



An experimental study of gender differences in agency relationships

Yaxiong Li^{a,b,*}, Erwann Sbai^a, Ananish Chaudhuri^{a,c}

^a University of Auckland

^b Xi'an Jiaotong-Liverpool University

^c CESifo Munich

ARTICLE INFO

JEL Classification Codes:

C71
C91
J30

Keywords:

Gender differences
Agency relationships
Trust
Reciprocity
Leadership

ABSTRACT

We look at gender differences in agency relationships between an employer and a worker. The employer can resort to two different ways of motivating the worker to exert effort. One of these depends on establishing a relational contract, which is based on mutual trust and reciprocity between the employer and the worker. A second type of contract relies more on extrinsic motivations in the form of fines for the worker if the worker is found to be shirking. Our *ex ante* hypothesis was that women would opt for the relational contract more than men. This conjecture is not borne out by the evidence. By and large, we do not find many significant differences between the genders, other than the fact that women tend to offer more generous contract terms, a fact that is in keeping with prior finding on gender differences in generosity.

1. Introduction

The well-documented “gender gap” in the workplace, has been and remains the subject of extensive research. One can think of the “gender-gap” as referring to two related, yet somewhat different, phenomena. The first refers to the well-known gender gap in earnings; the fact the median earnings for women are less than that for men. The OECD defines the gender wage gap as the difference between median earnings for males and females relative to the median earnings for males. Looking at data for 2015–2018, we find that the average gender wage gap for all OECD countries is 13.5%, however there are substantial cross-country variations. At the most unequal end we have Korea (with a difference of 35%), followed by Japan (25%), Israel (22%) and then Canada, USA and Finland (around 18%). At the other extreme, we have Belgium, Greece, Costa Rica, Denmark and Italy, all hovering around 5% with Ireland, Norway and Sweden at about 6% and New Zealand at a little less than 8%.

Alongside this difference in earnings, there is also a gender gap in leadership roles: as one moves up the hierarchy of organizations, one finds fewer women. Sandberg (2013) notes, that, at the time of her writing, of the 195 independent countries in the world, only 17 were led by women; women held just 20% of seats in parliaments globally and about 14% of executive officer positions, 17% of board seats, and constituted 18% of elected congressional officials in the US.

Early research in the area, such as Altonji and Blank (1999), attributed this phenomenon mainly to differences in human capital (including the motherhood penalty and breaks in work experience), occupational choice as well as taste-based and/or statistical discrimination. However, in recent years, a large body of experimental economics research, surveyed comprehensively in Bertrand (2011), suggests that observed gaps may also arise as a result of gender differences in psychological attributes and preferences between men and women. Such differences may include gender differences in risk preferences, in attitudes towards competition and negotiation and in other-regarding preferences. This in

* Corresponding author.

E-mail addresses: sherry.li@auckland.ac.nz (Y. Li), e.sbai@auckland.ac.nz (E. Sbai), a.chaudhuri@auckland.ac.nz (A. Chaudhuri).

turn may also have implications for gender differences in occupational choice or work-place strategies.

Babcock and Laschever (2003) and Niederle and Vesterlund (2007) argue that greater female aversion to competition may explain why one finds fewer women occupying positions of power. Bohnet (2016) provides an excellent overview of the issues involved and suggests a series of institutional changes in order to achieve parity between the genders.¹ Our study is intended to add to this experimental economics literature by looking at whether men and women have different preferences and adopt different strategies in a game designed to simulate worker-employer relationships.

The experimental literature in this area is large. We refer the interested reader to Eckel and Grossman (2008) and Croson and Gneezy (2009) for comprehensive reviews primarily with regards to differences in risk attitudes and other-regarding preferences. Babcock and Laschever (2003) discuss gender differences in a variety of labor market transactions including the decision to enter into salary negotiations. For a selection of findings on the causes and consequences of gender differences from the psychology literature see Walters, Stuhlmacher, and Meyer (1998) or Sax (2005).

This research agenda has obvious overlaps with research in leadership; specifically whether men and women tend to adopt different styles when it comes to dealing with employees. Rosener (1990), building on concepts introduced by Burns (1978), argues that men typically tend to be “*transactional*” leaders and see job performance as a series of transactions with subordinates involving rewards for services rendered and punishments for inadequate performance. Women on the other hand are seen as being more “*transformational*”, relying less on explicit rewards and punishments and more on a democratic and participative style.

Eagly and Johnson (1990) undertake a meta-analysis of 162 studies on leadership and find little difference between male and female leadership styles. They find some support for the view that women adopt a more democratic style while men tend to adopt a more authoritative style.² Eagly, Karau, and Makhijani (1995) undertake a further meta-analysis which extends the analysis of leadership styles to the issue of leadership effectiveness. They report that men and women are equally effective as leaders except that men tend to be more effective in occupations that are typically defined in more masculine terms such as the military while women are more effective in occupations defined in primarily feminine terms such as nursing³.

In this study, we intend to contribute to this issue of *transformational* as opposed to *transactional* leadership using the well-known gift-exchange game paradigm. (Fehr, Gächter, and Kirchsteiger, 1997; Fehr, Kirchler, Weichbold, and Gächter, 1998; Fehr, Kirchsteiger, and Riedl, 1993, 1998, 1996; Fehr, Klein, and Schmidt, 2007.) We report results from experiments simulating an employer-employee relationship, where the interactions are mediated by two types of labor contracts: The first contract relies on mutual trust and reciprocity between the two; we refer to this as a “*relational*” contract. A *relational* contract is one that relies purely on *intrinsic* incentives in the form of mutual trust and reciprocity between the employer and the worker. There is no enforcement mechanism built into such relational contracts. Such relational contracts are designed to serve as a proxy for a *transformational* leadership style. A second “*transactional*” contract relies on explicit penalties for the worker if found to be shirking. These contracts emphasize *extrinsic*

incentives in the form of active monitoring of worker effort and explicit penalties if worker is caught shirking. These contracts are designed to resemble *transactional* types of leadership.⁴

In doing so, we complement results reported in Chaudhuri et al. (2015), who also look at similar intrinsic incentive based *relational* contracts and extrinsic incentive based *transactional* contracts. There are two ways this current study differs from Chaudhuri et al. (2015). First, in that earlier paper reciprocity was one-sided in that the employer has the option of reposing trust on the worker and the worker could reciprocate that trust or not. But given that prior research finds women to be more reciprocal than men (see Eckel and Grossman, 2008 for instance), we modify the relational contract to allow for two-sided trust and reciprocity as opposed to only reciprocity from the worker to the employer. In this study, both employers and workers have the opportunity to engage in reciprocity. The worker can reciprocate the employer’s trust by providing the effort asked for even in the absence of any explicit enforcement mechanism. Further, the employer can reward the worker’s reciprocity with an *ex post* bonus to the worker. Such an *ex post* bonus is not incentive compatible for the employer and any *ex ante* promise to pay such an *ex post* bonus is merely “cheap talk”. Fehr and his colleagues, in voluminous work alluded to above, argue that the possibility of such mutual reciprocity on both sides creates “strongly reciprocal contracts” and has the potential to be significantly welfare improving over contracts that allow for only one sided reciprocity, that from the worker to the employer only.

Among other things, Chaudhuri et al. (2015) found that when it came to transactional contracts with explicit penalties for shirking workers, women employers tended to mitigate the stick of the penalty with the carrot of higher rents toward workers; a strategy that led to lower earnings for the women employers in those contracts. However, in that study, the assignment to a particular institution was exogenous and the employers had no choice in this matter. Given the greater generosity on the part of our female employers in the context of penalty contracts, we felt that a plausible conjecture is that female employers were not comfortable with the transactional contracts and therefore, felt the need to concede greater rent to make up for imposing penalties for non-compliance.

This leads to our second modification where we explicitly allow the employers to choose the type of institution they wish to implement. In other words, the employer is no longer exogenously assigned to one of the two types of institutions but can actually choose over multiple rounds whether they wish to choose a *relational* contract (based on intrinsic motivations and mutual reciprocity) or a *transactional* contract (based on extrinsic incentives, exogenous enforcement mechanisms and penalties for workers caught shirking).

