Insult versus accident: A study of the effect of cultural construals on learning to coordinate*

Benjamin A. Brooks† Karla Hoff‡ Priyanka Pandey§

November 10, 2013

Abstract

High-caste and low-caste men in India repeatedly played a coordination game. Compared to low-caste men, high-caste men coordinated far less efficiently. They were also 29 percentage points less likely to keep trying for efficient coordination after getting the "loser's payoff"—the payoff to a player who attempts efficient coordination when his partner does not. We explain both findings in a model of learning where high-caste, but not low-caste men, see the loser's payoff as an insult rather than an accident. These findings provide evidence that cultural construals can impede coordination and society's ability to adapt to change.

Keywords: Culture, coordination, stag hunt, learning, fairness equilibrium.

JEL codes: C72, O12, O17, Z1.

---

*We greatly benefitted from discussions at an early stage of this project with Paul Attewell, Gary Charness, Ernst Fehr, Rocco Macchiavello, and Matthew Rabin. We thank seminar participants at Indiana University, Princeton University, the World Bank, the University of California–San Diego, the University of California–Santa Barbara, the University of Memphis, and Warwick University. We also thank Cristina Bicchieri for detailed comments on our draft. Finally, thanks to Sonal Vats for excellent research assistance and the World Bank Research Support Budget for financial support.

†Department of Economics, Princeton University, babrooks@princeton.edu

‡The World Bank, khoff@worldbank.org

§The World Bank, ppandey@worldbank.org

1 Introduction

Why would a society fail to coordinate on broadly beneficial institutions? Why would efficient Nash equilibria fail to emerge? Examples abound of societies whose institutions seem to provide low welfare compared to available alternatives.[1] A major cost of miscoordination is missed opportunities for development (Matsuyama 1996, Ray 2000, Hoff 2001). The process of learning to coordinate in repeated play has attracted much attention (a survey is Camerer 2003, ch. 6), yet because nearly all of the experimental work has been undertaken in developed countries, little is known about why some societies seem to be less able than others to coordinate on efficient institutions.

What has been described as "institutional inertia" in the Indian state of Uttar Pradesh illustrates that a society can fail to establish institutions that address even glaringly obvious problems (Drèze and Gazdar 1997). In pre-independence India, a system of customary obligations provided street maintenance, sanitation, and drainage. That system eroded with the decline of the political power of landlords and the emancipation of the laboring classes, but fifty years later in Uttar Pradesh "the challenge of creating [a system]... to address those needs has been largely unmet" (ibid., p. 96). Drèze and Sharma (1998) describe other failures of coordination in the village of Palanpur, Uttar Pradesh:

(1) *Drainage*: [The] expansion [among the better-off households] in the ownership of private hand-pumps... is the cause of a major public nuisance—the uncontrolled flow of large quantities of waste water from private courtyards to public lanes... Palanpur villagers agree that drainage can be substantially improved if each hand-pump owner digs a hole outside her or his courtyard, and directs waste water to that hole... However, an individual has little incentive to dig a hole unless his or her neighbours do the same. In order to be effective, this measure has to be taken by all the residents in a village, or at least all the residents of a particular lane. This has not happened...

(2) *Early sowing*: [Farmers] acknowledge that, if everyone sowed earlier, yields would be higher. But... those who sow early are particularly vulnerable to seeing their seeds being eaten up by hungry birds... Coordinated sowing is one possible answer... [The]

[1] For example, North 1990 2006, Greif 1994, Hoff and Stiglitz 2002, Guiso, Sapienza and Zingales 2008.

> constraint is the absence of any forum for reaching a collective decision of this kind, and of any practice of cooperation among farmers. . . (pp. 71-73)

We view these situations as failures to coordinate on efficient institutions. All parties would gain from switching to a better drainage system or moving up the sowing season, yet the efficient conventions did not emerge.

Economists have identified several generic impediments to efficient coordination: large numbers of individuals (Schelling 1978, Van Huyck, Battalio and Beil 1990), pessimistic beliefs (Merton 1968, Hoff and Stiglitz 2002, Guiso, Sapienza and Zingales 2008, Tabellini 2008), and social distance (Chen and Chen 2011).

In this paper, we argue that subjective construals, which may lead people to think and feel and act in culture-specific ways, are an additional factor that can help or hinder individuals' ability to coordinate on efficient outcomes. We focus on two channels through which culture affects coordination. First, culture is a source of *mental frameworks* for interpreting strategic situations: For instance, is a given action an accident, an insult, or something else? Second, culture provides *rules of thumb* that guide behavior: For example, what is the culturally appropriate response to a perceived insult?[2] We argue that when players miscoordinate, these elements of culture influence what kind of convention emerges over the course of the two players' subsequent interactions. A mental framework that leads a player to interpret miscoordination as a punishable insult hinders his ability to coordinate with his partner on an efficient equilibrium.

Our notion of mental frameworks is related to the concept of *frames* in economics and psychology. A frame (or framing device) in economics is an aspect of the presentation of a situation, whereas a mental framework is in the mind—it is brought by the individual to the situation. The key idea is that events and actions do not always speak for themselves but instead may depend on framing for their meaning: the way in which a situation is presented,

[2]The role we assign to *rules of thumb* is related to the notion of culture as a "toolkit" in Swidler (1986) and Henrich and Ensminger (2011), and also to work by Nunn and colleagues, e.g. Alesina, Giuliano and Nunn (2011).

as well as the mental framework that an individual brings to the situation, influences its meaning and significance for the individual (Goffman 1974). Results from experiments that manipulate frames demonstrate that a change in the way a given game is described can lead to a large change in behavior, even though there is no change in the possible actions or material outcomes (cf. Andreoni 1995; a popular account is Ariely 2009). The novelty of our paper is to provide evidence that a cultural factor—in particular, the mental frameworks associated with a "culture of honor"—can make it difficult for individuals to learn to coordinate.

To evaluate the impact of culture on coordination, we implemented experiments in Uttar Pradesh. The participants were members of castes at either the top or the bottom of the Hindu caste hierarchy—castes that have occupied the same broad status positions for centuries.[3] Individuals' caste status is determined by birth rather than individual choice or competition. Thus, differences between castes will be the outcome of the nature of the communities rather than of a self-selection process.[4]

The high- and low-caste participants in our experiment speak the same language and live in the same villages, yet come from communities with radically different positions in Indian social structures. High castes were historically the dominant landowners and the priests, whereas low castes were the untouchables. This makes Uttar Pradesh a nice setting in which

[3]The distribution of caste in the high-status sample was Brahmin (55 percent), Thakur (44), and Lala (1). For the low-status sample, it was Chamar (56), Rawat (38), Dhanuk (3), and Pasi (2). Two percent of the low-status sample identified themselves by the generic term for low caste, *harijan*.

[4]A caveat is that some authors suggest that an individual's caste identity could change in response to incentives, such as a policy favoring a caste group or occupation (Srinivas 1962, Rao and Ban 2007, Cassan 2010). However, other authors view mobility as rare in the caste system. Most people remain in one caste their entire life and marry within their caste. For the high castes (the "General castes") and the low castes (the "Scheduled castes" or SCs), the overlap of caste and occupation persists (Dube 1996, Vaid 2007). Our reading of the literature is that although some caste mobility exists, it is more likely among the middle castes, whose members did not participate in our experiments. Moreover, mobility would be a greater concern if we were looking at the effect of caste on a long-term outcome, but an individual's participation in one of our experiments occurs on only one day. The region in which we implemented our experiments is a rural region that has been part of the same state since India's independence. In rural areas caste hierarchies would be, if anything, more entrenched than in urban areas. The settlements for high castes and low castes in all of the villages from which we drew our subjects were clearly demarcated. Our observation in the field in 2005-06 was that the degrading treatment of low castes based on caste norms continued even then. This is consistent with a large-scale survey in Uttar Pradesh, which also documents the erosion of stigmatizing practices against (SCs) between 1990 and 2007 (Kapur et al. 2010). It would be hard to imagine that a high-caste person was an SC (or an SC person a high-caste person) in the recent past.

to conduct a cross-cultural study, while avoiding the risk that observed variation across populations is due to mistranslation of the game. As a result of post-independence reforms in land rights and education, there is today considerable within-caste variation in land ownership and education. Our sample contains not only poor low-caste subjects, but also many poor high-caste subjects (see Figure 4). By controlling for individual characteristics such as wealth and education, we can identify the effect of cultural differences on behavior.

Participants played multiple rounds of a coordination game known as the Stag Hunt. Each participant observed the past choices of his partner before choosing his own action. This game has been studied extensively in the literature on coordination and learning (cf. Van Huyck, Battalio and Beil 1990). Its name comes from a parable told by Rousseau: Each of several hunters has a choice between hunting a stag and hunting a hare. Regardless of what other hunters do, choosing *Hare* always results in a meal, but it is meager. In contrast, if other hunters choose *Hare*, the stag hunt is unsuccessful and a player who chooses *Stag* does not eat at all. Only if everyone chooses *Stag* is the hunt successful, with each player obtaining a rich meal. The agents face a coordination problem, as everyone choosing *Stag* and everyone choosing *Hare* both constitute equilibria, and the former Pareto dominates the latter.

The philosopher Brian Skyrms (2003, pp. xii-xiii) writes that "If one simple game is to be chosen as an exemplar of the central problem of the social contract... [t]he most appropriate choice is... the stag hunt... Social institutions enable, and to a large part consist of, correlated interactions." As the political scientist Russell Hardin (1989, pp. 116-118) emphasizes, laws and institutions can succeed only insofar as individuals coordinate on behavior consistent with those laws:

> Clearly we can live by conventions for which no one has ever voted. Indeed, one might say that the English Constitution is only a constitution so-called. it is really a set of conventions that have developed over time and that now seemingly prevail, just as the peculiar rules of the English road prevailed in England before they were made the law of the road...... The [US] Constitution of 1787 worked in the end because enough of the relevant people worked within its confines long

> enough to get it established in everyone's expectations that there was no point in not working within its confines.

We interpret the ability of players to arrive at the efficient convention of the Stag Hunt as indicative of their ability to create efficient social institutions.

In the Stag Hunt experiment, we formed three kinds of pairs: *HH*, *LH*, and *LL*, where *H* represents a member of a high caste and *L* represents a member of a low caste. A player is told nothing about his partner except that he is a man from his own village with a given caste status, *L* or *H*, described respectively by the official terms Scheduled Caste or General Caste. Thus, the experiment has variation in both caste status and social distance. The main question is, how does caste status affect the rate of coordination on the efficient outcome (*Stag*,*Stag*)?

The striking finding is that *HH* pairs are less than half as likely to coordinate on the efficient outcome as *LL*. Only 32 percent of the *HH* pairs coordinate on the efficient outcome in the last period of the pairings. In contrast, the comparable figure for *LL* pairs is 73 percent. *LH* pairs achieved efficiency at an intermediate level of 50 percent (players had two partners in turn; these percentages are an average of the final outcomes over the two pairings—see Figure 2).

What could explain the lower efficiency in coordination of the *HH* pairs? We propose that it is a consequence of the culture of honor that exists among high castes. Cultures of honor have been created independently in many places around the world (Nisbett and Cohen 1996) and have two key elements. The first is a mental framework in which putting someone in a position inferior to that warranted by his social status is construed as an insult. The second is a set of rules for how to respond to insults: In a culture of honor, individuals have an obligation to take revenge when honor has been violated (Cohen and Vandello 1998).[5]

[5] Pitt-Rivers (1966) coined the term "pecking-order theory of honor" for the notion of honor formulated by Hobbes (1651, p. 58): "honour consisteth only in the opinion of power." Pitt-Rivers (36-37) goes on to distinguish two meanings of the word honor:

> ...honour which derives from virtuous conduct and that honour which situates an individual socially and determines his right to precedence. The two senses appear to be so far removed

Case studies of north India suggest that high castes have a culture of honor, and low castes do not (see Section 2).

It is not immediately obvious how a culture of honor would impede coordination. The equilibria of the Stag Hunt give the players *equal* payoffs, and so in equilibrium the question of precedence does not arise. However, the question of precedence emerges while the players are *learning* to coordinate. Suppose Harry and Henry are playing the Stag Hunt. Harry chooses *Stag* and Henry chooses *Hare*. Thus, Harry gets a low reward, the "loser's payoff." Harry could view Henry's action as an accident of miscoordinated expectations and continue to try for (*Stag*,*Stag*) in the next period. But if Harry has a culture of honor, then Henry has violated Harry's honor, for Henry has shown him less love, concern, or fear than Harry expects.[6] Harry may retaliate by choosing *Hare* in the next period, and this can lead to more events of miscoordination and push play towards the inefficient convention.

Evidence for our explanation is that the caste difference in behavior occurs primarily in the period *after* a player receives the loser's payoff. Receiving the loser's payoff is the only one-period history after which there is a statistically significant difference in behavior between H and L. In this situation, compared to L players, H players are 29 percentage points less likely to play *Stag* (71 percent versus 42; see Table 2). This difference is even larger when we control for covariates of caste.

Further evidence comes from our follow-up study, in which participants played binary a binary choice Investment Game (also known as the Trust Game). This experiment provides evidence that the caste status H or L is not correlated with trusting behavior or trusting beliefs. If anything, we find that high-caste subjects are more trusting than their low-caste

> from one another that one may ask why they were, and still are expressed by the same word, why the languages of Europe are so determined to avoid clarity in this matter.

The two meanings are not expressed by the same word in Hindi. Honor in the second sense is *izzat* (respect), as we discuss in Section 2.

[6]This thinking would be consistent with the characterizations of honor in Hobbes (ibid., ch. 10):

> To show any sign of love or fear of another is honour; for both to love and to fear is to value. To contemn, or less to love or fear than he expects, is to dishonour; for it is undervaluing.
>
> To be sedulous in promoting another's good. . . is to honour. . . To neglect is to dishonour.

counterparts.

Our interpretation is that the caste difference in outcomes in the Stag Hunt stems from a difference in cultural construals between high- and low-caste players. After obtaining the loser's payoff, a high-caste player construes his partner's action as an insult, whereas a low-caste player construes it as an accident. We do not see a large caste difference in the Dictator Games because in order to commit an insult, the partner must have an active, strategic role.

The Stag Hunt experiment also provides a test of the theory that social distance impedes coordination. As mentioned above, the proportion of *LH* pairs that coordinate on the efficient outcome is intermediate between the proportions for *HH* and *LL*. We find no evidence that social distance, as measured by caste status, impedes efficient coordination. But a Hindu villager's "in-group"—within which he marries, socializes, and dines—is his subcaste (*jati*) and there are many *H* subcastes and *L* subcastes. Our experiment is not a test of the effect of in-group versus out-group relationships on coordination.

