Treatment	-0.68*** (0.17)	-26.49** (11.84)
Constant	3.16*** (0.12)	100.54*** (8.37)
<b>Panel B: Indirect Effect on State Institution</b>		
Post Treatment	0.54*** (0.18)	-3.24 (13.86)
Constant	7.51*** (0.13)	151.62*** (9.80)
Observations	74	74
Respondents	37	37

*Notes:* This table examines the effect of the pilot state-negative treatment on main outcomes. All regressions include two observations per respondent, a pre-treatment observation and a post-treatment observation. Outcomes are regressed on a post-treatment indicator. Panel A reports results for outcomes concerning the state forum, and Panel B reports the same outcomes for the non-state forum. Other Usage asks respondents how likely (0-10) they believe others are to use a forum and Investment Allocation is the amount allocated to the forum in the Investment Game. The fund game was not included in this pilot and respondents were not asked about their own expected usage. This treatment was only administered during the pilot phase, hence the lower observation count. All regressions include individual fixed effects. Standard errors reported in parentheses. \* p<0.10, \*\* p<0.05, \*\*\* p<0.01

Table A8: **Effect of Non-State Negative Prime**

	(1) Expected Usage	(2) Fund Allocation	(3) Invest Allocation
<b>Panel A: Indirect Effect on State Institution</b>			
Post Treatment	0.65*** (0.15)	11.97** (4.62)	6.83 (4.37)
Constant	2.81*** (0.11)	99.52*** (3.27)	110.24*** (3.09)
<b>Panel B: Direct Effect on Non-State</b>			
Post Treatment	-0.69*** (0.15)	-13.17*** (4.67)	-12.30*** (4.54)
Constant	6.97*** (0.11)	124.66*** (3.30)	130.28*** (3.21)
Observations	312	498	504
Respondents	156	249	252

*Notes:* This table examines the effect of the pilot non-state negative treatment on our main outcomes. All regressions include two observations per respondent, a pre-treatment observation and a post-treatment observation. Outcomes are regressed on a post-treatment indicator. Panel A reports results for outcomes concerning the State Forum, and Panel B reports the same outcomes for the Non-State Forum. Expected usage asks respondents how likely (0-10) they believe others are to use a forum, Fund Allocation is the amount the respondent allocated to the forum in the Fund Game, and Investment Allocation is the amount allocated to the forum in the Investment Game. This treatment was only administered during the pilot phase and the earliest round of the main surveys, hence the lower observation count. All regressions include individual fixed effects. Standard errors reported in parentheses. \* p<0.10, \*\* p<0.05, \*\*\* p<0.01

Table A9: Effect of Non-State Positive Treatment (Pilot)

	(1) Other Usage	(2) Invest Allocation
<b>Panel A: Indirect Effect on State Institution</b>		
Post Treatment	-0.33 (0.21)	-10.00 (11.13)
Constant	3.53*** (0.15)	83.33*** (7.87)
<b>Panel B: Direct Effect on State Institution</b>		
Post Treatment	0.20 (0.37)	30.00** (11.75)
Constant	7.40*** (0.26)	143.33*** (8.31)
Observations	30	30
Respondents	15	15

*Notes:* This table examines the effect of the pilot non-state positive Treatment on our main outcomes. All regressions include two observations per respondent, a pre-treatment observation and a post-treatment observation. Outcomes are regressed on a post-treatment indicator. Panel A reports results for outcomes concerning the state forum, and Panel B reports the same outcomes for the non-state forum. Other Usage asks respondents how likely (0-10) they believe others are to use a forum and Investment Allocation is the amount allocated to the forum in the Investment Game. The fund game was not included in this pilot and respondents were not asked about their own expected usage. This treatment was only administered during the pilot phase, hence the lower observation count. All regressions include individual fixed effects. Standard errors reported in parentheses. \* p<0.10, \*\* p<0.05, \*\*\* p<0.01