Our *ex ante* conjecture is that women will show a preference for the relational contract. However, we do not find evidence in favor of this

¹ In a complementary paper to this study, Li, Sbai and Chaudhuri (2020) show that greater female reluctance to vie for leadership roles may be the outcome of internalized beliefs regarding greater backlash from followers toward female leadership as opposed to male leadership. However, the authors do not find evidence for such backlash.

² The authors suggest that these differences may arise in part from the fact that women, being out-numbered by men as leaders, face greater resistance from employees and feel the need to seek greater employee input.

³ See Moran (1992) for a succinct overview of much of this work.

⁴ Laboratory gift exchange experiments are admittedly highly stylized. But the issue of relational versus transactional approaches in the work-place is complex and field studies are subject to multiple confounds. This often makes it hard to distinguish which features are universal and which are inextricably tied to a particular corporate culture. As a result, Camerer (2003) and Charness and Kuhn (2011) argue that, while stylized, lab studies can provide a reasonable and tractable model of labor market interactions. The particular experiments in this paper have the added advantage that they mitigate some of the artificiality of lab experiments and provide greater context by using terms such as employer, worker, wage, effort and fines. A final caveat here is that laboratory studies typically report greater gender differences in behavior than are found in actual field studies. Eagly and Johnson (1990) point out that lab studies where (1) people deal with strangers, typically for one-off or short-lived interactions and (2) there are fewer social cues on which to anchor behavior tend to exacerbate gender differences. We have noted the difficulties posed by field studies. One way to think of lab studies is that they may help identify areas where such differences may arise and also establish bounds on how significant those differences may turn out to be.

conjecture. Both male and female employers show an overwhelming preference for the relational contracts over the transactional contracts. We do replicate the Chaudhuri et al. (2015) result that women employers tend to be more generous and concede higher rent to the workers than they have to. This tends to lower female employer earnings. Our results echo the Chaudhuri et al. (2015) results in the sense that most of our findings are null findings. We do not find evidence of strong gender differences across a range of variables. Our results corroborate and replicate the earlier results of the lack of significant gender differences. These null findings are still of interest since they suggest that the differential outcomes for men and women in the workplace need not be attributed to differences in preferences. By and large, men and women do not seem to differ much in their choice of strategies. This potentially suggests a larger role for conscious and unconscious bias and possibly calls for a more pro-active stance on implementing institutional design changes along the lines suggested by Bohnet (2016) and Eckel et al. (2021) in order to achieve parity between the genders.

We proceed as follows. In Section 2 we provide an overview of the experimental design and procedures. In Section 3 we present our results. In Section 4 we provide some context for our results. We make some concluding remarks in Section 5.

2. Experimental Design and Procedures

2.1. Experimental Design

We apply a modified version of the principal-agent model from Fehr et al. (1997, 2007). The employer offers a wage $w \in [0.01, \dots, 10]$, and the worker exerts an effort level ranging in integers from one to ten, $e \in [1, \dots, 10]$, which generates output V . Output V is a function of worker effort e , i.e. $V = V(e) = e$. The value of effort monotonically increases in effort level such as $V(e) = 1$. Effort generates value for the employer but is costly to the worker where the cost is measured in monetary terms, $C = C(e)$ with $C'(e) > 0$, $C''(e) \geq 0$. We normalize the price of output to 1 so that revenue is equivalent to output, and worker's outside option is normalized to zero. Table 1 provides detailed parameters for the relationship between value of effort and cost of effort for each effort level. All payoffs denoted in the table represents actual monetary payoffs that participants will earn in our experiment. All payoffs are in New Zealand dollars.

Employers and workers interact for 10 rounds. We implement two types of matching protocol: *Fixed matching* for the entire session and *random re-matching* between every round. At the beginning when participants login, they are randomly assigned to the role of either an employer or a worker. These roles remain unchanged for all rounds. In the fixed matching treatment, each employer is matched with the same worker for the entire session of the experiment. In the random re-matching treatment, employers and workers are randomly re-matched with a different person each round.

A fixed matching protocol simulates long term relations between a pair; this allows for signaling future intentions and reputation building. This type of matching better represents work relationships in organizations with an established work-force, say a large corporation or university. Random re-matching, on the other hand, is designed to simulate a series of one-off (or short-term) interactions that often characterize some industries such as fast-food and retail with high rates of employee turn-over. In each round, employers move first by making a contract offer, they can choose a relational contract or a transactional contract.

Table 1
Output and Cost of Effort (\$).

Worker Effort Level:	1	2	3	4	5	6	7	8	9	10
Employer Value of Effort:	\$1.00	\$2.00	\$3.00	\$4.00	\$5.00	\$6.00	\$7.00	\$8.00	\$9.00	\$10.00
Worker Cost of Effort	\$0.01	\$0.10	\$0.20	\$0.40	\$0.60	\$0.80	\$1.00	\$1.30	\$1.60	\$2.00

We explain these contracts in the next section.

2.2. Relational contracts

This is a three-stage game. In stage 1, the employer makes a contract offer to the worker and the worker chooses to accept or reject the contract. The contract contains a wage rate w , a suggested effort level e^* and a proposed bonus amount b^* from the employer. The proposed bonus is a payment between \$0 and \$10 to the worker that is specified in stage 1. The actual payment, however, is selected after the worker's effort decision is revealed and the employer can observe whether or not the suggested effort level is met. The point to note is that both the suggested effort e^* and the proposed bonus b^* are non-binding; the workers are free to choose any level of effort and the employer can pay any amount of bonus, which can be different from the proposed bonus. There is no explicit enforcement mechanism to monitor worker effort or to detect worker shirking. Similarly, there is no mechanism to hold the principal to paying a previously announced bonus *ex post*, making any *ex ante* bonus announcement nothing more than "cheap talk". The interaction here then relies on intrinsic motivations in the form of mutual trust and reciprocity between employers and workers.

In Stage 2, the worker chooses to accept or reject the contract. If the worker rejects the contract, then both employer and worker earn zero in that round. On the other hand, if the worker accepts the contract, then the worker decides an actual effort level to put in. This effort level can be greater than, less than or equal to the effort level asked for by the employer. There is no way to enforce this effort and no penalties for shirking. At stage 3, the employer observes the worker's actual effort level and then choose an actual bonus amount $b \in [0, \dots, 10]$ to pay the worker. The actual bonus payment may be any amount between \$0.00 and \$10, and may be greater than, equal to, or less than what the employer had proposed earlier in Stage 1. The worker cannot enforce the employer to pay a bonus for fulfilling the terms of the contract even if such a bonus was promised by the principal.

Any time a contract is accepted the employer's payoff is given by the value of the output minus the wage and the bonus (if any). The worker's payoff is the wage plus the bonus (if any) minus the cost of effort. If we assume that both the employer and the worker are self-interested and wish to maximize monetary returns, and especially if the two are engaged in a purely one-shot interaction, then we would expect that *ex post*, the employer has no incentive to pay any positive amount of bonus. Consequently, if we also assume that the worker knows that the employer is self-interested, the worker has no incentive to provide non-minimal effort. Anticipating this, and assuming that the employer is also self-interested and knows that the worker is self-interested too, the employer should offer the lowest possible wage. Therefore, the employer should offer a wage rate of 0.01, the worker should choose the lowest effort level of 1, and employer should pay zero bonus. The corresponding earning for the employer would be $\pi = V(e) - w - b = \$1 - \$0.01 - \$0 = \0.99 , and the payoff for the worker would be $u = w - C(e) - b = \$0.01 - \$0.01 - \$0 = \0 . This is in keeping with the equilibrium prediction in one-shot plays of the game as well as the subgame perfect equilibrium in repeated interactions with random re-matching which simulate one-shot games.

On the other hand, if the employer offers a positive rent by paying the worker a premium over the cost of effort, defined by $r = w - C(e)$ and the worker reciprocates this positive rent by choosing non-minimal effort, then both the employer and the worker are better off than in the

self-interested equilibrium of the game as described above. This would be particularly true in fixed matchings that allow for signaling future actions and reputation building. This suggests that while we should observe minimal trust and reciprocity in games with random re-matching, we should observe higher trust and reciprocity in games with fixed matching.