We formalize the connection between coordination and the culture of honor with a game-theoretic model in which a player's preferences depend on how the player construes his partner's intentions. If a player feels insulted, he prefers that his partner receives a low payoff. We incorporate these preferences into the standard learning model of stochastic fictitious play (an excellent survey is Fudenberg and Levine 1998). We find that if players react strongly to insults, then efficient coordination is unsustainable in the long run.

Our work is part of the resurgence of interest in the effect of culture on economic outcomes (e.g., Greif 1994, Fisman and Miguel 2007, Hoff, Kshetramade and Fehr 2008, Tabellini 2008, Algan and Cahuc 2009, Greif and Tabellini 2010, Alesina, Giuliano and Nunn 2011, Henrich and Ensminger 2011, Fernández 2013; surveys are Sapienza, Zingales and Guiso 2006, Nunn 2012, and Algan and Cahuc 2013 (in press)). Our paper is also closely related to the idea of fairness equilibria in Rabin (1993). Rabin considers players who care about the intentions of others. Our paper is distinguished from his in that it starts with a particular empirical puzzle—the caste difference in efficient coordination. We infer from the evidence, as well as from studies discussed in Section 2, that there are cross-cultural differences in what counts as bad intentions. Our theoretical model is different from Rabin's because we are concerned with how construals of intentions affect the process of learning to play a repeated game, rather than how such construals affect play in one-shot games.

The remainder of the paper is structured as follows. Section 2 discusses research from economics, psychology, and anthropology that bears on the culture of honor in north India. Sections 3 and 4 present the experimental design and quantitative analysis, respectively. Section 5 builds a model of learning in which a player can be in an "insulted" state. Section 6 investigates alternative explanations of our results and presents the outcome of the binary-choice Dictator Games. Section 7 concludes.

2 Mental frameworks and the culture of honor

When people make decisions, they are using a set of tools that are shaped by biology, culture, and experience. These tools include *mental frameworks* that influence how individuals perceive, assign meaning to, and respond to the objects and events around them. Such frameworks mediate between reality and individuals' minds.

Some mental frameworks are virtually universal in humans. For example, people are prone to apply causal thinking to situations that have no inherent causal meaning. Kahneman (2011) describes a classic experiment by the psychologists Heider and Simmel that demonstrates the tendency of humans to assign motive and intention in situations when it is unwarranted. Viewers are shown a brief film in which there is

> a large triangle, a small triangle, and a circle moving around a shape that looks like a schematic view of a house with an open door. Viewers see an aggressive large triangle bullying a smaller triangle, a terrified circle, the circle and the small triangle joining forces to defeat the bully... The perception of intention and emotion is irresistible; only people afflicted by autism do not experience it. *All this is entirely in your mind, of course.* (pp. 76-77, italics added)

The italicized text emphasizes that the viewers have tools for processing information, and these tools are part of an explanation of why they see what they do. The triangles and the circle are not really agents; they are perceived as agents because of the mental frameworks within which people interpret their experiences.

Many mental frameworks are not universal; different cultures or even subcultures within a culture organize their "optical predispositions" in different ways (Zerubavel 1999, p. 31). For instance, experiments demonstrate cross-cultural differences (a) in the parts of the brain that individuals use to process information and the systematic errors they make, (b) in the systems of classification that define the objects of thought and evaluation, (c) in what individuals remember of a stimulus, and (d) in what meanings they assign to a given event; a survey of cultural psychology is Heine (2012).

The influence of culture on construals, and of construals on behavior, is borne out in a celebrated set of studies of the culture of honor by Cohen, Nisbett, and colleagues. A defining characteristic of the culture of honor is that "men have to take action against insults or else lose status before their family and peers" (Cohen et al. 1999). Defined in this way, the culture of honor is broadly characteristic of the US South but not of the US North. For example, in a field experiment, Cohen et al. (1996) sent employers letters from job applicants who had committed a crime. Southern employers responded more positively than their northern counterparts when the crime was in defense of honor, whereas there was no regional difference for other kinds of crime. This finding is consistent with one Southern journalist's view of North-South cultural differences: "Northern men are spineless wimps with no honor. They will not defend themselves, their women, or their culture, assuming they have a culture" (Nethaway 1996, p. 65, cited in Cohen et al. 1999, p. 272). To investigate how the culture of honor shapes cognition and behavior in a controlled environment, Cohen et al. (1996) conducted a laboratory experiment with white male college students who were from either the South or the North. Individually, each subject was put in a situation in which he had to inconvenience another person (an actor), who responded by calling the subject an "asshole." Southerners were more likely than Northerners (a) to experience a surge in cortisol and testosterone—biomarkers of stress and readiness for aggression, respectively, (b) to feel that their masculinity was threatened, and (c) to behave aggressively in a subsequent interaction with a stranger. Further, Nisbett et al. (1993) and Grosjean (2011) find non-experimental evidence that higher homicide rates are associated with a culture of honor, and Nisbett finds that the rates are higher in the South only for argument-related homicides.

The high castes of north India have been characterized as possessing a version of the culture of honor. A typical member of the martial high castes (called Thakurs or equivalently Rajputs) in Uttar Pradesh in the 1950s was described as "brave, mettlesome, and very quick to perceive and resent an insult. It is part of his code that a slight to his prestige should be avenged" (Hitchcock 1958, p. 12). Drèze and Gazdar (1997, p. 105) find that "[a]mong the martial castes [there is] an obsession with 'honour' ... [and their values] have had a strong influence on other dominant landowning castes," i.e. the high castes. The Thakurs

> pride themselves on not compromising their honour (*izzat*), even at the cost of great hardship. For Thakur men, honor lies primarily in not doing anything that may put them in a position of subordination or moral debt... Thakurs do not engage in wage employment in the villages, as this would put them in a subordinate position (Drèze and Sharma 1998, pp. 32-33).

The sociologist Steve Derné (1992) interviewed members of middle-class households in a north Indian town in 1986-87. A 35-year old upper-caste Hindu man reported that "if a man loses his honor [*izzat*], society kicks him away and he becomes utterly worthless" (ibid., p. 265). The need to maintain honor leads to constant vigilance. The anthropologist David Mandelbaum (1993) reported that although it is hard-earned, "*izzat* does not keep well; it has to be continually reaffirmed in practice, reinforced in action, defended against challenge and rewon and advanced in competition" (ibid., p. 23). In this region, the contenders for power might have learned to defend their honor as a matter of survival. As Drèze and Gazdar (p. 103) note, "land ownership has often had to be won or defended through violent means in the past," since the fertile Gangetic plains of north India were frequently exposed to raids and invasion. The cultural patterns persist even though they are no longer functional because they are perpetuated by social mechanisms. As Derné reports, (pp. 277-9), men "see their family honor as important for their success. . . Men who dishonor themselves jeopardize marriage prospects for themselves, their children, and their brothers and sisters."

Concern with honor in the sense of *izzat* is more prevalent among high castes than among low castes. The concept of honor implies that identity is linked to institutional roles (Berger, Berger and Kellner 1975, p. 90). For the high castes, honor is linked to the role of landowner and to the institutions of caste and clan that maintain their power in the community. Low castes historically had no power or wealth to defend and no means to accumulate it by marriage, and thus little to gain from a culture of honor. Restrictions barring land ownership to low castes were removed only 60 years ago with Indian independence (Galanter 1984, p. 15). Many low-caste individuals have to work as wage laborers to survive. It is still a routine event for a low-caste person to be publicly insulted, and for claims by low-caste persons for equal status with high-caste persons to evoke a violent response.[7] In sum, because the social structure assigns the low castes institutional roles far inferior to those of the high castes, the low castes do not have either the need for or the opportunity to develop the culture of honor that is typical of high castes (see Khare 1984).[8]

The example by Heider and Simmel discussed in this section demonstrates that mental frameworks may lead people to interpret reality with much more flourish and spin than is warranted by the facts. Mental frameworks can trigger emotions that have directive force on behavior, and it seems that individuals frequently cannot suppress these emotions, even when failing to do so is costly. Cultures may differ with respect to what individuals perceive as insults, what people view as an appropriate response to an insult, and what behavior the insults tend to trigger. It should come as little surprise that mental frameworks have

[7]See New York University School of Law Center for Human Rights and Global Justice and Human Rights Watch (2007).

[8]A measure of the historical differences in culture between high and low castes is the difference in the female-male ratio and gender relations. Drèze and Sen (2002, Table 7.1) calculate that low castes had much above-average female-male ratios in 1901 compared with those of the high castes. The ratio was 0.970 for the low castes, compared to 0.887 for the high castes. By 1981, however, the gap had narrowed, which may reflect the gradual spreading to other castes of the patriarchal norms of the high castes.

consequences for coordination and social efficiency. This is the subject we explore in the rest of the paper.

3 Experimental design

3.1 Design and research questions

In order to investigate the effect of caste culture on coordination, we implemented a repeated Stag Hunt with low-caste and high-caste men. Figure 1 shows the stage game.

	Stag (6)	*Hare* (2)
Stag (6)	(10,10)	(3,7)
Hare (2)	(7,3)	(7,7)

Figure 1: Payoffs of the period game (in rupees)

Every period of play, each player was given an endowment of six rupees, denoted ₹6, and asked to choose between contributing the entire endowment to a common pool or only contributing ₹2. On the other hand, if a player contributed ₹2, he received ₹3 back, regardless of the choice of the other player, for a total payoff of ₹7. If a player contributed ₹6, he received ₹10 if the other player also contributed ₹6, but received only ₹3 if the other contributed only ₹2. In this case, the player's payoff of ₹3 was lower than the partner's ₹7. We will refer to this as the "loser's payoff." The loser's payoff would be a loss in the sense of (Kahneman and Tversky 1979).

It is clear how to map this game to the Stag Hunt: Playing *Stag* corresponds to contributing ₹6, and *Hare* corresponds to contributing ₹2. The equilibria are (*Stag*,*Stag*) and (*Hare*,*Hare*), and the former Pareto dominates the latter.

For the remainder of the paper, we eschew the labeling of actions as "6" and "2" in favor of the more colorful and descriptive *Stag* and *Hare*. However, we wish to emphasize that the experiment was explained to the subjects with neutral language that should not have introduced any particular interpretation of actions or outcomes. Moreover, the same individual explained the game to all of the players, H and L, in all of the sessions.

Experimental subjects were drawn from the extreme top and bottom of the caste hierarchy. Recall that H denotes a member of a high caste and L denotes a member of a low caste. Every subject played the game in Figure 1 ten times and received feedback after each period about the partner's choice. Five of the periods were played with a member of a caste of the same status (the "single-caste treatment"), denoted by either HH or LL. The other five were played with a member of a caste of the opposite status (the "mixed-caste treatment"), denoted by LH. We refer to a set of five periods played by two individuals as a "fixed pairing." It is important to note that a subject had the same partner for all five periods of a fixed pairing. We used a counterbalanced design to avoid confounding the treatment and the order of play. In Cohort I, subjects were organized into LL and HH pairs for periods 1-5, followed by LH for periods 6-10. In Cohort II, the order was reversed: LH for periods 1-5, and LL and HH for periods 6-10. Subjects were randomly assigned to a cohort.

The experiment was designed to answer two questions. First, *does a community's culture affect the ability to establish an efficient convention?* Any number of cultural differences—in trust, social preferences, and the way that individuals conceptualize the outcomes (e.g. is the loser's payoff an insult?) could lead to differences in coordination outcomes.

Second, *does social distance affect the ability to coordinate?* By social distance, we mean the perceived status difference between individuals. A decrease in social distance leads, in general, to an increase in empathy, as in Bernhard, Fehr and Fischbacher (2006), Charness and Gneezy (2008), and Chen and Li (2009). Chen and Chen (2011) find that inducing group identities in the lab leads to the selection of more efficient outcomes in coordination games. However, an offsetting consideration is that the social distance between the two players may affect the players' degree of aversion to obtaining the loser's payoff. Just as workers in non-unionized firms appear to make only within-firm comparisons of wages and not across-firm comparisons (Bewley 1999), a high- (low-) caste individual concerned with his relative position may feel that concern much more if the other player is high- (low-) caste. Loewenstein, Thompson and Bazerman (1989) find that aversion to disadvantageous inequality is greater when social distance is smaller. It is thus plausible that *LH* pairs would establish an efficient convention more easily than *HH* and *LL* pairs, since the sting of the loser's payoff might be less sharp in the former treatment than the latter.

3.2 Experimental procedures

As noted above, a single individual explained the game to the subjects. Before the game began, he explained that (a) each subject would interact with a single anonymous person from his village over five periods of the game, and (b) then he would interact with another anonymous person over five more periods. One monitor was assigned to each player. The monitor stayed with the player until the game was over, and the player could not talk with anyone else until the end of the game. However, the monitor never saw the decisions of the individual or his opponent. This was done to ensure that the monitor could not indicate a judgment of approval or disapproval.

At the beginning of period 1 and again in period 6, the monitor assigned to a subject gave him a "game box," a plastic utility box with three rows and five columns in which the subject was able to see his own payoffs in each period and *the full history* of the game with the current partner. Just before the first round with a given player, the monitor told the player that he did not know who the other player was, but he knew his caste status. He communicated the caste status using the terms "General Caste"—the official term for castes at the top of the status hierarchy—and "Scheduled Caste"—the official term for the castes formerly known as "untouchables."[9]

During the game, players used plastic tokens to represent rupees. After the tenth period, players received their payoffs in private and in cash, with an exchange rate of one rupee per

[9] General Castes and Scheduled Castes each make up about one-fifth of the population of the state of Uttar Pradesh.

token. We also administered a brief sociodemographic survey on their education, wealth, and subcaste. .

3.3 Subjects

We went each day to a different village in the district of Lucknow, the block of Bakshi Ka Tulab, to recruit subjects. Within a village, we drew about 12 subjects, of which an equal number were from high castes and low castes. Subjects were recruited using systematic sampling. As an example, if 30 low-caste households lived in the village, every fifth household was chosen, starting from the first that was encountered. Every selected household volunteered one adult male to participate in the game. With our procedure, there was no possibility of contagion among subjects within a village. In the *LH* treatment, *L* and *H* players were kept in two separate locations, within their own neighborhoods. In the *HH* and *LL* treatments, partners were kept in separate areas.

