CSAE Working Paper WPS/2014-07

Do employers trust workers too little?
An experimental study of trust in the labour market*

Stefano A. Caria† Paolo Falco‡

Dec 2013

Abstract

We conduct a field experiment to investigate employers' trust in workers. A sample of *real entrepreneurs* and *workers* from urban Ghana are respectively assigned to the roles of *employers* and *employees*. Employers have the option to hire (trust) an employee, who can in turn choose whether to exert effort (trustworthiness) in a real-effort task. By comparing employers' expectations to workers' revealed trustworthiness, we are able to detect potential misperceptions leading to sub-optimal hiring. We further devise two randomized treatments to test for the existence of expectation bias against specific worker categories and estimate the elasticity of employers' beliefs with respect to new information. We find that employers significantly underestimate workers' trustworthiness and this reduces their profit. Employees are aware of employers' sub-optimal trust. Expectations are largely inelastic with respect to news and negative signals have a stronger (downward) effect than positive ones. Our results suggest that raising employers' expectations would have a strong impact on hiring.

Keywords: trust, trustworthiness, expectations, effort, hiring, microenterprise, learning, discrimination, experiment, African labour markets.
JEL: J23, J71, O15, C9.

*We are grateful to Marcel Fafchamps, Abigail Barr, Helene Bie Lilleør, Marco Piovesan, Anna Folke Larsen and attendants of the CESS Seminar (Oxford), the University of Copenhagen Seminars, SEEDEC 2013 and of the Rockwool Foundation Seminar (Copenhagen) for invaluable comments. We are also indebted to the CSAE research assistants and to our team of enumerators in Ghana, led by Mr Moses Awoonor-Williams. We are very grateful to the Bill and Melinda Gates Foundation for financial support. All errors are our own.

†Centre for the Study of African Economies, Department of Economics, Oxford University, *Email:*stefano.caria@economics.ox.ac.uk

‡Centre for the Study of African Economies, Department of Economics, Oxford University, *Email:*paolo.falco@economics.ox.ac.uk

Centre for the Study of African Economies
Department of Economics · University of Oxford · Manor Road Building · Oxford OX1 3UQ
T: +44 (0)1865 271084 · F: +44 (0)1865 281447 · E: csae.enquiries@economics.ox.ac.uk · W: www.csae.ox.ac.uk

"If you want something done right, do it yourself."

C. Etienne, French Dramatist, 1778-1845

1 Introduction

Trust plays an important role in the process of development. At the macro level, a growing literature shows that distrust among economic agents has negative ripercussions on economic growth (e.g. Algan and Cahuc (2010), Beugelsdijk et al. (2004), Knack and Keefer (1997), La Porta et al. (1997) and Zak and Knack (2001)). At a micro level *expectations of trustworthiness* have been shown to be a key driver of risky investments, trade and other economic interactions. (Barr (2003), Ashraf et al. (2006), Binzel and Fehr (2013)). Expectations of trustworthiness, however, may diverge from the true moments of the stochastic process they refer to. Acquiring information on agents' trustworthiness amounts to risky experimentation, whose costs may prove too high for resource constrained principals. Moreover, the information produced has public-good characteristics and may be under-provided by private initiative. Formal theoretical studies have shown that individual experimentation may not converge to the truth (e.g. Rothschild (1974)) and information-sharing does not guarantee complete learning under certain network structures (Bala and Goyal (1998)). Moreover, a growing empirical literature supports the idea that agents' expectations may be biased and unrealistic in a number of domains. McKenzie et al. (2007) find that Tongan migrants to New Zealand under-estimate returns from migration. Binzel and Fehr (2013) find evidence of pessimistic expectations of peers' trustworthiness among slum dwellers in Cairo. Delavande et al. (2011a) provide observational evidence that individuals give excessive weight to the negative experiences of others when forming their own expectations. Crucially, unrealistic expectations may have severe negative consequences on agents' choices. Jensen (2010) shows that perceived returns to secondary school among students in the Dominican Republic are extremely low, despite high measured returns, resulting in sub-optimal human capital investment. He shows that a randomized intervention providing students with information on actual returns to schooling was able to increase school attainments significantly.

In the labour market, low expectations of workers' trustworthiness may depress hiring and overall employment. In many developing countries the vast majority of enterprises is run by a sole proprietor who hires no workers and hiring largely occurs through social networks and

personal connections, which, it has been argued, serve as a screening and monitoring device to minimise moral hazard (de Mel et al. (2010) Heath (2013)). Most strikingly, several wage-subsidy progrmmes, recently lauched in different developing countries, show puzzingly low take-up rates (de Mel et al. (2010), Galasso et al. (2004), Levinsohn et al. (2012)), indicating that the perceived marginal product of labour is extremely low for firms. Figure 1 illustrates this point by depicting the hiring decision of a typical micro-entrepreneur. For an initial level of expected trustworthiness θ_L, the expected Marginal Revenue Product of labour is too low to justify hiring at the going wage. As expectations increase to θ_H, the marginal revenue product shifts outwards and hiring the first worker becomes optimal.

<< **Figure 1 here** >>

The objective of this study is to investigate the decision of an employer to *entrust a worker with a job*. We pursue this goal by designing a trust experiment with a real-effort task, played between *real entrepreneurs* and *real workers* in urban Ghana. Employers have the option to trust an anonymous employee by investing an initial capital endowment to pay his/her wage. If hired, the employee can choose whether or not to exert effort to complete a trivial task that requires no skills or ability (and whose completion, is therefore, solely a function of effort). The employer's payoff is directly linked to the employee's performance, while the employee's wage, if hired, is fixed. By uncoupling the worker's pay from his/her performance and by removing the employer's ability to monitor and punish low effort, we are able to elicit trust and trustworthiness. Most importantly, by eliciting employers' expectations *directly* (by means of simple questions requiring no knowledge of probability) and comparing them to actual workers' performance, we are able to identify potential misperceptions of workers' trustworthiness leading to sub-optimal hiring. We further devise two randomized treatments to study the extent to which expectations are biased against specific worker categories and to estimate the elasticity of expectations with respect to new information.

Our main goal is not to separate expectations from other drivers of trust; and we are aware that pessimistic expectations are only 'one' of many potential mechanisms that may depress hiring in the labour market (ranging from macro effects, such as low aggregate demand, to risk-preferences and misperceptions in other domains, such as low expected

ability). Yet, convincing evidence shows that expectations are one fundamental determinant of trusting (Barr (2003), Ashraf et al. (2006)) and we aim to test this claim in the context of labour relationships. A lab-experiment is the ideal setting to stage our quest, since it helps us isolate the mechanism of interest and remove other confounding factors.

The paper answers four related questions: (i) Do employers have realistic expectations of workers' trustworthinesss? (ii) Do employers have overly pessimistic or optimistic expectations for certain categories of workers (i.e. women and youth)? (ii) How flexible are expectations; and can access to information change them? (iv) Does raising expectated trustworthiness raise employers' trusting, and if so, by how much?

Our findings show that employers significantly underestimate workers' trustworthiness, while workers correctly predict employers' trust (indicating that workers are aware of employers' pessimistic expectations). Low expectations have significant negative repercussions on hiring and cause a loss of profit for employers. Second, we find that employers' expectations of female workers' trustworthiness are biased downwards (largely among male employers), while women in fact reveal to be more trustworthy than men. Third, expectations are largely inelastic with respect to new information, but, interestingly, negative signals have a stronger effect than positive ones, making undesirable equilibria of low expectations and low trust particularly difficult to escape. Finally, we estimate the causal impact of expectations on trust and we conclude that raising employers' expectations would have a strong positive impact on their propensity to hire. These results support the idea that public policy, in the form of training, information campaigns and enhanced systems for information sharing, may be useful to correct employers' misperceptions and encourage employment.

The remainder of the paper is structured as follows. In section 2 we outline our experimental design. In section 3 we describe our sampling strategy and provide some key descriptive statistics. In section 4 we present our results. Section 5 discusses our findings and our plans for future work.

