Authors
Ederer, Florian
Manso, Gustavo

Publication Date
2009-08-04

Is Pay-for-Performance Detrimental to Innovation?

Florian Ederer and Gustavo Manso*

August 4, 2009

Abstract

Previous research in economics shows that paying the agent based on performance induces the agent to exert more effort thereby enhancing productivity. On the other hand, research in psychology argues that performance-based financial incentives may inhibit creativity and innovation. In a controlled laboratory experiment, we provide evidence that the combination of tolerance for early failure and reward for long-term success is effective in motivating innovation. Subjects under such an incentive scheme explore more and are more likely to discover a novel business strategy than subjects under fixed-wage and standard pay-for-performance incentive schemes. We also find evidence that the threat of termination can undermine incentives for innovation, while golden parachutes can alleviate these innovation-reducing effects. Our results suggest that appropriately designed incentives are useful in motivating creativity and innovation.

*Ederer and Manso are at the University of Californiat at Los Angeles and the Massachusetts Institute of Technology (E-mails: ederer@ucla.edu, manso@mit.edu). We would like to thank Dan Ariely, Nittai Bergman, Bruno Biais, Arthur Campbell, Ernst Fehr, Bob Gibbons, Lorenz Götte, Bengt Holmstrom, Muriel Niederle, Enrico Perotti, Thomas Philippon, Sebastien Pouget, Drazen Prelec, Panle Jia, Andrei Shleifer, Johannes Spinnewijn, Antoinette Schoar, Jean Tirole, Eric Van den Steen, Christian Zehnder (our discussant) and seminar participants at Chicago GSB, Gerzensee, LSE, Toulouse, University of Munich, the Harvard-MIT Organizational Economics Seminar, the Kauffman Summer Legal Institute, and the NBER Corporate Finance Meeting for helpful comments as well as Hareem Ahmad and Yasmin Sanie-Hay for outstanding research assistance. Financial support from the NBER Innovation, Policy, and the Economy group is gratefully acknowledged.

1 Introduction

Previous research in economics advocates that paying the agent based on his performance induces the agent to exert more effort thereby improving productivity. There is ample evidence supporting this thesis in different types of studies. For example, Lazear (2000) shows that the productivity of windshield installers in Safelite Glass Corporation increased when management changed their compensation from fixed wages to piece-rate pay. Shearer (2004) finds similar evidence in a randomized field experiment with Canadian tree planters. Dickinson (1999) shows that subjects in a laboratory experiment type more letters when their compensation is more sensitive to performance. As in the above examples, most of the existing evidence of the effect of financial incentives on performance comes from studying simple routine tasks, in which effort is the main determinant of productivity.

In contrast, a substantial body of experimental and field research in psychology provides evidence that, in tasks that require exploration and creativity, pay-for-performance may actually undermine performance. McGraw (1978), McCullers (1978), Kohn (1993) and Amabile (1996) summarize the findings of this line of research by stating that pay-for-performance encourages the repetition of what has worked in the past, but not the exploration of new untested approaches. These studies conclude that in tasks that involve creativity and innovation, monetary incentives should not be used to motivate agents.

Can performance-based financial incentives motivate innovation in creative tasks? Using a task that involves innovation through experimentation, we study subject performance under different incentive schemes in a controlled experimental setting. The paper provides evidence that incentive schemes that tolerate early failure and reward long-term success lead to more innovation and better performance than fixed wages or standard pay-for-performance incentive schemes. This result stands in contrast to previous research in psychology, which suggests that financial incentives inhibit innovation.[1] It also stands in contrast to principal-agent models of repeated effort, according to which incentive schemes that tolerate early failure should produce lower effort and productivity than standard pay-for-performance incentive schemes.[2]

Innovation is the production of knowledge through experimentation (Arrow (1969), Weitzman (1979)). As pointed out by March (1991), the central concern that arises when learning through experimentation is the tension between the exploration of new untested approaches and the exploitation of well-known approaches. Manso (2007)

[1] See Kohn (1993), Amabile (1996) for surveys of the psychology literature on this theme.

[2] See, for example, Rogerson (1985), Holmstrom and Milgrom (1987) and Sannikov (forthcoming). These models do not incorporate learning from experimentation which is a central component of the innovation process and also of the task used in our experiment.

incorporates this tension into a principal-agent model to study incentives for creativity and innovation. He shows that the optimal incentive scheme that motivates innovation exhibits substantial tolerance for early failure and reward for long-term success.

In our experiment, subjects control the operations of a lemonade stand for 20 periods. In each period of the experiment, subjects make decisions on how to run the lemonade stand and observe the profits produced by their inputs. Subjects must choose between fine-tuning the product choice decisions given to them by the previous manager ("exploitation") or choosing a different location and radically altering the product mix to discover a better strategy ("exploration").

To study the impact of different incentive schemes on productivity and innovation, we consider three different treatment groups. The only difference across these treatment groups is the compensation offered to subjects. Subjects in the first group receive a fixed-wage in each period of the experiment. Subjects in the second treatment group are given a standard pay-for-performance (or profit sharing) contract, receiving 50% of the profits produced during the 20 periods of the experiment. Subjects in the third treatment group are allocated a contract that is tailored to motivate exploration. Their compensation is 50% of the profits produced during the last 10 periods of the experiment.

Our main hypothesis is that subjects under the exploration contract are more likely to find the optimal business strategy than subjects under the fixed-wage and standard pay-for-performance contracts. Two features of the exploration contract encourage subjects to explore. First, tolerance for early failure permits subjects to fail at no cost in the first 10 periods while they explore different strategies. Second, the perspective of being paid for performance later on encourages subjects to learn better ways of performing the task.

