Incentives and Compensation Schemes: An Experimental Study*

Marina Agranov† and Chloe Tergiman‡

June 2012

Abstract

When the performances of agents are correlated (because of a common random component) contracts that use information on both the relative and absolute performance of agents theoretically outperform rank-order tournaments and piece-rate schemes. Although the theoretical advantage of such contracts has long been noticed in the literature, the empirical papers that study this question have produced mixed results. Using a controlled laboratory experiment, we shed light on the performance of these contracts. We focus on the simplest version of such a contract: the Relative Piece-Rate (RPR) and compare its performance with that of tournaments and piece-rate schemes. We find that when the RPRs are imposed on the agents, they exert higher effort levels and give the principal (weakly) higher profits than tournament and piece-rate contracts. Second, we find that agents have little aversion to self-select into the RPR scheme when other alternatives are available. In our paper we fix the environment agents face and vary only their wage scheme. Our results support the theoretical predictions and suggest that principals can benefit from using RPR schemes in places where piece-rate and tournament contracts would otherwise be used.

Keywords: incentives, compensation schemes, laboratory experiments, moral-hazard.

*We are particularly indebted to Guillaume Frechette and also Andrew Schotter for the time and guidance they generously gave us at every stage of this project. We thank Heski Bar-Isaac, Alessandro Lizzeri, Viktor Raskin, Mikael Rechtsman, Ariel Rubinstein, John Duffy, Charlie Holt, Tim Salmon and Lise Vesterlund for constructive criticism. The Center for Experimental Social Science at New York University provided us with the facilities and funds to conduct our experiments. We would also like to thank the participants and organizers of the following conferences: the French Economic Association on Behavioral Economics and Experiments in Lyon, the University of Amsterdam CREED, the World ESA meeting in Rome, and the North American ESA meeting in Arizona. Finally, we are very grateful to two anonymous referees for their many comments and useful suggestions.

†California Institute of Technology and Center for Experimental Social Sciences.

‡University of British Columbia and Center for Experimental Social Sciences.
1 Introduction

When the performances of risk-averse agents are correlated (because of a common random component), using relative performance information can improve the trade-off between incentives and risk-sharing. For example, when the common shock is a big part of the uncertainty, the optimal rank-order tournament (T), which pays a higher prize to the agent that produced the higher output, guarantees the principal a higher expected profit than the optimal independent piece-rate contract (IPR), in which agents are paid linearly based only on their absolute performance. However, even when tournaments theoretically outperform piece-rate contracts, they are “informationally quite wasteful if performance levels of agents can be measured cardinally rather than ordinally” (Holmstrom (1982)). Indeed, if the exact outputs of agents are observable, there is a better way of using this information in contract design.

The theoretical advantage of contracts that make use of both relative and absolute performance information of agents has long been known in the literature. However, the empirical studies that investigate the efficacy of these contracts in the labor markets have produced mixed results. Much of this literature uses data from CEO pay and the results are far from being clear (summarized in section 1.1). In this paper we aim to complement the empirical literature by conducting a controlled laboratory experiment, which evaluates the performance of such a contract. For simplicity and in line with the theoretical literature\(^1\), we concentrate on a contract that is linear in both agents’ performance. We call it the Relative Piece-Rate (RPR) scheme. In the RPR, each agent receives payment in proportion to his output as in the IPR scheme, but he is also penalized in a linear way for every unit of output produced by the other agent. We compare the RPR scheme with Tournaments and IPRs because they are often used in practice\(^2\) and, hence, form a natural baseline for comparison.

There are several reasons why a laboratory experiment is a useful approach in the


\(^2\)Standard examples of rank-order tournaments include promotion tournaments for managers, bonuses for salespeople at the end of the year, a tournament for a limited number of tenure positions that assistant professors compete for, etc. The IPR is often used in industries such as fruit picking or when people are paid on commission. The percentage of workers who receive their compensation in the form of bonuses or piece-rates is substantial. Lemieux, MacLeod and Parent (2009) report that the overall fraction is about 40%, and within sales and services the fraction of employees receiving their pay this way is about 75% and 20%, respectively.
current situation. We are interested in testing the performance of the RPR scheme and comparing it with the performance of the Tournament and IPR schemes. The key advantage of using the lab to answer this question is the ability to control many aspects of the environment, which is necessary for the comparison to be fair and adequate. These aspects are often impossible to observe or infer from labor market data. Moreover, while empirical data is rich, it often reflects a variety of environmental factors that affect the performance of the agents and are extremely hard to separate out. The lab environment is clean and gives the experimenter the freedom to fix all the parameters of the environment and vary only the scheme imposed on the agents. This allows for a rigorous comparison across wage scheme designs. Over the past 20 years, laboratory experiments have proven useful in understanding many important questions posed by labor economics (see survey by Charness and Kuhn (2010)). However, as far as we know, we are the first ones to try to resolve the conflicting empirical evidence regarding the RPR-like schemes using a laboratory experiment.

One important and unique feature of our experimental design is that we fix the environment that the principal faces, and test the performance of the RPR, T and IPR contracts with the optimal parameters that are determined by this environment. The environment we chose is standard in the literature: it represents one example of a broad class of environments, in which according to the theory RPR performs better than T and IPR contracts. Our choice to fix the environment is in sharp contrast with past experimental papers. Indeed, up to now, the literature has primarily focused on whether agents targeted the effort prediction of a wage scheme. Put differently, other studies chose the parameters of the wage scheme so that the predicted effort levels were the same across wage schemes. This means that the prior literature aimed to compare how well schemes could elicit a fixed (identical) effort level. This was without regard to whether the coefficients in these wage schemes (1) corresponded to the same environment, or (2) were optimal given an environment. In fact, if the experimenter holds the effort level prediction constant across wage schemes, then the parameters in those wage schemes cannot both be optimal and correspond to the same environment. To our knowledge,
we are the first to keep the environment constant across wage schemes. We can therefore compare wage schemes from the viewpoint of a profit-maximizing principal facing a particular environment. In other words, while the results from these past experiments provide insight into subjects’ behavior, they cannot address the question, “How should a principal, facing a specific environment, choose to pay his employees?” which is the main focus of this paper.

We designed two experiments. The purpose of Experiment 1 was to study how well the RPR scheme performs relative to the IPR and T, both in terms of the agents’ efforts and of the profits earned by the principal when a wage scheme is imposed on the subjects. The results from Experiment 1 support the main theoretical predictions. The profits of the principal under the optimal RPR scheme are (weakly) higher than in the T and IPR.\(^5\) Also, workers exert strictly higher efforts under the optimal RPR scheme when compared to the optimal T and IPR schemes. This is true when these contracts are exogenously imposed on agents. However, it is possible that people have an intrinsic aversion to some types of contracts, and if presented with the choice of several contracts, would not self-select into the RPR.

