A Revealed Preference Approach to the Elicitation of Political Attitudes: Experimental Evidence on Anti-Americanism in Pakistan*

Leonardo Bursztyn†
Michael Callen‡
Bruno Ferman§
Ali Hasanain¶
Noam Yuchtman∥

February 23, 2014

Abstract

We develop a revealed preference method of eliciting political attitudes that overcomes many concerns associated with subjective responses to direct survey questions. We implement this method in an experiment in Pakistan that aims to understand the expression of anti-American attitudes. Following a standard personality survey, we offer subjects a bonus payment for the survey’s completion. Around one quarter of subjects are willing to forgo a 100 Rupee payment (roughly equal to one-fifth of a day’s wage) in order to avoid anonymously checking a box indicating gratitude towards the United States government for providing funds. Unbeknownst to subjects, we use their rejection of payment as an indicator of their expression of anti-American ideology; this measure, which we validate in the paper, mitigates concerns about experimenter demand effects or social stigma distorting subjects’ responses. Indeed, we find that the social context in which ideology is expressed matters: when individuals believe their choice to accept or reject the 100 Rupee payment will be observable by other subjects, they reject the payment significantly less often than in the private setting. The willingness to forgo payment is responsive to the payment size as well: significantly fewer subjects are willing to give up a 500 Rupee payment than a 100 Rupee payment. Using a discrete choice framework, we estimate an individual’s cost of publicly rejecting payment, in monetary terms, to be nearly 175 Rupees. We also estimate the fraction of individuals in our sample who would privately express anti-American views at zero monetary cost to be around 31%, quite close to what we find using a list experiment and anonymous direct survey questions. These findings both provide evidence on the determinants of the expression of anti-American views, and more generally provide a novel method for eliciting attitudes, which can be useful in other settings.

---

*We would like to thank Sylvain Chassang, Ernesto Dal Bó, Stefano DellaVigna, Paola Giuliano, Matthew Kahn, Dean Karlan, Asim Khwaja, David Levine, Aprajit Mahajan, Edward Miguel, Petra Persson, Daniel Posner, Jacob Shapiro, Nico Voigtlander, Romain Wacziarg, Reed Walker, and seminar participants for helpful comments and suggestions. Saad Gulzar provided excellent research assistance. We also thank the Center for Global Management at UCLA Anderson for financial support.

†UCLA Anderson and NBER, bursztyn@ucla.edu.
‡UCLA Political Science, mjcallen@ucla.edu.
§São Paulo School of Economics, FGV, bruno.ferman@fgv.br.
¶Lahore University of Management Sciences, hasanain@lums.edu.pk.
∥UC Berkeley Haas and NBER, yuchtman@haas.berkeley.edu.
1 Introduction

Economists have devoted increased attention in recent years to investigating individuals’ beliefs and attitudes, from preferences for redistribution, to trust, moral values, religious beliefs, and political ideology.\(^1\) Often, empirical studies of individuals’ attitudes are forced to rely on subjective, stated preference measures in response to direct questions. Such measures have well-known problems, however: especially when questions are sensitive, or responses are stigmatized, respondents may answer in ways that are socially acceptable, rather than answer truthfully.\(^2\) Even when questions are not sensitive, responses may be affected by the fact that respondents know that their attitudes are being studied; they may also be affected by a desire to answer in a way that seems to be desired by the experimenter.\(^3\) Finally, one may be concerned that responses to direct survey questions do not reveal a preference, because there is no incentive to respond in the way that one actually feels. In this paper, we develop an indirect method for eliciting political attitudes, based on revealed preference. Not only are subjects in our study faced with a financial cost to express an anti-American attitude, but they also are unaware of the fact that their attitudes are being elicited when they act to reveal their preferences. We use this method to study anti-American attitudes in Pakistan, focusing on the responsiveness of political expression to price and to social visibility.

To implement our methodology, we conducted an experiment in Pakistan in July, 2013, with 1,152 participants.\(^4\) During each experimental session, groups of Pakistani men, aged between 18 and 35, were brought into a room where they were asked to complete a standard, “Big Five”

---

\(^1\) Theoretical work on beliefs, attitudes, and identity includes Alesina and Angeletos (2005), Akerlof and Kranton (2000), Bénabou and Tirole (2006), Piketty (1995); empirical work includes Alesina and Fuchs-Schündeln (2007), Giuliano and Spilimbergo (Forthcoming), and Luttmer (2001) on preferences for redistribution, Guiso et al. (2009), Nunn and Wantchekon (2011) on trust, and Alesina and Ferrara (2002), Di Tella et al. (2007) on beliefs related to the workings of capitalist society. Recent work on ideological expression among economists includes Ponticelli and Voth (2011), who study the impact of governments’ austerity policies on political unrest and Campante and Chor (2012), who present evidence that human capital and economic conditions interact to shape political behavior.

\(^2\) Recent work eliciting potentially stigmatized attitudes includes Díaz-Cayeros et al. (2011), Coffman et al. (2013), Blair et al. (2013a), and Blair et al. (2013b). We compare our method of preference elicitation to other methods, including methods used to reduce concerns about revealing stigmatized attitudes, in section 4.2.2, below.

\(^3\) Zizzo (2010) reviews the problem of experimenter demand effects as they relate to economic experiments. Crowne and Marlowe (1964) discuss the implications for survey research and Mayo (1933) provides the classical example of subjects changing their behavior merely because they are being observed (Hawthorne effects) based on experiments in the Western Electric factory at Hawthorne, Illinois. A recent reevaluation of the Hawthorne study is presented in Levitt and List (2011).

\(^4\) All survey personnel in the field were Pakistani, and no mention was made of the involvement of American faculty in designing and analyzing survey responses.
personality survey. The intervention of interest occurred after subjects had completed the survey, unbeknownst to them. In return for completing the survey, study participants could elect to receive a “bonus” payment (above a show-up fee they had received upon arrival). Receiving the bonus payment required checking a box in a form that indicated (from the subject’s perspective): “I gratefully thank the [funding agency] for its generosity and I accept the bonus payment offer.” Rejecting the payment required checking a box in the same form that indicated (again, from the subject’s perspective): “I choose not to accept the bonus payment offer.” The experiment randomly varied three separate components of the form, at the individual level, in a $2 \times 2 \times 2$ design:

**The identity of the funding agency:** The funding agency was either the U.S. government or the Lahore University of Management Sciences (LUMS), a leading Pakistani university.\(^5\)

**The amount of money offered:** Subjects were either offered a bonus payment of 100 Pakistani Rupees (Rs.) or of 500 Rs.; both payments represented a sizable fraction of a day’s wage.\(^6\)

**The expectation of privacy:** Respondents were led to believe that their decisions of whether to accept the payment would be either “public” or “private” with respect to the other subjects in the room. In the public condition, subjects were told: “If you choose to accept the bonus payment, in order to receive this additional payment, you will be asked to turn the letter in to the survey coordinator in the front of the room, so other participants will see you turn in the letter.” In the private condition, subjects were told, “If you choose to accept the bonus payment, in order to receive this additional payment, your decision will be completely private; you will simply replace the letter in envelope 2 and submit it with your other survey materials at the end of the study, so no other participants will know your choice.”\(^7\)

Subjects’ willingness to reject payment in order to avoid expressing gratitude to the U.S. government is our measure of expressions of anti-American attitudes that are held strongly enough that

---

\(^5\) Funds for bonus payments in fact came from the (public, so government-funded) University of California or from LUMS.

\(^6\) Irfan (2008) finds that in 2003–2004, the average monthly wage among Pakistani men was 4,278 Rs., or around 200 Rs. per day. Today, the daily wage for a manual worker is roughly 400–500 Rs.

\(^7\) In fact, all subjects would turn in their materials in exactly the same way: all subjects turned in their envelopes in the front of the room, and were seen doing so by other participants (as emphasized in the public condition). But, no subject’s decision regarding the bonus payment was ever observed by any other participant, because all survey materials were submitted inside subjects’ envelopes. The goal of the intervention was to manipulate subjects’ expectations, without telling them anything that was factually incorrect in order to minimize the use of deception in the study.
subjects are willing to forgo payments of a particular size. To place this decision within a simple economic framework, we consider political expression to be a function of three different components. First, political expression can be explicitly aimed at changing the world—individuals derive utility from changing (expected) outcomes for themselves, for their children, or for others whom they care about.\(^8\) Second, because many political acts occur publicly, social influence may shape individuals’ choices of political expression—this may be due to a desire to conform (Bernheim, 1994), to a desire to send a signal to a particular group, or because of the utility derived from social activity.\(^9\) Third, individuals may engage in costly political actions for reasons of ideology—they would express their views, even at some cost, even in the absence of social influence, and even in the absence of any potential impact on the world around them.\(^10\)

In our experiment, the first, “consequential” determinant of political expression is shut down, since accepting or rejecting the money offer is not likely to have any real-world impact. We use experimental variation in an individual’s private financial cost associated with rejecting the U.S. government money offer, and the social visibility of that rejection decision, to estimate, in monetary terms, the importance of the second and the third determinants of political expression in our setting: social influence and an individual’s (anti-American) ideology. Of course, subjects may wish to reject payment for reasons other than anti-Americanism, for example, because they do not want to feel indebted to another party. We thus compare subjects’ rates of rejecting money from the U.S. government to rates of rejecting money from LUMS in order to “difference out” a propensity to reject bonus payments from a relatively neutral funder.\(^11\)

It is important to highlight several virtues of our method of eliciting subjects’ ideology. First, we elicit individuals’ ideological views in a setting in which subjects are unaware of the elicitation. Not only was no subject aware of the purpose of the study, but also, the action through which

---

\(^8\)This is the case in rational voting models (Downs, 1957, Ledyard, 1984, Palfrey and Rosenthal, 1983, and Palfrey and Rosenthal, 1985), and warm glow models (Andreoni, 1990).

\(^9\)This is true even of the (often) private act of voting (DellaVigna et al., 2013, Gerber et al., 2008, Gerber et al., 2013).


\(^11\)We discuss other possible confounding factors leading to the rejection of the bonus payment in Section 4.1, below.
individuals’ preferences were revealed appeared, from the subjects’ perspective, simply to be part of the process of receiving payment for completing the survey. Because the choice of whether to accept the bonus payment does not appear to be of scientific interest to the researcher, we are able to observe subjects’ (relatively) natural behavior, reducing concerns about experimenter demand effects or Hawthorne effects (though these concerns are not completely eliminated, as subjects’ choices are still made in an artificial setting). Second, we create a setting in which a meaningful financial cost, the magnitude of which is experimentally varied across individuals, is imposed on an individual’s 

private political expression. This provides a revealed preference measure of subjects’ ideology in the absence of consequential or social reasons for expressing one’s ideology. The experimental variation in the financial cost of expressing an anti-American attitude also allows us to price individuals’ willingness to express their views. Third, by experimentally varying expectations of anonymity, we can manipulate the anticipated social costs of expressing one’s ideology, allowing us to study how social context affects political expression.

