Material Incentives and Effort Choice: Evidence from an Online Experiment Across Countries[*]

Elwyn Davies[†] Marcel Fafchamps[‡]

May 2025

Abstract

We conduct an interactive online experiment framed as an employment contract. Subjects from the US, India, and Africa are matched within and across countries. Employers make a one-period offer to a worker who can either decline or choose a high or low effort. The offer is restricted to be from a variable set of possible contracts. High effort is always efficient. Some observed choices are well predicted by self-interest, but others are better explained by conditional reciprocity or intrinsic motivation. Subjects from India and Africa follow intrinsic motivation and provide high effort more often. US subjects are more likely to follow self-interest and reach a less efficient outcome on average, but workers earn slightly more. We find no evidence of stereotypes across countries. Individual characteristics and stated attitudes toward worker incentives do not predict the behavioral differences observed between countries, consistent with cultural differences in the response to labor incentives.

JEL Codes: J31, D9, O12, O57

Keywords: work incentives; labor contracts; intrinsic motivation; conditional cooperation; cross-country comparisons

[*]We thank Jean-Marie Baland, Abigail Barr, Vittorio Bassi, John Strauss, Jonathan Schulz, Frank Schibach, Moses Shayo, David Laitin, Edo Gallo, participants to the 2019 New School of Economics conference on relational contracting, participants to the June 2020 Conference on Economic Development and Institutions, and seminar participants at the University of Southern California for useful comments. We thank Max Eber for excellent research assistantship on the Facebook part of the study. Funding for this study was provided by the Stanford Institute for Innovation in Developing Economies (SEED). IRB approval was obtained from Stanford University (Protocol 37472).

[†]The World Bank. Email: edavies@worldbank.org

[‡]Stanford University, NBER, and IZA. Email: fafchamp@stanford.edu

1 Introduction

In this paper we study how the terms of labor contracts affect effort using an experimental framework that focuses on three main behavioral patterns that have been documented in various settings: responding to incentives; reciprocating cooperation; and following intrinsic motivation. We do this using an multi-country interactive online experiment framed as a non-repeated job contract between a worker and an employer. The employer needs a task to be performed, and this requires the worker to provide effort (e.g., physical effort, time, attention, care, dedication). The employer can choose the wage offer, which can be high or low, and can decide to either condition the wage on effort or to pay a fixed wage agreed at the time of hiring. This situation captures many short-term employment contracts as well as many types of one-shot commissioned work (e.g., plumbing repair, taxi driver, coding, gig economy).

The canonical solution to this agent-principal problem is to offer a wage contingent on effort. This gives the worker an incentive to do the task well as long as the material incentive covers the effort cost. This solution, however, may not always work in practice. The literature has proposed other mechanisms to induce effort by combining material, behavioral, and social incentives (Rebitzer and Taylor, 2011; MacLeod, 2011). The identification of which mechanism works best ultimately remains an empirical question. Of particular interest for our paper is the finding that highly leveraged material incentives fail to elicit significantly more effort (e.g., Dellavigna and Pope, 2018a, 2018b). One possible explanation for this pattern is that workers are dedicated and conscientious even in the absence of strong material incentives – a feature either ascribed to reciprocal altruism towards the employer (e.g., Fehr et al. 1998; Fehr and Gächter 2000; Charness 2004; Landry et al. 2011) or to intrinsic motivation (e.g., Charness and Dufwenberg 2006; Cassar and Meier 2018; Cassar 2019).[1] In either case, employers would find it unnecessary or even unprofitable to introduce highly leveraged incentives (e.g., Hwang and Bowles 2014). There could also be other reasons why an incentivized worker provides low effort, such as spite or disdain for the employer, lack of attention and care, or poor work ethics. Similarly, incentivizing the worker may not be needed: the worker would do the work even for low pay; or would do the work for a high enough pay without the need to be monitored. In such cases, the employer could have the task done at lower cost, without incentivizing or monitoring the worker. Our objective in this paper is to study whether such situations do arise, in a controlled environment.

What we do *not* do in this paper is to try to ascertain *why* workers provide effort under different contractual arrangements. There are many causal mechanisms – other than self-interest – that can influence workers' choice of effort (e.g., Fields 2011). A non-exhaustive list based on the existing behavioral economics literature includes: guilt-aversion; social preferences; group identity; self-image and social image; and preference over process. Each of these causal mechanisms

[1] In a related paper, DellaVigna et al. (2022) vary piece rates in addition to gift treatments. They do not detect any effect of the gifts on productivity but find a sizable positive impact on labor supply, a finding that the authors also interpret as evidence of reciprocation. Experimental evidence further indicates that non-material incentives can increase effort (e.g., DellaVigna and Pope 2018a; Ashraf et al. 2015).

can affect effort in a given contract, and they can do differently depending on their *prescriptive content*, i.e., what behavior they dictate. Guilt-aversion can produce any behavior depending on the nature of the injunctive norms workers hold – and these norms can depend on the form of the contract. Social preferences can similarly take many forms, some of which are altruistic and others destructive (e.g., spite, paternalism, desire for control). Group identity can help cooperation between worker and employer but also hinder it, for instance if groups are inimical to each other or if one group is seen with disdain. Self-image and social image considerations depend on what image of themselves people seek to promote: some people may wish to appear diligent, others may wish to appear rebellious or individualistic. Finally, preferences over process can take many forms. Some workers may favor certain incentives (e.g., a bonus) but reject others (e.g., docking pay). They may even resent performance pay because it is seen as manipulative, disrespectful, or distrusting (e.g., Hwang and Bowles 2014).

Two key insights emerge from this rapid overview. First, for efficiency and equity reasons, it is critical that we know the extent to which workers and employers deviate from canonical model. Second, given that the reasons why workers and employers may deviate from the canonical model all depend on their prescriptive content, we expect variation not only across individuals but also across populations. Consequently, we also would like to know the extent to which different populations vary in the way they deviate from the canonical economic model of worker-employer relationship – and whether this variation can be predicted based on individual characteristics or may contain a population-specific component.

To make progress on both issues, we conduct an online interactive experiment framed as a series of one-shot work contracts between pairs of participants.[2] The experiment is designed around four behavioral 'archetypes' about what the worker does in response: always provide effort (A1); provide effort only when paid a high wage (A2); provide effort only when incentivized (A3); never provide effort (A4). These archetypes map into behavioral patterns discussed in the literature (e.g., Fehr and Gächter 2000, Hwang and Bowles 2014): intrinsically motivated for A1, conditional reciprocator for A2, selfish-rational for A3, and unconditionally unmotivated for A4. This is why we sometimes refer to these archetypes using the same vocabulary.

Our focus is on documenting the prevalence of these archetypes in various populations around the world. This is achieved by conducting the experiment online with subjects from seven countries across three continents (e.g., Horton et al. 2011; Charness and Kuhn 2011; Charness et al. 2013). The study countries were chosen to cover as varied a set of countries as was possible within our limited budget and the constraints imposed by the online format. We also investigate whether participants assigned the role of employer choose to offer labor contracts that best-respond to the behavior of those assigned the role of workers.

Subjects play eight incentivized games in pairs – four as employer and four as workers, in random order. They never play twice with the same person, so as to rule out the repeated play considerations already studied elsewhere (e.g., Fehr et al. 1997; Bell and Freeman 2001; Baker

[2]Prisse and Jorrat (2022) have shown that online experiment yield results similar to those obtained in lab experiments.

et al. 2002; Gachter and Falk 2002; Brown, Falk and Fehr 2004, 2012; Davies and Fafchamps 2021). Each game is framed as a one-shot employment contract between two subjects, one of which is assigned the role of employer and the other the role of workers. This framing serves to trigger wage employment heuristics among subjects. Employers choose to offer a job or not, and they select a contract type: low or high wage, conditional on effort or not, with full commitment or not. These labor contracts aim to capture, in a stylized way, the range of choices available to employers and faced by workers, such as fixed wage or performance pay, with or without 'wiggle room' for the employer in terms of respecting their contractual obligation to pay a bonus. Contingent contracts come in two versions: a bonus contract where the worker receives an extra payment for high effort; and a malus contract where the worker incurs a pay cut for low effort. They yield identical payoffs but are framed differently, which may trigger different heuristics.[3] Games differ in the type of contracts included in the choice set of the employer. This choice set is not revealed to the worker, who chooses to accept the offer or not and then selects an effort level – i.e., high or low. High effort always maximizes aggregate efficiency but is costly to the worker. While this design is vaguely reminiscent of Brown et al. (2004, 2012), here employers choose a contract, not a wage level.

The first set of participants were recruited on Amazon Mechanical Turk in the US and India. At the time, this was the only way of recruiting and paying experimental subjects online. To broaden the sample population, a second set of participants were recruited from ten countries through targeted ad campaigns on Facebook. Most of analysis is based on subjects coming from seven countries covering three main regions of the world: USA, India, and Africa.[4] In most cases, subjects play with someone from their own country. But some subjects were assigned to play with someone from another country, in which case they are told they are playing against someone from that country. Because our focus is on heuristics and beliefs, we want subjects to interact with each other in real time so as to make human interaction more vivid. Given the time constraints imposed by interactive online experiments – i.e., 15 minutes of interactive play at most – this rules out using real effort tasks. Participants also fill an online questionnaire before the game.

We first examine whether the contract acceptance and effort levels chosen by workers vary by contract offer. In each of the three regions we study (US, India, and Africa), we find sizable proportions of subjects in each of our four archetypes, with a lot of similarities between regions. The most common archetypes are those that do not respond to either the wage level or the presence of an incentive, that is, A1 and A4. They account for around 60% of the US sample and 80 to 90% of the India and Africa samples, and are more heavily loaded on always providing effort. Archetypes A2 and A3 account for a larger fraction of observed choices in the US samples, but they never represent more than 25% of play for A2 and 20% of play for A3. In the non-US

[3] Docking pay, for instance, is illegal in some countries, while the size of bonuses is socially constrained in others (e.g., Ockenfels et al. 2015).

[4] We also collected data from three other high-income countries, but the Facebook ad campaigns there yielded sample sizes that are too small for analysis.

samples, archetype A3 only accounts for less than 10% of observed choices. From this evidence, we conclude that material incentives are not what motivates workers in general, although they matter a bit more in our US samples. Since US participants are less likely to provide effort if not incentivized, we also find that efficiency is lower in the US samples but worker payoffs are slightly higher. These behavioral differences across regions are partly matched by differences in beliefs about the effectiveness of worker incentives and about the acceptability of sanctions for workers who shirk or are incompetent.

In a subset of our sample, we investigate whether subjects behave differently when they are matched with someone from another country. In contrast to Banuri et al. (2022), we find no evidence that subjects choose systematically lower effort when matched with a foreign participant. We also investigate whether subjects assigned the role of employer expect workers from another country to work less or more than workers from their own country. We find little or no evidence of such stereotypes in our data.

In the last Section of the paper, we use a counter-factual thought experiment to examine whether differences in behavior across regions could have been predicted from subject characteristics – keeping in mind that, in the experiment itself, subjects are not given any information about other subjects, apart from their country. The question we investigate is whether subjects from one region could predict the behavior of subjects in another region by extrapolating the correlation between behavior and individual characteristics that is present in their own region. The possibility is strongly ruled out by our results. While the evidence at our disposal does not allow us to ascertain what causes the behavioral differences across regions, we cannot rule out is that they capture cultural differences, as discussed in Henrich et al. (2001, 2006, 2010), Guiso et al. (2006), Falk et al. (2018), and Schulz et al. (2018, 2019).

This paper first contributes to a large empirical literature on the effectiveness of labor incentives (e.g., Weiss 1987; Lazear 2018; Grosnell et al. 2020). Dellavigna and Pope (2018a, 2018b) provide a summary of the experimental evidence on the effect of various worker incentives in advanced economics. Experimental results confirm that material incentives increase effort, but the marginal effectiveness of additional incentives is low. Our own results generalize these findings to include parts of the world that are seldom included in such studies. We find strong similarities across our sample populations, irrespective of region.

Second, we contribute to a small but growing literature on behavioral predictions. A good illustration of this line of research is the work of Dellavigna and Pope (2018a, 2018b) who test whether experts are able to predict the effect of various worker incentives in advanced economics. While they find that experts predict some key patterns in the data, a significant share of experts – economists and non-economists alike – mistakenly expect a small piece-rate to crowd out incentives to provide effort. More work is needed in this area because being able to predict other people's behavior with some accuracy is essential for social and economic interactions. Our approach to this question reverts around the idea that people form mental models of behavior and use them to predict how others will behave. These predictions typically vary with people's characteristics. We contribute to this literature by offering a novel approach for testing whether predictions based on correlation patterns found in one population help predict the behavior of another population. If they do not, we interpret this as a possible indication of cultural differences – in the sense that a decision maker faced with a new population is less able to 'understand' their behavior. We believe this approach offers a new way of analyzing the difficulties that firms experience when operating in unfamiliar markets (e.g., Bloom and Van Reenen 2007, 2010).

Third, we add relevant cross-country evidence to a growing literature on the relationship between culture and economic outcomes through its influence on people's expectations and preferences (Guiso et al, 2006). Empirical work at the macro level has provided evidence that culture matters for financial development (Guiso et al, 2004), performance of large organizations (La Porta et al, 1997), and venture capital investment (Bottazzi et al, 2016).[5] In controlled experimental settings, culture affects play in dictator (Henrich et al, 2006), ultimatum (Henrich et al, 2001), coordination (Jackson and Xing, 2014), minimum effort, and prisoner's dilemma games (Chen et al, 2014). Other authors have compared the rigidities that affect labor markets in different countries (e.g., Nickel 1997, Field 2010). Our study documents differences in the way human subjects from different regions of the world use respond to wage level and incentives in a game framed as a short-term employment contract. We do not, however, aim to identify why cross-cultural difference may be present, as this would require a completely different research design – e.g., focusing on the role of social norms (e.g., Enke 2019, Chuah et al 2023) or identity/social proximity (e.g., Bicchieri 2022).

It is important to note that our experimental design deliberately rules out making contributions to some other literatures – partly to avoid repeating other work, but mostly to ensure a clean identification of our research question. The literature has also shown that dynamic incentives matter. For instance, Kajackaite and Werner (2015) show that a minimum performance requirement has no direct effect on output in a real effort experiment, but has undesired effects in the dynamics of controlled agents' working performance. In a similar vein, Sliwka and Werner (2017) find that agents work harder under increasing wage profiles only if they do not know these profiles in advance. The authors interpret these findings as evidence of reciprocal altruism in a game with varying reference points. In our experiment, we deliberately eschew any type of dynamic considerations: subjects play a sequence of one-shot games with systematically different opponents.

In addition, we do our best to eliminate fairness considerations between workers from contaminating our results. The literature has shown that the effectiveness of material incentives can be reduced by considerations of fairness between workers (e.g., Fehr and Schmidt 2007; Abeler et al. 2010; Andreoni et al 2020). For instance, workers may find highly leveraged incentives unfair, which could crowd out intrinsic motivation and other non-material incentives (e.g., Benabou and Tirole 2006). In support of these ideas, Breza et al. (2018) find that incentivizing Indian casual workers through relative pay has a negative effect on performance. Cullen and

[5] Chu and Fafchamps (2022) provide qualitative evidence of culture clash between Chinese employers and local workers in Ethiopia.

Perez-Truglia (2018) find that workers in an international multinational reduce their effort level when informed that other workers similar to them earn more than them in the firm.[6] Bandiera et al. (2010) find that when workers who know each other well are placed in the same piece-rate work team, more productive workers reduce their effort so as to not overshadow their workmate. In all these examples, workers seem to have a distaste for pay differences across similar or proximate workers: in Cullen et al. (2019) and Breza et al. (2018), workers reduce effort when they learn that their peers are paid more; and in Bandiera et al. (2010) high performance workers exert lower effort to reduce the pay difference with less able colleagues.[7] To ensure that fairness considerations do not contaminate our findings, we restrict our attention to situations involving a single worker and employer. We also ensure that each subject gets an equal opportunity to take the roles of employer and worker.

We start in Section 2 by presenting the experimental design in detail. The conceptual framework underlying the study is discussed in Section 3 and implementation details are discussed in Section 4. Results from the experiment itself are the object of Section 5. In Section 6 we investigate whether individual behavior in the experiment can be predicted based on individual characteristics and answers to questions about the acceptability and effectiveness of various incentives. Section 7 concludes.

2 Experimental design

We design an experiment to test the different conjectures discussed in the introduction. To this effect, we create an online game between two subjects, one of whom is assigned the role of employer and the other the role of worker. Each subject plays four such games in sequence and is re-matched with a different partner each time. Subjects then switch roles and play four more games in the other role.

2.1 General design

The behavioral archetypes we are interested in are potentially linked to ethical considerations. If subjects were to play an individual decision game against a computer, moral considerations would likely disappear and we would expect conditional cooperation and intrinsic motivations to be crowded out by selfish-rational reasoning. For this reason, it is essential that there be an interpersonal element to our experimental design, and this requires using an interactive game.

A second consideration is that we want to study behavioral archetypes across regions of the

[6]In agreement with this interpretation, Ockenfels et al. (2015) find that a multinational corporation operating in the US and Germany has a much more compressed bonus scale for its managers in the latter than in the former. The authors ascribe this difference to the fact that, in Germany, the size of bonuses paid is public to the workers and this, the authors speculate, reduces their effectiveness in eliciting effort and thus the need to pay large bonuses.

[7]Bandiera et al. (2013) nonetheless show that rank tournaments among workers do increase productivity, thereby drawing a distinction between the effect of tournaments per se and the way they are rewarded – an observation that also appears in Ockenfels et al. (2015).

world. This means opting for an interactive online experiment, something that is notoriously difficult to achieve. In any case, the timing of our study overlaps partly with the Covid-19 pandemic, making an online experiment the only feasible option during our study period. Running an interactive international online experiment requires keeping the duration of the experiment short. This rules out asking subjects to undertake real performance tasks. We therefore opt for a one-shot 'gift exchange' game (e.g., Brown, Falk and Fehr 2004) framed as an employment contract (e.g., Brown et al. 2012; Davies and Fafchamps 2021).

Since it has already been shown that populations from different countries may play repeated employment games differently (Davies and Fafchamps 2021), we focus here on non-repeated contracts: subjects only play once against each other. Since subjects tend to earn more in the employer role than in the worker role, we ask subjects to play both roles, in random sequence, to guarantee equal treatment across them.

2.2 Contracts

The structure of the basic stage game resembles a one-sided Prisoner's Dilemma (PD): the subject in the role of employer makes an irrevocable wage offer to the subject in the role of worker, after which the worker chooses an effort level that determines the employer's payoff. In such a game, intrinsically motivated subjects may cooperate. But selfish-rational players will not: the standard Nash equilibrium of the game is non-cooperative.

Taking this simple game as starting point, we examine whether increasing the wage triggers increased effort, which would be suggestive of conditional cooperation. We also introduce an incentive contract that conditionally compensates workers for the cost of high effort and induces selfish-rational subjects to cooperate. Finally, we add a reneging option in some incentive contracts in order to mimic situations observed in many countries, whereby employers deviate from the worker's employment contract by docking pay or cancelling an expected bonus. These different contracts are chosen because they mimic, in a simple way, commonly observed employment contracts.

In each stage game, employers have a limited choice of contract offers that they can make. Table 1 summarizes the payoffs corresponding to each of the possible contracts and effort levels. Payoffs are presented in points. As is immediately apparent from the Table, across all contracts and effort levels, employers typically earn more than workers. This is a deliberate choice intended to capture the fact that, on average, employers earn more than the majority of their employees. Since subjects take turns in the two roles, this guarantees equal opportunity to all participants.

From the first row of the Table, we see that when no offer is made, workers and employers receive a reservation payoff of 20 points. In both cases, this reservation payoff is chosen to equate the expected earnings from an alternative low-wage-low-effort match, which corresponds to a standard reservation utility concept.[8] This implies that workers cannot (pay to) punish

[8]For workers, this reservation payoff is composed of a fixed payment of 10 points per round, plus an extra 10 points meant to capture the worker's expected earnings from a low-wage-low-effort elsewhere. For the employer, it represents the payoff from hiring another worker at a low wage and receiving low effort.

employers for making unattractive offers – and similarly for employers. This deliberate design choice avoids any contamination of play by punishment strategies – a topic that has already been studied extensively and need not be revisited here.

The next two rows present the payoff levels in fixed wage contracts. These contracts only differ in the wage offer made by the employer, which is 10 in the low fixed wage contract and 20 in the high fixed wage contract. The worker's payoff is equal to this wage plus the fixed payoff of 10 per round, minus the cost of effort, which is 0 for low effort and 5 for high effort. The employer's payoff is the income from production – which depends on effort - minus the wage paid to the worker. We note that the employer's payoff increases by 30 points with high effort – which is the marginal return to high effort – while the worker's payoff falls by 5 points. This means that high effort is efficient, i.e., it always maximizes joint payoffs. The reservation payoffs of the worker and employer are equal to the low-wage-low-effort case, which represents the inefficient Nash equilibrium of a one-sided PD game in which the employer chooses a wage offer and the worker choose effort. The experiment investigates the conditions under which the subjects can achieve an efficient outcome.

Table 1. Payoffs in the stage game

Contract	Worker payoff		Employer payoff	
No offer	20		20	
	Low effort	High effort	Low effort	High effort
Low fixed wage	20	15	20	50
High fixed wage	30	25	10	40
Bonus/Malus	20	25	20	40
Bonus/Malus WD(*)	20 *or 30*	15 *or 5*	20 *or 10*	40 *or 50*

Notes: (*) WD=with discretion. In this case, payoffs depend on whether the employer reneges on the contract, either to pay more or to pay less to the worker.

In the Bonus and Malus contracts, effort is incentivized: the wage paid to the worker is 10 points for low effort and 20 points for high effort. There are two versions of this incentivized contract that are equivalent in terms of payoff but are framed differently. In the Bonus contract, the worker is promised a wage increase of 10 points when choosing high effort; otherwise, the worker receives a wage of 10 points. In the Malus contract, the worker is told the wage of 20 points will be reduced by 10 points if they choose low effort. The reason for including both contracts in the experiment is to investigate whether workers suffer from a framing effect. Indeed, in many countries docking pay for poor performance is not allowed by law, but offering a bonus for good performance is allowed – suggesting that Bonus and Malus contracts are not equally acceptable, and this could crowd out intrinsic incentives.

We also include Bonus and Malus contracts 'with discretion' (WD) in which the employer can, ex post, renege on the link between effort and the wage. In the Bonus WD contract, the employer can decide *not* to pay a bonus after high effort – but also to pay a bonus after low effort. Similarly, in the Malus WD contract, the employer can pay a high wage after low effort – but also pay a low wage after high effort. As noted earlier, the possibility to renege on the promise of a high wage is introduced to mimic situations that frequently arise in many settings and is thus familiar to subjects. It also offers the advantage of allowing both the employer and the worker to defect, thereby broadening the range of behaviors that we are able to observe.

It is immediately apparent that low effort is always selfish-rational for worker in a fixed wage contract, while high effort is always selfish-rational in Bonus/Malus contract without employer discretion. Similarly, paying less after high effort is always selfish-rational for employer in the two WD contracts with ex-post discretion. Examining the extent to which subjects follow these predictions is the central objectives of our experimental design.

2.3 Stage games

There are six different stage games that subjects can play. They differ in the choice of contracts available to the employer, shown in Table 2. In treatments T_l and T_h, the employer can either make no offer, or make an unconditional offer of a low fixed wage (in T_l) or a high fixed wage (in T_h). In treatment T_b, the employer can make no offer, offer a low fixed wage, or offer a bonus contract that increases the wage in case of high effort. Conversely, in treatment T_m, the employer can make no offer, offer a high fixed wage, or offer a malus contract that decreases the wage in case of low effort. Treatments T_{bd} and T_{md} are similar to T_b and T_m, except using bonus and malus contracts that allow the employer to renege on the conditionality of the contract ex post, either by paying a low wage for high effort (in T_{bd}) or a high wage for low effort (in T_{md})

Table 2. Choice of contracts available to the employer in each treatment

Available offers to employer:	T_l	T_h	T_b	T_m	T_{bd}	T_{md}
High fixed wage		$\surd$		$\surd$		$\surd$
Low fixed wage	$\surd$		$\surd$		$\surd$	
Bonus for high effort			$\surd$			
Malus for low effort				$\surd$		
Bonus with discretion					$\surd$	
Malus with discretion						$\surd$
No offer	$\surd$	$\surd$	$\surd$	$\surd$	$\surd$	$\surd$

Subjects play four randomly selected games among the possible six. Each game has the following structure. The employer moves first. In most treatments, the employer can choose between two different contracts. In all treatments, the employer can also decide not to offer any contract. If no contract is offered, the game ends. If a contract is offered, it is the worker's turn to move. The worker can decide to refuse the contract, in which case the game ends. If the worker accepts the contract, the worker then chooses a level of effort, either high or low. Except in the two WD contracts, this ends the game and the payoffs are those presented in Table 1. In the Bonus and Malus WD contracts, the employer can decide to renege on the contract after observing the worker's effort choice. Reneging in the Bonus WD contract allow the employer to avoid paying a bonus after high effort; in the Malus WD contract, it allows not imposing the penalty (i.e., malus) after low effort.

The experiment is specifically designed to minimize dynamic play considerations. Strategic repeated play *across* games is prevented by rematching subjects before each game, keeping full anonymity across all games, and precluding communication across subjects to eliminate reputation effects. While we cannot rule out that subjects learn over the eight games that they play as employer and worker, they never play the same game twice in the same role, thereby reducing what they can learn about one particular strategic environment. The order of play between games is also randomized, making it harder for subjects to keep track.[9] This design helps ensuring that the behaviors we observe are driven by heuristics, which is what we aim to measure in this experiment.

3 Implementation

We conducted the online lab experiment using an identical experimental design, instructions to participants, screen visuals, and online interaction process with individuals recruited from ten countries and two languages (English and French). The sample of countries was chosen to cover a wide range of GDP per capita and rule of law (see Appendix Figure A1). It also covers a large fraction of the world: the US and India together account for 25% of the world's population, while adding Africa as a whole encompasses half the world population today.

One set of participants were recruited using Amazon Mechanical Turk (MTurk), another using Facebook (FB). To account for this difference in recruitment process, we present results separately for the MTurk and Facebook participants. In three of the targeted Facebook countries (France, Australia, and Canada), we were unable to recruit a large enough sample of participants and for this reason these countries are dropped from the analysis presented here.[10] A detailed presentation of the sample recruitment process on MTurk and Facebook is presented in the Online Appendix.

The experiment was implemented in two batches, depending on the way subjects were recruited. The first batch of participants were recruited through MTurk, which limited recruitment to two countries. After successful implementation with this population, we expanded to geographical reach of the experiment and the representativeness of the subject population by recruiting participants via Facebook. Neither of these two recruitment methods guarantees a representative sample of the study country's population, which is a drawback. [11] Unfortunately,

[9] By the same logic, it would be difficult for subjects to follow a contagious equilibrium strategy a la Kandori (1992): each game is different, and signal extraction about types or equilibrium strategies is probably beyond the computational capacities of the overwhelming majority of players, especially given the fact that the entire experiment lasts approximately 15 to 20 minutes.

[10] The very high recruitment cost imposed by Facebook made it impossible to reach a large enough sample while staying within our grant budget.