In Experiment 2 we allow (new) subjects to self-select into either compensation scheme. The purpose of Experiment 2 was to evaluate whether or not subjects exhibited certain biases against or in favor of the RPR (or any other scheme), and whether these biases were enforced or mitigated by experience. With our design, subjects should be indifferent between all three wage schemes in terms of their expected utility. Therefore, if they like or dislike one scheme more than the other it would be for reasons that are outside of utility maximization. The results of Experiment 2 indicate that subjects do not seem averse to the RPR. On the contrary, not only do subjects choose to participate in the RPR over the T and IPR before they have any experience, but they remain with the RPR and even switch to it after they have gained experience in all three wage schemes.\(^6\)

All together, our findings from Experiment 1 and Experiment 2 confirm that the RPR

---

\(^5\)The RPR always generates higher profits for the principal than the T contract. With respect to the IPR contract, in one environment RPR gives strictly higher profits and in the other, the RPR and IPR are “tied.”

\(^6\)In Experiment 2, we described all three wage schemes to the subjects. They then chose one wage scheme to participate in. After they participated in that scheme, subjects were asked to participate in all three wage schemes. After they gained experience in all three wage schemes, they were again offered the choice of which payment scheme to participate in.
scheme outperforms T and IPR schemes and suggest that a principal could significantly increase his profits with a frequent use of RPR in place of IPRs or Ts.

The rest of the paper is organized as follows. The literature review is in Section 2. In Section 3, we describe the environment and state the theoretical results. The design and results from Experiment 1 are in Section 4. Experiment 2 is described in Section 5. Finally, Section 6 concludes.

2 Related literature

The early theoretical work on the performance of piece-rate contracts versus rank-order tournaments includes Lazear and Rosen (1981), Green and Stokey (1983) and Nalebuff and Stiglitz (1983). Later on, Holmstrom (1982) and Mookherjee (1984) studied the conditions under which relative performance contracts (not restricted to tournaments) perform better than independent contracts. Since then, there has been a lot of theoretical work in agency theory that has investigated the relative performance of contracts under different circumstances (see, for example, Meyer and Vickers (1987), Holmstrom and Milgrom (1991) and Levitt (1995)). More directly related to the present paper is Tsoulouhas (2010) that compares tournaments, piece-rates and “hybrid” tournaments that resemble our RPR scheme in a theoretical framework involving n players.\(^7\)

While compelling theoretical arguments can be made in favor of relative performance evaluation schemes, empirical studies that assess their effectiveness have produced mixed results: some studies document that CEO’s pay is negatively related to industry performance (see Gibbons-Murphy (1990) and Janakiraman et al. (1992)); other studies show that relative performance is not an important source of managers’ contracts (see Jensen-Murphy (1990), Barro-Barro (1990) and Aggrawal-Sanwick (1999)); while others find that CEO’s compensation increases with industry performance (see Joh (1999)) or that CEO’s dismissals are more likely when the industry is doing poorly (see Jenter-Kanaan (2006)), which is at odds with the idea of the relative performance evaluation.\(^8\)

\(^7\)In Tsoulouhas (2010), in a tournament agents receive a bonus dependent on how much better they do than the average in addition to a flat payment. The weights on absolute and average group performance are equal in a standard tournament, whereas they are unequal under a “hybrid” tournament.

\(^8\)Knoeber (1989), Knoeber and Thurman (1994 and 1995) have studied these types of schemes in the context of the broiler chicken industry and find support for these types of schemes.
The experimental literature that tests the behavioral performance of various incentive schemes consists of both laboratory and field studies. Among the laboratory studies are Bull, Schotter and Weigelt (1987), Harbring and Irlenbusch (2003), van Dijk, Sonnemans and van Winden (2001) and Wu, Roe and Sporleder (2006). A consistent result in these experiments is that IPR schemes do well at eliciting the theoretically predicted level of effort, whereas tournament contracts have varied results. While our focus is mainly on the performance of the RPR scheme, consistent with the existing literature, our findings indicate that the tournament contract does not always perform as theory predicts. More precisely, our results indicate that the tournament scheme performs the worst (compared to IPR and RPR) when the common uncertainty is relatively large, which is exactly when it has theoretical advantages.

In terms of field experiments, the closest work to ours is Bandiera, Barankay and Rasul (2005) who conducted a field experiment in a fruit picking plant and compared three wage schemes: an individual piece-rate, a relative compensation scheme where individual pay and group performance are negatively correlated, and a relative compensation scheme where individual pay and group performance are positively correlated (group incentives). The authors observe that effort is higher under the individual pay schemes. However, the parameters of the wage schemes used in this field experiment were not necessarily optimal, and it is unclear what the results would have been if this had been the case. While an experiment conducted in the field has its obvious and important advantages, the laboratory experiment has a useful contribution to our understanding of the effect of incentives on workers’ behavior by controlling all parameters of the environment.

Bull, Schotter and Weigelt (1987), for example, report that subjects respond on target to the different IPR schemes, but for some tournaments their effort choices are off with subjects choosing extremely low effort levels (a behavior associated with “dropping out”), which they do not do in the IPR schemes. Harbring and Irlenbusch (2003) focus on the effect of tournament design on effort levels. They do not make point predictions but study how well subjects understand in which direction their effort should go to maximize their payoff when faced with a different tournament structure. Wu, Roe and Sporleder (2006) find that for a fixed total uncertainty, increasing the relative importance of the common shock results in subjects choosing below equilibrium effort levels.
3 Theory

In this section we describe the environment and derive the optimal parameters of our three wage schemes: the relative piece-rate (RPR), the independent piece-rate (IPR) and the tournament (T). We then show that the relative performance of these schemes depends on the characteristics of the environment; however, the RPR always dominates both the IPR and T schemes in terms of the principal’s profits and effort levels of agents.

Consider a standard moral-hazard setting à la Holmstrom (1979). A risk-neutral principal hires 2 agents (i and j) to produce output. The principal’s profit consists of the total output produced by workers minus the cost of hiring them. The output of an agent i has an additive structure: \( y_i = e_i + \epsilon_i + \eta \), where \( e_i \) is the effort level privately chosen by agent i, \( \epsilon_i \) is her private production shock and \( \eta \) is a common production shock that affects both agents equally. The effort levels of agents are not observable and therefore not contractible. The principal observes only the output produced by each agent.

Agents are risk-averse. The utility of agent i is \( u_i(x) = -e^{-rx} \) where r represents the coefficient of absolute risk aversion. Effort is costly and the cost function associated with effort \( e_i \) is given by \( \frac{(e_i)^2}{c} \). The private shocks \( \epsilon_i \) and \( \epsilon_j \) are independently drawn from \( U[-A,A] \) and the common shock \( \eta \) is drawn from \( U[-B,B] \).

There are three possible wage schemes that the principal can employ:

- **RPR scheme:** Agent i gets wage \( i = \alpha + \beta \cdot y_i + \gamma \cdot y_j \)
- **IPR scheme:** Agent i gets wage \( i = \alpha' + \beta' \cdot y_i \)
- **Tournament scheme:** The agent with the highest output wins a big prize \( W_1 \), the other agent gets a small prize \( W_2 \).