2.3. Transactional contracts

This contract allows the employer to impose a fine to penalize workers if they are shirking. The worker is shirking if the actual effort level is less than the suggested effort level, i.e. $e < e^*$. The employer must invest in monitoring technology in order to detect shirking workers. The monitoring technology is not perfect and can only detect shirking situations with a probability $p = 0.33$. By investing in this monitoring technology, the employer incurs a fixed cost k , and can impose a fine on the worker. The fine is paid to the employer only if shirking is verified.

This interaction also proceeds in three stages. In stage 1, the employer offers a wage w and chooses a specified effort level e^* . At this stage, the employer also decides whether to invest in monitoring technology at a cost of k and the level of fine f to impose on the worker. The cost of implementing the monitoring technology is fixed and equal to \$1. The level of fine ranges from 0 to 1.3 in one decimal place increments, i.e. $f \in [0, \dots, 1.3]$. The fine can be collected only if shirking is detected, which happens with a 0.33 probability.

In the second stage, the worker chooses to accept or reject the contract after observing the employer's choices in stage 1. This includes not only the wage rate and effort level specified, but also whether the employer invested in a monitoring technology or not, and the level of fine the employer chose to impose. If the worker accepts the contract, then the worker chooses an effort level e . If the contract is rejected by the worker, then both employer and worker earn zero.

The third stage is brief. In this stage, the verification process takes place; the monitoring technology verifies, with a probability of 0.33, whether shirking happened or not; i.e. whether the actual effort is below the employer's asked for effort level. It is worth noting the following. In the context of transactional contracts, it is always the case that the employer can see the worker effort and therefore knows whether the worker shirked or not. But, the key issue is whether this shirking can be proved or not via the monitoring technology. Consider CCTV surveillance, designed to pick up shirking but not with certainty. So, even if the worker shirks the employer may or may not be able to *prove* that shirking. In the context of the actual experimental software, in stage 3, the employer gets to learn whether the verification technology has actually caught the worker shirking or not. The shirking penalty is imposed on the worker if and only if the monitoring technology actually catches the worker shirking (which happens with a probability less than one) but not otherwise.⁵

It is easy to show that, in this setting, payoff for the employer is maximized by investing in the monitoring technology, imposing the maximal fine of 1.3 units and asking for an effort level of 4 units. The employer will offer that wage to the worker that just about compensates the worker for the cost of this effort, which is \$0.4, or a little bit more. In this case, the employer will earn \$2.60 and worker will earn zero. This follows from Fehr et al. (1997, 2007). We also provide a simple derivation of this in Appendix C. This outcome is better than the purely self-interested outcome in relational contracts. Self-interested motivations suggest that relational contracts should elicit only the smallest

⁵ E.g., as teachers in academia, many of us have experience with situations when we are sure that a particular student cheated on an assignment but depending on the nature of the assignment, proving this is often difficult. The idea here is similar. We thank an anonymous referee for asking us to highlight this distinction, which is subtle but important.

Table 2
Number of subjects in different treatments.

	Fixed matching	Random re-matching
Employers	N=33 Male=19 Female=14	N=38 Male=22 Female=16
Employees	N=33 Male=13 Female=20	N=38 Male=20 Female=18
Total	N=66 Male=32 Female=34	N=76 Male=42 Female=34

Table 3
Average proportion of contracts chosen across genders.

	Fixed matching	Random re-matching
Male employers	Transactional=12% (Relational=88%)	Transactional=13% (Relational=87%)
Female employers	Transactional=17% (Relational=83%)	Transactional=21% (Relational=79%)
Non-parametric Wilcoxon ranksum test z	-0.94	-2.54
p-value	0.35	0.01
Observations	Overall=33 Male=19 Female=14	Overall=38 Male=22 Female=16

possible effort levels and therefore the employer should be monetarily better off with transactional contracts.

2.4. Task and Questionnaire

Following the experiment, participants are asked to take part in the Holt and Laury (2002) lottery choice experiment. This is shown in Appendix D. We use this task to elicit risk preferences, which is one of the co-variables in further analysis. The switching point from option A to option B is used to measure each individual's risk preference. For example, a risk-neutral person would choose option A for the first four rows and then switch to option B in the fifth row. A risk-loving individual would switch to option B before the fifth row, while a risk-averse individual would switch after the fifth row. Participants are informed that this is a separate task. At the end of the session, the computer will randomly choose a row and will pay for either Option A or Option B, depending on what the subject chose for that specific row.

Finally, participants are asked to fill out a demographic questionnaire (Appendix B) used to collect information regarding participant's gender, field of study, year in the undergraduate program, age, income, whether they were born in New Zealand, and their ethnicity. This demographic questionnaire is similar to the one Statistics New Zealand uses to collect Census data.

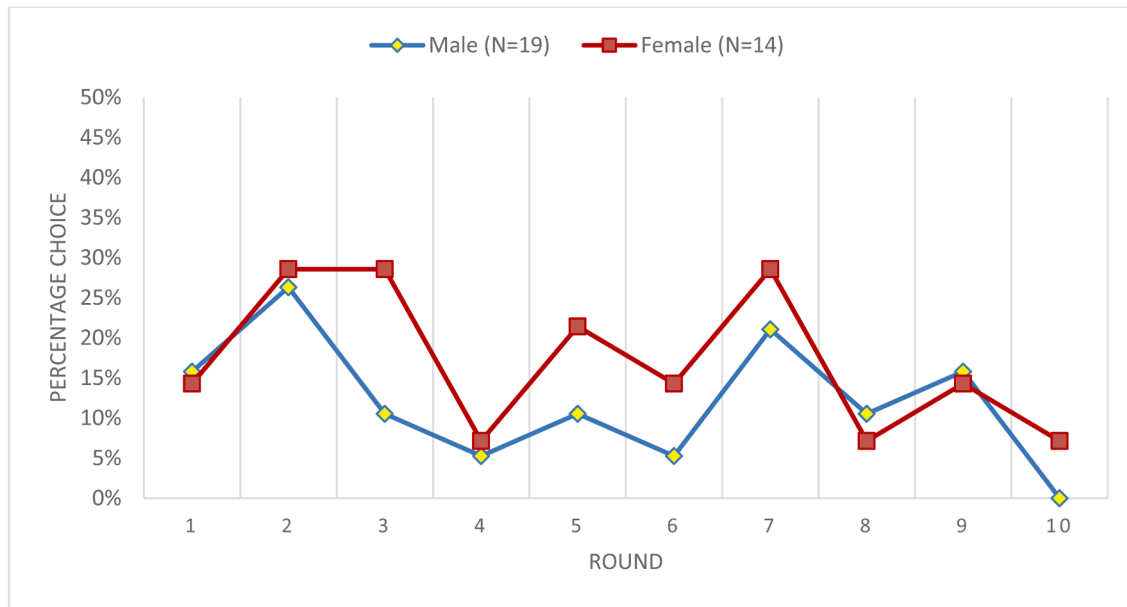
2.5. Experimental Procedures

A total of 142 subjects took part in this experiment. Table 2 provides summary statistics for our experiment. All sessions in this experiment were conducted in the DECIDE laboratory at the University of Auckland using Veconlab⁶, an online software developed by Charles Holt at the University of Virginia. Participants were recruited via an email announcement and they were students from undergraduate courses without any prior experience with the principal-agent game. We have 33 pairs in the fixed matching treatment and 38 pairs in the random re-matching treatment.

Participants are directed to computer cubicles once they enter the lab. There are dividers between each cubicle so that each participant is

⁶ (Holt, 2009) <http://veconlab.econ.virginia.edu/admin.htm>.