[11] Shared by the majority of laboratory or RCT experiments.

there were no feasible alternatives at the time we started the research.[12] By the time fieldwork ended, the Covid-19 pandemic was in full force, which made it impossible to expand the study population to in-person samples – but this should not be a concern (Prissé and Jorrat 2022).

3.1 Understanding of the games by the participants

In the Online Appendix, we present screen shots for all the stages of the online game. As is clear from these pictures, we put a lot of effort into making the interface friendly and easy to use. As a result, we received no complaints from subjects about the game interface. To minimize the cognitive burden of the game and allow subjects to concentrate on strategic issues rather than mental arithmetic, we calculate all payoffs for subjects directly on the screen and, in some cases, we depict payoffs graphically in colored bar charts. This stands in contrast with many laboratory experiments that expect subjects to undertake complex calculations on the fly, such as Bayesian updating or compound interest calculations.

Before allowing participants to engage with the experiment proper, we tested their understanding of the game. In both the MTurk and the Facebook samples, participants' knowledge was tested using a small quiz at the beginning of the experiment. Subjects were asked to answer questions about payoffs from various actions. Those who answered a question incorrectly were offered a second chance to answer it correctly. Subjects who do not manage to provide a correct answer were not allowed to proceed with the experiment.

There are small differences in the share of correct answers to the first question between samples and regions. For both the US and India, the share of correct responses is about 10 percent lower for the Facebook sample than for the MTurk sample. For both the MTurk and Facebook samples, Indian participants have a lower share of correct answers than US participants (by 7 and 5 percent, respectively). But this difference is only significant at the 5 percent level for the MTurk sample. In the Facebook sample, African participants had a 4% higher proportion of correct answers than US participants, but this difference is not statistically significant at the 5 percent level. From this we conclude that, although there are slight differences in average understanding across our sample countries and regions, these differences are small in magnitude and only weakly significant. In any case, subjects who were unable to answer correctly were not allowed to participate in the experiment.

The analysis below also allows us to ascertain whether participants played randomly, which would be another manifestation of a poor understanding of – or interest in – the game. Since most decisions involve choosing between two options, this means testing whether subjects chose between the two with equal probability. Random play is rejected in essentially all cases.

[12] The population of available subjects on Prolific Academic was very limited at the time. Qualtrics was already offering representative survey samples in some countries and we approached them. Unfortunately, their policy at the time was to pay every participant a fixed fee and they refused to allow us to invite their survey respondents to our online experiment. Remotely recruiting participants by other methods – e.g., by posting ads on the radio or in newspapers – would have created an enormous potential for fraudulent participation, and great difficulties in paying small amounts to hundred of subjects across country boundaries due to international money laundering regulations. These difficulties largely remain today.

3.2 Sessions

Unlike in lab experiments for which subjects pre-commit a set amount of time and do not get paid if they leave before the end, online experiments attract participants from a much broader subject pool, partly because they take much less time and effort, and partly because they do not require committing to stay until the end. Drop-outs therefore occur for a variety of reasons, most of which have little to do with the experiment itself: people get distracted, they lose their connection, or they lose interest. As a result, a common difficulty in online experiments is that some participants do not stay until the end of the session. This is of particular concern in an interactive experiment like ours, in which subjects interact online in real time.

We reduce the impact of dropouts in two ways. First, we seek to induce participants to stay longer by paying them the payoffs they earn in each period – instead of selecting a few periods at random, as is common in lab experiments. Subjects therefore get more if they stay longer. Second, we rely on an adaptive matching algorithm to minimize the loss of observations due to dropouts. In each of the eight periods of the experiment, the algorithm reassigns, in real time, a participant to an alternative partner if its preassigned partner has dropped out. This feature proved to be an essential factor in the success of the project.

To show the incidence of dropouts, in Appendix Table A1 we present the number of participants that stayed n period in the experiment, with n going from 1 to 8. We see that a large fraction of participants (45.8%) stay for the whole 8 periods and 71% stay for six periods or more. The rest (29%) stay for five periods or less. In the analysis we include all the choices made by participants irrespective of the number of periods they stay in the experiment.

3.3 Qualtrics survey

Each subject starts by filling an online questionnaire on Qualtrics (see Online Appendix). In addition to collecting basic information about each subject, the questionnaire also gathers information about their work experience. Six vignette-style questions focus on the acceptability of different incentive schemes, and six questions elicit subjective beliefs about the reliability of incentivized and unincentivized workers in the US, India and South Africa – the latter country being included as additional information.

We present in Table 3 a breakdown of individual characteristics by country or region and by mode of recruitment. There are significant differences across samples, with MTurk subjects more likely to report being employed and less likely to be students. There are also differences in education levels (Indian and African subjects are more educated than US subjects), forms of employment (e.g., African subjects have less experience being in wage employment), gender (US subject are more likely to be female),[13] and age (US subjects are older). This notwithstanding, there is also considerable variation within each sample. In Section 5, we investigate in detail

[13] Our samples include a larger proportion of men than women, particular in India and Africa. Since participation rates in wage work are lower for women than for men, particularly in India and Africa, this makes our samples more representative of attitudes towards wage work.

whether differences in subject characteristics can account for the differences in average behavior. In terms of balance within the experiment, we find no difference across countries in treatment, mix of employer and worker role, or in being matched with a partner from the same country.

Table 3. Average characteristics of subjects across sub-samples

Variable	US Mturk	India Mturk	US FB	India FB	Africa FB
Age in years	38.2	31.2	36.4	30.0	31.1
Male	58.3%	73.9%	62.7%	76.1%	64.3%
Education					
Completed primary	7.0%	6.0%	2.8%	0.0%	5.0%
Some secondary	18.0%	6.3%	16.2%	6.9%	8.3%
Completed secondary	43.3%	33.2%	35.2%	24.5%	35.1%
Post-secondary	31.7%	54.5%	45.8%	68.6%	51.6%
Current employment status					
Full-time student	0.0%	0.0%	18.3%	16.4%	21.5%
In fixed-term/short-term wage employment	3.0%	12.8%	6.3%	5.7%	11.5%
In permanent wage employment	68.0%	50.0%	43.0%	44.7%	25.9%
In self-employment	19.3%	33.8%	12.0%	28.3%	23.9%
Unemployed/not-working	9.6%	3.4%	20.4%	5.0%	17.2%
Ever wage employed (if currently not)	93.3%	85.2%	73.9%	79.9%	49.4%
Ever in permanent employment (if currently not)	68.0%	50.0%	43.0%	44.7%	25.9%
Ever self-employed (if currently not)	53.7%	71.3%	32.4%	62.9%	57.9%
Number of observations	460	352	142	159	703

Note: Based on answers to the Qualtrics survey that all subjects fill before taking part to the online experiment. MTurk and FB refer to subjects recruited via Amazon Mechanical Turk and Facebook, respectively.

The questionnaire includes a few attitudinal questions about the acceptability of various types of work incentives. The first question is a vignette of the following form: "Worker A is hired to perform a task for which he/she claims to be qualified. After a week on the job, it becomes clear that A is unable to perform the task. Worker A is laid off by the employer." Respondents are then asked to rate the employer's decision from 0 to 10 in terms of acceptability, with 0 being fully unacceptable and 10 fully acceptable. Questions 2 and 3 follow the same format but vary the employer's response, i.e., to cut the worker's wage by 30% or to eliminate the worker's 30% bonus. The next three questions follow the same sequence, but the vignette focus on a worker who is caught shirking.

The questionnaire also includes questions on whether workers "can be trusted to exert high effort if their earnings and continued employment depend on their performance on the job". This question is asked separately about workers from the US, India and South Africa. A similar question is asked for when the workers' earnings and continued employment *do not* depend on their performance on the job. Answers to these questions are discussed in detail in Section 5.

4 Testing strategy

The empirical analysis is divided into two main parts. In the first part we examine non-strategic choices, that is, choices that are unaffected by expectations of others' future choices. We use these choices to classify subjects into four archetypes. In the second part we focus on strategic choices that require subjects to form expectations about how others will act.

Our testing strategy is divided into two main parts. The first part is devoted to players' actions that are not followed by another action by the other player. These actions are non-strategic in the sense that they do not require strategic thinking or perspective taking to form a belief about what the other player will do. These actions are those that allow us to assign actions to one of our four archetypes: does the worker provide high effort in the four contracts without reneging by the employer; and does the employer renege in the other two contracts.

Theoretical predictions for worker effort are summarized in Table 4. To provide intuition, we have associated a name from the behavioral literature with each archetype, in the sense that the motivations associated with those names would predict the archetype behavior. We do, however, reckon that other motivations can cause behaviors similar to these archetypes.

With this caveat in mind, we now present the worker choices associated with each of the four archetypes. Given the structure of the contracts, rational workers pursuing their material interest behave as predicted by game theory and follow archetype A3. If offered a low wage contract, they either reject it or provide low effort – both actions that give them the same payoff. When offered a high wage contract, they accept the contract but provide low effort. And they provide high effort in bonus and malus, the two contracts with wage contingent on effort. Conditional reciprocators, in contrast, choose an action that is beneficial for the other player if the other player has chosen an action that benefits them (e.g., Fehr and Gächter 2000). Given the sequential nature of our experimental design, this means choosing high effort if the employer has chosen a high wage contract. Conditional cooperators also pick high effort in the contingent contracts since these are reciprocal by design. We also consider subjects who pick an action for its intrinsic value to them, irrespective of what the other player chooses or what their material interest is. In our case, they can either choose high effort in all four contracts, in which case we call them intrinsically motivated (i.e., they choose an action that is beneficial to the other player); – or they choose low effort or reject all four contracts[14], in which case we call them intrinsically *un*-motivated (i.e., they choose an action that is not beneficial to the other player). Since subjects play both worker and employer roles over the duration of the experiment, and are randomized across the different contract offers, we can use their choices to estimate the proportion of actions falling into each archetype.

[14] Recall that which both yield the same payoff for the players.

Table 4. Worker choice conditional on archetype in the non-reneging contracts

Archetype:	A1	A2	A3	A4
	Intrinsically motivated	Conditional reciprocator	Selfish rational	Unconditionally unmotivated
Low wage contract	high effort	reject or low effort	reject or low effort	reject or low effort
High wage contract	high effort	high effort	low effort	reject or low effort
Bonus/malus contract	high effort	high effort	high effort	reject or low effort

We do something similar for employers in the two reneging contracts. Model predictions are summarized in Table 5. In archetype A1, the employer chooses the high wage since it is most beneficial to the worker. In contrast, an A4 employer would choose the low wage conditional on having make a bonus or malus contract offer. It is, however, more likely that an unmotivated A4 employer did not offer such a contract in the first place, and in which case they would not appear in the reneging data. In archetype A2, the prescribed behavior is conditional cooperation which, in this case, dictates to follow the contract (i.e., not to renege). In archetype A3, the selfish-rational choice is to always pay the low wage. In this case, we cannot distinguish between archetypes A3 and A4. If employers were randomly assigned to the Bonus WD and the Malus WD contracts, there should be no difference in reneging between them. In practice, employers self-select into these contracts, which can induce differences in the proportions of archetypes across the two contracts, a point we revisit below.

Table 5. Employer choice conditional on archetype in the reneging contracts

	A1	A2	A3	A4
Wage paid if effort is:	Intrinsically motivated	Conditional cooperator	Selfish rational	Unconditionally unmotivated
low	high [renege]	low	low	no offer or low
high	high	high	low [renege]	no offer or low [renege]

In the second part of the main analysis, we examine choices that potentially rely on subjects' beliefs about what the other player will do next. Here, actions are determined both by archetypes and beliefs. Since we do not observe beliefs, we cannot infer archetypes. But we can nonetheless draw some indirect inference regarding beliefs, conditional on archetypes – the proportion of which we estimated in the first part.

We examine in some detail the choice of offer made by employers, which is the more interesting decision. To recall, in treatment T_l the employer has the choice between no offer and offering a fixed low wage; in treatment T_h, the choice is between no offer and a fixed high wage; in treatment T_b, the employer can offer a fixed low wage or a bonus contract – or make no offer; and in treatment T_m, the choice is between a malus contract, a fixed high wage, or nothing.

Treatments with ex-post employer discretion T_{bd} and T_{md} are the same as T_b and T_m, respectively except that, in the bonus and malus contracts, the employer has full ex post discretion to pay the low wage or the high wage irrespective of worker effort.

By design, these choices depend not only on employer archetype but also to some extent on whether employers expect workers to accept the contract and provide high effort. Let's first discuss A3 (e.g., selfish rational) employers. For these, some choices weakly dominate others. In T_l the employer is indifferent between no offer and low-wage-low-effort, which is the best response of an A2 or A3 worker. But making an offer dominates if there is a small chance that the worker follows archetype A1 and chooses high effort. It follows that offering the fixed wage contract weakly dominates. In contrast, in T_h the employer benefits from offering nothing if the worker is A1 (selfish-rational) but benefits more from offering the high wage contract if the worker is A1 or A2. In T_b and T_m the safe option for the employer is to offer the incentive contract, since it incentivizes even selfish-rational A3 players to provide high effort and hence it protects the employer against shirking. But in T_b the employer makes a large gain of 50 if there is a high chance that the worker is A1 and will provide high effort anyway. In contrast, in T_m the fixed high wage is always dominated by the malus contract for a selfish-rational A3 employer.

For A2 (conditional cooperator) employers, the optimal offers are identical to A3 archetypes in T_l and T_h. But in T_b and T_m, offering the incentivized contract is preferred since it directly implements conditional cooperation. Turning to A1 employers, let us first recall that we have defined the A1 archetype as choosing the cooperative action unconditionally. In the case of employers, this means offering a high wage – an action that is possible in treatments T_h and T_m. In T_b the employer cannot guarantee a high wage but can nonetheless pay a high wage to those workers who provide high effort. For this reason, offering the T_b is a dominating choice for an A1 (intrinsically motivated) employer. The situation is different in T_l, since the employer cannot achieve the payment of a high wage – and may thus be indifferent between offering a low wage contract or nothing. Not making an offer may, however, dominate if the A1 employer internalizes the loss that an intrinsically motivated worker would incur from accepting a low wage contract.[15] Making no offer in T_b and T_m is not compatible with either of the subject types we have considered here – and is thus not rationalized here.

5 Empirical results

We now turn to the empirical analysis. We present results broken down into five distinct samples: the US and India samples recruited on Amazon Mechanical Turk (MTurk); and the US, India, and African samples recruited via Facebook (FB). We document what is common and what is different across the different samples. Differences may arise for a variety of reasons – e.g., social norms, cultural attitudes, stereotypes, or beliefs about the choices made by other

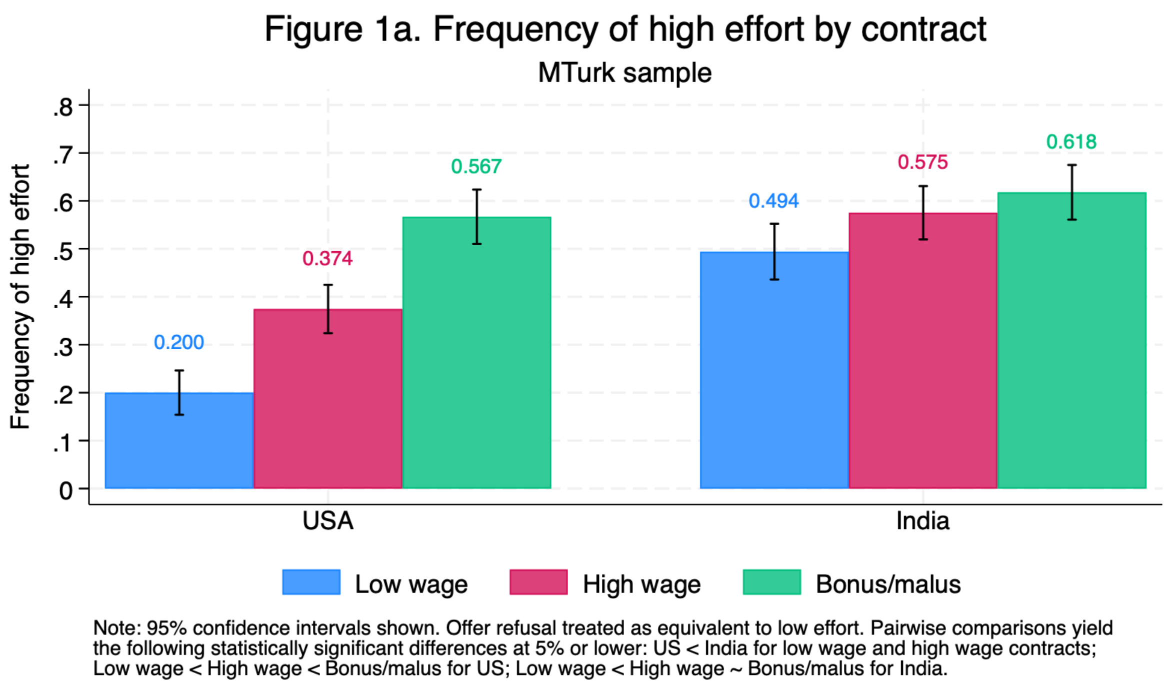
[15]See Buchmann et al. (2022) for an empirical design that investigates this type of behavior, which the authors call 'paternalistic discrimination'.

players. Our experiment is not designed to identify *which* of these differences predicts how subjects behave in the experiment. We simply aim to document whether subjects recruited from different populations across the world behave in similar way when placed in employer-employee contractual situations.

5.1 Effort choice by workers

We first examine the effort choice made by subjects assigned the role of worker, conditional on receiving one of four contract offers: low wage; high wage; and a full-commitment bonus or malus contract (which we combine here). Contracts that allow employers ex-post discretion are discussed later since they require workers to form expectations about the employer's likelihood of reneging.

Observed effort levels are summarized in Figures 1a for the MTurk sample and in Figure 1b for the Facebook sample. Rejecting the offer is assimilated to low effort for the purpose of the Figures.[16] BonusWD and malusWD are omitted from the Figures because, in these contracts, the worker's choice depends on expectations of employer ex post discretion. In these figures as in most subsequent figures, 95% confidence intervals are represented by a black line on top of each bar. Pairwise t-test results are summarized in words at the bottom of each figure.

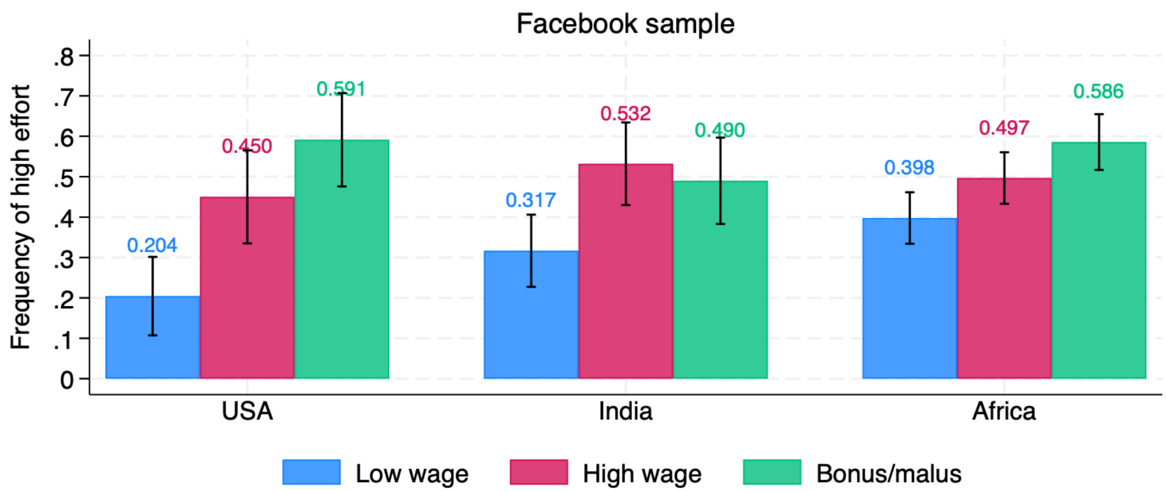


In Figure 1a we note a sharp difference among US MTurk subjects in the chosen effort depending on contract terms. The difference in effort choices is much less marked among Indian MTurk subjects. The blue bar measures the fraction of subjects in archetype A1 – i.e., they accept the contract and provide high effort in the low wage contract (see Table 4). They represent

[16] Appendix Figures A1a&b and A2a&b present separately the offer acceptance levels and the effort levels conditional on acceptance.

20% of subjects in the US sample compared to 49.4% in the India sample (Table 6). The gap between the red and blue lines gives the fraction of actions in archetype A2. It is 17.4% in the US and 8.1% in the India sample. The difference between the green bar and the red bar is the proportion of archetype A3 play, which is about 19.2% in the US and about 4.3% in the India sample. These are large differences. Finally, the proportion of archetype 4 play is given by the fraction of subjects who do not provide high effort in the bonus/malus contracts. These are 43.3% in the US MTurk sample and 38.2% in the India MTurk sample. These are large proportions, suggesting that much of the observed worker choices in the experiment fall outside simple behavioral models used in labor economics. We cannot, however, rule out the possibility that these high proportions are driven by inattention. Introducing an in-person real effort task would alleviate this concern but doing so was not possible due to logistical and financial constraints.

Figure 1b. Frequency of high effort by contract



Note: 95% confidence intervals shown. Offer refusal treated as equivalent to low effort.
T-test pairwise comparisons yields the following statistically significant differences at 5% or lower:
Low wage < High wage < Bonus/malus for US and Africa;
Low wage < High wage ~ Bonus/malus for India.
For low wage: US < India ~ Africa; For high wage: India ~ Africa;
Other pairwise comparisons are not statistically significant.

Comparable results for the Facebook sample are shown in Figure 1b, where we have divided the sample into three geographical regions: US; India; and Africa (i.e., Morocco, Senegal, Malawi, Kenya, and South Africa). The proportion of A1 play in the US Facebook sample is 20.4% – virtually identical to that of the US MTurk sample. This proportion rises to 31.7% in India, and 39.8% in the Africa sample. Corresponding proportions of A2, A3, and A4 play are given in Table 6. In a couple of cases, we find small negative numbers, a result potentially attributable to sampling error.

What is clear from the Figures and Table 6 is that the proportion of A1 play is much lower in the two US samples than in samples from the other two world regions covered by our study. We also find that the proportion of A3 play is higher on average in the US sample than in the low- and middle-income samples included in our study. What could account for these differences is a topic we revisit in the last Section of the paper. We also find that between a third and half of the subjects' choices follow archetype A4, without any systematic correlation with sample origin.

Table 6. Breakdown of worker types inferred from offer acceptance and effort

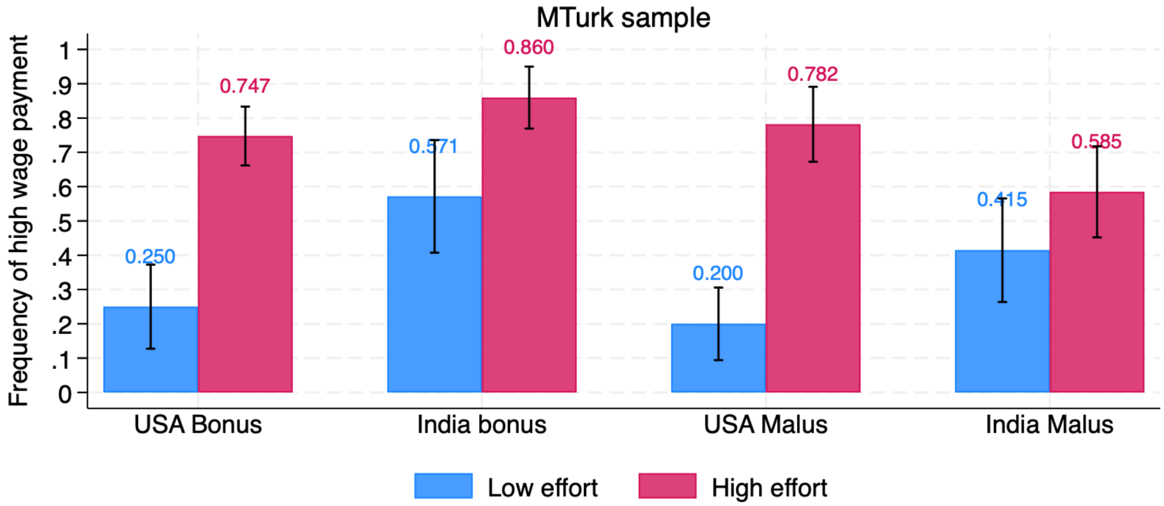
Archetype:	A1	A2	A3	A4
Sample:	Intrinsically motivated	Conditional reciprocator	Selfish rational	Unconditionally unmotivated
US MTurk	20.0%	17.4%	19.2%	43.3%
India MTurk	49.4%	8.1%	4.3%	38.2%
US FB	20.4%	24.6%	14.1%	40.8%
India FB	31.7%	21.5%	(-4.2%)	51.0%
Africa FB	39.8%	9.9%	8.9%	41.4%

Notes: MTurk and FB refer to subjects recruited via Amazon Mechanical Turk and Facebook, respectively.

5.2 Reneging by employers

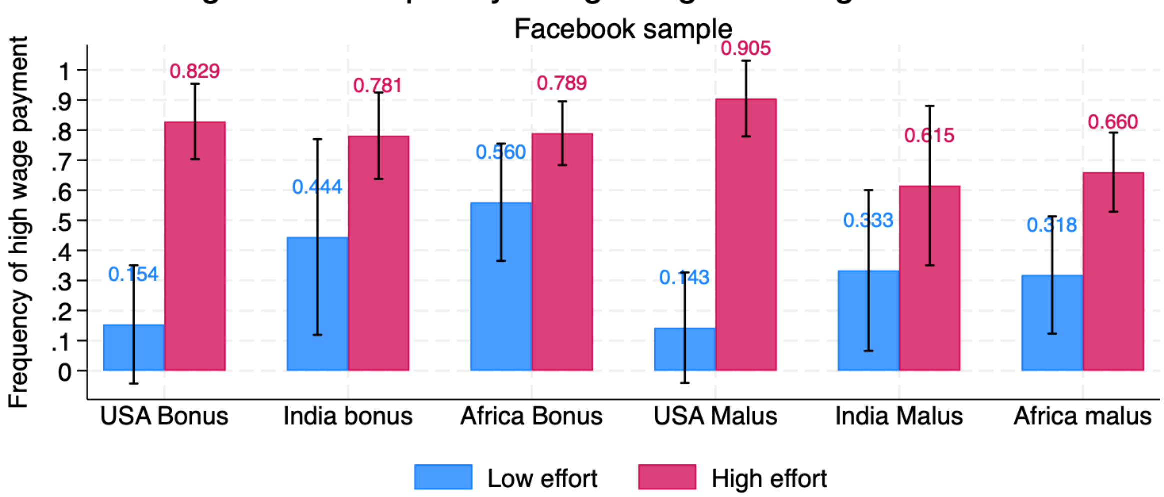
Next, we examine the reneging choices made by employers in the malus and bonus contracts with a reneging option. Since these are final choices, expectations do not affect them. Figure 2a shows the wage paid by employers in the MTurk sample as a function of worker effort in the two contracts offering ex post reneging. The blue bars show the proportion of employers who renege by paying a high wage when effort is low. These can be interpreted as the proportion of A1 play. We again find a higher proportion of A1 play in the India MTurk sample – 57.1% for Bonus WD and 41.5% for Malus WD contracts – compared to 25% and 20% in the US MTurk sample, respectively (see Table 7). A3 and A4 archetypes pay a low wage after high effort. For the Bonus WD contract, the proportion of A3 and A4 play is higher in the US sample (25.3%) than in the India sample (14.0%). We nonetheless find different proportions for the Malus WD contract: 21.2% of US subjects renege on the contract to pay a low wage after high effort, compared to 41.5% in the India MTurk sample. The rest of the choices follow archetype A2, which constitute the bulk of the employers in the US MTurk sample, but a smaller proportion in the India MTurk sample.