Our results provide strong support to the main hypothesis stated above. Subjects under the exploration contract end the experiment in the best location 80% of the time, while subjects under the fixed-wage and the pay-for-performance contracts end the experiment in the best location only 60% and 40% of the time respectively. To explain these differences we look at the reasons behind the poor performance of subjects under the fixed wage and pay-for-performance contracts. Even though subjects under the fixed-wage contract explore a lot, they are not as systematic in their exploration as subjects who are given an exploration contract. For example, when we analyze the notes subjects take in a table we provide to them at the beginning of the experiment, we find that only 55% of the subjects under the fixed-wage contract carefully keep track of their choices and profits while under the exploration contract 82% of the subjects keep track of their choices and profits using the table. Subjects under the pay-for-performance contract, on the other hand, tend to direct their effort towards fine-tuning the previous manager's product mix, instead of searching for better locations. During the first 10 periods of the experiment, subjects under the exploration contract choose a location other than the business district 80% of the time, while subjects under the pay-for-performance contract do so only 50% of the time.

We also compare the overall profits of subjects under the different contracts. Subjects under the exploration contract obtain higher average profits than subjects under the fixed-wage and pay-for-performance contracts. This result does not arise in a theoretical model in which the agent is risk-neutral. A risk-neutral agent who is paid 50% of his total profits should deliver higher profits than a risk-neutral agent who is paid 50% of his total profits in the last 10 periods of the experiment. This leads us to study the effects of different attitudes towards risk on the observed outcome under the different contracts. We find that risk aversion plays an important role in explaining differences in the exploration behavior and performance of the subjects under the pay-for-performance contract. Under the pay-for-performance contract, more risk-averse subjects are less likely to find the optimal strategy and they obtain lower average profits than less risk-averse subjects. Other explanations, such as pessimism about exploration, are also possible.

Finally, to study the effects of termination on innovation and performance, we introduce two new treatment groups: a termination treatment group and a termination with golden parachute treatment group. Subjects in both groups receive the exploration contract and are also told that the experiment will end early if their profits in the first 10 periods are lower than a certain threshold. Subjects in the termination with golden parachute treatment group are told that they will receive a reparation payment if the experiment ends after 10 periods. Our hypothesis is that subjects in the termination treatment are less likely to find the optimal location than subjects in the exploration treatment. We further hypothesize that subjects in the termination with golden parachute treatment group are more likely to find the optimal location than subjects in the pure termination treatment group. This hypothesis is supported by the data, since only 45% of the subjects in the termination treatment group find the optimal location, while approximately 65% of the subjects in the termination with golden parachute treatment group find the optimal location.

A common approach to the study of incentives using laboratory experiments is to give subjects a cost function and require them to choose an effort level (Bull, Schotter and Weigelt (1987), Fehr, Gachter and Kirchsteiger (1997), Nalbantian and Schotter (1997). More recently, however, researchers in the experimental economics literature have conducted studies in which subjects have to exert real effort. In these studies, subjects perform routine tasks such as typing letters Dickinson (1999), decoding a number from a grid of letters Sillamaa (1999), cracking walnuts Fahr and Irlenbusch (2000), solving two-variable optimization problems van Dijk, Sonnemans and van Winden (2001), and stuffing letters into envelopes Falk and Ichino (2006). These tasks, however, are inadequate to study incentives for innovation. In this paper, we introduce a task which involves real effort and also incorporates the trade-off between exploration and exploitation, essential in innovation activities.

Other papers in economics have found that pay-for-performance does not always increase performance. For example, Gneezy and Rustichini (2000) find that the effect of monetary incentives can be, for small amounts, detrimental to performance. Their interpretation is that a small compensation per unit of output may insult subjects leading them to exert less effort than if they were paid a fixed wage. In Fehr et al. (1997) and Fehr and Rockenbach (2003), the introduction of explicit incentives reduces the performance of workers in a firm-worker relationship because reciprocity was compromised. The focus of these papers differs from ours. They are concerned with tasks in which effort is the main input of the worker, and creativity is not an important determinant of performance.

Some other papers study the tension between exploitation and exploration in an experimental setting. In their analysis of an finite-horizon bandit problem Meyer and Shi (1995) show that subjects underexperimented with promising options and overexperimented with unpromising options. Banks, Olson and Porter (1997) study infinite-horizon bandit problems. They find that subjects use cut-off strategies and that discount rates and success probabilities affect subject behavior in the direction predicted by the theoretical model. Using a single-agent tournament game Merlo and Schotter (1999) demonstrate that learning and performance are lower in a setting where subjects are learning while they receive compensation than in a setting where subjects are learning before they receive compensation.

Finally, several recent papers study the effects of incentives on innovation. For example, Acharya and Subramanian (2007) investigate whether debtor-friendly bankruptcy laws foster innovation. Sapra, Subramanian and Subramanian (2008) and Atanassov (2007) study whether takeover pressure affects managers investment in innovation, while Aghion, Reenen and Zingales (2008) analyze the effects of competition and institutional ownership on innovation. **?** study whether stringent labor laws that restrict the dismissal of employees encourage innovation. Azoulay, Zivin and Manso (2007) study whether funding policies with tolerance for early failure and long horizons to evaluate results motivate creativity in scientific research. These papers provide support for the thesis that tolerance for early failure and reward for long-term success motivate innovation. However, because they use naturally occurring data, the variation in the incentive schemes is not exogenous and therefore estimation of the coefficients may be inconsistent. In our paper, we are able to study the effects of incentives on innovation by exogenously varying compensation schemes in a controlled laboratory environment.

2 Experimental Design

We establish an environment in which we can measure the effects of different incentive schemes on innovation and performance. For this purpose we conduct experiments in which participants have to solve a real task in which the trade-off between exploration and exploitation is central.

2.1 Procedures and Subject Pool

The experiments were programmed and conducted with the software z-Tree Fischbacher (2007) at the Harvard Business School Computer Laboratory for Economic Research (HBS CLER). Participants were recruited from the HBS subject pool using an online recruitment system. A total of 379 subjects participated in our experiments.

After subjects complete the experiment we elicit their degree of risk aversion and ambiguity aversion. We describe the exact procedures, which are standard, in the appendix. Subjects are then privately paid. A session lasted, on average, 60 minutes.

During the experiment, experimental currency units called francs were used to keep track of monetary earnings. The exchange rate was set at 100 francs = $1 and the show-up fee was $10. Subjects on average earned $24.

