

Casual Contact and Ethnic Bias: Experimental Evidence from Afghanistan

Luke N. Condra[†]
Sera Linardi[‡]

Abstract[∞]

What determines how contact with the out-group affects behavior? We show experimentally that casual interethnic contact in a post-conflict society can increase ethnic bias. Day laborers in Kabul, Afghanistan, were equally altruistic toward their in-group and their out-group when out-group members were not physically present. When out-group members were physically present in an environment where no guidance for interaction or explicit incentives for cooperation were given, out-group altruism was decreasing in time among those who did not speak the out-group's language, suggesting that this contact highlights differences in descent-based attributes and increases in-group identification. We provide evidence for a psychological mechanism leading to in-group bias in casual, everyday interactions in a post-conflict society, with implications for studying the nature of ethnic bias in political and economic behavior. Results suggest that interethnic interaction does not automatically improve ethnic relationships, and attention should be paid to the conditions under which interaction occurs.

[†] Assistant Professor, Graduate School of Public and International Affairs, University of Pittsburgh, Pittsburgh, PA 15260. Email: lcondra@pitt.edu.

[‡] Associate Professor, Graduate School of Public and International Affairs, University of Pittsburgh, Pittsburgh, PA 15260. Email: linardi@pitt.edu.

[∞] We thank the editor and three anonymous referees for their helpful suggestions. We are particularly indebted to Mohammad Isaqzadeh for his help in collecting the data for this project and for sharing invaluable insights with us. We also thank EMERGENCY and Emanuele Nannini for their cooperation. For their helpful suggestions we thank Claire Adida, Michaël Aklin, Michael Callen, Liz Carlson, James Fearon, Steph Haggard, Uri Gneezy, Ryan Grauer, Matthew Jackson, Malte Lierl, Laura Paler, Michael Poznansky, Amanda Robinson, William Spaniel, Meredith Wilf, and Rick Wilson. We thank Katherine Yoon for her valuable research assistance. Previous versions of this article were presented at the 4th Annual Conference of Network Science and Economics, 2017 North American ESA Annual Meeting, BABEEW 2017, the 2017 Behavioral Models of Politics Conference, and MPSA 2016. This research was approved by the University of Pittsburgh's Institutional Review Board (PRO14050489). Support for this research was provided by the Hewlett International Grant Program, the Global Studies Center, and the Central Research Development Fund, all of the University of Pittsburgh.

The effects of contact on political and economic attitudes and behavior have been widely studied in the social sciences. Much of this work is motivated by the Contact Hypothesis, which posits that under certain conditions, interpersonal contact can reduce out-group bias (Allport 1954). Meta-analyses of empirical work investigating this theory finds considerable evidence in support of its central tenet (Pettigrew and Tropp 2008), as improved racial attitudes have been seen among interethnic college roommates (Shook and Fazio 2008), racially mixed Outward Bound groups (Green and Wong 2008), or multi-day peace workshop participants (Malhotra and Liyanage 2005). However, scholars have recently pointed out two significant gaps in that literature. First, typical day-to-day interethnic contact rarely resembles the extended interpersonal interaction described above (Enos 2017), but instead is “casual” contact which, as Allport himself expected (1954, 263), may actually increase prejudice. Second, there are very few studies investigating this kind of contact in conflict or post-conflict areas, or among minorities, where past experience of hostility may affect current interactions. Since intergroup contact has arguably become the foremost strategy for reducing prejudice (Paluck, Green, and Green 2017), filling this gap has significant implications for post-conflict peacebuilding and reconciliation.

The Contact Hypothesis posits that, where cooperation is supported by custom and authority, contact is voluntary, and group members are of equal status and are pursuing joint goals, contact between the groups reduces intergroup prejudice and increases empathy toward the out-group (Allport 1954, Pettigrew 1997). Indeed, much of the recent experimental work supportive of the Contact Hypothesis takes place in settings where strong norms regarding cooperation can be maintained and relationships have time to form (e.g., an ethnically integrated school (Alexander and Christia 2011)). But we know much less about the effects of casual, involuntary, and undirected interethnic contact, such as when riding a crowded bus or waiting in line, because it is difficult to randomly expose people to non-coethnics in a natural setting and precisely measure the outcome,

especially in societies experiencing conflict. Our lab-in-field experiment in Afghanistan attempts to make a contribution by doing exactly this.

Why do people hold ethnic biases and why would these biases be negatively affected by casual contact? Sambanis and Shayo (2013) suggest that ethnic biases are partly driven by psychological motivation: emotional attachment or antipathy due to perceived similarities to (and differences from) an ethnic group, leading individuals to care about the welfare of others in that group (Tajfel and Turner 1986). Extending this framework, we argue that sharing the same physical space with strangers that are non-coethnics exposes an individual to a continuum of descent-based attributes (Chandra 2006, Horowitz 1985) that differentiate groups from one another, such as facial features, accents, scents, and gestures. Research in neurobiology shows that these attributes are well-known sensory stimuli that activate the limbic system, the part of the brain responsible for the unconscious assessment of threat and familiarity (Leukel 1976, Brück, Kreifelts, and Wildgruber 2011, Shah et al. 2001). The fact that previous exposure to such stimuli affects whether an individual will perceive them as threatening suggests it is useful to study these dynamics among ethnic minorities with some history of conflict with a majority group. Where physical proximity induces a sense of threat, casual contact could deepen the lines of difference between in-group and out-group, instead of breaking down barriers. Since much of this process happens unconsciously (Bertrand and Duflo 2017), group biases due to casual contact are observationally distinct from those that one might hold in the absence of physical exposure. And heeding Allport, we argue that the longer the exposure, the worse the effect will be.¹

¹ “Theoretically, every superficial contact we make with an out-group member could by the ‘law of frequency’ strengthen the adverse associations that we have” (Allport 1954, 264).

Our experiment investigates the experience of individuals from minority ethnic groups in post-conflict societies when, by chance, they come into contact with the majority group. Like other post-conflict societies, Afghanistan is facing the complex task of reconciling former adversaries after a history of violent interethnic conflict between the politically dominant Pashtun majority ethnic group and non-Pashtun minorities (Barfield 2010). We look specifically at out-group bias of day laborers from non-Pashtun minorities as they waited alongside Pashtun day laborers to work and get paid. We ran three treatments: *In-NoContact*, *Out-NoContact*, and *Out-Contact*. Subjects in the *In-NoContact* treatment make an anonymous charitable giving decision to benefit their in-group, while subjects in the two *Out* treatments are asked to give to their out-group. This giving decision provides a behavioral measure of out-group bias costlier than stated attitudes. Casual interethnic contact was implemented by controlling the ethnic mix in the enclosed waiting area while leaving the nature of the interaction entirely to the discretion of the subjects: no Pashtuns were present in the waiting area of the *NoContact* treatments and the waiting area of the *Contact* treatment maintains a constant ratio of Pashtuns to non-Pashtuns. Duration of casual contact is randomly allocated by retrieving subjects from the waiting area in random order.

There was no statistically significant difference in giving between the *In-NoContact* and *Out-NoContact* treatments, but giving dropped by 25% in *Out-Contact*. The difference in out-group giving in *NoContact* vs. *Contact* suggests that sharing physical space with a dominant out-group triggers a reaction among minorities in post-conflict societies that goes beyond generally held out-group biases. Furthermore, echoing Allport, we find that the reduced giving in the *Out-Contact* group was driven by waiting longer in the interethnic waiting area, which potentially increases the frequency of casual contact. However, this result is only observed among those who do not speak the out-group's language (Pashto). We investigate several possible mechanisms that might explain this heterogeneous treatment effect. After eliminating several possibilities (language as a vehicle of

positive interethnic interaction, free-riding, and day-specific shocks), we find evidence consistent with the role of language as (1) a sensory stimulus and (2) a signal of an individual's history of previous interethnic integration.

This article makes several contributions. First, in studying casual contact, we direct attention back to Allport's (1954, 262 ff.) distinctions between contact types and warning that they may have different effects. Far from arguing against the Contact Hypothesis, we simply recognize that in settings characterized by conflict like the one we study here, carefully controlled interethnic contact supported by institutions is not the norm, and that undirected, casual "sidewalk contacts" (Jacobs 1961, 51) between ethnic groups are much more prevalent. We explore the implications of this distinction here.

Second, we build on work that emphasizes the important role that casual contact plays in politics: inducing interaction that does not meet the conditions of the Contact Hypothesis (e.g., forced integration, or resettlement of ethnic communities without first altering prevailing norms), runs the risk of effecting increased out-group prejudice (Enos 2017). Our results suggest that under some conditions, and among some ethnic minorities in post-conflict settings, attitudes toward the out-group can worsen with interethnic contact because casual interactions involve exposure to stimuli that are viscerally perceived as threatening. For policy whose aim is to reconcile and improve intergroup attitudes, our results support the argument that the theory's conditions should be met for contact to achieve desired ends; we should not assume that integration by itself will have positive effects, an insight useful in the context of debates over ethnic integration or segregation following violent conflict (cf. Kaufmann 1996, Whitt 2014).

Third, we provide evidence consistent with what others have termed a psychological (i.e., non-instrumental) mechanism through which out-group bias manifests itself (Sambanis and Shayo 2013), discussed in literature on ethnic politics but for which we have scant evidence. Understanding this

psychological mechanism is crucial since it implies that in-group biases can persist even after instrumental reasons for discrimination have been removed.

Fourth, we connect psychological responses to sensory stimuli in our experiment and explain observed behavior through insights from neurobiology, an approach that we think could be fruitfully explored in other experimental work. Finally, we make a methodological contribution to experimental approaches of studying intergroup bias by highlighting how such bias can be masked in the absence of physical contact, when ethnic identity is cued via names, photos, or similar means.

Casual Contact and Out-group Bias

How does contact affect prejudicial and discriminatory attitudes and behavior toward racial and ethnic out-groups? Gordon Allport is credited with the original articulation of the Contact Hypothesis, but recognized that contact, per se, was not enough to reduce prejudice (Allport 1954, 263). Instead, he and other proponents of the theory argued that for contact to reduce prejudice, five conditions should hold: personal interaction occurs face-to-face, there is equal status between groups, there are common goals, there is no competition between groups, and there exists institutional, societal support for reducing prejudice (see review in Pettigrew 1997).