Figure 2a. Frequency of high wage in renegable contracts



Note: 95% confidence intervals shown. Following the contract means blue bar=0 and red bar=1.
T-test pairwise comparisons yields the following statistically significant differences at 5% or lower:
Low vs high effort: significant for USA (both contracts) and India (bonus contract only);
Bonus vs malus contract: significant difference for India high effort only;
US vs India: significant difference for all at 10% level or lower.

Figure 2b. Frequency of high wage in renegable contracts



Note: 95% confidence intervals shown. Following the contract means blue bar=0 and red bar=1.
T-test pairwise comparisons yields the following statistically significant differences at 5% or lower:
Low vs high effort: significant for all locations except malus-renege in India (small sample);
Bonus vs malus contract: difference never significant at the 10% level.
Countries: Africa pays high wage for low effort more often than USA in bonus contract,
but pays low wage more often for high effort in malus contract.

Similar general patterns are observed in the Facebook sample (Figure 2b and Table 7). The inferred proportion of A1 play in the US Facebook sample is 15.4% and 14.3% for the Bonus WD and Malus WD contracts, respectively – compared to 44.4% and 33.3% in the India Facebook sample and 56.0% and 31.8% in the Africa sample, respectively. The proportion of A3 or A4 play is slightly smaller in the US Facebook sample for the Bonus WD contract, but much smaller for the Malus WD contract, an issue we revisit below. This again leaves a much larger proportion of A2 play in the US sample than in the India and African Facebook samples.

Table 7. Breakdown of employer types inferred from reneging behavior

Archetype:	A1		A2		A3	
	Intrinsically motivated		Conditional cooperator		Selfish rational	
Sample:	BonusWD	MalusWD	BonusWD	MalusWD	BonusWD	MalusWD
US MTurk	25.0%	20.0%	49.7%	58.8%	25.3%	21.8%
India MTurk	57.1%	41.5%	28.9%	17.0%	14.0%	41.5%
US FB	15.4%	14.3%	67.5%	76.2%	17.1%	9.5%
India FB	44.4%	33.3%	33.7%	28.2%	21.9%	38.5%
Africa FB	56.0%	31.8%	22.9%	34.2%	21.1%	34.0%

Notes: MTurk and FB refer to subjects recruited via Amazon Mechanical Turk and Facebook, respectively. A3 may include some A4 archetypes as well.

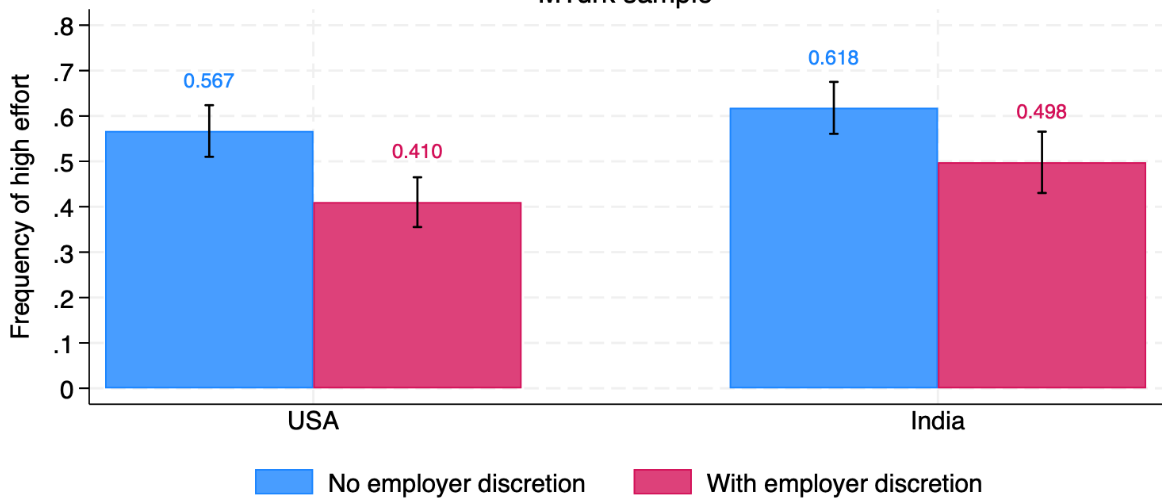
The difference between the behavior of employers in the Malus WD and Bonus WD contracts is a priori puzzling. But it could be due to a different self-selection in these contracts by employers. In the Malus WD treatment, the alternative contract that the employer can offer is a high fixed wage, something that would appeal to A1 archetypes – who like to benefit the other player. In contrast, in the Bonus WD treatment, the alternative is a low fixed wage, a contract that would appeal to A3 archetypes hoping to extract a large surplus from A1 workers. We revisit this point below when we examine the choice of offers made by employers.

5.3 Effort choice in the bonus and malus contracts with a reneging option

We now turn to strategic choices, that is, decisions that potentially depend on how subjects expect their partner to act in the one-shot game. We first examine the effort choice of workers in the two contracts with employer ex-post discretion. In Figure 3a, we compare the proportion of workers who accept the contract and choose high effort among all workers offered a bonus or malus contract. We immediately see that high effort falls in both the US and India MTurk samples. In the US sample, the proportion of high effort decrease from 56.7% go 41.0% – a fall of 15.7 percentage points. The fall is 12 percentage points in the India MTurk sample. In both cases these differences are statistically significant. A similar picture is given in Figure 3b for the Facebook sample – with falls of 12.5, 0.1, and 14.6 percentage points in the US, India, and Africa Facebook samples. Apart from the India Facebook figure, all indicate a large fall in efficiency when we remove the guarantee that workers who provide high effort cannot be cheated ex post by employers.

Figure 3a. Frequency of high effort in bonus/malus contracts

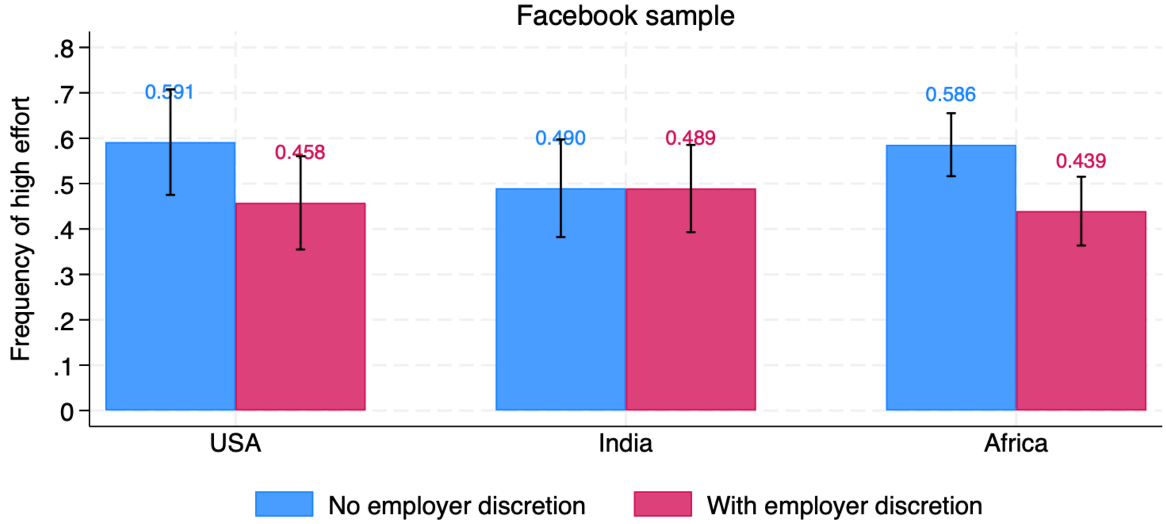
MTurk sample



Note: 95% confidence intervals shown. Offer refusal treated as equivalent to low effort.
T-test pairwise comparisons yields the following statistically significant differences at 5% or lower:
higher effort in India than US when employer has discretion to renege;
reduction in effort when employer has discretion is significant in both countries.

Figure 3b. Frequency of high effort in bonus/malus contracts

Facebook sample



Note: 95% confidence intervals shown. Offer refusal treated as equivalent to low effort.
T-test pairwise comparisons yields the following statistically significant differences at 5% or lower:
no significant difference between countries in any of the bars;
reduction in effort when employer has discretion is significant in the US and Africa.

While the fall in effort that is observed in all study samples is large, it does not imply that workers expect all employers to behave in a selfish way. Indeed, if they did, we would expect a much lower proportion of workers choosing high effort, e.g., no more than those who were classified as A1 in Table 6. To illustrate, we see 41% of US MTurk workers choose high effort in discretionary contracts compared with 20% of subjects being identified in Table 6 as A1 for that sample.

What can account for the 'excess effort' in renege contracts? We believe the response lies in the fact that employers can deviate from the contract in two ways: by paying a low wage after high effort; or by paying a high wage after low effort. The former is predicted for A3 archetypes (e.g., selfish-rational employers); the latter is predicted for A1 archetypes (e.g., intrinsically motivated employers) (see Table 5). In *both* cases the incentive effect on worker effort is *negative*: in the first case the worker expects no reward from incurring the cost of high effort, while in the second the worker does not need to incur the cost of high effort to receive a high wage. As shown in Table 7, both types of employer behaviors are present in our study, in different proportions across samples. This reduces the loss that a worker may rationally expect to incur by choosing high effort. If all employers followed archetype A3, this loss would be 10. In the experiment, the expected loss is much smaller: it varies from a high of 6.3 on average in the US samples to a low 2.7 and 2.9 for the India and Africa samples, respectively. This, combined with the presence of A1 workers, can account for the relatively high supply of effort even when the contract terms are not externally enforced.

5.4 Robustness checks

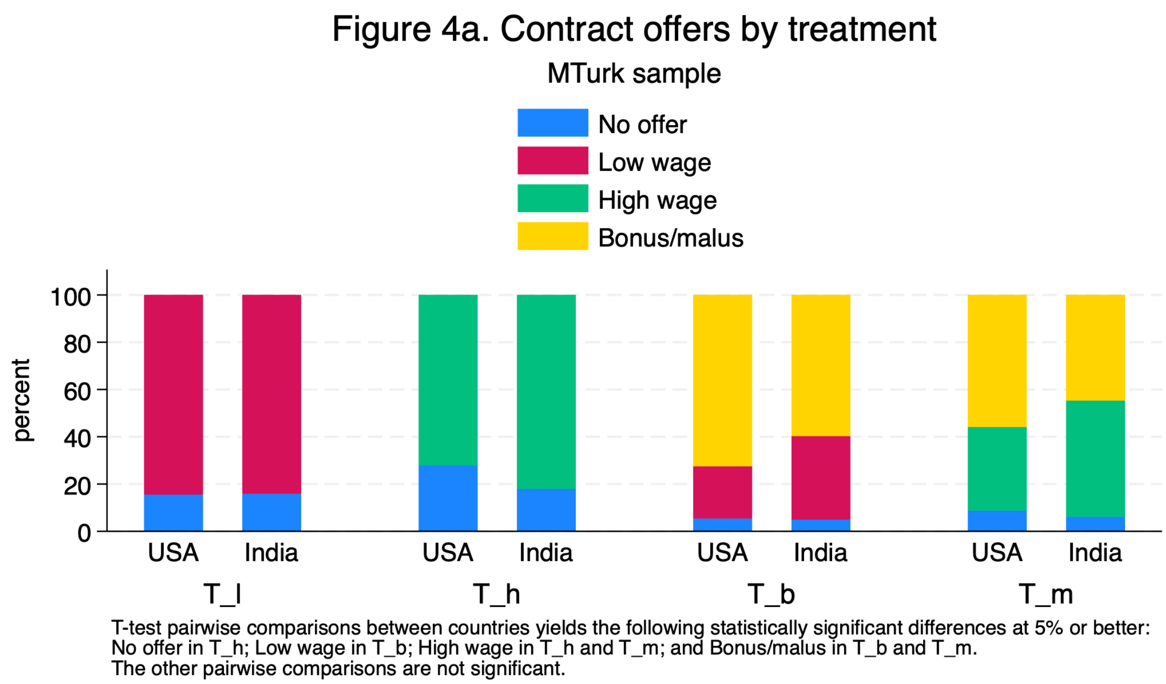
Before proceeding with the rest of our analysis, we check the reliability of that the main results reported so far. We first consider the possibility of a bias induced by the removal of those observations we concluded were dubious. This only affects the Facebook sample, since there were no dubious participants in the MTurk sample. To demonstrate that the observation removal process is not driving our results, we re-estimated Figures 1b, 2b and 3b using the full sample. The results, presented in the Appendix, show no qualitative change in our findings: while there may be a slight move towards random 50/50 play in some cases, we find no evidence that our results are an artifact of dropping these observations.

Second, we can also rule out that our findings are a result of pure random play by either the full sample or the restricted sample without dubious observations. To see this, we use the fact that each bar in Figures 1a to 3b summarizes a dichotomous choice made by participants. If these choices had been made at random, we should observe average play to be around 50% for each of the two choices – which means all the bars in the six Figures would not be statistically different from 50%. Using the 95% confidence interval of each bar shown in each of the Figures, we immediately see that most bars (i.e., averages) are statistically different from 50%: the 0.5 horizontal line falls outside the confidence interval. This is true across samples and regions. Furthermore, there are ample – and often significant – within-region differences in play depending on the treatment, confirming that subjects are not simply playing at random. This is true for both the MTurk and Facebook sample, and for all three regions of interest, the US, India, and Africa.

Finally, we test whether pooling observations across African countries is warranted in our case. To this effect, we regress on country dummies the choices made by African participants in each of the Africa-related bars in Figures 1b, 2b, and 3b . The p-value of the F-test of significance of the regression are shown in Appendix Table A1. Only one of the nine regressions yields an F-test that is marginally significant at the 10% level. From this we conclude that, for the purpose of this paper, pooling across the Africa sample is acceptable in the sense that the choices made by participants from the five countries are not statistically different from each other.

5.5 Contract offer

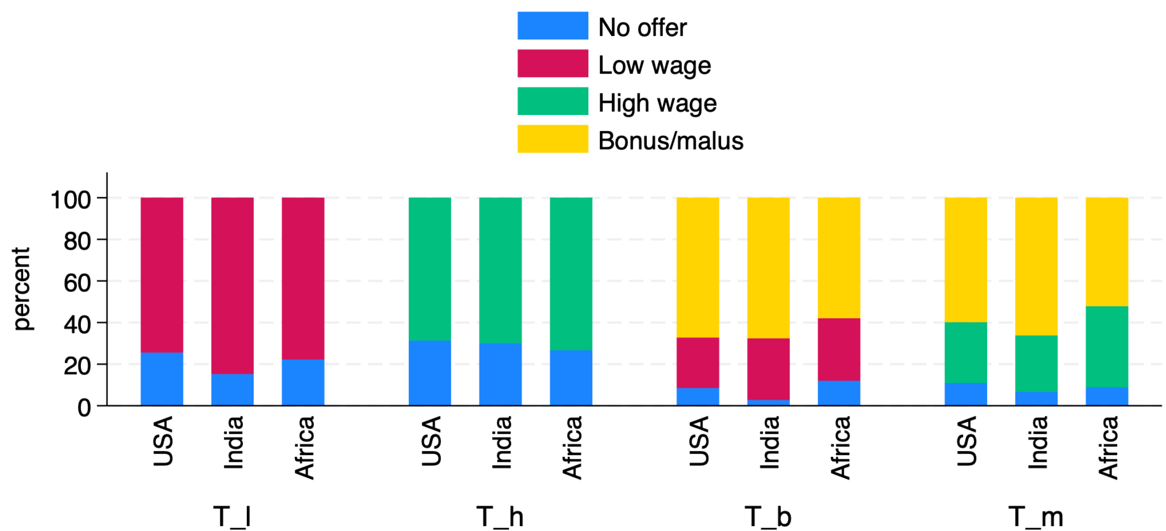
Next, we turn to the choice of contract offer made by employers, depending on treatment. We show in Figures 4a a breakdown of the offers made by employers in the MTurk sample across the four treatments T_l to T_m. A similar breakdown is provided in Figure 4b for the Facebook sample. In the low wage treatment, most employers make an offer. This is consistent with our theoretical predictions (and the findings from Figure 1a, 1b, and Table 6) that a large proportion of workers provide high effort in the low wage contract, particularly in the India and Africa samples. We indeed observe in Figure 4b a slightly higher frequency of no-offer in the US sample than in the India and (to a lesser extent) the Africa sample.



In treatment T_h, we find that a larger proportion of employers opts for no offer, with a slightly lower proportion in the India and US samples. This implies that the majority of employers in all samples take the risk of offering a fixed wage in the hope that this kindness will be reciprocated by workers – as is indeed the case in Figures 1a, 1b and Table 6. In other words, most employers expect that a sufficient proportion of workers falls outside the A3 and A4 archetypes to warrant taking the risk. This is something that our experimental design implicitly encourages since the loss to the employer (relative to no offer) is 10 if the worker chooses low effort while the gain is 20 if the worker chooses high effort. This means that even a risk-neutral, selfish-rational employer would offer a fixed high wage if the proportion of high-effort workers exceeds one third. This proportion is exceeded in all samples (see Figures 1a and 1b). But it is exceeded more in the non-US samples, which may help explain why the frequency of no-offers is slightly lower outside the US samples. These differences, however, are small and not statistically significant.

Figure 4b. Contract offers by treatment

Facebook sample

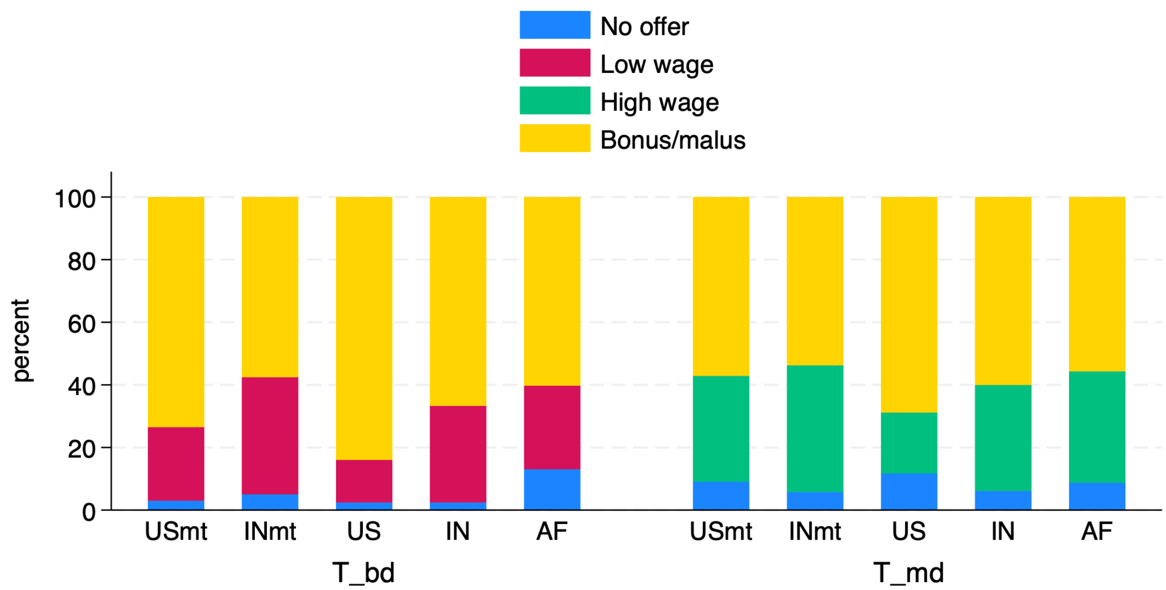


T-test pairwise comparisons between countries yields the following statistically significant differences at 5% or better: No offer in T_b and Bonus/malus offer for India vs Africa. All the other pairwise comparisons are not significant.

Turning to the two treatments T_b and T_m with incentivized contracts, we find a small fraction of choices that fall in the A4 archetype in all samples, but particularly in the India samples. A large fraction of employers opt for incentivized contracts. This is especially true in the T_b treatment where the alternative is a low fixed wage, which is less likely to elicit high effort than an incentivized contract. We nonetheless find a sizable fraction of employers offering a low fixed wage instead. As argued above, this choice suggests these employers are attempting to take advantage of intrinsically motivated A1 workers, the proportion of which is higher in the India and Africa samples than in the US samples (see Figures 1a and 1b). We also find that it is more common for employers to offer the fixed high wage in T_m than in T_b. This behavioral pattern is difficult to reconcile with A2 or A3 archetypes on the part of employers, since both motives militate in favor of the incentive contract. It is more in line with the intrinsic motivation A1 archetype to pay workers well – and perhaps expect high effort in return. This seems to be particularly true in the India MTurk sample and the Africa Facebook sample.

Figure 5 summarizes, for the MTurk and Facebook samples combined, the offers made in the two treatments T_{bd} and T_{md} in which the employer has full ex-post discretion to pay the low or high wage in the bonus and malus contracts. In these treatments, the strategic choices made by employers are more complex since they have not only to anticipate the reaction of workers, but also their own future decision to exert ex-post discretion. A3 (selfish rational) employers may be tempted to offer bonus or malus contracts to motivate workers, only to renege later.

Figure 5. Offers in the treatments with ex post discretion



Note: USmt & INmt=US & India MTurk samples; US, IN, AF=US, India, and Africa Facebook samples

We see from Figure 5 that the offers made by employers in T_{bd} and T_{md} are fairly similar to those they make in treatments T_b and T_m.[17] In principle, this could arise because employers do not intend to renege. But we already know from Section 3.2 that a large fraction of employers does exert ex-post discretion – some to pay less than what is due and some to pay more. An alternative interpretation for our findings could thus be that these two motivations more or less balance each other in our sample.

5.6 Efficiency and equity

We complete this Section by examining the joint efficiency achieved by subjects in the different treatments. Joint efficiency is normalized to 0 in the absence of trade – e.g., either if no offer is made by the employer, or if the offer made is rejected by the worker. We set joint efficiency to be 10 with low effort, which is the marginal product of labor paid to the worker in the form of a wage. With high effort, efficiency rises to 35, which is the marginal product of high effort (40) minus its cost (5).[18] We examine which treatment achieves the highest joint efficiency and whether efficiency varies across country samples.

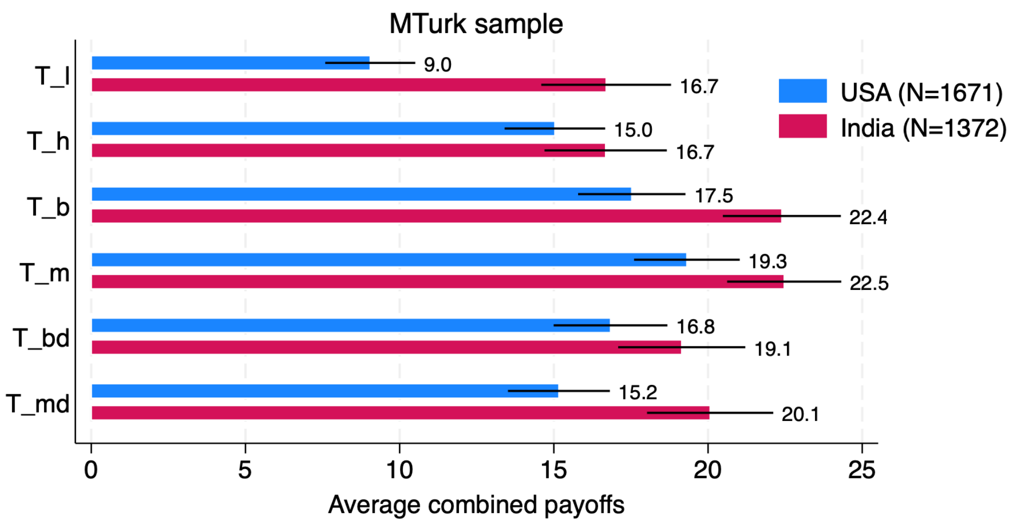
Results are presented in Figure 6a for the MTurk sample. The country shown is that of the worker. More efficiency is achieved on average with the two incentivized contracts, T_b and T_m.

[17]We nonetheless observe a slight increase in the proportion of malus contracts offered by employers in the US Facebook sample and, to a smaller extend, in the Africa sample. A similar movement is not observed in the MTurk sample and, in the India sample, we observe no such changes.

[18]To verify that our findings do not depend on the numerical assumptions we have made, we repeat the analysis using an index equal to 0 for no trade, 1 for low effort, and 2 for high effort. Virtually identical results are obtained. They are available from the authors upon request.

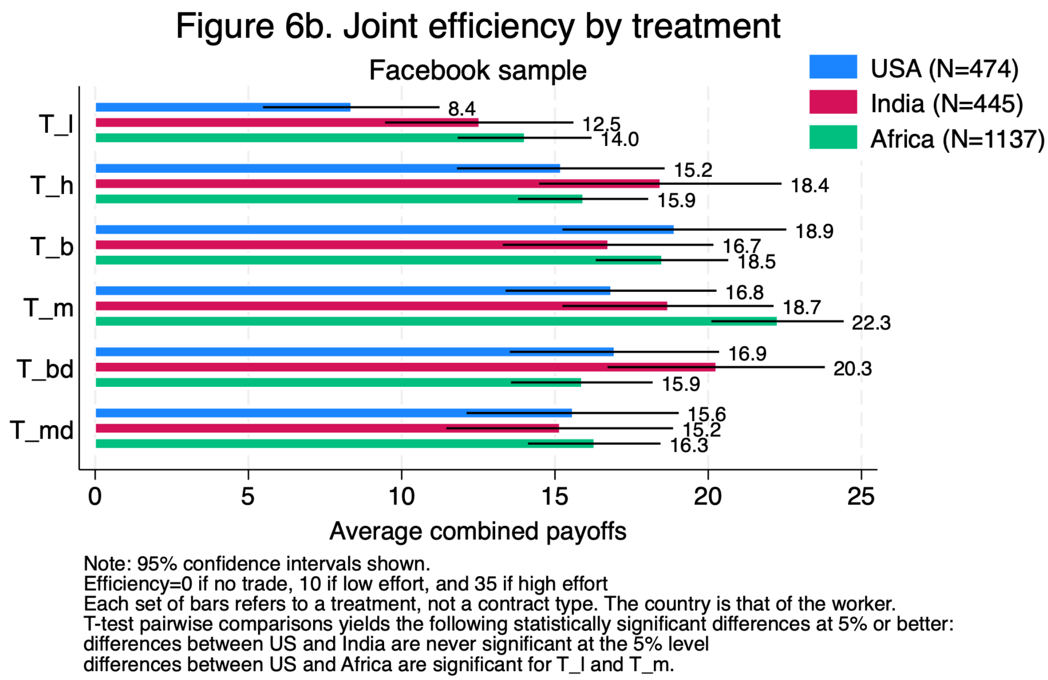
Allowing for ex post discretion by the employer reduces efficiency, as could be expected, but the difference is not large. Efficiency is lower in the fixed high wage contract and lowest in the fixed low wage contract. These results are in line with predictions from standard economic theory. But the differences are much smaller – and the level of efficiency much higher – than would be predicted by assuming selfish-rational behavior and rational expectations on the part of both employers and workers. Furthermore, all treatments – including the incentivized contracts – remain well below the full efficiency of 35. We also note that efficiency is, on average, lower in the US sample – significantly so in four of the six treatments. This is most noticeable in the low wage treatment because of a higher likelihood of rejection and a lower choice of effort among US MTurk subjects assigned the role of worker. It is also in that sample that incentivized contracts yield the highest efficiency gain relative to a fixed low wage.