We find that when individuals express their ideology privately a significant minority—around one quarter of subjects—are willing to forgo 100 Rs. to avoid taking an action that would undermine their ideology: checking a box and thus thanking the U.S. government for its generosity.\(^\text{12}\) We also find that social context significantly affects individuals’ public expression of their ideology: when subjects believe that their decision to accept the payment will be observed by the other study participants, significantly fewer individuals reject the bonus payment—the rejection rate falls by nearly 10 percentage points.\(^\text{13}\) Next, we find that individuals’ willingness to check the box thanking the U.S. government is responsive to the size of the payment. While 25% of subjects are willing to forgo a 100 Rs. payment rather than check the box indicating gratitude toward the U.S., only around 10% of subjects are willing to forgo a 500 Rs. payment (this difference is highly statistically significant). We apply a discrete choice framework that exploits the experimental variation in the

\(^{12}\)In a sense, individuals’ expression of their ideology functions much like an element of their identity (Akerlof and Kranton, 2000); individuals are willing to pay a private financial cost rather than undertaking action that undermines their ideology (identity).

\(^{13}\)In our analysis below, we present results comparing rejection rates for the U.S. government vs. LUMS as the funding agency. The results are very similar to the raw rejection rates presented here. By differencing out LUMS rejection rates across conditions we account for rejection for reasons other than anti-Americanism and for other sources of private/public differences in rejection rates.
financial cost and social context of political expression to estimate several parameters of interest. We estimate an individual’s cost of publicly rejecting payment, in monetary terms, to be nearly 175 Rs. Next, we estimate the fraction of individuals in our sample who would privately express anti-American views at zero monetary cost to be around 31%, quite close to what we find using a “list experiment” to measure attitudes and using anonymous direct survey questions.

Survey evidence (discussed in further detail below) supports our interpretation of rejection of payment from the U.S. government as an expression of anti-American ideology. We find, first, that individuals who rejected the bonus payment from the U.S. government report significantly more negative views of the U.S. government, and of aid provided by the U.S. government, in response to anonymous direct survey questions that followed our main intervention. Importantly, individuals who rejected the U.S. payment offer are no more likely to report negative views of Japan’s government or of aid from Japan, a relatively neutral foreign country. Survey evidence, too, is consistent with our findings of “moderating” effects of public expression: individuals who reject the bonus payment from the U.S. government view the other experimental subjects—correctly—as less “extremist” than themselves. Furthermore, the moderating effects of public expression do not appear to result from our construction of artificial social contexts: moderation effects are strongest among subjects who report knowing most of the other participants in their session, suggesting that the effects of public expression arise from naturally occurring social ties.

Below, we compare our approach to other methods of eliciting sensitive or stigmatized attitudes and behaviors. A large literature in sociology, statistics, and political science describes experimental survey techniques used to address biases that can arise when respondents prefer not to confide their true views to the interviewer. Our methodology offers two advantages over these other approaches. First, experimental survey techniques require a large sample in order to estimate a

---

14 While this moderating effect of the majority may be specific to our study, the finding that private ideological preferences and public expression can differ has implications for understanding ideology, and its expression in different social contexts. Our findings also suggest that even subjects with relatively “extreme” views can moderate the expression of their attitudes due to social pressure.

15 Warner (1965) introduced the “randomized response technique”, Raghavarao and Federer (1979) formalized the “list experiment” (also called the “unmatched count” and the “item count technique”), and Sniderman and Piazza (1993) provide, to our knowledge, the first example of an endorsement experiment. Recent papers in political science include Bullock et al. (2011a); Fair et al. (2012); Blair et al. (2013a). We discuss experimental survey techniques in further detail in Section 4.2.2, below.
quantity of interest (say the share of illegal drug users) for a population. They cannot provide reliable information for a specific individual. This severely limits their usefulness in measuring the impacts of experimental interventions on individual attitudes. Second, while these other methods should reduce social desirability bias, they may not eliminate it completely. They still require respondents to volunteer information that may be socially stigmatized, and respondents have no incentive to truthfully reveal their attitudes—beyond a desire to honestly respond to the surveyor—even if given sufficient “cover” for the expression of stigmatized attitudes.

In addition to our contributions to academic literatures on political expression and on preference elicitation methodology, our findings have important practical implications. They contribute to a growing body of empirical evidence on, and economic analysis of, social, political, and economic outcomes in Central Asia, an area of geopolitical importance. Our results indicate that there is a significant minority of literate young men who are willing to pay a sizable financial cost solely to express their ideological views. Furthermore, they suggest that individuals’ political expression in private may significantly differ from—and be much more extreme than—views expressed publicly. This suggests that bringing extremist groups into the light of day, rather than forcing members to interact in the shadows, could affect political rhetoric, and perhaps policy positions as well.

We next describe the design and implementation of our experiment in Section 2. We then present our results in Section 3, and discuss them in Section 4. We offer concluding thoughts in Section 5.

2 Experimental Design and Implementation

Our experiment was implemented in two stages: first, a set of pilot studies that served as a “proof of concept” that our design could be implemented safely and successfully; then, the main study.

16 Beath et al. (2012) study the impact of U.S. government assistance on Afghans’ views on security and on the Afghan government, NGO’s and foreign military forces. See Gentzkow and Shapiro (2004) for an overview of anti-Americanism in the Muslim world.

17 Of course, one must keep in mind the caveat that the external validity of our findings may be limited; we discuss this in detail below.
2.1 Piloting

We developed our protocol in a series of pilots. First, in November 2012, we ran a small pilot and focus group discussion with 20 undergraduate students at the Lahore University of Management Sciences (LUMS). Next, before running the full experiment, we ran a larger pilot study in the field with 143 subjects. The exercise comprised 6 separate sessions, with approximately 24 subjects per session. 71 subjects participated on June 24th, 2013, in Islamabad and 72 subjects participated on June 25th, 2013, in Peshawar. Anticipating the necessity of having Pakistanis conduct the main experiment, we used the larger pilot to train our lab coordinators, allowing us to avoid the direct involvement of any foreigners in the implementation of the main experiment.\textsuperscript{18}

Data from the pilot allowed us to refine our experimental design and to establish that we could carry out the activity safely with minimal risk to enumerators or participants. We committed in advance to using data from the pilot studies only for these purposes, and do not include them in our main analysis.\textsuperscript{19}

2.2 Timeline and Site Selection

We implemented our experiment simultaneously in three cities, Peshawar, Islamabad, and Dera Ghazi Khan, between July 7th and July 16th, 2013. We selected these dates so that half of our sessions would be completed prior to Ramadan and half would be completed during Ramadan, which began on July 11, 2013. It is important to emphasize that all of the research staff who participated in the implementation of the main study were Pakistanis from the cities where we conducted the experiment.

An important objective of our project is to measure the degree of anti-Americanism among populations directly affected by the war on terror—this is where anti-American views are likely to be of greatest importance. To access these populations, we ran our experiment in areas either directly affected by the United States-led invasion of Afghanistan (Peshawar) or in cities that have

\textsuperscript{18}Our concern was that elicitation anti-American attitudes by a team including Americans would compromise the validity of our findings.

\textsuperscript{19}Results, available from the authors upon request, were qualitatively similar.
substantial numbers of refugees from conflict-affected areas (Islamabad and Dera Ghazi Khan).\textsuperscript{20}

Peshawar and Islamabad have large Pashtun populations and Dera Ghazi Khan has a large Balochi population, which make them especially interesting locations for the study of anti-American attitudes. Pashtuns are an ethnic majority in Southern and Eastern Afghanistan and in Northern in Western Pakistan. Both the Afghan and the Pakistani Taliban draw their support primarily from Pashtuns in this region and the vast majority of the fighting related to the U.S.-led invasion of Afghanistan has happened in predominately Pashtun areas. At the time of the study, Balochistan was home to a very active secessionist movement, and the capital, Quetta, is home to the Quetta Shura which is the primary faction of the Afghan Taliban. In scouting locations for our initial pilot, we determined that direct implementation of the experiment either in rural Khyber Pakhtunhwa or in the Federally Administered Tribal Areas (FATA) involved too much risk to respondents and to enumerators, so we opted to work in large urban areas with large migrant populations, which are generally safer.

2.3 Subject Recruitment and Screening

We contracted with local survey firms to recruit men aged between 18 and 35 from neighborhoods with large migrant populations in Islamabad and Peshawar. In both cities, we asked the recruiters to target migrants from the Federally Administered Tribal Areas (FATA), Khyber Pakthunhwa (KP), and Balochistan.\textsuperscript{21} In Dera Ghazi Khan, we first selected a tehsil randomly, then selected a union council randomly, and then used a simple right-hand sampling rule to contact potential participants. We ran 22 sessions in Peshawar, 10 sessions in Islamabad, and 16 sessions in Dera Ghazi Khan (Appendix Figure A.1 presents a map of the laboratory locations).

Upon contacting a potential subject, recruiters asked him to read aloud a short script in order to verify literacy, and an additional literacy test of comparable difficulty was administered when

\textsuperscript{20}Peshawar lies between Kabul, Afghanistan, and Islamabad on the Khyber pass and is the capital of Khyber Pakhtunhwa Province (formerly Northwestern Frontier Province). Dera Ghazi Khan and Islamabad both lie close to the provincial border of Khyber Pakhtunhwa and have large migrant populations.

\textsuperscript{21}While we did not record the birth place of subjects to preserve anonymity, in these cities our recruiters drew subjects from neighborhoods primarily populated by migrants from the Swat and Malakand agencies (agencies are administrative units in FATA). Both of these agencies, located in FATA, have seen substantial levels of insurgent conflict in recent years.
a subject reached the study site. Potential subjects who failed either test were not allowed to participate. Subject literacy was crucial for our experimental design, as the entire study required subjects to comprehend printed text. Both literacy test scripts are reproduced in the Appendix.\footnote{Appendix Figure A.2 provides Urdu translations of the two literacy screening tasks.}

2.4 Enrollment

After subjects arrived at the study site, they were directed to a waiting room, provided with an informed consent form to read, and asked to wait until they were called to participate. We relied on verbal informed consent to assure subjects that personally-identifiable information on their participation and choices was not being collected. The study coordinator called subjects to enroll one at a time; subjects then received a chit with a randomly assigned subject number, between 1 and 24, from a research assistant.\footnote{Individual stations were ordered sequentially by subject number inside the lab. Subject numbers were provided in random order to reduce the chance that subjects would be acquainted with the person sitting next to them—a concern if acquainted subjects entered the study site together, and station assignments were made in a non-random order. In practice, a research assistant handed each subject a chit, numbered from 1 to 24, from a shuffled deck. The number on the chit became a subject’s participant identification number.} After receiving their number, subjects then went to the enrollment desk outside of the laboratory (Appendix Figure A.3 provides a picture of the enrollment desk). At the desk, subjects read the second literacy script aloud, and received a payment envelope with their subject number printed on it.\footnote{Only one potential subject passed the first reading comprehension test but failed the second; this subject was replaced from the pool of recruits.} After completing the enrollment procedure, a research assistant led subjects into the laboratory and seated them at the individual lab station corresponding to their subject number.