Table A10: **Effect of the Social Multiplier Treatment**

	(1) Expected Usage	(2) Fund Allocation	(3) Invest Allocation
<b>Panel A: Direct Effect on State Institutions</b>			
Post × Social	0.10 (0.12)	4.07 (3.49)	-2.24 (3.50)
Post Treatment	0.58*** (0.08)	14.96*** (2.44)	15.49*** (2.45)
Constant	4.77*** (0.04)	98.78*** (1.23)	118.99*** (1.24)
<b>Panel B: Indirect Effect on Non-State</b>			
Post × Social	0.03 (0.11)	-1.76 (3.37)	-0.28 (3.74)
Post Treatment	-0.47*** (0.08)	-11.92*** (2.36)	-12.47*** (2.62)
Constant	6.63*** (0.04)	98.28*** (1.19)	116.46*** (1.32)
Observations	1648	1650	1650
Respondents	824	825	825

*Notes:* This table examines whether the state-positive treatment has a differential impact depending on respondents belief about the availability of the information to society at large. A random subset of the respondents who received the state-positive treatment received an additional reminder that this information had been shared with others in their area as well. All regressions include two observations per respondent, a pre-treatment observation and a post-treatment observation. Outcomes are regressed on a post-treatment indicator, a Social-Multiplier indicator, and an indicator for their interaction. The coefficient on their interaction represents the effect of the Social-Multiplier minus the effect the standard state-positive treatment. Panel A reports results for outcomes concerning the state forum, and Panel B reports the same outcomes for the non-state forum. Expected Usage asks respondents how likely (0-10) they are to use each forum, Fund Allocation is the amount the respondent allocated to the forum in the Fund Game, and Investment Allocation is the amount allocated to the forum in the Investment Game. The observation numbers are lower as this social multiplier randomization was only introduced in a later round of surveys. All regressions include individual fixed effects. Standard errors reported in parentheses. \* p<0.10, \*\* p<0.05, \*\*\* p<0.01

Table A11: **Alternative Forum Perception Questions (Cross-Subject)**

	(1) Deliver Justice	(2) Enforce Verdict
<b>Panel A: Direct Effect on State Institutions</b>		
SP Dummy	0.07*** (0.02)	0.06** (0.02)
Constant	0.41*** (0.02)	0.57*** (0.02)
<b>Panel B: Indirect Effect on Non-State</b>		
SP Dummy	-0.06** (0.02)	-0.03 (0.02)
Constant	0.49*** (0.02)	0.37*** (0.02)
Observations	2160	2160

*notes:* This table examines the effect of the state-positive treatment on our main outcomes independent of any potentially confounding social-experimenter effect. The sample is limited to only post-treatment observations, and a regression is run of outcomes on a state-positive treatment indicator. The comparison category are respondents who received the Social-Experimenter Treatment, so the coefficient on the state-positive indicator shows the effect of SP-treatment beyond that which can be explained by the SE effect. Deliver Justice asked respondents “*Between the Panchayat and the court, which is most effective in delivering justice?*” and Enforce Verdict asked “*Between the Panchayat and the court, which is more able to enforce the verdict?*” Respondents were able to answer Court, Panchayat, both, or neither. Panel A reports effects on the share of respondents who answers only court, while Panel B shows effect on share who answered only Panchayat. Note that the two other categories (both and neither) create a middle category and allow for these variables to be non-mutually exclusive. Without these categories, the effects reported in Panel B would mechanically be equal in magnitude to the effect in Panel A but with an opposite sign. Strata fixed effects included in all regressions. Standard errors reported in parentheses. \* p<0.10, \*\* p<0.05, \*\*\* p<0.01

Table A12: **Dropping Observations with Poor Understanding (Forum Perceptions)**

	(1) Effectiveness	(2) Enforce	(3) Access	(4) Trust	(5) Allegiance
<b>Panel A: Direct Effect on State Institutions</b>					
Post Treatment	0.51*** (0.06)	0.25*** (0.04)	0.40*** (0.06)	0.58*** (0.06)	-0.02 (0.05)
Constant	4.11*** (0.04)	6.78*** (0.02)	3.10*** (0.04)	5.19*** (0.04)	7.34*** (0.03)
<b>Panel B: Indirect Effect on Non-State</b>					
Post Treatment	-0.08** (0.03)	-0.10*** (0.03)	0.13*** (0.05)	-0.11*** (0.03)	-0.18*** (0.04)
Constant	5.38*** (0.02)	6.24*** (0.02)	7.79*** (0.03)	6.49*** (0.02)	7.10*** (0.03)
Observations	1610	1606	1608	1610	1614
Respondents	805	803	804	805	807