A: Choice of transactional contract over time under fixed matching



B: Choice of transactional contract over time with random re-matching

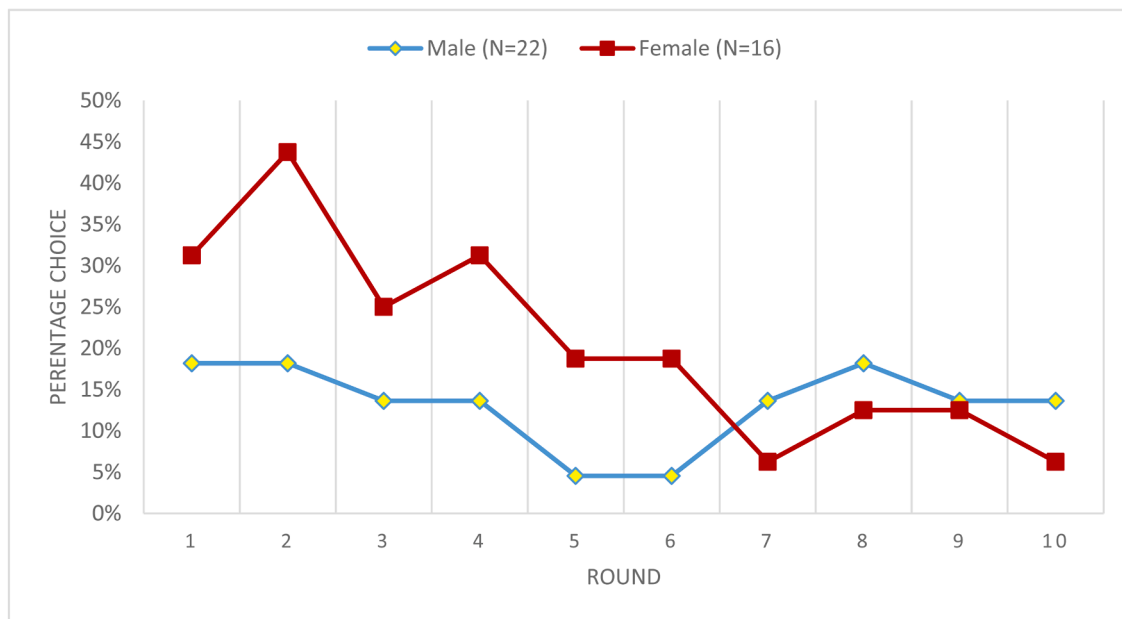


Fig. 1. (A) Choice of transactional contract over time under fixed matching. (B) Choice of transactional contract over time with random re-matching.

separated from another and is unable to see any other participant's computer screen. This prevents any opportunities to observe the decisions made by any other participants during the experiment. They are also cautioned against communicating with others. At the beginning of the experiment, the experimenter distributes the game instructions (included in Appendix A) to participants and reads it out loud. The instructions include a summary of the game with some numerical examples and practice questions regarding payoffs from relational and

transactional contracts. Participants are given 10 to 15 minutes to read the instructions and work on the numerical practice questions. All participants' answers are checked, and any questions participants might have are answered. Participants are told that the experiment consists of 2 parts: there are 4 practice rounds to help them get familiar with the game followed by 10 actual rounds, where they will be earning money. Once the actual game starts, the computer program will keep track of their total earnings for all ten rounds, and these will be shown as

“cumulative earnings” on a results page. The earnings from the practice rounds are also shown on the screen but participants know that this will not be added to their total earnings.

At the end of each round, participants can observe their individual earnings for that round as well as their cumulative earnings. However, they do not have any information about the earnings of any other players, including the players they are paired with in any round. However, they can use the parameters given in the instructions to calculate their pair member’s earning, should they wish to do so.

Participants know that they are randomly assigned the role of either an employer or a worker when they log in to the Veconlab website. This role will remain unchanged for the entire duration of the experiment, that is, for both 4 practice rounds and 10 actual rounds. Each participant is also assigned a subject identification number (ID), and they are unable to learn the identity of other participants. At the completion of the 10 rounds of the game, participants are asked to take part in the Holt-Laury lottery choice game and fill out the demographic questionnaire.

Each session lasts approximately 90 minutes. At the end of the session, participants are paid their earnings in cash from the principal-agent game and from the lottery choice task, plus a show-up fee. Participants are told that their earnings are private information and they were free to leave after collecting their payment.⁷ Average earnings for the 10 rounds are \$23 for employers and \$26 for workers, not including payment from the lottery choice experiment and the \$5 show-up fee.

2.6. Hypotheses

Based on our review of prior findings in the literature, we propose the following hypotheses.

Hypothesis 1 : Female employers are more likely to choose a higher proportion of trust-based relational contracts as opposed to penalty-based transactional contracts.

Hypothesis 2: Given existing evidence in favour of greater female generosity, female employers will concede higher rent to the workers in both types of contracts; this will result in lower earnings for female employers.

Hypothesis 3: Given existing evidence in favour of greater female reciprocity, when it comes to relational contracts, compared to male employers, female employers will either pay a higher bonus or pay the bonus more often or both.

⁷ A common difficulty with this game is that employers often end up with negative earnings. This is true for many of the studies we have cited including the one most closely related, Chaudhuri et al. (2015). Given that taking money away from subjects is not an option, experimenters typically restrict earnings to zero in the event of negative earnings. But this limited liability creates potential confounds because once in negative territory, subjects may not pay adequate attention. There is no longer any incentive to make careful decisions at this point since they cannot earn anything less than zero. This leads to a loss of reward salience. To get around this (and any other loss aversion type arguments), we provide an additional \$10 dollars to the employer without providing any further information regarding the reason for this additional endowment. The relevant paragraph in our instruction states: *The employers will be provided with an additional endowment of \$10.00 at the start of this experiment. This amount will be added to the earnings of the employers at the end of the experiment. This is in addition to the show-up fee of \$5.00 paid to every participant.* Given that this is a lump-sum transfer paid at the conclusion of the experiment, rather than based on individual decisions, this should not distort decisions. A potential worry is that workers may suffer from a sense of inequity but as the game progresses, the workers get to see that on average the employer earnings are not very high. We believe that this mitigates any inequity concerns. In any event, one needs to trade-off between inequity concerns and concerns caused by negative earnings. We believe that the latter is a bigger worry than the former.

Hypothesis 4: Given existing evidence in favour of greater female reciprocity, controlling for rent, we expect female workers to shirk less in both types of contracts.

3. Results

In this section we report on our findings. We will report four separate results, each built around one of the hypotheses stated above.

Result 1: Female employers are no more likely to choose trust-based relational contracts than male employers; in fact, they are more likely to choose penalty-based transactional contracts under random re-matching protocol though this difference dissipates over time. There are no significant gender differences in contract choice under fixed matching protocol.

Our first hypothesis was that female employers will show a preference for the relational contract. This is not borne out; if anything, compared to male employers, female employers show a greater preference for the transactional contract under random re-matching. Both men and women choose the relational contract much more frequently than the transactional contract. This is true with both fixed matching and random re-matching. Table 3 shows the average proportion of each contract type as well as results of non-parametric ranksum tests where we take the proportion of transactional contracts chosen by each employer over ten rounds as the unit of observation. We find that, on average, women choose transactional contract significantly more than men under random matching treatment, but not under the fixed matching treatment. Figs. 1A and 1B shows the time-series of contract choices over time.⁸

In Table 4, we provide further corroborating evidence in the form of random effects regression results on contract choice, where we control for other relevant co-variables. The dependent variable is equal to 1 if the transactional contract is chosen, zero otherwise. The regressors include a dummy for Female (=1 if employer is female, 0 not); Round, an interaction term between Female and Round; how much the employer earned in the previous round and finally Lag Shirk, a dummy for whether the paired employee shirked in the previous or not. Lag Shirk =1 if the employee shirked in the previous round, and 0 otherwise.⁹ When we have fixed matching, one employer-employee pair can be treated as an independent observation. Conversely, in random re-matching, the only independent observation is the session itself. Given this, in Table 4 and for other regressions reported below, we cluster errors at the level of individual subjects. We have only three sessions in each treatment, therefore clustering on sessions may also lead to imprecise estimates.

We present the results for fixed matching and random re-matching separately. In each case we present three separate models. The first one includes only the regressors noted above. The second one controls

⁸ It is certainly noteworthy and somewhat curious that the vast majority of contracts chosen are relational rather than transactional. It has been noted in the prior literature that contrary to the theoretical prediction, transactional contracts do not generate higher earnings for the employer, rather relational contracts do. It is our conjecture that the tendency to choose the relational contract in our study was exacerbated by the presence of practice rounds where it is likely that the employers became conscious of the earnings differentials in the two types of contracts. In the absence of those practice rounds we would have likely seen more transactional contracts being chosen in the earlier rounds of the actual experiment. Some evidence for this conjecture is provided by the fact that there is a tendency among female employers to choose transactional contracts more in the earlier rounds (particularly with random re-matching) but even these employers move away towards relational contracts over time.