2.2 The Task

Subjects take the role of an individual operating a lemonade stand. The experiment lasts 20 periods. In each period, subjects make decisions on how to run the lemonade stand. These decisions involve the location of the stand, the sugar and the lemon content, the lemonade color and the price. The choices available to the subjects as well as the parameters of the game are given in the appendix.

At the end of each period, subjects learn the profits they obtained during that period. They also learn customer reactions that contain information about their choices. Customer feedback is implemented by having the computer randomly select one choice variable to provide a binary feedback to the subject.[3] For example, if the computer selects sugar content and the subject has chosen a sugar content that is above the optimal level for the particular location chosen by the subject, the feedback takes the form: "Many of your customers told you that the lemonade is too sweet."

Subjects do not know the profits associated with each of the available choices. Attached to the instructions, however, there is a letter from the previous manager which is reproduced in the appendix. The letter gives hints to the subjects about a strategy

[3]This feedback is only relevant to the location in which the subject chose to sell.

that has worked well for this manager and offers an accurate description of a good business strategy for one particular lemonade stand location. The strategy suggested by the previous manager involves setting the stand in the business district, choosing a high lemon content, a low sugar content, a high price and green lemonade. The manager's letter also states that the manager has tried several combinations of variables in the business district location, but that he has never experimented setting up the stand in a different location. It further suggests that different locations may require a very different strategy.

The participants in the experiment thus face the choice between fine-tuning the product choice decisions given to them by the previous manager (exploitation) or choosing a different location and radically altering the product mix to discover a more profitable strategy (exploration). The strategy of the previous managers is not the most profitable strategy. The most profitable strategy is to set the lemonade stand in the school district, and to choose a low lemon content, a high sugar content, a low price and pink lemonade. The payoffs in the game were chosen in such a way that without changing the default location the additional profits earned from improving the strategy in the business district are relatively small. On the other hand, changing the location to the school required large changes in at least two other variables to attain an equally high profit as suggested by the default strategy.

In addition to the previous manager's letter, the instructions contain a table in which subjects can input their choices, profits, and feedback in each period. Subjects are told that they can use this table to keep track of their choices and outcomes. We use the information subjects record in this table as one measure of their effort during the experiment.

2.3 Treatment Groups and Predictions

We initially implement three treatment conditions in order to examine how different incentive schemes affect innovation success, exploration behavior, time allocation and effort choices. The only difference between the groups is the way subjects are compensated. The compensation language used in each of the treatment groups is as follows:

Treatment Group 1 (Fixed-Wage):

"You will be paid a fixed wage of 50 francs per period."

Treatment Group 2 (Pay-for-Performance):

"You will be paid 50% of the profits you make during the 20 periods of the experiment."

Treatment Group 3 (Exploration):

"You will be paid 50% of the profits you make during the last 10 periods of the experiment."

The first two treatment groups are motivated by previous research in economics and psychology. The third treatment group is motivated by previous theoretical research (see Manso (2007)), which argues that tolerance for early failure and reward for long-term success is optimal to motivate innovation. Under the exploration contract, subjects that perform poorly in the first 10 periods and perform well in the last 10 periods receive a higher compensation than subjects that perform well in the first 10 periods and poorly in the last 10 periods.

Our experiment allows us to address a number of hypotheses. Our main hypothesis concerns the extent to which the different payment schemes considered in our treatment groups affect the exploration activity of subjects. In particular, we hypothesize that subjects under the exploration contract condition should find the optimal business strategy more often than subjects in the other treatments.

Main Hypothesis: *Subjects under the exploration contract get closer to the optimal business strategy than subjects under the fixed-wage and pay-for-performance contracts.*

The main hypothesis addresses the key question of our research agenda. To effectively motivate innovation and exploration subjects should be given a compensation contract that tolerates early failure but rewards success in later periods. Tolerance for early failure allows subjects to explore different strategies early on without being concerned with losses in terms of their compensation. At the same time, the long-term reward induces subjects to exert effort to learn better ways of performing the task.

What are the alternative hypotheses in this setting? As pointed out by the psychology literature, financial incentives reduce intrinsic motivation, an important ingredient for innovation. According to this view, subjects under the fixed wage contract should get closer to the optimal business strategy than subjects under the other two treatment groups, which have their compensation tied to performance. On the other hand, according to dynamic principal-agent models in which the main concern is to induce the agent to exert effort, subjects under the fixed wage or exploration contracts should engage in shirking, while subjects under the pay-for-performance contract should provide effort during the 20 periods of the experiment.

Our main hypothesis naturally leads us to two sub-hypotheses which deal with the problems of the two other contracts we consider in this study. Relative to subjects under the exploration contract, subjects under the pay-for-performance contract engage in less exploration, while subjects under the fixed-wage contract exert less effort.

Exploration Sub-Hypothesis: *Subjects under the exploration contract are more likely to explore than subjects under the pay-for-performance contract who are more likely to focus on exploitation activities.*

Since the compensation of subjects under the pay-for-performance contract depends on their performance from the very first period, we hypothesize that they will explore less than subjects under the exploration contract. A subject under the pay-for-performance contract who uses his first few periods to explore different strategies is likely to obtain lower profits and consequently lower compensation during those periods.

While the exploration hypothesis explains the differential effects of exploration and pay-for-performance contracts it does not predict how subjects under the fixed-wage contract behave. Subjects under the fixed wage contract are guaranteed a fixed compensation and therefore do not face any costs from failing while they explore different strategies. Under a fixed-wage contract, however, subjects do not have explicit incentives for performance and we would therefore expect them to minimize the costly contemplation effort necessary to find the best business strategy.

Shirking Sub-Hypothesis: *Subjects under the fixed-wage contract exert less effort than subjects under the exploration contract.*

Since their compensation is independent of performance, subjects under the fixed-wage contract do not have incentives to perform well. Since performance in the task requires effort in the form of costly contemplation of choices and outcomes, we hypothesize that shirking will be more prevalent in the fixed-wage contract.

Note that while we predict that subjects under the exploration contract are more likely to explore than subjects under the pay-for-performance contract and less likely to shirk than subjects in the fixed-wage contract, it need not be the case that they also produce better average performance than subjects under these two other contracts.