Figure 6a. Joint efficiency by treatment



Note: 95% confidence intervals shown.
Efficiency=0 if no trade, 10 if low effort, and 35 if high effort
Each set of bars refers to a treatment, not a contract type. The country is that of the worker.
T-test pairwise comparisons yields the following statistically significant differences at 5% or better:
differences between US and India are significant in all treatments except T_h and T_bd.

Similar but less significant findings are observed in the Facebook sample (Figure 6b): the incentivized contracts achieve a higher level of efficiency on average, but the efficiency gain is small relative to either fixed high wage or contracts that allow ex post reneging by the employer. The fixed low wage contract is, however, much less efficient, especially in the US sample relative to the India and Africa samples. As a result of this, average efficiency across treatments is higher in the India and Africa than in the US sample.



Turning to equity, we compare worker and employer payoffs averaged over all treatments. Results are summarized in Table 8. As expected from Figures 6a and 6b, the total of workers' and employers' payoffs is, on average, lower in the US MTurk and Facebook samples than in the India and Africa samples. This shortfall is entirely borne by employers, whose payoffs are lower in the US samples, while worker payoffs are slightly higher. This difference in relative payoffs is driven by the distribution of behavioral types in the different country samples, notably the larger proportion of A3 archetype workers in the US samples and the larger fraction of A1 archetype workers in the India and Africa samples.

Table 8. Mean payoffs

	Worker	Employer	Total
US	12.57	26.10	35.44
India	11.94	29.76	38.93
Africa	12.15	28.18	37.14

Note: Means have been corrected for possible imbalance in treatments across samples. The MTurk and Facebook samples are combined for India and the US.

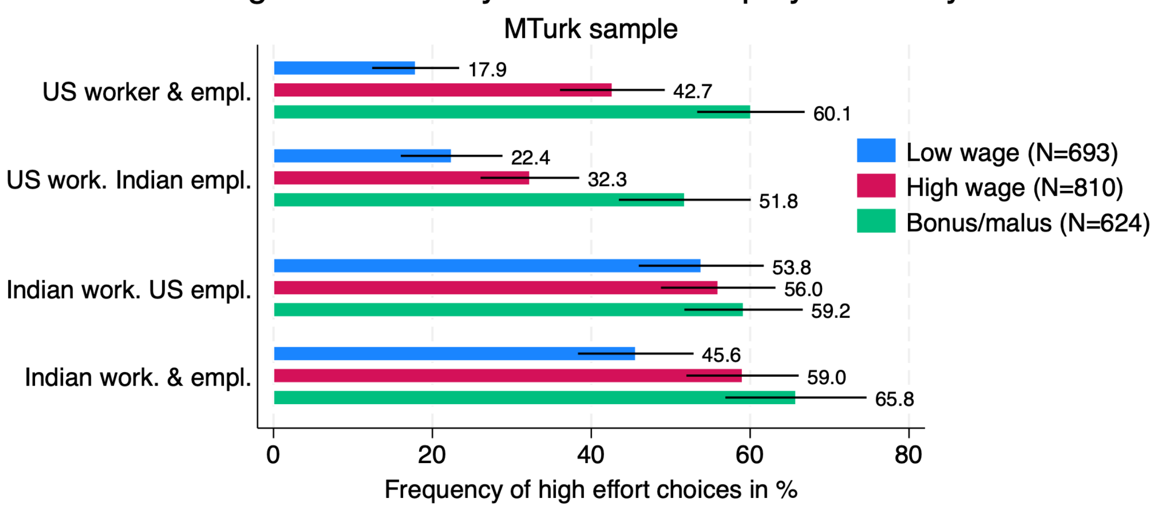
6 Cross-country prejudice and stereotypes

We have seen that the sampled populations from different parts of the word behave slightly differently. Here we examine whether this may be driven by prejudice and stereotypes. This analysis is conducted with the MTurk sample only. Given the results we obtained in the MTurk sample and the added complexity of running an online experiment involving subjects in totally different time zones, the pairing of subjects from different countries was not replicated with the Facebook sample.

For the purpose this analysis, we define *prejudice* as being less positively inclined towards someone of a different culture or origin, which in the context of our experiment means being less willing to cooperate with them, either conditionally or unconditionally. This idea is similar to in-group favoritism documented by Banuri et al. (2022). To see whether behavior in our experiment may be affected by prejudice, we examine whether workers provide effort differentially when the employer is from their own country or from somewhere else. To the extent that effort choice follows conditional cooperation or intrinsic motivation, we expect prejudice to manifest itself as lower effort when the employer is from another country.

Figure 7 depicts the effort choices made by MTurk subjects. Regarding US workers (the upper-half of the Figure), we see that, if anything, intrinsically motivated effort is higher for subjects matched with an Indian employer – the blue bar is slightly higher. We do, however, notice a dramatic fall in conditional cooperation, captured by the difference between the red bar and the blue bar: US workers are much less likely to provide high effort in a fixed high wage contract with an Indian employer, compared to a US employer. This difference carries over to the incentivized contracts, implying that a larger proportion of the effort choices made by US workers are not rationalized by our three behavior categories: many US workers provide low effort in incentivized contracts, even though it is against their interest.

Figure 7. Effort by worker and employer country

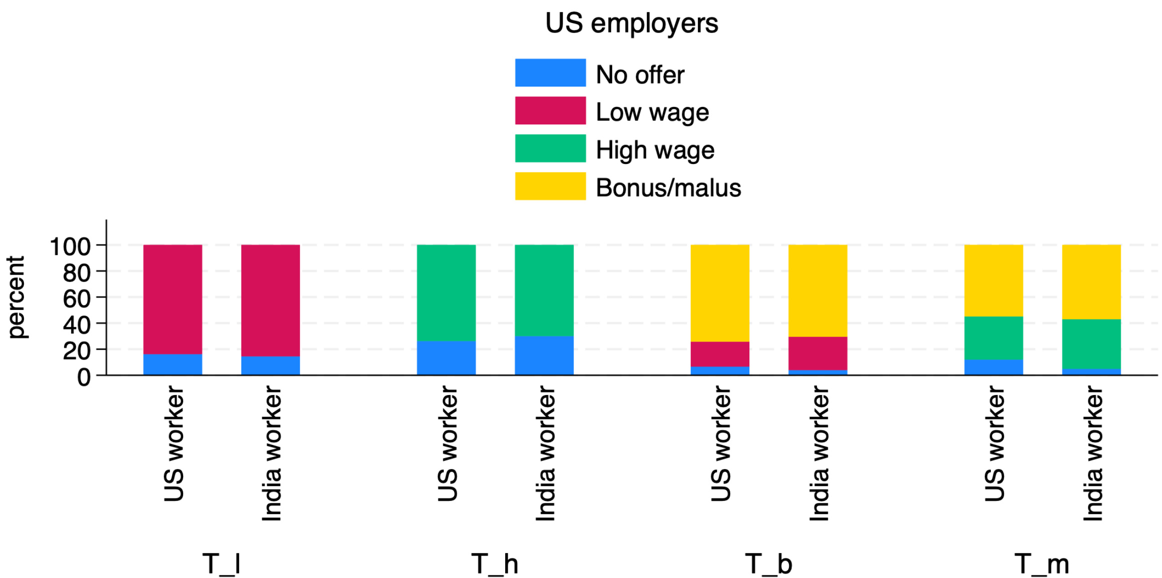


Note: 95% confidence intervals shown. Offer refusal treated as equivalent to low effort.
T-test pairwise comparisons yields the following statistically significant differences at 5% or better:
differences between US and Indian workers are always significant except for
the bonus/malus contract with a US employer.
no significant difference in effort with employer country for Indian workers;
US workers provide significantly less effort to Indian than US employers in the high wage contract.

Turning to Indian workers (the lower half of the Figure), we observe a somewhat similar pattern: the blue bar is higher when the worker is matched with a US employer, indicating higher intrinsic motivation. But conditional cooperation (the difference between the red and the blue bars) falls when matched with a US employer, and high effort also falls in the bonus and malus contracts, even though it is against the interest of the Indian worker.[19] Taken together, these patterns suggest that the differences in effort choice we observe between the US and Indian worker samples are not simply driven by prejudice – there also exist differences in the proportion of conditional cooperators, selfish-rational, and intrinsically motivated subjects between the two sample populations – a topic we revisit in the next Section.

The second hypothesis we wish to investigate is whether our subjects have *cultural stereotypes* that may affect their behavior in a cross-cultural context.[20] For the purpose of this analysis, we define stereotypes as having different behavioral expectations with respect to people of a different origin. To investigate whether our subjects have cultural stereotypes regarding effort choice, we examine whether contract offers differ systematically by the country of the worker. Cultural stereotypes could be positive or negative. Positive stereotypes mean that subjects expect workers from the other country to be more likely to choose high effort even when unincentivized – and conversely for negative stereotypes. If subjects from India, say, believe that workers from the US are more likely to choose high effort when unincentivized, we should observe more offers of fixed wage contracts by India employers to US workers. If, on the other hand, India subjects believe that US workers need to be incentivized to provide high effort, we should observe a higher proportion of bonus and malus contract offers in the T_b and T_m treatments.

Figure 8a. Contract offers by treatment



Note: None of the pairwise differences by worker country are statistically significant at the 5% level except one (no offer in T_m).

Figures 8a and 8b presents a breakdown of contractual offers by the country of the worker

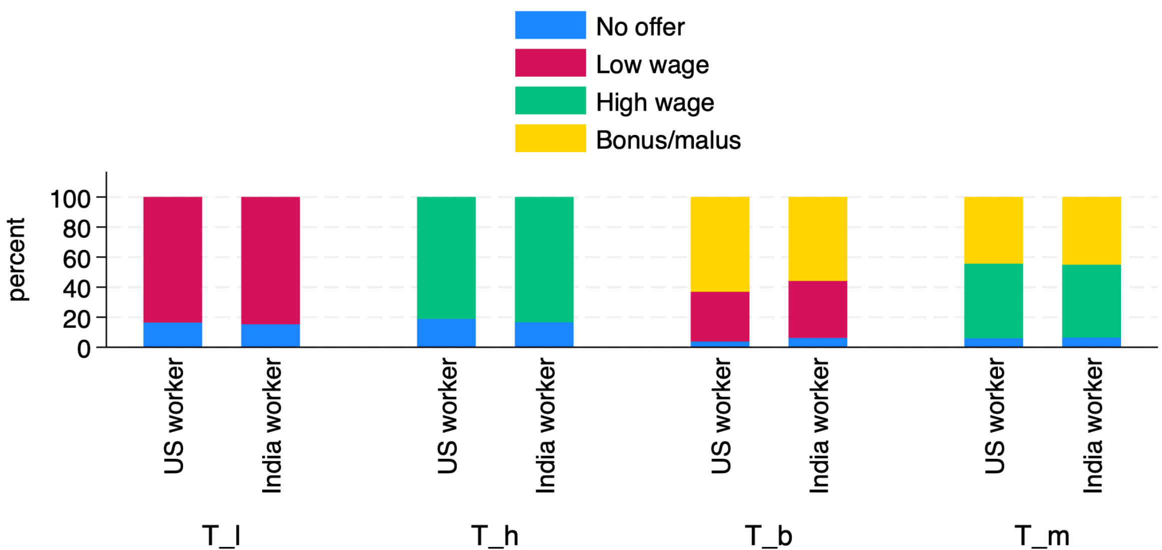
[19]We do, however, observe that effort levels in fixed wage contracts remain significantly higher among Indian workers matched with US employers than among US workers matched with someone from the US. Similarly, Indian workers supply much more effort than US workers when matched with an Indian employer.

[20]Unlike Banuri et al. (2022), our games offer little opportunity for employers to favor in-group workers. This make favoritism towards workers unlikely in our setting, which is why we ignore it here for ease of exposition.

in treatments T_l, T_h, T_b, and T_m. Figure 8a shows the behavior of subjects in the US sample; Figure 8b does the same thing for subjects in the India sample. In Figure 8a, the only evidence of stereotypes is in treatment T_h, where a slightly higher proportion of US subjects do not make a fixed high wage offer to India workers. We do not, however, find systematic evidence of increased reliance on incentive contracts: in T_b, a slightly smaller proportion of Indian workers are offered a bonus contract, while in T_m, the proportion is slightly higher. Things are similar in Figure 8b, where we see even fewer differences between the offers made to Indian and US workers. Taken together, these findings do not constitute strong evidence that, for the MTurk sample, stereotypes affect the offer of contracts to online workers from another culture. In both Figures, none of the differences between regions is statistically significant. This does not, however, imply that stereotypes could not be prevalent in other contexts.

Figure 8b. Contract offers by treatment

Indian employers



Note: None of the pairwise differences by worker country are statistically significant at the 5% level.

7 Predicting types across samples

So far, we have documented differences in behavior and inferred type across different samples. In this Section, we investigate the extent to which differences across samples can be accounted for by differences in the observable characteristics of sampled individuals. We have already documented in Table 3 the differences in average characteristics that we observe across samples. In this Section we examine whether these differences in characteristics can predict observed differences in the prevalence of archetypes across samples, or whether there remain differences in behavior that cannot be explained or predicted by differences in characteristics. We focus on the behavior of subjects when they are assigned the role of workers because it maps directly into types.

Our research question is best formalized through the following thought experiment. Imagine two distinct populations A and B who do not interact with each other. When an agent i of population A meets another agent j of the same population, i is able to predict j's archetype from j's observed characteristics based on the observable correlation between archetypes and characteristics in population A. For instance, if educated agents in population A are more likely to behave as conditional cooperators than uneducated agents, then i predicts a higher likelihood that j is a conditional cooperator if j is educated. Agents in population B do the same but based on the correlation between archetypes and characteristics in *their* own population.

If populations A and B share the same correlation between characteristics and archetypes, then agents in population A can predict the archetype of agents in population B as well as agents from that population. If this is the case, the differences in characteristics between the two populations predict the differences in behavior across them. To continue with our earlier example, if there are more educated agents in population A, agents in that population will, on average, behave more like conditional cooperators. If, however, these correlations differ across the two populations, there is an element of mutual surprise: agents in A cannot accurately predict the behavior of agents in B based on their characteristics and, similarly, agents in B cannot accurately predict behavior in A. This inability to accurately predict behavior across populations captures what many people mean by 'cultural differences' across populations – i.e., that *members from another population behave in a way that is different from what observers expect based on behavior patterns in their own population.* This definition is the one we adopt in the rest of this Section where we examine the data for evidence of such cultural differences.

Formally, let each subject in population A be of archetype θ^w with some probability p_A^w, where w takes one of four possible values – e.g., selfish-rational, conditional cooperator, intrinsically motivated, and non-rationalized. The frequency of archetype w in population A is denoted F_A^w, and similarly for F_B^w. In general $F_A^w \neq F_B^w$. We assume that each member i of population A forms correct expectations about the archetype θ_j of any other member j based on j's vector of observable characteristics X_j. This expectation can be obtained by estimating a predicting regression $\theta_j^w = P_A^w(X_j) + u_j^w$, where $P_A^w(X_j)$ is a particular predicting rule[21] for archetype θ^w and only observations on $j \in A$ are used to estimate $\hat{P_A^w}(X_j)$. We make the same assumptions for subjects in population B to define $\hat{P_B^w}(X_j)$ for each archetype w.

Now take an individual i from population A and ask this person to predict the archetype θ_k of an individual k with characteristics X_k in population B. Let this prediction be denoted as $Pr_A[\theta_k = \theta^w | X_k, k \in B]$. Since, by assumption, i has no behavioral information on population B, i forms expectation $Pr_A[\theta_k = \theta^w | X_k, k \in B]$ by predicting the archetype of an individual with characteristics X_k in his own population A. Similarly for an individual in population B asked

[21] E.g., a linear regression or a random forest projection.

to predict the archetype of individual m in population A. This gives the following relationship:

$$Pr_A[\theta_k = \theta^w | X_k, k \in B] = \hat{P}_A^w(X_k)$$
$$Pr_B[\theta_m = \theta^w | X_m, m \in A] = \hat{P}_B^w(X_m)$$

Let us now average $\hat{P}_A^w(X_k)$ over all the k individuals in population B and denote the result $\hat{P}_A^w(X_B)$. It is the probability that an individual in population A believes an average individual in population B is of archetype w. We do the same for all m individuals in population A to obtain $\hat{P}_B^w(X_A)$. With this formalism, we can now define the cultural difference D_{AB}^w between A and B as:

$$D_{AB}^w = \hat{P}_A^w(X_B) - \hat{P}_B^w(X_A)$$

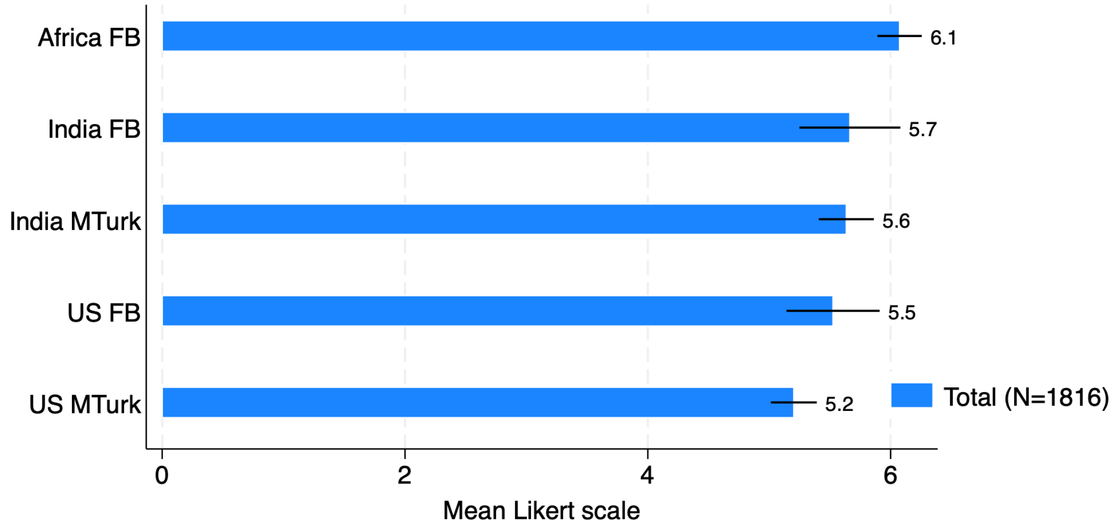
We want to test whether $D_{AB}^w = 0$ in our data by obtaining estimates of $\hat{P}_A^w(X_B)$ and $\hat{P}_B^w(X_A)$.

To do so, we start by providing survey evidence that the different populations in our study hold different expectations about behaviors associated with each of the three major archetypes we are interested in, namely, A1 (intrinsically motivated), A2 (conditional cooperator), and A3 (selfish rational). As mentioned in Section 2, all subjects in our study fill a questionnaire before being invited to the online experiment. In addition to a series of questions on age, gender, occupation, and education, respondents answer two sets of questions regarding labor markets. One set focuses on beliefs on whether workers in three countries – US, India, and South Africa – provide high effort whether their wage depends or not on their performance on the job. Believing that high effort will be provided without performance pay suggests expecting a high fraction of workers to be A1 (intrinsically motivated). The difference between the confidence with which respondents expect high effort with and without performance pay proxies for expecting workers to be A2 or A3.

Responses are summarized in Figure 9a. We see that Africa subjects and, to a lesser extent, Indian subjects are more likely than US subjects to expect workers to provide effort without the need for incentives. This is especially true for subjects recruited on MTurk. From Figure 9b, we see that US MTurk subjects are also more likely than African FB subjects to believe that offering work incentives induces more effort. These differences could be correlated with individualistic ethics (Bazzi et al. 2020; Enke 2019; Adams et al. 2019), but also with differences reported in Table 6.

Figure 9a. Beliefs in worker productivity

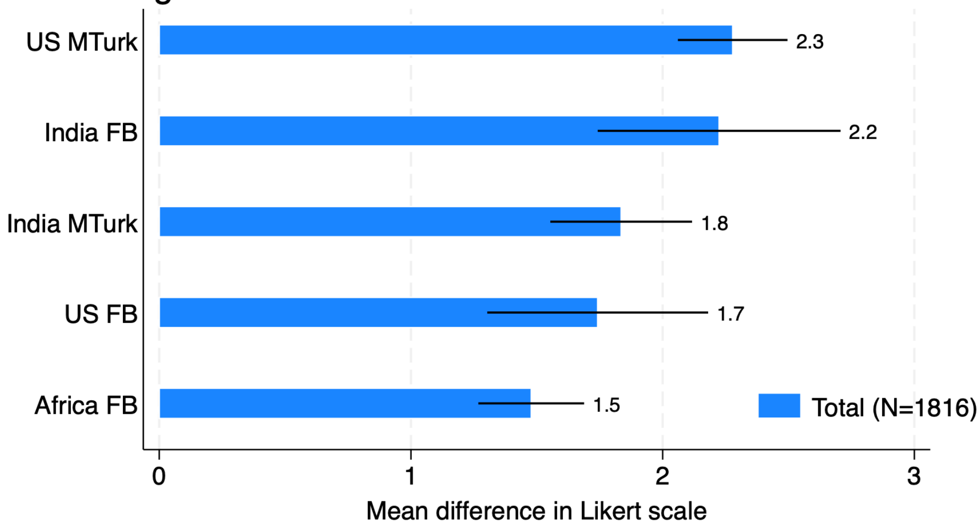
without material incentives



Note: 95% confidence intervals shown. Each bar reports the mean of answers to a 0-10 Likert scale question.
T-test pairwise comparisons at 5% significance or better yields the following statistically significant rankings:
Africa FB > India MTurk and India FB > US MTurk ; Africa FB > US FB.

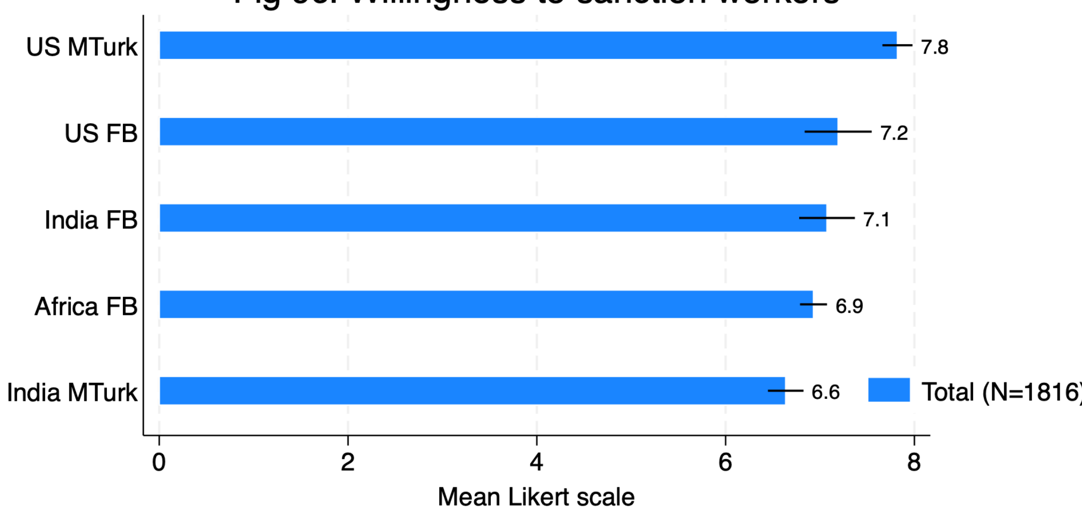
The second set of questions focuses on the acceptability of various forms of ex post punishment for low effort or skill misreporting. Three types of punitive measures are considered in the survey questions: firing the worker; reducing pay by 30%; and cancelling a 30% bonus. We regard being in favor of punitive measures against shirking or skill misreporting as indicative of an individualistic, selfish-rational attitude, i.e., archetype A1. Since answers to the six questions are correlated within individual, we construct an aggregate index by summing answers to all six. Variation in this index across subject populations is summarized in Figure 9c. We again see that US subjects are more likely than African or Indian subjects to regard punishment for low effort as acceptable. This is consistent with US subjects also believing more in the effectiveness of incentives than Indian and African subjects.

Fig 9b. Beliefs in the effectiveness of work incentives



Note: 95% confidence intervals shown.
Each bar reports the mean difference in answers to a Likert scale question from
Figures 9a and a similar question with material incentives.
T-test pairwise comparisons at 5% significance or better yields the following statistically significant rankings:
US MTurk > India MTurk, US FB, and Africa FB; India FB and India MTurk > Africa FB

Fig 9c. Willingness to sanction workers



Note: 95% confidence intervals shown.
Each bar reports the mean answers to separate Likert scale questions
about the acceptabilty of firing, reducing wage, or not giving a bonus
to a worker who is caught shirking or lacks self-reported skills.
T-test pairwise comparisons at 5% significance or better yields the following statistically significant rankings:
US MTurk > USA FB ~ India FB ~ Africa FB > India MTurk.

Having established that our subjects do hold behavioral expectations that vary across world regions, we now investigate whether these differences – and differences in other individual characteristics – can account for differences in behavior in our experiment. To do this, we turn to a formal investigation of whether $D^w_{AB} = 0$. To keep the analysis straightforward, we focus on four dependent variables θ^w_j per subject, each constructed from the subject's acceptance and effort choices. We first create a dummy equal to 1 if the effort choice of the subject in a particular game is compatible with archetype A1 (intrinsically motivated play) according to Table 4. We do the same for A2 (conditional reciprocal play), for A3 (selfish-rational play), and for A4 (unconditionally unmotivated). We then average these dummies across games for each subject to construct indices going from 0 to 1, one for each of the four archetypes.[22]

We then regress these four dummies θ_j^w on a vector of predictors X_j to obtain the predictions $\hat{P_A^w}(X_j)$ in a given population A. The predictors include the individual characteristics presented in Table 3, as well as the variables used to produce Figures 9a, b and c, since these variables capture differences in beliefs and expectations that may be known or partly observable to individuals from the same population. Results are illustrated in Tables B1 to B4 in Appendix, separately for each of the three subsamples: US, India, and Africa. In the case of random forest, we *linearly* regress the predictions $\hat{P_A^w}(X_j)$ obtained with the random forest estimator on the same predictors X_j to allow for comparison with OLS. Four conclusions emerge from this exercise. First, the *linear* fit is uniformly better for the random forest estimator than for the simple linear projection of the archetypes on the predictors – but neither of the two methods yields a particularly tight fit. Second, for the same world region and archetype, the set of significant predictors is nearly always the same for OLS and RF, indicating that we have robust predictors. Third, for a given world region, the set of significant predictors varies a lot across archetypes, implying a lot of heterogeneity in archetype behavior. And last but not least: for a given archetype, the set of significant predictors varies a lot across world regions, which already suggests differences in archetype predictability across them.