Lab stations consisted of a chair with a clipboard; laboratory materials were placed on the chairs, which were positioned approximately four feet apart to prevent subjects from observing each other’s choices (Appendix Figure A.5 provides a picture of the experiment site in Islamabad and Appendix Figure A.6 provides a picture of the experiment site in Peshawar). We randomly assigned survey versions to lab station numbers using a simple computer program (Appendix Figure A.4 provides the mapping between survey versions and lab stations). All sessions involved exactly 24 subjects, resulting in a total of 1,152 Pakistani men participating in the main study. After a session, research assistants ensured that subjects exited the building; they were bussed off site.
immediately and were not allowed to interact with other subjects waiting to participate in the study.

2.5 Experiment

At the beginning of a session, the lab director read a set of instructions aloud. After explaining the laboratory protocol, the instructor took the subjects through three specific example questions. Each subject had a printed version of these three questions. After completing the instructions, the lab director fielded questions. The director then indicated that no questions would be answered during the experiment, allowing subjects one final opportunity to ask questions before the experiment commenced. It is important to emphasize that no details were provided by the lab director regarding the payment process; research assistants were told to reveal no more than that payment for completing the study would occur at the end of the session. To increase subjects’ confidence that they would be paid, subjects were provided their show-up fee of 300 Rupees when they began the first activity in the study.

The experiment involved four separate activities, each of which required completing a form contained in a separate envelope, numbered in order. Upon the completion of an activity, subjects were instructed to close their envelope and place it below their chair before proceeding. Furthermore, they were told not to return to previously completed activities, and that subjects who did not comply would be asked to leave. The primary purpose of strictly disallowing participants from returning to previous activities was to ensure that they could not change their responses in the revealed preference activity (activity 2) after completing the stated preference activity (activity 4).

When all subjects were had completed the four activities, the lab director and research assistant collected all laboratory materials except for subjects’ chits and payment envelopes, which subjects had received upon enrollment. The envelopes (on which were written subjects’ participant identification numbers) were taken into a separate room with the laboratory materials for the calculation of payments for each subject. Payments were placed in envelopes, which were then exchanged for

25 The purpose of the three example questions was to familiarize the subjects with the kinds of multiple choice questions that they would have to answer during the survey.
26 We disallowed questions because we did not want subjects’ inquiries to contaminate the research design.
the chits handed out to subjects at the beginning of the session.

**Activity 1**

The experiment began as a standard personality survey which was contained in envelope 1 along with subjects’ “show up” payment. In addition to a few demographic questions, participants completed a Big 5 personality assessment. The version of the Big 5 assessment used by our team was adapted to use in Pakistan and validated by psychologists at Quaid-i-Azam University.

**Activity 2**

After completing the survey, subjects opened the second envelope. This envelope contained a note offering an *additional* payment (above that for showing up) in return for checking a box on a letter to the funding agency. Checking the box indicated receipt of the payment and gratitude to the agency for providing funding. This activity measures whether respondents are differentially willing to forgo a payment, depending on the funder (U.S. government or the Lahore University of Management Sciences), the amount of the payment (100 Rupees or 500 Rupees), and respondents’ perception of whether the payment would be public or private. Specifically, envelope 2 contained a piece of paper with the following information printed on it (italics indicate text that varied depending on the experimental condition, with slashes dividing the conditions):

“Thank you for completing the survey. In return for completing the survey, we can offer you a bonus payment. Funding for the bonus payment comes from \{Lahore University of Management Sciences / the U.S. government\}. We can pay you \{100 Rupees / 500 Rupees\} for completing the survey, but in order to receive the bonus payment you are required to acknowledge receipt of the funds provided by \{Lahore University of Management Sciences / the U.S. government\} and thank the funder. If you choose not to accept the payment, you will forgo the bonus payment of \{100 Rupees / 500 Rupees\}, but not the payment of 300 Rs. for your participation.”

After this passage, in half of the forms, the language emphasized that subjects’ choices would
be private, while in the other half, the language implied that subjects’ choices to accept the bonus payment would be observed by other subjects, as follows:

“If you choose to accept the bonus payment, in order to receive this additional payment, {you will be asked to turn the letter in to the survey coordinator in the front of the room, so other participants will see you turn in the letter / your decision will be completely private; you will simply replace the letter in envelope 2 and submit it with your other survey materials at the end of the study, so no other participants will know your choice}. Once you have made your decision on the next page, please place the letter into envelope 2, whether or not you chose to accept the bonus payment.”

In addition to the sheet of paper with instructions, envelope 2 contained the bonus payment acceptance/rejection letter, with the following options (the funding agency in the letter matched the agency mentioned in the instructional form):

□ I gratefully thank {Lahore University of Management Sciences / the U.S. government} for its generosity and I accept the bonus payment offer.

□ I choose not to accept the bonus payment offer.

It is worth noting that in practice, no subject’s choice of whether to accept the bonus payment was actually public. All participants turned their acceptance/rejection letter in to the survey coordinator at the front of the room, having replaced their letter into envelope 2, and submitting it with the other survey materials (note that we did not actually provide false information about what would be required of participants, as the language in both “public” and “private” conditions was literally true). The language in the “public” treatment arm was designed to suggest (without lying to subjects) that the decision to accept the bonus payment would not be private, but subjects in the “public” condition may have expected their decision to be private because they knew that the letter would be enclosed in an envelope. Because we can only imperfectly manipulate expectations of privacy, we view this exercise as providing a lower bound estimate on the effect of making the

---

27 We opted not to have any decisions made in public to ensure the safety of participants.
decision to accept the bonus payment public.

Activity 3

In activity 3, participants filled out a self response survey that began by asking subjects to guess how many of the other participants where willing to accept the bonus payment. This question was incentivized: subjects were informed that the three individuals who guessed closest to the actual number would receive an additional 300 Rupees. Next, the survey collected information on the number of other participants the respondent knew.

We then ran a “list experiment,” a method used to measure attitudes toward sensitive topics, for example, policy positions on sensitive issues. List experiments provide individual respondents with some degree of plausible deniability (“cover”) for their expression of an unpopular, embarrassing, or stigmatized view, and thus increase the likelihood that such expression will occur (though truthful expression is not incentivized).

The list experiment works as follows: first, respondents are (randomly) assigned into a control group and one or more treatment groups. Subjects in all conditions are asked to indicate the number of policy positions they support from a list of positions on several issues. Support for any particular policy position is never indicated, only the total number of positions articulated on the list that a subject supports. In the control condition, the list includes a set of contentious, but not stigmatized, policy positions. In the treatment condition, the list includes the contentious policy positions from the control list, but also adds the policy position of interest, which is stigmatized. The degree of support for the stigmatized position at the population level is determined by comparing the average number of issues supported in the treatment and control conditions.

In our study, we randomly assigned our subjects to a control group or to one of two treatment groups, with each group containing 384 subjects. In the control condition, we asked respondents:

The following are four policies some government officials express support for. Please report HOW MANY of the four you support. You do not need to indicate which ones you support, just how many.
• Providing the poor with free electricity generators

• Establishing an independent state in Kashmir that is not part of India and not part of Pakistan

• Ensuring that civilians (President or Prime Minister) control the military

• Reducing the number of people eligible for the Benazir Income Support Program, but increasing payments to those eligible.\textsuperscript{28}

In the treatment conditions, subjects were asked a question that is identical other than the inclusion of an additional stigmatized item. In the first treatment condition, we added the policy position:

• refusing humanitarian aid from the U.S. government.

In the second treatment group, we added the position:

• supporting the activities of Pakistan Tehreek-e-Insaf (PTI).\textsuperscript{29}

Activity 4

Envelope 4 contained another survey, which asked subjects direct questions to elicit their stated preference support for: (i) aid provided by the Japanese government to Pakistan; (ii) the Japanese government overall; (iii) aid provided by the United States; and (iv) the United States government overall. We also asked a question regarding willingness to take risk using a simple Likert scale approach; we asked about subjects’ political awareness; and, about their support for Japan and the U.S. \textit{relative} to other subjects in the room.

Payments

When all subjects had completed the survey, they were asked to come, in order of their subject number, to the front of the room. They gave their payment envelope and materials packet to the session coordinator and were asked to return to their seat to await payment. After collecting all 24

\textsuperscript{28}The Benazir Income Support Program is a popular targeted unconditional cash transfer program.

\textsuperscript{29}PTI is the most anti-American of the major parties in Pakistan.
packets, two research assistants went into a separate room and calculated total subject payments. The payments were sealed in an envelope, with the cash payments wrapped in a thick debriefing handout so that subjects could not tell how much each had been paid. This was important to ensure that subjects could not be identified as having taken the U.S. government funding offer based on the thickness of the payment envelope.

Subjects were then called to the front of the room, were paid by providing their chit with the subject number on it in exchange for the payment, and were sent out of the lab into a waiting bus—there were no opportunities for subjects who had completed the study to communicate with subjects who had not yet participated. As soon as all subjects were paid and had exited, the subsequent session began immediately.

3 Empirical Analysis

This section presents our core empirical results. We first present the experimental subject’s decision viewed through the framework of a simple utility expression; this will help structure and clarify our empirical analysis of the experiment. Second, we explore the role of social pressure by examining differences in rejection rates between subjects in the private and public conditions. Third, we study the sensitivity of private ideological expression to the size of the payment offer. We then use the experimental variation in the financial cost of political expression and in the perception of privacy to estimate parameters in the utility expression. This allows us to predict private rejection rates of the U.S. government offer at values different from the offers in our experiment (i.e., different from 100 Rs. and 500 Rs.). We provide an estimate of rejection rates at zero financial cost, and also estimate the cost at which virtually everyone would accept the U.S. government offer. We also calculate the cost, in monetary terms, of expressing anti-American attitudes in public.