The empirical literature on the Contact Hypothesis is enormous. Pettigrew and Tropp (2006) provide a meta-analytic review of it in an effort to identify and synthesize effects of contact. They conclude that contact seems to reduce prejudice in the vast majority of studies reviewed, and there is evidence supportive of the argument from recent experimental work motivated by the theory (e.g., Carrell, Hoekstra, and West 2015, Green and Wong 2008, Shook and Fazio 2008).

However, while scholars have pointed to various reasons to question the strength and consistency of the empirical evidence in favor of the theory (Paluck, Green, and Green 2017), we focus on just one issue that motivates our study. Since the Contact Hypothesis theorizes about how extended,

directed, and frequent intergroup interactions under specified conditions can reduce prejudice, empirical research has understandably investigated mostly along these lines. This means that we know relatively little about how intergroup contact of the kind that occurs in everyday life affects attitudes and behavior. One review worries about “the prioritizing of the study of interactions occurring under rarefied conditions”, and that “the contact literature has become detached from (and sometimes irrelevant to) everyday life in divided societies” (Dixon, Durrheim, and Tredoux 2005, 697). This is particularly problematic in conflict or post-conflict areas because the vast majority of intergroup contact will not conform to the conditions of the Contact Hypothesis. Moreover, since post-conflict reconciliation often requires attention to minority group grievances (Bloomfield, Barnes, and Huyse 2003), this constitutes another gap and reason for study. We do not know the extent to which casual contact works to reduce tensions between formerly or actively warring groups.

With this in mind, we focus on what Allport (1954, 263) referred to as “casual” contact, which is “*brief*– the type that may happen by chance when groups occupy the same place, such as passing another person on the street or seeing her on the train” (Enos 2017, 15). This is distinct from more intimate relationships that require extended and frequent contact, such as acquaintance and friendship.

Why would these understudied casual, brief interactions have a different effect than extended contact? The few studies that look at brief contact have found negative effects. For example, Enos (2014) randomly exposes white Bostonians to Spanish-speaking Hispanic confederates while waiting for their commuter train and finds a statistically significant increase in whites’ support for exclusionary immigration policies toward Mexicans after three consecutive days of brief physical exposure. Similarly, Adida, Laitin, and Valfort (2016) find that increasing the number of Senegalese Muslims that are physically present in a dictator game session decreases native French donors’ offers to their Senegalese Muslim partners.

To explain why short-lived, casual contact might elicit negative responses, we begin with an argument made by Sambanis and Shayo (2013), who assert that ethnic identification involves balancing the desire to identify with a higher-status ethnic group against an assessment of one's salient similarities with (and differences from) that group. In casual contact with non-coethnics (e.g., waiting for the bus or transacting at a market), one is exposed to a continuum of visible (e.g., skin color) and non-visible (e.g., accent) differences in descent-based attributes (Chandra 2006, Horowitz 1985) but without the extended interaction necessary to uncover similarities or common goals. Neuroscientific research into parochial altruism and implicit racial bias suggests that these differences may also serve as sensory stimuli that then affect unconscious processing of threat perception. The impact of these stimuli on the individual will depend on past experience with the stimuli (Damasio 1994) and the intensity of the exposure (Metcalf and Mischel 1999), which suggests the importance of studying minorities in such contexts in post-conflict societies.

Following others who have noted the neurobiological roots for implicit biases and parochial altruism (Bertrand and Duflo 2017, Whitt 2014), we argue that casual contact with a non-coethnic can activate an otherwise dormant psychological mechanism of in-group bias. We begin with the link made in neuroscience between sensory stimuli – such as scents (Leukel 1976), sounds (Brück, Kreifelts, and Wildgruber 2011), and expressions (Shah et al. 2001) – and the activation of the limbic system, the part of the brain that has long been connected to unconscious assessment of threat and familiarity. For example, the activation of amygdala, the fear-processing center of the limbic system, is linked to implicit bias about race and ethnicity (Cunningham et al. 2004), while hormones produced in the hypothalamus have been shown to induce self-sacrifice for one's in-group welfare

and defensive aggression toward competing out-groups (De Dreu et al. 2010).² This unique function of the limbic system may have emerged from its evolutionary role, where it categorizes environmental stimuli as aversive or attractive to generate survival impulses in animals (see review in Dalgleish (2004)).

Research has shown that unconscious and conscious processing involves different areas of the brain. When the prefrontal cortex is in control, an individual is in what is labeled the “cold” state, able to engage in instrumental and strategic reasoning (Metcalf and Mischel 1999, Lieberman 2007, Shafir, Simonson, and Tversky 1993). However, when the more primitive limbic system is in control, such as when visceral cues are present,³ the individual is in the “hot” state: emotional and reactive in ways that even the individual cannot foresee (and may even oppose in his cold state). In a telling study, individuals who could smell chocolate chip cookies wagered for them more recklessly compared to individuals for whom the cookies were merely described (Ditto et al. 2006). This research suggests that individuals may behave differently toward the out-group when sensory stimuli associated with group-defining attributes are present (i.e., casual contact) vs. a situation where stimuli are not present and the out-group is cued not through contact, but through photos, names, or textual description.

These insights from neuroscientific research map well onto prevailing frameworks in comparative politics. An individual’s sensitivity to the visceral stimuli present in his immediate social

² Neuroscientists’ investigation of oxytocin revealed that giving decisions are associated with both altruism and emotionally identifying with another person (Zak, Stanton, and Ahmadi 2007).

³ Other examples include visuals that induce sexual arousal (Ariely and Loewenstein 2006). Research on the influence of visual, olfactory, and auditory stimuli on decision-making has been heavily applied to marketing.

and physical environment would dynamically affect his assessment of his similarity to (and differences from) a particular ethnic group (Tajfel and Turner 1979). This may explain why, despite the fact that some descent-based attributes are hard to change, there are short-run shifts in the salience of identity (Eifert, Miguel, and Posner 2010, Michelitch 2015). In addition, the neuroscience findings that stimuli are processed differently depending on “the particular history of [one’s] interactions with the environment” (Damasio 1994, 133) echoes Sambanis and Shayo’s (2013, 295) point that “while patterns of social identification shape conflict risk, conflict also shapes patterns of social identification,” suggesting that those who either have a history of conflict or have had less exposure to the stimuli are more likely to react strongly to this sort of in-person contact.

Experimental Setting and Design

This review of various literatures suggests that we investigate the effects of undirected, casual contact on out-group bias in a setting characterized by interethnic conflict, that this contact and behavior should mimic as much as possible conditions in a real world context, and that we be able to compare behavior in the presence of casual contact with behavior absent contact, since casual contact with the out-group may introduce visceral, multi-sensory cues that trigger the emotional (“hot”) state.

These conditions are reflected in our charitable giving experiment in Afghanistan. First, we measure ethnic minorities’ willingness to give anonymously to a charitable cause benefiting a majority that has historically dominated them. Afghanistan’s social and political history makes it a good case to investigate these dynamics. Pashtuns, which are split into different tribes, are the largest ethnic group, comprising approximately 40% of the population (Barfield 2010). There are a number of non-Pashto-speaking minority groups, including Hazaras and Tajiks, which speak Dari.⁴ While these

⁴ Three features of Afghanistan’s ethnic groups are particularly relevant for our experimental design: (1) like many countries, it exhibits interethnic subordination between “unranked” groups (see

groups are distinct from one another, the country's history is characterized by violent political conflict along Pashtun/non-Pashtun lines. Pashtuns founded the modern Afghan state in the 1880s through the repression and subjugation of non-Pashtun minorities, and controlled the Afghan state and dominated the country's politics until the civil war in 1992 (Barfield 2010). Interethnic violence escalated during the civil war, and almost all ethnic factions engaged in arbitrary killing, torture, and sexual violence against civilian members of other ethnic groups (Oxfam 2009). Second, we investigate how casual contact affects behavior under extremely suboptimal conditions for the contact hypothesis: subjects are not working to achieve common goals, there is relative inequality between groups, and institutions are weak or unable to monitor cooperation among groups. In order to retain as much resemblance as possible to the type of casual contact discussed earlier, we did not direct or encourage interaction in any way in any of our experimental conditions. Third, we manipulated the ethnic mix in the waiting area in order to observe giving with and without casual contact with members of the ethnic (out-group) majority. Giving to the out-group in both cases is compared to a baseline of giving to the in-group, which allows us to compare the size of bias against the out-group under a condition of casual contact vs. when the out-group is not present.

Our experiment takes place in Kabul, which is much more ethnically diverse than rural areas due to increased urban migration. Thus, while Kabul's neighborhoods tend to be ethnically segregated, the city's residents are more used to seeing members of other ethnic groups in public

Horowitz 1985, 25 ff.), (2) the Pashtun/non-Pashtun ethnic divide is largely synonymous with language (Pashto vs. Dari), and (3) the groups have settled in distinct regions: Pashtuns historically reside in the south and east. This collinearity of ethnicity with language and region is not unique to Afghanistan, as scholars have observed that ethnic groups congregate in geographic space (Bates 1974, Bates 1983) and that language is often synonymous with ethnic identity (Laitin 2000).

places than their rural counterparts, and Kabul's non-Pashtuns are more likely than rural non-Pashtuns to speak Pashto. These features may make interethnic relations more harmonious than elsewhere in the country, though how exactly local demographic geography moderates the effects of interventions on behavior and attitudes is difficult to predict (Enos and Gidron 2016). We reason that Kabul's uniquely multi-ethnic composition should bias against finding evidence of in-group bias running through a psychological channel, and implies that any in-group bias resulting from physical contact with the out-group would be amplified in rural Afghanistan.