122 subjects played the Stag Hunt in September 2005.[10] An experimental session lasted approximately four hours. Mean earnings were equivalent to 1.5 times the daily unskilled wage (₹77 of the possible maximum of ₹100, which was ₹10 per period for 10 periods).

4 Results of the Stag Hunt

In this section, we present the results of the Stag Hunt experiment. Before giving our econometric analysis, we discuss four examples of histories of play. This will give the reader a feel for the learning over the course of the game.

In Table 1, we report the history of play for four individuals, whom we call Harry, Henry, Louis, and Luke. Initially, Harry and Henry are partners, and Louis and Luke are partners.

Consider first Harry and Henry, who start in a state of miscoordination in which Henry receives the loser's payoff. This is followed by (*Hare*,*Hare*), and after one more episode of

[10]Two subjects completed only periods 1-5, so for periods 6-10 there are 120 observations.

Table 1: Examples of the history of play in the Stag Hunt. A player's name is listed directly above or below the player's actions. Efficient coordination in the stage game is circled with solid lines, and inefficient coordination is circled with dashed lines.

Period									
1	2	3	4	5	6	7	8	9	10
		Harry					Harry		
Hare	Hare	Stag	Hare	Hare	Hare	Stag	Stag	Stag	Hare
Stag	**Hare**	**Hare**	**Hare**	**Hare**	STAG	STAG	STAG	STAG	STAG
		Henry					LOUIS		
		LOUIS					**Henry**		
HARE	STAG	STAG	STAG	STAG	**Hare**	**Hare**	**Hare**	**Hare**	**Hare**
Hare	*Hare*	*Stag*	*Stag*	*Stag*	*Hare*	*Stag*	*Hare*	*Hare*	*Hare*
		Luke					*Luke*		

miscoordination they appear to converge on the inefficient equilibrium. Louis and Luke, on the other hand, begin with inefficient coordination, but by the end of the first fixed pairing are playing (*Stag*,*Stag*).

At this point, they swap partners. Louis, who just concluded three rounds of (*Stag*,*Stag*), continues to play *Stag*. Louis' new partner Harry, on the other hand, is coming off of two consecutive rounds of (*Hare*,*Hare*), and starts off with *Hare*. Nonetheless, Louis persists with *Stag*, and by period 10 they have achieved three periods of efficient coordination.

In contrast, Luke starts off with *Hare* just as he did in the first pairing. Henry, who perhaps recalls the inefficient coordination of periods 1-5, also plays *Hare*. After one more round, they appear to establish the inefficient convention.

This example nicely demonstrates some general features of the data. The pattern of play varied widely across trials, with some pairs quickly achieving efficient coordination, and others miscoordinating and then settling on inefficiency. Moreover, outcomes were frequently very different between the two fixed pairings, with inefficiency followed by efficiency or vice versa. Another feature is that there is a fair amount of noise in individual behavior, as evidenced by Harry's play of *Hare* in the last period, even though the three previous outcomes were (*Stag*,*Stag*). Finally, these histories are suggestive of the general pattern that we document below: Harry and Henry belong to high-status castes and as a pair they never had an outcome of (*Stag*,*Stag*). In contrast, Louis and Luke are low caste, and as a pair they were able to move from (*Hare*,*Hare*) to (*Stag*,*Stag*).

We now consider, in turn, pair outcomes and individual behavior.

4.1 Pair outcomes

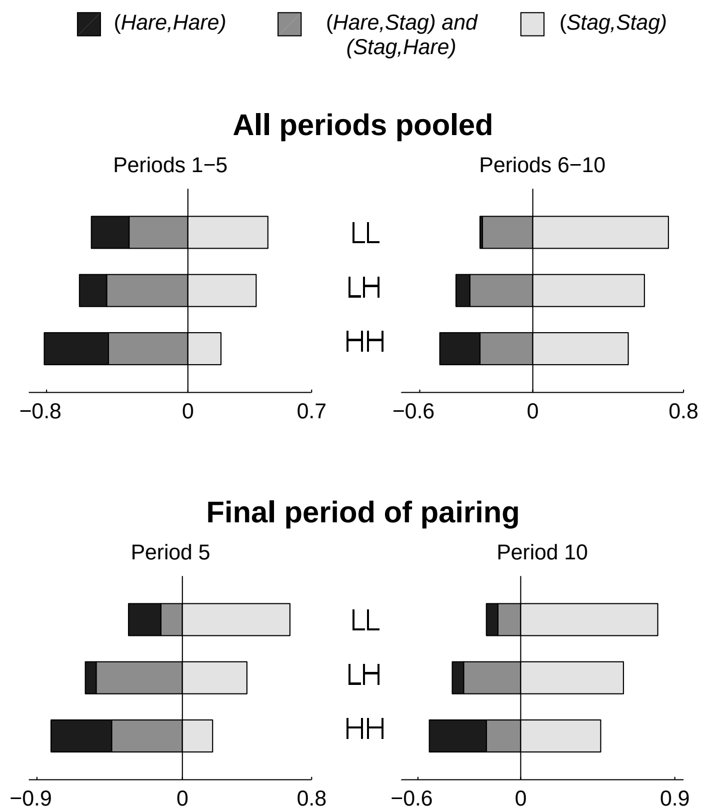
Figure 2 shows the average outcomes for the three types of pairs—*LL*, *LH*, and *HH*. The top panel pools all periods of a fixed pairing, and the bottom panel shows results for the final period of fixed pairings. The number of periods spent in (*Stag*,*Stag*) is highest for *LL*, intermediate for *LH*, and lowest for *HH*. For the number of periods spent in (*Hare*,*Hare*), the pattern is reversed: it is higher for *HH* than for *LL*, though *LH* spend more periods in miscoordination.

In particular, for Cohort I, who had the single-caste treatment for periods 1-5, the percentage of *LL* pairs with an outcome of (*Stag*,*Stag*) was 45 percent (34 out of 75 periods), compared to only 19 percent (15 out of 80) for *HH* pairs. Cohort II's results were similar for periods 6-10. The comparable figures were 72 percent (54 out of 75) for *LL* and 51 percent (38 out of 75) for *HH*. The difference between *LL*, *LH*, and *HH* in both panels is significant (χ^2 test, $p < 0.0001$).

Figure 2 suggests that social distance does not impede coordination. In periods 1-5, the proportion of *LH* with outcome (*Stag*,*Stag*) was 39 percent (58 out of 150), which was greater than the proportion for *HH*, and the difference is significant ($p < 0.0001$). The difference between *LH* and *LL* is not significant. The same general pattern occurred in periods 6-10, with the difference between *LH* and *HH* significant at 5 percent.

All the above results continue to hold if we consider only the *last* period of fixed pairings, shown in the lower panel. As before, *LL* were the most likely to have outcome (*Stag*,*Stag*), and *HH* were the least likely. For example, in period 10, 80 percent (12 out of 15) of *LL* pairs

Figure 2: Outcomes for fixed pairs in the Stag Hunt. Black bars are proportional to the number of pairs with outcome (*Hare*,*Hare*), gray bars are proportional to the number of pairs with outcomes (*Hare*,*Stag*) or (*Stag*,*Hare*), and light gray bars are proportional to the number of pairs with outcome (*Stag*,*Stag*).



had outcome (*Stag*,*Stag*), compared to 47 percent (7 out of 15) of *HH* pairs. For *LH*, the comparable figure of 60 percent (18 out of 30) was again intermediate between those of the two single-caste treatments, which suggests that what matters for efficient coordination is the caste status of the individual, not the status difference between the players. The outcome differences between *HH* and *LL* are significant in the last period of a pairing ($p < 0.0001$ for period 5, and $p < 0.01$ for period 10).

We next consider whether individuals are learning an efficient convention over time. Van Huyck, Battalio and Beil (1990) found that the frequency of efficient coordination between fixed pairs in a Stag Hunt increased over the first five periods, and then did not deviate from that high level over the next five periods of a ten-period game. In their experiment, by period five, 11 of 14 pairs had established an efficient convention in the sense that after that period, deviation from efficiency was rare and not sustained.[11] In Figure 3, we report the period-by-period outcome shares for each pair type.

Consider first the time trends for the first five periods, shown in the top panel. Among *LL* pairs, the proportion with outcome (*Stag*,*Stag*) increased over five periods from 27 to 67 percent, a significant difference ($p < 0.001$). In contrast, among *HH* pairs, there was no apparent time trend to the proportion with outcome (*Stag*,*Stag*). A plurality of *LH* pairs were in the miscoordination or inefficient outcomes at period 5, so there is no evidence of emergence of a convention.

For periods 6-10, shown in the bottom panel, an efficient convention was established among *LL* pairs from the initial period. In every period, between 67 and 80 percent of pairs had outcome (*Stag*,*Stag*). Note also that no *LL* or *HH* pairs had an outcome of (*Hare*,*Hare*) in the initial period. This remained true for *LL* pairs with a single exception. However, over time the *HH* pairs learned to avoid the disequilibrium outcome of the stage game by settling on the inefficient equilibrium (*Hare*,*Hare*).[12] Thus, the fraction of pairs with that outcome rose over time, from zero in the initial period, to 20 percent for the next two periods, to 33 percent in the last two periods. For *LH*, a plurality were in the efficient equilibrium in period 10, though this share is not significantly different from 50 percent at the 5 percent level of significance.

To sum up, most *LL* pairs established the efficient convention; most *HH* pairs did not. In period 1, The behavior of *HH* pairs differed only marginally from that of *LL* pairs ($p = 0.08$) but evolved differently over time. The outcome for *LH* pairs was intermediate between the outcomes for *LL* and *HH*.

[11] It was because of their finding that in our experimental design, we limited the number of periods for fixed pairs to five. The need to have messengers running between locations made each period time-consuming, so that a large number of periods was infeasible.

[12] We know that the *H* players who move *to* (*Hare*,*Hare*) are generally moving *from* (*Stag*,*Hare*) or (*Hare*,*Stag*), since almost 90 percent of *H* players in (*Stag*,*Stag*) in a given period of a fixed pairing continue to play *Stag* in the next period. See Table 2.

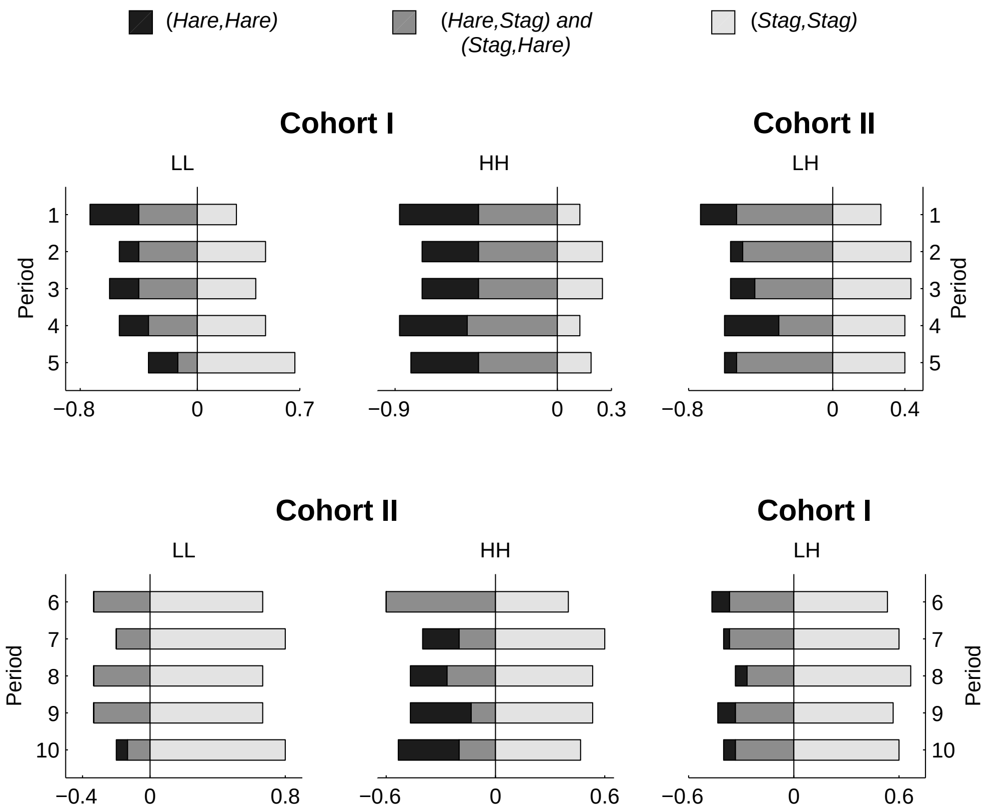


Figure 3: Period-by-period outcomes for fixed pairs in the Stag Hunt.

4.2 Individual behavior

We now trace the caste gap in outcomes for fixed pairs to the caste gap in individual responses to the loser's payoff. We start with summary statistics of individual behavior, and then present a regression analysis that controls for covariates such as wealth and education.

4.2.1 Preliminary analysis

It is natural to think that players use the history of the game in order to decide their next action. Studying how they use this history can help us disentangle the sources of the difference in pair outcomes. Conditioning on the entire history is both impractical and likely to be unedifying, given the huge number of possible histories. As an alternative, we look at individual responses after one-period histories. Certainly, the outcome of the last period is the freshest in a player's mind, and is likely to be the most salient component of the history for determining future play. If we condition on one-period histories, how do players' choices vary with caste status?

Table 2 presents this information. In the expression (x, y), x denotes the player's own action and y is the partner's action. Column 1 reports the proportion of individuals that chose *Stag* in the initial period of play and columns 2-5 report that proportion for all subsequent periods, conditional on the outcome in the preceding period. For instance, the last column shows that after the event (*Stag*,*Stag*)—the Pareto-dominant equilibrium of the stage game—players in all three pair types played *Stag* at virtually the same rate, which was nearly 90 percent.

The largest caste difference occurs after the event (*Stag*,*Hare*), in which the player received the loser's payoff. The figures are reported in column 4: 71 percent of L players chose *Stag* in the next period, but the corresponding percentage for H players is 42! This difference is statistically significant ($p < 0.0001$), and represents the largest gap between H and L in the rate of choosing *Stag* after any one-period history, and the only H—L gap that is statistically significant according to a two-sided t-test at $p < 0.05$.

We also note the large difference after history (*Stag*,*Hare*) between H in *HH* and H in LH, which is statistically significant ($p < 0.05$). This is additional evidence that the difference in efficient coordination rates is attributable to caste characteristics, which may be compounded when there are two H players rather than one.