3.1 A Simple Framework

Let us assume that individual $i$ derives utility from expressing attitude $j$ through three channels. First, expressing attitude $j$ may provide an individual with utility for instrumental reasons; that is,
because expression changes the world (for individual $i$ or for others) in ways that bring individual $i$ utility. Second, individual $i$ might derive utility for *intrinsic* reasons, directly from the act of expressing attitude $j$. Finally, expression might provide utility for *social* reasons when it is observed by others; individual $i$ may derive additional utility or disutility from the public expression of attitude $j$, compared to its private expression (because public expression allows an individual to send a signal to others, because group expression may result in a different experience, etc.). These components of the utility function of individual $i$ can be expressed as follows:

\[
u_i = a \times \Pr(\text{expressing } j \text{ consequential}) \times 1[\text{express } j]
+ b \times 1[\text{express } j]
+ c \times 1[\text{express } j] \times 1[\text{public expression}]
+ \varepsilon_i
\]  

(1)

where $\Pr(\text{express } j \text{ consequential})$ is the probability that expressing $j$ changes the world in a way that brings utility to individual $i$, and $\varepsilon_i$ is an individual-specific preference parameter. The attitude $j$ of interest is anti-Americanism, which is expressed by subjects when they reject the U.S. bonus payment offer. In our experiment, we provide a context in which $\Pr(\text{express } j \text{ consequential}) \approx 0$. Ticking the box was likely to have been viewed by subjects as having negligible real world consequences that might affect utility for *instrumental* reasons. Experimental variation will allow us to estimate parameters $b$ (utility for *intrinsic* reasons) and $c$ (utility for *social* reasons).

In our application, the expression we consider is the rejection of a payment from the United States. If a subject is anti-American, then $b > 0$; the individual derives some positive utility from rejecting payment. If social pressure exacerbates this effect, then $c > 0$ (this would be the case, for example, if the individual felt especially good about signaling to others his anti-American attitude), and if it moderates this effect, then $c < 0$ (this would be the case if the individual was embarrassed about his anti-American views). If a subjects is pro-American, then $b < 0$ (he loses utility from expressing an anti-American view), and if social pressure exacerbates (moderates) this effect then
c < 0 (c > 0). We estimate these parameters in Section 3.6, but first look at reduced form results from the experiment.

### 3.2 Sample Characteristics and Basic Results

Table 1 presents the characteristics of our experimental sample. One can see that all of our participants were men, which was by design. In addition, participants were, on average, young and relatively well-educated. The latter is again by design, as literacy was required to implement our study. Around one half of the subjects were engaged in an economic activity at the time of the intervention. Around two-thirds of subjects were Pashtun, and about 10 percent were either Punjabi or Baluchi. Table 2 displays the sample sizes in each treatment cell and Table 3 compares the characteristics of subjects across experimental conditions. We find that respondent characteristics, including demographics, education levels, and Big 5 personality traits are balanced across conditions.

Figure 1 depicts our core outcomes of interest: rejection rates of the bonus payment offered in activity 2 by donor and payment amount, for decisions that were expected to be private, then for those expected to be public.\(^{30}\) In the top figure (1.A), we see that subjects offered a payment of 100 Rupees in the private condition are significantly more likely to reject a U.S. government payment than a payment from LUMS. However, for subjects who received a 500 Rupee offer, also in the private condition, there is only a minor and statistically insignificant difference in rejection rates between the U.S. government and LUMS offers. When subjects believe that their decisions will be made public, seen in the bottom figure (1.B), both price and the identity of the funder impact the decision much less than in private. In particular, the rejection of a 100 Rupee payment from the U.S. occurs significantly less often in public than in private, indicating that the effect of making expression public is to increase acceptance of the payment. We next analyze these results in greater depth.

\(^{30}\)Appendix Table A.1 provides rejection rates and standard errors for each treatment cell.
3.3 Measuring Private Attitudes

We begin by considering rejection rates among subjects offered the low payment option (100 Rupees) in the “private” condition. Subjects in this group will provide us with an indication of whether individuals are willing to pay a significant, but not immense cost, simply to privately express their ideological position. The first entry in the first row of Table 4 presents the raw rejection rate for the U.S. government offer among subjects offered 100 Rupee bonus payments in the private condition. In this group, 25.2% of subjects who receive the U.S. government offer chose to reject the offer.

Of course, it might be the case that some of these subjects would have rejected money from any funding agency, not only from the U.S. government. In order to account for this possibility, we present in the second entry in row 1 of Table 4 the rejection rate among subjects offered 100 Rupees in the private condition from LUMS. The proportion of subjects who rejected an offer from LUMS is 8.4%. Individuals who rejected the LUMS offer represent an estimate of the proportion of individuals who rejected the U.S. government offer not because of anti-Americanism, but because they would reject a bonus payment even from a relatively neutral entity. We can subtract this fraction from the overall rate of rejection of the U.S. government offer to find that the proportion of subjects who reject an U.S. government offer, but would have accepted an offer from LUMS is 16.8% (p<0.001).\footnote{Note that 16.8% may represent a lower bound for the fraction of people who are anti-American, as some of those who rejected LUMS might be anti-American as well. Indeed, LUMS is Pakistan’s top-ranked university, has an international orientation, and is patterned after universities in the United States. Given this, subjects may associate LUMS with the United States, biasing our results toward finding no anti-Americanism when we compare U.S. government offer rejection rates to LUMS offer rejection rates. Of course, if subjects would have rejected payment from any government, then this would be generate results consistent with those shown here. We explore whether attitudes towards foreign governments in general might drive our results, along with other alternative hypotheses, in Section 4.1, below.}

In Appendix Table A.2, we report regressions reflecting the difference in rejection rates for U.S. government and LUMS offers controlling for session fixed effects and a set of subject covariates. The estimated treatment effects and standard errors remain virtually unchanged, suggesting the implementation of the laboratory protocol across rounds and experimental sites was successful.\footnote{Implementation is of special concern in our case. As outsiders (including the co-author from Eastern Pakistan), our presence could have affected subjects’ behavior, preventing us from directly monitoring the experiment.}
3.4 The Role of Social Context

We now examine the role of social context in shaping individuals’ public expression of their ideological positions. We investigate a second dimension of randomization incorporated into our design: variation subjects’ perceptions of whether their choices to accept the bonus payment offer would be publicly observed by other participants at the end of the session.

We present our social pressure effects in the second row of Table 4. Column 1 presents the difference in rejection rates of the 100 Rs. offer from the U.S. government under the public and private conditions. The proportion of subjects who reject the U.S. government offer when they believe this information will be public is 8.2 percentage points lower than when they make this decision in the private condition (p=0.093). In column 2, we consider the same difference for subjects who received a 100 Rs. offer from LUMS—this difference provides information about the effects of making a decision regarding the bonus payment public, in the absence of a clearly ideological component of the choice (for example, being publicly identified as accepting a relatively large payment may result in embarrassment). The difference between the public and private LUMS conditions is very small, is not statistically significant (p=0.439), and actually points toward higher rejection rates in public for the LUMS offer. Therefore, the effects of social pressure we find for subjects who received an U.S. offer are not simply capturing a general increase in acceptance rate in the public condition, for instance due to a fear of looking foolish in front of others by rejecting money. In column 3, we consider the U.S. government offer public versus private difference after differencing out the rejection rates for the LUMS offer. We now estimate a 10.9 percentage point lower rejection rate for the U.S. government offer in the public condition (p=0.069). These results indicate that there exist social effects and their direction is toward moderation: fewer subjects reject the U.S. offer when they believe their choice will be made public to other participants.33

A central consideration when evaluating our estimated effects of social pressure is whether these effects are consistent with subjects’ beliefs about the views of the other subjects around them. Indeed, we find evidence that respondents were able to correctly infer the positions of others.

33 In Appendix Table A.3, we present regression results for the effect of public condition on rejection rates controlling for session fixed effects and a set of of subject covariates.
in the room regarding the U.S. government offer. The evidence we collect suggests that respondents
generally perceive other participants as politically moderate (in terms of their views on the U.S.);
they also believe that most other participants accepted the U.S. government 100 Rs. offer. This
is true even for those who originally reject the offer: they are aware that they are a minority and
that most other participants are politically moderate.

To measure subjects’ beliefs about other subjects’ willingness to accept the bonus payment, we
included additional components in the study after the decision of whether to accept the bonus pay-
ment offer. After closing and sealing the second envelope (containing the letter of acknowledgment
for additional payment), subjects were asked to complete two additional envelopes with additional
questions. The third envelope in the experiment included an incentivized elicitation of individuals’
beliefs about the number of other participants (from 0 to 23) accepting the money offer in the room
(all sessions included exactly 24 participants). Among respondents who received the “private” of-
fer, with lower payments and U.S. government-funded money (whose estimates would therefore
have been closest to estimates of other subjects’ ideological positions), the average (median) guess
was that 80% (95.6%) of other participants in the room accepted the money. Among respondents
who rejected the offer above, the average guess was that 62.3% (median 87%) of other respondents
accepted the offer.

In the fourth and last envelope, subjects were directly asked to compare their views to those
of others in the room regarding: (i) the U.S. government; and (ii) accepting U.S. aid. Regarding
views of the U.S. government, among those who accepted the “100 Rs.-private-U.S. donor” payment,
17% of subjects viewed themselves as strictly more anti-American than the other respondents in
the room; among those who rejected that offer, that number goes up to 57.2%. Moreover, only
14.3% of respondents rejecting that offer report viewing others in the room as strictly more anti-
American than themselves. When we look at views on accepting U.S. aid, the numbers are very
similar: among those who accepted the offer, 18.2% view themselves as strictly more likely to refuse
U.S. aid, whereas the percentage is 58.3% among those who reject the money offer (and only 16.6%
of those rejecting the money offer view themselves as strictly less likely to refuse U.S. aid than

34Admittedly, one worries about this elicitation is affected by the decision regarding the bonus payment, so it is
best viewed as a suggestive complement to the incentivized estimates of other subjects’ behavior.
others in the room).

The results we find in the exercises contained in envelopes 3 and 4 paint a consistent picture. Rejectors of the U.S. government bonus payment offer believed that a majority of the other subjects would accept the payment and also self-identified as being in an anti-American minority. Thus, if there were among our subjects a desire to conform to the median view, then potential rejectors should have been less likely to do so in public—precisely what we find. Of course, *ex ante*, one might have hypothesized that a minority of extremists might have pushed more moderate individuals to express *more* anti-American attitudes in public. While this might occur in some settings, our finding of moderating effects of public expression are certainly of interest given the theoretical ambiguity.

Finally, it is important to consider the variation in social context arising from respondents’ familiarity with each other from interactions outside the study. We included in the third envelope a question asking subjects how many people they knew in the room, where subjects had to pick from 5 categories: no other participant, between 1 and 6 other participants, between 7 and 12, between 13 and 18, between 19 and 23 (this was asked just after subjects estimated the number of other subjects who accepted the bonus payment). We find that 58.8% of respondents reported knowing at least one other person—this suggests that although the study occurred in an artificial setting, the social environment was not entirely foreign. We can exploit the non-random variation in the number of people each participant knows in the lab session to provide suggestive evidence on the heterogeneity of social influence effects. We find that social pressure effects are stronger among individuals who reported to know *more* participants in the session. For instance, among those who reported knowing 19 or more people (nearly everyone else) in the room, public decisions reduce the difference between U.S. and LUMS rejection rates by significantly more than among individuals who know fewer than 19 others in the room (and other splits of the subject population reveal qualitatively similar difference).
3.5 Sensitivity of Political Expression to Payment Size

In Section 3.3, we showed that privately, individuals are significantly more likely to reject an offer of 100 Rs. when they are told that the money is funded by the U.S. government, as opposed to LUMS: some respondents are willing to pay a positive, financial cost to privately express their ideological position. We now examine whether individuals are sensitive to the cost of that expression. In order words, what is the price elasticity of demand for ideological expression?