The city's public markets serve as microcosms of the nature of interethnic interactions that take place throughout Kabul: though the markets themselves are ethnically heterogeneous, individuals tend to congregate in ethnically homogeneous groups within the markets and choose whether (and how) to interact with others. Our subjects are male day laborer recruited from these markets. While our experiment is conducted on a convenience sample, our subjects are drawn from the very population about which we wish to make inferences: typical Afghan men (low income, undereducated, sporadically employed in a low-skill occupation, ethnically diverse).

We recruited day laborers early in the morning as they waited for job opportunities. Our personnel approached potential subjects in the market area and asked if they would participate in a survey about the labor force in Kabul. In exchange, they were paid 500 AFN (about \$8.60 at the time, a little over a day's wage). Subjects who provided oral consent were transported to the enumeration location, where they waited in a walled-in courtyard to be paid. Enumeration of the first group of subjects began around 7 AM, and the last group of subjects had finished and exited the location by about noon, a span of approximately 5 hours. To simulate the type of contact we are interested in, we did not force or direct interactions among waiting subjects in any of the treatment conditions. Our goal was to allow participants to interact as they saw fit and thus, despite the physical proximity, social interaction could be positive, negative, both, or non-existent. We estimate that the

walled courtyard where subjects waited is roughly 1000 square meters – small enough to enable subjects to hear the language spoken by others around them, but large enough that they could choose who to sit next to and interact with while they waited.

The core of our lab-in-the-field experiment involves subjects making an anonymous charitable giving decision, which is consistent with other experiments that investigate the role of other-regarding preferences in in-group bias and view giving as informative for prospects of reconciliation in post-conflict settings (Whitt and Wilson 2007, Whitt 2014). While we wanted to exert the experimental control that lab experiments allow, we also wanted to ensure that the giving task was as close as possible to subjects’ previous experiences.⁵ While our subjects may not have encountered situations where they had to donate to anonymous individuals of varying ethnicity, they are familiar with charitable giving to nonprofits due to the Islamic norms surrounding charity (Lambarraa and Riener 2012). Hence, even though charitable donation allows for the possibility of free-riding,⁶ we cooperated with Emergency Hospital, a well-known international nonprofit that provides medical care to Afghans (Mogelson 2012), as the receiving charity. Two of Emergency’s main hospitals in Afghanistan are located in different, highly ethnically homogeneous provinces: Anabah, in the

⁵ Gneezy and Imas (2017) argue that lab-in-the-field experiments can maximize the benefits of lab and field experiments (control and applicability of results, respectively) while minimizing the costs of each (subjects not drawn from the relevant population and lack of evidence on causal mechanisms, respectively).

⁶ In other words, subjects may not give because they think that others are giving. We test for this in the Exploring the Mechanism section below. Concerns about free-riding would not have arisen had we used the dictator game, but given the background of our subjects, we felt that it would not correspond well to their everyday life.

Panjshir Valley of the north, and Lashkar-gah, Helmand, in the south. We can confidently assume that our study subjects were aware that these provinces are dominated by a particular ethnic group – non-Pashtuns and Pashtuns, respectively.⁷ By alternating the beneficiary of the contribution between these two hospitals that are run by the same nonprofit, we vary the ethnic identity of those who would benefit from the hospital’s services while holding constant the attributes of the public good itself (e.g., type of services, reputation, etc.).

We run three treatments: (1) giving to in-group with no out-group contact (*In-NoContact*), (2) giving to out-group with no out-group contact (*Out-NoContact*), and (3) giving to Out-group with contact (*Out-Contact*). The beneficiary hospital in the *In-NoContact* treatment is in a non-Pashtun, Dari-speaking province (Panjshir), while in both *Out-NoContact* and *Out-Contact* it is in the Pashto-speaking province (Helmand). In the *In-NoContact* and *Out-NoContact* treatments, we recruited only non-Pashtuns, so subjects were waiting only with other non-Pashtuns before making their contribution. In the *Out-Contact* treatment, we recruited Pashtuns and non-Pashtuns at a ratio of 1:2 (this linguistic mix was maintained in each session) and had everyone wait in the same courtyard. Subjects were called in from the courtyard in groups of 5–7 to be enumerated. As illiteracy is high in Afghanistan, we paired each subject with a coethnic enumerator to orally administer the survey and record their answers. As subjects were called, interviewers confirmed their language out loud, so group members were aware of the linguistic mix in their session. To maintain privacy, each subject–enumerator pair sat behind a privacy partition. Once all subjects in the session had finished the survey, each received payment in an envelope. Then, a solicitor told subjects that they had the (unexpected) opportunity to contribute to Emergency, by reading the text below in Dari:

⁷ It is well-known among Afghans that these regions are highly ethnically homogeneous, an attribute reflected in systematic settlement data (Wucherpfennig et al. 2011).

Thank you for completing this survey. To thank you for your participation, we would like to compensate you with 500 AFN for your time. On the table in front of you there is an envelope with 500 AFN inside. Today you also have the opportunity to donate part of your compensation to EMERGENCY in Helmand [Panjshir] Province, in the south [north]. As you know, EMERGENCY is committed to caring for the victims of war. The organization provides treatment to our Pashtun [Tajik] brothers and sisters who are victims of war in the south [north]. We invite you to contribute to the treatment of victims of war in the region. If you would like to donate to EMERGENCY, please leave the money you wish to donate in the envelope. Remove the money you wish to keep for yourself from the envelope. If you decide not to make any donations, you can leave the envelope empty. Before you leave the room, please drop the envelope in the box by the door.

The solicitor and enumerators then left the room to give the subjects privacy. Subjects decided privately how much to contribute and then left with the money they decided to keep.⁸

Due to heightened security concerns in the country at the time, it was imperative that the experiment was completed in as few days as possible. While the ideal would have been to take the day's recruits and place them randomly into one of our three different treatment conditions upon arrival at our enumeration location, this would have required two separate and identical waiting areas, an ethnically mixed one (for non-Pashtuns in *Out-Contact*) and an ethnically homogeneous one (for non-Pashtuns in *In-NoContact* and *Out-NoContact*) to retain the integrity of the treatments. Since we could not do this at our location, we alternated between *In-NoContact* and *Out-NoContact* sessions on day 1 and ran *Out-Contact* on day 2. This has two potential consequences. First, it may have affected the balance of demographics across treatment conditions (Table 1), which we control for in our regressions. Second, one might worry that the results are an artifact of day-specific shocks. However, the intensity of treatment is randomly assigned: we populated the sessions by selecting the n th person from the day's registration list of recruits, where n differed each day. Thus, subjects' wait

⁸ Envelopes were unobtrusively numbered to match surveys with contributions. The location of the exit allows no opportunity for exiting subjects to communicate with waiting subjects.

times randomly varied from minutes to several hours. This provides a way to test that differences across treatment conditions are due to exposure to treatment.

In the Exploring the Mechanism section, we lay out the observable implications of our primary mechanism, as well as alternative mechanisms (such as free-riding and day-specific shocks) and test them against the data. The evidence is not consistent with these alternative mechanisms and instead shows that the behavioral differences observed are indeed treatment effects (contact vs. no contact with the out-group).

Experimental Results

Summary statistics of key variables for our subjects are presented in Table 1 (Panel A). The average age was 33, and most were married. Only half were educated at all, and those who had received any education had, on average, 6.8 years of schooling. About half speak more than one language (average 1.6), and Pashto is the most common second language. Subjects had spent almost a decade as day laborers, earning about 4,346 AFN (\$75) per month, which provides them with \$2.50 a day for their family. The education level and earnings are in line with the best available demographic survey evidence across the entire population (Hopkins 2014), suggesting that our subjects are representative of the average Afghan on these observable characteristics.

Table 1 (Panels B, C, and D) shows subject demographics across the *In-NoContact*, *Out-NoContact*, and *Out-Contact* sessions, respectively. Statistical tests reveal that subjects' demographics across the treatments are quite similar (SI Table 1), though we find that subjects in *Out-Contact* are younger than those in both other groups. There is a slight difference in language ability and earnings between *Out-Contact* and *Out-NoContact*, but not between *Out-Contact* and *In-NoContact*. No differences are observed in education, marital status, or work experience, suggesting that overall, the recruitment process distributed subjects with similar observables across the three treatments.

Demographics of the Pashtun subjects in *Out-Contact* (Table 1, Panel E) are also similar in most attributes to the non-Pashtun subjects, with the exception that the Pashtuns are slightly less likely to have had any education.

Non-Pashtun subjects gave an average of 20 AFN in *In-NoContact*, 21.4 AFN in *Out-NoContact*, and 16 AFN in *Out-Contact*. Since Pashtun subjects in *Out-Contact* gave less than non-Pashtuns in every treatment condition (10 AFN), the average amount raised across all subjects in *Out-Contact* is much less than the other two.⁹ Figure 1 shows the kernel density plots of non-Pashtuns' contributions in the three treatment conditions. A substantial number of subjects gave nothing, especially in *Out-Contact*. The distribution of small contributions is similar in *In-NoContact* and *Out-NoContact* but starkly different in *Out-Contact*. Large contributions are not present in *In-NoContact*, but are present and distributed similarly in *Out-NoContact* and *Out-Contact*. Shapiro-Wilk tests reject the null hypothesis that the contributions are normally distributed, suggesting the need for non-parametric tests. Wilcoxon rank-sum tests show that while non-Pashtuns' giving in *Out-NoContact* is not statistically different from their giving in *In-NoContact*, their giving in *Out-Contact* is different from *In-NoContact* and *Out-NoContact* (Table 2). This is true whether we use individuals or sessions as the unit of analysis. Kolmogorov Smirnov tests return similar results.

As the dependent variable is censored from below at zero, it is recommended that analysis of charitable giving account for this fact (Hill, Griffiths, and Lim 2011, 615). Table 3 therefore displays estimates of the main treatment effects from a tobit regression; the results are qualitatively similar with an OLS regression (SI Table 3). The base model in column 1 regresses an individual's contributed amount on the treatment dummy variables of interest (*Out-NoContact* and *Out-Contact*) (SI Table 2). Only two controls are included: enumerator fixed effects and a 'session' variable that

⁹ The lower contribution could be due to the Pashtuns' relatively lower earnings (Table 1).

indicates the length of the waiting time. Since the longer subjects wait in the courtyard, the less chance they have of finding employment opportunities that day, ‘session’ serves as a rough proxy of the opportunity cost of participating in the experiment. Consistent with others’ findings that contributions in experiments decrease in later rounds (Alexander and Christia 2011, Chuah et al. 2014), we find that the ‘session’ variable is negative and significant and so we include it in all models. Heteroskedastic robust standard errors are clustered at the session level to account for within-session correlation in all regressions.