We now turn to the test of differences in means between $\hat{P}_B^w(X_A)$ and $\hat{P}_A^w(X_B)$. We do so for three pairwise comparisons: US vs India; US vs Africa; and India vs Africa. Two predicting models are used in the analysis: ordinary least-squares (OLS) and random forest (RF).[23] These methods are chosen because they tend to bracket the in-sample performance of predictive estimators: quite low for OLS and quite high for RF.[24]

[22]Since it is possible for several dummies to be simultaneous equal to 1 in a particular contract, the sum of the four indices can exceed 1. To illustrate, suppose that in the first game the subject chooses high effort in a low fixed wage contract. This is compatible with A1 but not with either of the other three – hence the A1 dummy equals 1 for that game, and the other three equal 0. In the second game, the subject refuses a high fixed wage contract. This is only compatible with the A4 archetype, which means we set the A4 dummy equal 1, and the other three to 0. In the third game the subject chooses high effort in a Bonus contract. This choice is compatible with archetypes A1 to A3 listed in Table 4, but not A4. Hence the A1, A2, and A3 dummies are set to 1, and the A4 dummy is set to 0. In the fourth and final game, the subject choose low effort in a high fixed wage contract. This is only compatible with the A3 archetype – hence that dummy is set to 1 and the other three to 0. We then average over the fours games for that subject. The values of the four indices are thus: 0.5 for A1 and A3, and 0.25 for A2 and A4.

[23]All predictors are normalized and standardized before applying OLS or random forest.

[24]In Online Appendix Table A2, we show that, when we pool two regions together and regress θ_j^w on X_j without including region dummies, we indeed obtain a reasonable but low fit with OLS and a very high fit with random forest. The Table also shows that both predictive methods perform equally poorly out-of-sample. To this effect, we obtain the predictions made for region B based on fitting a predictor to region A, and similarly for the predictions for A produced by fitting a predictor to region B. We then pool these out-of-sample predictors across both regions and regress each of the four dummies θ_j^w on the combined predictor vector. The unadjusted R^2 of each regression and each pairwise comparison is shown in the last two columns of Appendix Table A3. We see that the fit is uniformly poor using either of the two predictors, OLS or RF. This is preliminary evidence that predicting behavior in region j based on observed behavior in region i performs poorly in our data.

Table 9. Can behavioral predictions from one region predict average behavior in another region?

	T-tests of the mean difference between regions				
US vs India	**Actual**	**In-sample predictions**		**Out-of-sample predictions**	
A1 index (Intrinsically motivated)	**Data**	**OLS**	**RF**	**OLS**	**RF**
Estimated mean difference between regions	-0.318	-0.116	-0.237	0.286	0.255
t-test value of difference in means	-5.44	-7.62	-6.17	14.20	18.81
Standard error of difference in means	0.059	0.015	0.038	0.020	0.014
A2 index (Conditional cooperator)					
Estimated mean difference between regions	-0.019	0.052	0.017	0.228	0.191
t-test value of difference in means	-0.31	5.32	0.43	16.81	14.32
Standard error of difference in means	0.060	0.010	0.039	0.014	0.013
A3 index (Selfish-rational)					
Estimated mean difference between regions	0.362	0.185	0.305	-0.231	-0.154
t-test value of difference in means	6.19	12.18	8.08	-16.30	-12.03
Standard error of difference in means	0.058	0.015	0.038	0.014	0.013
A4 index (Unconditionally unmotivated)					
Estimated mean difference between regions	-0.004	-0.039	-0.018	-0.056	-0.056
t-test value of difference in means	-0.07	-3.76	-0.45	-3.52	-3.90
Standard error of difference in means	0.061	0.010	0.040	0.016	0.014
US vs Africa	**Actual**	**In-sample predictions**		**Out-of-sample predictions**	
A1 index (Intrinsically motivated)	**Data**	**OLS**	**RF**	**OLS**	**RF**
Estimated mean difference between regions	-0.312	-0.080	-0.236	0.539	0.314
t-test value of difference in means	-5.61	-5.60	-6.37	25.68	21.68
Standard error of difference in means	0.056	0.014	0.037	0.021	0.015
A2 index (Conditional cooperator)					
Estimated mean difference between regions	0.034	0.071	0.054	0.186	0.164
t-test value of difference in means	0.60	7.99	1.49	14.99	12.91
Standard error of difference in means	0.056	0.009	0.036	0.012	0.013
A3 index (Selfish-rational)					
Estimated mean difference between regions	0.363	0.204	0.299	-0.208	-0.138
t-test value of difference in means	6.61	16.84	8.44	-14.20	-10.28
Standard error of difference in means	0.055	0.012	0.035	0.015	0.013
A4 index (Unconditionally unmotivated)					
Estimated mean difference between regions	0.036	-0.071	0.003	-0.321	-0.184
t-test value of difference in means	0.65	-8.88	0.09	-23.40	-10.21
Standard error of difference in means	0.056	0.008	0.036	0.014	0.018
India vs Africa	**Actual**	**In-sample predictions**		**Out-of-sample predictions**	
A1 index (Intrinsically motivated)	**Data**	**OLS**	**RF**	**OLS**	**RF**
Estimated mean difference between regions	0.006	-0.025	0.006	-0.024	0.014
t-test value of difference in means	0.11	-1.59	0.16	-1.28	1.06
Standard error of difference in means	0.058	0.015	0.038	0.019	0.014
A2 index (Conditional cooperator)					
Estimated mean difference between regions	0.053	0.015	0.039	-0.042	-0.035
t-test value of difference in means	0.92	1.91	1.06	-3.42	-2.75
Standard error of difference in means	0.057	0.008	0.037	0.012	0.013
A3 index (Selfish-rational)					
Estimated mean difference between regions	0.001	0.058	0.013	0.127	0.033
t-test value of difference in means	0.02	5.29	0.35	9.22	2.77
Standard error of difference in means	0.058	0.011	0.037	0.014	0.012
A4 index (Unconditionally unmotivated)					
Estimated mean difference between regions	0.040	-0.028	0.017	-0.149	-0.182
t-test value of difference in means	0.70	-2.90	0.45	-9.13	-12.72
Standard error of difference in means	0.057	0.010	0.038	0.016	0.014

Notes: For this Table, data from the MTurk and Facebook samples are combined. Each cell of the Table presents the result of a $t-$test of equality of means across two regions of the world. Since we have three regions (US, India, Africa), there are three pairwise comparisons. The means that are compared are those of four behavioral indices (see text for details). The first number is the mean of the first region minus the mean of the second region -- i.e., a positive number means that the first region has a larger value. Each panel compares a pair of regions to each other. The second number is the value of the t-test. The third number is the standard error of the difference in means. The column labeled Actual Data represents the t-test of difference in means in the data itself. The other four columns are based on the following procedure: (1) regress the dependent variable on a vector of predictors separately in each region and obtain the in-sample predictors; (2) obtain the in-sample predictions -- e.g., the OLS predicted values for the US based on the US regression -- and the out-of-sample predictions -- e.g., the OLS predicted values for India based on the US regression; (3) stack the in-sample predictions into a pooled vector of in-sample predictions, and stack the out-of-sample predictions into a pooled vector of out-of-sample predictions; and (4) estimate a pooled t-test by region for either the in-sample predictions or the out-of-sample predictions. In the OLS columns, predictions are obtained using least squares; they are obtained using random forest in the RF column.

Our main test results are presented in Table 9. Each cell of three figures represents the result of a t-test of equality of means across two regions of the world. Since we have three regions (US, India, Africa), there are three pairwise comparisons, each over four behavioral indices θ_j^w.

The first column presents an unconditional t-test of the *actual* θ_j^w between regions A and B. A positive value means that the average of the index is higher in region A than B. This column replicates the results already discussed in Section 4.1, albeit on subject-level observations.

The interest of the Table is in the comparison between test results for in-sample and out-of-sample predictions.[25] In the in-sample predictions, we compare the mean of $\hat{P}_A^w(X_A)$ to the mean of $\hat{P}_B^w(X_B)$, that is behavioral predictions made for region A by applying a prediction model to data from that region, to similar in-sample predictions made for region B. We expect the difference between these average predictions to be roughly similar to the mean difference in the actual data.

The interesting part is the out-of-sample comparison, which focuses on $\hat{P}_B^w(X_A) - \hat{P}_A^w(X_B)$. If individual characteristics predict differences in behavior *and* the predictions made in one region are the same as those in another region for an individual with identical characteristics, then the difference $E\left[\hat{P}_B^w(X_A) - \hat{P}_A^w(X_B)\right]$ should have the same sign as $E\left[\theta_A^w - \theta_B^w\right]$. If this is the case, then the magnitude of the gap between $E\left[\hat{P}_B^w(X_A) - \hat{P}_A^w(X_B)\right]$ and $E\left[\theta_A^w - \theta_B^w\right]$ gives an idea of how much of the behavioral difference between A and B can be attributed to differences in characteristics between the two sample populations. In this case, someone from region A arriving in region B would be able to predict behavior in B based on the characteristics of individuals in that region. On the other hand, if $E\left[\hat{P}_B^w(X_A) - \hat{P}_A^w(X_B)\right]$ and $E\left[\theta_A^w - \theta_B^w\right]$ do not have the same sign, this means that differences in characteristics across subjects in the two population do not predict behavioral differences between them. Consequently, someone from A arriving in B would be surprised by the behavior of people in A, and vice versa for people from B arriving in A.

The results presented in the last two columns of Table 9 show that, in all but a few cases, $E\left[\hat{P}_B^w(X_A) - \hat{P}_A^w(X_B)\right]$ either has a sign opposite to $E\left[\theta_A^w - \theta_B^w\right]$, or $E\left[\hat{P}_B^w(X_A) - \hat{P}_A^w(X_B)\right]$ is large and significant although $E\left[\theta_A^w - \theta_B^w\right]$ is small in magnitude and not statistically significant. This implies that subjects A could not use the observational characteristics of subjects in B to make unbiased predictions of their behavior, and vice versa for subjects in B. This is another way of saying that differences in individual characteristics between regions cannot account for the differences in behavior observed in our samples. While we cannot say what *causes* these behavioral differences, we can be confident that they are not predicted by a long list of individual characteristics. It is conceivable that data on other individual characteristics would be able to predict the observed cross-country differences. But at this point, we do not know what these are, and how observable they would be to an employer.

[25] The standard error of the difference in means is included in the Table as a visual guide on the role that prediction precision plays in the test: because OLS predictions are less variable than RF predictions (i.e., lower R^2), comparing means obtained via OLS may yield a more significant test than when using RF.

8 Conclusion

In this paper we reported the results from an online experiment framed as a series of one-shot employment contracts between an anonymous employer and an anonymous worker. Subjects assigned the role of employer make an employment offer they select among a restricted set of contracts. Workers choose whether to accept the contract, and an effort level if they do. High effort is always efficient but it is costly to workers. Participants are recruited among individuals registered on Amazon Mechanical Turk in the US and India, and through Facebook ad campaigns in seven countries covering three main regions of the world: USA, India, and Africa.[26]

We focus on four behavioral archetypes that are inspired by the literature: A1 (intrinsically motivated), A2 (conditional reciprocator), A3 (selfish-rational), and A4 (unconditionally unmotivated). We find that A1 and A4 (behavior that is not contingent on the wage level) account for the overwhelming majority of choices made in the experiment. A smaller fraction of choices follows the A2 archetype, and an even smaller fraction follows A3. There are significant differences between the three population samples in the relative proportions of the four archetypes, but they are present in all sample. These differences in behavior between regions are matched, on average, by differences in beliefs regarding the effectiveness of worker incentives and in the acceptability of employer sanctioning workers who shirk or misrepresent their skills.

In a subsample of participants, we examine whether subjects behave differently when matched with someone from another country. First, we look at prejudice, defined here as being less positively inclined towards someone of a different region by providing less effort. We find no evidence that subjects choose systematically lower effort when matched with a foreign employer. We also investigate whether our subjects have cultural stereotypes regarding effort choice, meaning that they expect workers from other regions to work less or more than those of their own. We find no strong evidence in our experiment that cross-country stereotypes affect the offer of contracts to online workers from another region.

In the last part of the paper, we investigate whether differences in lab behavior across the three main regions covered by our sample could be predicted from differences in subjects' observable characteristics. To this effect, we conduct a thought experiment in which subjects from one region extrapolate to another those correlations between behavior and individual characteristics that are present in their own region. We find strong evidence that differences in individual characteristics between regions cannot account for the differences in behavior observed in our samples. While we cannot say what causes these behavioral differences, we cannot rule out the possibility that they capture cultural differences such as those discussed in Kirkman et al. (2001), Henrich et al. (2001, 2006, 2010), Guiso et al. (2006), Falk et al. (2018), and Schulz et al. (2018, 2019).

More work is needed to disentangle the mental processes that underlie the four algorithms discussed in this paper.

[26]We also collected data from three developed countries but the sample sizes were too small for most of the analysis presented here.

References

Abeler, J., Altmann, S., Kube, S., & Wibral, M. (2010). "Gift-Exchange and Worker's Fairness Concerns: When Equality Is Unfair", *Journal of the European Economic Association*, 8: 1299–324.

Adams, G., Estrada-Villalta, S., Sullivan, D., & Markus, H. R. (2019). "The psychology of neoliberalism and the neoliberalism of psychology". *Journal of Social Issues*, 75(1): 189-216.

Andreoni, James, Deniz Aydin, Blake Barton, B. Douglas Bernheim, and Jeffrey Naecker (2020). "When Fair Isn't Fair: Understanding Choice Reversals Involving Social Preferences", *Journal of Political Economy*, 128(5): 1673-711

Ashraf, Nava, Oriana Bandiera, and Kelsey Jack (2014). "No Margin, No Mission? A Field Experiment on Incentives for Public Services Delivery." *Journal of Public Economics*, 120: 1-17.

Baker, G., Gibbons, R., & Murphy, K. J. (2002). "Relational Contracts and the Theory of the Firm". *Quarterly Journal of Economics*, 117(1), 39–84. http://www.jstor.org/stable/2696482

Bandiera, Oriana, Iwan Barankay & Imran Rasul (2010). "Social Incentives in the Workplace," *Review of Economic Studies*, 77(2): 417-58.

Bandiera, Oriana, Iwan Barankay & Imran Rasul (2013). "Team Incentives: Evidence From A Firm Level Experiment," *Journal of the European Economic Association*, 11(5): 1079-114.

Banuri, Sheheryar, Catherine Eckel and Rick K. Wilson (2022), "Does cronyism pay? Costly ingroup favoritism in the lab", *Economic Inquiry*, 1–19. Available from: https://doi.org/10.1111/ecin.13080

Bazzi, Samuel, Martin Fiszbein and Mesay Gebresilasse (2020). "Frontier Culture: The Roots and Persistence of "Rugged Individualism" in the United States", *Econometrica*, 88(6): 2329-68

Bell, Linda, and Richard B. Freeman (2001). "Working Hard." In *Working Time in Comparative Perspective: Patterns, Trends, and the Policy Implications of Earnings Inequality and Unemployment*, Vol. 1, Ging Wong, and Garnett Picot, eds. Kalamazoo, MI: W.E. Upjohn Institute for Employment Research, pp. 71–105.

Bénabou, Roland and Jean Tirole (2006). "Incentives and Prosocial Behavior," *American Economic Review*, 96(5): 1652-78, December

Bicchieri, Cristina, Eugen Dimant, Simon Gächter, and Daniele Nosenzo (2022). "Social proximity and the erosion of norm compliance," *Games and Economic Behavior,* 132: 59-72, 2022,

Breza, Emily, Supreet Kaur and Yogita Shamdasani (2018). "The Morale Effects of Pay Inequality", *Quarterly Journal of Economics*, 133(2): 611–63

Bloom, N., & Van Reenen, J. (2007). "Measuring and Explaining Management Practices across Firms and Countries", *Quarterly Journal of Economics*, 122(4): 1351-408

Bloom, N., & Van Reenen, J. (2010). "Why Do Management Practices Differ across Firms and Countries?", *Journal of Economic Perspectives*, 203-224.

Bottazzi, L., Da Rin, M., & Hellmann, T. F. (2016). "The Importance of Trust for Investment: Evidence from Venture Capital", Review of Financial Studies, 29(9): 2283–2318

Brown, M., Falk, A., & Fehr, E. (2004). ”Relational Contracts and the Nature of Market Interactions”, *Econometrica*, 72(3): 747-80.

Brown, M., Falk, A., & Fehr, E. (2012). ”Competition and Relational Contracts: The Role of Unemployment as a Disciplinary Device”, *Journal of the European Economic Association*, 10(4): 887-907.

Buchmann, Nina, Colin Sullivan, and Carl Meyer (2022). “Paternalistic Discrimination”, Stanford University (mimeograph)

Cassar, Lea and Stephan Meier, (2018). "Nonmonetary Incentives and the Implications of Work as a Source of Meaning," *Journal of Economic Perspectives*, 32(3): 215-38, Summer.

Cassar, Lea (2019). "Job Mission as a Substitute for Monetary Incentives: Benefits and Limits," *Management Science*, 65(2): 896-912, February.

Charness, G. (2004). ”Attribution and Reciprocity in an Experimental Labor Market”, *Journal of Labor Economics*, 22(3): 665–688.

Charness, G., & Dufwenberg, M. (2006). ”Promises and Partnership”, *Econometrica*, 74(6): 1579-160

Charness, G., Gneezy, U., & Kuhn, M. A. (2013). ”Experimental Methods: Extra-Laboratory Experiments Extending the Reach of Experimental Economics”, *Journal of Economic Behavior and Organization*, 91: 93-100.

Charness, G., & Kuhn, P. (2011). ”Lab Labor: What Can Labor Economists Learn from the Lab?”, *Handbook of Labor Economics*, Volume 4: 229-330.

Chen, Y., et al. (2014). ”Which Hat to Wear? Impact of Natural Identities on Coordination and Cooperation”, *Games and Economic Behavior*, 84: 58-86.

Chu, James and Marcel Fafchamps (2022). ”Procedural Fairness and Antagonistic Labor Relations: The Case of Manufacturing Firms in Ethiopia”, *World Development*, 159: 106037, November.

Chuah, Swee-Hoon, Simon Gächter, Robert Hoffmann, and Jonathan H.W. Tan (2023). “Who discriminates? Evidence from a trust game experiment across three societies,” *Journal of Economic Psychology*, 97, 2023,

Cohen, Daniel & Piketty, Thomas & Saint-Paul, Gilles (ed.), (2014). *The Economics of Rising Inequalities*, Oxford University Press, Oxford

Cullen, Z. and Perez-Truglia, R. (2022). “How Much Does Your Boss Make? The Effects of Salary Comparisons”, *Journal of Political Economy*, 130(3): 766-822

Davies, E., & Fafchamps, M. (2021). ”When No Bad Deed Goes Punished: A Relational Contracting Experiment in Ghana”, *Journal of Business and Economic Organization*, 191: 714-37

DellaVigna, Stefano, and Devin Pope (2018a). ”What Motivates Effort? Evidence and Expert Forecasts”, *Review of Economic Studies*, April, 85: 1029–69.

DellaVigna, Stefano, and Devin Pope (2018b). ”Predicting Experimental Results: Who Knows What?”, *Journal of Political Economy*, December, 126: 2410-56.

DellaVigna, Stefano, John A. List, Ulrike Malmendier, and Gautam Rao (2022). "Estimating Social Preferences and Gift Exchange at Work." *American Economic Review*, 112 (3): 1038–74.

Enke, Benjamin (2019). ”Kinship, Cooperation, and the Evolution of Moral Systems”, *Quarterly Journal of Economics*, 134(2): 953-1019.

Falk, Armin, Anke Becker, Thomas Dohmen, Benjamen Enke, David Huffman, and Uwe Sunde (2018). ”Global Evidence on Economic Preferences”, *Quarterly Journal of Economics*, 133(4): 1645-92

Fehr, E., Gächter, S., & Kirchsteiger, G. (1997). ”Reciprocity as a Contract Enforcement Device: Experimental Evidence”. *Econometrica*, 65(4), 833–860. doi:10.2307/2171941

Fehr, E., Kirchler, E., Weichbold, A., & Gächter, S. (1998). ”When Social Norms Overpower Competition: Gift Exchange in Experimental Labor Markets”, *Journal of Labor Economics*, 16(2): 324–51. doi:10.1086/209891

Fehr, E., Klein, A., & Schmidt, K. M. (2007). ”Fairness and Contract Design”, *Econometrica*, 75(1): 1–19.

Fehr, Ernst, and Simon Gächter (2000). "Fairness and Retaliation: The Economics of Reciprocity." *Journal of Economic Perspectives*, 14 (3): 159-181.

Fields, Gary S. (2011). “Labor market analysis for developing countries”, *Labour Economics*, 18(1): S16-S22

Gächter, S., & Falk, A. (2002). ”Reputation and Reciprocity: Consequences for the Labour Relation”, *Scandinavian Journal of Economics*, 104(19): 1–27. doi:10.2307/3441139

Gosnell, Greer K., John A. List, and Robert D. Metcalfe (2020). “The Impact of Management Practices on Employee Productivity: A Field Experiment with Airline Captains”, *Journal of Political Economy*, 128(4): 1195-233.

Guiso, L., Sapienza, P., & Zingales, L. (2004). ”The Role of Social Capital in Financial Development”, *American Economic Review*, 526-556.

Guiso, L., Sapienza, P., & Zingales, L. (2006). ”Does Culture Affect Economic Outcomes?”, *Journal of Economic Perspectives*, 20(2): 23-48.

Henrich, J., et al. (2001). ”In Search of Homo Economicus: Behavioral Experiments in 15 Small-scale Societies”, *American Economic Review*, 73-78.

Henrich, et al. (2006). ”Costly Punishment across Human Societies”, *Science*, 312(5781): 1767-70.

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). ”The Weirdest People in the World?”, *Behavioral and Brain Sciences*, 33(2-3): 61-83.

Horton, J. J., Rand, D. G., & Zeckhauser, R. J. (2011). ”The Online Laboratory: Conducting Experiments in a Real Labor Market”, *Experimental Economics*, 14(3): 399-425.

Hwang, S.-H. and Bowles, S. (2014), “Optimal Incentives with State-Dependent Preferences”. *Journal of Public Economic Theory*, 16: 681-705.

Jackson, M., & Xing, Y. (2014). ”Culture-Dependent Strategies in Coordination Games”, *Proceedings of the National Academy of Science*, 111: 10889-96.

Kajackaite, Agne & Werner, Peter, (2015). "The Incentive Effects of Performance Requirements – A Real Effort Experiment," *Journal of Economic Psychology*, 49(C): 84-94.

Kandori, M. (1992). "Social Norms and Community Enforcement", *Review of Economic Studies*, 59(1): 63-80.

Kirkman, B. L., & Shapiro, D. L. (2001). "The Impact of Team Members' Cultural Values on Productivity, Cooperation, and Empowerment in Self-Managing Work Teams", *Journal of Cross-Cultural Psychology*, 32(5): 597-617.

La Porta, R. et al. (1997). "Trust in large organizations", *American Economic Review*, 333-338.

Lazear, Edward P. (2018). "Compensation and Incentives in the Workplace", *Journal of Economic Perspectives*, 32(3): 195–214

Landry, Craig and Lange, Andreas and List, John A. and Price, Michael K. and Rupp, Nicholas G. (2011). "The Hidden Benefits of Control: Evidence from a Natural Field Experiment", NBER WP 17473, Available at SSRN: https://ssrn.com/abstract=1922771 or http://dx.doi.org/10.2139/ssr

MacLeod, W. B. (2011). "Great Expectations: Law, Employment Contracts, and Labor Market Performance". *Handbook of Labor Economics*, Volume 4, 1591-696.

Ockenfels, Axel, Dirk Sliwka and Peter Werner (2015). "Bonus Payments and Reference Point Violations", *Management Science*, 61(7): 1496-513.

Prissé, Benjamin and Diego Jorrat (2022). "Lab vs online experiments: No differences", *Journal of Behavioral and Experimental Economics*, 100, 101910

Rebitzer, J. B., & Taylor, L. J. (2011). "Extrinsic Rewards and Intrinsic Motives: Standard and Behavioral Approaches to Agency and Labor Markets", *Handbook of Labor Economics*, Volume 4: 701-72.

Schulz, Jonathan, Duman Bahrami-Rad, Jonathan Beauchamp, and Joseph Henrich (2019). "The Church, intensive kinship, and global psychological variation", *Science*, 366(6466): eaau5141

Schulz, J., A. Edmond, S. Dsouza, R. Kim, J. Henrich, A. Shariff, J. Bonnefon, and I. Rahwan (2018). "The Moral Machine Experiment", *Nature*, 563: 59-64

Sliwka, Dirk & Werner, Peter, (2017). "Wage Increases and the Dynamics of Reciprocity", *Journal of Labor Economics*, 35(2): 299-344

Weiss, Andew (1987). "Incentives and Worker Behavior: Some Evidence", *Cooperation, Incentives and Risk Sharing*, ed. Haig Nalbantian; Rowman and Littlefield Publishers, 1987

Online Appendices

9 Sample recruitment

9.1 Recruitment via Amazon Mechanical Turk (MTurk)

The experiment was first implemented online using subjects recruited from Amazon Mechanical Turk (MTurk).[27] The main reason for using MTurk is to be able to easily pay subjects conditional on their performance in the experiment while ruling out multiple play by the same subject. Indeed, paying online subjects in multiple countries is fraught with difficulties and, at the time of the experiment started, survey outfits such as Qualtrics were refusing to pay participants from their subject pool anything other than a fixed fee.[28] Since the two largest pools of subjects on MTurk are from India and the US, the MTurk online experiment used a 2x2 country design summarized in Table A1.

After filling an online survey on Qualtrics and giving informed consent for their participation to the experiment, subjects were offered a choice of time windows at which they were invited to join a specific online session of their choosing. Each session is designed for 8 players so as to ensure random rematching. As soon as the desired number of online subjects is reached in a session, the sequence of games is initiated. Subjects are nonetheless identified to the researcher by their MTurk identifier, which ensures that no subject is allowed to play more than one session.

Table A1. Breakdown of the MTurk sample

	Subjects from:	
Matched with subject from:	US	India
US	279	254 (*)
India	266 (*)	205
Total	545	459

Note: (*) except for 9 subjects (5 from the US and 4 from Indian) who are matched with someone from the other country only part of the time.

Sessions were organized over a period of several months between the summer of 2017 and the spring of 2018. In total 2,260 individuals filled the online questionnaire. Of those, 1,004 participated in the experiment. All subjects were expected to play 8 games in total. In practice

[27] Amazon Mechanical Turk is an online jobbing site. Individuals are free to register and post their job-related information and bank account data. They can then apply to various 'gigs' advertised in the platform, most of which take place online. MTurk has been used as convenient subject recruitment pool in a large number of online experiments, of which Jackson and Xin (2014) is a good example. But to our knowledge, it had not yet been used for experiments in which subjects interact with each other. In recent years, Prolific Academic has tended to replace MTurk as the sampling pool favored by experimentalists. But at the time we conducted our own experiment, Prolific Academic was not sufficiently developed to allow the sample sizes we were aiming for.