The principal chooses the parameters for each of the schemes to maximize his expected profits: \( E \left[ \sum_{i,j} (y_i - \text{wage}_i) \right] \) subject to standard incentive compatibility and individual rationality constraints. He takes as given the parameters of the environment, which are the coefficient of absolute risk aversion of agents r, the distributions of the two shocks \( \epsilon_i \) and \( \eta \), the structure of the cost function of effort and the outside option of the agents which we denote by \( \bar{v} \). The agents choose their effort levels simultaneously after learning
which wage scheme is used and what the exact parameters of that scheme are. Shocks are then realized and payments are distributed. The payment of agent $i$ is her wage minus the cost of effort: $\text{wage}_i = \text{wage}_i - \frac{e_i^2}{c}$.

The framework described above can be extended to more than 2 agents. The extension of the IPR and the T scheme is rather straightforward, while if the RPR scheme is employed then agent $i$ gets

$$\text{wage}_i = \hat{\alpha} + \hat{\beta} \cdot y_i + \hat{\gamma} \cdot \bar{y}_{-i}$$

where $y_i$ is the output of agent $i$ and $\bar{y}_{-i}$ is the average output of all other agents not including agent $i$. That is, each worker receives payment in proportion $\hat{\beta}$ to his own output and in proportion $\hat{\gamma}$ to the average output of the others.

The RPR Scheme

The principal chooses the parameters $(\alpha, \beta, \gamma)$ to maximize his expected profits given an incentive compatibility (IC) and an individual rationality (IR) constraint. Agent $i$ chooses her optimal effort level $e_i$ given the parameters of the RPR scheme. We use the approximation of expected CARA utility in terms of certainty equivalent to solve for the optimal effort level of an agent. The certainty equivalent of agent $i$ when $(e_i, e_j)$ are the effort levels chosen by agents is

$$CE_i(e_i, e_j) = \alpha + \beta \cdot e_i + \gamma \cdot e_j - \frac{r}{2} \left[ (\beta^2 + \gamma^2) (\sigma^2 + \sigma^2) + 2 \beta \gamma \sigma^2 \right] - \frac{(e_i)^2}{c}$$

$$\Rightarrow e_i^* = \frac{\beta c}{2}$$

---

10See Tsoulouhas (2010).

11The optimal linear RPR schemes are symmetric with respect to the agents since the principal cares about the sum of agents’ outputs. This is contrary to Levitt (1995), who studies optimal linear RPR schemes when the principal cares only about the highest output produced by his agents. Levitt shows that asymmetric RPR schemes sometimes do better than symmetric ones even if agents are ex-ante identical.

12For an individual with CARA utility and normal risks, the certainty equivalent formula holds exactly. Otherwise, the relationship is approximate and we have $E[u(\tilde{x})] \approx u(CE(\tilde{x})) = u(E\tilde{x} - \frac{r}{2} \cdot \text{var}(\tilde{x}))$, where $\tilde{x}$ is the risky alternative, $CE(\tilde{x})$ is the certainty equivalent of $\tilde{x}$, $E\tilde{x}$ is the expectation of $\tilde{x}$ and $\text{var}(\tilde{x})$ is the variance.
The principal's maximization problem is:

$$\max_{(\alpha, \beta, \gamma)} [(1 - \beta - \gamma)\mathbb{E}(y_i + y_j) - 2\alpha]$$

subject to:

$$e_i^* = \frac{\beta c}{2} \quad \text{and} \quad e_j^* = \frac{\beta c}{2}$$

$$CE_i(e_i^*, e_j^*) = \bar{v} \quad \text{and} \quad CE_j(e_i^*, e_j^*) = \bar{v}$$

With $\rho = \frac{\sigma_2^2}{\sigma_f^2 + \sigma_2^2}$, we get the optimal parameters of the RPR scheme:

$$\alpha = \bar{v} + \frac{r}{2} \left( \beta^2 + \gamma^2 \right) \left( \sigma_\eta^2 + \sigma_f^2 \right) + \beta \gamma r \sigma_\eta^2 - \frac{\beta^2 c}{4} - \frac{\beta \gamma c}{2}$$ (1)

$$\beta = \frac{c}{c + 2r \left( \sigma_\eta^2 + \sigma_f^2 \right) (1 - \rho^2)}$$ (2)

$$\gamma = -\rho \beta$$ (3)

There are two features of the optimal RPR scheme that are worth noting. First, the RPR scheme partially filters out the common shock through the negative $\gamma$ coefficient: the high output produced by agent $i$ is compensated less if agent $j$ also produced a high rather than low output. Second, the behavior of agent $i$ impacts the wage received by agent $j$ and vice-versa: each unit of output produced by agent $i$ ($y_i$) lowers the wage of agent $j$ by $\rho \beta$.

In a more general case, when a principal hires more than two agents, the following two features of the optimal RPR scheme remain true: $\beta > |\gamma|$ and $\gamma < 0$. Thus, an agent who produces a high output is compensated more when the average output is low rather than high. This means that doing well when others are doing badly is better than doing well when everybody is also doing well.

To sum up, given the optimal parameters of the RPR scheme, agents choose effort levels $e_i^* = e_j^* = \frac{\beta c}{2}$ and expect a utility level of $u(\bar{v})$. The principal’s expected payoff is

$$\frac{c^2}{2c + 4r \frac{\sigma_f^2}{\sigma_f^2 + \sigma_\eta^2} \left( \sigma_f^2 + 2\sigma_\eta^2 \right)} - 2\bar{v}.$$  

**The IPR Scheme**

Applying the same technique as in the RPR case, we arrive at the following optimal
parameters for the IPR scheme:

$$\beta' = \frac{c}{c + 2r(\sigma_i^2 + \sigma_n^2)}$$ and $$\alpha' = \bar{v} - \frac{(\beta')^2 c}{4} + \frac{r}{2} (\beta')^2 (\sigma_n^2 + \sigma_i^2)$$

Given these parameters, agents choose effort levels of $$e_i^* = e_j^* = \frac{\beta' c}{2}$$ and expect a utility level of $$u(\bar{v})$$. The principal’s expected payoff is $$\frac{c^2}{2c+4r(\sigma_i^2 + \sigma_n^2)} - 2\bar{v}$$.

The Tournament Scheme

The optimal prize structure is\(^\text{13}\):

$$W_1 = \frac{c^2 + 4Ac}{4(c + 2A^2r)} + \bar{v}$$ and $$W_2 = \frac{c^2 - 4Ac}{4(c + 2A^2r)} + \bar{v}$$

Given this prize structure, agents choose effort levels of $$e_i^* = e_j^* = \frac{c^2}{2(c+2A^2r)}$$ and expect a utility level of $$u(\bar{v})$$. The principal’s expected payoff is $$\frac{c^2}{2(c+2A^2r)} - 2\bar{v}$$.

Theoretical comparison of three wage schemes

The main feature of the RPR scheme is that the payment of agents is a function of both their absolute and relative performance. In fact, each agent receives payment in proportion to his/her output as in the IPR scheme, but is also penalized in a linear way for every unit of output produced by the other agent ($$\gamma < 0$$). In an environment in which agents are risk-averse and are both subject to the common shock, the relative component of the RPR payment allows the principal to (partially) filter this common uncertainty and to balance the incentives and risk-sharing.