In period 1, H players played *Stag* at a lower rate than L players, and the difference is significant at $p < 0.05$. However, the difference between *HH* and *LL* is not significant ($p = 0.16$). Thus, the difference seems consistent with multiple explanations: Do H and L have different initial beliefs or levels of trust, or are L players being deferential to H in the LH treatment? We will revisit this issue in the regression analysis.

Table 2: Proportion of players who chose *Stag*

	Period 1	If the preceding period's outcome was:			
		(*Hare*,*Hare*)	(*Hare*,*Stag*)	(*Stag*,*Hare*)	(*Stag*,*Stag*)
Player is low caste:	**0.57**	**0.52**	**0.58**	**0.71**	**0.88**
In *LL*	0.47	0.50	0.66	0.68	0.86
In *LH*	0.67	0.53	0.47	0.72	0.89
Player is high caste:	**0.37**	**0.31**	**0.46**	**0.42**	**0.87**
In *HH*	0.34	0.26	0.47	0.32	0.88
In *LH*	0.40	0.40	0.46	0.56	0.86
Observations:	122	154	181	181	452

4.3 Regression analysis

Caste is correlated with education and wealth. Thus, it is important to consider whether these covariates can explain the differences in individual behavior. We collected data in post-play interviews on land ownership, education, and house type.[13] Figure 4 shows the distribution of these covariates by caste status. Within our sample, H individuals are more likely to own land, have a high school education, and live in a house that is not built out of mud. However, there is still substantial within-group variation, and the within-group distributions overlap. For example, 14.5 percent of L have completed high school, compared to 17.7 of H; 63 percent of L individuals live in a mud house, compared to 19 of H.

We use regression analysis to show that the caste difference in playing *Stag* is robust to the addition of controls for individual characteristics. We estimate choice probabilities conditional on the outcome in the preceding period, which are analogous to the summary statistics presented above. The choice probability is allowed to depend on the player's own caste and partner's caste, as well as on the covariates for education and wealth.

[13] Using data from the 1997-98 Survey of Living Conditions in Uttar Pradesh, we find that land ownership, housing, wealth, and education have a large and significant impact on adult per capita consumption. Together, these variables explain a substantial part (between 30 and 40 percent) of the variation in consumption for both high- and low-caste individuals (Hoff, Kshetramade and Fehr 2011).

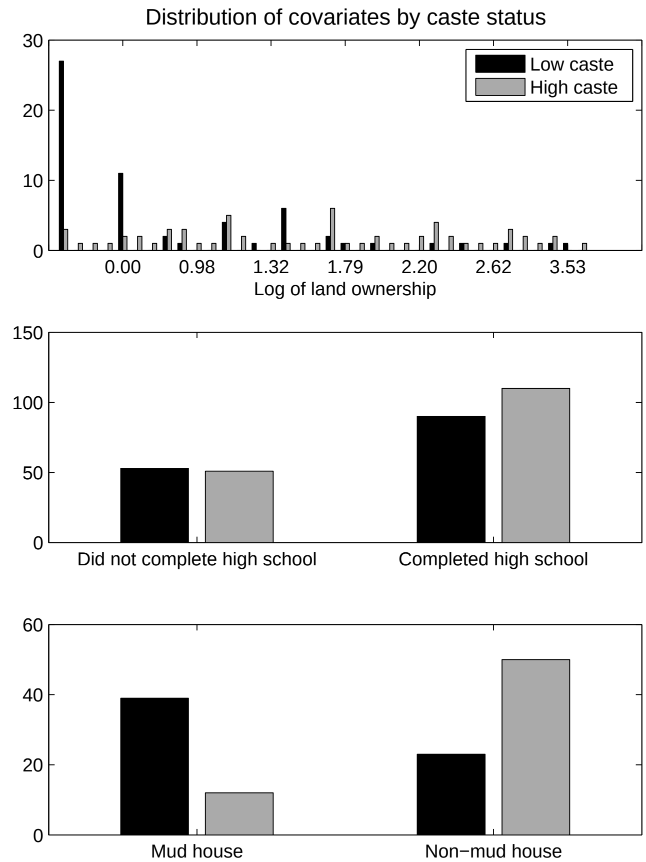


Figure 4: Distribution of covariates by caste status

In the regression analysis, we also attempt to better control for the influence of history on the choice of action. By conditioning only on the last period's outcome rather than the full history, we are omitting a large amount of information that could be important for understanding a player's choice. For example, some individuals may be more trusting or have more pro-social preferences, which would be correlated with efficient outcomes. Furthermore, a player's experience with the first partner may bias the outcome of the second pairing. We define a player's *type* to be the action he chooses in the first period of play. In particular, type is 1 if the player chose *Stag* in the fist period of the game, i.e. the first period of the first pairing. This measure of type is exogenous from the perspective of subsequent play. There are three types that could influence a player's behavior. First, his own type and his current partner's type have direct influence and influence through the history of play. Also, the influence of the first pairing history on play in the second pairing can be proxied by the *first* partner's type, and the second partner's history can be proxied for by the second partner's first partner's type. We should note that all of our results are robust to the omission of these proxies.

Thus, we study the following conditional probit model:

$$\begin{aligned} &\text{Prob}(a_t = \mathit{Stag} | h_{t-1}) \\ &\quad = \Phi\Big(\gamma_1[HH] + \gamma_2[H \text{ in } LH] + \gamma_3[L \text{ in } LH] \\ &\qquad + \delta_1[\text{Own type}] + \delta_2[\text{Partner's type}] + \delta_2[\text{First partner's type and } t > 5] + \\ &\qquad + \delta_3[\text{Second partner's first partner's type and } t > 5] + X\beta + \epsilon\Big) \end{aligned} \tag{1}$$

where X is a vector of controls for education and wealth, and ϵ is a random noise term.

The omitted category in all regressions is a player in LL. Also, in all regression tables, we report the marginal effect of a variable. Thus, the reported coefficient γ_1 represents the marginal difference in probability of playing *Stag* for a player in HH, relative to a player in LL. Similarly, γ_2 is the marginal difference between H in LH versus a player in LL, and so on. To compute the difference between the single-caste and mixed-caste treatments, we would look at γ_3 for L, and $\gamma_2 - \gamma_1$ for H.

Table 3 reports the results of probit regressions for observations in periods 2-5 and 7-10. All regressions cluster disturbance terms.[14] Column 1 pools all observations over these periods. After controlling for education and wealth, we find that players in HH are 21 percentage points less likely to choose *Stag* than are players in LL.

In columns 2-5, we run the same regression conditional on the last period outcome. We see that the caste effect in the single-caste treatment is insignificant in all regressions *except*

[14] Players were organized into fixed pairings in groups of four, such that the initial pairings were (A, B) and (C, D) and the second pairing was (A, C) and (B, D). Standard errors are clustered by these four-tuples of players.

Table 3: Regression analysis for periods 2-5 and 7-10.

	Dependent variable = 1 if plays *Stag* conditional on one-period history.				
	Outcome in preceding period				
	All	(*Hare*,*Hare*)	(*Hare*,*Stag*)	(*Stag*,*Hare*)	(*Stag*,*Stag*)
	(1)	(2)	(3)	(4)	(5)
HH	-0.211**	-0.198	-0.152	-0.413***	0.0147
	(0.0988)	(0.203)	(0.134)	(0.127)	(0.0391)
H in *LH*	-0.0722	-0.0663	-0.131	-0.180	0.0165
	(0.0883)	(0.233)	(0.132)	(0.140)	(0.0502)
L in *LH*	-0.000403	0.0761	-0.139	0.0826	0.0154
	(0.0457)	(0.180)	(0.132)	(0.125)	(0.0318)
Own type	0.155***	-0.0695	0.180	-0.0516	0.114**
	(0.0436)	(0.110)	(0.116)	(0.0640)	(0.0452)
Current partner's type	0.105***	-0.0192	0.0825	0.00266	0.00648
	(0.0389)	(0.111)	(0.0901)	(0.0944)	(0.0379)
First partner's type and $t > 5$	0.117***	0.475***	0.0986	0.132*	0.0624*
	(0.0271)	(0.166)	(0.151)	(0.0740)	(0.0341)
Second partner's first partner's type and $t > 5$	0.0137	-0.136***	0.0103	-0.0106	0.0126
	(0.0100)	(0.0336)	(0.0229)	(0.0199)	(0.00768)
High school	-0.0398	-0.135	-0.0803	0.193**	-0.0681
	(0.0763)	(0.151)	(0.133)	(0.0955)	(0.0605)
Non-mud house	0.00763	0.0555	0.0331	0.0723	0.0162
	(0.0500)	(0.118)	(0.108)	(0.0917)	(0.0466)
Land	-0.00154	-0.00405	-0.00217	0.00455	-0.000956
	(0.00594)	(0.00686)	(0.00739)	(0.00520)	(0.00239)
Observations	968	154	181	181	452

Robust standard errors in parentheses

*** Significant at the 1 percent level, ** Significant at the 5 percent level , * Significant at the 10 percent level in column 4,[15] in which the previous period's outcome was (*Stag*,*Hare*). For observations in this column, the player received the loser's payoff in the previous period. In this situation, a high-caste player in *HH* is 41 percentage points less likely to continue playing *Stag* than a low-caste player in *LL*. The result is significant with $p < 0.001$. This supports the hypothesis that the loser's payoff is particularly salient to the high-caste players, and can explain the caste difference in pair outcomes.

Compared to the benchmark case of *LL*, a smaller and also insignificant reduction in the probability of playing *Stag* occurs for high-caste individuals in *LH*.[16] However, before concluding that it is the partner's caste that affects the choice of action estimated in column 4, we have to ask whether the observed difference between *H* in *HH* versus *LH* reflects a difference in the history of play preceding period $t - 1$ in the single- versus mixed-caste treatment. For instance, do high-caste players more often experience the loser's payoff in *HH* than in *LH*?

We find that they do not. Considering all observations in column 4 of Table 3, and looking back over all histories with a given opponent, we find that 43.6 percent of *H* players in *HH* received the loser's payoff ($n = 32$) compared to 45.7 percent of *H* in *LH* ($n = 47$). The difference is statistically insignificant. This finding suggests that *H* deviate from the comparison group *LL* more in the the single-caste treatment than in the mixed-caste treatment. Contrary to one's intuition based on fractionalization studies, social distance does not exacerbate, but instead mitigates the tendency of *H* to play *Hare* after obtaining the loser's payoff.

In nearly all of the regressions we have run, the covariates on education, the quality of one's house, and land are jointly and individually insignificant. The covariates are jointly significant only for column 2 of Table 3. The absence of any pattern of significance among these controls leaves us with no reason to believe that it is the covariates of caste, rather than

[15] In particular, the coefficients on *L* in *LH* are insignificant, so the apparent "deference" of *L* players to *H* players in the summary statistics is not robust to the addition of controls.

[16] The difference between *H* in *HH* and *H* in *LH* is significant at $p < 0.01$ in pooled periods 2-5 and 7-10, but this difference is insignificant for periods 2 and 7.

Table 4: Regression analysis for period 1.

Dependent variable = 1 if plays *Stag*		
	(1)	(2)
HH	-0.125	-0.165
	(0.127)	(0.138)
H in *LH*	-0.0671	-0.111
	(0.133)	(0.142)
L in *LH*	0.203	0.215*
	(0.128)	(0.129)
High school		0.169
		(0.129)
Non-mud house		0.0519
		(0.0820)
Land		0.00356
		(0.00418)
Observations	122	122

Robust standard errors in parentheses
*** Significant at the 1 percent level,
** Significant at the 5 percent level,
* Significant at the 10 percent level

caste culture, that explains the difference in behavior. We note that education is individually significant and positive in column 4. This may indicate a mitigating effect of education on the reaction to the loser's payoff.

Finally, we return to the issue of whether or not high- and low-caste behavior varies in the initial period of play. Are low-caste players more trusting; that is, do they harbor more optimistic views about their partner's choice of action? We estimate a similar model as (1), but only for period 1, and reput the results in Table 4. While *H* players are less likely than players in *LL* to play *Stag* in the initial period, the coefficients on *HH* and *H* in *LH* are individually and jointly insignificant with the inclusion of covariates. We note that *L* in *LH* are more likely to play *Stag* in *LH* than in *LL* (the difference is marginally significant), which is counter to the fractionalization hypothesis. However, all differences are modest relative to the variation conditional on one-period histories. Thus, it is fair to say that the most prominent difference between H and L occurs *after* the first period of play and, in particular, after the loser's payoff.

4.4 Summary

Our empirical findings can be condensed into three results:

Result 1. *After receiving the loser's payoff, high-caste players are significantly less likely to play* Stag *than low-caste players. The caste gap in the probability of playing* Stag *in the initial period and after other one-period histories is of smaller magnitude and generally not significant.*

Result 2. HH *pairs are less likely to coordinate on the efficient outcome than* LH*, who are in turn less likely to establish the efficient convention than* LL*.*

Result 3. *Conditional on a player's own caste status, greater social distance from the other player never reduces the probability of playing* Stag*, and in some cases increases it.*

5 Learning with insults

In this section, we interpret Result 1 as evidence that the preferences of high-caste players are state-dependent. We build a simple model of state-dependent preferences and investigate its implications for how individuals learn how to play the Stag Hunt, in a fixed pairing. We show that the caste gap in the play of *Stag* after the loser's payoff (Result 1) and the ranking in efficiency of coordination ($HH < LH < LL$, Result 2) can be explained by state-dependent preferences.

5.1 Model

In each period $t = 1, 2, \ldots$, individuals $i = 1, 2$ play a Stag Hunt. Players have a type H or L. We will consider models in which the pair's types are $\mathscr{P} \in \{LL, LH, HH\}$.

State-dependent preferences

Players can be in one of two states—normal (N) or insulted (I). In state N, preferences over outcomes are as given in Figure 1: player i's utility is precisely his monetary payoff π_i. However, after receiving the loser's payoff, an H player enters state I, in which he believes that his partner has insulted him. In state I, utility changes to $\pi_i - \beta\pi_{-i}$,[17] with $\beta > 0$. In words, an insulted player receives utility from his own monetary payoffs and disutility from his partner's monetary payoffs.[18] Player i's state in period t is $s_t^i \in \{N, I\}$.

The learning model

To study the impact of these preferences on coordination, we embed them in the well-known learning model of stochastic fictitious play. With fictitious play, each player keeps track of the history of observed outcomes. This historical distribution of play is used to "forecast" what others' actions will be in the next period. The players use an approximate best response to this forecast, which is interpreted as a best response when the payoffs are randomly perturbed. For games like the Stag Hunt, play (and forecasts) converge to a long-run steady-state in which forecasts are correct, and play is close to either the efficient or inefficient Nash equilibrium of the period game.