3 Results

In this section we present the results obtained in our experiments comparing the outcome across the three main treatments (fixed-wage contract, pay-for-performance contract and exploration contract). There were 51, 46 and 47 subjects in each of these three treatments.

3.1 Innovation, Exploration Behavior and Effort Choice

We first focus on the exploration behavior of subjects across the three different conditions. Our first result shows that the prediction that the exploration contract leads to more innovation than the other two contracts is confirmed by the data.

Result 1 (innovation): *Subjects under the fixed-wage and pay-for-performance contracts are significantly less likely to choose to sell at the school (highest profit location) in the final period of the experiment than subjects under the exploration contract. Subjects under the exploration contract come closest to finding the optimal business strategy.*

Initial supporting evidence for Result 1 comes from Figure 1 which shows the proportion of subjects under the fixed-wage, pay-for-performance, and exploration contract conditions choosing to sell lemonade in a particular location in the final period. Consistent with our exploration hypothesis, subjects under the exploration contract setting are more likely to sell at the school which is the location with the highest profits in the final period of the experiment than subjects under the fixed-wage and pay-for-performance conditions. Whereas in the exploration contract condition more than 80% of subjects choose to sell lemonade at the school, only 40% of subjects choose to do so in the pay-for-performance condition and 60% choose to do so under the fixed-wage contract. Using Wilcoxon tests for independent samples we can show that these differences are highly significant between the exploration contract and the fixed-wage contract (p-value 0.0042) and the exploration and the pay-for-performance contract (p-value 0.0001). The difference is less marked between the fixed-wage and the pay-for-performance contract (p-value 0.0865).[4]

We also examine how close subjects come to finding the optimal strategy over the course of the experiment. This can easily be measured by examining the maximum per period profit achieved by subjects throughout the course of the experiment. Per period profit is a more comprehensive measure than location choice. It captures the multi-dimensional aspect of the task which involves the choice of several variables. On average, subjects under the exploration contract achieve the highest maximum per period profits (145 francs) while subjects under the fixed-wage (128 francs) and the pay-for-performance (117 francs) contracts perform worse on this dimension. The same pattern holds for final period profit where the respective values are 140 (exploration), 120 (fixed wage) and 111 francs (pay-for-performance). As before the differences in maximum per

[4]In addition, we estimated a logit model where the dependent binary variable takes the value 1 if the final location choice is the school which is the optimal location choice in the experiment, and 0 otherwise. The independent variables are binary variables for the three different contracts. As before, the coefficient estimates show that subjects under the pay-for-performance (p-value 0.0001) and fixed-wage contract (p-value 0.0054) are significantly less likely to choose to sell in the school in the final period of the experiment than subjects in the exploration contract. The negative effect on finding the optimal location in which to sell is particularly pronounced for the pay-for-performance contract while the difference between fixed-wage and pay-for-performance contracts is not as significant (p-value 0.0865).

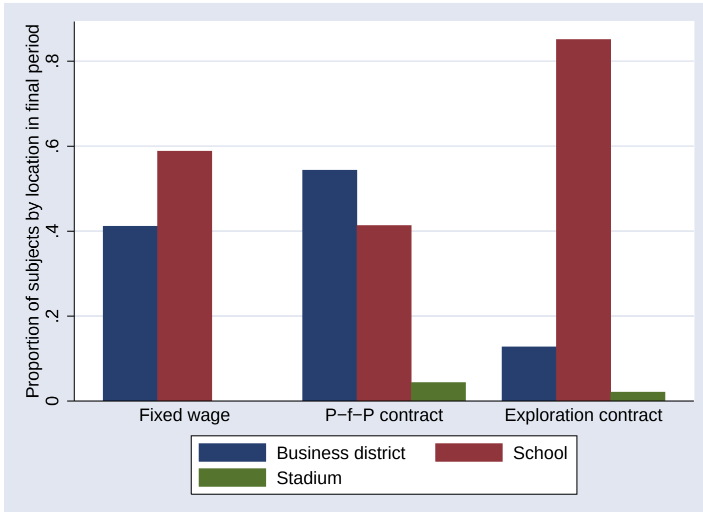

Figure 1: Proportion of subjects by location in the final period of the experiment for the fixed-wage, pay-for-performance and exploration contracts.

period profit as well as final period profit between the exploration contract and the other two contracts are highly significant (p-values of 0.013 and 0.0001 for maximum profit, p-values of 0.009 and 0.0001 for final period profit) while the difference between the fixed-wage and the pay-for-performance contract is not statistically significant (p-value 0.1144 for maximum profit, p-value 0.28 for final period profit).

To explain why subjects under the exploration contract are more likely to find the optimal location and business strategy than subjects under the other two contracts, we analyze different measures of exploration and effort. The next result shows that subjects under the exploration contract explore more than subjects under the fixed-wage contract while subjects under the pay-for-performance contract explore the least.

Result 2 (exploration behavior): *Subjects under the pay-for-performance contract explore less than subjects under the fixed-wage contract and the exploration contract with the latter exploring the most.*

Using the different choice variables available to the agents we can construct several measures of exploration activity. We first analyze location choice behavior. Subjects in the pay-for-performance condition explore locations other than the default location (business district) less often than subjects under the other two contracts with subjects under the exploration contract choosing to explore the most often. While subjects under the exploration contract choose a location other than the default location in 82% and 85% of cases in the first and the last 10 periods, subjects under the fixed-wage contract choose to do so only in 60% and 63% of cases and the proportions are as low as

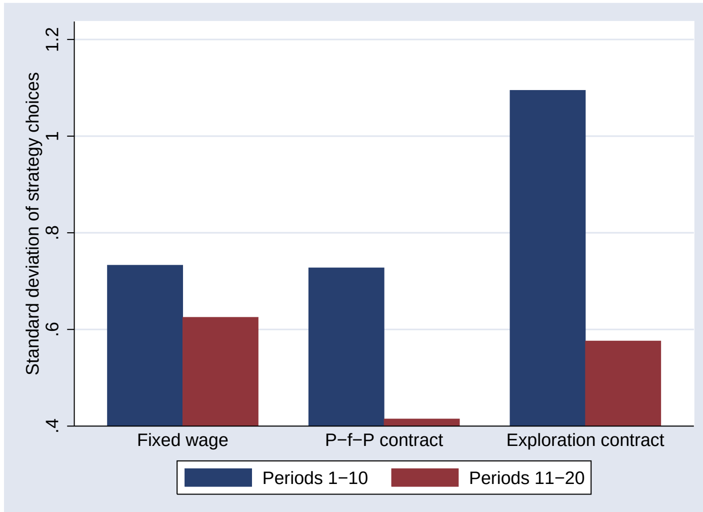

Figure 2: Average subject-specific standard deviation of strategy choices for the three continuous variables (sugar content, lemon content, price) in periods 1-10 and 11-20 of the experiment for the fixed-wage, pay-for-performance and exploration contracts.