In order to answer this question, we offered to a random half of respondents bonus payments of 500 Rs., instead of 100 Rs. By comparing the change in rejection rates for both money funded by the U.S. government and LUMS, we can measure individuals’ responsiveness to changes in the cost of privately expressing their ideological position. The third row of Table 4 presents our findings. Column 1 reports the difference in rejection rates of the U.S. government offer in the private condition when the bonus payment is 500 Rs. Relative to the 100 Rs., private U.S. government offer, an increase in the offer to 500 Rs. decreases the rejection rate by 15.5 percentage points, from 25.2% to 9.7% (p=0.001). The results after subtracting rejection rates for LUMS offer (column 3) indicate a decrease of 12.7 percentage points (p=0.0128). Therefore, the change is coming mostly from the offers of high payments from the U.S. government. For LUMS payment, there is virtually no change in rejection rates from 100 Rs. to 500 Rs. offers (column 2).

Figure 1 summarizes our overall findings, by plotting the raw rejection rates for the U.S. government offers privately and publicly, for both 100 Rs. and 500 Rs. payments.

3.6 Uncovering Preference Parameters

We began with a simple utility expression to describe the quantities we sought to measure in the experiment. Our findings suggests that both intrinsic utility and social utility are relevant to the decision to express preferences. In the language of our utility expression (Equation 1), b and c are both negative.

In this section, we directly estimate estimate these parameters. Subjects obtain both direct

---

35 In Appendix Table A.4, we present regression results for the effect of higher payments on rejection rates controlling for session fixed effects and a set of of subject covariates.
utility from the money they receive and and some utility from the decision to accept or reject the U.S. offer. We normalize the non-monetary benefit from accepting the U.S. offer to be equal to zero, and assume that rejecting the offer affects subjects through the *intrinsic utility* and *social utility* channels. We can write the utility levels associated with accepting ($u^a$) or rejecting ($u^r$) the U.S. offer publicly, for individual $i$:

$$u^a_i = \alpha p + \gamma p 1[public],$$

$$u^r_i = \beta_1 + \beta_2 1[public] + \epsilon^r_i.$$  \hfill (2)

The coefficient $\alpha$ measures the marginal utility of receiving money in private; $\gamma$ measures the extra marginal utility of money when receiving money in public; $\beta_1$ measures the private utility from rejecting the U.S. offer; $\beta_2$ measures the utility effect due to social influence from rejecting the offer, and $p$ corresponds to the size of the payment. We assume that $\epsilon^r$ is an individual-specific source of variation in preferences for *privately* accepting the U.S. offer, and that $\epsilon^r$ is distributed according to a logistic distribution.

We can write the public rejection decision for individual $i$ as a binary choice:

$$r_i = 1[(\alpha p + \gamma p 1[public] - \beta_1 - \beta_2 1[public] - \epsilon^a_i) < 0].$$  \hfill (3)

This setup maps to our utility expression Equation 1 above, with $\beta_1 = b$ and $\beta_2 = c$. Using a logit regression, we estimate that $b = -1.089$ and $c = -0.495$, which suggests that on average participants in the experiment are pro-American and that social pressure reinforces this position.\footnote{The standard errors corresponding to our estimates of $b$ and $c$ are 0.193 and 0.296, respectively. The p-value of the hypothesis $H_0 : b = 0$ is 0.000 and of $H_0 : c = 0$ is 0.094.}

We can also use our estimate of the private marginal utility of money to estimate both parameters in money terms: the average participant derives a benefit equivalent to 382.49 Rs. from accepting the U.S. offer, and pays a “social” cost of 173.74 Rs. from publicly rejecting the offer.

Using our logit estimation, we also predict private rejection rates of the U.S. offer at different
prices exploiting the experimental variation we observe between 100 Rs. and 500 Rs. At price zero, private rejection rates of the U.S. offer would be 30.9%. In Figure 2, we plot the predicted private rejection rates as a function of price. We observe that it would cost about 1356 Rs. to get the individual in the 99th percentile (in the distribution of willingness to pay to reject the offer) to accept our U.S. money offer.

4 Discussion

This section addresses a set of potential objections to the idea that our measure is capturing anti-American attitudes; provides a comparison of this technique with other approaches to the measurement of stigmatized attitudes from the existing literature, and discusses external validity.

4.1 Examining Alternative Explanations

In this subsection, we examine evidence related to alternative stories that could generate patterns similar to our findings.

Using Stated Views to Validate our Revealed Preference Measure. As mentioned above, following subjects’ decisions of whether to accept the bonus payment, they were asked to answer a number of questions, which included their stated views on: (i) aid provided by the U.S. government, (ii) the U.S. government overall, (iii) aid provided by the Japanese government, and (iv) the Japanese government overall. For each one of these four questions, respondents were asked to express their views by picking a number from 1 to 5, with 1 corresponding to very negative views, and 5 to very positive views. Subjects were also asked to compare their views on the four aforementioned topics relative to the other participants in the room (also on a scale from 1 to 5). The answers to the questions on stated views about U.S. aid and the U.S. government also suggest the presence of anti-American sentiment for a significant minority of the sample: 26.4% of respondents have a negative view of U.S. aid (i.e., picked either 1 or 2 as their answer to corresponding question) and 29.8% of respondents have a negative view of the U.S. government overall.

37 Japan was picked as a plausibly neutral, but still rich and foreign, funding nation.
We can also use the answers to these direct survey questions to validate our revealed preference ideology measure. The results are presented for respondents receiving private offers of 100 Rs. in Table 5.\(^{38}\) In column 1 from Panel i, we observe that respondents that privately reject the U.S. money offer are 68.3 percentage points more likely (significantly at the 1\% level) than respondents who accept it to report having a negative view on U.S. aid (i.e., they pick either 1 or 2 as their answer to that question). Column 2, row 1 shows that respondents who privately reject the U.S. money offer are 58.6 percentage points more likely (also significant at 1\%) to report having a negative view on the U.S. government than those who accept the offer. Columns 3 and 4, row 1 show similar results if instead we use a dummy on whether the respondent believes he has a negative on U.S. aid and on the U.S. government relative to the other participants in the room.

It is important to note that we look at this set of results correlating revealed and stated preferences as merely suggestive, since we find evidence that participants on average report to be more pro-American after receiving the U.S. offer, and more so when the offer is of 500 Rs. as opposed to 100 Rs. These findings suggest that subjects might be “reciprocating”, in terms of stating a more positive view on the U.S. after receiving a money offer from the U.S. government.\(^{39}\)

This finding provides direct evidence that receiving an aid payment from the U.S. will increase the stated support of beneficiaries. It is highly plausible that this change merely reflects experimenter demand effects, though it could reflect a genuine change in opinions.

In columns 1-4 from Panel ii, we reproduce the analysis, but instead we look at whether there is a differential likelihood of having a negative view on U.S. aid or the U.S. government depending on whether respondents accepted or rejected the LUMS money offer. The results suggest that, if anything, individuals rejecting the LUMS offer have a more positive view on U.S. aid and the U.S. government. This gives us extra confidence that we are not simply showing that respondents who reject money offers are more likely to report being anti-American.

In columns 5-8, we present another set of placebo exercises. If our revealed preference measure is indeed capturing an anti-American ideology, one would not expect the rejection of the U.S.

\(^{38}\)In Appendix Figure A.7, we replicate Table 5 for other subgroups of our treatment conditions.

\(^{39}\)The correlation between receiving a U.S. offer and stating positive views about the U.S. would only be problematic for our validation if it interacted with whether or not the participant accepted the U.S. offer. Unfortunately, we do not have data to address this potential concern.
money offer to correlate with views on another, foreign country that is viewed more neutrally. To assess this hypothesis, we also asked individuals their views on Japanese aid and on the Japanese government. We observe that accepting or rejecting the U.S. money offer does not correlate with having a negative view on Japanese aid or the Japanese government. Furthermore, accepting or rejecting the LUMS offer also does not correlate with views on Japan.

We also conducted a list experiment in which we asked subjects to report how many of a series of statements they agreed with, but not which statements they agreed with. This provides subjects with plausible deniability with respect to supporting a particular statement. In one experimental condition, subjects were presented with a “control” list of statements. In the “treatment” list, a single anti-American statement was added to the list. The difference in the average number of statements supported between experimental and control groups provides an estimate of the fraction of individuals who agree with the anti-American statement. Results from the list experiment are consistent with the rejection of the bonus payment being an expression of anti-American ideology: as we found using our revealed preference measure, the list experiment reveals a significant minority of subjects affirming an anti-American statement.  

Unfortunately, we do not have the power to identify a statistically significant difference in the treatment effect estimated in the list experiment between individuals who rejected and individuals who accepted the bonus payment from the U.S. government. However, the sign of the difference is consistent with our interpretation of the revealed preference measure.

**Distaste for accepting money offer.** Some subjects might have privately rejected the U.S. offer not because they dislike the U.S. but instead because they dislike accepting that type of offer. As discussed above, in order to deal with that, we difference out the private rejection rates from the LUMS offer. This procedure most likely generates a lower bound for our estimates since it assumes that no subject who rejected the LUMS offer would have rejected the U.S. offer for being anti-American. Moreover, we also difference out the LUMS public rejection rates from the U.S. public rejection rates to deal with factors other than anti-American sentiment that might affect the

---

40In our sample, 22.6% of respondents agreed with a government policy of refusing humanitarian aid from the U.S. government, and 43.2% agreed with a government policy of support to activities of an anti-American party focusing on this issue. This last number is considerably higher than the fraction of individuals with stated anti-American views, confirming that the cost of manifesting such views is the lowest in the list experiment, due to plausible deniability.
public decision in addition to the private decision (e.g., embarrassment to publicly accept money). As discussed below, our results are robust to subtracting LUMS rejection rates.

**Anti-Americanism vs. general anti-foreign sentiment.** The lack of correlation found in the previous subsection between rejection of the U.S. money offer and stated views regarding the Japanese government and humanitarian aid from Japan suggests that our revealed preference measure is capturing a specific ideology, anti-Americanism, as opposed to a general anti-foreign sentiment.