⁹ As a robustness check and also for the sake of completeness, we also re-did all of these regressions by replacing Lag Shirk with Lag Effort, the effort exerted by the employee in the previous round. The results are unchanged and therefore, we have chosen to report to report the results with the Lag Shirk regressor here. The results with Lag Effort are available upon request.

Table 4
Random effect probit regression for contract choice with errors clustered on subjects.

	Fixed matching			Random re-matching		
	(1)	(2)	(3)	(1)	(2)	(3)
Female	0.551 (0.628)	0.454 (0.695)	0.315 (0.760)	1.862*** (0.737)	2.622*** (0.815)	2.753*** (0.945)
Round	-0.093* (0.055)	-0.092* (0.056)	-0.091 (0.056)	0.013 (0.092)	0.076 (0.111)	0.074 (0.109)
Female*Round	-0.028 (0.077)	-0.029 (0.080)	-0.030 (0.080)	-0.209* (0.114)	-0.284** (0.132)	-0.284** (0.131)
Lag Earnings	0.009 (0.099)	0.002 (0.103)	0.001 (0.106)	-0.137* (0.079)	-0.020 (0.072)	-0.012 (0.070)
Lag Shirk	0756** (0.337)	0.661* (0.335)	0.661* (0.373)	-0.896** (0.417)	-0.415 (0.408)	-0.384 (0.403)
Risk Averse		0.285 (0.543)	0.154 (0.523)		0.217 (0.426)	0.077 (0.523)
Constant	-1.513*** (0.565)	-1.525*** (0.620)	-1.352* (0.757)	-1.335** (0.652)	-2.425*** (0.832)	-2.132** (0.835)
Demographic Control	No	No	Yes	No	No	Yes
Log likelihood	-100.078	-92.694	-89.097	-109.047	-86.479	-84.663
Wald χ^2	13.368	11.658	14.189	22.929	18.715	22.015
Prob > χ^2	0.02	0.07	0.16	0.00	0.01	0.06
Number of observations	297	270	243	342	297	297

Dependent variable: Contract Choice = 1 if transactional contract chosen; 0 if relational contract chosen.

Notes: Standard errors in parentheses; ***, ** and * denote significance at 1%, 5% and 10% respectively.

for risk preferences using the Holt-Laury lottery choice task and the third specification controls for demographic characteristics as well.¹⁰ Very few of the coefficients are significant with fixed matching except for the one for Lag Shirk, which is positive and marginally significant. This makes intuitive sense; when the employee shirked in the previous round, the employer responded by moving to a transactional contract in the next round. This result is significant at 5% in the first specification but becomes only marginally significant after we control for risk preferences and demographics.

Turning to the results for the random re-matching treatment, we find that consistent with the ranksum tests in Table 3, compared to male employers, female employers are more likely to choose a transactional contract under the random re-matching treatment. However, this difference declines over time as indicated by the negative and significant coefficient for the interaction term of female and round. This finding is certainly contrary to our hypothesis that women will choose relational contracts over transactional contracts. But at the same time it is also the case that we would expect employers of both genders to choose the transactional contract more often with random matching. This is because relying on intrinsic incentives makes more sense in the context of long-term and repeated interactions which allow for signaling and reputation building while when it comes to short-term relations, it is expected that more employers will choose to rely on extrinsic motivations. It is noteworthy that female employers seem more attuned to these underlying incentives than male employers.

Looking at the random re-matching protocol, we find that in the first specification the coefficients for lagged earnings and lag shirk (whether the worker shirked in the previous round or not) are both negative and significant. It is important to note that with random re-matching subjects are not playing one another for more than one round and any effects here need to be interpreted in the context of the whole session. The result here suggests that when earnings or shirking rises, employers are less likely to choose the transactional contract in the current round. This is not surprising since prior work in the area generally find that both

¹⁰ We lose quite a few observations when we control for risk preferences and demographics. This is mostly because of subjects who make inconsistent choices on the Holt-Laury lottery choice task. This is a well-known drawback of this task. Therefore, when we control for risk aversion, we exclude these subjects who made inconsistent choices resulting in a loss of observations. We also lose a few observations for those subjects who did not complete the demographic survey.

effort levels and earnings are usually higher with relational contracts rather than transactional ones. This suggests that at the level of the session an increase in shirking in one round or an increase in earnings in one round led to an increase in the choice of relational contracts. We also note that once we control for risk preferences and demographics, both these coefficients become non-significant, and that there is a change in the level of estimated effects associated with the female dummy. This further suggests that some of the results in the first specification are explained partly by the interaction between gender and risk preferences and may be the result of an omitted variable bias.

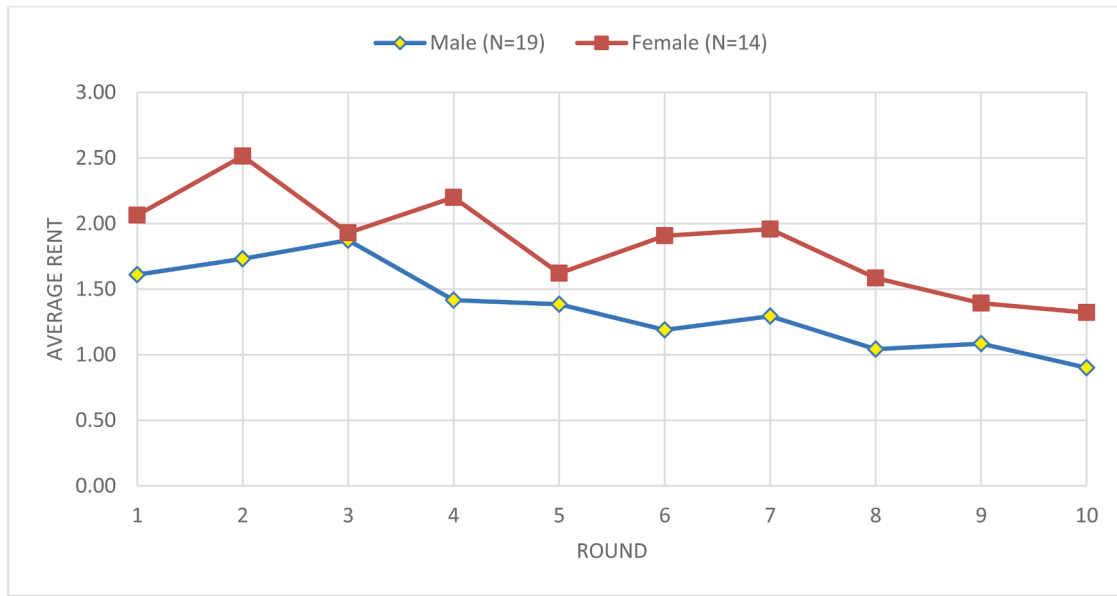
The results in Table 4 suggest that we do not find systematic gender differences in the choice of contracts; if anything, women were marginally more likely to choose a transactional contract, compared to men. However, this difference also dissipates over time. Importantly, there is one area where we find some evidence in keeping with our initial hypothesis. If we look within the transactional contracts at the patterns of fine usage, we find that on average, women impose a fine of \$1.10 compared to \$1.23 imposed by men. Men tended to impose the optimal fine in 88% (46/52) of the contracts while women did so in only 65% of contracts (37/57). This is a significant difference using two-sample proportions test ($z=2.96$; $p<0.01$). However, as noted already, such transactional contracts constitute a small proportion of choices.

Result 2: Female employers offered higher rent to workers under random re-matching; we find that this results in lower earnings for the female employers under that protocol. There are no gender differences in either rent offered or earnings under fixed matching.

This result corroborates our second hypothesis that female employers will offer higher rents. As explained above, for any contract, the employer offers a wage rate and suggests an effort level. Given that the worker's reservation wage has been normalized to zero, the employer has to pay the worker the cost of effort (or epsilon more) in order for the worker to accept the contract. Any amount higher than that implies worker receives rent from the employer. Figs. 2A and 2B show average rent offered by male and female employers over 10 rounds in both fixed matching treatment and random re-matching treatment. We can see that female employers on average offer a higher rent in all 10 rounds compared to male employers.