The combination of absolute and relative performance information in agents’ payments is what distinguishes an RPR scheme from (1) an IPR scheme, in which agents are incentivized based solely on their absolute performance and from (2) a Tournament scheme, which makes use only of the relative performance of agents. When we compare the performance of these three wage schemes for a given environment, we obtain the following results:

- The expected payoff of the principal and the effort level of agents are always higher

\(^{13}\)We concentrate on the symmetric equilibrium in which both agents exert equal effort $$e_1^* = e_2^*$$. For the detailed solution we refer the reader to Bull, Schotter and Weigelt (1987).
under the RPR scheme when compared to the IPR and the Tournament schemes.

- When the common uncertainty is high relative to the idiosyncratic one ($\sigma_\eta^2 > 2 \cdot \sigma_\epsilon^2$), the Tournament gives the principal higher expected profits than the IPR scheme does. The effort of workers follows the same ordering. When the common uncertainty is small relative to the idiosyncratic one ($\sigma_\eta^2 < 2 \cdot \sigma_\epsilon^2$) the IPR gives the principal higher expected profits than the Tournament scheme does. The effort of workers follows the same ordering.

4 Experiment I

In this paper we design two experiments. In Experiment I, we test the performance of the RPR contract and compare it to the IPR and Tournament contracts both in terms of the principal’s profits and the effort levels exerted by the agents when the wage contracts are imposed on the agents. The purpose of Experiment I is to identify possible behavioral issues that would impact the performance of the RPR. Other-regarding preferences (such as Fehr-Schmidt (1999), Bolton-Ockenfels (2000) or Charness-Rabin (2002)) might be one possible reason for this concern; after all, each unit of output of agent $i$ directly decreases the compensation of agent $j$. If that’s the case, then a principal might refrain from using RPR schemes because he will end up with lower expected payoffs than if he uses other contracts like the simple piece-rate.

We test the performance of the RPR scheme in two environments: the high variance private shock (HV) (in which $\sigma_\eta^2 < 2 \cdot \sigma_\epsilon^2$) and the low variance private shock (LV) (in which $\sigma_\eta^2 > 2 \cdot \sigma_\epsilon^2$). In both environments we expect that under the RPR scheme the principal’s payoffs and agents’ efforts will be highest. In the HV environment we expect the IPR to dominate the T. We expect the opposite in the LV environment. These expectations from Experiment I are summarized in the hypotheses below:

4.1 Hypotheses

**Hypothesis 1:** The ordering of the effort levels the agents choose is RPR > IPR > T in the HV environment. In the LV environment this ordering is RPR > T > IPR.
Moreover, for any given contract, agents choose higher effort levels in the LV than in the HV environment.

Hypothesis 2: In both the HV and the LV environments principal’s payoff is the highest when the RPR contract is used than when the T or the IPR contracts are used.

4.2 Procedures

In this section we describe the details of Experiment 1. All sessions were run at the experimental lab of the Center for Experimental Social Science (CESS) at New York University. Students were recruited from the general undergraduate population via e-mail solicitations. All sessions were conducted through computer terminals, and the software was written in ZTree (see Fischbacher (2007)).

Each recruited subject participated in only one session: 184 subjects participated in Experiment 1. In each experimental session, the subjects experienced only one of the three wage schemes (IPR, T or RPR). Each session was organized according to a within subject block design: each subject went through 20 rounds of the HV environment, followed by 20 rounds of the LV environment, and then again 20 rounds of the HV environment and 20 rounds of LV (HLHL).14 For each wage scheme we also conducted an LHLH session to check for order effect and found none.15 The within-subject design allows us to study whether the subjects individually respond to an exogenous change in the environment they face and adjust their behavior accordingly.

In the HV environment, the private shock was drawn from a uniform distribution on the interval \([-24, 24]\). In the LV environment, the private shock was drawn from \(U[-14, 14]\). The variance of the common shock was kept constant in both environments and was drawn from a uniform distribution on the interval \([-27, 27]\). The cost parameter

\[14\] Recall that HV means that the variance of the private shock was high relative to the common shock, whereas LV variance treatment means that variance of the private shock was low relative to the common shock.

\[15\] A Kolmogorov-Smirnov test for equalities of distributions on the average effort choice for each subject shows no pair-wise difference between sessions for a given game and stage (all p-values are greater than .1). The summary statistics are in Table 2. The LHLH session was Session III for the IPR and RPR and Session IV for the T. These summary statistics show that in terms of means, these sessions are similar to the HLHL sessions, and that the direction of change between treatments is in line with that found for the HLHL sessions. In the remainder of the paper, the data are pooled across sessions, and regressions are clustered at the session-level.
was fixed at $c = 110$, the outside option was held constant for all environments and contracts at level $\bar{v} = 7$, and the risk aversion parameter $r$ was set to $0.1$.\textsuperscript{16} Table 1 displays the optimal contracts as well as the theoretical predictions for the effort levels of the agents and the expected profits of the principals for these parameters.

<table>
<thead>
<tr>
<th>HV environment</th>
<th>Contract</th>
<th>Parameters</th>
<th>Effort Level</th>
<th>Principal’s Profit</th>
</tr>
</thead>
<tbody>
<tr>
<td>RPR</td>
<td>$wage_i = 13.63 + 0.65 \cdot y_i - 0.36 \cdot y_j$</td>
<td>35.75</td>
<td>23.62</td>
<td></td>
</tr>
<tr>
<td>IPR</td>
<td>$wage_i = 4.21 + 0.56y_i$</td>
<td>30.80</td>
<td>18.68</td>
<td></td>
</tr>
<tr>
<td>T</td>
<td>$W_1 = 31.16$ and $W_2 = 7.71$</td>
<td>26.87</td>
<td>14.87</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>LV environment</th>
<th>Contract</th>
<th>Parameters</th>
<th>Effort Level</th>
<th>Principal’s Profit</th>
</tr>
</thead>
<tbody>
<tr>
<td>RPR</td>
<td>$wage_i = 20.75 + 0.82 \cdot y_i - 0.65 \cdot y_j$</td>
<td>45.10</td>
<td>33.36</td>
<td></td>
</tr>
<tr>
<td>IPR</td>
<td>$wage_i = 1.04 + 0.64 \cdot y_i$</td>
<td>35.20</td>
<td>23.26</td>
<td></td>
</tr>
<tr>
<td>T</td>
<td>$W_1 = 36.66$ and $W_2 = 15.95$</td>
<td>40.55</td>
<td>28.55</td>
<td></td>
</tr>
</tbody>
</table>

Table 1: Parameters of the RPR, IPR and T contracts

In each round of Experiment 1 subjects went through the following stages: at the beginning of a round, subjects were reminded of the payment structure (IPR, T or RPR) as well as the environment that they were facing (high or low private shock variance). They were randomly and anonymously matched into pairs. Subjects then had to choose an effort level and could choose any number between 0 and 55, with at most a 2 decimal point precision.\textsuperscript{17,18} After everyone made their choice, the subjects were shown their output, as well as the output of the person they were paired with, and finally the number of tokens they earned for that round. In our experiment, subjects were not given any information on their “partner’s” choice of effort, only on the output. This parallels the information available to a principal in a moral hazard context. The subjects then

\textsuperscript{16}When computing the optimal parameters in each wage scheme, the principal takes $\bar{v}$ as given and ensures that in expectation the agents receive $\bar{v}$, as detailed in Section 3. In the laboratory experiment, subjects did not have an explicit outside option.