2 Experimental Design

The experiment was conducted in the city of Accra over a period of 2 weeks in 2013.[1] It took place in a central location that was easily reached from different parts of the city. Each respondent was pre-assigned to one of two experimental roles: *employer* or *employee* (the exact details of how the assignment was carried out are discussed in the next session) and randomly allocated to one of 30 experimental sessions. Each session had 10 employers and 10 employees in total.[2] Two separate locations, sufficiently distant from each other, were equipped to host the employer and the employee's room respectively.[3] Inside each room, respondents sat at separate individual desks, which were equipped with custom-designed wooden screens that ensured respondents' choices in the game could not be observed by their peers. The screens in the employees' room were higher, to ensure that employees could not see each other, minimising peer effects on productivity (see (Falk and Ichino, 2006) and (Mas and Moretti, 2009) for recent evidence on the effect of peers' pressure on workers' effort).[4] They also ensured that respondents did not feel observed or monitored by the experimenter. All the features of the employees' room were outlined to employers to ensure full common knowledge of the experimental setup.

2.1 The Game

Each employer j owns a firm producing y and in each round of the game has an initial endowment of R, which he/she can keep or invest in a worker's salary. Production of y occurs through a real effort task by an anonymous worker i, according to the following production function:

$$y_j = f(e_i) = e_i \quad where \quad e_i = 0, 1 \ (low/high) \tag{1}$$

[1] Plus four days of additional piloting conducted in the two months prior to the experiment.

[2] There were few exceptions, where some respondents dropped out and the session was not full.

[3] The distance between the rooms, which were located on two different floors of a large compound, helped us preventing respondents in different roles from getting to know each other prior to the experiment. Moreover, the regular flow of people in and out of the compound, which hosts a number of different offices and businesses, made it virtually impossible for respondents to infer information about the pool of participants in the other room based on the features of the people they met upon arrival.

[4] Moreover, by preventing workers from knowing which one of their peers was employed in each round, we were able to exclude direct effects of the unemployment rate in each session on effort, as predicted by a model of unemployment as a worker discipline device (Shapiro and Stiglitz, 1984).

This binary setup is the key to a simple elicitation of employers' expectations and it ensures high levels of understanding in the game. If j hires a worker (i.e. invest/trust), production is a binary outcome that strictly depends on employee's success in the assigned task. The triviality of the task ensures that success is strictly a function of effort (see the next section for a detailed discussion). When production is successful (i.e. $e_i = 1 \rightarrow y = 1$), the employer receives a value of $p > R$ and the worker receives a wage of W. When production fails (i.e. $e_i = 0 \rightarrow y = 0$), the employer does not receive any money, while still having to pay W to the employee. Our chosen payoffs were such that $R = W$ (i.e. a trusting employer loses all his endowment when his/her employee fails). In addition to these observable payoffs, we assume, for the sake of exposition, that each employee has an additional *unobservable* cost of effort equal to c_i and, for simplicity, he/she attaches an unobservable value of b_i to being trustworthy (e.g. intrinsic motivation). These last two terms are unknown to the employer, who has beliefs on their relative strength. These beliefs play a crucial part in the employers' decision to trust.

The game tree in Figure 2 outlines the sequence of choices and corresponding payoffs faced by employers and employees in each round. The game was played twice, with a second round announced as a surprise at the end of the first one. The results from the first round (i.e. whether a hired worker revealed to be trustworthy or not) were not announced to employers until the very end of the game.

<< **Figure 2 here** >>

Figure 3 shows the same game tree with the chosen monetary payoffs (in Ghana Cedis) replacing the game parameters. The payoffs were large with respect to hourly earnings in the reference population, hence generating strong incentives in the game.[5]

<< **Figure 3 here** >>

After the two rounds of the hiring game, respondents took part in a *risk game* and in a *dictator game*, before completing a short end-questionnaire and finally receiving their prizes.

[5] In 2012 (the latest available year of data at the time of the experiment) the median of net *daily* earnings for workers in the reference population was approximately 6 Cedis (calculated using Oxford GHUPS data). The endowment given to employers for *each* round of the hiring game *and* for a risk game played at the end was equal to 5 Cedis. A dictator game was also played and it provided an additional endowment of 3 Cedis. The result was a *total endowment of 18 Cedis* for the three games combined. The entire experiment lasted approximately 2 hours.

The risk-game follows Gaechter et al. (2010) and it consists of 5 consecutive dichotomous choices between a lottery with two possible outcomes, each carrying a probability of 50%, and a safe option worth a fixed amount. For each of the five choices the respondent was asked whether he/she would like to play the lottery or receive the fixed amount. By progressively decreasing the value of the negative outcome, the lottery becomes increasingly riskier, leading respondents to 'switch' to the safe option.[6] Figure 10 in Appendix C shows the visual aid used to explain the lotteries. For comparability with our main trust game, the value of the safe option was equalised to the initial endowment employers could invest to hire a worker. Moreover, in order to mimic the hiring decision as closely as possible, the risk-game was framed in the *loss domain*, with employers receiving an initial endowment they could either *keep* or *risk*, in the same way as they could keep their endowment or hire a worker in the trust game. Our measure of risk-preferences, therefore, is more precisely a measure of 'risk aversion in the loss domain'. In the current draft, we construct this measure by simply counting the number of choices (out of 5) when the respondent preferred the safe option over the risky lottery.[7] In the dictator game employers received an endowment of 3 Cedis, which they were asked to split between themselves and a randomly chosen anonymous employee.[8] The results of the dictator game are not relevant to the analysis in this draft and will be the subject of future work.

2.2 Treatments

In addition to the basic version of the game (control), we conducted two variants (treatments) of the experiment over random sub-samples of the employer population (who were assigned to specific 'treated' sessions). This section describes the main features of the two treatments. A complete outline of the experiment protocol is provided in Appendix C.

[6] Using a standard strategy method, we asked respondents to make a choice for each of the 5 decisions, knowing that only one of them would be randomly acted out in the end determining their prize.

[7] Given that the large majority of the respondents played according to a monotone 'switching rule' (i.e. once they switched from the lottery to the safe option they never switched back in subsequent choices in the sequence of five), this method is a consistent way of categorising workers. Further checks will be necessary to test the robustness of our results to the exclusion of workers who do not switch monotonously in the choice sequence.

[8] In half of the session the receiving employee would be randomly chosen among those who were *successful* in the task (i.e. exerted high effort) in the task. In the other half, the receiver would be chosen among *unsuccessful* workers (i.e. those who exerted low effort in the task). Employers were informed about this. The split was devised to study the role of reciprocity in altruistic choices.

Treatment 1 (T1): Information provision. At the end of round 1 employers in T1 sessions received a note showing them whether a random employee out of all those hired in previous sessions completed the task (trustworthy) or not (untrustworthy). In doing so, the treatment effectively conveyed to employers the same information they could gather through their own experience of hiring a worker. Crucially, we provided this information both to employers who had been willing to hire in round 1 and to those who had not, overcoming a fundamental problem of endogenous selection whereby new information only naturally accrues to employers who are willing to experiment. For the sake of comparability, the employers who hired in the first round did not know the performance of their worker at the time of receiving the signal (that was only revealed at the end of the experiment).[9] In half of T1 sessions the random signals did not reveal any worker characteristics. In the other half, they revealed if the worker whose performance was being reported was male/female or young/old. This additional variation was meant to test whether employers' expectations are more or less elastic towards women and young workers (two groups that commonly attract the efforts of policy-makers and will also be the focus of Treatment 2). A detailed discussion of the signal distribution is provided in the next sections.

Treatment 2 (T2): Changing the composition of the workers' pool. In T2 sessions the sample of respondents invited to take part as 'employees' was drawn so that 80% of them would belong to one of two categories of interest. In half of T2 sessions, 80% of invited workers were Women and, in the other half, 80% were Young (below the age of 25).[10] Employers in these sessions were informed of the peculiar sample composition at the beginning of the game (while respondents in non-T2 sessions were informed of the average gender and age

[9] In addition, respondents in the role of workers were informed between round 1 and round 2 that some employers had received positive information on workers' performance (and may have updated their beliefs accordingly). This was intended to test for the potential effect of changing employers' expectations on employees' effort. This additional feature is not relevant to the analysis in the current draft and the performance of treated employees in round 2 of these sessions will be excluded from the subsequent discussion.