In Table 5, we look at regression analysis for rent offered to examine the possible differences between male and female employers because the non-parametric test does not control for covariates. We apply random effects model with robust standard errors clustered on individuals to account for any individual-specific components of choice behavior. We

A: Average rent offered over 10 rounds under fixed matching



B: Average rent offered over 10 rounds under random re-matching

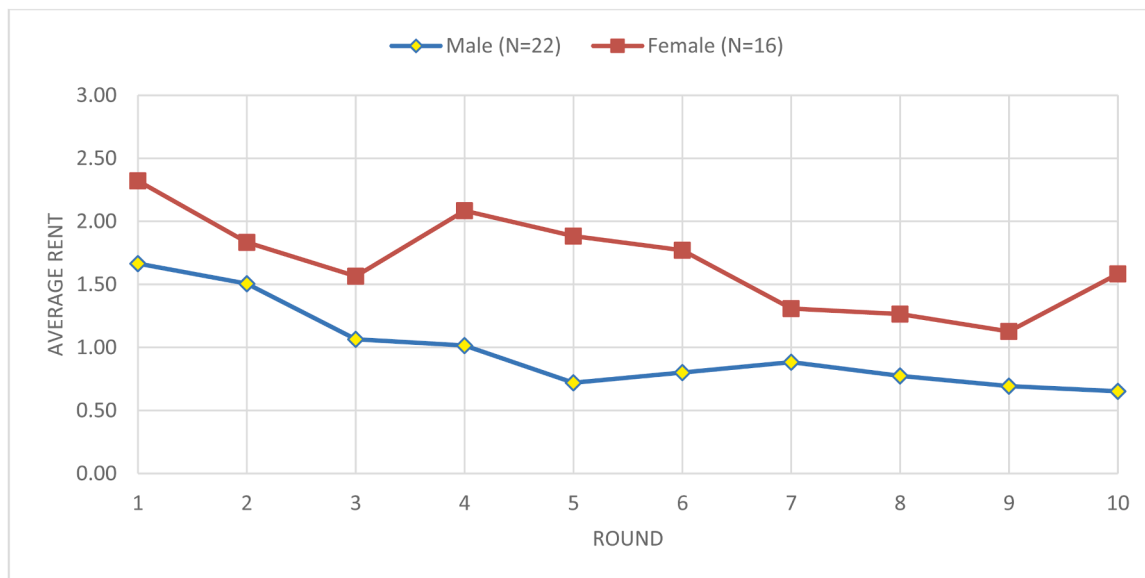


Fig. 2. (A) Average rent offered over 10 rounds under fixed matching. (B) Average rent offered over 10 rounds under random re-matching.

use a random effects specification here because each decision regarding rent offered involves the specific choice behavior of each subject.¹¹

The dependent variable is the rent offered by employer i in round t . The independent variables include (1) female (a dummy variable =1 for female and =0 for male); (2) round; (3) an interaction term between female dummy and round to detect any gender specific trends in rent over time; (4) lag cumulative earnings; (5) lag worker effort; (6) risk-

averse (=1 for risk aversion and =0 otherwise); (7) a dummy variable for contract type (=1 for transactional contract, =0 for relational contract); (8) a vector of other demographic variables. The coefficient for round is negative and significant in both treatments indicating that the rent offered by employers decline over time. This result is similar to ones earlier ones reported in Chaudhuri et al. (2015) and Chaudhuri and Sbai (2011). The female dummy is only significant under random matching treatment, consistent with our results for ranksum tests. Employers choosing the transactional contract under the fixed matching treatment are more likely to make a higher rent offer. The coefficient for transactional contract is marginally significant (at 10%) in the fixed matching protocol. This suggests that both genders offer higher rent in this type of

¹¹ We also use other models along with fixed effects and correlated random effects for robustness check. When controlling for fixed effects or correlated random effects, the coefficient for transactional contract is no longer significant while other coefficients and significance level remain similar.

contract but as noted the effect is only marginal and therefore, we do not elaborate on this further.

In Table 6, we present results for random effects regressions for earnings per round. We do not find any significant gender differences in earnings under fixed matching treatment, however, we do find that that earnings are lower in transactional contracts compared to the relational contracts.¹² With random re-matching, female employers earn less compared to male employers though we note that the interaction term between the female dummy and round is positive and significant. This suggests that over time this difference in earnings reduces. The result is not surprising given that female employers tend to make higher rent offers to the workers. It is also the case that under random re-matching women tend to choose transactional contracts more often and these contracts perform worse in terms of earnings. Over time women move away from these penalty-based contracts and this results in a narrowing of the earnings gap between men and women. The coefficient of the rent variable is negative and significant under both fixed matching treatment and random matching treatment suggesting that contracts with higher rent offers resulted in lower earnings for the employer. We do not find any clear gender differences in earnings of the worker for either of the treatments.

Result 3: Male employers tend to propose a higher bonus and pay a higher bonus on average.

This result deals with our third hypothesis that within relational contracts, women will tend to be more generous by offering higher bonuses. This conjecture turns out to be incorrect. In Table 7, we present ranksum tests for proposed bonus and actual bonus in round 1 for both fixed matching treatment and random matching treatments. We look at Round 1 choices only since these can be considered to be independent observations. On average, men proposed \$2.25 in round 1 while women proposed \$1.5, and, on average men paid an actual bonus of \$1.81 in round 1 while women paid an actual value of \$0.75. Both the proposed bonus and the actual bonus are marginally higher (significant at 9%) for men compared to women under the fixed matching treatment. However, men pay a lower actual bonus under random matching treatment than women.

Result 4: There is no significant gender difference in shirking under relational contracts. However, we find that female workers are more likely to shirk under the transactional contract, but this difference disappears over time.

This result has to do with our fourth and final hypothesis suggesting that female employees, being more reciprocal, in general, will tend to shirk less, controlling for the level of rent offered by the employer. We use a random effects probit regression model with shirking as the dependent variable. We define the dummy variable shirking = 1 if actual effort is less than suggested effort and otherwise shirking = 0. We will present the results for relational and transactional contracts separately since the set of regressors are different in the two cases as explained below.

In Table 8 we present results for relational contracts both with fixed matching and random matching. The independent variables include (1) the rent offered by the employer, which can be viewed as an intrinsic incentive for the workers to choose the suggested effort level; (2) female; (3) round; (4) an interaction term between the female dummy and the round; (5) the proposed bonus; (6) an interaction term between the female dummy and the proposed bonus; (7) lag earnings; (8) risk-averse (=1 for risk aversion and =0 otherwise); (9) a vector of other demographic variables. However, with the fixed matching protocol we

¹² As noted already, it is a common finding in the literature that contracts relying on mutual trust and reciprocity often lead to higher efficiency compared to penalty-based contracts. This finding is true for some of the original papers by Fehr and his co-authors, such as Fehr et al. (1997). This finding is also reported in the closely related Chaudhuri et al. (2015) study. Therefore, we refrain from elaborating further on this.

provide two separate specifications, 1A and 1B. Specification 1A includes the regressors described above but for Specification 1B we include two additional regressors. In the fixed matching protocol for relational contracts, a worker not only knows whether a bonus is being offered as part of the current period's contract but also whether a bonus was promised and fulfilled for the previous period's contract. We define a variable "lag bonus honored" (equal to 1 in all cases where a bonus was offered and fulfilled, zero otherwise). In Specification 1B, we add this lag bonus honored variable and also add an interaction term between this variable and the female dummy.

For both specifications 1A and 1B, the amount of rent offered does not have a significant effect on the decision to shirk or not. The amount of shirking increases over time. The amount of bonus offered by the employer does make a difference and higher bonuses promised (even though this is akin to cheap-talk) leads to lower shirking. But, this positive impact of the bonus on shirking is lower for female employees, as shown by the negative coefficient for the interaction term of female and proposed bonus. Looking at specification 1B, it does not seem to matter whether the bonus promised was fulfilled or not (the coefficient for lag bonus honored is not significant). Compared to male employees, the negative impact of proposed bonus on shirking is less influential for female employees in both specification 1A and 1B. Looking at relational contracts under random re-matching, we find that the likelihood of shirking among female employees declines over time.