51% and 48% for subjects in the pay-for-performance contract. The tolerance for early failure of the exploration contract relative to the fixed-wage and pay-for-performance contracts encouraged individuals to attempt new untried approaches in the first 10 periods. Using Wilcoxon tests for independent samples reveals that this difference in location choice behavior between the different contracts is statistically significant. In the first 10 periods subjects under the exploration contract choose to explore a different location more often than subjects under the fixed-wage contract (p-value 0.0053) and the pay-for-performance contract (p-value 0.0001). The difference in exploration behavior as measured by location choice in the first 10 periods is not statistically significant between subjects under the fixed-wage and the pay-for-performance contracts (p-value 0.1482), but subjects under the fixed-wage contract choose to explore significantly more often than subjects under the pay-for-performance contract in the last 10 periods of the experiment (p-value 0.0985).

This particular form of exploration activity is also reflected in Figure 2 which shows the average subject-specific standard deviation in strategy choices for the three continuous choice variables (sugar content, lemon content and price) during the first and last 10 periods of the experiment. This standard deviation measure captures variation in all the variables of this multi-dimensional choice problem. There are several features of note.

First, the variability of action choices significantly declines over the course of the experiment in the pay-for-performance (p-value 0.0005) and the exploration contracts

(p-value 0.0001). This occurs because in periods 11 to 20 the beneficial learning effects of exploration relative to exploitation are no longer as large as at the beginning of the experiment since the time horizon is shorter. In contrast, the variability of action choices only decreases slightly in the fixed-wage contract and this decline is not statistically significant (p-value 0.2194). Since agents are not penalized for low profits, exploration behavior in the fixed-wage contract is exclusively driven by intrinsic motives and subjects may therefore continue to explore even though the additional benefits of exploration are small.

Second, the variability of action choices in the first 10 periods is significantly higher in the exploration contract than in the pay-for-performance (p-value 0.0012) and the fixed-wage contracts (p-value 0.0027). This shows that subjects under the exploration contract experiment and consciously make very different action choices in a directed attempt to find more promising strategies. In contrast, in the pay-for-performance contract the standard deviation of action choices is much lower as subjects opt to fine-tune the default values. This is also true for subjects under the fixed-wage contract who explore less than subjects under the exploration contract during the first 10 periods. However, because subjects in the other two treatments explore much less in the later periods of the experiment when their compensation is directly linked to their performance, the variability of action choices of subjects under the fixed-wage contract is higher (though not always significantly so) than in the pay-for-performance (p-value 0.0246) and the exploration contracts (p-value 0.6567). The relatively high exploration behavior of subjects under the fixed-wage contract in the last 10 periods of the experiment also explains why they are more likely to find the highest-profit location than subjects under the pay-for-performance contract who explore the least over the entire course of the experiment among the three contract treatment groups.

We also expect the variability of profits to mirror the variability of action choices. This is indeed the case. First, the variability of profits significantly declines over time with the decline in variability being particularly marked for the exploration contract and the pay-for-performance contracts. Second, the variability of profits in the first 10 periods is significantly higher for subjects under the exploration contract than subjects under the other two contracts, while there is no significant difference in profit variability across subjects under the three contracts in the last 10 periods.

Furthermore, we use Cox hazard rate models to analyze the dynamics that govern the strategy choices of individuals in the experiment. In particular, this allows us to test whether the different treatment conditions also influence whether, once they have decided to explore, subjects continue to explore and what other factors contribute to making them persist in their exploration activities. We classify subjects as having entered an explorative phase as soon as they choose a location other than the default

	Cox Hazard Rate Models			
	Period 1-20	Period 1-10	Period 1-20	Period 1-10
	b/se	b/se	b/se	b/se
rh				
Fixed Wage	0.217	0.564***		
	(0.136)	(0.204)		
Pay-for-Performance	0.334***	0.632***		
	(0.126)	(0.201)		
Termination			0.487***	0.861***
			(0.110)	(0.185)
Parachute			0.297**	0.580***
			(0.123)	(0.199)
t				
Profits	-0.001***	-0.001***	-0.001***	-0.001***
	(0.000)	(0.000)	(0.000)	(0.000)
Pseudo-R^2	0.011	0.013	0.014	0.015
N	2418	1068	2995	1458

Table 1: Estimates from a Cox hazard rate model reporting the hazard rates for exiting an exploration phase with the exploration contract as the baseline. Separate estimations are shown for the entire 20 periods of the experiment and the first 10 periods. Robust standard errors are reported in brackets. Statistical significance at the ten, five and one percent level is indicated by *, ** and ***.

location (business district) suggested by the previous manager. An explorative phase ends when subjects make only small changes to strategy choices relative to the previous period or switch back to the default location.[5] As can be seen from column 1 of Table 1, the hazard rate of ending an explorative phase is significantly higher under the pay-for-performance contract than under the exploration contract. The hazard rate is also higher in the fixed-wage contract although this effect is not statistically significant. Moreover, higher profits significantly decrease the hazard rate as subjects are encouraged to persist in their exploration effort. Column 2 of Table 1 shows that the estimates for the first 10 periods are qualitatively similar.