Column 1 shows that there is no statistically significant difference in the amount contributed between *In-NoContact* and *Out-NoContact*, but there is a substantively and statistically significant drop in contribution size in *Out-Contact* of 12.53 AFN ($p < 0.05$). Even after adding extensive demographic controls listed in Table 1 to the regression model (Column 2), these results still hold, with an estimated decrease in contribution size of 11.79 AFN for *Out-Contact*.

Result 1: Contribution size in Out-NoContact is not statistically different from contributions in In-NoContact.

Result 2: Contributions in Out-Contact are lower than contributions in In-NoContact.

Calling in subjects for survey enumeration in random order creates exogenous variation in the duration of physical exposure to a particular ethnic mix (no out-group in *In-NoContact* and *Out-NoContact*, or a 1:2 ratio of out-group to in-group in *Out-Contact*). If the drop in giving among subjects is due to the sensory stimuli brought on by the presence of the out-group, we would expect the negative reaction in *Out-Contact* to increase with prolonged exposure (within limits of being habituated to the stimuli). Figure 2a shows that despite starting at a similar level to the other two treatments in the first half of the day, contributions in *Out-Contact* decreased significantly relative to the other two by the later half of the day, suggesting that the difference between *Out-Contact* and *In-NoContact* is not due to differences across subject pools, but to the treatment itself. This is confirmed

in the regressions in Table 3, where the full model from Column 2 is estimated separately for Early and Late sessions. Giving in *Out-NoContact* is marginally higher earlier (Column 3, 5.59 AFN, $p < 0.10$) but not later, due to a slight drift downwards (Column 4, 1.34, $p > 0.10$). Giving in *Out-Contact* starts out statistically similar to *In-NoContact* (Column 3, -7.15, $p > 0.10$), but is significantly lower by the later sessions (Column 4, -16.99 AFN, $p < 0.05$).

Result 3: Contributions in Out-Contact are not statistically different from contributions in Early In-NoContact sessions, but become significantly lower by the Late In-NoContact sessions.

This is consistent with our hypothesis that casual contact mimicking everyday interactions (undirected, inadvertent) is more likely to activate the psychological mechanism of in-group bias than non-physical cues of social identity. Next, we test further observable implications of this mechanism and discuss several alternatives to the argument that would challenge our interpretation of the main results.

Exploring the Mechanism

In this section, we test the observable implications of our favored mechanism along with those of three other competing mechanisms that might explain the main results. We take each in turn, first explaining the mechanism's logic and then using the data to gauge the strength of empirical support for it. Pettigrew and Tropp (2006, 767) assert that the effects of “[f]actors that curb contact’s ability to reduce prejudice...are likely to be moderated by the degree to which group membership is salient during contact.” Evidence from other experiments shows that subjects’ reactions to cueing the out-group are moderated by characteristics of both the individual and the setting (Glaeser, Laibson, and Scheinkman 2000, Haile, Sadrieh, and Verbon 2008), which informs our analysis.

Sensory Stimuli Mechanism

Which characteristic should we focus on? If the drop in giving in *Out-Contact* was due to a decrease in altruism toward the out-group (e.g., due to negative emotional arousal), it would be tied to out-group stimuli that are present in casual contact but missing from the textual cues in *Out-NoContact*: smells, sounds, expressions and movements. While the differences in such stimuli experienced by our non-Pashtun subjects are difficult to define, the strong ethnolinguistic identity of the out-group in this case and the political context of Afghanistan ensure that the *Out-Contact* condition will introduce the spoken language of the out-group to what is otherwise a social environment that features only the language of in-group members. According to Damasio (1994) and Sambanis and Shayo (2013), the perception of these stimuli, and the resulting patterns of social identification and behavior, vary depending on an individual's background. In particular, the sound of the out-group's spoken language is more likely to be perceived as a threat to those who do not understand it. We also expect the negative reaction of a subject who cannot speak the out-group's language to increase the longer he has to wait in an environment where he hears the out-group's spoken language, driving down out-group altruism over time. This is our *sensory stimuli mechanism*.

Figure 2b illustrates subjects' average contributions split by their ability to speak the out-group's language. Subjects who do not speak the out-group's language appear to give less as the out-group cues strengthen: giving is highest in *In-NoContact*, slightly lower in *Out-NoContact* and much lower in *Out-Contact*. Yet giving from subjects who speak the out-group's language is the same in *In-NoContact* and *Out-Contact*, and higher in *Out-NoContact*. To test this more rigorously, and to explore the effect of time, we estimate the full model from Table 3 (Columns 2-4), adding interactions of the treatment variables (*Out-NoContact* and *Out-Contact*) and a dummy variable (*Speaks Out-group Language*).¹⁰ Table 4 (Columns 1-4) summarizes the differential effect of the

¹⁰ SI Table 4 shows the full tobit regression results; SI Table 5 shows an OLS version.

treatments by language ability, using the linear combination of the coefficients (Treatment + Treatment x *Speaks Out-group Language*).

The results suggest that language ability is indeed a strong predictor of how one responds to out-group cues. Column 1 shows that subjects who cannot speak the out-group's language do not discriminate against the out-group without casual contact, but do so when exposed to cues of identity through casual contact (*Out-Contact*: -18.42 AFN, $p < 0.01$). However, subjects who can speak the out-group's language do not discriminate against the out-group in the presence of casual contact, and even show out-group favoritism absent that casual contact (*Out-NoContact*: 11.31 AFN, $p < 0.10$), which is similar to the behavior of the Bosnian students attending an ethnically integrated school and playing in the ethnically mixed condition in Alexander and Christia (2011). Comparing the coefficients in Columns 2 and 3 illustrates the trajectory of giving for these two groups of subjects. In the early sessions, subjects who cannot speak the out-group's language give similarly to the in-group and out-group, regardless of whether they have casual contact with the out-group (Column 2). However, by the later sessions, the contributions of subjects who cannot speak the out-group's language become significantly lower when they have casual contact with the out-group (Column 3). This is not true for subjects who can speak the out-group's language, or for subjects who cannot speak the out-group's language and do not have casual contact with the out-group (Column 4), suggesting that the mechanism driving the drop in contributions in *Out-Contact* uniquely affects subjects who cannot speak the out-group's language who had casual contact with the out-group for a relatively longer period of time.

Result 4: The drop in contributions in Out-Contact is driven by subjects who do not speak the out-group's language. These subjects contribute equally to in-group and out-group absent casual contact with the out-group (Out-NoContact) or when casual contact is short-lived (Out-Contact, Early).

Result 5: The increase in contributions in Out-NoContact is driven by subjects who speak the out-group's language. These subjects do not discriminate after casual contact with the out-group (Out-Contact) regardless of the duration of that contact.

Our findings are consistent with the idea that the *Out-Contact* environment introduces spoken language of the out-group, a sensory stimulus that may be unconsciously classified as threatening by subjects who do not speak that language, thus decreasing their altruism toward the out-group.

Communication Mechanism

The ability to speak the out-group's language can also affect a subject's altruism toward the out-group through two other mechanisms, which can interact with the sensory stimuli mechanism - communication and previous integration (see below). Subjects who can speak the out-group's language can more easily converse with members of the out-group, and the *Out-Contact* treatment might have affected giving by providing an opportunity for in-group and out-group to get to know one another. This could facilitate more positive attitudes toward the out-group (Allport 1954). We call this the *communication mechanism*. It has two observable implications. First, giving in *Out-Contact* should increase in time among subjects who can speak the out-group's language, but remain unchanged for non-speakers. Second, due to the lack of opportunity to interact, giving in *Out-NoContact* should not vary according to one's ability to speak the out-group's language.

Neither prediction of the communication mechanism was borne out: despite the lack of opportunity to interact in *Out-NoContact*, giving was not equal across speakers and non-speakers of the out-group's language, and despite the opportunity to interact in *Out-Contact*, giving did not increase with time for subjects who could speak the out-group's language. This suggests that interactions that improve interethnic relationships may not necessarily occur even when people who can communicate with each other occupy the same physical space.

Previous Integration Mechanism

Subjects who can speak the out-group's language may be fundamentally different from subjects who do not learn the out-group's language, since the former have been more integrated with the out-group in their everyday lives, and hence are likely to exhibit less in-group bias. This dovetails with Alexander and Christia (2011), who show that only Bosnian students who previously had been (randomly) placed in an ethnically segregated school decreased their contributions in a public goods game in response to textual out-group cues; those who attended an ethnically mixed school demonstrated no out-group discrimination and even showed out-group favoritism in the presence of the same cues. If subjects who can speak the out-group's language feel more altruistic toward the out-group, they will give more to the out-group in both *Out-NoContact* and *Out-Contact* compared to subjects who do not speak the out-group's language.

We call this the *previous integration mechanism*, and it appears to have more support. It predicts that subjects who can speak the out-group's language are more integrated with the out-group and hence are more altruistic toward them. Subjects who can speak the out-group's language give more than subjects who cannot in both treatments, though surprisingly, they only show out-group favoritism in the absence of casual contact.

Day-specific Shocks

One might be concerned that the similarities in contribution in *In-NoContact* and *Out-NoContact*, and their differences with *Out-Contact*, are not due to the experimental cues but to *day-specific shocks*. This could occur in several ways. First, subjects recruited on the first day could be different than subjects recruited on the second day (e.g., higher income, more education, etc.). Second, unobserved characteristics that affected subjects' inherent altruism (e.g., neighborhood, or activity in the marketplace) may have differed across days. Third, there could be day-specific shocks in time

that affected the time trend of contribution, such as unobserved changes in weather that worsened subjects' mood.