**“Feeling insulted” effect.** It might be that some subjects felt insulted by the bonus payment amounts, thinking that they were too small, especially coming from the U.S. and for the low payment offer. Moreover, the correlation between rejection rates of the U.S. offer and stated views on the U.S. could also be consistent with subjected feeling insulted, since the stated opinions were elicited after the payment intervention. We do not believe that “feeling insulted” could be driving our findings for a number of reasons. First, the show-up fee to participate in the experiment was 300 Rs., so participants were willing to take a bus and come participate in a survey for that amount. It seems unreasonable that they will find 100 Rs. unreasonably small just for completing the personality survey. Second, as benchmark for the amounts, as discussed before, the average day’s wage in Pakistan was about 200 Rs. in 2004 (Irfan, 2008), which also suggests that the bonus payment amounts were not small. In fact, it would unnatural to offer bonus payment amounts much larger than the amounts we offered, given that the survey subjects had just completed was not particularly long and challenging.

**Time cost to read instructions and check the box.** Our results cannot be explained by some subjects not being willing to pay the time cost to go over the bonus payment instructions and check the box. Not only were the payment amounts sufficiently large, but rejecting the money offer also required checking a box. A small number of subjects (less than 3% of our sample) did not check any box and were not paid any bonus payment. Our results are robust to dropping these observations and assigning them to either the acceptance or rejection categories (the latter being our baseline specification). Moreover, not checking any box does not correlate with receiving a U.S. offer (the p-value of the correlation is 0.6).
**Risk aversion.** We also do not believe that our findings could be explained by risk aversion. First, no signature was ever required; individual choices therefore could not be matched to the subjects who made them. To further deal with this concern, subjects’ risk aversion was also measured in envelope 4, using a five-point Likert scale. We create a dummy that is equal to one if individuals reported to be either “very unwilling” or "unwilling” to take risks (which corresponds to about 56% of the sample). First, we note that there is no effect of receiving a U.S. offer on reported risk aversion (results available upon request). We then note that individuals who are risk averse according to this definition are, if anything, more likely to reject the U.S. money offer both in our baseline case (100 Rs. payment privately) and in the pooled sample, although not significantly so (results available upon request). Finally, it is important to note that we find no effect of receiving a U.S. money offer on response rates to subsequent sensitive questions, such as stated views on the U.S. government and U.S. humanitarian aid. Response rates were always very high and virtually identical for both U.S. and LUMS offers. Since individuals could in principle stop answering the forms or simply skip a few questions, it is reassuring to observe that receiving a U.S. offer has no effect on subsequent response rate.

### 4.2 Comparison with Other Methods of Eliciting Attitudes

#### 4.2.1 Comparison with Measurements from Stated Preferences and List Experiments

As discussed above, we predict that at a financial cost equal to zero, private rejection rates of the U.S. offer would be 30.9% in our sample. We can compare this number to other estimates of the share of individuals with anti-American views in our sample, coming from direct survey questions, and from list experiments conducted after our intervention of interest. Because receiving the U.S. government bonus payment offer may have directly affected attitudes toward the U.S., in our analysis of responses to the survey questions and to the list experiments we focus on individuals who received the LUMS bonus payment offer.

First, we look at stated attitudes in response to direct survey questions. The share of partici-
pants who received the LUMS offer who report having a negative view on accepting humanitarian aid provided by the U.S. government is 33.3%; the share of those reporting a negative view of the U.S. government is 36.3%. Next, among subjects receiving the LUMS offer, we estimate from our list experiments that 22% of subjects support “refusing humanitarian aid from the U.S. government,” and 55.1% support “supporting the activities of Pakistan Tehreek-e-Insaf (PTI),” the most anti-American of the major parties in Pakistan. Thus, our estimated share of individuals expressing anti-American attitudes at zero financial cost lies between the various estimates coming from alternative elicitation methods.

4.2.2 Advantages of our Approach, and Caveats

Our method of eliciting attitudes has several advantages when compared with existing methods widely used in the social sciences, such as stated preference survey measurement, randomized response techniques, endorsement experiments, and list experiments. We briefly summarize each technique, and describe the advantages of our approach. Then, we discuss important caveats one should keep in mind regarding our methodology.

Randomized Response Techniques: The randomized response technique, first proposed by Warner (1965), allows respondents to maintain confidentiality while responding to sensitive questions. The respondent is asked to roll a die (or use some other randomization device) that only the respondent can see, and to provide the stigmatized answer (for example “yes” to a question about drug use)—regardless of the respondent’s true answer—if the die shows particular numbers (say 1 or 2). The respondent is asked to respond truthfully if the die shows other numbers (3, 4, 5, or 6). This provides the respondent plausible deniability when providing a stigmatized response. Critically, given that the probabilities of different rolls of the die are known, it is straightforward to estimate the share of a survey population that would answer affirmatively (and truthfully) to the stigmatized position.\(^{42}\) The main drawbacks of this approach, relative to our technique, is that it cannot be easily used as an outcome for evaluating the effects of an intervention targeting the

\[^{42}\text{If } S_A = p \times \mu_A + (1 - p) \text{ where } S_A \text{ is the share of recorded affirmative responses, } p \text{ is the probability of the honest response outcome and } \mu_A \text{ is the actual share of population with the stigmatized property, then we can calculate } \hat{\mu}_A = \frac{S_A - p}{p}. \text{ Assuming perfect compliance to the protocol, } \mu_A \text{ should be a consistent estimator for } \mu_A.\]
stigmatized behavior (e.g., anti-drug programs) and they still require that respondents adhere to protocols perfectly to provide a consistent estimate. Furthermore, that approach does not allow for individual-level measurements of the variable of interest, and is not an incentivized measure.

**List Experiments:** We have described the mechanics of list experiments in Section 2.5 above. Similar to randomized response techniques, the main drawbacks are that they cannot be easily used as an outcome in a regression and they require adherence to the protocol. These techniques only lessen, but do not eliminate, the problem of social desirability bias as they still require that respondents volunteer some (noisy) sensitive information. Additionally, these techniques can be sensitive to the selection of items placed on the list. Finally, list experiments also do not allow for individual-level measurements of the variable of interest and are again not incentivized.

**Endorsement Experiments:** The final technique that is increasingly being used to measure sensitive political attitudes (Blair et al., 2013a; Bullock et al., 2011b) is the endorsement experiment. This approach asks respondents agreement with a mostly non-controversial policy position (e.g., “children should be vaccinated against polio”) and experimentally varies the identity of a political group supporting that position (e.g., the Taliban or the United Nations). The main advantage of our technique over this approach is that it is not clear exactly what this approach is measuring and, again, it is difficult to use as an outcome in a regression and could still be sensitive to social desirability bias.

**Direct Stated Views Elicitation:** As observed above, our evidence suggests that stated preferences measured without other prior interventions and with privacy and anonymity could do a good job at estimating the prevalence of an attitude within a sample. However, even under these conditions, stated view elicitations are not incentivized, and one might want a revealed preference measure. Perhaps more importantly, when we study individual-level interventions, these interventions whose consequences for attitudes/views we wish to measure will often have a direct link to experimenter demand effects in the preference elicitation stage of the study. Indeed, we do observe that individuals receiving an offer from the U.S. are more likely to express a more pro-American view when asked directly, especially when they receive the 500 Rs. offer. Our approach, by making individuals’ preference-revealing actions occur in a state of unawareness, avoids this problem. To
the extent that the action was not completely natural, and so there may have been some awareness of (or beliefs about) the experimenters’ intent, our method at least has the advantage of making conformity to the experimenter’s “demand” costly (assuming that the direction that seems to be desired is toward rejecting money).

However, our method has some disadvantages. Some machinery is involved, which will certainly not be appropriate for measuring many attitudes or political preferences of interest. Depending on funding availability and willingness to use deception, the approach may be very difficult to apply for certain political entities of interest (for example, if one wished to study attitudes toward the American Nazi Party). The privacy of actions in our study was crucial, and required individuals to be literate to complete the study; this limits our method to use on literate populations. Finally, one could not measure multiple attitudes for the same individual using this method.

4.3 External Validity

Our sample consists of literate, young men, and is therefore not representative of Pakistan’s population as whole. However, our sample involves a broad representation from across Pakistani ethnic groups, drawn from three distinct study sites, and we find the same patterns of results across all main ethnic groups represented and across all three sites (results available upon request). Therefore, although our sample is not representative, our results may broadly hold across a range of literate, young men.

Perhaps our most surprising finding is that moderation of expression can be achieved via social pressure. While public expression may not always be more moderate than private expression, it is worth emphasizing that this result seems to arise in our context from natural social ties: social pressure effects toward moderation are particularly strong in our study for the subjects who reported knowing most other participants in their session. This suggests that our moderation effects were most likely not a consequence of social networks artificially created in the lab.

We also emphasize that the attitude expressed in our setting may have different characteristics from other ideologies. It may be that public expression in some settings leads to greater extremism. It may also be that attitude towards the U.S. is more or less price elastic than other ideologies.
However, our method of eliciting subjects’ ideology is not restricted to our setting and could be used to study political attitudes and the expression of those attitudes in a broad range of other contexts.

5 Conclusion

Understanding the expression of anti-American ideology is a challenging task, not least because the measurement of a potentially sensitive attitude is fraught with difficulties. We have presented a novel methodology for eliciting individuals’ ideological positions that offers advantages over both direct survey questions and survey experimental techniques used to encourage truthful revelation. Relative to direct survey questions, our method elicits subjects’ ideological views without directly asking about them, and without subjects’ being aware of it—this reduces concerns about untruthful response, experimenter demand, or Hawthorne effects. While other methods, for example randomized response techniques, list experiments, and endorsement experiments, can provide estimates of sensitive attitudes for a population, they cannot do so for an individual. This can limit their use as an outcome in experiments aimed at measuring the effects of interventions which target individuals. Moreover, our approach is incentive-compatible in the sense that respondents must forgo a payment to express their political position, making the expression of ideology in our study a revealed preference.

Using our method, we show that a significant minority of Pakistani men in our sample are willing to forgo a sizable payment simply to avoid checking a box that affirms gratitude toward the U.S. government for providing the funds. This behavior is private, and is unlikely to be of “real world” consequence suggesting that rejection of payment is an expression of anti-American ideology. Experimental variation in the financial cost of expressing anti-American attitudes allows us to estimate individuals’ willingness to pay to express their views: we find that expression is, indeed, sensitive to price. In addition, experimental variation in perceptions of privacy points to an important social component to ideological expression, with public expression in our setting appearing much more moderate than private expression.
These results can assist our understanding of an important ideological current in a pivotal part of the world. Our results suggest that even individuals with extreme views might suppress those views in some social contexts, a finding that may have important policy implications. More generally, understanding where political and social values and beliefs come from, and how and why they are expressed, is of interest to a broad range of social scientists. Our findings suggest that ideological expression is partly a function of the private economic costs and benefits of that expression, as well as of the social environment in which it occurs.