Finally, answers in the open-ended post-experimental questionnaire in which all sub-

[5]In particular, an explorative phase is defined as ending when a subject switches back to the default location or when a subject does not change location and lemonade color and also does not change lemon content, sugar content and price by more than 0.25 units. As a robustness check we also used other definitions thresholds for the end of an exploration phase. The resulting magnitudes and significance levels are very similar.

jects were asked to describe their strategies and the effect the compensation scheme had on their choices also reflected the described exploration pattern. Subjects under the exploration contract spontaneously argued that the tolerance for early failure of the compensation scheme as well as the strong rewards for success in later periods influenced their strategic choices, causing them to experiment with untested locations and action choices early on and then to choose and fine-tune the best available strategy beginning in period 11.

So far, our results have largely focused on exploration behavior. However, we also predicted that subjects under the fixed-wage contract should exert less effort than subjects under the other two contracts since their compensation does not depend on their performance in the experiment. As Result 3 shows, this is indeed the case.

Result 3 (time allocation and effort choice): *Subjects under the fixed-wage contract spend less time making and evaluating decisions and exert less effort recording their previous choices and outcomes in the experiment than subjects under the pay-for-performance and exploration contracts.*

A principal deciding whether to pay agents a fixed wage might worry that absent any intrinsic motivation and implicit incentives the agent will choose to minimize costly effort. Similarly, in our experiment—where subjects have to mentally focus and record past choices to try to maximize their performance—subjects whose compensation does not depend on their performance may choose to minimize costly and time-consuming contemplation and deliberation effort. Indeed, many subjects under the fixed-wage contract claimed in the post-experimental questionnaire that they attempted to minimize the time and effort necessary to complete the experiment since their performance did not affect their compensation. This pattern is also borne out in our experiment data.

While subjects under the fixed-wage contract spend only an average of 24 seconds on the decision screen (where subjects enter their strategy choices), subjects under the exploration and the pay-for-performance contracts spend 31 and 30 seconds respectively. That is, over the entire duration of the experiment, subjects under the exploration and the pay-for-performance contract condition spend almost 30% more time on the decision screen than subjects under the fixed-wage condition and these differences are statistically significant (p-values of 0.0014 and 0.0175) over the course of the entire experiment as well as in subperiods. Moreover, subjects in the exploration contract treatment spend significantly more time on the decision screen than subjects in the fixed-wage treatment (p-value 0.022) even during the first 10 periods of the experiment when they receive no compensation while this difference in time spent between the exploration and pay-for-performance contracts is not significant (p-value 0.8477). This evidence stands in contrast to dynamic principal-agent models of repeated effort, such as Rogerson (1985), Holmstrom and Milgrom (1987) and Sannikov (forthcoming), which predict that the exploration contract should induce more shirking during the first ten periods of the experiment than the pay-for-performance contract since under the exploration contract a subject's compensation is not tied to his performance during the first ten periods of the experiment. These models fail to incorporate the learning produced by the exploration of new strategies, which potentially enhances performance in later periods, and may thus provide incentives for the agent to exert effort in early periods, even when his compensation does not depend on productivity in those early periods. The results above suggest that experimentation and learning can indeed be important components in incentive problems, and should be taken into account when designing compensation schemes for innovative tasks.

Furthermore, in addition to spending less time making decisions, subjects under the fixed-wage contract also exert less effort by entering less information into the sheet given to them than subjects under the pay-for-performance and exploration contracts. Figure 3 shows that across the three contracts there is a considerably smaller proportion of subjects under the fixed-wage contract who fill out half or more of the fields in the decision table than in the other two contract treatments. This difference in effort choice is statistically significant between the exploration contract and the fixed-wage contract (p-value 0.0053) as well as between the pay-for-performance and the fixed-wage contract (p-value 0.0804). Subjects who are given the exploration contract or the pay-for-performance contract, record their past choices significantly more frequently in the table. The reward for success is sufficient to motivate subjects under the exploration or the pay-for-performance contract to exert more effort in the experiment. In the first 10 periods of the experiment subjects under the exploration contract are significantly more likely to record information than subjects in the fixed-wage contract (p-value 0.0111) thereby refuting once more the shirking prediction of the standard repeated moral-hazard model. The difference in effort exerted during the first 10 periods between subjects under the exploration contract and the pay-for-performance contract is not significant (p-value 0.5782).

The difference in effort choice between the exploration contract and the pay-for-performance contract is positive but not statistically significant (p-value 0.29). On the one hand, subjects under the pay-for-performance contract are given more powerful incentives overall since their compensation depends on performance both in the first and the last 10 periods of the experiment. On the other hand, since subjects under the exploration contract choose to experiment with very different strategies in the first 10 periods as we showed in Result 2, they need to exert more effort when evaluating their decisions than subjects under the pay-for-performance contract. This is also visible in Figure 3 which shows that effort declines in the pay-for-performance contract. This

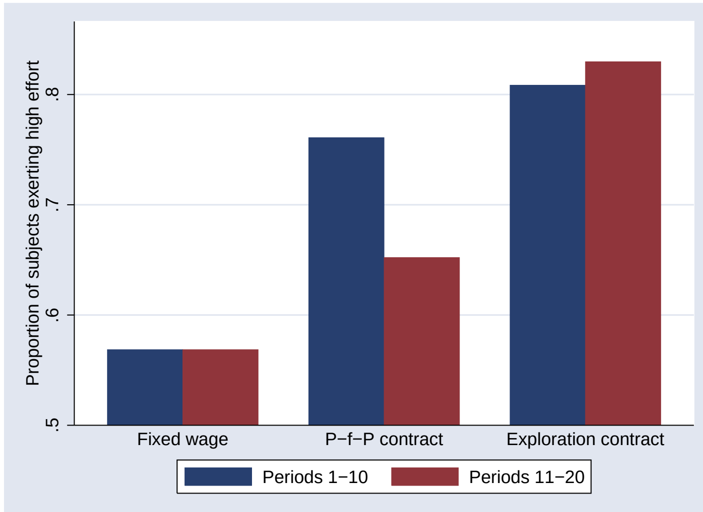

Figure 3: Proportion of subjects who complete more than half of the fields in the decision record table for the fixed-wage, pay-for-performance and exploration contracts.

occurs since subjects in the pay-for-performance contract essentially stop exploring and experimenting with different choices very early in the experiment and therefore they barely change their choices in the last 10 periods. Since there is little change, they do not have to record their choices as carefully as subjects in the exploration contract treatment.