Our earlier analysis in the Experimental Results section lessens any worries about the first two. The demographic breakdown in Table 1 suggests that subjects across all three treatments are similar across a range of observables, and we control for any remaining differences in subsequent regressions. We further find that no in-group bias was observed in the early sessions. Given that wait time was randomly assigned, this suggests that the decrease in giving was not due to the second type of shock. This leaves the third possibility. A shock that affects the time trend of contribution on one of the days (such as rising temperature or changes in the solicitor's energy level) would change the contribution time trend for all subjects on that day; note that demographic features that relate to integration with the out-group – such as the ability to speak the out-group's language – would not matter for giving.

The possibility that our findings in the Experimental Results section are due to day-specific shocks is unsupported by the data. Instead of seeing day-specific time trends across all subjects, as would be predicted by shocks occurring in the course of the day (such as changes in weather or in the administration of the experiment), we find that contribution trends were driven by demographic features that relate to one's ability to be comfortable around the out-group – the ability to speak their language. Specifically, we find that the significant drop in contribution in time can be seen only among subjects that are unable to speak the out-group's language; subjects who can speak it exhibit the same time trend in their contribution across both days, suggesting that there were no differences in their experience across days.

Free-riding

Another potential explanation for the drop in giving in *Out-Contact* is free-riding: pure altruists may have decreased their giving because they believed that others were giving more. According to this explanation, subjects give less in *Out-Contact* not because cues of ethnic identity that come via casual contact affect their altruism, but because they learn that they no longer need to give as much to maintain the same level of charity provision as in the *Out-NoContact* treatment, since the out-group will be giving more to the hospital than subjects. The implications of this explanation for the time trend are unclear. Subjects could update their beliefs on the need to give after finding out that out-group members are also giving, which would imply that giving in *Out-Contact* should be lower from the beginning. But subjects may only learn that out-group members are generous because they have more opportunity to interact with them. This would suggest that the decrease over time is driven by those who can speak the out-group's language.¹¹

However, free-riding in any form is unlikely, since it crucially depends on the out-group being more generous to the out-group hospital than the subjects, which, as we have seen earlier, is not the case.¹² Concerns over this mechanism decrease further in light of the fact that the drop in giving over time came from subjects who could not speak the out-group's language. While we cannot completely rule out this alternative explanation, the evidence we have does not support it.

Overall, our results suggest that the *Out-Contact* treatment effects are driven by the sensory stimuli mechanism that is moderated by previous integration with the out-group. We caution that our reasoning and analysis of these heterogeneous treatment effects is post hoc and should be

¹¹ We thank an anonymous referee for pointing out that the decrease in giving in time does not preclude free-riding, since non-Pashtuns may have had to learn that the Pashtuns were giving more.

¹² Above, we note that the out-group gave on average 10 AFN, which is lower than subjects' giving in all treatment conditions (≥ 16 AFN).

considered speculative, but the evidence is consistent with studies of similar behavior in countries with a history of interethnic conflict. For example, Sambanis and Shayo (2013) argue that identification with a higher-status ethnic group is facilitated by attributes in common between the individual and the higher-status group, suggesting that subjects who can speak the out-group's language might even exhibit reverse discrimination and identify with the out-group in some situations.

Conclusion

We conducted a lab-in-the-field charitable giving experiment in Kabul to investigate how casual contact affects ethnic bias and to explore the mechanisms driving behavior. Our study reiterates an often-overlooked point in the theoretical literature about the effect of contact on attitudes and behavior toward the out-group: sustained, directed, and positive interactions can promote reduction of prejudice and bias, but “casual contact has left matters worse than before” (Allport 1954, 264). We hypothesize that stimuli associated with descent-based attributes of social identity are revealed most markedly in casual in-person contact (relative to no contact at all), and that associating these stimuli with a threat (in part based on previous experience with the out-group) results in decreased identification with the out-group. The experimental evidence supports this argument. While subjects were just as willing to contribute to help their own group (*In-NoContact*) as the out-group without contact (*Out-NoContact*), contributions decreased markedly after casual physical contact with the out-group (*Out-Contact*). The fact that contributions were decreasing with time of exposure to the out-group, and that a common language with the out-group moderated the negative effect of casual contact, is consistent with a sensory stimuli mechanism of behavior.

This study provides three main lessons. First, we show the effects of a relatively understudied type of interethnic interaction – undirected, involuntary, and relatively short in duration. Our results

suggest that while visceral cues of the dominant out-group are powerful drivers of behavior for ethnic minorities, attributes that enable positive interaction with the out-group – such as a shared language – can moderate this psychological channel. Which attributes will act as moderators may vary by context and may be hard to detect ex ante, but just as research shows that the effects of diversity on discrimination is context-dependent (Enos and Gidron 2016), linking sensory stimuli to behavior as we do here should provide scholars with analytical leverage in anticipating how the effects of casual contact on out-group bias may vary across individuals and settings.

Second, we show that the psychological mechanism plays a significant role in motivating in-group bias (Sambanis and Shayo 2013). We provide a framework for understanding how this psychological mechanism is related to (and activated by) stimuli uniquely present in casual in-person interaction, and report evidence consistent with our argument. These findings help us better understand the microfoundations of discrimination by incorporating the insights of neuroscience into experiments (Bertrand and Duflo 2017, 312-313). Accounting for the psychological bases of group identification may provide new insights for institutional design and efforts to resolve conflict (Sambanis and Shayo 2013).¹³

Third, our results suggest that cueing the out-group through physical exposure may have very different effects on subjects' behavior relative to out-group cueing through photographs, names, or textual description. The latter is much more common in the experimental literature, sometimes for

¹³ As one scholar observes: “Research on the role of the amygdala and related regions in the temporal lobe suggests that dependence on rational incentives is unlikely to resolve fears and hatreds and that leaders of opposing sides are limited in their ability to convince their followers to accept peace without somehow ameliorating emotional biases, which may reside deep within the brain” (Blank 2005, 18).

ethical, logistical, or other reasons. But this variation in effects has implications for research in comparative politics and political economy that increasingly employs experimental, non-physical cues or primes of the out-group, such as studies of vote choice (Dunning and Harrison 2010, Adida 2015), politician responsiveness to constituents' needs (McClendon 2016), and citizen policy preferences (McCauley 2014). While the physical and non-physical cues may have different effects, this suggests complementary roles for both types of cues in eliciting group identification in the sense that they examine how the individual behaves in different states. Comparing the effects of these cues under different conditions is a fruitful avenue of future research.

References

- Adida, Claire L. 2015. "Do African Voters Favor Coethnics? Evidence from a Survey Experiment in Benin." *Journal of Experimental Political Science* 2 (1):1-11.
- Adida, Claire L., David D. Laitin, and Marie-Anne Valfort. 2016. "One Muslim Is Enough! Evidence from a Field Experiment in France." *Annals of Economics and Statistics* 121/122:121-160.
- Alexander, Marcus, and Fotini Christia. 2011. "Context Modularity of Human Altruism." *Science* 334 (6061):1392-1394.
- Allport, Gordon. 1954. *The Nature of Prejudice*. Cambridge, MA: Addison-Wesley.
- Ariely, Dan, and George Loewenstein. 2006. "The Heat of the Moment: The Effect of Sexual Arousal on Sexual Decision Making." *Journal of Behavioral Decision Making* 19 (2):87-98.
- Barfield, Thomas J. 2010. *Afghanistan: A Cultural and Political History*. Princeton: Princeton University Press.
- Bates, Robert. 1983. "Modernization, Ethnic Competition, and the Rationality of Politics in Contemporary Africa." In *State versus Ethnic Claims: African Policy Dilemmas*, edited by Donald Rothchild and Victor A. Olorunsola, 152-171. Boulder, CO: Westview Press.
- Bates, Robert H. 1974. "Ethnic Competition and Modernization in Contemporary Africa." *Comparative Political Studies* 6 (4):457-484.
- Bertrand, Marianne, and Esther Duflo. 2017. "Field Experiments on Discrimination." In *Handbook of Economic Field Experiments*, edited by Abhijit V. Banerjee and Esther Duflo, 309-394. Amsterdam: North-Holland.
- Blank, Robert H. 2005. "The Brain, Aggression, and Public Policy." *Politics and the Life Sciences* 24 (1/2):12-21.
- Bloomfield, David, Teresa Barnes, and Luc Huyse, eds. 2003. *Reconciliation after Violent Conflict*. Stockholm: International IDEA.
- Brück, Carolin, Benjamin Kreifelts, and Dirk Wildgruber. 2011. "Emotional Voices in Context: A Neurobiological Model of Multimodal Affective Information Processing." *Physics of Life Reviews* 8 (4):383-403.
- Carrell, Scott E., Mark Hoekstra, and James E. West. 2015. "The Impact of Intergroup Contact on Racial Attitudes and Revealed Preferences." *NBER Working Paper No. 20940*.
- Chandra, Kanchan. 2006. "What is Ethnic Identity and Does it Matter?" *Annual Review of Political Science* 9:397-424.
- Chuah, Swee-Hoon, Robert Hoffmann, Bala Ramasamy, and Jonathan H. W. Tan. 2014. "Religion, Ethnicity and Cooperation: An Experimental Study." *Journal of Economic Psychology* 45:33-43.
- Cunningham, William A., Marcia K. Johnson, Carol L. Raye, J. Chris Gatenby, John C. Gore, and Mahzarin R. Banaji. 2004. "Separable Neural Components in the Processing of Black and White Faces." *Psychological Science* 15 (12):806-813.
- Dalgleish, Tim. 2004. "The Emotional Brain." *Nature Reviews Neuroscience* 5 (7):583-589.
- Damasio, Antonio R. 1994. *Descartes' Error*. New York: G. P. Putnam.
- De Dreu, Carsten K. W., Lindred L. Greer, Michel J. J. Handgraaf, and et al. 2010. "The Neuropeptide Oxytocin Regulates Parochial Altruism in Intergroup Conflict Among Humans." *Science* 328 (5984):1408-1411.
- Ditto, Peter H., David A. Pizarro, Eden B. Epstein, Jill A. Jacobson, and Tara K. MacDonald. 2006. "Visceral Influences on Risk-Taking Behavior." *Journal of Behavioral Decision Making* 19 (2):99-113.