[10] While the composition of the 'invited' workers' pool was designed to reflect these proportions, the sample of workers who 'attended' the session may have differed. This was due to respondents changing their availability in unpredictable ways, forcing the research team to accommodate their needs by re-allocating them to different sessions. It was impossible, under these circumstances, to strictly maintain the designated proportions. However, since employers were separated from and unable to see the workers' pool, and since we do not believe employers had any reason to think the pool of attendees should be 'systematically different' from the pool of invited workers, this issue should not affect our results.

composition of the invited worker pool in non-treated sessions). The rest of the design was identical to control sessions.

2.3 The Real Effort Task

We designed a trivial task, such that any worker who was willing to make an effort should succeed (i.e. reveal trustworthy). The task did not rely on any specific know-how or skills, ruling out the influence of human capital, and success was driven by workers' effort (as opposed to ability). This was explicitly and repeatedly explained to the employers, who were fully aware that any worker could attain success if he/she was willing to work steadily without shirking. We chose a real effort task, as in Fahr and Irlenbusch (2000), van Dijk et al. (2001), Falk and Ichino (2006), as opposed to a chosen effort task (e.g. Fehr et al. (1993), Altmann et al. (2012)) for one fundamental reason. We are interested in the potential mismatch between expected and revealed trustworthiness in a real-effort setting (akin to real employment); and, while shown to be correlated in laboratory settings, real effort and chosen effort may diverge substantially. This may be due, for instance, to the direct utility/disutility of work and effort. Indeed, in a lab-experiment designed to directly compare the outcomes of chosen and real-effort experiments, Bruggen and Strobel (2007) show that effort is significantly higher in real-effort tasks. Our chosen task was the following:

"Starting from a bag with 3 types of beans, your job will be to sort the beans into three smaller bags, each containing only one type, in 10 minutes"

Each bag contained 350 Grams of beans in total (1/3 of each type). This was our best estimate of a *minimum attainable amount under steady effort and no shirking.*[11]

The task was trivial, but it required constant effort and attention. In order to minimise uncertainty about one's optimal effort to succeed, workers were explicitly told in the instructions that by 'applying a steady level of effort throughout the period, they should be able to complete the task'. Moreover, in order to rule out the possibility that workers

[11]The chosen quantity was equal to the *minimum* amount of mixed beans sorted in 10 minutes by a sample of trusted survey staff, *who were instructed to work at a regular pace and were constantly monitored.* These subjects were of mean age close to the survey average, half male and half female. Direct observation of respondents in pilot sessions confirmed that completion of the task was attainable for a wide range of subjects under steady effort.

may work 'out of boredom', the experiment venue was equipped with a TV screen showing a popular local show and workers had an implicit choice between working on the task and watching TV. This choice was not recorded explicitly (i.e. respondents were not explicitly asked whether they wanted to work or watch TV), as we believe its measurement would have been severely biased. The respondents' propensity to choose leisure when observed/asked by the experimenter might have been lower than when he/she was behind the surrounding screen and was not observed.

Given the features of the task, we confidently conclude that workers' success is a valid proxy for effort and trustworthiness. It should be noted, however, that if ability (or other workers' characteristics besides their willingness to make an effort) play a significant role in their rate of success, our estimated level of trustworthiness will be a lower bound for actual trustworthiness. In other words, if some workers tried their best (i.e. were trustworthy), but due to lack of skills were unable to complete (and will be mistakenly classified as untrustworthy), overall estimated trustworthiness will be biased downwards. This possibility only strengthens the results in the next sections. Indeed, despite this potential downward bias, we will show high levels of revealed trustworthiness, which are significantly underestimated by employers.

2.4 Eliciting Expectations

The focus of this paper is on employers' expectations of workers' trustworthiness and on employees' expectations of workers' trusting. In this section we outline our procedure for eliciting these expectations.

In each round of the experiment, before or after the hiring decision was made (the order was randomized across sessions to control for potential 'order effects'), we asked employers the following question: *"Out of the 10 workers in the other room, one of whom will be assigned to you by chance, how many do you think will complete the task successfully?"* The reverse question was asked to employees: *"Out of the 10 employers in the other room, to one of whom you will be assigned by chance, how many do you think will choose to hire a worker?"* Answers to these questions were gathered by means of visual aids consisting of 10

tokens (small yellow plastic disks) that respondents were asked to distribute between a blue and a red circle printed on a game card (respectively representing success and failure). Our design was inspired by a methodological study by Delavande et al. (2011a), and by the prior work of Manski (2004) and Attanasio and Kaufmann (2009). Moreover, having experimental sessions with exactly 10 respondents per room allowed us to simplify the elicitation strategy used in much existing work, veering away from questions about 'probabilities' and towards simpler questions about 'frequencies'.[12]

The elicitation questions were not independently incentivised. The main rationale behind this decision was to avoid that respondents could use these questions to hedge against their choices in the trust game, which could have severely distorted the elicited belief distribution (and created a spurious correlation between beliefs and trusting).[13] Second, we strongly wished to minimise complexity, especially after the results of pilot sessions had clearly indicated that we should attempt to contain the cognitive burden placed upon our respondents (whose average levels of numeracy are low). Introducing an additional scoring rule for the elicitation questions appeared to go directly against this principle and we had strong reasons to believe it might have increased complexity to the point of lowering overall understanding. Third, a recent review of the literature on belief-elicitation in developing countries concludes that non-incentivised questions are effective in capturing expectations over a wide range of outcomes and that no conclusive evidence exists to indicate the superiority of incentivised tasks (Delavande et al. (2011b)). In fact, there may be reasons to believe that monetary incentives attributed by means of a scoring rule may distort the elicited distribution depending on the rule that is applied. A methodological study by Gächter and Renner (2010) with subject from developed countries (UK and Germany) in the context of a public good game shows that while monetary incentives may increase the accuracy of beliefs, they may also alter the relationship between beliefs and behaviour (with mixed

[12] Under the former, more common, approach, a typical question would be "On a scale from 1 to 10, how likely is it that a worker/employer...?". Under our approach, the question of interest becomes "Out of 10 workers/employers in the other room, how many...? The latter does not rely on employees grasping basic probability concepts and it should hence deliver more accurate answers.

[13] A way to resolve this concern and maintain the monetary incentive would have been to devise a randomized rule whereby *either* the elicitation question *or* the trust game would determine people's payoffs at the end of the game. This option would not have been easy to implement in light of the results from pilot sessions, which strongly suggested we should rather attempt to minimise complexities in order to attain satisfactory levels of understanding.

evidence on the direction of the resulting bias[14]). For this reason they conclude that if a researcher is interested (as we are) in the relatonship between beliefs and behaviour, belief elicitation should *not* be incentivised. In addition, looking at their results in detail, it appears that the increased precision of beliefs resulting from the incentive is symmetric around the mean (zero) of the error distribution. In other words, by introducing the monetary incentive respondents are equally *less* likely to *over*-estimate and *under*-estimate the object of interest. The implication for our work is that by incentivising the elicitation we could have tightened our estimate of expected trustworthiness, but we wouldn't have changed its central tendency. Our results could have thus been even stronger (both statistically and in magnitude, as imprecision may have led to classic attenuation bias in our regressions), but they would not have changed qualitatively. This improvement, however, would have come at the cost of hitting a binding constrain on the level of complexity we could allow without compromising understanding and the incentive might have caused spurious behavioural changes. This is a risk we chose to avoid at the design stage. Overall, it should be remarked that the available evidence on the role of incentives is too scant to provide clear guidance and further work in this area would be needed. Fourth, and perhaps most importantly, upon testing our experiment in pilot sessions, we observed that *unincentivised expectations strongly predicted behavour in the incentivised trust game* (a finding that is strongly confirmed by the results in the next sections), lending support to the conclusion that people truthfully revealed their belief despite the elicitation itself was not incentivised.