We also note that time allocation and effort choice in the fixed-wage is strictly greater than zero since some of the subjects are sufficiently motivated by intrinsic rewards to exert effort. An inspection of effort choices by subjects in the fixed-wage treatment reveals a bimodal distribution. Subjects either fully record or do not record any of their past choices. Moreover, subjects in the fixed-wage treatment who exert more effort are more likely to successfully innovate: 65% of them end up selling at the school in the final period compared to 47% of the subjects who exert less effort, but this difference is not statistically significant (p-value 0.2047). However, maximum profits are significantly higher for subjects in the fixed-wage treatment who exert more effort (p-value 0.0298).[6]

3.2 Average Performance

Having confirmed that the innovation success, exploration behavior, time allocation and effort choice across the different contracts is in accordance with our theoretical predictions, we now turn to analyzing average performance differences. In particular,

[6]For a study of the effect of intrinsic motivation on innovation productivity, see for example Sauermann and Cohen (2008).

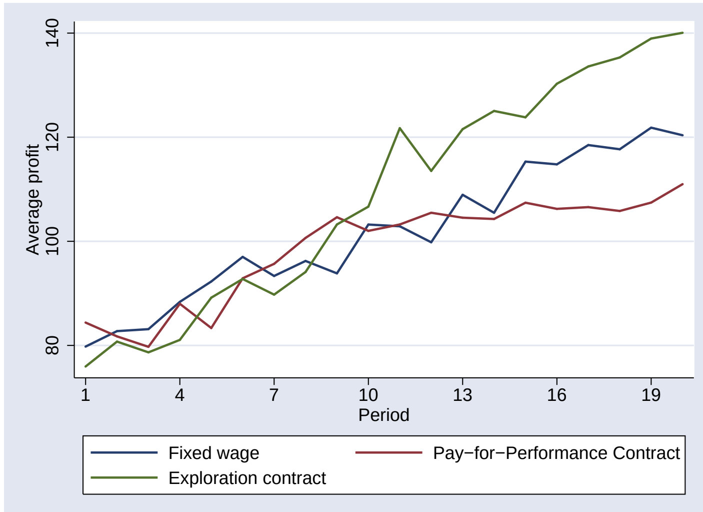

Figure 4: Average per period profits over the course of the experiment for subjects under the fixed-wage, pay-for-performance and exploration contracts.

we show that subjects' overall performance in the experiment as measured by average profit is highest in the exploration contract.

Result 4 (performance): *Subjects under the exploration contract produce higher average profits than subjects under the pay-for-performance and fixed-wage contracts.*

Preliminary evidence for Result 4 comes from inspecting the average profit for the three contracts. This performance measure is highest in the exploration contract (111 francs) and the difference in performance between the exploration contract and the pay-for-performance (96 francs) and the fixed-wage contract (102 francs) is statistically significant (p-values of 0.0009 and 0.0253). This difference in performance exists despite the fact that the average wage received by subjects under the exploration contract is lower than in the other two contracts.

We can also investigate the evolution of profits over time in Figure 4. From Result 1 we know that subjects under the exploration contract undertake thorough innovation efforts to find the best strategy in the first 10 periods. It is therefore not surprising that the variation in profits in the first 10 periods is also highest in the exploration contract. However, in terms of average profits the three contracts are virtually indistinguishable during the first 10 periods of the experiment. It is only after period 10 that the performance under the different contracts begin to diverge as subjects under the exploration contract revert to and subsequently fine-tune the best strategy they found during the first 10 periods of the experiment.

The result that profits are higher under the exploration contract than under the pay-for-performance contract does not arise in a model with a risk-neutral agent such as Manso (2007). If a risk-neutral agent is paid for overall performance, he should deliver higher performance than a risk-neutral agent who is only paid for performance in the last 10 periods of the experiment. The differences in performance documented in Result 4 naturally lead us to investigate what factors cause the departure from the theoretical predictions. As the following result shows, attitudes toward risk play an important role in explaining some of the heterogeneity in exploration behavior and performance of experimental subjects.

Result 5 (risk aversion): *Under the pay-for-performance contract more risk-averse subjects are significantly less likely to explore and to choose to sell in the optimal location in the final period of the experiment. They also produce significantly lower profits. Attitudes to risk have a similar (though statistically insignificant) effect in the exploration contract, while no systematic effects of risk are found for the fixed-wage contract.*

We now incorporate the subjects' different attitudes toward risk into our analysis. Using the data from the separate risk aversion experiment we classify subjects into more and less risk-averse groups. Figure 5 provides a first indication for the sign and magnitude of the effect of risk aversion on the likelihood of finding the best strategy. In this figure we use our risk aversion measures to further analyze the final period location choice as we did in Figure 1. We separately present final location choices for more and less risk-averse subjects for each of the three contracts. In the pay-for-performance contract, more risk-averse subjects are less likely to find the optimal location as they are less likely to explore than the less risk-averse subjects. This innovation-reducing effect of risk is statistically significant in the pay-for-performance contract treatment (p-value 0.0170) but it is not statistically significant in the other two treatments. This lower rate of innovation success caused by risk aversion is driven by the lower levels of exploration under the pay-for-performance contract since in this treatment the proportion of location choices other than the default location (p-value 0.0075) as well as the variability of action choices (p-value 0.0181) are significantly lower for subjects with higher risk aversion. However, in the exploration contract where subjects' failure is tolerated in early periods of the experiment and compensation has a much smaller risky component, as we would expect, the effect is smaller in magnitude and not statistically significant. The same is true in the fixed-wage contract where compensation entails no risk.[7]

Since more risk-averse subjects under the pay-for-performance contract are less likely to explore and therefore less likely to sell lemonade in the optimal location in the final

[7]Qualitatively similar results hold for a similarly constructed ambiguity aversion measure which we elicited using the experiment described in the appendix. The effects are of the same sign as the effects of risk aversion, but they are generally smaller in magnitude and in some cases not statically significant.