Beyond this setting, our methodology could be used to elicit views of a variety of organizations, from religious organizations, to international organizations, to corporations. We believe that the further development and implementation of credible attitude elicitation methods is important for future research on beliefs, values, and ideology.
References


Figures and Tables

Figure 1: Rejection rates by donor, payment amount, and privacy condition

Figure 1.A: Private decision

Figure 1.B: Public decision

Notes: The top figure (1.A) presents raw rejection rates (on the x-axis) and 95% confidence intervals among subjects who received an offer of bonus payment in the private condition, by donor and by payment amount. The bottom figure (1.B) presents raw rejection rates among subjects who received an offer of bonus payment in the public condition, by donor and by payment amount.
Notes: This figure presents predicted private U.S. offer rejection rates as a function of prices using our logit model from Subsection 3.6.
<table>
<thead>
<tr>
<th>Variable:</th>
<th>Mean (1)</th>
<th>Standard Deviation (2)</th>
<th>p10 (3)</th>
<th>p90 (4)</th>
<th>N (5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Currently engaged in economic activity?</td>
<td>0.504</td>
<td>0.500</td>
<td>0</td>
<td>1</td>
<td>1,121</td>
</tr>
<tr>
<td>Age</td>
<td>23.7</td>
<td>5.0</td>
<td>19</td>
<td>32</td>
<td>1,072</td>
</tr>
<tr>
<td>Gender (male=1)</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1,152</td>
</tr>
<tr>
<td>Single</td>
<td>0.692</td>
<td>0.462</td>
<td>0</td>
<td>1</td>
<td>1,093</td>
</tr>
<tr>
<td>Years of education</td>
<td>11.9</td>
<td>2.8</td>
<td>9</td>
<td>16</td>
<td>1,085</td>
</tr>
<tr>
<td>Ethnic groups</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Punjabi</td>
<td>0.101</td>
<td>0.301</td>
<td>0</td>
<td>1</td>
<td>1,063</td>
</tr>
<tr>
<td>Pashtun</td>
<td>0.641</td>
<td>0.480</td>
<td>0</td>
<td>1</td>
<td>1,063</td>
</tr>
<tr>
<td>Baluchi</td>
<td>0.091</td>
<td>0.288</td>
<td>0</td>
<td>0</td>
<td>1,063</td>
</tr>
<tr>
<td>Urdu</td>
<td>0.005</td>
<td>0.068</td>
<td>0</td>
<td>0</td>
<td>1,063</td>
</tr>
<tr>
<td>Seraiki</td>
<td>0.007</td>
<td>0.081</td>
<td>0</td>
<td>0</td>
<td>1,063</td>
</tr>
<tr>
<td>Religion</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shia</td>
<td>0.053</td>
<td>0.224</td>
<td>0</td>
<td>0</td>
<td>1,057</td>
</tr>
<tr>
<td>Sunni</td>
<td>0.853</td>
<td>0.354</td>
<td>0</td>
<td>1</td>
<td>1,057</td>
</tr>
<tr>
<td>Muslim (unspecified)</td>
<td>0.076</td>
<td>0.265</td>
<td>0</td>
<td>0</td>
<td>1,057</td>
</tr>
<tr>
<td>Big 5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Openness to experience</td>
<td>3.020</td>
<td>0.424</td>
<td>2.500</td>
<td>3.583</td>
<td>1,147</td>
</tr>
<tr>
<td>Conscientiousness</td>
<td>4.110</td>
<td>0.563</td>
<td>3.417</td>
<td>4.750</td>
<td>1,146</td>
</tr>
<tr>
<td>Extraversion</td>
<td>3.590</td>
<td>0.512</td>
<td>2.917</td>
<td>4.250</td>
<td>1,147</td>
</tr>
<tr>
<td>Agreeableness</td>
<td>3.805</td>
<td>0.566</td>
<td>3.083</td>
<td>4.500</td>
<td>1,146</td>
</tr>
<tr>
<td>Neuroticism</td>
<td>2.901</td>
<td>0.530</td>
<td>2.200</td>
<td>3.583</td>
<td>1,148</td>
</tr>
<tr>
<td>Visibly religious</td>
<td>0.239</td>
<td>0.426</td>
<td>0</td>
<td>1</td>
<td>1,152</td>
</tr>
</tbody>
</table>

Notes: this table presents the mean, standard error, 10th percentile, 90th percentile, and number of observations for each variable.
### Table 2: Sample Size by Treatment Cells

<table>
<thead>
<tr>
<th>Donor</th>
<th>Payment</th>
<th>Private</th>
<th>Public</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Low</td>
<td>144</td>
<td>144</td>
<td>288</td>
</tr>
<tr>
<td>LUMS</td>
<td>High</td>
<td>144</td>
<td>144</td>
<td>288</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>288</td>
<td>288</td>
<td>576</td>
</tr>
<tr>
<td></td>
<td>Low</td>
<td>144</td>
<td>144</td>
<td>288</td>
</tr>
<tr>
<td>U.S. Government</td>
<td>High</td>
<td>144</td>
<td>144</td>
<td>288</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>288</td>
<td>288</td>
<td>576</td>
</tr>
<tr>
<td></td>
<td>Low</td>
<td>288</td>
<td>288</td>
<td>576</td>
</tr>
<tr>
<td>Total</td>
<td>High</td>
<td>288</td>
<td>288</td>
<td>576</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>576</td>
<td>576</td>
<td>1152</td>
</tr>
</tbody>
</table>

Notes: this table presents the number of observations in each treatment cell.
Table 3: Covariates Balance

<table>
<thead>
<tr>
<th></th>
<th>LUMS Low payment</th>
<th>LUMS High payment</th>
<th>U.S. government Low payment</th>
<th>U.S. government High payment</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Currently engaged in economic activity?</td>
<td>0.468 (0.042)</td>
<td>0.518 (0.043)</td>
<td>0.500 (0.042)</td>
<td>0.521 (0.042)</td>
<td>0.97</td>
</tr>
<tr>
<td>Age</td>
<td>23.2 (0.4)</td>
<td>23.6 (0.4)</td>
<td>23.3 (0.4)</td>
<td>24.2 (0.5)</td>
<td>0.63</td>
</tr>
<tr>
<td>Gender (male=1)</td>
<td>1 -</td>
<td>1 -</td>
<td>1 -</td>
<td>1 -</td>
<td>-</td>
</tr>
<tr>
<td>Single</td>
<td>0.696 (0.039)</td>
<td>0.691 (0.039)</td>
<td>0.748 (0.038)</td>
<td>0.674 (0.040)</td>
<td>0.90</td>
</tr>
<tr>
<td>Years of education</td>
<td>12.1 (0.2)</td>
<td>11.9 (0.3)</td>
<td>11.5 (0.2)</td>
<td>12.0 (0.3)</td>
<td>0.55</td>
</tr>
<tr>
<td>Ethnic groups</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Punjabi</td>
<td>0.090 (0.025)</td>
<td>0.096 (0.025)</td>
<td>0.101 (0.027)</td>
<td>0.119 (0.028)</td>
<td>1.00</td>
</tr>
<tr>
<td>Pashtun</td>
<td>0.634 (0.042)</td>
<td>0.640 (0.041)</td>
<td>0.643 (0.042)</td>
<td>0.622 (0.042)</td>
<td>1.00</td>
</tr>
<tr>
<td>Baluchi</td>
<td>0.082 (0.024)</td>
<td>0.103 (0.026)</td>
<td>0.093 (0.026)</td>
<td>0.096 (0.025)</td>
<td>0.88</td>
</tr>
<tr>
<td>Urdu</td>
<td>0.022 (0.013)</td>
<td>0.000 (0.000)</td>
<td>0.000 (0.000)</td>
<td>0.007 (0.007)</td>
<td>0.17</td>
</tr>
<tr>
<td>Seraiki</td>
<td>0.015 (0.011)</td>
<td>0.000 (0.000)</td>
<td>0.008 (0.000)</td>
<td>0.015 (0.010)</td>
<td>0.13</td>
</tr>
<tr>
<td>Religion</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shia</td>
<td>0.037 (0.016)</td>
<td>0.083 (0.024)</td>
<td>0.040 (0.017)</td>
<td>0.045 (0.018)</td>
<td>0.66</td>
</tr>
<tr>
<td>Sunni</td>
<td>0.844 (0.031)</td>
<td>0.812 (0.034)</td>
<td>0.849 (0.032)</td>
<td>0.895 (0.027)</td>
<td>0.67</td>
</tr>
<tr>
<td>Muslim (unspecified)</td>
<td>0.096 (0.025)</td>
<td>0.090 (0.025)</td>
<td>0.095 (0.026)</td>
<td>0.038 (0.017)</td>
<td>0.37</td>
</tr>
<tr>
<td>Big 5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Openness to experience</td>
<td>3.062 (0.032)</td>
<td>3.072 (0.036)</td>
<td>3.021 (0.038)</td>
<td>2.990 (0.035)</td>
<td>0.56</td>
</tr>
<tr>
<td>Conscientiousness</td>
<td>4.110 (0.047)</td>
<td>4.110 (0.044)</td>
<td>4.095 (0.058)</td>
<td>4.064 (0.043)</td>
<td>0.93</td>
</tr>
<tr>
<td>Extraversion</td>
<td>3.655 (0.044)</td>
<td>3.572 (0.038)</td>
<td>3.543 (0.049)</td>
<td>3.566 (0.040)</td>
<td>0.14</td>
</tr>
<tr>
<td>Agreeableness</td>
<td>3.812 (0.047)</td>
<td>3.792 (0.045)</td>
<td>3.785 (0.049)</td>
<td>3.797 (0.040)</td>
<td>0.82</td>
</tr>
<tr>
<td>Neuroticism</td>
<td>2.902 (0.047)</td>
<td>2.952 (0.041)</td>
<td>2.898 (0.046)</td>
<td>2.880 (0.046)</td>
<td>0.89</td>
</tr>
</tbody>
</table>

Notes: this table reports the mean level of each variable, with standard error in parenthesis, for each treatment cell. For each variable, column 9 reports the p-value of a joint test that the mean levels are the same for all treatment cells.
<table>
<thead>
<tr>
<th></th>
<th>U.S. government (1)</th>
<th>LUMS (2)</th>
<th>U.S. – LUMS (3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline (private, low payment)</td>
<td>0.252***</td>
<td>0.084***</td>
<td>0.168***</td>
</tr>
<tr>
<td></td>
<td>[0.036]</td>
<td>[0.023]</td>
<td>[0.043]</td>
</tr>
<tr>
<td>Effects of social pressure (public, low payment) – (baseline)</td>
<td>-0.082*</td>
<td>0.027</td>
<td>-0.109*</td>
</tr>
<tr>
<td></td>
<td>[0.048]</td>
<td>[0.035]</td>
<td>[0.060]</td>
</tr>
<tr>
<td>Effect of high payment (private, high payment) – (baseline)</td>
<td>-0.155***</td>
<td>-0.028</td>
<td>-0.127**</td>
</tr>
<tr>
<td></td>
<td>[0.044]</td>
<td>[0.030]</td>
<td>[0.053]</td>
</tr>
</tbody>
</table>