\textsuperscript{17}We restricted the effort level to be between 0 and 55 because implementing risk-aversion requires that the support of the wage be bounded. The upper limit of 55 was chosen so the theoretically optimal effort levels would not be obvious focal points.

\textsuperscript{18}Prior experimental findings suggest that the choice of bounds by the experimenter may impact the choices made by subjects in the laboratory. For example in the BDM mechanism, Bohn, Linden and Sonnegard (1997) show that in some circumstances, the higher the upper bound, the higher the average reservation price. In our experiment, the range of possible choices for effort remains constant over all schemes. Unless there is reason to believe that this would have an asymmetric effect across the different wage schemes, this rules out that these bounds explain the difference in our subjects’ choices across payment schemes.
moved to the next round, and were re-matched with another participant. At the end of the session, one round for each block was randomly chosen (four rounds in total), and the tokens the students earned for that round were then converted into a probability of winning a big prize of $4. So subjects could win up to $16 if they won in all four blocks, and could win $0 if they did not win any of the big prizes. In addition, subjects received a participation fee of $5. Each session lasted about 90 minutes, and at the end we asked subjects to fill out a questionnaire in which we asked them questions about themselves, their understanding of the game and how they made their choices during the experiment. On average, subjects earned slightly over $20, including the participation fee. The instructions for the RPR treatment are in the online Appendix.

The design of Experiment 1 has four important properties:

- We fix the environment (defined by the distributions of private and common shocks, the risk aversion parameter of the agents, the costs of effort and the outside option) and test the three wage schemes with the optimal parameters for that environment. This allows the comparison of the relative efficiency of wage schemes both in terms of the principal’s payoffs and in terms of how well agents target the theoretically predicted effort levels.

- We use a within subject design to study each wage scheme separately (and therefore can look at individual responses to changes in the environment), and a between subject design to compare across wage schemes.

- We induce risk aversion. We induced CARA utility with risk aversion coefficient of 0.1 using the method of Berg, Daley, Dickhaut, and O’Brien (1986) and Roth and Malouf (1979). The general procedure is the following: subjects play for

---

19Subjects were matched anonymously and randomly after each choice of effort level. So, while two subjects may have encountered each other more than once over the course of the experiment, because the matching was anonymous, the possibility of colluding was minimized.

20Since subjects were not aware of what was to happen in the following parts of the experiment, this procedure is identical to the original procedure of inducing risk-aversion with two prizes. Indeed, in each block, subjects could win either a big prize of $4 or a small prize of $0.

21Aside from the questions pertaining to the game itself, the questionnaire also asked for a subject’s major and year but did not ask for a subject’s gender.

22Thus, subjects earned slightly over 13 dollars an hour, well exceeding minimum wage earnings in New York ($7.25/hour), which is common practice for laboratory experiments.

23Estimations of CARA coefficients have generally been between 0.0085 and 0.14 - see Saha, Shamway and Talpaz (1994) (page 175) who summarize studies on risk aversion estimations.
tokens, and those tokens are converted into a probability of winning a big prize. The conversion procedure converts the tokens to reflect the preferences that the experimenter would like to induce. For example, if the experimenter wants to induce risk neutrality, he/she would make the conversion linear. For risk-aversion, the conversion would be concave, thus making earlier tokens more valuable. Since we induce a specific risk-aversion, we are able to use the optimal coefficients in each of the wage schemes.  

- We keep the common shock as part of the environment for each wage scheme. The common shock is part of the payment for the RPR and the IPR schemes; it is, thus, not an option to remove it from these treatments. We keep it for the Tournament treatment as well for two reasons. The first is behavioral: the magnitude of the common uncertainty could be important in practice. The second is to keep the information subjects receive at each round identical across all wage schemes.  

4.3 Results

We organize results from Experiment 1 around the hypotheses formulated in Section 4.1.

HYPOTHESIS 1 “Effort levels of the workers”

The effort levels chosen by agents is the main concern of the principal as it is the major determinant of output and profits. We start by discussing whether the effort levels chosen by laboratory subjects correspond to the theoretically predicted ones and obey the ordering predicted by theory.

Figure 1(a) (1(b)) shows the average effort per period across all subjects in an HV (LV) environment. The data from blocks of the same environment are appended at the

---

24 This method of inducing risk attitude is not uncontroversial. While some papers find that risk-averse subjects tend to behave as expected value maximizers (see, for example, Berg, Daley, Dickhaut and O’Brien (1986), Prasnikar (1993, 2002) and Rietz (1993)), others have found that this method does not change the behavior of subjects (see Walker, Smith and Cox (1990), Cox and Oaxaca (1995), or Selten, Sadrieh and Abbink (1999)). It is important to note that under risk neutrality, all wage schemes collapse and become identical in terms of profits for the principal. Thus, if subjects had behaved in a risk-neutral way in our experiments, this would have led to results opposite to those in the data, which we present in section 4.3. In the present paper it is clear that subjects behaved far from risk-neutrality.

25 Most experimental papers that study the tournament scheme leave the common shock out of the game as it theoretically has no impact (see Bull, Schotter and Weigelt (1987)). Wu and Roe (2005) and Wu, Roe and Sporleder (2006) are the only papers that include the common shock in tournament experiments.
In the first few rounds of the experiment the average effort per period is similar in all three wage schemes. However, as more rounds are played, the average efforts become distinct across schemes. In the experimental literature, it is a common observation that subjects take time to get used to a game and understand its mechanism. Our experiment is not an exception. For that reason we restrict our statistical analysis to the last 20 rounds of each environment.

In our analysis we will present results from both random effects regressions clustering at the session level as well as non-parametric tests using session-level averages.

In Table 2 we report the average efforts of the subjects for each scheme by session as well as the theoretical levels as predicted by the theory. In Table 3 we report the results of the Wilcoxon rank-sum test that tests whether the observed effort levels in different wage schemes come from populations with the same distribution. This comparison is of special importance since it allows us to see which of the three contracts elicits the higher effort levels from agents.