3 Sampling and Descriptive Statistics

3.1 Sample

One of the key objectives of this study is to measure the potential misperceptions of workers' trustworthiness among employers in Ghana. While we acknowledge the potential limitations of a comparison between laboratory outcomes and real-world decisions, the best way to achieve this goal was to assign *real entrepreneurs* and *real workers* to the experimental roles of employers and employees respectively. It seems reasonable to assume that respondents' behaviour in the lab experiment is based on the priors they develop in their actual labour

[14]While Gächter and Renner (2010) show that monetary incentives *increase* respondents' contributions in a public good game, they report a result by Croson (2000) showing the opposite.

market experience; and in order to give our experiment the best chances to elicit those priors, the most appealing strategy was to assortatively match respondents to their roles based on real-life experience (a similar approach is used by Barr and Zeitlin (2011) who study the incentives of primary school providers in Uganda by means of a lab experiment). Here is how we pursued this goal in practice. Using data from the *Ghana Household Urban Panel Survey (GHUPS)*, a representative survey of the urban Ghanaian population conducted by the Centre for the Study of African Economies (University of Oxford), we extracted a random sample of workers *who owned a business in at least one of the 2 most recent survey waves (2010, 2012)*. We assigned those respondents to the role of *employers*. The sample of *employees*, on the other hand, was randomly drawn from the general working-age population.[15]

3.2 Descriptive Statistics

Table 1 and 2 describe the main characteristics of our samples of Employers and Employees respectively.

<< **Table 1 here** >>

<< **Table 2 here** >>

As expected, given our sampling strategy, employers are significantly older, more likely to be married and, given the structure of the Ghanaian economy (where female participation in entrepreneurial activities is high), the majority of them are women. Employees, on the other hand, are balanced across gender, they are younger on average and significantly more likely to be currently unemployed. Figure 4 plots the distribution of educational attainments for the two groups. It shows that employees have slightly higher educational attainments, most likely the reflection of belonging to a younger generation.

<< **Figure 4 here** >>

Finally, Table 3 tests the balance of covariates across the different treatment groups to which employers were assigned. It shows that random assignment largely succeeded in achieving balance. In order to improve upon the precision of our estimates, the analysis in the next section will control for the most important employer characteristics.

[15] Since the subset of eligible GHUPS respondents in Accra was not sufficient to attain the desired sample size, we recruited additional respondents by means of visits to random households in areas of Accra that were near but did not overlap with GHUPS enumeration areas. These additional participants were assigned to the two experimetnal roles using the same criteria described above.

<< **Table 3 here** >>

4 Results

4.1 Are employers' expectations realistic?

We begin by addressing the central question in our investigation: *Do employers have 'realistic' expectations of workers' trustworthiness?* Our experimental design allows us to answer this question directly, by comparing employers' elicited expectations to the revealed trustworthiness of workers (captured by their rate of success in the task).

$$H_0 : \underbrace{\{E_{j,t}[e_{i,R=1,2}] - Pr[e_{i,R=1,2} = 1]\}}_{\text{(Exp. - Actual Trustworthiness in Round 1,2)}} = 0 \tag{2}$$

Result 1: Employers significantly underestimate workers trustworthiness. Figure 5 shows our main result very clearly. In both rounds of the game the proportion of workers who reveal to be trustworthy by carrying out the task correctly is significantly higher than it is expected by employers on average. Moreover, the cumulative distribution of employers' expectations, plotted in Figure 6, reveals that the vast majority of employers underestimate employees' trustworthiness.

<< **Figure 5 here** >>

<< **Figure 6 here** >>

Result 2: Workers correctly predict employers' trust. Second, we compare employees' elicited expectations of employers' propensity to trust with the actual rate of trusting. The results are shown in Figure 7. Interestingly, they reveal that workers' average expectations correctly predict employers' trusting, since the share of employers who choose to hire is not significantly different from workers' predicted proportion.

<< **Figure 7 here** >>

Insofar as behaviour in the lab is driven by beliefs developed in the labour market, these two results point to an interesting asymmetry between employers' and employees' learning. Employees appear to be aware of employers' low expectations and sub-optimal hiring. On

the other hand, the costs and risks of experimentation may be preventing employers from forming correct beliefs about workers' productivity. Further theoretical work will be required to place a structure on the process leading to this asymmetry.

Next, we explore the heterogeneity of employers' misperceptions by exploiting the exogenous changes in sample composition we induced in Treatment 2. At the beginning of each T2 session, employers were informed that the majority of the workers invited (precisely 80%) belonged to a specific category (as opposed to Control and T1 sessions where the sample was balanced). In particular, our focus was on Gender and Age. In half of the T2 session, employers were informed that 80% of invited employees were women and in the other half they were told that 80% of the employees were young (below 25). This was the only design variation with respect to the Control sessions. Hence, by comparing employers' expectations between T2 and T0 we can identify the effect of workers' characteristics on employers' beliefs. We do so by estimating the following model:

$$E_j[e_i = 1] = \alpha + \beta_W T_{2,W} + \beta_Y T_{2,Y} + \gamma X_j + u_j \tag{3}$$

where:

$T_{2,W} = 1$[T2 with majority of FEMALE workers]
$T_{2,Y} = 1$[T2 with majority of YOUNG workers]
X_j is a vector of control variables.

The results, reported in Table 4, are restricted to round 1 to conserve space. Moreover, by focusing on round 1, we are able to compare T2 to Control *and* T1 sessions jointly (since round 1 was identical in the latter two), maximising the available sample size. The same patters, however, emerge in round 2.

Result 3: *Male employers significantly underestimate female employees' performance.* When we run the estimation over the entire sample (col 1), we find that *employers have significantly lower expectations when the majority of workers in the session are women.* The result is not very strong (only statistically significant at the 10% level). However, when we split the estimation by gender (of the employer), we find a much stronger effect on the

expectations of male employers (col 2), which are significantly lower when the sample of workers is predominantly female. The expectations of female employers, on the other hand, are not affected by the worker pool composition (col 3). The result is particularly striking when confronted with revealed trustworthiness by worker gender. Figure 8 shows that female workers are, in fact, significantly more trustworthy than men.

<< **Table 4 here** >>

<< **Figure 8 here** >>

The results in Table 4 also suggest that male employers have a positive expectations bias towards young workers, which is absent among female employers. This bias, however, is not reflected in revealed employees' trustworthiness, which shows no differential between older and younger workers (see Figure 9).

<< **Figure 9 here** >>

4.2 Do expectations matter for trusting?

Having presented our main results on employers' misperceptions of employees trustworthiness, we now test whether the expectations we elicited have any predictive power on the (incentivised) choice to trust a worker or not. This test follows in the footsteps of Ashraf et al. (2006), who find a significant effect of expected trustworthiness on trusting in a monetary trust game. We estimate the following model of trust, pooling observations from the two rounds:[16]

$$Pr(H_{j,R} = 1) = \alpha + \beta E_{j,R}[e_i = 1] + \gamma X_j + u_{j,R} \tag{4}$$

where:

$Pr(H_{j,R} = 1) \equiv$ Probability of Trusting (Hiring) in round R.

$E_{j,R}[e_i = 1] \equiv$ Employer j's expectations in round R.

X_j is a vector of control variables.

[16] And accordingly clustering the standard errors at the specific session-round level.

Result 4: Employers' expectations are a strong predictor of trusting. Our results largely confirm the conclusions by Ashraf et al. (2006). Table 5 shows that expected trustworthiness strongly and significantly affects the probability of trusting in all the proposed specifications. Increasing an employers' expectations by 10 percentage points increases the probability of trusting by nearly 3 percentage points on average. We also find that our elicited measure of risk-preferences significantly predicts trust *independently* of expectation. This second result is very important, as it shows that we are clearly identifying two different effects: one related to employers' perceptions of the moments of the trustworthiness distribution and a second one, which depends on the curvature of their utility function in the loss-domain. It should be noted, once again, that the loss-domain is the most relevant one for the choice at hand, since employers who choose to hire stand to 'lose' an endowment they could otherwise keep for themselves.

<< **Table 5 here** >>

An important conclusion emerging from this part of the analysis is that *by underestimating workers' trustworthiness employers lose profits*, since pessimistic expectations lead to sub-optimal hiring.