Forecasts

Let $a_t^i \in \{\mathit{Stag}, \mathit{Hare}\}$ be player i's action in period t. Player i's forecast of the probability that player $-i$ will play *Stag* in the next period is p_t^i, which starts at the initial value p_1^i. p_t^i

[17] We use $-i$ to denote $j \neq i$.

[18] This corresponds to spiteful preferences in the sense of Fehr and Schmidt (1999) and Charness and Rabin (2002).

is defined by the formula

$$p_t^i = \frac{1}{t}\left(p_1^i + \sum_{\tau=1}^{t-1}[a_\tau^{-i} = Stag]\right) \tag{2}$$

where $[a_\tau^{-i} = Stag]$ is 1 if player $-i$ played *Stag* in period τ, and 0 otherwise. This formula can be written recursively

$$p_t^i - p_{t-1}^i = \frac{1}{t}\left([a_{t-1}^{-i} = Stag] - p_{t-1}^i\right) \tag{3}$$

Best responses

A player's action is a noisy best response to the forecast. In particular, the probability that player i plays *Stag* when his state is s and his forecast p is given by exponentially weighted utilities

$$b_s(p) = \frac{\exp\left(u(Stag|p,s)\right)}{\exp\left(u(Stag|p,s)\right) + \exp\left(u(Hare|p,s)\right)} \tag{4}$$

where $u(a|p,s) = pu(a, Stag|s) + (1-p)u(a, Hare|s)$ is a player's expected utility from the action a, with the expectation taken with respect to the forecast over the other player's action. The smoothing is a reduced form for the best response when payoffs are randomly perturbed every period. We show in the Appendix that $b_N(p) > b_I(p)$ for all p and $\beta \in (0,1)$. For future reference, our notation is summarized in Table 5.

Our analysis now turns to the characterization of long-run behavior.

5.2 Long-run behavior

Suppose that players' forecasts were *fixed* at (p^i, p^{-i}). In the long-run, what is the average probability that player i chooses *Stag*? If player i is L, then every period he must be in state N and he chooses *Stag* with probability $b_N(p^i)$. But if he is H, then the answer depends on the relative frequency of states I and N. The more frequent is I, the lower will be the

Table 5: Notation

$\mathscr{P}$	A pair type, either *LL*, *LH*, or *HH*.
a_t^i	Player i's action in period t, either *Stag* or *Hare*.
s_t^i	Player i's state in period t, either N or I.
β	Disutility of other player's monetary payoffs in state I.
p^i	A player's forecast of their partner's action. The historical average of play.
b_s	Probability of playing *Stag* in state s.
$\Pi_{\mathscr{P}}(s^1, s^2)$	The long-run distribution of states, when forecasts are (p^1, p^2) (implicit).
$\bar{b}_{\mathscr{P}}(p^i, p^{-i})$	The long-run average probability of playing *Stag*.
$p^C_{\mathscr{P}}(p^{-i})$	The forecast of player i that is consistent with the long-run average play when the other player's forecast is p^{-i}.

average probability of playing *Stag*. Let $\bar{b}_{\mathscr{P}}(p^i, p^{-i})$ be the average probability with which a player i chooses *Stag* when forecasts are fixed at (p^i, p^{-i}) and in the pair type $\mathscr{P}$.

Recall that for each pair type, players' states evolve as a Markov chain. Thus, the long-run frequencies of states converge to the steady-state distribution of the chain. For pair type LL, the only state pair that can arise is (N, N). But for LH, state pairs (N, I) and (N, N) can arise, and for model HH, we can have the three pairs (I, N), (N, I), and (N, N).[19] The transition probabilities are all positive so there is a unique steady-state frequency of the pair (s^1, s^2), which we denote $\Pi_{\mathscr{P}}(s^i, s^{-i}, p^i, p^{-i})$. For example, $\Pi_{LL}(N, N, p^i, p^{-i}) = 1$. Thus, to return to the initial question, the long-run average probability of player i choosing *Stag* is

$$\bar{b}_{\mathscr{P}}(p^i, p^{-i}) = \Pi_{\mathscr{P}}(I, N, p^i, p^{-i})b_I(p^i) + (1 - \Pi_{\mathscr{P}}(I, N, p^i, p^{-i}))b_N(p^i)$$

For example, $\bar{b}_{LL}(p^i, p^{-i}) = b_N(p^i)$.

In our thought experiment, we held the forecasts fixed, but actually they are evolving

[19]Since a player only enters state I after receiving the loser's payoff, and only one player can receive the loser's payoff at a time, it is impossible for *both* players to be in state I.

over time according to (3). The literature on stochastic approximation has shown that in the long-run, the forecasts trend towards an average probability of *Stag* that is precisely $\bar{b}_{\mathscr{P}}$. Moreover, since the weight of $\frac{1}{t}$ is decreasing, the difference equation (3) "slows down", and asymptotically the movement of forecasts approaches the dynamical system:

$$\frac{dp^i}{dt} = \bar{b}_{\mathscr{P}}(p, p^{-i}) - p^i \tag{5}$$

Hence, we can characterize the long-run behavior of the learning game by characterizing the long-run behavior of this dynamical system.

A long-run *equilibrium* of the game satisfies $p^i = \bar{b}_{\mathscr{P}}(p^i, p^{-i})$ for $i = 1, 2$. Under our main assumption given below, for each pair type $\mathscr{P}$, the forecasts converge to a long-run equilibrium. This is shown in Proposition 1 in the Appendix.

Here is our main assumption:

Assumption 1. *For* $\mathscr{P} = \mathrm{LH}, \mathrm{HH}$, $p^i - \bar{b}_{\mathscr{P}}(p^i, p^{-i})$ *is strictly increasing in* p^i *and* $\bar{b}_{\mathscr{P}}(p^i, p^{-i})$ *is strictly increasing in* p^i*.*

When a player's forecast increases, there are two effects. First, both b_N and b_I increase which raises the probability of playing *Stag* in each state. Second, the distribution of states may change. In particular, an increase in p^1 will make player 1 more likely to play *Stag* in each state, which would tend to decrease the probability of player 2 being in state I. This will raise player 2's average probability of playing *Stag*, which will decreasing the steady state probability that player 1 is in state I, and further increase the probability that player 1 plays *Stag*. The first part of Assumption 1 says that this channel cannot be too strong.

Also, when p^1 goes up, player 1 is more exposed to the outcome (*Stag*,*Hare*). In principle, even though player 1 plays *Stag* more in every state, the frequency of state I could increase so much that the *average* probability of playing *Stag* decreases. We rule out this case as well.[20]

[20]The second part of the Assumption 1 in fact goes against our results, since we are limiting the potential

In the Appendix, we prove that Assumption 1 is satisfied for model LH, and for model HH it is satisfied in all of our numerical examples. Under this assumption, for each p^{-i}, there will exist a unique p^i that is "consistent" with p^{-i}, in the sense that $p^i = \bar{b}(p^{-i}, p^i)$. The consistent forecast is denoted $p^C_{\mathscr{P}}(\cdot)$. Since forecasts are consistent in equilibrium, equilibrium forecasts are characterized by the conditions $p^C_{\mathscr{P}}(p^1) = p^2$ and $p^C_{\mathscr{P}}(p^2) = p^1$.

5.3 Comparison of models

We are now ready to compare the outcomes of pairs LL, LH, and HH. We show in the Appendix that (a) for all pair types $p^C_{\mathscr{P}}(p)$ is strictly increasing in p and for pair types LH and HH, $p^C_{\mathscr{P}}$ is strictly decreasing in β and (b) $p^C_{LL} > p^C_{LH} > p^C_{HH}$. The interpretation of (a) is that as a player's forecast increases, the consistent forecast of his partner goes up as well, and as β goes up, the consistent forecast goes down. (b) says that consistent forecasts are higher for LL than for LH, which are higher than those of HH.

Since these functions are strictly increasing, they have inverses. Equilibrium forecasts for pair $\mathscr{P}$ lies at the intersection of $p^C_{\mathscr{P}}$ and $(p^C_{\mathscr{P}})^{-1}$. In Figure 5, we give examples of the three curves together for $\beta = 0.2$.

As the models with pairs LL and HH are symmetric, the forecasts in equilibrium will be the same. We denote by $\widehat{p}_{LL}$ and $\widehat{p}_{HH}$ the highest equilibrium forecasts for LL and HH, respectively. Since model LH is asymmetric, the forecasts will not be the same. However, since $p^2_{LH} = b_N(p^1_{LH})$, and the latter is strictly increasing, it still makes sense to define the highest forecasts in an equilibrium for LH, which we denote $(\widehat{p}^1_{LH}, \widehat{p}^2_{LH})$. Since these forecasts are consistent, they also describe the highest average probability of playing *Stag* for each pair. We also denote by $\underline{p}_{LL}$ the forecast in the *lowest* equilibrium of model LL.

Proposition 2 in the Appendix gives a characterization of the solution of the model. The key points are summarized in the following results:

for the I state to reduce the probability of playing *Stag*. We suspect that our results would hold even if this assumption is relaxed, though it simplifies the analysis greatly.

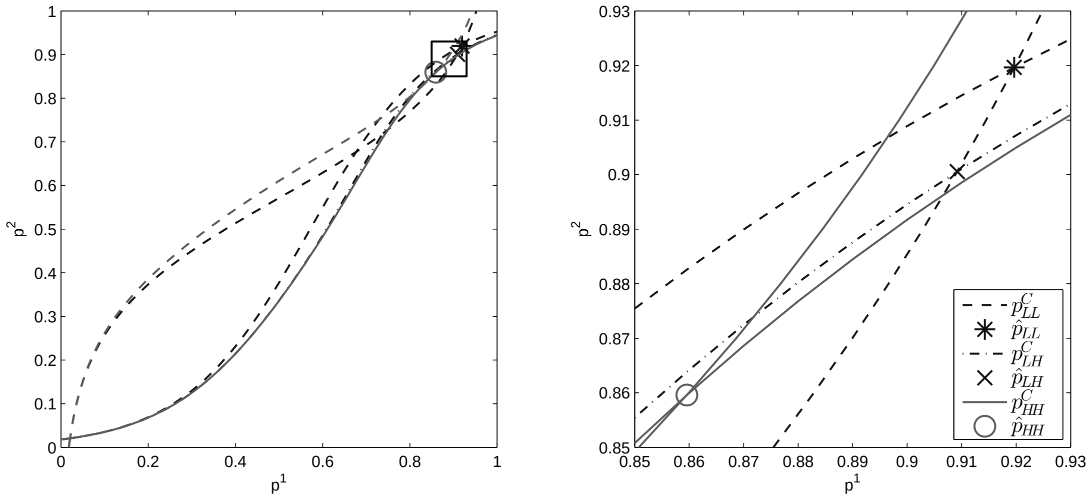


Figure 5: Comparison of curves for $\beta = 0.2$. The dashed lines are $b_n = p^C_{LL}$, the dot-dashed line is p^C_{LH}p and the solid lines are p^C_{HH}. The best equilibrium for LL is $*$, the best for LH is $\times$, and the best for HH is $\circ$. Note that small changes in β can produce dramatic changes in the best equilibrium; a small increase in β will mean that p^C_{HH} and $(p^C_{HH})^{-1}$ have no intersection with high forecasts.

Result 4. *If $\beta > 0$, the highest probability of playing* Stag *in* LL *is higher than that in* LH*, which is higher than that in* HH*, i.e.* $\widehat{p}_{\text{LL}} > \widehat{p}^2_{\text{LH}} > \widehat{p}^1_{\text{LH}} > \widehat{p}_{HH}$.

Thus, we can order the outcomes in the sense that the long-run probability of playing *Stag* in *HH* cannot be higher than the probability in *LH*, which cannot be higher than the probability in *LL*.

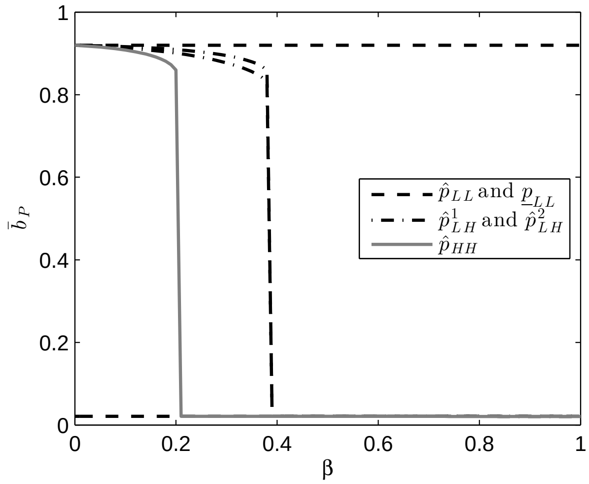
Result 5.

(1) For $\beta > 0$, the highest probability of playing Stag *is strictly decreasing in β for pair types* LH *and* HH.

(2) There exist cutoffs $0 < \beta_{\text{HH}} < \beta_{\text{LH}} < 1$ such that

- $\widehat{p}_{\text{HH}} \geq \underline{p}_{\text{LL}}$ *when* $\beta \leq \beta_{\text{HH}}$ *and* $\widehat{p}_{\text{HH}} < \underline{p}_{\text{LL}}$ *when* $\beta > \beta_{\text{HH}}$
- $\widehat{p}^i_{\text{LH}} \geq \underline{p}_{\text{LL}}$ *when* $\beta \leq \beta_{\text{LH}}$ *and* $\widehat{p}^i_{\text{LH}} < \underline{p}_{\text{LL}}$ *when* $\beta > \beta_{\text{LH}}$.

Figure 6: The best and worst equilibria of LL are the dashed lines. The best equilibrium of LH are the dot-dashed lines ($\widehat{p}^1_{LH} < \widehat{p}^2_{LH}$) and the best equilibrium of HH is the solid line. We see that for small β, all models have equilibria in which *Stag* is played frequently.



This result says that if β is small, all three models will have an equilibrium in which *Stag* is played with relatively high probability. For example, for $\beta = 0.2$, $(\widehat{p}^1_{LH}, \widehat{p}^2_{LH}) \approx (0.91, 0.90)$ and $\widehat{p}_{HH} \approx 0.86$. This is in comparison to LL, for which $\widehat{p}_{LL} \approx 0.92$.