Theoretically, we expect the following rankings of the effort levels: RPR > IPR > T in the HV environment and RPR > T > IPR in the LV environment. The main prediction is borne out: in both environments, the RPR scheme elicits higher effort levels than the IPR or the T schemes. The IPR scheme performs better than T in both environments.\footnote{This result is not surprising given Figure 2 where we see that laboratory subjects drastically un-}
The behavior we observe in the Tournament scheme in the Low Variance treatment is surprising and unexpected. It is theoretically precisely in situations where the uncertainty is predominantly common that the Tournament should be advantageous. However, we observe no such thing: instead, subjects exert the least effort relative to the other schemes. At the population level, there is no statistical difference in behavior across treatments. However, this masks important individual differences in behavior that may explain what is going on. As can be seen in Figure 2, a majority of subjects in the Tournament scheme do respond to the incentives in the correct direction; that is, they increase their effort when the private uncertainty is low relative to the common one (i.e. in the Low Variance treatment). However, those that do not, decrease their choices by so dershoot effort levels in the Tournament scheme, which theoretically should outperform the IPR in LV environment.

Table 2: Session-level average effort by variance.

<table>
<thead>
<tr>
<th></th>
<th>High Variance</th>
<th></th>
<th>Low Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Data</td>
<td>Theory</td>
<td>Data</td>
</tr>
<tr>
<td>IPR</td>
<td>Session I</td>
<td>33.78</td>
<td>35.34</td>
</tr>
<tr>
<td></td>
<td>Session II</td>
<td>31.88</td>
<td>30.8</td>
</tr>
<tr>
<td></td>
<td>Session III</td>
<td>29.07</td>
<td></td>
</tr>
<tr>
<td>RPR</td>
<td>Session I</td>
<td>38.79</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Session II</td>
<td>37.59</td>
<td>35.8</td>
</tr>
<tr>
<td></td>
<td>Session III</td>
<td>36.52</td>
<td></td>
</tr>
<tr>
<td>T</td>
<td>Session I</td>
<td>23.54</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Session II</td>
<td>28.35</td>
<td>26.9</td>
</tr>
<tr>
<td></td>
<td>Session III</td>
<td>26.02</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Session IV</td>
<td>24.21</td>
<td></td>
</tr>
</tbody>
</table>
Table 3: Rank-sum Tests for Effort Levels between Wage Schemes

<table>
<thead>
<tr>
<th>Environment</th>
<th>Observed Effort Levels (p-value)</th>
</tr>
</thead>
<tbody>
<tr>
<td>HV</td>
<td>RPR &gt;**IPR (p = 0.0495)</td>
</tr>
<tr>
<td></td>
<td>RPR &gt;**T (p = 0.0339)</td>
</tr>
<tr>
<td></td>
<td>IPR &gt;**T (p = 0.0339)</td>
</tr>
<tr>
<td>LV</td>
<td>RPR &gt;**IPR (p = 0.0495)</td>
</tr>
<tr>
<td></td>
<td>RPR &gt;**T (p = 0.0339)</td>
</tr>
<tr>
<td></td>
<td>IPR &gt;**T (p = 0.0339)</td>
</tr>
</tbody>
</table>

>** means higher at 5% (2-sided)

much that the overall average stays relatively stable. There are, in fact, stark differences between the types of choices of those who did not react to the change in environment. In the Tournament scheme, those who did not switch average choice all chose the minimum possible effort (zero). This “dropping out” behavior, while representing only a small percentage in our experiment, has been observed in prior literature, in a larger proportion of subjects.\(^{27}\) In the RPR scheme, on the other hand, those who do not change direction are predominantly choosing very high levels of effort (on average 52.8 out of a maximum of 55).

We investigate whether subjects respond correctly to a change in incentives within each scheme and test for the overall statistical difference in effort levels between the HV and LV environments. We test the equality of matched pairs of observations (average effort in each session and treatment) using the Wilcoxon matched-pairs signed-rank test. The null hypothesis that both distributions are the same is rejected at the highest level of confidence possible with three sessions for both the RPR and the IPR scheme (\(p = 0.1088\)). For the T scheme, the results are not significant, but the direction of the change is correct (\(p = 1\)).\(^{28}\) Figure 2 below shows the fraction of subjects whose effort was higher in the LV than in the HV environment (as the theory predicts). At least half the subjects responded to the incentives in the correct direction in each scheme (50% in RPR, 67% in IPR and 61% in the Tournament).

\(^{27}\)For example, in Schotter and Weigelt (1992), subjects who are disadvantaged in the competition (i.e. who have a higher marginal cost of effort) are observed to drop out of T even when in equilibrium they should not.

\(^{28}\)Random Effects regressions with clustering at the session level show similar results: for both the RPR and IPR schemes, subjects exert higher levels of effort in the LV environment (significant at the 5% level) but in the T scheme there is no treatment effect.
<table>
<thead>
<tr>
<th></th>
<th>High Variance</th>
<th>Low Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>IPR</td>
<td>-5.76*** (1.01, 0.00)</td>
<td>-4.41*** (1.09, 0.00)</td>
</tr>
<tr>
<td>T</td>
<td>-12.02*** (1.47, 0.00)</td>
<td>-13.92*** (.99, 0.00)</td>
</tr>
<tr>
<td>Constant</td>
<td>37.58*** (.53, 0.00)</td>
<td>39.35*** (.62, 0.00)</td>
</tr>
</tbody>
</table>

R-sq=.1003  
Wald $\chi^2$=88.46(pvalue=0.00)  
R-sq=.1360  
Wald $\chi^2$=225.88(pvalue=0.00)

Coefficient and in parenthesis: robust standard errors and p-values  
*** - significant at 1%

Nb. obs: 3,680; Nb. groups: 184; Obs per group: 20.

Table 4: Predicting effort levels: random effect regression, clustering at the session level.

In all schemes, an increase in effort leads to a for certain increase in cost. The increase in benefits associated with an increase in cost is different depending on the scheme. The tournament is the only scheme for which an increase in effort only increases the probability of receiving a higher wage. This is in contrast to the IPR and RPR schemes where an increase in effort increases one's wage directly. This may explain our subjects’ reluctance to exert as high an effort as they should theoretically. Beyond that, it can also explain why subjects who do increase their choices do so only in less dramatic fashion (on average 4 points) compared with the IPR and the RPR (both over 7 points on average). In other words, the tournament seems more risky: subjects will definitely incur the higher costs if they increase their efforts, but they will only receive higher wages with a certain probability.

Finding 1. Consistent with Hypothesis 1, in both environments the RPR scheme elicits higher effort choices than the IPR and the Tournament contracts do. Contrary to the theory, however, the ranking of the Tournament and IPR are reversed relative to the theory in the LV treatment: the Tournament fares less well than the IPR when it should
be the opposite.

**HYPOTHESIS 2 “Profits of the principal”**

Hypothesis 1 questioned whether subjects target the effort levels predicted by the theory when different contracts (RPR, IPR and T) are used. We have established that the RPR scheme elicits higher effort choices than the IPR or the T schemes. These effort decisions influence the final profits received by the principal via two channels: (1) the total output produced and (2) the wages that the principal has to pay out. Therefore, it is not obvious what the comparison between the profits in each wage scheme will be, as higher effort levels chosen by workers may increase both the output produced and the workers’ wages.

Recall that in our experiments subjects performed only the role of workers. However, given the effort levels chosen by the workers, we can compute the profits of the principal that would have hired these workers. In Table 5 we report the average expected profits of the principal, which are calculated based on the value of the expected private and common shocks (both are zero in our parameterization).