- Dixon, John, Kevin Durrheim, and Colin Tredoux. 2005. "Beyond the Optimal Contact Strategy: A Reality Check for the Contact Hypothesis." *American Psychologist* 60 (7):697-711.
- Dunning, Thad, and Lauren Harrison. 2010. "Cross-Cutting Cleavages and Ethnic Voting: An Experimental Study of Cousinage in Mali." *American Political Science Review* 104 (1):21-39.
- Eifert, Benn, Edward Miguel, and Daniel N. Posner. 2010. "Political Competition and Ethnic Identification in Africa." *American Journal of Political Science* 54 (2):494-510.
- Enos, Ryan. 2014. "Causal Effect of Intergroup Contact on Exclusionary Attitudes." *PNAS* 111 (10):3699-3704.
- Enos, Ryan D. 2017. *The Space Between Us*. New York: Cambridge University Press.
- Enos, Ryan D., and Noam Gidron. 2016. "Intergroup Behavioral Strategies as Contextually Determined: Experimental Evidence from Israel." *The Journal of Politics* 78 (3):851-67.
- Glaeser, Edward L., David I. Laibson, and Jose A. Scheinkman. 2000. "Measuring Trust." *Quarterly Journal of Economics* 115 (3):811-846.
- Gneezy, Uri, and Alex Imas. 2017. "Lab in the Field: Measuring Preferences in the Wild." In *Handbook of Economic Field Experiments*, edited by Marianne Bertrand and Esther Duflo, 439-464. Amsterdam: North-Holland.
- Green, Donald P., and Janelle S. Wong. 2008. "Tolerance and the Contact Hypothesis: A Field Experiment." In *The Political Psychology of Democratic Citizenship*, edited by Eugene Borgida, Christopher M. Federico and John L. Sullivan, 228-46. London: Oxford University Press.
- Haile, Daniel, Abdolkarim Sadrieh, and Harrie A. A. Verbon. 2008. "Cross-Racial Envy and Underinvestment in South African Partnerships." *Cambridge Journal of Economics* 32 (5):703-724.
- Hill, R. Carter, William E. Griffiths, and Guay C. Lim. 2011. *Principles of Econometrics*. 4th ed. Hoboken, NJ: Wiley.
- Hopkins, Nancy, ed. 2014. "Afghanistan in 2014: A Survey of the Afghan People." *The Asia Foundation*, Available from: <http://asiafoundation.org/resources/pdfs/Afghanistanin2014final.pdf>.
- Horowitz, Donald L. 1985. *Ethnic Groups in Conflict*. Berkeley: University of California Press.
- Jacobs, Jane. 1961. *The Death and Life of Great American Cities*. New York: Random House.
- Kaufmann, Chaim. 1996. "Possible and Impossible Solutions to Ethnic Civil Wars." *International Security* 20 (4):136-75.
- Laitin, David D. 2000. "What Is a Language Community?" *American Journal of Political Science* 44 (1):142-55.
- Lambarraa, Fatima, and Gerhard Riener. 2012. "On the Norms of Charitable Giving in Islam: A Field Experiment." *Courant Research Centre Discussion Papers*, Available from: <https://ideas.repec.org/p/ags/iaae12/126795.html>.
- Leukel, Francis. 1976. *Introduction to Physiological Psychology*. 3rd ed. St. Louis: Mosby.
- Lieberman, Matthew D. 2007. "Social Cognitive Neuroscience: A Review of Core Processes." *Annual Review of Psychology* 58:259-289.
- Malhotra, Deepak, and Sumanasiri Liyanage. 2005. "Long-Term Effects of Peace Workshops in Protracted Conflicts." *Journal of Conflict Resolution* 49 (6):908-924.
- McCauley, John F. 2014. "The Political Mobilization of Ethnic and Religious Identities in Africa." *American Political Science Review* 108 (4):801-16.
- McClendon, Gwyneth. 2016. "Race and Responsiveness: A Field Experiment with South African Politicians." *Journal of Experimental Political Science* 3 (1):60-74.
- Metcalf, Janet, and Walter Mischel. 1999. "A Hot/Cool-System Analysis of Delay of Gratification: Dynamics of Willpower." *Psychological Review* 106 (1):3-19.

- Michelitch, Kristin. 2015. "Does Electoral Competition Exacerbate Interethnic or Interpartisan Economic Discrimination? Evidence from a Field Experiment in Market Price Bargaining." *American Political Science Review* 109 (1):43-61.
- Mogelson, Luke. 2012. "Pacifists in the Cross-Fire: The Kabul Hospital that Treats All Sides." *New York Times*, May 18.
- Oxfam. 2009. "The Cost of War: Afghan Experiences of Conflict, 1978-2009." Report. Available from: <https://www.oxfam.org/sites/www.oxfam.org/files/afghanistan-the-cost-of-war.pdf>.
- Paluck, Elizabeth Levy, Seth A. Green, and Donald P. Green. 2017. "The Contact Hypothesis Reevaluated." *SSRN paper*. Available from: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2973474.
- Pettigrew, Thomas F. 1997. "Generalized Intergroup Contact Effects on Prejudice." *Personality and Social Psychology Bulletin* 23 (2):173-185.
- Pettigrew, Thomas F., and Linda R. Tropp. 2006. "A Meta-Analytic Test of Intergroup Contact Theory." *Journal of Personality and Social Psychology* 90 (5):751-783.
- Pettigrew, Thomas F., and Linda R. Tropp. 2008. "How Does Intergroup Contact Reduce Prejudice? Meta-analytic Tests of Three Mediators." *European Journal of Social Psychology* 38:922-34.
- Sambanis, Nicholas, and Moses Shayo. 2013. "Social Identification and Ethnic Conflict." *American Political Science Review* 107 (2):294-325.
- Shafir, Eldar, Itamar Simonson, and Amos Tversky. 1993. "Reason-based Choice." *Cognition* 49 (1-2):11-36.
- Shah, N. J., J. C. Marshall, O. Zafiris, A. Schwab, K. Zilles, H. J. Markowitsch, and G. R. Fink. 2001. "The Neural Correlates of Person Familiarity." *Brain* 124 (4):804-15.
- Shook, Natalie J., and Russell H. Fazio. 2008. "Interracial Roommate Relationships: An Experimental Field Test of the Contact Hypothesis." *Psychological Science* 19 (7):717-723.
- Tajfel, Henri, and John Turner. 1979. "An Integrative Theory of Intergroup Conflict." In *The Social Psychology of Intergroup Relations*, edited by Stephen Worchel and William Austin, 33-48. Monterey, CA: Brooks/Cole.
- Tajfel, Henri, and John Turner. 1986. "The Social Identity Theory of Intergroup Behavior." In *Psychology of Intergroup Relations*, edited by Stephen Worchel and William Austin, 7-24. Chicago: Nelson Hall.
- Whitt, Sam. 2014. "Social Norms in the Aftermath of Ethnic Violence: Ethnicity and Fairness in Non-costly Decision Making." *Journal of Conflict Resolution* 58 (1):93-119.
- Whitt, Sam, and Rick K. Wilson. 2007. "The Dictator Game, Fairness and Ethnicity in Postwar Bosnia." *American Journal of Political Science* 51 (3):655-68.
- Wucherpfennig, Julian, Nils B. Weidmann, Luc Girardin, Lars-Erik Cederman, and Andreas Wimmer. 2011. "Politically Relevant Ethnic Groups across Space and Time: Introducing the GeoEPR Dataset." *Conflict Management and Peace Science* 28 (5):423-37.
- Zak, Paul J., Angela A. Stanton, and Sheila Ahmadi. 2007. "Oxytocin Increases Generosity in Humans." *PLoS One* 2 (11):e1128.

Table 1. Summary Statistics on Demographics for Subjects across All Treatments

Variable	Obs.	Mean	Std. Dev.	Min.	Max.
<i>Panel A: All Non-Pashtun Subjects</i>					
Age	213	33	12.13	18	79
Languages Spoken	213	1.62	0.72	1	4
Married	212	0.79	0.41	0	1
No Education	213	0.50	0.50	0	1
Years of Education (if educated)	96	6.83	3.13	2	14
Monthly Income (AFN)	213	4353	2216	88	12000
Time in occupation (years)	212	10.33	8.69	0	50
<i>Panel B: Non-Pashtuns contributing to Non-Pashtuns(In-NoContact)</i>					
Age	67	36	13.35	18	79
Languages Spoken	67	1.58	0.78	1	4
Married	66	0.788	0.41	0	1
No Education	67	0.478	0.50	0	1
Years of Education (if educated)	31	7.03	3.44	2	14
Monthly Income (AFN)	67	4427	2385	88	10000
Time in occupation (years)	67	9.51	8.81	0	50
<i>Panel C: Non-Pashtuns contributing to Pashtuns (Out-NoContact)</i>					
Age	73	34	11.54	18	64
Languages Spoken	73	1.52	0.65	1	3
Married	73	0.78	0.42	0	1
No Education	73	0.48	0.50	0	1
Years of Education	33	6.58	3.04	2	13
Monthly Income (AFN)	73	3905	1796	88	10000
Time in occupation (years)	73	11.67	9.81	0	50
<i>Panel D: Non-Pashtuns contributing to Pashtuns (Out-Contact)</i>					
Age	73	29	10.59	18	62
Languages Spoken	73	1.74	0.73	1	4
Married	73	0.81	0.40	0	1
No Education	73	0.53	0.50	0	1
Years of Education (if educated)	32	6.91	2.99	2	12
Monthly Income (AFN)	73	4732	2379	88	12000
Time in occupation (years)	72	9.72	7.19	0	30
<i>Panel E: Pashtuns contributing to Pashtuns (Out-Contact)</i>					
Age	33	30	9.79	18	55
Languages Spoken	33	1.55	0.56	1	3
Married	33	0.91	0.29	0	1
No Education	33	0.73	0.45	0	1
Years of Education (if educated)	9	6.11	2.42	3	11
Monthly Income (AFN)	33	4078	1951	88	9000
Time in occupation (years)	33	10.70	8.00	0	35