*Notes:* This table examines the effect of the state-positive treatment on the outcomes in Table 6, dropping observations that could potentially be low-quality. After playing each game we had both the surveyor and respondent rank the respondent's understanding of the game on a scale of 1-10. These regressions drop respondents who received below a 5 at any point. All regressions include two observations per respondent, a pre-treatment observation and a post-treatment observation. Outcomes are regressed on a post-treatment indicator. Panel A reports results for outcomes concerning the state forum, and Panel B reports the same outcomes for the non-state forum. Service Effectiveness measures how effective they believe the forum to be at providing service to the ordinary man, Enforcement Effectiveness measures how effective respondents believe the forum to be at enforcing it verdicts, Access measures how difficult respondents believe it is for any ordinary man to access the forum, Trust measures how much respondents trust forum officials, and Allegiance measures how much effort the respondent would make in order to strengthen their connection to the forum. All regressions include individual fixed effects. Standard errors reported in parentheses. \* p<0.10, \*\* p<0.05, \*\*\* p<0.01

Table A13: **Dropping Observations on the Boundary (Forum Perceptions)**

	(1) Effectiveness	(2) Enforce	(3) Access	(4) Trust	(5) Allegiance
<b>Panel A: Direct Effect on State Institutions</b>					
Post Treatment	0.34*** (0.05)	0.37*** (0.05)	0.28*** (0.06)	0.54*** (0.06)	-0.10 (0.08)
Constant	5.17*** (0.04)	5.75*** (0.03)	4.73*** (0.05)	5.38*** (0.04)	6.16*** (0.06)
<b>Panel B: Indirect Effect on Non-State</b>					
Post Treatment	-0.04 (0.03)	-0.08*** (0.03)	0.15** (0.08)	-0.11*** (0.04)	-0.07 (0.05)
Constant	5.45*** (0.02)	5.72*** (0.02)	5.91*** (0.05)	5.59*** (0.03)	6.23*** (0.04)
Observations	1010	1140	434	880	476
Respondents	505	570	217	440	238

*Notes:* This table examines the effect of the state-positive treatment on the outcomes in Table 6, dropping observations that we're potentially constrained from changing their initial outcome due to the bounded nature of our outcomes. These regressions omit observations where the respondent answered the maximum (e.g. 10) or minimum (e.g. 0) to the relevant forum in the baseline. All regressions include two observations per respondent, a pre-treatment observation and a post-treatment observation. Outcomes are regressed on a post-treatment indicator. Panel A reports results for outcomes concerning the state forum, and Panel B reports the same outcomes for the non-state forum. Service Effectiveness measures how effective they believe the forum to be at providing service to the ordinary man, Enforcement Effectiveness measures how effective respondents believe the forum to be at enforcing its verdicts, Access measures how difficult respondents believe it is for any ordinary man to access the forum, Trust measures how much respondents trust forum officials, and Allegiance measures how much effort the respondent would make in order to strengthen their connection to the forum. All regressions include individual fixed effects. Standard errors reported in parentheses. \* p<0.10, \*\* p<0.05, \*\*\* p<0.01



Table A14: **Heterogeneity by State/Non-State Prior**

	(1) Expected Usage	(2) Fund Allocation	(3) Invest Allocation
<b>Panel A: Direct Effect on State Institutions</b>			
Post × Initial State	-1.69*** (0.15)	-11.80*** (4.33)	-4.31 (4.51)
Post × Initial Non-State	-0.08 (0.16)	-0.93 (4.44)	-1.12 (4.63)
Post	1.54*** (0.16)	20.99*** (4.43)	16.35*** (4.62)
Constant	4.06*** (0.03)	105.14*** (0.91)	115.83*** (0.95)
<b>Panel B: Indirect Effect on Non-State</b>			
Post × Initial State	-0.09 (0.14)	-2.32 (4.38)	-5.30 (4.50)
Post × Initial Non-State	-1.28*** (0.14)	-7.05 (4.50)	-12.06*** (4.62)
Post	0.33** (0.14)	-4.53 (4.49)	-1.96 (4.61)
Constant	6.64*** (0.03)	103.79*** (0.93)	115.48*** (0.95)
Observations	3810	3812	3812
Respondents	1905	1906	1906