4.3 Does access to information change expectations?

Finally, we set out to answer the most policy-relevant questions in our analysis: *Can access to information change expected trustworthiness? And how large is the causal impact of changing expectations on trust?* For this part of the investigation we exploit the design features of Treatment 1. Between the first and the second round of the experiment, employers in Treatment 1 received a random (and private) information signal, which informed them of the trustworthiness of a random worker out of those who had been hired in previous sessions. In half of T1 sessions, the signal carried no worker characteristics. In the other half, it revealed whether the randomly drawn worker was male or female, and whether he/she was old or young (above or below 25). This added feature was designed to explore potential heterogeneity in the elasticity of beliefs with respect to new information.[17] The treatment was designed

[17]The distribution of signals we used was calibrated on the results of a pilot we ran prior to the main experiment, which revealed a level of trustworthiness among employees of around 60%. In order to have a balanced distribution of signals, we deviated slightly from that proportion and provided 50% of the time a positive signal and a negative one in the remaining 50% of the cases. Two employers in each session received

to provide employers with *the same information they could obtain by hiring a worker* (i.e. experimentation). At the time of receiving the signal the results of the first round had not been revealed to the employers.[18] This meant that *all* employers had the same information prior to treatment and they all received a random signal, whether or not they had hired in round 1. We could therefore overcome, by experimental design, a classic problem of endogeneity whereby employers who are less prone to trusting are also the ones who are less likely to receive information from the market, and their expectations might be the harderst to change.

In order to gauge the impact of information signals on expectations, we estimate the following model:

$$E_{j,2}[e_{i,2}=1] = \alpha + \eta E_{j,2}[e_{i,2}=1] + \sum_{\kappa} \beta_\kappa S_{T1,\kappa} + \gamma X_j + \nu_j \tag{5}$$

where:

- $E_{j,R}[e_{i,2}=1] \equiv$ employers' expectations in each round, R = (1,2).
- $S_{T1,\kappa} = 1[Signal\ in\ T1 = \kappa],\ where\ \kappa = Positive,\ Negative\ or\ No\ Signal.$[19]

<u>Result 5:</u> *Expectations are inelastic to new information and negative signals have a stronger (downward) effect than positive ones.* The results are reported in Table 6. In the first two columns we pool data from all T1 sessions (splitting the different types of positive signals in col 2). In the remaining columns we confine the estimation to sessions with homogenous (col 3) and heterogeneous (col 4-5) signals, but the results are unchanged.[20] They show that positive signals have no impact on employers' expectations, while negative signals significantly decrease expected workers' trustworthiness from one

'no signal', to test for the sheer effect of being in a treated session. When we further cross-cut the analysis by gender and age (for half of T1), it was impossible to provide a distribution of signals that reflected the real distribution, while maintaining sufficient statistical power to identify heterogenous effects. In those sessions, we provided respondents with information from a synthetic distribution of only positive signals from workers who were either female or young. Since the signals were private knowledge, employers could not possibly be aware of this deviation. Moreover, given the overall high levels of employees' trustworthiness, we were not worried this minor deception could lead to significant financial losses for employers who decided to trust as a result of receiving a positive signal (which, anyway, only happened in extremely rare cases).

[18] They were only revealed at the very end of the game, to avoid any wealth effects in round 2 caused by knowledge of having won (or lost) in round 1.

[19] Positive signals were further split into (Positive + Female worker) and (Positive + Young worker) as explained in a previous note. The rationale for having workers in T1 who received no signal was to test for the sheer effect of taking part in a treated session.

[20] In col 3 we confine the estimation to sessions with homogenous signals, where no gender or age info was attached to the signals. In col 4-5 we concentrate on sessions where signals were heterogenous by gender and

round to the next. We also find that receiving no signal (while being in a T1 session) does not have an impact on expectations. This is an important robustness check, as it allows us to rule out the possibility that our results may be driven by a spurious 'session effect'.

Crucially, a negative signal appears to decrease employers' expectations by 7-8 percentage points. The result, however, is not statistically strong in all specifications, pointing to the overall conclusion that employers' expectations are rather inelastic with respect to new information.[21] Yet, the interesting asymmetry between positive and negative signals strongly indicates that the scarring effects of early disappointments in an entrepreneur's career may be hard to overcome and that undesirable equilibria of low expectations and low trust may be hard to exit. It also suggest that workers may have a weak incentive to signal their trustworthiness in the labour market if employers' beliefs are hard to change, potentially leading to a vicious cycle of low employee performance and low employer expectations. Further theoretical work will be necessary to characterise the mechanisms at play and the nature of such equilibria.

<< **Table 6 here** >>

Finally, we are aware that the amount of information provided in Treatment 1 is rather limited. Our design was driven by the idea of providing employers with the same information they could obtain from hiring their first worker (i.e. one instance of experimentation). One can easily think of more powerful interventions that may have a stronger impact on workers' expectations. For instance, one may provide information on *average* worker performance, which would be much more informative than a single observation and could have a stronger impact on expectations. However, before designing a more powerful treatment, it would be useful to know *what is the impact of changing expectations on the rate of hiring/trusting*. The last step in our analysis will be to gauge the magnitude of this impact. Crucially, we will once again achieve identification by exploiting the random assignment of signals in Treatment 1, which can be used as an instrument for expected trustworthiness.

age. Since all signals in heterogenous sessions were positive, the impact of negative signals is only identified in col 3.

[21] Or that the signal generated by a single instance of experimentation is too weak to change beliefs, as discussed below.

Result 6: Raising expected trustworthiness has a strong impact on trusting. The estimates from our IV estimation are reported in Table 7. Column 1 and 2 show the results from a naïve OLS regression of the probability of trusting in round 2.[22] In column 3, we instrument the expectations variable and, interestingly, we find that its effect grows considerably in magnitude, while remaining highly significant. The results show that increasing expected trustworthiness by 10 percentage points would increase the probability of trusting a worker by more than the same amount. These estimates suggest that a policy intervention aiming to improve employers' expectations of workers' trustworthiness may have a strong impact on employment.

<< Table 7 here >>

5 Conclusions

A growing literature documents the crucial role of trust in the process of development. Little emphasis, however, has been placed on the role that trust between employers and employees plays in the labour market. From a theoretical point of view, low expectations of workers' trustworthiness may depress hiring and decrease employment. From an empirical standpoint, little is known about the extent to which employers' expectations and willingness to trust an employee are a realistic reflection of workers's behaviour. This study is an attempt to determine whether employers' expectations are aligned with workers' actual trustworthiness, and to what extent such expectations can be changed through new information, stimulating higher trust.

We pursue this goal by means of an original lab experiment consisting of a binary trust-game with a real-effort task, played between *real entrepreneurs* and *real workers* in urban Ghana. In our design entrepreneurs have the option to trust an anonymous employee by investing an initial capital endowment to pay his/her wage. If hired, the employee can choose whether or not to complete a trivial task that requires no skills or ability (and whose completion, is therefore, solely a function of effort). The employer's payoff is directly linked to the employee's performance, while the employee's wage is fixed. By uncoupling the

[22] We restrict the analysis to round 2 because exogenous T1-signals were only provided after the first round.

worker's pay from his/her performance and by removing the employer's ability to monitor and punish low effort, we are able to elicit trust and trustworthiness. Most importantly, by eliciting employers' expectations *directly* (by means of simple questions requiring no knowledge of probability concepts) and comparing them to actual workers' performance, we are able to identify potential misperceptions leading to sub-optimal hiring. We further devise two randomized treatments to study the extent to which expectations are biased against specific worker categories and to estimate the elasticity of expectations with respect to new information.

Our main finding is that *employers significantly underestimate workers' trustworthiness, while workers' correctly predict employers' propensity to (under)trust them.* This points to the interesting conclusion that the costs and risks of experimentation may lead employers to an undesirable equilibrium of pessimistic expectations and sub-optimal trusting. When we explore the heterogeneity of expectations with respect to workers' characteristics (gender and age), we find that employers' expectations of female workers' trustworthiness are biased downwards among male employers, while women in fact reveal to be more trustworthy than men. No bias is instead detected against young workers. Third, we find that expectations are largely inelastic with respect to new information and negative signals (i.e. signals of workers' untrustworthiness) have a stronger (downward) impact on employers' expectations than positive ones. This important asymmetry may crucially sustain the aforementioned bad equilibrium of low expectations and low trust. It also indicates that negative experiences early in an employer's career may have sustained scarring effects on his/her propensity to hire later on. Finally, we estimate that raising employers' expectations would have a strong impact on their propensity to trust. This result supports the idea that public policy, in the form of training, information campaigns and enhanced systems for information sharing, may be useful to correct employers' misperceptions and encourage hiring.

References

Algan, Y. and P. Cahuc (2010): "Inherited Trust and Growth," *American Economic Review*, 100, 2060–92.