Table 2. Average Contributions across Treatment Conditions

	In- NoContact	Out- NoContact	Out- Contact	In-NoContact vs Out-NoContact	Out-NoContact vs Out-Contact	In-NoContact vs Out-Contact
Individual Level						
N	67	73	73			
Amount (AFN)	20	21.3	16	$p=0.99$	$p=0.04$	$p=0.05$
Session Level						
N	10	11	17			
Amount (AFN)	19.9	21.2	15.9	$p=0.65$	$p=0.05$	$p=0.08$

Notes: p -values correspond to Wilcoxon rank-sum tests

Table 3. Effects of Contact with Out-group on Contributions (Tobit)

VARIABLES	All	All	Early	Late
	(1) Amount	(2) Amount	(3) Amount	(4) Amount
Out-NoContact	1.33 (3.31)	2.15 (3.32)	5.59* (3.32)	1.34 (4.48)
Out-Contact	-12.53** (5.41)	-11.79** (5.63)	-7.15 (7.21)	-16.99** (7.22)
Session	-0.75*** (0.29)	-0.61** (0.29)	-1.10 (0.68)	-0.69 (0.64)
Constant	34.56*** (6.11)	32.51* (18.41)	54.36** (26.93)	10.90 (23.18)
Demographic Controls	No	Yes	Yes	Yes
Log likelihood	-760.04	-754.27	-394.17	-345.53
F-statistic	3.54	4.32	25.52	9.96
Observations	213	212	106	106

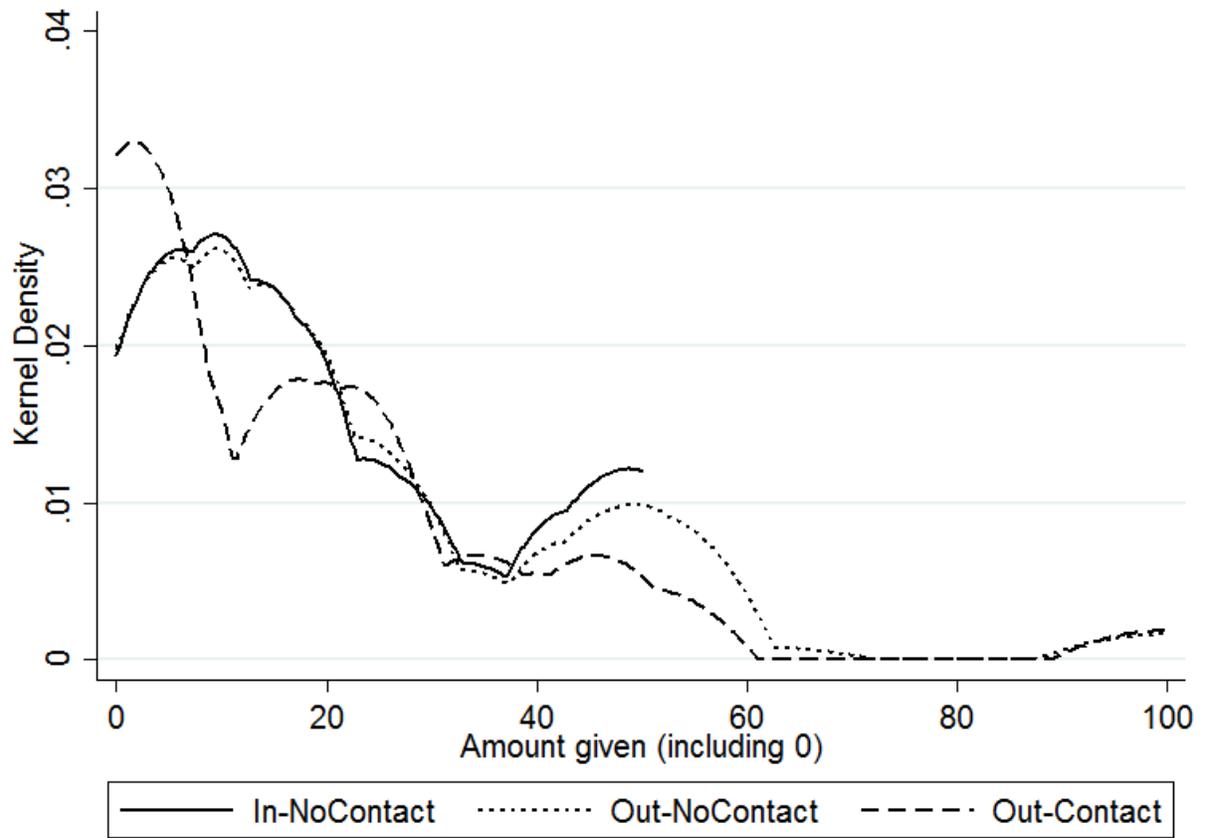
Notes: All models control for enumerator fixed effects. Robust standard errors clustered on session. See SI Table 2 for full results. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 4. Treatment Effects by Ability to Speak Pashto (Early vs. Late) (Tobit)

Effect of treatment	ALL	EARLY	LATE
	<u>Subject does not speak Pashto</u>		
Out-NoContact	-2.71 (3.85)	2.24 (3.26)	-3.34 (6.99)
Out-Contact	-18.42*** (6.75)	-9.18 (9.68)	-26.76*** (7.01)
<u>Subject speaks Pashto</u>			
Out-NoContact + Out-NoContact * demog. dummy	11.31* (6.71)	11.74 (11.27)	9.22 (8.58)
Out-Contact + Out-Contact * demog. dummy	-5.00 (7.37)	-4.11 (10.13)	-5.29 (11.31)

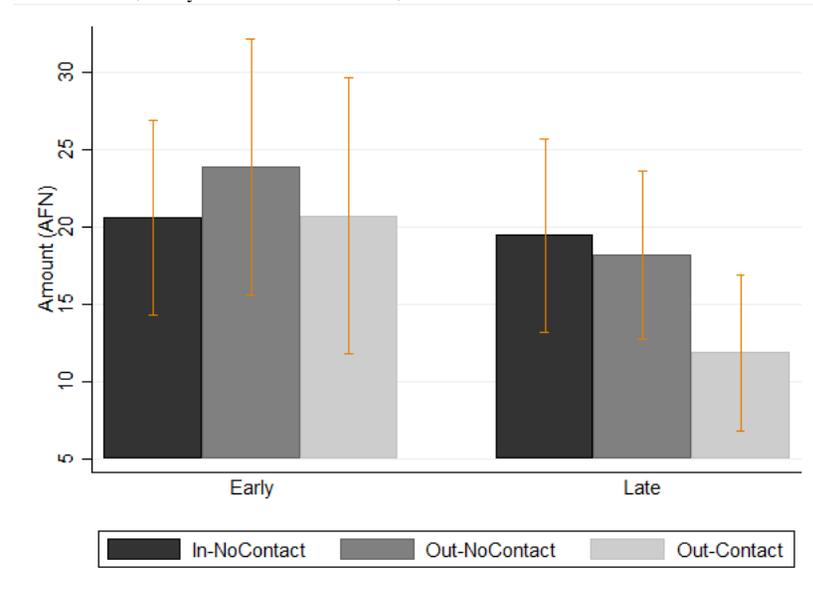
Notes: Results based on models in Table 3 Columns 2-4. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Figure 1. Kernel Density of Subjects' Contributions across Treatment Conditions



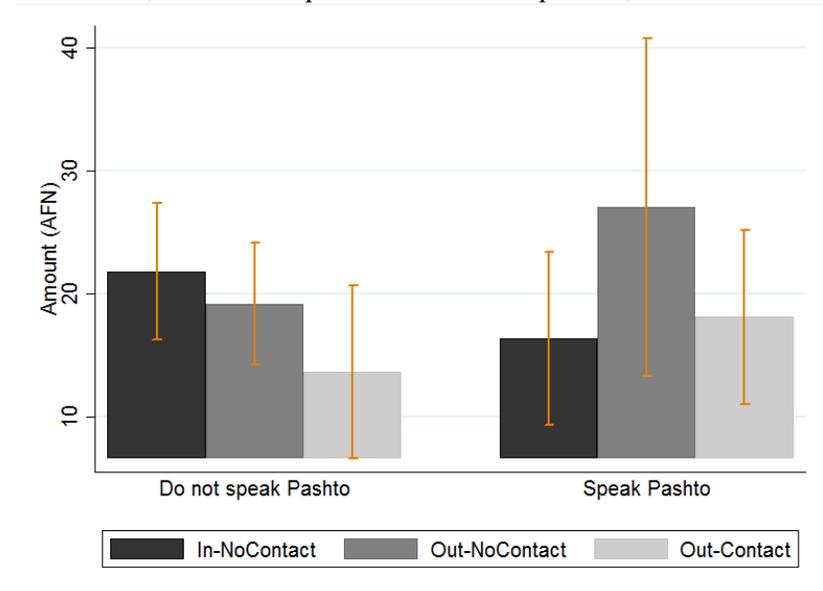
Notes: Only data on non-Pashtuns' contributions depicted.

Figure 2a. Average Contributions
(Early vs. Late Sessions)



Notes: Only data on non-Pashtuns' contributions depicted. Bars depict subjects' mean contribution amounts, by treatment condition, with associated 95% confidence intervals.

Figure 2b. Average Contributions
(Non-Pashto speakers vs. Pashto Speakers)



Notes: Only data on non-Pashtuns' contributions depicted. Bars depict subjects' mean contribution amounts, by treatment condition, with associated 95% confidence intervals.

Casual Contact and Ethnic Bias:
Experimental Evidence from Afghanistan

Online Appendix

Institutional Review and Ethics

The experimental protocol was approved by the University of Pittsburgh's IRB (PRO14050489). The protocol also was reviewed by the Asia Foundation's Director of Survey and Research, at the University of Pittsburgh's request, and had "no concern with this study in terms of cultural sensitivities or local norms." The IRB granted approval for researchers to gain oral (instead of written) consent to participate in the experiment, given high rates of illiteracy in Afghanistan. Authors held CITI human subjects training/certification and were present in the field throughout the experiment and participated in and oversaw recruitment and enumeration of subjects. The survey instrument appears at the end of this appendix.