Notes: the first row of column 1 presents the rejection rate for subjects who received U.S. 100 Rs. U.S. government offer in private condition. The second row of column 1 presents the difference in rejection rates for those who received 100 Rs. offer from U.S. government in public versus in private condition. The third row of column 1 presents the difference in rejection rates for those who received 500 Rs. versus 100 Rs. offer from U.S. government in private condition. Column 2 replicates column 1 for subjects who received an offer from LUMS instead of from U.S. government. Column 3 presents the differences between columns 1 and 2. Robust standard errors in brackets. * significant at 10%; ** significant at 5%; *** significant at 1%.
Table 5: Revealed and Stated Preferences

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Rejected</td>
<td>0.683***</td>
<td>0.586***</td>
<td>0.405***</td>
<td>0.412***</td>
<td>-0.023</td>
<td>0.069</td>
<td>-0.058</td>
<td>-0.046</td>
</tr>
<tr>
<td></td>
<td>(0.090)</td>
<td>(0.108)</td>
<td>(0.118)</td>
<td>(0.118)</td>
<td>(0.118)</td>
<td>(0.094)</td>
<td>(0.100)</td>
<td>(0.106)</td>
</tr>
<tr>
<td>N</td>
<td>139</td>
<td>141</td>
<td>135</td>
<td>135</td>
<td>140</td>
<td>140</td>
<td>140</td>
<td>141</td>
</tr>
</tbody>
</table>

Panel i: U.S. Offer

Panel ii: LUMS Offer

| Rejected       | -0.292**                          | -0.250*                                  | -0.216                                         | -0.297*                                          | 0.083                        | 0.083                                  | 0.143                                        | 0.023                                        |
|                | (0.136)                           | (0.132)                                  | (0.201)                                        | (0.162)                                          | (0.101)                      | (0.125)                                | (0.127)                                       | (0.144)                                       |
| N              | 141                               | 141                                      | 129                                           | 129                                              | 142                         | 141                                    | 137                                           | 137                                           |
| p-value        | 0.000                             | 0.000                                    | 0.026                                         | 0.004                                            | 0.597                        | 0.663                                  | 0.164                                        | 0.679                                        |

Notes: panel i reports differences in stated preference views between subjects who rejected and those who accepted the U.S. 100 Rs. private offer, and the number of observations used in this comparison. For column 1, subjects were asked about their views towards aid provided by the U.S. government to Pakistan: “very negative”, “negative”, “neither negative nor positive”, “positive”, or “very positive”. We coded a “negative views about U.S. aid” equal to 1 for subjects who answered “very negative” or “negative”. Column 2 uses a question about subjects’ views towards U.S. government in general. Columns 3 and 4 use questions about how subjects political views about U.S. aid and U.S. government in general compare to that of others in the room. Columns 5 to 8 replicate columns 1 to 4 using political views about Japan instead of U.S.. Panel ii replicates panel i for differences in stated preference views between subjects who rejected and those who accepted LUMS offer. The last line in this table reports the p-value of a test that the difference in panel i is equal to the difference in panel ii. Sample is restricted for subjects in private condition with 100 Rs. offer. * significant at 10%; ** significant at 5%; *** significant at 1%.
Appendix Figures and Tables
Figure A.1: Map of Laboratory Locations
Panel A: First Screening Test

Panel B: On-site Screening Test

Figure A.2: Urdu Versions of Screening Tests
Figure A.3: Enrollment Desk Outside of the Lab in Islamabad
Figure A.4: Survey Version to Session-Participant Number Mapping
Figure A.5: Experiment Session in Islamabad
Figure A.6: Experiment Session in Peshawar
Figure A.7: Revealed vs Stated Preferences

<table>
<thead>
<tr>
<th>Full sample</th>
<th>Low payment</th>
<th>High payment</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Privacy status</td>
<td>Privacy status</td>
</tr>
<tr>
<td></td>
<td>Pooled</td>
<td>Private</td>
</tr>
</tbody>
</table>

*Panel i: U.S. Offer*

*Panel ii: LUMS Offer*

1: Negative views about U.S. aid  
2: Negative views about U.S. government  
3: Relatively more negative views about U.S. aid  
4: Relatively more negative views about U.S. government  
5: Negative views about Japan aid  
6: Negative views about Japan government  
7: Relatively more negative views about Japan aid  
8: Relatively more negative views about Japan government

Notes: each figure presents the differences in stated preference views between subjects who rejected and those who accepted U.S. offer as in Table ?? for subjects in different treatment cells. Panel i restricts the sample to subjects who received an offer from U.S. government, while panel ii restricts the sample to subjects who received an offer from LUMS.
<table>
<thead>
<tr>
<th>Donor</th>
<th>Payment</th>
<th>Private or Public condition?</th>
<th>Private</th>
<th>Public</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td></td>
<td></td>
<td>0.084</td>
<td>0.111</td>
<td>0.098</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>[0.023]</td>
<td>[0.026]</td>
<td>[0.018]</td>
</tr>
<tr>
<td>LUMS</td>
<td>High</td>
<td></td>
<td>0.056</td>
<td>0.097</td>
<td>0.077</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>[0.019]</td>
<td>[0.025]</td>
<td>[0.016]</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td></td>
<td>0.070</td>
<td>0.104</td>
<td>0.087</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>[0.015]</td>
<td>[0.018]</td>
<td>[0.012]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0.252</td>
<td>0.170</td>
<td>0.211</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>[0.036]</td>
<td>[0.032]</td>
<td>[0.024]</td>
</tr>
<tr>
<td>US Government</td>
<td>High</td>
<td></td>
<td>0.097</td>
<td>0.148</td>
<td>0.122</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>[0.025]</td>
<td>[0.030]</td>
<td>[0.019]</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td></td>
<td>0.174</td>
<td>0.159</td>
<td>0.167</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>[0.022]</td>
<td>[0.022]</td>
<td>[0.016]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0.168</td>
<td>0.140</td>
<td>0.154</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>[0.022]</td>
<td>[0.021]</td>
<td>[0.015]</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td></td>
<td>0.077</td>
<td>0.122</td>
<td>0.100</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>[0.016]</td>
<td>[0.019]</td>
<td>[0.013]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0.122</td>
<td>0.131</td>
<td>0.127</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>[0.014]</td>
<td>[0.014]</td>
<td>[0.010]</td>
</tr>
</tbody>
</table>

Notes: this table presents the rejection rate in each treatment cell. Standard errors in brackets.
<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>US government</td>
<td>0.168***</td>
<td>0.169***</td>
<td>0.176***</td>
</tr>
<tr>
<td></td>
<td>[0.043]</td>
<td>[0.043]</td>
<td>[0.046]</td>
</tr>
<tr>
<td>Mean (LUMS offer)</td>
<td></td>
<td>0.084</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>[0.023]</td>
<td></td>
</tr>
<tr>
<td>Session FE</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Covariates</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>N</td>
<td>286</td>
<td>286</td>
<td>243</td>
</tr>
</tbody>
</table>

Notes: this table reports the coefficients of regressions of rejection on a US government donor dummy. Column 1 reports coefficients of a regression with no controls. Column 2 reports coefficients of a regression using session fixed effects. Column 3 reports coefficients of a regression including session fixed effects and a set of subject covariates. The sample in these regression include subjects who received 100 Rs. offer in the private condition. Sample size in the regression presented in column 3 is smaller due to missing values for some covariates. Robust standard errors in brackets. * significant at 10%; ** significant at 5%; *** significant at 1%.
Table A.3: **Social pressure: regression results**

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Public × US government</td>
<td>-0.109*</td>
<td>-0.107*</td>
<td>-0.142**</td>
</tr>
<tr>
<td></td>
<td>[0.060]</td>
<td>[0.060]</td>
<td>[0.065]</td>
</tr>
<tr>
<td>Public</td>
<td>0.027</td>
<td>0.028</td>
<td>0.072</td>
</tr>
<tr>
<td></td>
<td>[0.035]</td>
<td>[0.036]</td>
<td>[0.038]</td>
</tr>
<tr>
<td>US government</td>
<td>0.168***</td>
<td>0.169***</td>
<td>0.178***</td>
</tr>
<tr>
<td></td>
<td>[0.043]</td>
<td>[0.043]</td>
<td>[0.045]</td>
</tr>
<tr>
<td>Mean (LUMS offer, private)</td>
<td>0.084</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[0.023]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Session FE</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Covariates</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>N</td>
<td>571</td>
<td>571</td>
<td>488</td>
</tr>
</tbody>
</table>

Notes: this table reports the coefficients of regressions of rejection on the interaction of a public condition dummy and a US government donor dummy, a public condition dummy, and a US government donor dummy. Column 1 reports coefficients of a regression with no controls. Column 2 reports coefficients of a regression using session fixed effects. Column 3 reports coefficients of a regression including session fixed effects and a set of subject covariates. The sample in these regression include subjects who received 100 Rs. offer. Sample size in the regression presented in column 3 is smaller due to missing values for some covariates. Robust standard errors in brackets. * significant at 10%; ** significant at 5%; *** significant at 1%.
Table A.4: **Price effects: regression results**

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>High payment × US government</td>
<td>-0.127**</td>
<td>-0.127**</td>
<td>-0.121**</td>
</tr>
<tr>
<td></td>
<td>[0.053]</td>
<td>[0.053]</td>
<td>[0.055]</td>
</tr>
<tr>
<td>High payment</td>
<td>-0.028</td>
<td>-0.027</td>
<td>-0.022</td>
</tr>
<tr>
<td></td>
<td>[0.030]</td>
<td>[0.033]</td>
<td>[0.033]</td>
</tr>
<tr>
<td>US government</td>
<td>0.168***</td>
<td>0.169***</td>
<td>0.180***</td>
</tr>
<tr>
<td></td>
<td>[0.043]</td>
<td>[0.042]</td>
<td>[0.044]</td>
</tr>
<tr>
<td>Mean (LUMS offer, low payment)</td>
<td>0.084</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[0.023]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Session FE</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Covariates</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>N</td>
<td>572</td>
<td>572</td>
<td>499</td>
</tr>
</tbody>
</table>

Notes: this table reports the coefficients of regressions of rejection on the interaction of a high payment dummy and a US government donor dummy, a high payment dummy, and a US government donor dummy. Column 1 reports coefficients of a regression with no controls. Column 2 reports coefficients of a regression using session fixed effects. Column 3 reports coefficients of a regression including session fixed effects and a set of subject covariates. The sample in these regression include subjects who received an offer in the private condition. Sample size in the regression presented in column 3 is smaller due to missing values for some covariates. Robust standard errors in brackets. * significant at 10%; ** significant at 5%; *** significant at 1%.