<table>
<thead>
<tr>
<th></th>
<th>HV environment</th>
<th>LV environment</th>
</tr>
</thead>
<tbody>
<tr>
<td>RPR</td>
<td>26.11 (13.81, .44)</td>
<td>23.82 (17.40, .55)</td>
</tr>
<tr>
<td>IPR</td>
<td>19.59 (8.49, .25)</td>
<td>23.08 (6.66, .19)</td>
</tr>
<tr>
<td>T</td>
<td>12.24 (21.99, .57)</td>
<td>−1.68 (23.21, .60)</td>
</tr>
</tbody>
</table>

Table 5: Expected Profits of the Principal (st. dev., st. error)
As we show in Table 5, in both environments the principal gets higher profits when he uses the RPR than the IPR or the T contracts. To test whether these differences are significant, we use the expected profits of the principal (one observation per wage scheme per session). We perform the Wilcoxon rank-sum test that asks whether the expected profits of the principal in different wage schemes come from populations with the same distribution and report these results in Table 6.

<table>
<thead>
<tr>
<th>Environment</th>
<th>Expected profits of the principal (p-value)</th>
</tr>
</thead>
<tbody>
<tr>
<td>HV</td>
<td>RPR &gt;**IPR ( (p = 0.0495) )</td>
</tr>
<tr>
<td></td>
<td>RPR &gt;**T ( (p = 0.0339) )</td>
</tr>
<tr>
<td></td>
<td>IPR &gt;*T ( (p = 0.0770) )</td>
</tr>
<tr>
<td>LV</td>
<td>RPR = IPR ( (p = 0.8270) )</td>
</tr>
<tr>
<td></td>
<td>RPR &gt;**T ( (p = 0.0339) )</td>
</tr>
<tr>
<td></td>
<td>IPR &gt;**T ( (p = 0.0339) )</td>
</tr>
</tbody>
</table>

\( >^* \) (\( >^* \)) the profits are significantly higher at 5% (10%)

Table 6: Wilcoxon rank-sum test of expected principal’s profits

As we see in Table 6, the principal gets (weakly) higher profits when the RPR scheme is used than when the IPR or T scheme is used. In the HV environment, the rank of the wage schemes in terms of profits is as the theory predicts. However, in the LV environment, the IPR continues to do better than the Tournament. Given that the subjects drastically undershoot effort predictions in the Tournament, this result is not surprising. It does however highlight the risk a principal faces when employing the Tournament: the cost to the principal (the high prize plus the low prize) is independent of the output produced by the agents. In the case where workers produce an effort below the predicted one, the principal still has to give out the same level of payments without being compensated for it by a high output.

This may be a weakness of the Tournament payment scheme: in this scheme, the principal not only bears the full cost of a negative shock, but he/she also bears the full cost if agents deviate from theory and put in below-theoretical efforts. In this sense, the tournament is more prone to behavioral issues than the IPR or the RPR where
agents’ wages are reduced if agents exert more effort. As far as we know, our paper is the only one to explicitly incorporate a common shock into the experimental design (previous experiments assumed this component away since it is theoretically irrelevant to an agent’s decision-problem). Because agents exert lower effort levels than the equilibrium prediction (in our experiment and in others), in situations where the common uncertainty is large (as in our Low Variance treatment), the principal is more penalized from deviations since the contract he offers promises high payments to the agents. Understanding the behavior of the subjects in the Tournament scheme requires further research.

Finding 2. Consistent with Hypothesis 2, in the HV environment the RPR scheme delivers strictly higher profits to the principal relative to both the IPR and T. In the LV environment, however, while the RPR scheme generates strictly higher profits relative to the Tournament scheme, there is no statistical difference between the RPR scheme and the IPR scheme.

5 Experiment II

In the previous section we have established that the RPR contract performs better than both the IPR and T contracts, as the theory predicts. In other words, our results indicate that when “forced” to use the RPR scheme, subjects perform well. There are, however, several recent studies that suggest that people may dislike competition and when given the choice would not opt into such a payment scheme.\(^{29}\) Thus, it is entirely possible that people exhibit some aversion to sorting into the RPR contracts when other alternatives are available in the labor market. We design Experiment 2 to investigate this possibility.

\(^{29}\) There is a growing literature that shows that some people opt out of competitive and risky environments when other alternatives are available. Morgan, Orzen and Sefton (2011) show that even when the prize is high, subjects choose to enter a contest at rates much below those predicted. Niederle and Vesterlund (2007) show that women may shy away from competition; Gneezy, Leonard and List (2008) establish that culture matters in determining gender differences in competitive behavior by studying patriarchal and matrilineal societies; Booth and Nolen (2009a, 2009b) show that girls from single-sex schools are more likely to behave as boys in terms of risk-taking. Also, see Eriksson and Villeval (2009) for a comparison of choices when people self-select into different wage schemes versus when schemes are exogenously imposed. This is, however, not true of all tournaments: more recently, Cason, Masters and Sheremeta (2010) show that in tournaments in which the prize is divided in proportion to output subjects exert higher effort levels and there is less “entry aversion.”
5.1 Hypotheses

In Experiment 2 subjects can self-select into one of the three schemes (RPR, IPR and T). The goal of this experiment is to test whether people have any biases against or in favor of the RPR scheme when other alternatives are available. We also test how experience affects those biases.

Theoretically, when the value of an outside option is kept fixed ($\bar{v} = const$) workers should be indifferent between the contracts. Therefore, they should be equally likely to select each of the three available schemes as they all guarantee the same expected utility. We start by testing if this is indeed the case when laboratory subjects have to choose a wage scheme without having had any experience. We then let subjects experience all three wage schemes and again let them self-select into one of them. This last part of Experiment 2 lets us ask whether our early sorting results depend on experience.

**Hypothesis 3:** Subjects are indifferent about which payment scheme to select into.

5.2 Procedures

In Experiment 2 we used the HV environment from Experiment 1 and held all the parameters of this environment fixed throughout the session. A subject participating in Experiment 2 went through the following stages:

- We described all three wage schemes.

- Stage 1: Subjects chose which wage scheme they wanted to play for a single (paid) round.\(^{30}\)

- Stage 2: Subjects were forced to play 20 rounds of each scheme. They were paid for a single round for each of the three schemes, determined randomly.

- Stage 3: Subjects chose which scheme they wanted to play for 20 additional rounds. They were paid for a single of those 20 rounds, determined randomly.

\(^{30}\)If an odd number of people chose T or RPR contracts, then one random subject was “paired” twice with the remaining person. However, he/she was paid only once as were the other subjects.
As in Experiment 1, at the beginning of each round, subjects were randomly re-matched. At the end of each round, subjects were told their output as well as the output of the person they were paired with. We ran two sessions of Experiment 2. A total of 47 subjects participated.31

Stage 1 aims to observe whether subjects have an initial bias for or against the RPR scheme, while Stage 3 will show whether this initial attitude changes after subjects have experienced all three available contracts. One important feature of our design is that unlike after each round of Stage 2, subjects were given no feedback after Stage 1, so that when they arrived at Stage 3, all subjects had the same learning experience. We believe that the decision of which scheme to play for Stage 3 was not impacted by any result from Stage 1 that could have been made more salient than those from Stage 2 since in Stage 1 a single round was played. Put differently, the only difference between Stages 1 and 3 is in the amount of experience subjects had.