[28] At the time that we initiated the experiment, O-Tree did not yet exist. This means that the experiment was coded directly by the researchers using a combination of PHP and JavaScript. Subjects were then channelled from Qualtrics – used for the survey and consent form – to the experimental interface and finally to MTurk – used for payment.

some subjects arrive late and some leave early, either because of internet connection issues or because they get distracted. Hence the number of subjects in an online session varies somewhat over the duration of the session. The experimental protocol is specifically programmed to take this into account by re-matching subjects on the fly to minimize disruption.[29] The total number of games played is 3060, meaning that subjects on average played 6.1 games instead of the maximum possible of 8.[30]

Subjects spent an average of 6 minutes on each batch of four games, with a median of 5 minutes. Each of the six treatments was played by between 482 to 554 pairs of subjects, depending on the treatment. In the analysis presented here, 25 participants are dropped for various data quality reasons (20 from the US and 5 from India), leaving a total sample of 979.

In terms of compensation, the exchange rate is US$0.03 per point in Table 1. Subjects are paid for each of the games that they play. This mode of compensation is chosen to incentivize staying until the end of the experiment. On average subjects who participated in the experiment received a compensation of 3.5US$,[31] which is considered normal for MTurk experiments. Those who only filled the questionnaire received a fixed fee of 2$. Despite differences in standards of living between the two countries, we use these same compensation amounts to avoid contaminating our design by introducing equity considerations when pairing subjects across countries.

9.2 Recruitment via Facebook

We recruited an additional 1260 subjects via Facebook. To allow comparison between the populations recruited by MTurk and Facebook, India and the US are also included in the Facebook targeted countries, together with eight other countries. The other targeted countries included five African countries (Kenya, Malawi, Morocco, Senegal, and South Africa) and three high-income countries (Australia, Canada, and France). The experiment was offered in French to subjects in France, Canada, Morocco, and Senegal, and in English to the other six. As in the MTurk experiment, subjects were first invited to fill in a survey in Qualtrics, at the end of which they were offered the option to choose a time slot to register for the online interactive experiment. The rest of the experimental design is identical to the MTurk sessions.[32] After extensive piloting in late 2019, the Facebook sessions took place in the first half of 2020, running through the first phase of the Covid-19 pandemic.

Recruitment was done through paid-for ads on Facebook. This proved difficult and costly. The payment of subjects turned out to be our hardest logistical challenge, given the lack of widely used platform for small payments across countries. We opted for PayPal because, to prevent money laundering, users are in principle not allowed to have PayPal accounts in different

[29] In these cases, some individuals are not rematched with someone from the intended country (see Table A1) to avoid losing observations.

[30] The first game is played by 414 pairs of players, the second by 433 pairs, games 3-6 by around 400 pairs, and games 7 and 8 by 308 or 309 pairs of players.

[31] $3.45 in India and $3.49 in the US.

[32] Here too we keep the compensation amounts the same across countries to avoid an online backlash if, say, it was discovered that we were paying more to subjects from rich countries.

countries or to open multiple PayPal accounts linked to the same bank account. Based on this understanding of PayPal's rules, we aimed to prevent subjects from participating multiple times by limiting each PayPal identifier to a single shot at the Qualtrics questionnaire and the online experiment. We subsequently discovered that, in some countries, PayPal seems to allow users to register multiple email addresses under the same PayPal account – e.g., for members of the same family or group. We also came to suspect that some users have multiple PayPal accounts.[33] Having discovered this problem, we spent a large amount of time checking all participants and rejecting a large proportion of dubious observations.[34] While the problem was most severe for the Qualtrics survey, it also affected the online experiment.[35] On the upside, we cannot reject the possibility that dubious observations come from different subjects who used the same PayPal account solely for payment purposes – their answers to the survey are not identical, for instance. Furthermore, there does not appear to be a significant difference in survey answers between those classified as reliable or dubious. This being said, in an abundance of caution all the dubious observations are omitted from the analysis presented here.

Table A2. Facebook sample size by country and data quality

	Verified	Dubious	All
USA Facebook	122	22	144
Australia (*)	15	0	15
Canada (*,**)	10	0	10
France (*)	17	38	55
Morocco	44	0	44
India Facebook	141	39	180
South Africa	99	129	228
Kenya	91	259	350
Senegal	35	27	62
Malawi	134	38	172
Total	708	552	1260

Notes: Dubious data quality refers to subjects who, based on their payment data and email name, are suspected of using multiple identifiers to play multiple times. We cannot, however, rule out that

[33]These suspicions are based on similarity in user names and interaction with us. We do not have information on the bank accounts associated with PayPal identifiers – except that, in a number of cases, our payments to subjects were rejected because PayPal was unable to deposit funds in the account number registered by the user.

[34]In addition to participants with rejected payments, we dropped all groups of participants that used the same PayPal account: while they may represent different members of the same household, they may also indicate that a single person participated to the experiment multiple times. We also scanned manually all PayPal identifiers and dropped all groups of participants using identifiers that were suspicious similar to each other, jds1, jds2, and jds3@gmail.com. While there may exist multiple individuals with the same initials in a given country, it is nonetheless unlikely that they would have all seen our Facebook ad and joined our experiment. It is more likely that they are the same person who created multiple email addresses to be paid multiple times for participating in the experiment.

[35]For the Qualtrics survey, we also rejected observations for which the time spent to fill the survey was too short, suggesting lack of attention.

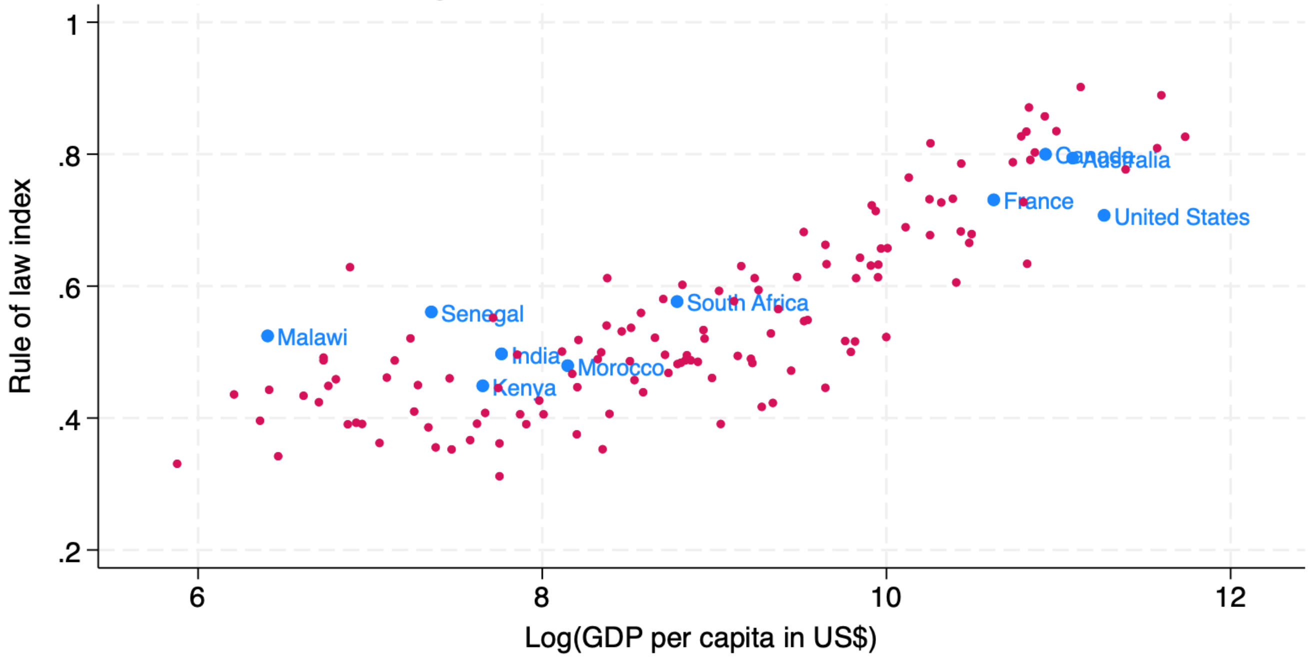
these are different subjects using the same PayPal identifier for payment purposes. (*) Due to the small number of verified observations, these countries are omitted from our empirical analysis. (**) In Canada, the invitation to participate was targeted at French speakers.

Table A2 shows a breakdown of the Facebook experimental sample across countries, and the number of observations that were rejected as dubious. There are more dubious observations in some countries than others, possibly because PayPal adapts its rules to local laws and regulations. The Table also shows large variation in the number of participants from different countries. This variation in sample size by country is largely due to the process by which Facebook allows us to target subjects by demanding that we pre-specify a budget for a particular country. Once the budget is set, the Facebook algorithm works out a 'price' per subject based on the proportion of recipients of the ad who engage with it by clicking on the link to the Qualtrics survey. In countries where this price is high, the budget we set beforehand only yields a small number of individuals. Furthermore, only those who complete the survey are invited to the online experiment, but the proportion of invited subjects who participate also varies across countries. Our capacity to attract subjects to the experiment also depends on them having (or opening) a PayPal account, something that is widespread in some countries but not in others. In France, Canada and Australia, the unit cost per subject turned out to be particularly high, thereby rapidly exhausting our budget for these countries. Because we only managed to recruit a small number of subjects there, they are dropped from the analysis. For ease of presentation, we combine the five African countries into a single category for the purpose of analysis.

Compensation levels were kept the same as for the MTurk sample (US$0.03 per point), so as to keep things comparable between the two samples. We chose to use the same remuneration level for all countries: in a cross-country online experiment, paying participants less simply because they live in a low-income country would have been unacceptable ethically, and could have created a significant backlash if publicized on social media. The compensation amounts offered to participants (around 3.5US$) are of the same order of magnitude as cell phone time costs.

Additional Figures and Tables

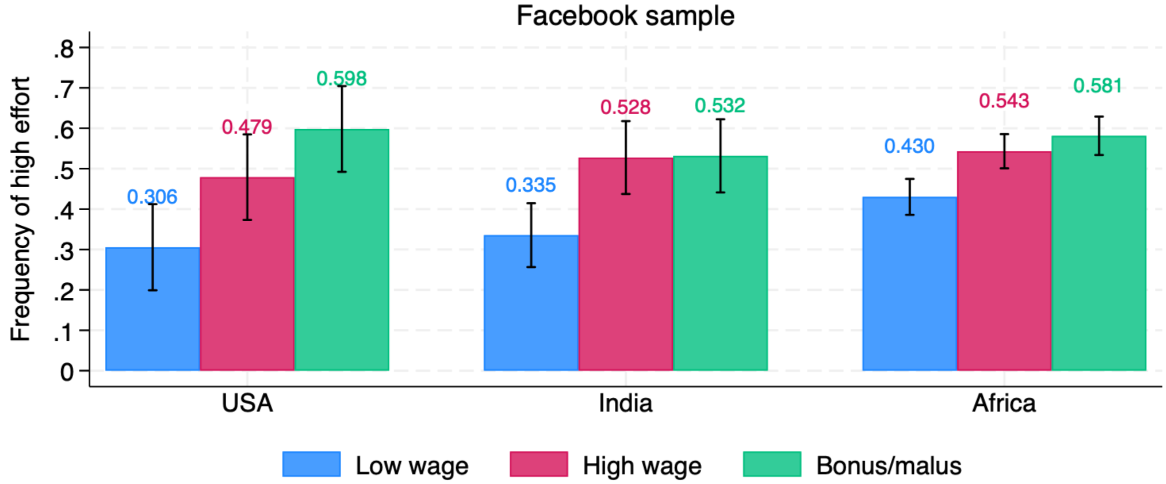
Figure A1. Rule and Law and GDP



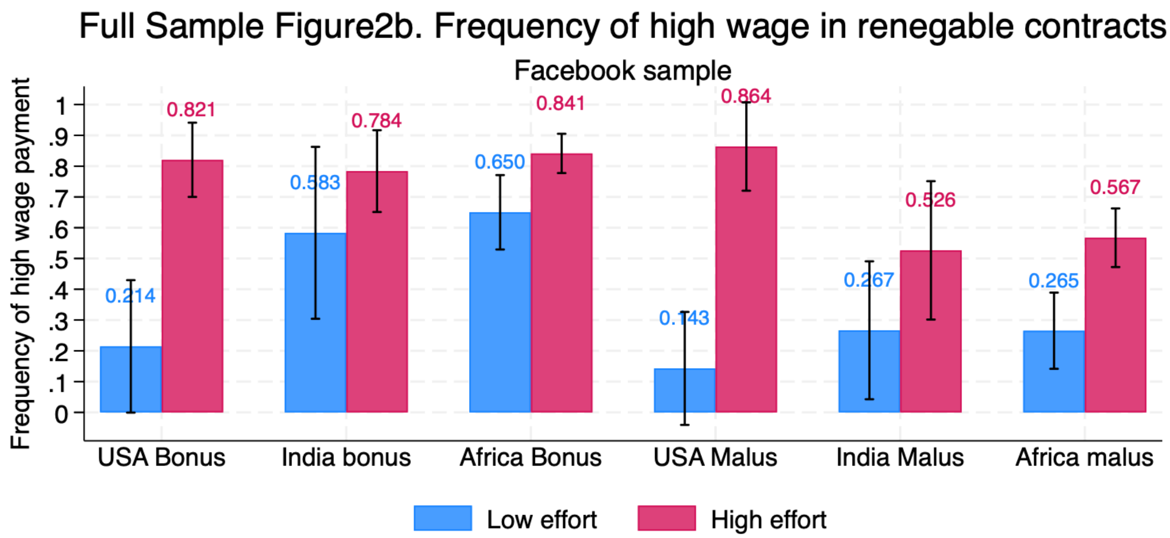
Year 2022 values of the GDP data are from the World Development Indicators of the World Bank.
The 2022 Rule of Law index is from the World Justice Project.

Full Sample Figure1b. Frequency of high effort by contract

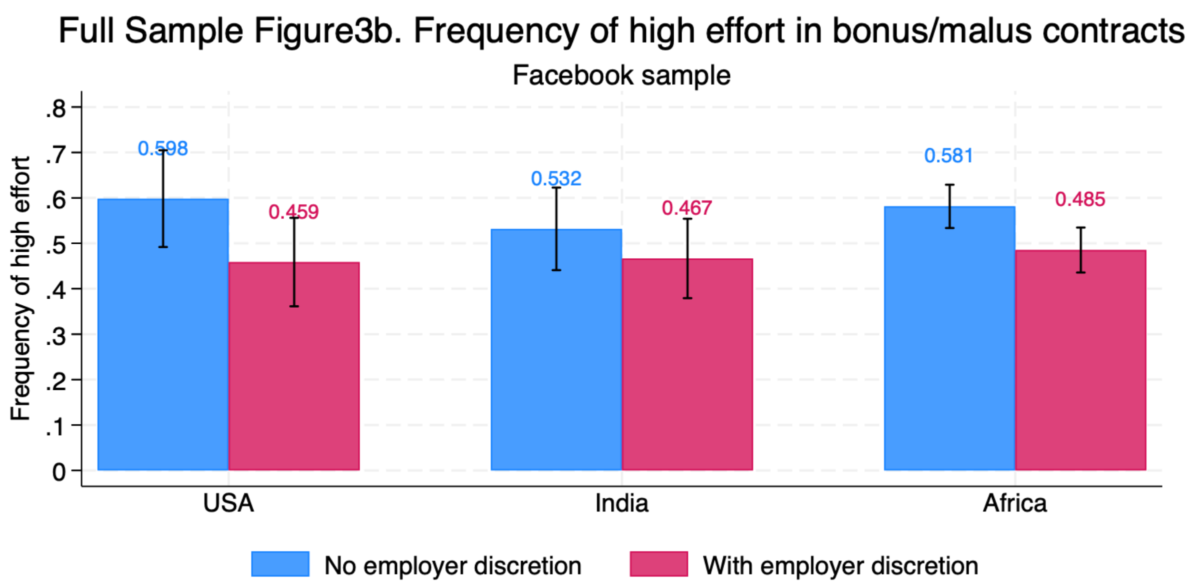
Facebook sample



Note: 95% confidence intervals shown. Offer refusal treated as equivalent to low effort.
T-test pairwise comparisons yields the following statistically significant differences at 5% or lower:
Low wage < High wage < Bonus/malus for US and Africa;
Low wage < High wage ~ Bonus/malus for India.
For low wage: US < India ~ Africa; For high wage: India ~ Africa;
Other pairwise comparisons are not statistically significant.



Note: 95% confidence intervals shown. Following the contract means blue bar=0 and red bar=1.
T-test pairwise comparisons yields the following statistically significant differences at 5% or lower:
Low vs high effort: significant for all locations except malus-renege in India (small sample);
Bonus vs malus contract: difference never significant at the 10% level.
Countries: Africa pays high wage for low effort more often than USA in bonus contract,
but pays low wage more often for high effort in malus contract.



Note: 95% confidence intervals shown. Offer refusal treated as equivalent to low effort.
T-test pairwise comparisons yields the following statistically significant differences at 5% or lower:
no significant difference between countries in any of the bars;
reduction in effort when employer has discretion is significant in the US and Africa.

Table A3. Drop-out rates during experiment

Number of rounds attended by the participant	Number of participants	Percentage	Cumulative percentage
1	142	6.5	6.5
2	96	4.4	10.9
3	114	5.2	16.1
4	120	5.5	21.6
5	163	7.5	29.0
6	266	12.2	41.2
7	285	13.0	54.2
8	1003	45.8	100

Table A4. Pooling test across African countries

	F-test	p-value	N.obs.
Figure 1b			
Effort in low wage contract	2.03	0.0888	609
Effort in high wage contract	0.65	0.6251	694
Effort in bonus contract	0.29	0.8819	461
Figure 2b			
High wage if low effort in bonus contract	0.00	n.a.	60
High wage if high effort in bonus contract	0.00	n.a.	126
High wage if low effort in malus contract	0.00	n.a.	49
High wage if high effort in malus contract	0.66	0.6209	104
Figure 3b			
High effort without employer discretion	0.29	0.8819	461
High effort with employer discretion	1.91	0.1083	474

Notes: Each test was obtained by regressing the dichotomous choice listed in column 1 on African country dummies. Only choices made by African participants are included in the regressions.

Table A5. R2 fit of wihin-sample and out-of-sample predictions					
		In-sample		Out-of-sample	
US vs India [MTurk sample]		OLS	RF	OLS	RF
	1. Selfish-rational index	0.089	0.940	0.005	0.004
	2. Conditional cooperator index	0.028	0.952	0.001	0.002
	3. Intrinsically motivated index	0.101	0.945	0.000	0.001
	4. Unconditional unmotivated index	0.023	0.946	0.006	0.001
US vs India all					
	1. Selfish-rational index	0.074	0.935	0.002	0.001
	2. Conditional cooperator index	0.027	0.950	0.001	0.001
	3. Intrinsically motivated index	0.069	0.948	0.001	0.000
	4. Unconditional unmotivated index	0.029	0.943	0.001	0.000
US vs Africa all					
	1. Selfish-rational index	0.057	0.944	0.000	0.000
	2. Conditional cooperator index	0.026	0.954	0.006	0.004
	3. Intrinsically motivated index	0.065	0.953	0.000	0.001
	4. Unconditional unmotivated index	0.022	0.943	0.001	0.002
India vs Africa all					
	1. Selfish-rational index	0.037	0.965	0.005	0.002
	2. Conditional cooperator index	0.019	0.969	0.000	0.002
	3. Intrinsically motivated index	0.072	0.960	0.022	0.019
	4. Unconditional unmotivated index	0.029	0.966	0.004	0.004

Note: The in-sample columns present the unadjusted R2 of a pooled regression of the dependent variable index on a vector of subject characteristics, without region or country dummy. OLS stands for ordinary least squares and RF for random forest. In the RF case, the column shows the R2 of a regression of the dependent variable on the random forest predictor. The out-of-sample columns is the outcome of the following three-step process: (1) regress the dependent variable on subject characteristics separately for each region, once by OLS and one by RF; (2) obtain the predictors from each regression for the whole sample; (3) create a new regressor containing, for region i, the out-of-sample predictors from the region j regression, and vice versa for the other region; (4) regress the dependent variable on these out-of-sample predictors in an OLS pooled regression. In the OLS (RF) column, the predictors are obtained using a least-squares (random forest) estimator.

Table B1. Predictors of Archetype 1 (Intrinsically motivated)												
	USA				India				Africa			
	OLS on sample data		OLS on RF predicted		OLS sample data		OLS on RF predicted		OLS sample data		OLS on RF predicted	
Male dummy	0.0769	*	0.0601	**	0.0209		0.0244		0.0463		0.0385	
Age	0.0460		0.0311		-0.0664		-0.0620		0.0305		0.0194	
Dummy ever employed	0.1609	**	0.1265	***	0.1390	**	0.1107	***	-0.0384		-0.0249	
Dummy ever self-employed	-0.0862	*	-0.0658	**	-0.0097		-0.0083		0.0397		0.0339	
Years in job (1)	0.0116		0.0126		0.0935		0.0689		-0.0050		0.0220	
Years unemployed	-0.0791		-0.0521		-0.3049	**	-0.2004	**	-0.0607		-0.0534	
Age squared	-0.0100		-0.0040		0.0379		0.0290		0.0308		0.0338	
OK to fire incompetent worker	-0.0563		-0.0263		-0.0230		-0.0144		0.0035		0.0092	
OK to cut pay of incompetent worker	0.0500		0.0292		0.0248		0.0265		0.0657		0.0494	
OK to cut bonus of incompetent worker	0.0236		0.0083		-0.0020		-0.0065		0.0641		0.0473	
OK to fire shirking worker	0.1717	***	0.1116	***	0.0044		0.0087		0.1348	***	0.1086	***
OK to cut pay of shirking worker	-0.0519		-0.0289		0.0974		0.0736	*	0.0007		0.0141	
OK to cut bonus of shirking worker	-0.1314	**	-0.0981	**	-0.0427		-0.0382		-0.0899	*	-0.0777	**
US workers perform with incentives	0.0960	*	0.0766	**	-0.0423		-0.0332		-0.0134		-0.0084	
US workers perform without incentives	0.1967	***	0.1572	***	0.0427		0.0268		0.0628		0.0342	
India workers perform with incentives	0.0826		0.0512		-0.0485		-0.0473		-0.0948	*	-0.0849	**
India workers perform without incentives	-0.0291		-0.0204		0.0925		0.0817	**	0.0223		0.0294	
S.Afr. workers perform with incentives	-0.1474	**	-0.0913	**	0.0474		0.0346		-0.0205		-0.0136	
S.Afr. workers perform without incentives	-0.0438		-0.0313		-0.0288		-0.0129		-0.0759		-0.0609	
Primary education or less dummy (2)	0.1350	***	0.1046	***	0.0351		0.0305		0.1373	***	0.1089	***
Some secondary education dummy (2)	0.0257		0.0220		0.1076	*	0.0831	**	0.1914	***	0.1399	***
Completed secondary education dummy (2)	0.0253		0.0287		0.0908	**	0.0720	**	0.0832	**	0.0573	**
Full-time student dummy (3)	-0.0615		-0.0430		-0.2774	*	-0.1969	**	0.0135		0.0063	
Short-term wage employment dummy (3)	0.0953		0.0701		-0.1709		-0.1143		0.0595		0.0408	
Permanent wage employment dummy (3)	-0.1327		-0.0923		-0.4655	**	-0.3192	**	-0.0212		-0.0498	
Self-employed dummy (3)	-0.0855		-0.0538		-0.2915		-0.1848		-0.0271		-0.0374	
Constant	-0.2525	***	-0.2359	***	0.0390		0.0539		0.0826		0.0716	*
Observations	602		602		511		511		703		703	
R-squared	0.0948		0.1255		0.0873		0.1410		0.0932		0.1414	
Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1												

Notes: All regressions are within sample. OLS on RF predicted is a regression of the predictions from a fitted Random Forest model on the regressors. (1) Years in job =0 if currently unemployed. (2) Post-secondary education is the omitted education category. (3) Unemployed is the omitted occupation category.

Table B2. Predictors of Archetype 2 (Conditional reciprocator)

	USA				India				Africa			
	OLS on sample data		OLS on RF predicted		OLS sample data		OLS on RF predicted		OLS sample data		OLS on RF predicted	
Male dummy	0.0291		0.0244		0.0372		0.0290		0.0624		0.0477	**
Age	-0.0256		-0.0229		0.0826		0.0660		-0.0180		-0.0201	
Dummy ever employed	0.1186		0.0849	*	0.0354		0.0284		0.0003		0.0038	
Dummy ever self-employed	0.0173		0.0145		-0.0093		-0.0096		-0.0313		-0.0188	
Years in job (1)	0.0203		0.0113		0.0506		0.0451		-0.0355		-0.0157	
Years unemployed	-0.0500		-0.0297		-0.0273		0.0143		0.0486		0.0348	
Age squared	-0.0214		-0.0126		0.0020		-0.0037		0.0258		0.0225	
OK to fire incompetent worker	-0.0269		-0.0046		0.0076		0.0128		0.0697		0.0559	*
OK to cut pay of incompetent worker	-0.0498		-0.0362		-0.0310		-0.0264		0.0160		0.0086	
OK to cut bonus of incompetent worker	0.1046	*	0.0703	*	-0.0198		-0.0081		0.0036		0.0155	
OK to fire shirking worker	0.1372	**	0.0979	**	0.0371		0.0302		0.0331		0.0275	
OK to cut pay of shirking worker	-0.0138		-0.0054		0.0323		0.0351		-0.0807		-0.0491	
OK to cut bonus of shirking worker	-0.0193		0.0024		0.0361		0.0130		0.0410		0.0193	
US workers perform with incentives	0.0177		0.0166		-0.0245		-0.0141		-0.0167		-0.0098	
US workers perform without incentives	0.0170		0.0250		-0.0769		-0.0548		-0.0065		-0.0129	
India workers perform with incentives	0.0774		0.0467		0.1007		0.0630		0.0152		0.0079	
India workers perform without incentives	-0.0042		0.0083		0.0838		0.0676		0.0916		0.0652	
S.Afr. workers perform with incentives	-0.0616		-0.0375		-0.0549		-0.0393		-0.0223		-0.0174	
S.Afr. workers perform without incentives	-0.0500		-0.0559		0.0255		0.0176		-0.1096	*	-0.0806	**
Primary education or less dummy (2)	0.0802	*	0.0594	**	-0.0008		0.0060		0.0461		0.0299	
Some secondary education dummy (2)	0.0533		0.0394		0.0543		0.0334		0.0379		0.0188	
Completed secondary education dummy (2)	0.0116		0.0115		0.0870	*	0.0671	**	-0.0000		-0.0051	
Full-time student dummy (3)	-0.0235		-0.0143		-0.0927		-0.0371		0.0317		0.0154	
Short-term wage employment dummy (3)	-0.0437		-0.0345		-0.1014		-0.0486		0.0236		0.0146	
Permanent wage employment dummy (3)	-0.2349	*	-0.1701	*	-0.1873		-0.0978		0.0587		0.0221	
Self-employed dummy (3)	-0.2200	*	-0.1624	**	-0.1097		-0.0374		0.0376		0.0049	
Constant	-0.0418		-0.0364		0.0622		0.0701		-0.0158		-0.0255	
Observations	602		602		511		511		703		703	
R-squared	0.0484		0.0672		0.0395		0.0597		0.0276		0.0399	

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Notes: All regressions are within sample. OLS on RF predicted is a regression of the predictions from a fitted Random Forest model on the regressors. (1) Years in job =0 if currently unemployed. (2) Post-secondary education is the omitted education category. (3) Unemployed is the omitted occupation category.