Under a transactional contract, our independent variables include: (1) the rent offered; (2) extrinsic incentive;¹³ (3) female; (4) round; (5) an interaction term between the female dummy and the round; (6) an interaction term between the female dummy and extrinsic; (7) values of the fine imposed; (8) risk-averse (=1 for risk aversion and =0 otherwise); (9) lag cumulative earnings; The results are presented in Table 9. Here we have combined the data for the fixed matching and random re-matching protocols because, as noted earlier, the majority of the time the employers chose the relational contract. This implies that there are few observations for the transactional contract and if we further divide this by the matching protocol then we have very few observations from which to estimate.

As with the relational contract, rent does not seem to have an effect on shirking. However, for the transactional contracts, the coefficient for extrinsic incentive is negative and significant which suggests that workers do pay attention to the payoff differential from not shirking as opposed to shirking. When the payoff differential increases, workers responded by lowering their likelihood of shirking. The coefficient on the female dummy is negative and significant indicating that female workers are more likely to shirk compared to male workers; though this effect decreases over time.

4. Discussion

In this section we provide a very quick overview of our major findings and their potential implications.

First, contrary to our hypothesis that women will prefer intrinsic motivation based relational contracts over men, employers of both genders preponderantly chose relational contract over transactional

¹³ In the transactional contract treatment, workers face a trade-off between shirking or not. If a worker shirks then he/she incurs a lower cost of effort but may be found out and therefore penalized. On the other hand, not shirking means incurring a higher effort cost but no possibility of being fined. The extrinsic incentive ($U^{ms} - U^s$) variable is defined as the expected payoff to the worker from not shirking minus the expected payoff when the worker does shirk. The value of this variable depends on the wage offered and the effort specified and follows from the theoretical model presented in the Appendix C. Suffice it to say that in the context of the transactional contracts, the rent acts as an intrinsic incentive to provide effort while this other variable provides a proxy for the extrinsic incentive to provide effort.

Table 5
Random Effects Regression on Rent Offered (errors clustered on subject).

	Fixed matching		Random re-matching	
	(1)	(2)	(3)	(4)
Female	0.685 (0.603)	0.829 (0.590)	1.088** (0.467)	0.935** (0.426)
Round	-0.107*** (0.032)	-0.107*** (0.033)	-0.053*** (0.018)	-0.053*** (0.018)
Female*Round	-0.008 (0.069)	-0.008 (0.070)	-0.041 (0.039)	-0.041 (0.040)
Lag Earnings	-0.017 (0.030)	-0.016 (0.030)	0.000 (0.019)	0.001 (0.019)
Transactional Contract	0.507* (0.297)	0.519* (0.298)	0.243 (0.197)	0.249 (0.196)
Risk Averse	-0.557 (0.366)	-0.518 (0.367)	0.085 (0.354)	0.030 (0.410)
Constant	2.118*** (0.461)	2.051*** (0.448)	1.112*** (0.229)	1.085*** (0.251)
Demographic control	No	Yes	No	Yes
Wald χ^2	33.30	38.70	31.04	41.70
Prob > χ^2	0.0000	0.0000	0.0000	0.0000
Number of observations	270	270	297	297

Dependent variable: Rent offered.

Notes: Standard errors in parentheses; ***, ** and* denote significance at 1%, 5% and 10% respectively.

Table 6
Random Effects Regression on Employer Earnings (errors clustered on subject).

	Fixed matching		Random re-matching	
	(1)	(2)	(3)	(4)
Rent	-0.292** (0.140)	-0.219* (0.121)	-0.600*** (0.152)	-0.609*** (0.145)
Female	-0.293 (0.869)	-0.660 (0.850)	-2.079*** (0.624)	-2.243*** (0.622)
Transactional Contract	-2.434*** (0.374)	-2.583*** (0.374)	-0.752** (0.328)	-0.688* (0.364)
Round	-0.024 (0.078)	-0.020 (0.076)	-0.218*** (0.050)	-0.218*** (0.050)
Female*Round	0.047 (0.125)	0.049 (0.125)	0.261*** (0.090)	0.264*** (0.090)
Suggested Effort	0.257*** (0.072)	0.276*** (0.091)	-0.064 (0.047)	-0.063 (0.046)
Risk Averse	0.075 (0.387)	0.189 (0.321)	-0.017 (0.257)	0.092 (0.303)
Constant	0.898 (0.878)	0.456 (0.771)	3.506*** (0.548)	3.436*** (0.577)
Demographic Control	No	Yes	No	Yes
Wald χ^2	79.07	149.42	67.23	96.51
Prob > χ^2	0.0000	0.0000	0.0000	0.0000
Number of observations	300	300	330	330

Dependent variable: Employer Earnings.

Notes: Standard errors in parentheses; ***, ** and* denote significance at 1%, 5% and 10% respectively.

Table 7
Ranksum tests for proposed and actual bonus by gender and treatment in Round 1.

Fixed matching treatment		Random re-matching treatment	
Proposed	Actual	Proposed	Actual
Male=2.25	Male=1.81	Male=3.28	Male=0.28
Female= 1.50	Female= 0.75	Female= 2.64	Female= 0.73
z=1.90	z=1.69	z=0.74	z=-1.66
p=0.06	p=0.09	p=0.46	p=0.10
Male=16, Femlae=12		Male=18, Female=11	

contracts. Furthermore, over time there is a move away from transactional toward relational contracts. This may be justified by the fact that the relational contracts outperform transactional contracts in terms of employer earnings, especially because one needs to subtract the costs of monitoring to get net earnings. It is likely that this lesson was highlighted for the employers from the practice rounds that they played prior

Table 8
Random Effects Probit regression on shirking under Relational Contract (standard errors clustered on the subject).

	Fixed matching		Random re-matching
	(1A)	(1B)	(2)
Rent	-0.192 (0.173)	-0.181 (0.160)	-0.035 (0.082)
Female	-0.119 (0.683)	0.927 (0.828)	0.553 (0.680)
Round	0.108* (0.058)	0.112* (0.068)	0.079 (0.054)
Female*Round	-0.006 (0.079)	-0.104 (0.086)	-0.182** (0.079)
Proposed Bonus	-0.443*** (0.144)	-0.395*** (0.139)	-0.003 (0.065)
Female*Proposed Bonus	0.323* (0.185)	0.262 (0.160)	-0.010 (0.100)
Lag bonus honored	-	-0.917 (0.574)	-
Female*Lag bonus honored	-	1.450** (0.621)	-
Lag Earnings	0.021 (0.107)	-0.048 (0.133)	0.025 (0.058)
Risk Averse	-0.597 (0.410)	-0.413 (0.316)	-0.331 (0.244)
Constant	-0.188 (1.143)	-0.325 (1.298)	0.032 (0.770)
Demographic Control	Yes	Yes	Yes
Wald χ^2	35.30	78.50	52.02
Prob > χ^2	0.0000	0.000	0.000
Number of observations	206	206	262

Dependent variable: Shirking = 1 if actual effort < suggested effort; 0 otherwise.

Notes: Standard errors in parentheses; ***, ** and* denote significance at 1%, 5% and 10% respectively.

to the money rounds. In the absence of those practice rounds we would probably have seen a large proportion of transactional contract choices in the earlier rounds.

Also contrary to our hypothesis, we find that female employers exhibit a higher propensity to choose the transactional contract early on in the game but only with random re-matching and they too move away towards the relational contract over time.

In keeping with our hypothesis, we find evidence that regardless of contract type, female employers offer higher rent to the workers. This is in line with prior findings regarding greater female generosity. Offering higher rents leads to lower earnings for those female employers. So, while female employers are offering higher rents, they are not receiving commensurately higher effort from the workers to make up for this. If one extrapolates to real-life firms, then the conclusion would be that female-led firms may end up with lower earnings. This may have implications for both the gender wage gap and the gender leadership gap. If female leaders are perceived as less effective, this may hamper their promotion prospects and if there are more men than women in higher paying positions then this will also show up as a gender wage gap.

Over time, and not surprisingly, the rent offered declines but the rate of decline is similar for both male and female employers as shown by the insignificant coefficient of the interaction term between round and the female dummy in Table 5. So, while the magnitude of the rent offered declines over time the positive female-male gap persists, which in turn explains the differential earnings for the gender alluded to immediately above.