Altmann, S., A. Falk, and M. Wibral (2012): "Promotions and Incentives: The Case of Multistage Elimination Tournaments," *Journal of Labor Economics*, 30, 149 – 174.

Ashraf, N., I. Bohnet, and N. Piankov (2006): "Decomposing trust and trustworthiness," *Experimental Economics*, 9, 193–208.

Attanasio, O. and K. Kaufmann (2009): "Educational Choices, Subjective Expectations, and Credit Constraints," NBER Working Papers 15087, National Bureau of Economic Research, Inc.

Bala, V. and S. Goyal (1998): "Learning from Neighbours," *Review of Economic Studies*, 65, 595–621.

Barr, A. (2003): "Trust and expected trustworthiness: experimental evidence from zimbabwean villages," *Economic Journal*, 113, 614–630.

Barr, A. and A. Zeitlin (2011): "Conflict of interest as a barrier to local accountability," *CSAE Working Paper Series*, 2011-13.

Beugelsdijk, S., H. L. de Groot, and A. B. van Schaik (2004): "Trust and economic growth: a robustness analysis," *Oxford Economic Papers*, 56, 118–134.

Binzel, C. and D. Fehr (2013): "Social distance and trust: Experimental evidence from a slum in Cairo," *Journal of Development Economics*, 103, 99–106.

Bruggen, A. and M. Strobel (2007): "Real effort versus chosen effort in experiments," *Economics Letters*, 96, 232–236.

Croson, R. T. A. (2000): "Thinking like a game theorist: factors affecting the frequency of equilibrium play," *Journal of Economic Behavior & Organization*, 41, 299–314.

de Mel, S., D. McKenzie, and C. Woodruff (2010): "Wage Subsidies for Microenterprises," *American Economic Review*, 100, 614–18.

Delavande, A., X. Giné, and D. McKenzie (2011a): "Eliciting probabilistic expectations with visual aids in developing countries: how sensitive are answers to variations in elicitation design?" *Journal of Applied Econometrics*, 26, 479–497.

——— (2011b): "Measuring subjective expectations in developing countries: A critical review and new evidence," *Journal of Development Economics*, 94, 151–163.

Fahr, R. and B. Irlenbusch (2000): "Fairness as a constraint on trust in reciprocity: earned property rights in a reciprocal exchange experiment," *Economics Letters*, 66, 275–282.

Falk, A. and A. Ichino (2006): "Clean Evidence on Peer Effects," *Journal of Labor Economics*, 24, 39–58.

Fehr, E., G. Kirchsteiger, and A. Riedl (1993): "Does Fairness Prevent Market Clearing? An Experimental Investigation," *The Quarterly Journal of Economics*, 108, 437–59.

Gächter, S. and E. Renner (2010): "The effects of (incentivized) belief elicitation in public goods experiments," *Experimental Economics*, 13, 364–377.

Gaechter, S., E. J. Johnson, and A. Herrmann (2010): "Individual-level loss aversion in riskless and risky choices," .

Galasso, E., M. Ravallion, and A. Salvia (2004): "Assisting the transition from workfare to work: A randomized experiment," *Industrial and Labor Relations Review*, 58, 128–142.

Heath, R. (2013): "Why do Firms Hire using Referrals? Evidence from Bangladeshi Garment Factories," *University of Washington, Unpublished Paper.*

Jensen, R. (2010): "The (Perceived) Returns to Education and the Demand for Schooling," *The Quarterly Journal of Economics*, 125, 515–548.

Knack, S. and P. Keefer (1997): "Does Social Capital Have an Economic Payoff? A Cross-Country Investigation," *The Quarterly Journal of Economics*, 112, 1251–88.

La Porta, R., F. L. de Silane, A. Shleifer, and R. W. Vishny (1997): "Trust in Large Organizations," *American Economic Review*, 87, 333–38.

Levinsohn, J., N. Rankin, G. Roberts, and V. Schoer (2012): "Wage subsidies and youth employment in South Africa. Evidence from a randomised control trial." *CSAE Annual Conference 2012*.

Manski, C. F. (2004): "Measuring Expectations," *Econometrica*, 72, 1329–1376.

Mas, A. and E. Moretti (2009): "Peers at Work," *American Economic Review*, 99, 112–45.

McKenzie, D., J. Gibson, and S. Stillman (2007): "A land of milk and honey with streets paved with gold : do emigrants have over-optimistic expectations about incomes abroad ?" Policy Research Working Paper Series 4141, The World Bank.

Rothschild, M. (1974): "A two-armed bandit theory of market pricing," *Journal of Economic Theory*, 9, 185–202.

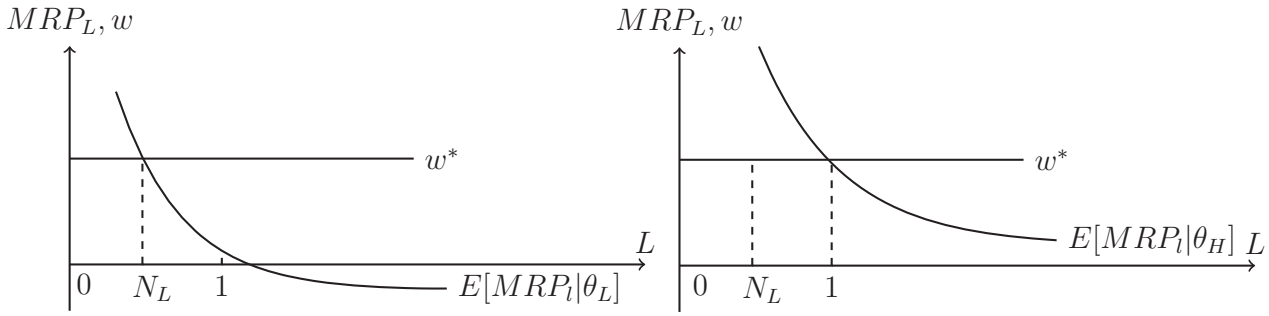
Shapiro, C. and J. E. Stiglitz (1984): "Equilibrium Unemployment as a Worker Discipline Device," *American Economic Review*, 74, 433–44.

van Dijk, F., J. Sonnemans, and F. van Winden (2001): "Incentive systems in a real effort experiment," *European Economic Review*, 45, 187–214.

Zak, P. J. and S. Knack (2001): "Trust and Growth," *Economic Journal*, 111, 295–321.

A Figures

Figure 1: The hiring decision of a micro-entrepreneur



(a) Low Expected Trustworthiness (b) High Expected Trustworthiness

Figure 2: The Game Tree

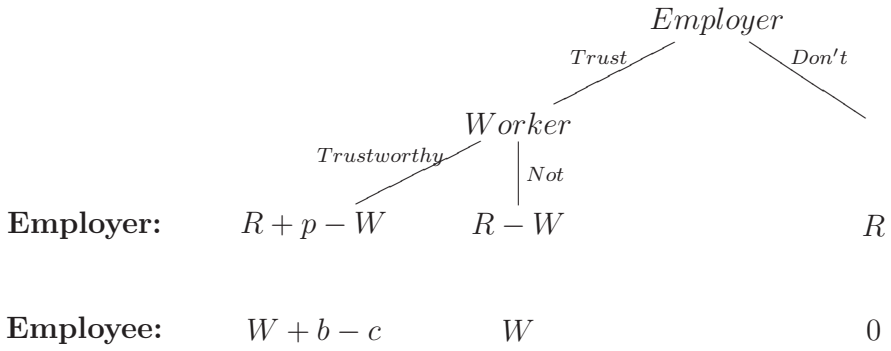


Figure 3: The Game Tree with Monetary Values (Ghana Cedis)