Table B3. Predictors of Archetype 3 (Selfish rational)												
	USA				India				Africa			
	OLS on sample data		OLS on RF predicted		OLS sample data		OLS on RF predicted		OLS sample data		OLS on RF predicted	
Male dummy	0.0053		0.0055		-0.0155		-0.0121		-0.0168		-0.0139	
Age	0.0029		0.0114		0.0638		0.0641		-0.1417	**	-0.1112	**
Dummy ever employed	-0.0353		-0.0244		-0.1312	**	-0.0992	***	0.0476		0.0350	
Dummy ever self-employed	0.0422		0.0314		-0.0261		-0.0181		-0.0447		-0.0283	
Years in job (1)	-0.0507		-0.0418		-0.0390		-0.0389		-0.0359		-0.0286	
Years unemployed	0.0460		0.0382		0.0693		0.0427		-0.0113		0.0026	
Age squared	-0.0190		-0.0217		-0.0299		-0.0148		0.0358		0.0164	
OK to fire incompetent worker	0.0880		0.0738	**	-0.0014		-0.0062		-0.0011		-0.0018	
OK to cut pay of incompetent worker	-0.1617	***	-0.1209	***	-0.0190		-0.0134		-0.0009		-0.0026	
OK to cut bonus of incompetent worker	0.0138		0.0101		-0.0331		-0.0225		-0.0151		-0.0105	
OK to fire shirking worker	-0.0645		-0.0323		-0.0176		-0.0223		0.0027		-0.0051	
OK to cut pay of shirking worker	0.0809		0.0511		-0.0209		-0.0180		-0.0751		-0.0602	*
OK to cut bonus of shirking worker	0.1430	**	0.1149	***	0.0719		0.0431		0.0352		0.0339	
US workers perform with incentives	-0.0437		-0.0385		-0.0211		-0.0022		-0.0422		-0.0261	
US workers perform without incentives	-0.1014	*	-0.0711	*	-0.0295		-0.0262		0.0225		0.0186	
India workers perform with incentives	-0.0491		-0.0367		0.1347	**	0.0975	**	0.0726		0.0504	
India workers perform without incentives	0.0428		0.0443		-0.0746		-0.0506		-0.0090		-0.0112	
S.Afr. workers perform with incentives	0.0670		0.0484		-0.0529		-0.0394		0.0373		0.0237	
S.Afr. workers perform without incentives	-0.0378		-0.0522		0.0065		-0.0042		-0.0277		-0.0164	
Primary education or less dummy (2)	-0.0887	**	-0.0713	***	-0.1330	***	-0.1059	***	-0.0168		-0.0174	
Some secondary education dummy (2)	-0.0059		-0.0018		0.0124		0.0066		-0.1083	**	-0.0801	***
Completed secondary education dummy (2)	-0.0400		-0.0328		-0.0652		-0.0590	**	-0.0896	**	-0.0669	**
Full-time student dummy (3)	-0.1074		-0.0769		-0.0297		-0.0218		-0.1391	*	-0.1046	**
Short-term wage employment dummy (3)	-0.0645		-0.0592		0.0146		0.0005		-0.0699		-0.0480	
Permanent wage employment dummy (3)	0.0277		0.0195		0.1166		0.0761		-0.0179		0.0106	
Self-employed dummy (3)	-0.0308		-0.0254		0.0678		0.0376		-0.0052		0.0086	
Constant	0.1786	***	0.1761	***	-0.0900		-0.0968	**	-0.1157	**	-0.0972	***
Observations	602		602		511		511		703		703	
R-squared	0.0709		0.1118		0.0674		0.1071		0.0481		0.0749	

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Notes: All regressions are within sample. OLS on RF predicted is a regression of the predictions from a fitted Random Forest model on the regressors. (1) Years in job =0 if currently unemployed. (2) Post-secondary education is the omitted education category. (3) Unemployed is the omitted occupation category.

Table B4. Predictors of Archetype 4 (Intrinsically unmotivated)

	USA				India				Africa			
	OLS on sample data		OLS on RF predicted		OLS sample data		OLS on RF predicted		OLS sample data		OLS on RF predicted	
Male dummy	-0.0687		-0.0520	*	-0.0187		-0.0223		-0.0297		-0.0292	
Age	0.0495		0.0258		-0.0299		-0.0088		0.0737		0.0573	
Dummy ever employed	-0.0852		-0.0605		0.0112		0.0048		-0.0155		-0.0152	
Dummy ever self-employed	-0.0048		-0.0010		0.0100		0.0039		0.0163		0.0057	
Years in job (1)	-0.0327		-0.0244		0.0295		0.0237		0.0253		0.0145	
Years unemployed	0.0234		0.0164		0.2544	*	0.1765	*	0.0162		0.0164	
Age squared	-0.0157		-0.0013		0.0162		0.0207		-0.0362		-0.0226	
OK to fire incompetent worker	-0.0892		-0.0886	**	-0.0220		-0.0214		0.0095		0.0034	
OK to cut pay of incompetent worker	0.1457	**	0.1029	***	0.0344		0.0208		-0.0471		-0.0408	
OK to cut bonus of incompetent worker	-0.0252		-0.0030		0.0262		0.0256		-0.0239		-0.0171	
OK to fire shirking worker	-0.0165		-0.0133		0.0307		0.0270		-0.1175	**	-0.0987	***
OK to cut pay of shirking worker	-0.0794		-0.0564		-0.1432	*	-0.1071	**	0.0264		0.0213	
OK to cut bonus of shirking worker	-0.0138		-0.0307		0.0523		0.0379		0.0608		0.0442	
US workers perform with incentives	-0.0998		-0.0726	*	0.0650		0.0384		0.0159		0.0069	
US workers perform without incentives	-0.0217		-0.0252		0.0212		0.0260		-0.0430		-0.0327	
India workers perform with incentives	0.0564		0.0629		-0.1319	*	-0.0918	**	0.0036		0.0056	
India workers perform without incentives	-0.0150		-0.0265		-0.0089		-0.0142		-0.0438		-0.0263	
S.Afr. workers perform with incentives	-0.0378		-0.0506		0.0617		0.0444		0.0102		0.0150	
S.Afr. workers perform without incentives	0.0575		0.0737		-0.0333		-0.0328		0.0933	*	0.0600	*
Primary education or less dummy (2)	-0.0333		-0.0281		0.0580		0.0457		-0.0554		-0.0434	*
Some secondary education dummy (2)	-0.0405		-0.0329		-0.1208	**	-0.0899	**	-0.0856	*	-0.0654	**
Completed secondary education dummy (2)	-0.0015		-0.0069		-0.0342		-0.0195		-0.0107		-0.0014	
Full-time student dummy (3)	0.1686		0.1446	**	0.3834	**	0.2757	***	0.0719		0.0652	
Short-term wage employment dummy (3)	0.0230		0.0323		0.1809		0.1202		-0.0504		-0.0371	
Permanent wage employment dummy (3)	0.0880		0.0645		0.4205	*	0.2938	**	-0.0119		0.0028	
Self-employed dummy (3)	0.1203		0.0946		0.2755		0.1839		-0.0438		-0.0330	
Constant	0.1362	**	0.1200	***	0.0380		0.0374		-0.0338		-0.0360	
Observations	602		602		511		511		703		703	
R-squared	0.0481		0.0796		0.0564		0.0822		0.0394		0.0642	

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Notes: All regressions are within sample. OLS on RF predicted is a regression of the predictions from a fitted Random Forest model on the regressors. (1) Years in job =0 if currently unemployed. (2) Post-secondary education is the omitted education category. (3) Unemployed is the omitted occupation category.

Informed consent and start-up questionnaire

Welcome

Welcome to this study! Thank you for your interest! This is a study by Stanford University on decision-making. We first ask you a few questions. Later on, qualified participants are invited for an online experiment. In this experiment we present you with a situation and ask you to make choices. In this experiment, your choices and the choices of the other participants will determine how much you earn. This is explained later on.

We ask your consent for both this qualifying survey as well as the online experiment.

A few things you should know:

- This questionnaire takes about 5 minutes. The experiment happens later on as a separate HIT. Separate invitations are sent out for this. The experiment lasts about 20 minutes.
- Your earnings are kept private and are paid through MTurk. By completing this survey you will earn $1. By participating in the experiment, you can earn additional money.
- We will not reveal who you are to other participants.
- We never deceive participants. For example, if we inform you that another participant is making a choice on which you can react, this is indeed the case. We keep our promises made to participants. For example, if we promise a certain payment, participants will indeed receive it.
- The data collected in this session will be used for the research study and might be published, both online or offline. No data that allows people to identify you will ever be published.
- Participation is voluntary and you can withdraw at any time. If anything makes you feel uncomfortable, let us know. Our study has been reviewed and approved by the Stanford University Human Subjects Research and IRB Committee. You can find the details at the bottom of this page.

In case you want to get in touch with us, you can message us through the internal messaging system of MTurk, or through stanfordseedstudy@gmail.com. Click on the Next button to proceed.

Informed consent

The principal researcher in this project is Marcel Fafchamps, Senior Research Fellow at Stanford University. If you agree to participate in this project, the research will be written up as one or more research articles. These articles, as well as other publications resulting from this project, might be published online as well as in print. If you have any questions regarding the research or your participation, you can contact the Protocol Administrator, who will answer your questions.

The administrator's phone number is + 1-650-736-1436. Email: jeduarte@stanford.edu. If at any time you have comments or concerns regarding the conduct of the research, or questions about your rights as a research subject, you should contact the Stanford University Institutional Review Board /IRB. The phone number for the IRB is + 1 650-723-2480. Or, you can write to the Research Compliance Office, Stanford University, 3000 El Camino Real, Five Palo Alto Square, 4th Floor, Palo Alto, CA 94306 or by sending an email to Mr. A. Bailey: afbailey@stanford.edu.

Page Break

Please read the following statement:

'I have read the description of the study and agree to take part in this study. I have had the opportunity to ask questions and have received satisfactory answers to these. I understand that I can withdraw from the study at any time, by indicating this to the researchers. I understand that the project has been reviewed and approved by Stanford University Human Subjects Research and IRB Committee. I understand that the collected data might be used in dissertations and other publications, both on line as well as printed, and that only data that cannot be used to identify me personally will be published. I am aware how to raise a concern and make a complaint.'

Please tick the following box to indicate that you agree with the above statement and to proceed with the survey.

o I agree (1)

Demographics and employment history

Before starting this study we would like to ask you a couple of questions.

Q6 What is your age in years? _______

Q7 What is your gender?

o Male (1)

o Female (2)

o Other (3)

o Decline to state (4)

Q8 What is the highest level of education you have reached?

o No education (1)

o Some Primary (2)

o Completed Primary (3)

o Some Secondary (4)

o Completed Secondary (5)

o Post-Secondary (6)

If the highest level of education you have reached is Post-Secondary:

Q32 How many years of post-secondary education did you complete? _____

Q33 What best represents your current employment status?

o in permanent wage employment (1)

- o in fixed-term/short-term wage employment (2)
- o self-employed without paid employees (3)
- o self-employed with paid employees (4)
- o unemployed/not-working (5)

If your current employment status is not unemployed/not-working:

Q9 What occupational category best describes your employment?

- o Forestry, fishing, hunting or agriculture support (1)
- o Real estate or rental and leasing (2)
- o Mining (3)
- o Professional, scientific or technical services (4)
- o Utilities (5)
- o Management of companies or enterprises (6)
- o Construction (7)
- o Admin, support, waste management or remediation services (8)
- o Manufacturing (9)
- o Educational services (10)
- o Wholesale trade (11)
- o Health care or social assistance (12)
- o Retail trade (13)
- o Arts, entertainment or recreation (14)
- o Transportation or warehousing (15)
- o Accommodation or food services (16)
- o Information (17)
- o Other services (except public administration) (18)
- o Finance or insurance (19)
- o Unclassified establishments (20)
- o Public sector (21)

Q34 For how long have you been working in your current job? (Answer in years) ________

If your current employment status is permanent wage employment or fixed-term/short-term wage employment:

Q41 For how long have you been working in wage employment in total (including your current and your past jobs)?

- o Less than 6 months (1)
- o Between 6 months and 1 year (2)
- o Between 1 and 2 years (3)
- o Between 2 and 3 years (4)
- o Between 3 and 5 years (5)
- o Between 5 and 10 years (6)
- o More than 10 years (7)

If your current employment status is self-employed with paid employees

Q35 How many employees does your firm have? ______

If your current employment status is unemployed/not-working

Q36 For how long have you been unemployed?

o Less than 6 months (1)

o Between 6 months and 1 year (2)

o Between 1 and 2 years (3)

o Between 2 and 3 years (4)

o Between 3 and 5 years (5)

o Between 5 and 10 years (6)

o More than 10 years (7)

If your current employment status is not in permanent wage employment:

Q37 Have you ever been in permanent wage employment in the past?

o Yes (1)

o No (2)

If Yes:

Q38 For how long in total have you been in permanent wage employment?

o Less than 6 months (1)

o Between 6 months and 1 year (2)

o Between 1 and 2 years (3)

o Between 2 and 3 years (4)

o Between 3 and 5 years (5)

o Between 5 and 10 years (6)

o More than 10 years (7)

If your current employment status is self-employed with or without paid employees:

Q39 Have you ever been self-employed in the past?

o Yes (1)

o No (2)

If yes:

Q40 For how long have you been self-employed?

o Less than 6 months (1)

o Between 6 months and 1 year (2)

o Between 1 and 2 years (3)

o Between 2 and 3 years (4)

o Between 3 and 5 years (5)

o Between 5 and 10 years (6)

o More than 10 years (7)

Attitudes

We will present you with 6 situations. Please indicate, on a scale from 0 to 10, whether you think the decision of the employer is fully acceptable (10) or fully unacceptable (0).

Q15 Worker A is hired to perform a task for which he/she claims to be qualified. After a week on the job, it becomes clear that A is unable to perform the task. Worker A is laid off by the employer. On a scale from 0 to 10, where 0 means fully unacceptable and 10 means fully acceptable, is the employer's decision acceptable or unacceptable? 0 1 2 3 4 5 6 7 8 9 10

Q46 Worker A is hired to perform a task for which he/she claims to be qualified. After a week on the job, it becomes clear that A is unable to perform the task. The salary of worker A is reduced by 30% by the employer. On a scale from 0 to 10, where 0 means fully unacceptable and 10 means fully acceptable, is the employer's decision acceptable or unacceptable? 0 1 2 3 4 5 6 7 8 9 10

Q43 Worker A is hired to perform a task for which he/she claims to be qualified. After a week on the job, it becomes clear that A is unable to perform the task. Worker A is denied a 30% bonus that other similar workers receive. On a scale from 0 to 10, where 0 means fully unacceptable and 10 means fully acceptable, is the employer's decision acceptable or unacceptable? 0 1 2 3 4 5 6 7 8 9 10

Q18 Worker A is hired to perform a task for which he/she claims to be qualified. After a month on the job, it becomes clear that A is able to perform the task, but is frequently caught shirking. Worker A is laid off by the employer. On a scale from 0 to 10, where 0 means fully unacceptable and 10 means fully acceptable, is the employer's decision acceptable or unacceptable? 0 1 2 3 4 5 6 7 8 9 10

Q17 Worker A is hired to perform a task for which he/she claims to be qualified. After a month on the job, it becomes clear that A is able to perform the task, but is frequently caught shirking. Worker A is denied a 30% bonus that other similar workers receive. On a scale from 0 to 10, where 0 means fully unacceptable and 10 means fully acceptable, is the employer's decision acceptable or unacceptable? 0 1 2 3 4 5 6 7 8 9 10

Q16 Worker A is hired to perform a task for which he/she claims to be qualified. After a month on the job, it becomes clear that A is able to perform the task, but is frequently caught shirking. The salary of worker A is reduced by 30% by the employer. On a scale from 0 to 10, where 0 means fully unacceptable and 10 means fully acceptable, is the employer's decision acceptable or unacceptable? 0 1 2 3 4 5 6 7 8 9 10

Statements

Q20 We will now present you with 6 statements. Please indicate on a scale from 0 to 10 whether you agree with the statement, where 0 means fully disagree and 10 means fully agree.

Q21 Workers from the United States can be trusted to exert high effort if their earnings and continued employment depend on their performance on the job.

Fully disagree Disagree Somewhat disagree Neither agree nor disagree Somewhat agree Agree Fully agree

Q22 Workers from the United States can be trusted to exert high effort even if their earnings and continued employment do not depend on their performance on the job.

Fully disagree Disagree Somewhat disagree Neither agree nor disagree Somewhat agree Agree Fully agree

Q23 Workers from India can be trusted to exert high effort if their earnings and continued employment depend on their performance on the job.

Fully disagree Disagree Somewhat disagree Neither agree nor disagree Somewhat agree Agree Fully agree

Q24 Workers from India can be trusted to exert high effort even if their earnings and continued employment do not depend on their performance on the job.

Fully disagree Disagree Somewhat disagree Neither agree nor disagree Somewhat agree Agree Fully agree

Q25 Workers from South Africa can be trusted to exert high effort if their earnings and continued employment depend on their performance on the job.

Fully disagree Disagree Somewhat disagree Neither agree nor disagree Somewhat agree Agree Fully agree

Q26 Workers from South Africa can be trusted to exert high effort even if their earnings and continued employment do not depend on their performance on the job.

Fully disagree Disagree Somewhat disagree Neither agree nor disagree Somewhat agree Agree Fully agree

Invitation to the experiment

Interest to participate in an experimental study As part of this study, we are conducting an online experiment. We ask participants in this experiment to make choices in a described situation. By participating in this experiment you can earn extra money. Your choices and the choices of the other participants determine your earnings. The experiment lasts about 20 minutes and the average expected earnings are $3.00-$4.00, which will be paid out through MTurk. The minimum earnings are $2.00. These experiments will happen at set times. If you are invited to this experiment, we will ask you to open the experiment website within 5 minutes of the set time. Invitations will be sent through the MTurk internal messaging system.

Q28 Would you be interested in participating in this experiment?

o Yes (1)

o No (2)

If Yes:

Q29 Please indicate your preferred times for participating in this experiment. To increase your chances of being able to participate, please select a minimum of two times you are likely to be available to participate in this experiment. This time is not guaranteed. You might receive an invitation for a session at a different time than posted. All times refer to Indian [US] time (UTC+5:30).

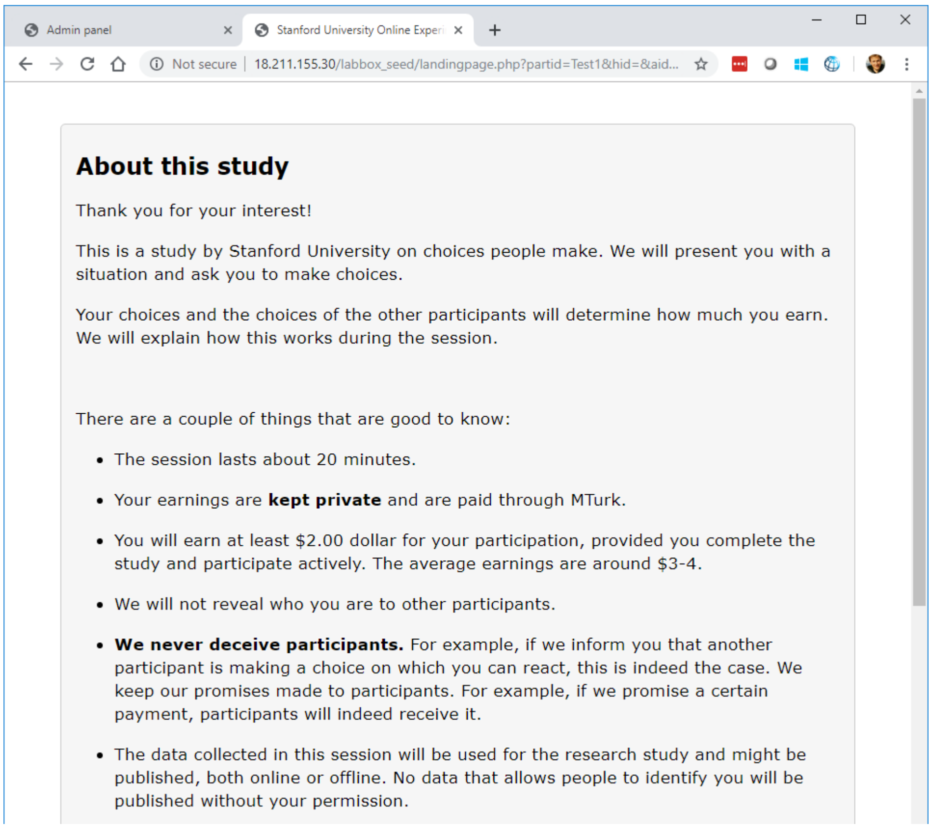
- o Saturday Nov 10 8.30 PM Delhi time (132)
- o Saturday Nov 10 9.30 PM Delhi time (133)
- o Sunday Nov 11 8.30 PM Delhi time (134)
- o Sunday Nov 11 9.30 PM Delhi time (135)

Q30 In our invitation, we will specify the time and date of the session. A separate HIT will be made available 5 minutes before the stated time in the invitation. You will need to accept the HIT and open the link within 5 minutes of the stated time. After these 5 minutes you can no longer participate.

Q47 You indicated that none of the times mentioned work for you. What time could work for you instead? _____

Thank you for your participation! Your completion code is CARFAX50. Please enter this in the MTurk field. You can now close this window.

Screen shots for the online interactive game



Welcome. Please press the button to proceed.

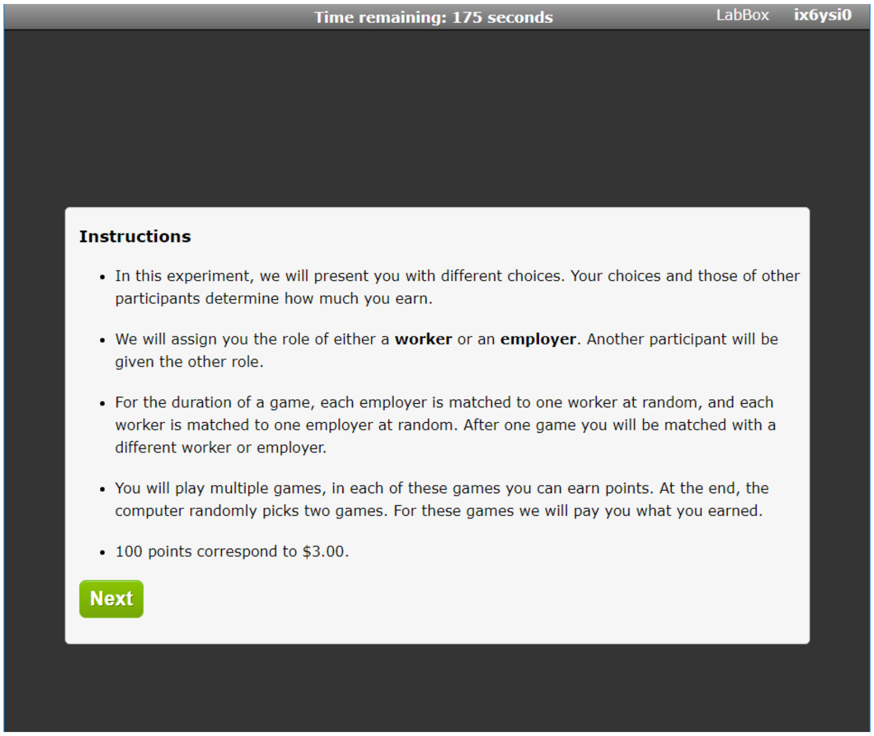
Click here to proceed

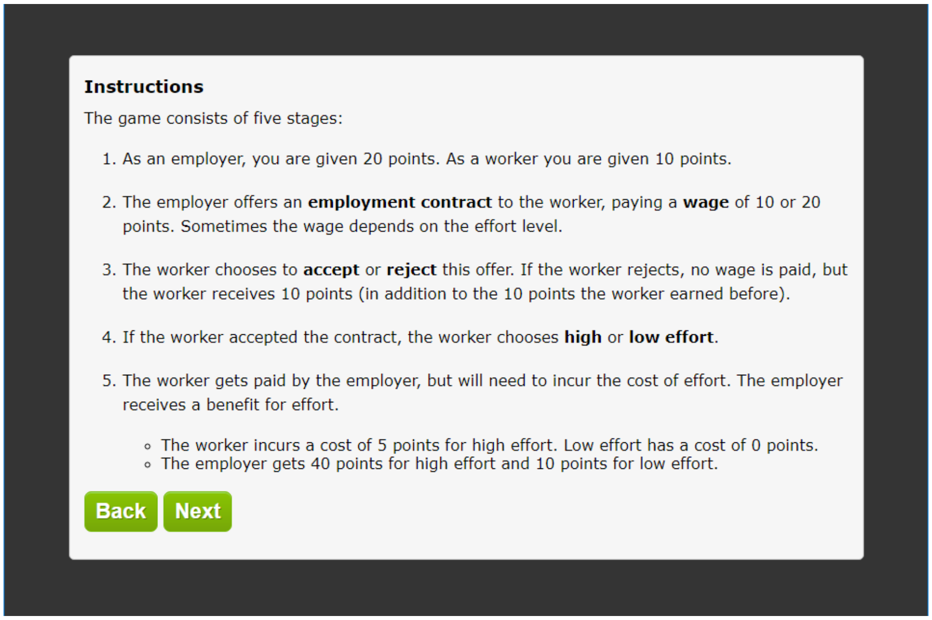
Good to know

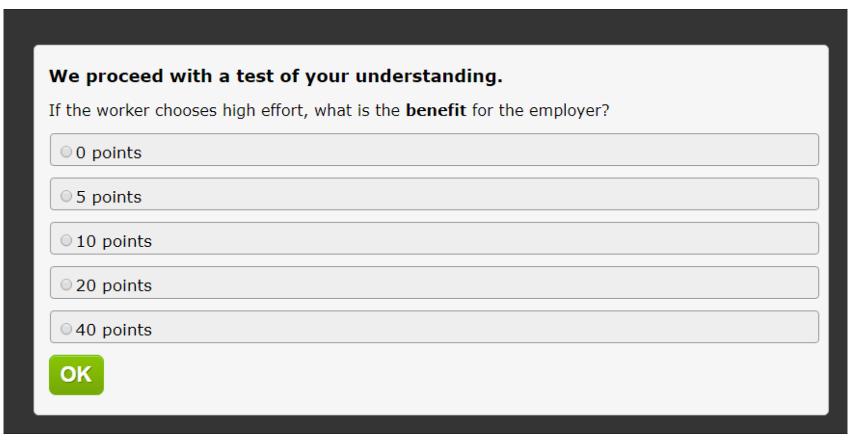
- If you accidentally close the experimental window, you can re-open it by coming back to this page. You can then resume where you have left.
- You can save this page as a bookmark.
- If you have any questions, you can get in touch with us through stanfordseedstudy@gmail.com or through the internal messaging system of MTurk.