However, as one can see from Figure 6, as β increases, eventually a critical β is reached at which the best outcome drops dramatically to be worse than the *inefficient* equilibrium of pair type LL, for which $p \approx 0.02$. The break occurs for model HH first, at the critical value $\beta_{HH} \approx 0.204$, and then for model LH at $\beta_{LH} \approx 0.383$. For β above these cutoffs, the equilibrium probability of *Stag* is very low.

Thus, the model shows that Results 1 and 2 of the experiment can both be explained by state-dependent preferences of the H players. Such preferences would cause lower probabilities of *Stag* after receiving the loser's payoff, which in turn affects the long-run probabilities of *Stag*. Since L have normal preferences, there is no reason why they cannot have high forecasts and hence high probabilities of *Stag* in the long-run. If high-caste players have a high β, then they will not be able to sustain a high probability of *Stag*. The fact that some pairs are able to coordinate on an efficient convention while others are not is consistent with variation in β among high-caste players.

6 Discussion

We have argued that the outcome of the Stag Hunt experiment can be explained by state-dependent preferences of high-caste men that are shaped by the culture of honor. It is only after a particular *strategic* outcome—the loser's payoff—that high-caste men become concerned about the distribution of payoffs. An alternative explanation of the lower rate of efficient coordination among H players is a caste difference in how much players "trust" their partner to take the efficient action. Indeed, some of the results in Table 2 are consistent with the hypothesis that H are less trusting than L: In period 1, before there is any history of play to influence decisions, 57% of L players chose *Stag*, as compared to only 37% of H.

A canonical game for assessing trust is the investment game of Berg, Dickhaut and McCabe (1995) (see Algan and Cahuc (2013) for a review of the trust literature). Two players are anonymously paired, and each player is given an initial endowment of money. One of the players (the principal) chooses a portion of his endowment to send to the the other player (the agent), and the principal is told that the experimenter will multiply the investment by three. For example, if the principal invests ₹1, then the agent receives ₹3. Finally, the agent chooses a non-negative amount of money to return to the principal. A self-interested agent will always seize the investment, so the unique Nash equilibrium calls for the principal to invest nothing. We interpret a positive investment as an indication that the principal trusts the agent to pay a return.

We implemented a binary choice version of this game in the district of Unnao, Uttar Pradesh in 2007. Unnao is approximately 40 miles from the district in which we conducted the Stag Hunt experiment. The players in a given pairing were drawn from different villages within the district. Each player was given an endowment of ₹50, which was equivalent to the daily wage for an unskilled worker. The principal had to choose between investing all or none of the ₹50. If the principal chose to invest, the agent received ₹150 and chose whether or not to return ₹100.

There were four treatments that differed in the caste statuses of the two players. In

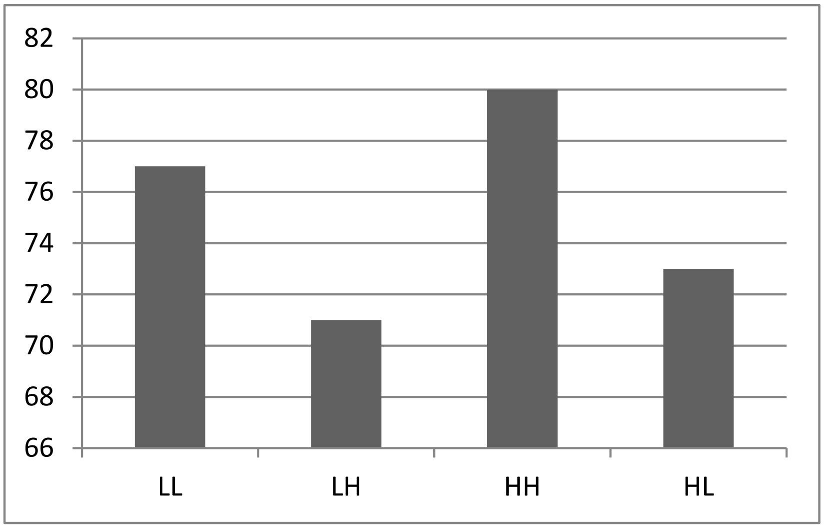


Figure 7: The proportion of principals who chose to invest.

treatment *LL*, both players were *L*, and in treatment *HH*, both players were *H*. In treatment *LH*, the principal was *L* and the agent was *H*, and finally in treatment *HL*, the principal was *H* and the agent was *L*. Caste statuses were indicated to the players by the use of names that were known to be exclusive to particular subcastes, as we verified in a pre-experiment test in the district. The numbers of pairs by treatment were: 26 (*LL*), 34 (*LH*), 30 (*HH*), and 30 (*HL*).

Figure 7 presents the proportion of principals who invested. There are no significant differences in this proportion across different treatments. In fact, trust as measured by the proportion of principals who invest is higher for *HH* (80%) than for the other treatments.

We also asked the principals directly: Do you believe that the agent will return ₹100 if you invest? Figure 8 reports the proportion that said yes, and distinguishes between the principals who invested and those who did not. Among principals who invested, there are again no significant differences in beliefs by the caste status of the principal or the agent. Finally, we asked the agents: Do you believe the principal will invest the ₹50? In a dprobit regression (not reported here), all coefficients for the treatments are close to zero and insignificant; the results are robust to controls for the agents' wealth.

In sum, the investment game experiment provides no evidence to support the hypothesis

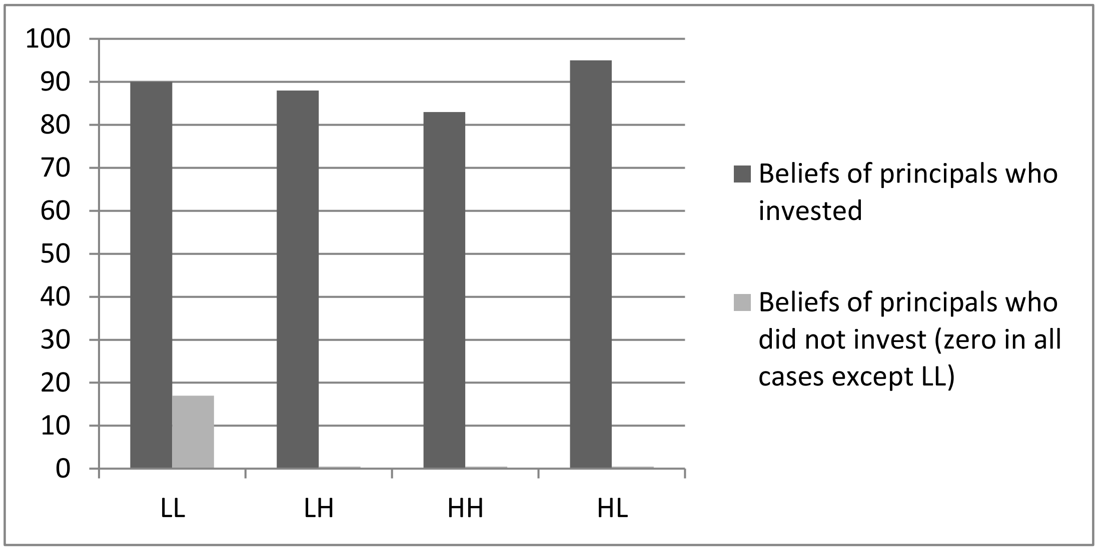


Figure 8: The proportion of principals who believed that the agent would return on the investment.

H are less trusting than L. As such, trust is not an explanation of the caste differences in behavior in the Stag Hunt experiment.

7 Conclusion

The broad question that this paper addresses is how conventions emerge. Much of what we value in society depends on coordination: for example, language, fiat money, standardization, politeness, and the rule of law. Development itself has been viewed as a process of coordination. To shed light on this process, we ran an experiment in India with low-caste and high-caste men. The low-caste pairs generally coordinated on a convention under which their endowment grew by 70 percent (₹6 to ₹10), whereas the high-caste pairs showed evidence of adopting a convention under which their endowment grew by only 17 percent(₹6 to ₹7). To explain this caste gap, we found evidence for a novel obstacle to efficient coordination: Because they are from a culture of honor, the high-caste individuals may interpret the loser's payoff and, more generally, miscoordination, as an insult, to which they respond by not trying as hard for efficient coordination in the future. In simplest terms, a mindset associated with the culture of honor is “cross me, and I'll punish you.” This mindset means that an accidental miscoordination can lead to more miscoordination and misunderstanding and an unraveling of cooperative behavior. The limited ability to coordinate of the high castes, which are the dominant castes in the region, is consistent with the observed failures of coordination and “institutional inertia” in the state of Uttar Pradesh.

Events and actions do not always speak for themselves but instead depend on framing for their meaning. In such cases, social behavior depends on subjective interpretations. Experiments that expose individuals to the same situation under alternative frames have demonstrated that frame shifts can induce large shifts in economic behavior. These findings make it plausible that a difference in mental frameworks, associated with a difference in culture, is also capable of producing a large difference in economic behavior.

We considered and found little support for two alternative explanations for the observed caste difference in coordination. First, there is a statistically insignificant difference in trust in the initial period of the game. To assess trust directly, we report results of an investment game experiment in a nearby district in Uttar Pradesh. We find no difference in trust between high- and low-caste men in the investment game. Second, all of our results are robust to controls for education and wealth, which suggests that non-cultural correlates of caste cannot explain the caste gap in coordination.

Our evidence that mental frameworks may powerfully impede coordination has policy implications. There is clear evidence from interventions that target individuals that interpretative frameworks (or “theories of mind”) are malleable.[21] There is also recent evidence that interventions that target whole communities can change mental frameworks across the communities. For instance, Beaman et al. (2009) find that exposure to women village lead-

[21] Algan et al. (2012) evaluate a randomized control trial in which disruptive young boys in the treated group were regularly placed in small groups with pro-social children. The intervention helped the subjects learn trust, empathy, and self-control. At age 26, the treated subjects, compared to the control group, were much more likely to be employed and much less likely to have a criminal record. Experimental studies also show that people who were taught to reframe, in a self-distancing perspective, a negative experience that had made them feel rejected or enraged, were less likely to experience high distress and less likely to reciprocate hostile behavior (Ayduk and Kross 2008 2010, Kross and Ayduk 2011).

ers through a policy of reservations for women in village elections in India improved men's ability to judge fairly the quality of female leaders. That is, exposure changed perceptions. Blattman, Hartman and Blair (2012) find that an advocacy and training program in mediation gave individuals new conceptual tools with which to understand conflicts—for instance, seeing the interaction as positive—sum instead of zero-sumwhich increased the resolution of land disputes. The frameworks within which people view the world are an under-studied aspect of economic behavior, and they are not set in stone.

A Proofs

In this section, we provide a detailed analysis of models LH and HH.

First we make some preliminary observations about b. Note that $u(Stag|p, I) = 10(1 - \beta)p + (3 - 7\beta)(1 - p)$ and $u(Hare|p, I) = (7 - 3\beta)p + 7(1 - \beta)(1 - p)$. Hence, $u(Stag|p, I) - u(Hare|p, I) = 7(1 - \beta)p - 4$ and so $b_I(p) = (1 + \exp(4 - 7(1 - \beta)p))^{-1}$. We note for future reference that $b_I(p)$ is decreasing in β, so that $\Delta(p) = b_N(p) - b_I(p)$ is increasing in β.

We begin by solving explicitly for $\Pi_{\mathscr{P}}$ and $\bar{b}_{\mathscr{P}}$ for models LH and HH. Henceforward, we drop the (p^2, p^1) from $\Pi_{\mathscr{P}}$ for economy of notational.

Lemma 1. Π_{LH} *and* $\bar{b}_{\mathrm{LH}}$ *are given by*

$$\Pi_{\mathrm{LH}}(\mathrm{N}, \mathrm{I}) = \frac{b_{\mathrm{N}}(p^2)(1 - b_{\mathrm{N}}(p^1))}{1 + \Delta(p^2)(1 - b_{\mathrm{N}}(p^1))} \tag{6a}$$

$$\bar{b}_{\mathrm{LH}}(p^2, p^1) = \frac{b_{\mathrm{N}}(p^2)}{1 + \Delta(p^2)(1 - b_{\mathrm{N}}(p^1))} \tag{6b}$$

Π_{HH} *and* $\bar{b}_{\mathrm{HH}}$ *are given by*

$$\Pi_{\mathrm{HH}}(\mathrm{N},\mathrm{I}) = \frac{b_{\mathrm{N}}(p^2)\left[1 - b_{\mathrm{N}}(p^1) + \Delta(p^1)(1 - b_{\mathrm{N}}(p^2))\right]}{(1+\Delta(p^2))(1+\Delta(p^1)) - b_{\mathrm{N}}(p^1)\Delta(p^2)(1+\Delta(p^1)) - \Delta(p^1)b_{\mathrm{N}}(p^2)(1+\Delta(p^2))} \tag{7a}$$

$$\bar{b}_{\mathrm{HH}}(p^2,p^1) = b_{\mathrm{N}}(p^2)\left[\frac{1+\Delta(p^1)\left[1 - b_{\mathrm{N}}(p^2) - b_{\mathrm{N}}(p^1)\Delta(p^2)\right]}{(1+\Delta(p^2))(1+\Delta(p^1)) - b_{\mathrm{N}}(p^1)\Delta(p^2)(1+\Delta(p^1)) - \Delta(p^1)b_{\mathrm{N}}(p^2)(1+\Delta(p^2))}\right] \tag{7b}$$

Proof. We explicitly calculate Π_{LH} from the identity:

$$\begin{aligned}
\Pi_{LH}(N,I) &= \big(b_N(p^2) + (b_I(p^2) - b_N(p^2))\Pi_{LH}(N,I)\big)\,(1 - b_N(p^1)) \\
\implies \Pi_{LH}(N,I) &= \frac{b_N(p^2)(1 - b_N(p^1))}{1 + \Delta(p^2)(1 - b_N(p^1))}
\end{aligned}$$

Hence, $\bar{b}_{LH}(p^2,p^1)$ is

$$\begin{aligned}
\bar{b}_{LH}(p^2,p^1) &= b_N(p^2)\left(1 - \frac{\Delta(p^2)(1 - b_N(p^1))}{1 + \Delta(p^2)(1 - b_N(p^1))}\right) \\
&= \frac{b_N(p^2)}{1 + \Delta(p^2)(1 - b_N(p^1))}
\end{aligned}$$