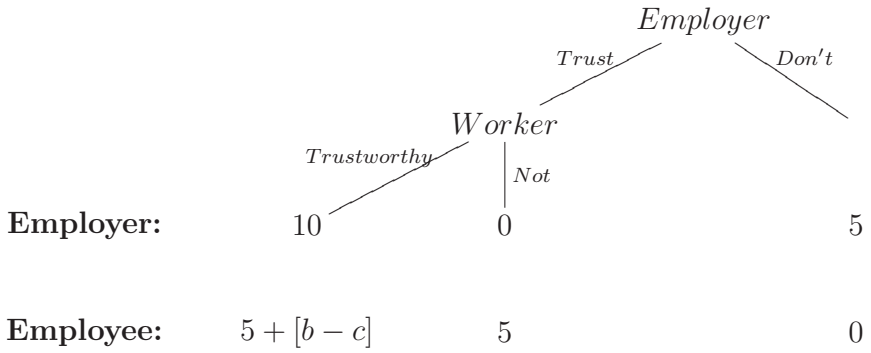


Figure 4: Education

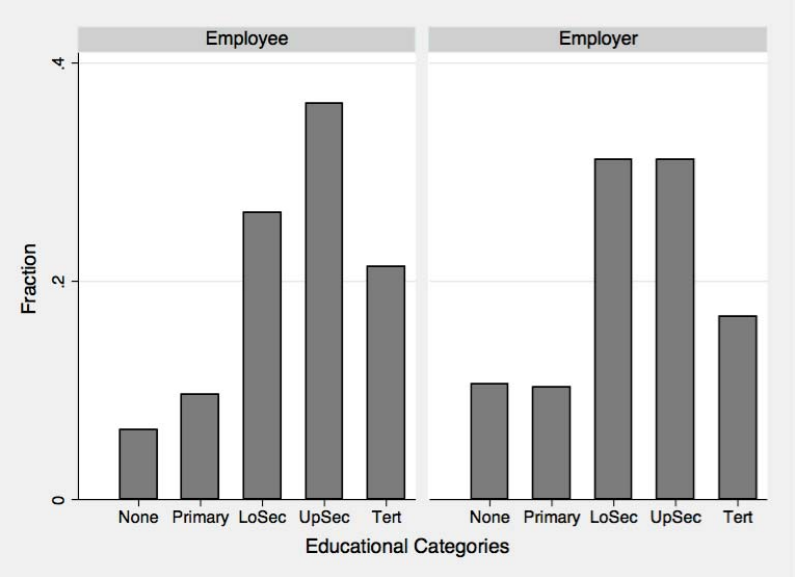


Figure 5: Expected and Revealed Trustworthiness

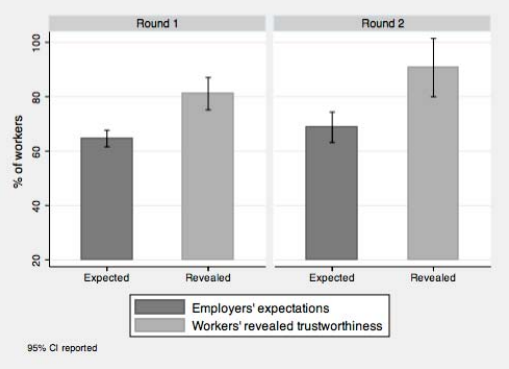


Figure 6: Cumulative distribution of Employers' expectations

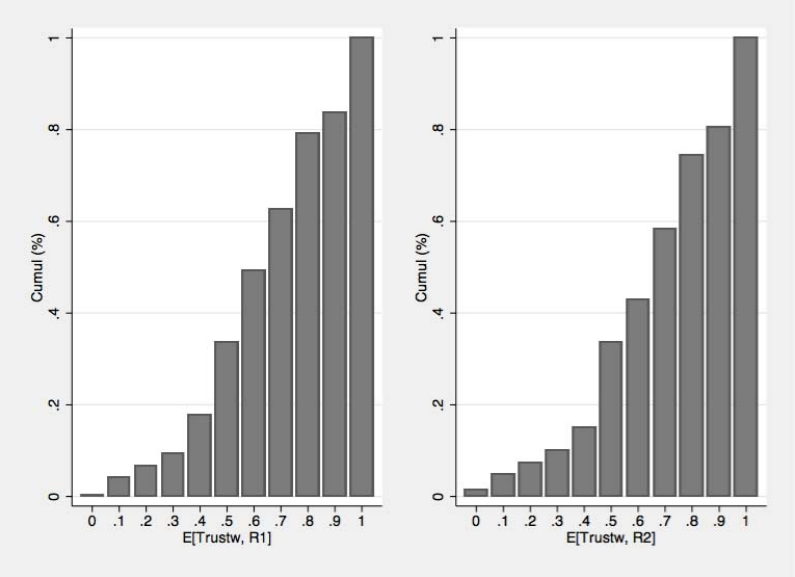


Figure 7: Expected and Revealed Trusting

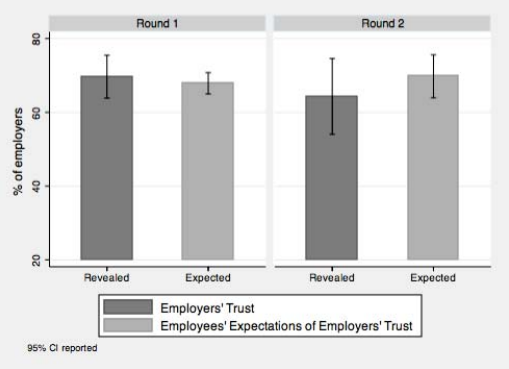


Figure 8: Revealed Trustworthiness by Gender

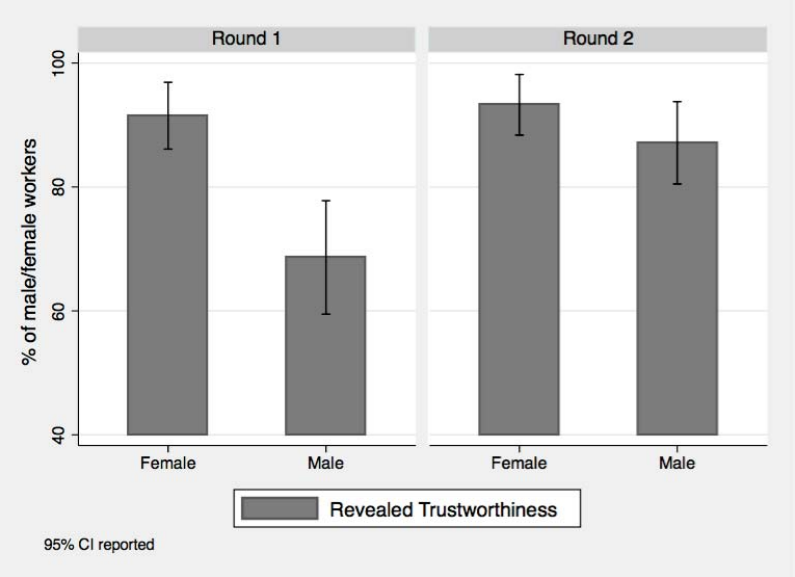
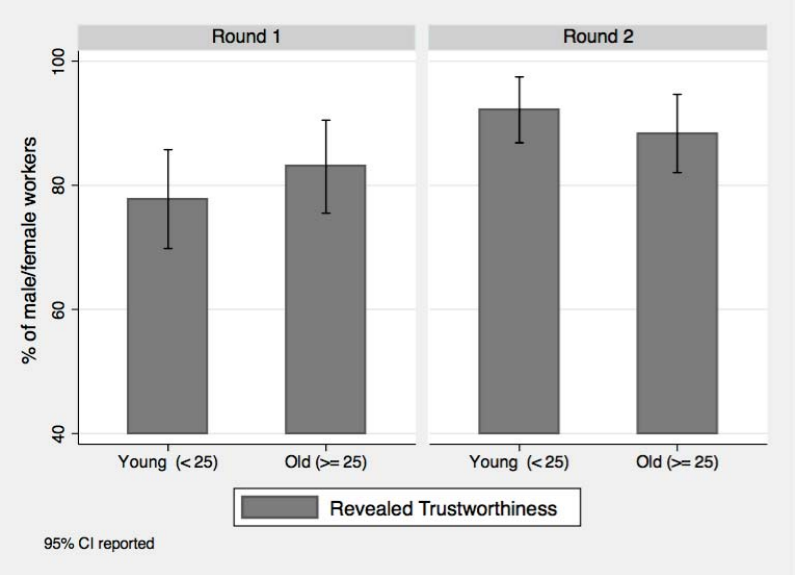


Figure 9: Revealed Trustworthiness by Age



B Tables

Table 1: Employers

Variable	Mean	Std. Dev.
Age	37.582	11.701
Married	0.562	0.497
NoSchool	0.106	0.309
Primary	0.103	0.304
LoSecondary	0.312	0.464
UpSecondary	0.312	0.464
Tertiary	0.168	0.374
Male	0.366	0.483
Unemployed	0.106	0.309
N	292	