There was no deception in the experiment. Authors delivered all funds that subjects donated in the experiment to EMERGENCY's Programme Coordinator in Kabul. (A letter confirming EMERGENCY's receipt of the funds is on file with the corresponding author.) We do not think that participation in the experiment constituted economic hardship for the (severely) economically poor subjects. Subjects were paid a little more than a normal day's wage for participation in the experiment (even though subjects gave at most a half day of their time). The choice presented - to contribute to a nonprofit charitable organization - is familiar to subjects in the context of charitable giving in Islam, and since their contribution choice was anonymous, they faced no external pressures to give.

Recruitment and Enumeration

As we note in the main text, our results are based on a convenience sample. While we describe the close similarity of our subjects' demographic characteristics to available national-level population statistics in the main text, since there is no accurate recent census data available, there is no way to know whether our subjects are representative of the broader Kabul (or Afghan) population. We recruited from markets across the city; to avoid duplicate recruitment, we did not revisit any market. Since day-laborers tend to wait for employment opportunities at only one market, and given the short time frame of the experiment, the chance of spillover effects seems small, though we cannot rule this out empirically. Day 1 subjects were recruited from Sar e Shamali, Sar e Kotal, and Pol e Sorkh markets; day 2 subjects were recruited from Karte Naw, Bagrami, and Doghabad. There was no subject attrition.

SI Table 1. Balance of Demographics across Treatment Conditions

<i>p-values of Differences in Means (two-tailed tests)</i>				
	non-Pashtuns	non-Pashtuns	non-Pashtuns	non-Pashtuns vs Pashtuns
	In-NoContact vs Out-NoContact	Out-NoContact vs Out-Contact	In-NoContact vs Out-Contact	Out-Contact
Age	0.39	0.01	0.01	0.57
Languages Spoken	0.61	0.06	0.22	0.18
Marital Status	0.92	0.68	0.77	0.19
No education	0.98	0.51	0.51	0.06
Years of education (if educated)	0.57	0.66	0.88	0.47
Monthly income	0.14	0.02	0.45	0.17
Years in occupation	0.17	0.17	0.87	0.54

SI Table 2. Effects of Casual Contact with Out-group on Contributions (Tobit)

	All	All	Early	Late
	(1)	(2)	(3)	(4)
VARIABLES	Amount	Amount	Amount	Amount
Out-NoContact	1.33 (3.31)	2.15 (3.32)	5.59* (3.32)	1.34 (4.48)
Out-Contact	-12.53** (5.41)	-11.79** (5.63)	-7.15 (7.21)	-16.99** (7.22)
Session	-0.75*** (0.29)	-0.61** (0.29)	-1.10 (0.68)	-0.69 (0.64)
Age		-0.01 (0.24)	-0.24 (0.35)	0.32 (0.31)
Languages Spoken		0.75 (4.12)	-3.63 (6.80)	2.54 (4.47)
Not married		-14.12** (6.66)	-32.69*** (11.84)	1.86 (7.48)
Years of Education		0.38 (0.54)	0.61 (0.71)	-0.06 (0.65)
Monthly Income (thousand AFN)		0.54 (0.98)	0.84 (1.24)	1.75 (1.32)
Time in occupation (years)		-0.16 (0.30)	-0.79* (0.44)	0.15 (0.33)
Log likelihood	-760.04	-754.27	-394.17	-345.53
Fstatistic	3.54	4.32	25.52	9.96
Constant	34.56*** (6.11)	32.51* (18.41)	54.36** (26.93)	10.90 (23.18)
Observations	213	212	106	106

Notes: Full results from models in Table 3. All models control for enumerator fixed effects. Robust standard errors clustered on session. *** p<0.01, ** p<0.05, * p<0.1

SI Table 3. Effects of Casual Contact with Out-group on Contributions (OLS)

	All	All	Early	Late
	(1)	(2)	(3)	(4)
VARIABLES	Amount	Amount	Amount	Amount
Out-NoContact	1.53 (2.39)	2.03 (2.41)	6.30** (2.70)	0.30 (3.68)
Out-Contact	-6.80* (3.67)	-6.27 (3.77)	-2.92 (5.29)	-9.31* (5.17)
Session	-0.53** (0.20)	-0.45** (0.22)	-0.59 (0.51)	-0.54 (0.54)
Age		-0.01 (0.19)	-0.21 (0.32)	0.29 (0.25)
Languages Spoken		0.09 (2.83)	-2.45 (5.87)	1.55 (2.66)
Not married		-9.21* (4.76)	-23.59** (8.57)	2.80 (5.49)
Years of Education		0.27 (0.40)	0.37 (0.54)	-0.03 (0.50)
Monthly Income (thousand AFN)		0.22 (0.78)	0.44 (0.98)	0.91 (1.00)
Time in occupation (years)		-0.11 (0.21)	-0.48 (0.33)	0.05 (0.27)
Constant	32.86*** (4.96)	32.98** (14.65)	47.61* (24.26)	16.49 (16.38)
Observations	213	212	106	106
R ²	0.08	0.11	0.21	0.18

Notes: OLS versions of models in Table 3. All models control for enumerator fixed effects. Robust standard errors clustered on session. *** p<0.01, ** p<0.05, * p<0.1

SI Table 4. Treatment Effects by Ability to Speak Pashto (Early vs. Late) (Tobit)

VARIABLES	(1) ALL	(2) EARLY	(3) LATE
Out-NoContact	-2.71 (3.85)	2.24 (3.26)	-3.34 (6.99)
Out-Contact	-18.42*** (6.75)	-9.18 (9.68)	-26.76*** (7.01)
Out-NoContact x Speaks Pashto	14.02* (7.96)	9.50 (12.88)	12.56 (11.59)
Out-Contact x Speaks Pashto	13.42* (8.09)	5.07 (12.65)	21.46* (11.35)
Session	-0.59** (0.30)	-1.04 (0.64)	-0.41 (0.79)
Age	5.78 (5.09)	6.92 (6.13)	-4.67 (6.95)
Speaks Pashto	-10.28 (6.31)	-8.37 (8.87)	-8.37 (10.51)
Not married	-16.28** (6.84)	-32.85*** (11.57)	2.97 (6.29)
Years of education	0.50 (0.51)	0.60 (0.70)	0.08 (0.65)
Monthly Income (thousand AFN)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)
Time in occupation (years)	-0.05 (0.28)	-0.75* (0.44)	0.35 (0.27)
Constant	33.03*** (9.41)	39.58*** (12.12)	23.62 (17.27)
Log likelihood	-751.95	-393.51	-344.27
F-statistic	3.90***	56.42***	24.09***
Observations	212	106	106

Notes: Full results of Table 4 (Tobit). All models control for enumerator fixed effects. Robust standard errors clustered on session. *** p<0.01, ** p<0.05, * p<0.1

SI Table 5. Treatment Effects by Ability to Speak Pashto (Early vs. Late) (OLS)

VARIABLES	(1) ALL	(2) EARLY	(3) LATE
Out-NoContact	-2.32 (3.06)	3.16 (2.81)	-4.44 (5.54)
Out-Contact	-11.76** (4.43)	-5.56 (7.67)	-16.28*** (4.66)
Out-NoContact x Speaks Pashto	12.02* (6.54)	8.82 (11.25)	11.74 (8.76)
Out-Contact x Speaks Pashto	10.23* (5.57)	5.95 (9.57)	14.76* (7.55)
Session	-0.43* (0.22)	-0.56 (0.48)	-0.27 (0.60)
Age	6.03 (4.30)	6.75 (5.58)	-1.39 (5.56)
Speaks Pashto	-8.72* (4.40)	-7.28 (6.88)	-7.56 (6.93)
Not married	-11.75** (5.00)	-24.08** (8.44)	1.59 (4.72)
Years of education	0.39 (0.37)	0.37 (0.56)	0.17 (0.52)
Monthly Income (thousand AFN)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)
Time in occupation (years)	0.01 (0.20)	-0.45 (0.29)	0.25 (0.25)
Constant	32.41*** (8.15)	35.33*** (11.04)	26.23* (14.20)
Observations	212	106	106
R ²	0.14	0.22	0.19

Notes: OLS version of SI Table 4. All models control for enumerator fixed effects. Robust standard errors clustered on session. *** p<0.01, ** p<0.05, * p<0.1

SI Table 6. Treatment Effects by Ability to Speak Pashto (Early vs. Late) (OLS)

<u>Effect of treatment</u>	<u>ALL</u>	<u>EARLY</u>	<u>LATE</u>
	<u>Subject does not speak Pashto</u>		
Out-NoContact	-2.32 (3.06)	3.16 (2.81)	-4.44 (5.54)
Out-Contact	-11.76** (4.43)	-5.56 (7.67)	-16.28*** (4.66)
	<u>Subject speaks Pashto</u>		
Out-NoContact + Out-NoContact * demog. dummy	9.70* (5.35)	11.97 (9.94)	7.30 (6.37)
Out-Contact + Out-Contact * demog. dummy	-1.52 (5.03)	0.38 (7.05)	-1.51 (7.85)

Notes: Results based on models in SI Table 5. All models control for enumerator fixed effects. Robust standard errors clustered on session. *** p<0.01, ** p<0.05, * p<0.1

16 What is the most challenging problem that you face in your work as a construction worker?			
1	Low wages	5	competition with Pakistani workers
2	Cannot find work every day	6	No support from government
3	Work is too hard	7	working days too long
4	Being away form family		

17 Which province are you from?	
Province Name:	

18 Which province are you from?	
Province Code: <input type="text"/> <input type="text"/>	

19 Which ethnic group do you belong to?	
1	Pushton
2	Tajik
3	Hazara
4	Uzbek
5	Turkmen
6	Aimagh
7	Baluch
8	Other:

20 How wealthy is your household comparedto other households in your neighborhood?					
1	Poor	3	Above Average	88	Don't Know
2	Below Average	4	Rich	99	Refuse to Answer