*Notes:* This table examines the effect of the state-positive treatment on our main outcomes concerning the state forum. All regressions include two observations per respondent, a pre-treatment observation and a post-treatment observation. Outcomes are regressed on a post-treatment indicator, the respondent's initial belief regarding the state forum (based on expected usage), the respondents' initial belief regarding the non-state forum (based on expected usage), and an interaction between each belief and the post indicator. Panel A reports results for outcomes concerning the state forum, and Panel B reports the same outcomes for the non-state forum. Expected Usage asks respondents how likely (0-10) they are to use each forum, Fund Allocation is the amount the respondent allocated to the forum in the Fund Game, and Investment Allocation is the amount allocated to the forum in the Investment Game. All regressions include individual fixed effects. Standard errors reported in parentheses. \* p<0.10, \*\* p<0.05, \*\*\* p<0.01

## Appendix B: Full Text of Information Treatments and Selected Questions

### Main Treatment Text:

- **State-Positive Treatment:** The legal system and judges have formed a new judicial policy. This policy was introduced in Multan and has resolved 6000 pending cases in 2 months. For this reason, Multan's number of pending cases have decreased by 20%. This policy has now been implemented in Sargodha and it is estimated that most pending cases could potentially be resolved within a year.
- **Social Experimenter Treatment:** So I've been thinking about the current state of affairs and how the state's been dealing with everything and while I don't really know how great a job state institutions are doing, in my personal opinion, I really like the state system.

### Piloted Treatment Text:

- **Non-State Positive Treatment:** Panchayats have been doing a great job of resolving people's issues at the local level. From our research we have found that panchayats manage to resolve 80% of cases where most cases are resolved in a week. We also found that a third of all cases in panchayats are resolved without any cost while in the remaining cases the costs are also very low, mostly around Rs. 1,000. 80% of people also believe that panchayats are very good at bringing justice to people and these people are very satisfied with their panchayats.
- **Non-State Negative Treatment:** Through our research we have found that in Kasur district, the number of disputes arising across baradaris and across villages has increased significantly. These disputes primarily relate to land, family or water theft issues. Our research has shown that where in the past the panchayat was 70% effective in resolving disputes; now that percentage has dropped to 35%. The reason for this is that disputes across baradaris and villages are not so easily resolved through panchayats. Most often, these disputes are not resolved, and even those cases where a resolution is reached, it is near impossible to implement/enforce the decision of the panchayat. Due to these reasons the number of disputes are increasing, it is taking longer, and no action has been taken to tackle this problem.
- **State Negative Treatment:** At the moment, the district courts of Okara there are 16,241 pending cases that have been lying dormant for a long time. Even if every judge in the district were to drop all new cases, it would take them 13 years to resolve these pending cases at the current rate.

### Perception Question Text:

- **Self-Usage:** What is the likelihood of you going to [STATE/NONSTATE]? 0 meaning not at all, and 10 meaning completely.
- **Other-Usage:** What do you think is the likelihood of others in your area going [STATE/NONSTATE]? 0 meaning not at all, and 10 meaning completely.
- **Effectiveness:** Generally speaking, how effective do you find [STATE/NONSTATE] in providing services to the ordinary man? 0 meaning not at all and 10 meaning completely?
- **Enforce:** Generally speaking, how effective do you find the [STATE/NONSTATE] in enforcing verdicts? 0 meaning not at all, and 10 meaning completely.
- **Access:** In your opinion, how difficult is it for an ordinary man to get access to the [STATE/NONSTATE]? 0 meaning not at all difficult, and 10 meaning completely.
- **Trust:** Generally speaking, how much do you trust [STATE/NONSTATE] officials? 0 meaning not at all and 10 meaning completely.
- **Allegiance:** How much effort would you like to make in order to strengthen these connections? 0 meaning no effort at all, and 10 meaning completely.