For model HH, the derivation is longer but not hard. We note that Π_{HH} can be derived from the system

$$\begin{aligned}
\Pi_{HH}(I,N) &= \Pi(I,N)b_I(p^1)(1 - b_N(p^2)) + \Pi_{HH}(N,I)b_N(p^1)(1 - b_I(p^2)) \\
&\quad + (1 - \Pi_{HH}(I,N) - \Pi_{HH}(N,I))b_N(p^1)(1 - b_I(p^2)) \\
\Pi_{HH}(N,I) &= \Pi_{HH}(I,N)(1 - b_I(p^1))b_N(p^2) + \Pi_{HH}(N,I)(1 - b_N(p^1))b_I(p^2) \\
&\quad + (1 - \Pi_{HH}(I,N) - \Pi_{HH}(N,I))(1 - b_N(p^1))b_N(p^2)
\end{aligned}$$

Details are left to the reader. □

For model LH, it is straightforward to see that Assumption 1 is satisfied. Note that

$$p^1 - \bar{b}_{LH}(p^2, p^1) = p^1 - \frac{b_N(p^2)}{1 + \Delta(p^2)(1 - b_N(p^1))}$$

For $p^1 > 0$, this expression has the same sign as $1 + \Delta(p^2)(1 - b_N(p^1)) - \frac{b_N(p^2)}{p^1}$, and in particular, they are zero at the same values. The derivative of this expression w.r.t p^1 is $-\Delta(p^2)b'(p^1, N) + \frac{b_N(p^2)}{(p^1)^2}$. This expression is positive iff $(p^1)^2 b'(p^1, N) < \frac{b_N(p^2)}{\Delta(p^2)}$, where the RHS is larger than 1. In fact, it is always true for the utilities we chose, as $(p^1)^2 b'(p^1, N) < 1$.[22]

We can also show that $\bar{b}_{LH}$ is strictly increasing in p^2. Differentiating (6b) with respect to p^2, the numerator will be

$$b'_N(p^2) + (1 - b_N(p^1)) \left[\Delta(p^2) b'_N(p^2) + \Delta'(p^2) b_N(p^2)\right]$$

We want to show that this expression is positive. Clearly if the term in brackets is positive, the whole expression is positive for all p^1. If the bracketed term is negative, then the expression is bounded below by $b'_N(p^2) + [\Delta(p^2) b'_N(p^2) + \Delta'(p^2) b_N(p^2)]$, so we show this bound is positive. Expanding $\Delta(p^2)$ and simplifying, the expression becomes

$$b'_N(p^2)[1 - b_I(p^2)] + b_N(p^2) b'_I(p^2)$$

Clearly the second term is positive, and $b_I < 1$, so we are done.

The algebra required to give a similar argument for model HH would be quite cumbersome and unedifying. We simply note that for all of the examples we report, Assumption 1 is satisfied. Having made explicit our sufficient conditions, we proceed with the rest of the analysis.

We will find many uses for Assumption 1, but one use is that it rules out cycles in the

[22] Consider $\log(p^1 b'(p, N)) = 2\log(p) + \log(7) + 4 - 7p - 2\log(1 + \exp(4 - 7p))$. The derivative of this expression is $\frac{2}{p} - 7 + 14\frac{exp(4-7p)}{1+exp(4-7p)}$. Clearly this is positive for p close to 0 and for $p = 1$ is approximately $-4.336 < 0$. Moreover, both terms involving p are strictly decreasing. Hence, there is a unique critical point at which the function is maximized, which is about 0.6961. At this critical value, $p^1 b'(p, N) \approx 0.7050$.

mean dynamical system.

Proposition 1. *Play converges to an equilibrium.*[23]

Proof. Kushner and Yin (2003) Chapter 8 Theorem 4.3 shows that the forecasts converge to a trajectory of the dynamical system (5). By Assumption 1, $\frac{\partial}{\partial p_i}\left(\bar{b}(p^i, p^{-i}) - p^i\right) < 0$. Hence, the divergence of the vector field for $(\dot{p}^1, \dot{p}^2)$ is negative, and the vector field is area decreasing. Therefore, no non-empty open sets are invariant under the mean dynamical system, and play converges to an equilibrium (cf. Benaïm and Hirsch 1999, Fudenberg and Levine 1998). □

Lemma 2. *For* $\mathscr{P} = \text{LH}, \text{HH}$,

(1) There exists a unique solution to the functional equation $p^C_{\mathscr{P}}(p^2) = \bar{b}_{\mathscr{P}}(p^2, p^C_{\mathscr{P}}(p^2))$. *For* $p^1 < p^C_{\mathscr{P}}(p^2)$, $p^1 < \bar{b}_{\mathscr{P}}(p^2, p^1)$, *and for* $p^1 > p^C_{\mathscr{P}}(p^2)$, $p^1 > \bar{b}_{\mathscr{P}}(p^2, p^1)$.

(2) $p^C_{\mathscr{P}}$ *is strictly increasing in* p^2.

Proof. $p^1 - \bar{b}_{\mathscr{P}}(p^2, p^1)$ is negative for $p^1 = 0$ and positive for $p^1 = 1$, and $\bar{b}_{\mathscr{P}}$ is smooth, so there exists a p^1 at which it is zero. By Assumption 1, the difference is strictly increasing so there can be at most one point at which it is zero. For $p^1 < p^C_{\mathscr{P}}(p^2)$, $p^1 - \bar{b}_{\mathscr{P}}(p^2, p^1) < 0$, and for $p^1 > p^C_{\mathscr{P}}(p^2)$, the converse holds.

By Assumption 1, $p^C_{\mathscr{P}}(p^2) = \bar{b}_{\mathscr{P}}(p^2, p^C_{\mathscr{P}}(p^2)) < \bar{b}_{\mathscr{P}}(p, p^C_{\mathscr{P}}(p^2))$ for $p > p^2$. Part (1) implies that $p^C_{\mathscr{P}}(p^2) < p^C_{\mathscr{P}}(p)$. □

Lemma 3. $\bar{b}_{\text{LH}}$, $\bar{b}_{\text{HH}}$, p^C_{HH} *and* p^C_{LH} *are all strictly decreasing in* β.

Proof. Since $\Delta(p) = b(p, N) - b(p, I)$ is increasing in β, $\bar{b}_{LH}$ is decreasing in β.

To see that $\bar{b}_{HH}$ is decreasing in β, observe that $\bar{b}_{HH}$ can be rewritten

$$\bar{b}_{HH} = b_N(p^2)\left[\frac{\overbrace{1 + \Delta(p^1)[1 - b_N(p^2) - b_N(p^1)\Delta(p^2)]}^{A}}{\underbrace{1 + \Delta(p^1)[1 - b_N(p^2) - b_N(p^1)\Delta(p^2)]}_{A} + \underbrace{\Delta(p^2)[1 - b_N(p^1) + \Delta(p^1)(1 - b_N(p^2))]}_{B}}\right]$$

[23] We use the term equilibrium for what has been called a Nash distribution equilibrium elsewhere in the literature.

Note that both A and B are positive, since $\bar{b}_{HH} < b_N(p^2)$. $\bar{b}_{HH}$ is decreasing in β iff $\frac{A}{A+B}$ is decreasing in β. Note that the only effect of β is through $\Delta(p^2)$ and $\Delta(p^1)$, both of which are increasing in β (as b_I is decreasing). Furthermore, A is decreasing in $\Delta(p^2)$ and B is increasing in $\Delta(p^2)$, so it will be sufficient for us to show that $\frac{A}{A+B}$ is decreasing in $\Delta(p^1)$. $\frac{A}{A+B}$ will be decreasing iff $(A+B)dA - A(dA+dB) = BdA - AdB$ is negative.

To see this, note that

$$\frac{dA}{d\Delta(p^1)} = 1 - b_N(p^2) - b_N(p^1)\Delta(p^2)$$
$$\frac{dB}{d\Delta(p^1)} = \Delta(p^2)(1 - b_N(p^2))$$

so

$$BdA - AdB = -\Delta(p^2)b_N(p^1)B + \Delta(p^2)(1 - b_N(p^2))(\Delta(p^2)\Delta(p^1) - 1)b_N(p^1)$$

Clearly $\Delta(p^2)\Delta(p^1) < 1$, so both terms are negative.

We now show that $p^C_{\mathscr{P}}$ is decreasing in β for both models. For this argument, we make β explicit. Suppose that $\beta < \beta'$. Then $p^C_{\mathscr{P}}(p^2, \beta) = \bar{b}_{\mathscr{P}}(p^2, p^C_{\mathscr{P}}(p^2, \beta), \beta) > \bar{b}_{\mathscr{P}}(p^2, p^C_{\mathscr{P}}(p^2, \beta), \beta')$ since $\bar{b}_{\mathscr{P}}$ is decreasing in β. This implies that $p^C_{\mathscr{P}}(p^2, \beta) > p^C_{\mathscr{P}}(p^2, \beta')$ by Lemma 3. □

Lemma 4. *For all $\beta \in (0,1)$, $\bar{b}_{\mathrm{LH}} > \bar{b}_{\mathrm{HH}}$ and $p^C_{\mathrm{LH}} > p^C_{\mathrm{HH}}$.*

Proof. $\bar{b}_{\mathscr{P}}$ can be rewritten

$$\bar{b}_{LH}(p^2, p^1) = \frac{b_N(p^2)}{1 + \Delta(p^2)(1 - b_N(p^1))} = \frac{b_N(p^2)}{C}$$
$$\bar{b}_{HH}(p^2, p^1) = b_N(p^2)\left[\frac{1 + \Delta(p^1)\left[1 + \Delta(p^2)(1 - b_N(p^1)) - b_N(p^2) - \Delta(p^2)\right]}{(1 + \Delta(p^1))\left[1 + \Delta(p^2)(1 - b_N(p^1))\right] - \Delta(p^1)b_N(p^2)(1 + \Delta(p^2))}\right]$$
$$= b_N(p^2)\left[\frac{1 + \Delta(p^1)\left[C - b_N(p^2) - \Delta(p^2)\right]}{C\left[1 + \Delta(p^1)(1 - b_N(p^2))\right] - \Delta(p^1)b_N(p^2)b_N(p^1)\Delta(p^2)}\right]$$

where $C = 1 + \Delta(p^2)(1 - b_N(p^1))$. Hence, $\bar{b}_{LH} > \bar{b}_{HH}$ iff

$$C\left[1 + \Delta(p^1) - \Delta(p^1)b_N(p^2)\right] - b_N(p^2)b_N(p^1)\Delta(p^1)\Delta(p^2) > C + C\Delta(p^1)\left[C - b_N(p^2) - \Delta(p^2)\right]$$
$$\iff C\Delta(p^1)\left[1 + \Delta(p^2) - C\right] > b_N(p^2)b_N(p^1)\Delta(p^1)\Delta(p^2)$$
$$\iff C\Delta(p^1)\Delta(p^2)b_N(p^1) > b_N(p^2)b_N(p^1)\Delta(p^1)\Delta(p^2)$$

which is always true, since $C > 1 > b_N(p^2)$.

As in the previous lemma, this implies that $p^C_{LH} > p^C_{HH}$, since

$$p^C_{HH}(p^2) = \bar{b}_{HH}(p^2, p^C_{HH}(p^2)) < \bar{b}_{LH}(p^2, p^C_{HH}(p^2))$$

which implies that $p^C_{HH}(p^2) < p^C_{LH}(p^2)$. □

Finally, we need the following technical result to characterize the equilibria:

Lemma 5. *Let $h^1, h^2, l^1, l^2 : [0, 1] \to (0, 1)$ be continuous and strictly increasing, with $h^1 \geq l^1$ and $h^2 > l^2$. Let x_h denote the largest solution to $h^1(x_h) = (h^2)^{-1}(x_h)$ and x_l any solution to $l^1(x_l) = (l^2)^{-1}(x_l)$. Then $x_l < x_h$.*

Proof. Let $\bar{x} = (h^2)^{-1}(1)$. Since $h^2 > l^2$, $(l^2)^{-1} > (h^2)^{-1}$, as $h^2(x) = l^2(x') = y$ implies that $x < x'$.

For $x > x_h$, it must be that $h^1(x) < (h^2)^{-1}(x)$. For $h^1(\bar{x}) < (h^2)^{-1}(\bar{x}) = 1$, so if $h^1(x) > (h^2)^{-1}(x)$, then by continuity there is an $x' \in (x, \bar{x})$ such that $h^1(x') = (h^2)^{-1}(x')$, and $x' > x_h$, which contradicts that x_h is the largest solution. Now, if $x_h \leq x_l$, then $(h^2)^{-1}(x_l) < (l^2)^{-1}(x_l) = l^1(x_l) < h^1(x_l)$, a contradiction. □

For models LL and HH, equilibria are symmetric, for if $p^1 < p^C_{HH}(p^1)$, then $p^1 < p^C_{HH}(p^C_{HH}(p^1))$, which contradicts the consistency of an equilibrium. The same argument applies if $p^1 > p^C_{HH}$, and if HH is replaced with LL. Thus, it makes sense to define $\widehat{p}_{HH}$ and $\widehat{p}_{LL}$ to be the highest forecasts in an equilibrium of models HH and LL, respectively.

Let $(\widehat{p}^2_{LH}, \widehat{p}^1_{LH})$ denote the highest equilibrium of model LH. This makes sense because the monotonicity of $p^C_{\mathscr{P}}$ ensures that one coordinate is maximized iff the other is maximized. Let $\underline{p}_{LL}$ denote the lowest equilibrium of model LL.

Here is the main result:

Proposition 2.

(1) $\widehat{p}_{\text{LL}} > \widehat{p}^1_{\text{LH}} > \widehat{p}^2_{\text{LH}} > \widehat{p}_{\text{HH}}$.

(2) $\widehat{p}_{\text{LH}}$ *and* $\widehat{p}_{\text{HH}}$ *are strictly decreasing in* β.

(3) There exist cutoffs $\beta_{\text{LH}}, \beta_{\text{HH}} \in (0,1)$ *such that for* $\beta > \beta_{\mathscr{P}}$, $\widehat{p}_{\mathscr{P}} > \underline{p}_{\text{LL}}$ *and* $\widehat{p}_{\mathscr{P}} < \widehat{p}_{\text{LL}}$, *and for* $\beta < \beta_{\mathscr{P}}$, $\widehat{p}_{\mathscr{P}} < \underline{p}_{\text{LL}}$. *Moreover,* $\beta_{\text{LH}} < \beta_{\text{HH}}$.