Table 2: Employees

Variable	Mean	Std. Dev.
Age	27.954	10.089
Married	0.278	0.449
NoSchool	0.064	0.245
Primary	0.096	0.295
LoSecondary	0.263	0.441
UpSecondary	0.363	0.482
Tertiary	0.214	0.411
Male	0.47	0.5
Unemployed	0.356	0.48
N	281	

Table 3: Covariates Balance across treatments

	T1	T0,T2	Diff	T2	T0,T1	Diff
Variable	Mean		t-value	Mean		t-value
Male	.36	.37	.23	.39	.35	-0.63
Age	36.7	38.0	0.9	37.1	37.8	0.53
Primary	.08	.11	.84	.10	.10	-.014
Lower Secondary	.27	.34	1.21	.33	.30	-.47
Upper Secondary	.33	.30	-.39	.40	.27	-2.37**
Tertiary	.15	.18	.48	.12	.19	1.42
Married	.49	.60	1.76*	.57	.56	-0.13
Unemployed	.06	.13	1.77*	.11	.10	-0.28
Student/Apprentice	.06	.05	-.34	.08	.04	-1.46
Trader	.37	.26	-1.95*	.23	.33	1.79*
OtherSelf	.13	.14	.15	.10	.15	1.19
N	98	194	292	97	195	292

Note: Confidence: *** ↔ 99%, ** ↔ 95%, * ↔ 90%

Table 4: Employers' Expectations towards different worker-types (Dep Var: $E_{R=1}$[Trustw])

	All	MaleEnt	FemalEnt
Young Work.s	.179	1.412	-.368
	(.577)	(.539)***	(.731)
Women Work.s	-.656	-1.621	-.100
	(.340)*	(.633)**	(.368)
Age	.011	.026	.010
	(.017)	(.024)	(.021)
Male	.120		
	(.314)		
Education (ys)	.051	.046	.059
	(.028)*	(.064)	(.029)**
Married	.352	.036	.209
	(.397)	(.744)	(.364)
Unemployed	.468	.728	.323
	(.332)	(1.050)	(.513)
Session Char.s	Yes	Yes	Yes
Const.	5.308	4.542	5.724
	(.665)***	(1.037)***	(.821)***
Obs.	292	107	185

Note: Confidence: *** ↔ 99%, ** ↔ 95%, * ↔ 90%.; Robust standard errors in parentheses (cluster. by Session); Session Char.s = Time and Timing of Expect.s Q.

Table 5: Trust and Expected Trustworthiness (Dep Var: Pr[Hire])

	OLS1	OLS2	Probit1	Probit2	Probit-Mfx
E[Trustw]	.030 (.008)***	.027 (.008)***	.083 (.023)***	.077 (.024)***	.027
Loss Aversion		-.048 (.012)***		-.142 (.037)***	-.049
Age		.001 (.002)		.003 (.006)	.001
Male		.027 (.035)		.087 (.106)	.030
Education (ys)		.004 (.006)		.011 (.017)	.004
Married		-.029 (.040)		-.085 (.119)	-.030
Unemployed		.042 (.057)		.124 (.177)	.042
Exp Q b/f choice		-.111 (.051)**		-.330 (.151)**	-.116
Round 2	-.016 (.052)	-.015 (.051)	-.043 (.150)	-.038 (.152)	-.013
Const.	.494 (.066)***	.594 (.125)***	-.042 (.177)	.279 (.358)	
Obs.	584	584	584	584	584

Note: Confidence: *** ↔ 99%, ** ↔ 95%, * ↔ 90%.; Robust standard errors in parentheses (cluster. by Session+Round).

Table 6: Expectations Updating (Dep Var: E[Trustw, R2])

	All1	All2	Homog	Heter1	Heter2
E[Trustw, R1]	.680 (.051)***	.679 (.052)***	.639 (.053)***	.676 (.062)***	.675 (.063)***
Neg Sig	-.760 (.331)**	-.682 (.380)*	-.782 (.413)*		
Pos Sig	.030 (.339)		.032 (.471)	-.202 (.559)	
Pos Sig (F)		.196 (.539)			.023 (.564)
Pos Sig (Y)		-.397 (.591)			-.477 (.652)
No Sig (Homog T1)	.542 (.387)	.468 (.459)	.609 (.485)		
No Sig (Heter T1)	-.257 (.374)	-.189 (.465)		.059 (.505)	.060 (.504)
Const.	2.303 (.590)***	2.296 (.592)***	2.221 (.673)***	2.059 (.743)***	2.052 (.738)***
Personal Char.s	Yes	Yes	Yes	Yes	Yes
Session Char.s	Yes	Yes	Yes	Yes	Yes
Obs.	292	292	195	194	194

Note: Confidence: *** ↔ 99%, ** ↔ 95%, * ↔ 90%.; Robust standard errors in parentheses (cluster. by Session); All1: includes all T1 sessions, all positive signals are pooled together; All2: includes all T1 sessions, different positive signals are separated; Homog: only T1 sessions with homogenous signals (i.e. no info on worker age and gender attached to signals); Heter1: only T1 sessions with heterogenous signals by gender and age, all positive signals are pooled together; Heter2: only T1 sessions with heterogenous signals by gender and age, different positive signals are separated.

Table 7: The impact of raising expectations on trust (Dep Var: Pr[Hire])

	R2	R2C	R2IV
E[Trustw, R2]	.036	.034	.171
	(.011)***	(.012)***	(.065)***
Age		-.0005	-.002
		(.002)	(.002)
Male		.053	.042
		(.045)	(.056)
Education (ys)		-.001	-.003
		(.007)	(.008)
Married		-.030	-.101
		(.056)	(.062)
Unemployed		.090	.013
		(.077)	(.122)
Exp Q b/f choice		-.123	-.062
		(.076)	(.082)
Const.	.434	.529	-.292
	(.086)***	(.156)***	(.423)
Obs.	292	292	292

Note: Confidence: *** ↔ 99%, ** ↔ 95%, * ↔ 90%.; Robust standard errors in parentheses (cluster. by Session).

C Experiment Protocol

C.1 Control

Round 1

1. j informed of average worker char.s (gender, age, etc.) to anchor expectations.
2. j's expectations of workers trustworthiness elicited.
3. j chooses hire/not hire a random and anonymous i.
4. i chooses effort (not revealed to j until the end).

Round 2 (announced as a 'surprise' at the end of R1)

1. j's expectations (re)elicited

   [they cannot be assumed equal to R1 (e.g. due to expected learning or fatigue)]
2. j chooses hire/not hire a (new) random i.
3. i chooses effort.

Post-Exper:

i.Risk Game

ii.Dictator Game

iii.End Questionnaire,

iv. Payoffs Revealed and Prizes Distributed

C.2 Treatment 1

Round 1

1. j informed of average worker char.s (gender, age, etc.) to anchor expectations.
2. j's expectations of workers trustworthiness elicited.
3. j chooses hire/not hire a random and anonymous i.
4. i chooses effort (not revealed to j).

Round 2 (announced as a 'surprise' at the end of R1)

1. ***T1: Inform*** j ***of the trustworthiness of a random*** i ***from prev. sessions.***
2. j's (new) expectations elicited.

3. j chooses hire/not hire a (new) random i.
4. i chooses effort.

Post-Exper:

i.Risk Game

ii.Dictator Game

iii.End Questionnaire,

iv. Payoffs Revealed and Prizes Distributed

C.3 Treatment 2

Round 1

1. **T2: j informed that 80% of workers invited to the session belong to a specific category (i.e. young, female)**.
2. j's expectations of workers trustworthiness elicited.
3. j chooses hire/not hire a random and anonymous i.
4. i chooses effort (not revealed to j)

Round 2 (announced as a 'surprise' at the end of R1)

1. j's expectations (re)elicited

   [they cannot be assumed equal to R1 (e.g. due to expected learning or fatigue)]
2. j chooses hire/not hire a (new) random i.
3. i chooses effort.

Post-Exper:

i.Risk Game

ii.Dictator Game

iii.End Questionnaire,

iv. Payoffs Revealed and Prizes Distributed

Figure 10: Risk Game Lotteries