Press Join to participate

Join


Time remaining: 175 seconds
LabBox
ix6ysi0
Instructions
In this experiment, we will present you with different choices. Your choices and those of other participants determine how much you earn.
We will assign you the role of either a worker or an employer. Another participant will be given the other role.
For the duration of a game, each employer is matched to one worker at random, and each worker is matched to one employer at random. After one game you will be matched with a different worker or employer.
You will play multiple games, in each of these games you can earn points. At the end, the computer randomly picks two games. For these games we will pay you what you earned.
100 points correspond to $3.00.
Next


Instructions
The game consists of five stages:
1. As an employer, you are given 20 points. As a worker you are given 10 points.
2. The employer offers an employment contract to the worker, paying a wage of 10 or 20 points. Sometimes the wage depends on the effort level.
3. The worker chooses to accept or reject this offer. If the worker rejects, no wage is paid, but the worker receives 10 points (in addition to the 10 points the worker earned before).
4. If the worker accepted the contract, the worker chooses high or low effort.
5. The worker gets paid by the employer, but will need to incur the cost of effort. The employer receives a benefit for effort.
The worker incurs a cost of 5 points for high effort. Low effort has a cost of 0 points.
The employer gets 40 points for high effort and 10 points for low effort.
Back
Next


We proceed with a test of your understanding.
If the worker chooses high effort, what is the benefit for the employer?
0 points
5 points
10 points
20 points
40 points
OK

Your answer was correct

The **benefit** for the employer is

- 10 points for low effort,
- 40 points for high effort.

OK

Game 1

Please wait.

We are currently matching you to another player.

Note: We will try to match you to another player as soon as possible, but we have to wait until other partipants are ready. The waiting time should not be more than a couple of minutes.

You are a worker

In the following game you are a worker. You are randomly matched with an employer.

The employer is from the **United States (USA)**.

In this game, you start with **10 points**.

The employer can now offer a contract to you.

OK

You are an employer

In the following game you are an employer. You are randomly matched with a worker.

The worker is from the **United States (USA)**.

For this game, we have given you **20 points** to start with.

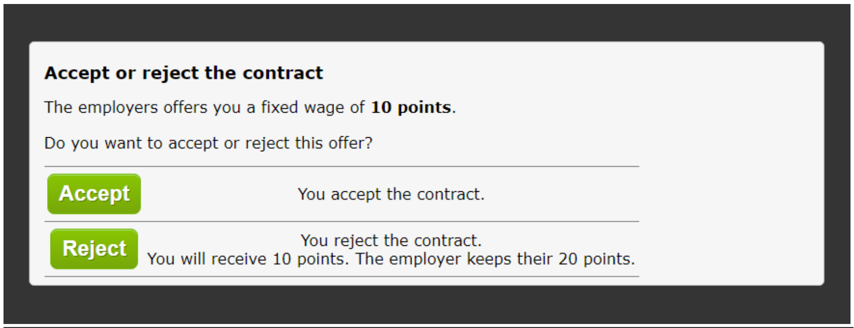
In the next screen you can offer a contract to this worker.

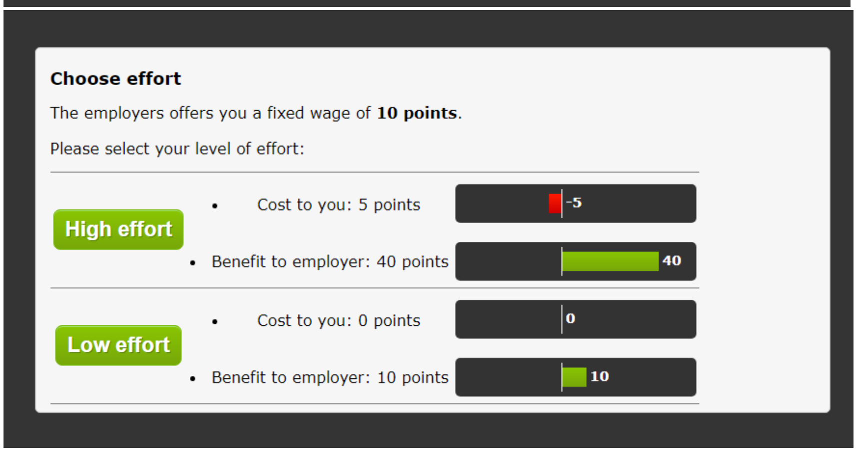
OK

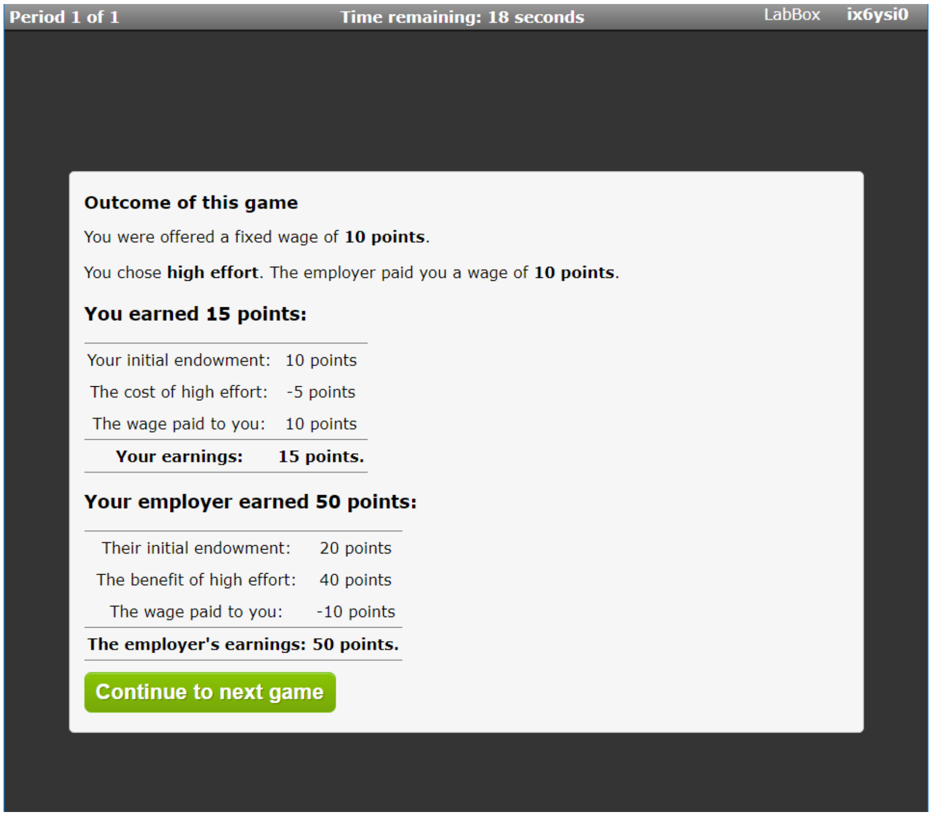
What contract do you want to offer to the worker?

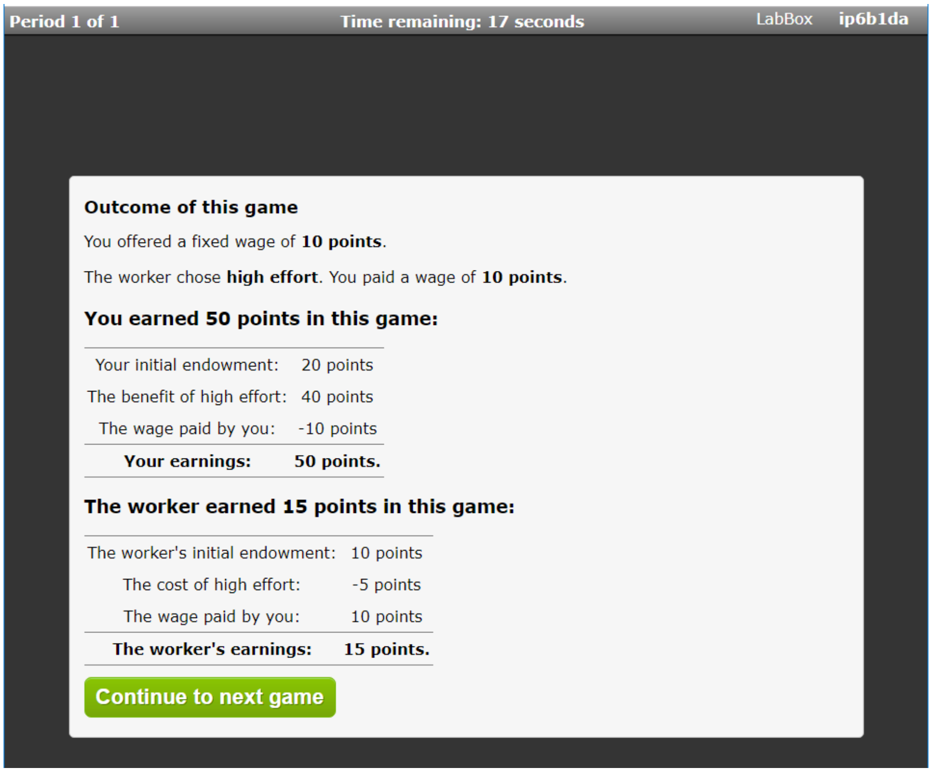
Your current balance in this game: 20 points.

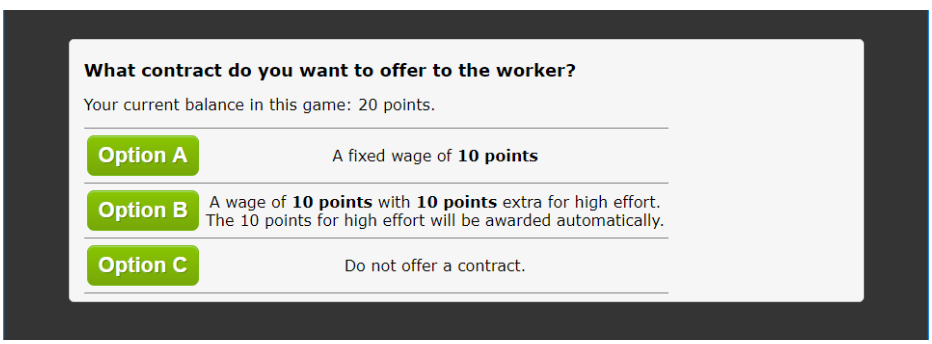
Option A	A fixed wage of **10 points**
Option B	Do not offer a contract.

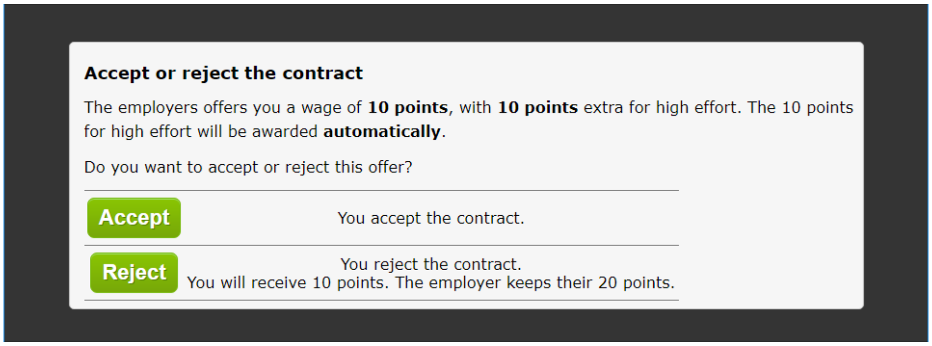

Accept or reject the contract
The employers offers you a fixed wage of 10 points.
Do you want to accept or reject this offer?
Accept
You accept the contract.
Reject
You reject the contract.
You will receive 10 points. The employer keeps their 20 points.

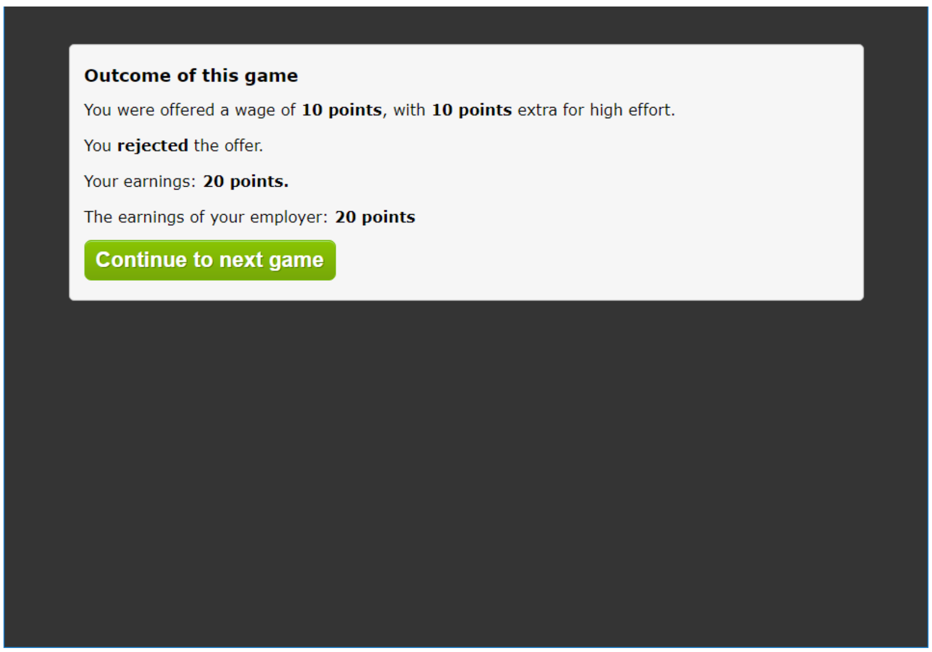

Choose effort
The employers offers you a fixed wage of 10 points.
Please select your level of effort:
High effort
Cost to you: 5 points
-5
Benefit to employer: 40 points
40
Low effort
Cost to you: 0 points
0
Benefit to employer: 10 points
10

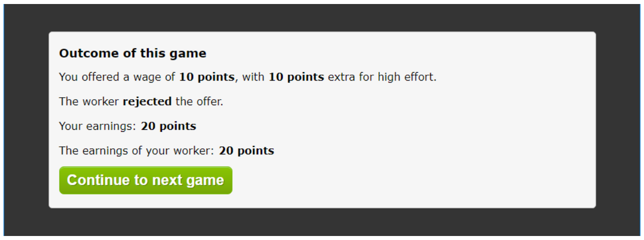

Period 1 of 1
Time remaining: 18 seconds
LabBox
ix6ysi0
Outcome of this game
You were offered a fixed wage of 10 points.
You chose high effort. The employer paid you a wage of 10 points.
You earned 15 points:
Your initial endowment: 10 points
The cost of high effort: -5 points
The wage paid to you: 10 points
Your earnings: 15 points.
Your employer earned 50 points:
Their initial endowment: 20 points
The benefit of high effort: 40 points
The wage paid to you: -10 points
The employer's earnings: 50 points.
Continue to next game

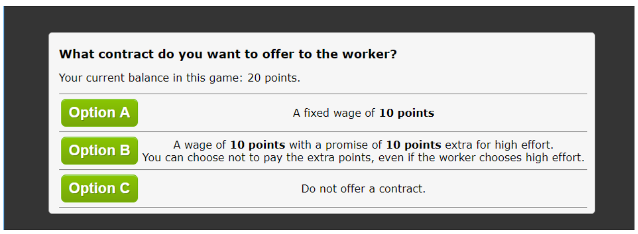

Period 1 of 1
Time remaining: 17 seconds
LabBox
ip6b1da
Outcome of this game
You offered a fixed wage of 10 points.
The worker chose high effort. You paid a wage of 10 points.
You earned 50 points in this game:
Your initial endowment: 20 points
The benefit of high effort: 40 points
The wage paid by you: -10 points
Your earnings: 50 points.
The worker earned 15 points in this game:
The worker's initial endowment: 10 points
The cost of high effort: -5 points
The wage paid by you: 10 points
The worker's earnings: 15 points.
Continue to next game

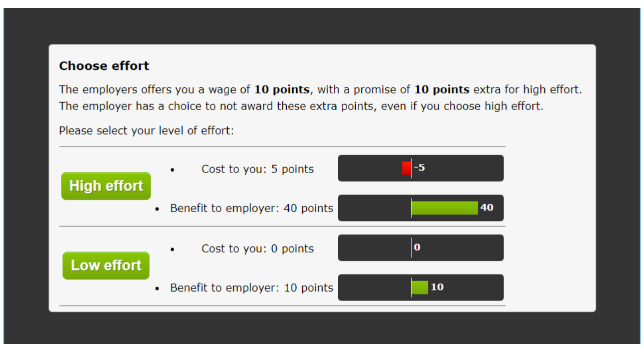

What contract do you want to offer to the worker?
Your current balance in this game: 20 points.
Option A
A fixed wage of 10 points
Option B
A wage of 10 points with 10 points extra for high effort.
The 10 points for high effort will be awarded automatically.
Option C
Do not offer a contract.

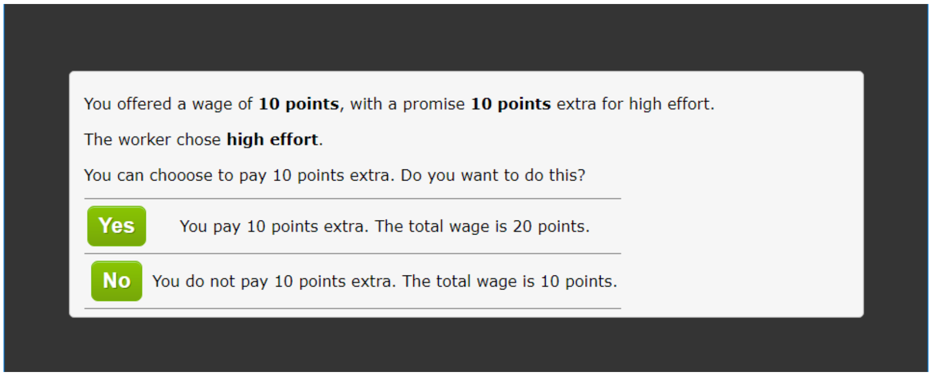

Accept or reject the contract
The employers offers you a wage of 10 points, with 10 points extra for high effort. The 10 points for high effort will be awarded automatically.
Do you want to accept or reject this offer?
Accept
You accept the contract.
Reject
You reject the contract.
You will receive 10 points. The employer keeps their 20 points.

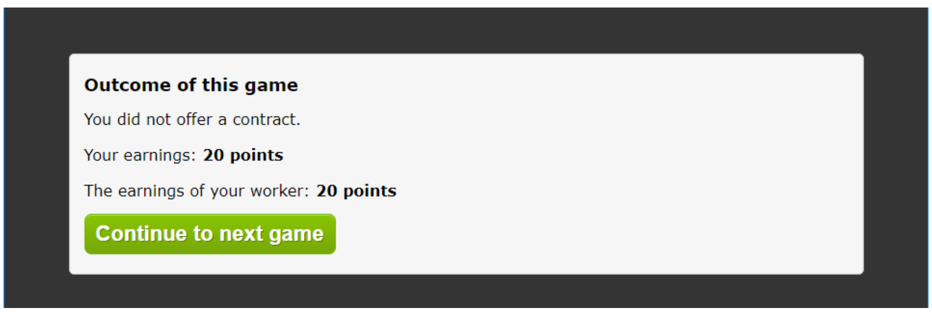

Outcome of this game
You were offered a wage of 10 points, with 10 points extra for high effort.
You rejected the offer.
Your earnings: 20 points.
The earnings of your employer: 20 points
Continue to next game

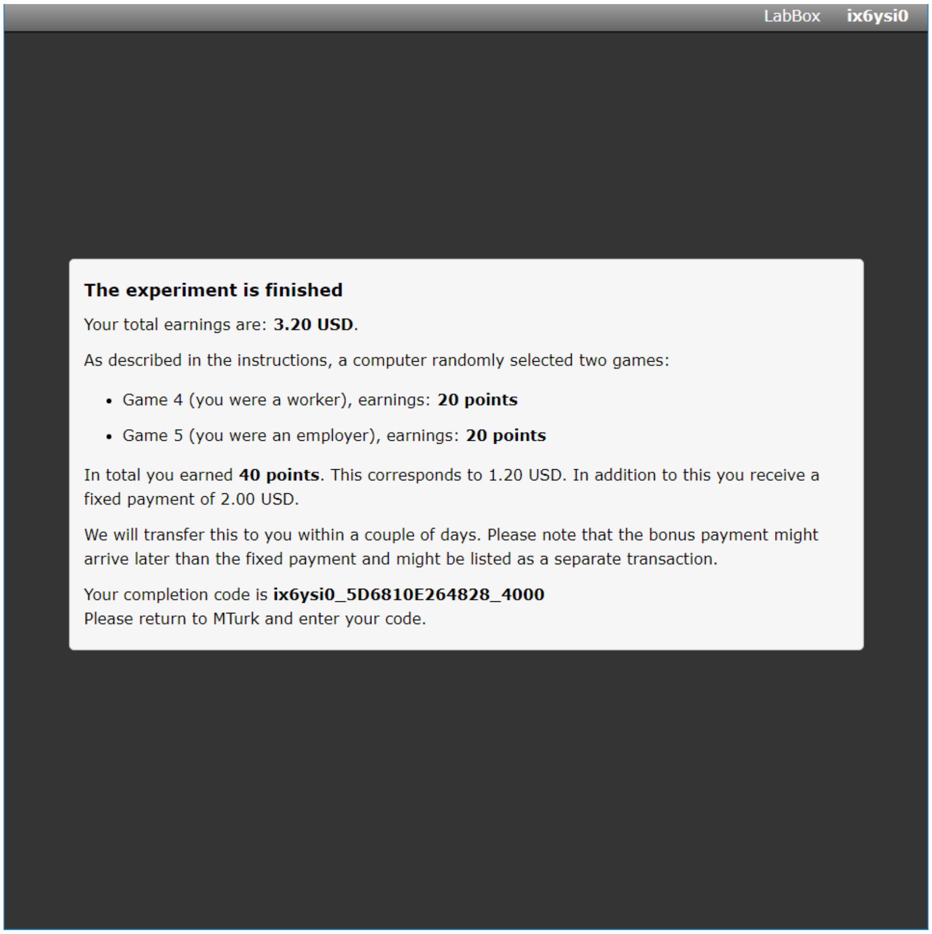

Outcome of this game
You offered a wage of 10 points, with 10 points extra for high effort.
The worker rejected the offer.
Your earnings: 20 points
The earnings of your worker: 20 points
Continue to next game

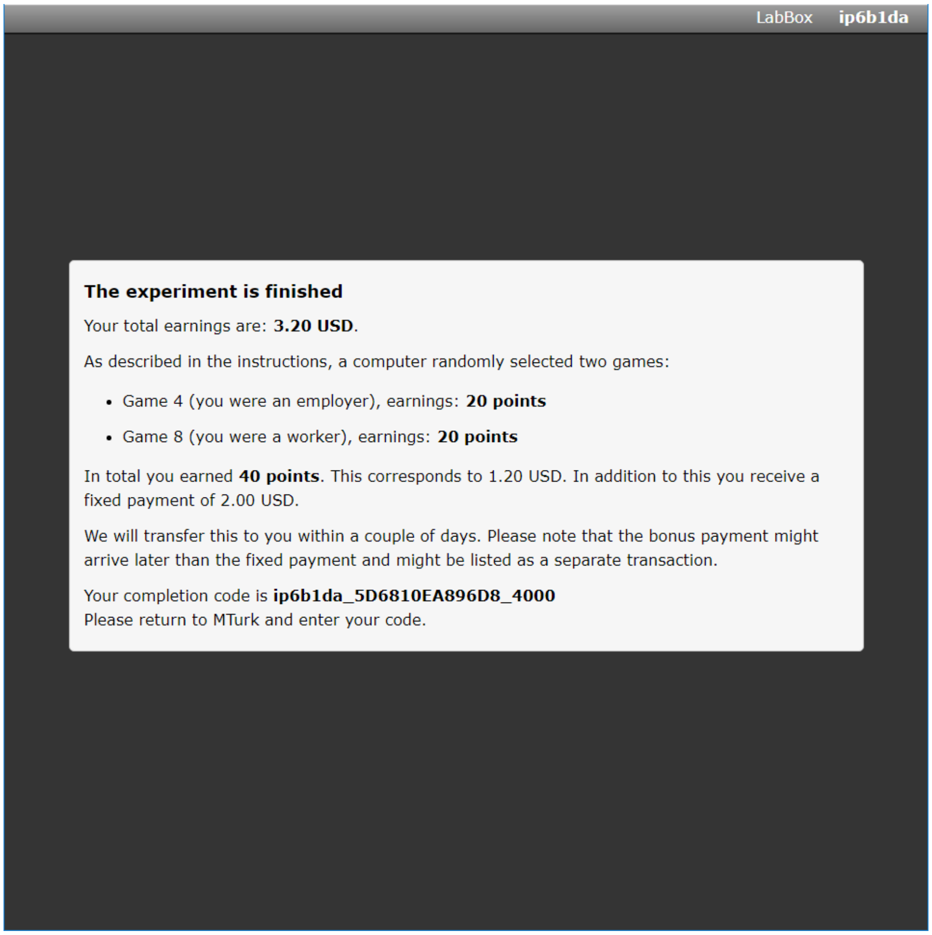

What contract do you want to offer to the worker?
Your current balance in this game: 20 points.
Option A
A fixed wage of 10 points
Option B
A wage of 10 points with a promise of 10 points extra for high effort.
You can choose not to pay the extra points, even if the worker chooses high effort.
Option C
Do not offer a contract.


Choose effort
The employers offers you a wage of 10 points, with a promise of 10 points extra for high effort.
The employer has a choice to not award these extra points, even if you choose high effort.
Please select your level of effort:
High effort
Cost to you: 5 points
-5
Benefit to employer: 40 points
40
Low effort
Cost to you: 0 points
0
Benefit to employer: 10 points
10


You offered a wage of **10 points**, with a promise **10 points** extra for high effort.
The worker chose **high effort**.
You can chooose to pay 10 points extra. Do you want to do this?
Yes
You pay 10 points extra. The total wage is 20 points.
No
You do not pay 10 points extra. The total wage is 10 points.


Outcome of this game
You did not offer a contract.
Your earnings: **20 points**
The earnings of your worker: **20 points**
Continue to next game


LabBox ix6ysi0
The experiment is finished
Your total earnings are: 3.20 USD.
As described in the instructions, a computer randomly selected two games:
• Game 4 (you were a worker), earnings: 20 points
• Game 5 (you were an employer), earnings: 20 points
In total you earned 40 points. This corresponds to 1.20 USD. In addition to this you receive a fixed payment of 2.00 USD.
We will transfer this to you within a couple of days. Please note that the bonus payment might arrive later than the fixed payment and might be listed as a separate transaction.
Your completion code is ix6ysi0_5D6810E264828_4000
Please return to MTurk and enter your code.


LabBox ip6b1da
The experiment is finished
Your total earnings are: 3.20 USD.
As described in the instructions, a computer randomly selected two games:
Game 4 (you were an employer), earnings: 20 points
Game 8 (you were a worker), earnings: 20 points
In total you earned 40 points. This corresponds to 1.20 USD. In addition to this you receive a fixed payment of 2.00 USD.
We will transfer this to you within a couple of days. Please note that the bonus payment might arrive later than the fixed payment and might be listed as a separate transaction.
Your completion code is ip6b1da_5D6810EA896D8_4000
Please return to MTurk and enter your code.