In terms of incentives, subjects could earn $4 for Stage 1, $4 for each set of 20 rounds in Stage 2, and finally another $4 for Stage 3. Thus, there was a large emphasis on Stage 1, as their single-round decision would be decisive for their payment.

5.3 Results

HYPOTHESIS 3 “Self-selection into wage schemes”

Recall that the value of the outside option was held fixed in all three wage schemes. Therefore, we expect that subjects will be absolutely indifferent between which contract to choose since it guarantees them the same expected utility. The choices of inexperienced subjects for Stage 1 are summarized in Table 7. As is obvious, subjects are not shying away from the RPR scheme. Indeed, more than half of them choose to participate in the RPR, giving up the tournament and IPR.

Even though theoretically subjects should be indifferent between the three wage contracts, the chi-square goodness of fit test rejects the hypothesis that the distribution of

31 The only difference between these two sessions was that in Stage 2 the order of contracts was RPR, IPR, and T for one session and IPR, RPR and T for another. There was no statistical or qualitative differences in the behavior in the two sessions, which rules out the possibility that our results are driven by the order of the contracts presented in Stage 2.
Table 7: Proportion of inexperienced subjects self-selecting in each contract

<table>
<thead>
<tr>
<th>Wage contract</th>
<th># of choices (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>RPR</td>
<td>26 (55.3%)</td>
</tr>
<tr>
<td>IPR</td>
<td>12 (25.5%)</td>
</tr>
<tr>
<td>Tournament</td>
<td>9 (19.2%)</td>
</tr>
</tbody>
</table>

Table 7: Proportion of inexperienced subjects self-selecting in each contract

subjects into schemes is \( \left( \frac{1}{3}, \frac{1}{3}, \frac{1}{3} \right) \) in Stage 1.

In Stage 2, subjects had an opportunity to experience all three wage schemes (20 rounds of each scheme). The data from Stage 2 allow us to check the robustness of our results from Experiment 1: subjects exert effort levels of 36.43 in the RPR, 33.16 in the IPR and 29.09 in the T scheme. These results are well in line with the theory. Thus, it seems that subjects understood the incentives of each wage scheme, as we anticipated given the results of Experiment 1.

By construction, the subjects’ expected returns from each wage scheme are equal. In the laboratory, on average, subjects responded to the wage schemes in the theoretically predicted manner. Given these two facts, we would expect to observe equal shares in each wage scheme for Stage 3. Indeed, inexperienced subjects in Stage 1 may have been influenced by ”fears” that some wage schemes are harder to understand or more difficult to do well in. After their performance in Stage 2, such fears should be mitigated. Table 8 shows us the proportion of subjects in each wage scheme.

Table 8: Choices of wage contracts in Stages 1 and 3

<table>
<thead>
<tr>
<th></th>
<th>Stage 3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>RPR</td>
</tr>
<tr>
<td>RPR</td>
<td>17 (36.2%)</td>
</tr>
<tr>
<td>IPR</td>
<td>5 (10.6%)</td>
</tr>
<tr>
<td>Tournament</td>
<td>6 (12.8%)</td>
</tr>
<tr>
<td>Total</td>
<td>28 (59.6%)</td>
</tr>
</tbody>
</table>

Table 8: Choices of wage contracts in Stages 1 and 3

The RPR scheme is the only one of the three schemes where more than half of the subjects who initially chose it chose it again: 17 out of the 26 subjects chose the RPR
scheme both in Stage 1 and in Stage 3. For both the IPR and the Tournament, more than half of the subjects moved away from their initial choices.

**Finding 3.** Subjects do not have any initial aversion to the RPR scheme. In fact, without any experience subjects seem to choose the RPR scheme more often than the IPR or the T schemes. Similarly, after experiencing all three wage schemes, subjects tend to favor the RPR scheme.

6 Conclusions

In this paper we study a moral hazard environment, in which a principal hires several workers and needs to decide how to compensate them for the output they produce. The workers are subject to two types of uncertainty: private and common. In such an environment, theory predicts that the principal should use both the absolute and relative performance of workers to balance incentives and risk. The RPR scheme is the simplest one that does the job: an agent receives payment in proportion to his output (as in standard piece-rate schemes) but is also penalized for producing less than the average output of other agents. Although the theoretical advantage of the RPR-like schemes has long been noticed in the literature, empirical papers that study this question have produced mixed results. Given the lack of established evidence in favor of the effectiveness of relative performance schemes and the natural concern that such schemes may perform worse than simpler schemes like piece-rate for example (due to an aversion toward competition or because of other-regarding preferences for example), the rare use of RPR schemes may have been understandable.\(^\text{32}\) In this paper we take the first step in investigating the performance of RPR schemes in the controlled laboratory experiment.

We do this by conducting a controlled laboratory experiment in which we fix the environment and vary only the wage contract imposed on the agents. First, we study the performance of the RPR scheme versus two alternative schemes, the piece-rate scheme (which uses only the absolute performance) and the tournament scheme (which uses only the relative performance). Our results indicate that the RPR scheme allows a principal to elicit higher effort levels from the workers and delivers (weakly) higher profits to the

\(^{32}\)For a survey of the literature on the dearth of RPR see Lazear and Oyer (2007).
principal. Second, we test how subjects sort themselves into the RPR scheme when other alternatives are available. Our results suggest that subjects do not seem to exhibit any ex-ante or ex-post dislike of the RPR scheme. To sum up, consistent with theoretical predictions, the RPR scheme seems to perform well both when it is exogenously imposed on agents and when it is available for self-selection on a par with other alternatives.

Most empirical papers (see Section 2) have tried to look for explicit evidence of RPR schemes and have concluded that they are seldom used in the labor market. However, incentives may be more implicit. For example, suppose that the reason people get fired is poor relative performance. That is exactly the way the RPR scheme works when payment is defined as the probability to keep a job (as opposed to the wage as in our study). So far, the empirical work that addresses this possibility has only used data on CEO turnovers, and the conclusions of the different studies diverge. Another possibility is that workers are rewarded in terms of their future wages and not their current ones. Indeed, Devaro (2006) shows the importance of relative performance (rather than absolute performance) when managers decide who to promote. So, it may be that current wages do not fit a model where performance is compared to that of one’s peers, but that progression within a firm is. In other words, RPR may be more prevalent than what we think and what has been studied empirically.

---

33 Morck, Shleifer and Vishny (1989) find that turnovers of CEOs are equally likely to occur in troubled and in healthy industries, suggesting that industry shocks are filtered from the dismissal decision. Jenter and Kanaan (2006) document that CEOs are significantly more likely to be dismissed from their jobs after bad industry and bad market performance.
References