Proof. (1) is a straightforward consequence of the fact that $b_N > \bar{b}_{HH}$ and $\bar{b}_{LH} > \bar{b}_{HH}$, and the technical lemma. Note that $\widehat{p}_{LH}$ satisfies $\widehat{p}^1_{LH} = \bar{b}_{LH}(\widehat{p}^2_{LH}) < b_N(\widehat{p}^2_{LH}) = b_N(b_N(\widehat{p}^1_{LH}))$. Thus, $\widehat{p}^1_{LH} < b_N(\widehat{p}^1_{LH}) = \widehat{p}^2_{LH}$.

Similarly, property (2) is implied by the technical lemma and the fact that $\bar{b}_{LH}(\cdot,\beta) > \bar{b}_{LH}(\cdot,\beta')$ and $\bar{b}_{HH}(\cdot,\beta) > \bar{b}_{HH}(\cdot,\beta')$ for $\beta' > \beta$.

We now show property (3). For $\beta = 1$, the curve $p^C_{LH}(p) < b_N^{-1}(p)$ for $p \geq \underline{p}_{LL}$. Take $\beta_{LH} = \inf\{\beta | p^C_{LH}(p) \geq b_N^{-1}(p) \text{ for some } p \geq p_{LL}\}$. We know the inf is non-negative, since for $\beta = 0$, models LH and HH are equivalent to model LL. Also, since the gap at $\beta = 1$ is strict, and p^C_{LH} is continuous (by the Implicit Function Theorem, as $\bar{b}_{LH}$ is continuously differentiable) and strictly decreasing in β, it must be that $\beta_{LH} < 1$.

Similarly, for $\beta = 1$, $p^C_{HH} < p^C_{LL} < b_N^{-1} < (p^C_{HH})^{-1}$ for $p > \underline{p}_{LL}$. Take $\beta_{HH} = \inf\{\beta | p^C_{HH}(p) \geq (p^C_{HH})^{-1}(p) \text{ for some } p \geq \underline{p}_{LL}\}$. Clearly, $\beta_{HH} < \beta_{LH}$, since at β_{LH}, we would have $p^C_{HH} < p^C_{LH} \leq b_N^{-1}(p) < (p^C_{HH})^{-1}$ for all $p \geq \underline{p}_{LL}$.

For $\beta > \beta_{\mathscr{P}}$, model $\mathscr{P}$ does not have an equilibrium above $\underline{p}_{LL}$. But some equilibrium exists, so it lies below p_{LL}. □

References

Alesina, Alberto F, Paola Giuliano, and Nathan Nunn. 2011. "On the origins of gender roles: Women and the plough." National Bureau of Economic Research.

Algan, Y, and P Cahuc. 2013. "Trust and Human Development: Overview and Policy." In *Handbook of Economic Growth*. Vol. 1, , ed. Philippe Aghion and Steven Durlauf.

Algan, Yann, and Pierre Cahuc. 2009. "Civic virtue and labor market institutions." *American Economic Journal: Macroeconomics*, 111–145.

Algan, Yann, Elizabeth Beasley, Richard E Tremblay, and Franck Vittaro. 2012. "The Long-Term Impact of Social Skills Training at School Entry: A randomized controlled trial." *Sciences Po Working Paper*.

Andreoni, James. 1995. "Warm-glow versus cold-prickle: the effects of positive and negative framing on cooperation in experiments." *The Quarterly Journal of Economics*, 110(1): 1–21.

Ariely, Dan. 2009. *Predictably irrational: The hidden forces that shape our decisions.* New York:Harper Collins.

Ayduk, Özlem, and Ethan Kross. 2008. "Enhancing the pace of recovery self-distanced analysis of negative experiences reduces blood pressure reactivity." *Psychological Science*, 19(3): 229–231.

Ayduk, Özlem, and Ethan Kross. 2010. "From a distance: Implications of spontaneous self-distancing for adaptive self-reflection." *Journal of personality and social psychology*, 98(5): 809.

Beaman, Lori, Raghabendra Chattopadhyay, Esther Duflo, Rohini Pande, and Petia Topalova. 2009. "Powerful women: does exposure reduce bias?" *The Quarterly Journal of Economics*, 124(4): 1497–1540.

Benaïm, Michel, and Morris W Hirsch. 1999. "Mixed equilibria and dynamical systems arising from fictitious play in perturbed games." *Games and Economic Behavior*, 29(1): 36–72.

Berger, Peter, Brigette Berger, and Hansfried Kellner. 1975. *The Homeless Mind: Modernization and Consciousness.* New York:Random House.

Berg, Joyce, John Dickhaut, and Kevin McCabe. 1995. "Trust, reciprocity, and social history." *Games and economic behavior*, 10(1): 122–142.

Bernhard, Helen, Ernst Fehr, and Urs Fischbacher. 2006. "Group affiliation and altruistic norm enforcement." *The American Economic Review*, 96(2): 217–221.

Bewley, Truman F. 1999. *Why wages don't fall during a recession.* Cambridge:Harvard University Press.

Blattman, Christopher, Alexandra Hartman, and Robert Blair. 2012. "Building institutions at the micro-level: Results from a field experiment in property dispute and conflict resolution." *Available at SSRN 2158966*.

Camerer, Colin. 2003. *Behavioral game theory: Experiments in strategic interaction.* Princeton:Princeton University Press.

Cassan, Guilhem. 2010. "British law and caste identity manipulation in colonial India: the Punjab Alienation of Land Act." Paris School of Economics Working paper.

Charness, Gary, and Matthew Rabin. 2002. "Understanding social preferences with simple tests." *The Quarterly Journal of Economics*, 117(3): 817–869.

Charness, Gary, and Uri Gneezy. 2008. "What's in a name? Anonymity and social distance in dictator and ultimatum games." *Journal of Economic Behavior & Organization*, 68(1): 29–35.

Chen, Roy, and Yan Chen. 2011. "The potential of social identity for equilibrium selection." *The American Economic Review*, 101(6): 2562–2589.

Chen, Yan, and Sherry Xin Li. 2009. "Group identity and social preferences." *The American Economic Review*, 99(1): 431–457.

Cohen, Dov, and Joe Vandello. 1998. "Meanings of violence." *The Journal of Legal Studies*, 27(S2): 567–584.

Cohen, Dov, Joseph Vandello, Sylvia Puente, and Adrian Rantilla. 1999. "'When You Call Me That, Smile!' How Norms for Politeness, Interaction Styles, and Aggression Work Together in Southern Culture." *Social Psychology Quarterly*, 257–275.

Cohen, Dov, Richard E. Nisbett, Brian F. Bowdle, and Norbert Schwarz. 1996. "Insult, aggression, and the southern culture of honor: 'an experimental ethnography.'." *Journal of Personality and Social Psychology*, 70(5): 945.

Derné, Steve. 1992. "Beyond institutional and impulsive conceptions of self: Family structure and the socially anchored real self." *Ethos*, 20(3): 259–288.

Drèze, Jean, and Amartya Kumar Sen. 2002. *India: Development and participation.* New Delhi:Oxford University Press.

Drèze, Jean, and Haris Gazdar. 1997. "Uttar Pradesh: The burden of inertia." In *Indian development: Selected regional perspectives.* , ed. Jean Drèze and Amartya Sen, 33–128. Oxford:Oxford University Press.

Drèze, Jean, and Naresh Sharma. 1998. "Palanpur: population, society, economy." In *Economic Development in Palanpur over Five Decades.* , ed. Peter Lanjouw and Nicholas Stern, 3–113. Oxford:Oxford University Press.

Dube, Leela. 1996. "Caste and women." In *Caste: Its 20th century avatar.* , ed. M. N. Srinivas, 1–27. New Delhi:Penguin Books.

Fehr, Ernst, and Klaus M Schmidt. 1999. "A theory of fairness, competition, and cooperation." *The Quarterly Journal of Economics*, 114(3): 817–868.

Fernández, Raquel. 2013. "Cultural change as learning: The evolution of female labor force participation over a century." *The American Economic Review*, 103(1): 472–500.

Fisman, Raymond, and Edward Miguel. 2007. "Corruption, norms, and legal enforcement: Evidence from diplomatic parking tickets." *Journal of Political Economy*, 115(6): 1020–1048.

Fudenberg, Drew, and David K Levine. 1998. *The theory of learning in games.* Vol. 2, Cambridge:MIT press.

Galanter, Marc. 1984. *Competing equalities: law and the backward classes in India.* Berkeley:University of California Press.

Goffman, Erving. 1974. *Frame analysis: An essay on the organization of experience.* Boston:Harvard University Press.

Greif, Avner. 1994. "Cultural beliefs and the organization of society: A historical and theoretical reflection on collectivist and individualist societies." *Journal of Political Economy*, 912–950.

Greif, Avner, and Guido Tabellini. 2010. "Cultural and institutional bifurcation: China and Europe compared." *The American Economic Review*, 100(2): 135–140.

Grosjean, Pauline. 2011. "A history of violence: the culture of honor as a determinant of homicide in the US South." Australian School of Business Working paper.

Guiso, Luigi, Paola Sapienza, and Luigi Zingales. 2008. "Long term persistence." National Bureau of Economic Research.

Hardin, Russell. 1989. "Why a Constitution?" In *The Federalist Papers and the New Institutionalism.* , ed. Bernard Grofman and Donald Wittman, 100–120. New York:Agathon Press.

Heine, Steven J. 2012. *Cultural psychology.* WW Norton New York.

Henrich, Joseph, and Jean Ensminger. 2011. "Theoretical foundations: The coevolution of social norms, intrinsic motivation, markets and the institutions of complex societies." *mimeo.*

Hitchcock, John T. 1958. "The Idea of the Martial Rājpūt." *The Journal of American Folklore*, 71(281): 216–223.

Hobbes, Thomas. 1651. *Leviathan.* Oxford:Oxford University Press. Reprinted in 1996.

Hoff, Karla. 2001. "Beyond Rosenstein-Rodan: The modern theory of coordination problems in development." In *Proceedings of the Annual World Bank Conference on Development Economics 2000.* 145–188. Washington, D.C.:The World Bank.

Hoff, Karla, and Joseph E Stiglitz. 2002. "After the big bang? Obstacles to the emergence of the rule of law in post-communist societies." *American Economic Review*, 94(3): 753–763.

Hoff, Karla, Mayuresh Kshetramade, and Ernst Fehr. 2008. "Spite and development." *The American Economic Review*, 98(2): 494–499.

Hoff, Karla, Mayuresh Kshetramade, and Ernst Fehr. 2011. "Caste and Punishment: the Legacy of Caste Culture in Norm Enforcement*." *The Economic Journal*, 121(556): F449–F475.

Kahneman, Daniel. 2011. *Thinking, fast and slow.* New York:Farrar, Straus and Giroux.

Kahneman, Daniel, and Amos Tversky. 1979. "Prospect theory: An analysis of decision under risk." *Econometrica*, 263–291.

Kapur, Devesh, Chandra Bhan Prasad, Lant Pritchett, and D Shyam Babu. 2010. "Rethinking inequality: Dalits in Uttar Pradesh in the market reform era." *Economic & Political Weekly*, 45(35): 39–49.

Khare, R. S. 1984. *The untouchable as himself: Ideology, Identity, and Pragmatism Among the Lucknow Chamars.* Cambridge:Cambridge University Press.

Kross, Ethan, and Ozlem Ayduk. 2011. "Making meaning out of negative experiences by self-distancing." *Current Directions in Psychological Science*, 20(3): 187–191.

Kushner, Harold J, and George Yin. 2003. *Stochastic approximation and recursive algorithms and applications.* Vol. 35, New York:Springer-Verlag.

Loewenstein, George F, Leigh Thompson, and Max H Bazerman. 1989. "Social utility and decision making in interpersonal contexts." *Journal of Personality and Social psychology*, 57(3): 426–441.

Mandelbaum, David G. 1993. *Women's seclusion and men's honor: Sex roles in North India, Bangladesh, and Pakistan.* Tucson:University of Arizona Press.

Matsuyama, Kiminori. 1996. "Economic development as coordination problems." In *The role of government in East Asian economic development: comparative institutional analysis.* , ed. Masahiko Aoki, Hyung-Ki Kim and Masahiro Okuno-Fujiwara. New York:Oxford University Press.

Merton, Robert K. 1968. *Social theory and social structure.* Free Press, New York.

Nethaway, Rowland. 1996. "Southern White Men on Honor." *Orlando Sentinel*, G5.

Nisbett, Richard E, and Dov Cohen. 1996. *Culture of honor: The psychology of violence in the South.* Colorado:Westview Press.

Nisbett, Richard E, et al. 1993. "Violence and US regional culture." *American Psychologist*, 48: 441–441.

North, Douglass C. 1990. *Institutions, institutional change and economic performance.* Cambridge:Cambridge University Press.

North, Douglass C. 2006. *Understanding the process of economic change.* Academic Foundation.

Nunn, Nathan. 2012. "Culture and the historical process." *Economic History of Developing Regions*, 27(sup1): 108–126.

Pitt-Rivers, Julian. 1966. "Honour and social status." In *Honour and shame: The values of Mediterranean society.* , ed. J.G. Peristiany, 19–77. Chicago:University of Chicago Press.

Rabin, Matthew. 1993. "Incorporating fairness into game theory and economics." *The American Economic Review*, 1281–1302.

Rao, Vijayendra, and Radu Ban. 2007. "The political construction of caste in South India." *The World Bank*.

Ray, Debraj. 2000. "What's new in development economics?" *The American Economist*, 44(2): 3–16.

Sapienza, Paola, Luigi Zingales, and Luigi Guiso. 2006. "Does culture affect economic outcomes?" National Bureau of Economic Research.

Schelling, Thomas C. 1978. *Micromotives and Macrobehavior.* New York:WW Norton.

Skyrms, Brian. 2003. *The stag hunt and the evolution of social structure.* Cambridge:Cambridge University Press.

Srinivas, Mysore Narasimhachar. 1962. *Caste in modern India and other essays.* London:Asia Publishing House.

Swidler, Ann. 1986. "Culture in action: Symbols and strategies." *American Sociological Review*, 51(2): 273–286.

Tabellini, Guido. 2008. "The scope of cooperation: Values and incentives." *The Quarterly Journal of Economics*, 123(3): 905–950.

Vaid, Divya. 2007. "An Empirical Exploration of the Relationship between Caste, Class and Mobility in India1." Yale University Working paper.

Van Huyck, John B, Raymond C Battalio, and Richard O Beil. 1990. "Tacit coordination games, strategic uncertainty, and coordination failure." *The American Economic Review*, 80(1): 234–248.

Zerubavel, Eviatar. 1999. *Social mindscapes: An invitation to cognitive sociology.* Cambridge:Harvard University Press.