We find that while the *ex ante* bonus announcement in relational contracts are “cheap talk” they do lead to a reduction in the degree of shirking on the part of the employees. There is an increase in shirking over time, especially with fixed matching and contrary to our hypothesis women will shirk less, we find that conditional on a given bonus women actually tended to shirk more in the fixed matching treatment but these differences do dissipate over time.

Table 9
Random Effects Probit regression on shirking under Transactional Contract.

	Fixed matching and Random re-matching
Rent	-0.094 (0.126)
Extrinsic	-1.460** (0.597)
Female	2.057** (0.943)
Round	0.190 (0.150)
Female*Round	-0.291** (0.133)
Female*Extrinsic	1.190 (0.733)
Fine	0.705 (0.518)
Risk Averse	-0.255 (0.403)
Lag Cumulative Earnings	-0.017 (0.036)
Constant	-1.224 (1.368)
Demographic Control	Yes
Wald	23.13
	0.0266
Observations	72

Notes: Standard errors in parentheses; ***, ** and* denote significance at 1%, 5% and 10%.

5. Concluding remarks

We designed this study to test the proposition that in terms of the work-place, female leaders tend to be more transformational in the sense of relying more on mutual trust and reciprocity based relationship compared to male leaders, who rely more on explicit carrots and sticks. Our views in this regard were formed on the basis of earlier work looking at gender differences in the context of labour markets. We conjectured that such differential preferences on the part of men and women may have interesting and interpretable implications for their relative success in the work-place. By and large, we fail to find dramatic differences in behaviour. Contrary to our *a priori* hypothesis, women showed a slight preference for transactional contracts if anything. When it comes to worker behavior, women actually tend to shirk a little more. But these differences are not major and dissipate over time.

We do find some evidence that regardless of contract type, female employers offer higher rent to the workers. This is in line with prior findings regarding greater female generosity. Offering higher rents leads to lower earnings for those female employers, though, the differences are not stark. By and large, we conclude that when it comes to work-place strategies, there are no dramatic differences in the strategies employed by men and women.

This, in turn, suggests that the differential outcomes for men and women observed in the work-place both in terms of the gender wage gap and the leadership gap possibly do not arise due to dramatic differences in psychological preferences or work-place strategies. Therefore, as pointed out by Bohnet (2016) and Eckel et al. (2021), such a result suggests a requirement for more pro-active work-place practices and institutional changes that work toward ensuring greater gender parity.

Acknowledgements

We are grateful for the financial assistance provided by the University of Auckland Faculty Research Development Fund. We thank Professor Charles Holt for providing us access to the Veconlab website at the University of Virginia for carrying out our experiments. We are indebted

to Bram Cadsby, Paul Frijters, John Hillas, Nikos Nikiforakis, Andreas Leibbrandt, Alan Rogers, Tony So, Fei Song, two anonymous referees, the Editor (Philip Grossman) and participants at various seminars and conferences for helpful comments and feedback. We thank Yijing Wang for research assistance. All errors are the responsibility of the authors.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.socec.2020.101650.

References

- Altonji, J. G., & Blank, R. M. (1999). Race and Gender in the Labor Market. In O. C. Ashenfelter & D. Card (Eds.), *Handbook of Labor Economics* (Vol. 3C, pp. 3143-3259). New York: Elsevier.
- Babcock, L., & Laschever, S. (2003). *Women Don't Ask: Negotiation and the Gender Divide*. Princeton, NJ: Princeton University Press.
- Bertrand, M. (2011). New Perspectives on Gender. In O. C. Ashenfelter & D. Card (Eds.), *Handbook of Labor Economics* (Vol. 4B, pp. 1545-1592). New York: Elsevier.
- Bohnet, Iris (2016). *What works: gender equality by design*. Cambridge: Harvard University Press.
- Burns, J. M. (1978). *Leadership*. New York: Harper and Row.
- Camerer, C. F. (2003). *Behavioral Game Theory: Experiments in Strategic Interaction*. Princeton, NJ: Princeton University Press.
- Charness, G., & Kuhn, P. (2011). Lab Labor: What Can Labor Economists Learn from the Lab? In: O. C. Ashenfelter & D. Card (Eds.), *Handbook of Labor Economics* (Vol. 4A, pp. 229-330). New York: Elsevier.
- Chaudhuri, A., Cruickshank, A., & Sbai, E. (2015). Gender differences in personnel management: experimental evidence. *Journal of Behavioral and Experimental Economics*, 58, 20-32. October 2015.
- Chaudhuri, A., & Sbai, E. (2011). Gender Differences in Trust and Reciprocity in Repeated Gift Exchange Games. *New Zealand Economic Papers*, 45(1-2), 81-95.
- Crosen, R., & Gneezy, U. (2009). Gender Differences in Preferences. *Journal of Economic Literature*, 47(2), 448-474.
- Eagly, A. H., & Johnson, B. T. (1990). Gender and Leadership Style: A Meta-Analysis. *Psychological Bulletin*, 108(2), 233-256.
- Eagly, A. H., Karau, S. J., & Makhijani, M. G. (1995). Gender and the Effectiveness of Leaders: A Meta-Analysis. *Psychological Bulletin*, 117(1), 125-145.
- Eckel, C. C., & Grossman, P. J. (2008). Differences in the Economic Decisions of Men and Women: Experimental Evidence. In C. R. Plott & V. L. Smith (Eds.), *Handbook of Experimental Economics Results* (Vol. 1, pp. 509-519). New York: Elsevier.
- Eckel, C. C., Gangadharan, L., Grossman, P. J., & Xue, N. (2021). The Gender Leadership Gap: Insights from Experiments. In A. Chaudhuri (Ed.), *Research Agenda in Experimental Economics*. Cheltenham, UK: Edward Elgar (Chapter 7)Forthcoming.
- Fehr, E., Gächter, S., & Kirchsteiger, G. (1997). Reciprocity as a Contract Enforcement Device: Experimental Evidence. *Econometrica*, 65(4), 833-860.
- 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), 324-351.
- Fehr, E., Kirchsteiger, G., & Riedl, A. (1993). Does Fairness Prevent Market Clearing? An Experimental Investigation. *Quarterly Journal of Economics*, 108(2), 437-459.
- Fehr, E., Kirchsteiger, G., & Riedl, A. (1996). Involuntary Unemployment and Non-Compensating Wage Differentials in an Experimental Labour Market. *Economic Journal*, 106, 106-121.
- Fehr, E., Kirchsteiger, G., & Riedl, A. (1998). Gift Exchange and Reciprocity in Competitive Experimental Markets. *European Economic Review*, 42(1), 1-34.
- Fehr, E., Klein, A., & Schmidt, K. M. (2007). Fairness and Contract Design. *Econometrica*, 75(1), 121-154.
- Holt, C. (2009). University of Virginia: Veconlab. Online experiments available at <http://veconlab.econ.virginia.edu/>.
- Holt, C., & Laury, S. (2002). Risk Aversion and Incentive Effects. *American Economic Review*, 92(5), 1644-1655.
- Li, Y., Sbai, E., & Chaudhuri, A. (2020). (Un)willing to lead? Men, Women and the Leadership Gap. *Working Paper*.
- Moran, B. B. (1992). Gender Differences in Leadership. *Library Trends*, 40(3), 475-491.
- Niederle, Muriel, & Vesterlund, Lise (2007). "Do Women Shy Away From Competition? Do Men Compete Too Much?". *The Quarterly Journal of Economics*, 122(3), 1067-1101.
- Rosener, J. B. (1990). Ways Women Lead. *Harvard Business Review*, 68(6), 119-125.
- Sandberg, S. (2013). *Lean In: Women, Work and the Will to Lead*. New York: Knopf Doubleday.
- Sax, L. (2005). *Why Gender Matters: What Parent and Teachers Need to Know about the Emerging Science of Sex Differences*. New York: Doubleday.
- Walters, A. E., Stuhlmacher, A. F., & Meyer, L. L. (1998). Gender and Negotiator Competitiveness: A Meta-analysis. *Organizational Behavior and Human Decision Processes*, 76(1), 1-29.