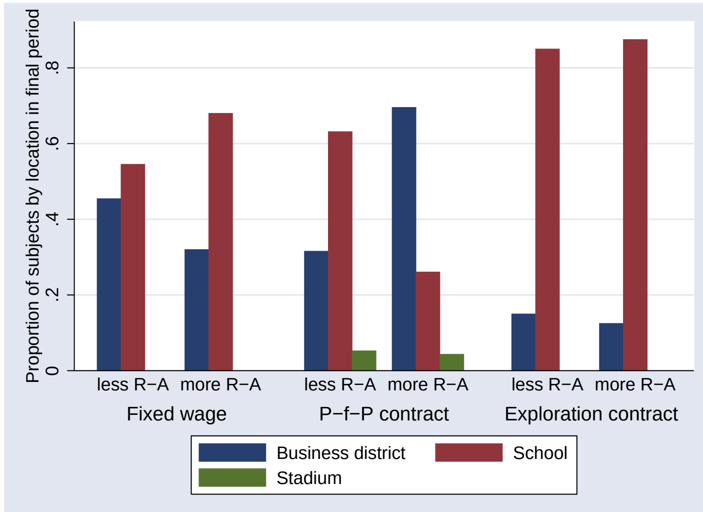

Figure 5: Proportion of subjects by location in the final period of the experiment for the fixed-wage, pay-for-performance and exploration contracts adjusting for differences in risk aversion.

period, they also produce lower profits as can be seen in Figure 6. This profit-reducing effect of risk aversion in the pay-for-performance contract is large in magnitude and statistically significant for maximum profit (p-value 0.0563) and final period profit (p-value 0.0382), but it is not statistically significant for average profit (p-value 0.1846). Furthermore, as in the case of the final period location choice, risk aversion also has a small negative but statistically insignificant effect on profit measures in the exploration and the fixed-wage contract treatment.

There could be reasons in addition to risk aversion for the difference in average profits across the three treatment groups. For example, in our experiment subjects are not given precise information about the profits associated with each of the available choices.[8] The differences in average profits across the three treatment groups could thus be due to subjects being pessimistic about the returns to exploration. The explanation is exactly the same as the one in the above two paragraphs with pessimism in place of risk-aversion.

[8]This is also the case in some psychology experiments which find that subjects under a fixed-wage contract perform better than subjects under a pay-for-performance contract. For a survey of this literature, see McGraw (1978), McCullers (1978), Kohn (1993) and Amabile (1996).

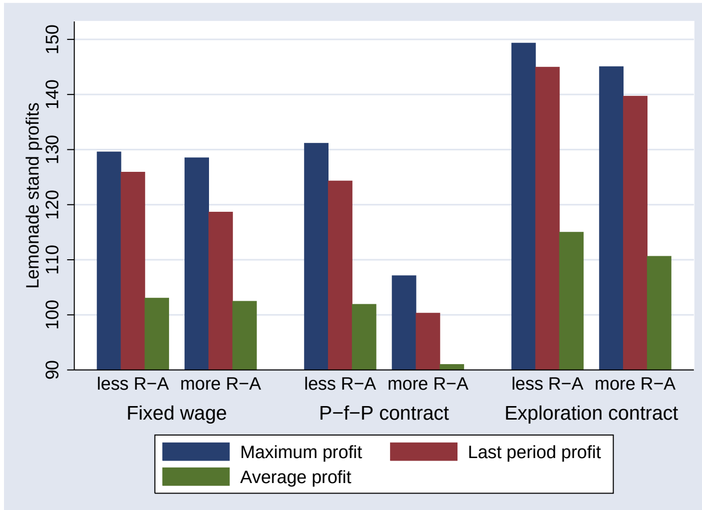

Figure 6: Maximum profit, last period profit and average per period profit of subjects under the fixed-wage, pay-for-performance and exploration contract adjusting for differences in risk aversion.

4 Termination

Having analyzed our main three treatment groups we now turn to investigating how the threat of early termination influences exploration behavior and performance. Early termination can undermine the exploration behavior induced by the exploration contract by eliminating the tolerance for early failure. We also show that this effect can be mitigated by the use of "golden parachutes"or reparation payments which subjects receive in case of early termination since these payments reintroduce tolerance for early failure.

We leave the design of the experiment unchanged, but introduce two new treatment groups that enable us to investigate the effects of termination and golden parachutes. The compensation for these two treatment groups is identical to the one in the exploration contract, except that we change the following sentences in the description of the compensation.

Treatment Group 4 (Termination):

"You will be paid 50% of the profits you make during the last 10 periods of the experiment. However, if the profits you make during the first 10 periods of the experiment are below 800 francs, the experiment will end early."

Treatment Group 5 (Termination with Golden Parachute):

"You will be paid 50% of the profits you make during the last 10 periods of the experiment. If the profits you make during the first 10 periods of the experiment are below 800 francs, the experiment will end early and you will receive a payment of 250 francs."

Pure termination inhibits exploration activities because it undermines a crucial aspect of an exploration contract, namely the tolerance for early failure.[9] While the threat of termination produces strong incentives for good performance, it also forces individuals to focus on producing good performance from the very beginning and thus reduces the incentives for exploration. In contrast, in the golden parachute treatment we expect subjects to explore a little more intensively than in the termination treatment at the beginning of the experiment despite the pending threat of termination since the golden parachute payment provides them with some insurance in case of failure. In particular, we have the following prediction:

Termination Hypothesis: *Subjects under the termination contract are less likely to find the optimal business strategy than subjects under the exploration treatment since the threat of termination has an exploration-deterring effect. However, subjects under the golden parachute treatment are more likely to find the optimal business strategy than subjects in the termination treatment since the reparation payment encourages exploration.*

In a setting where exploration is a key ingredient for achieving good performance, the threat of early termination is predicted to have adverse effects on innovation success and exploration. However, these effects are predicted to be mitigated by the use of golden parachutes. As we will show, these predictions are also borne out in our experiment data. We begin our analysis by showing that the threat of termination reduces the probability that subjects successfully innovate because the threat of early termination reduces exploration activities. Furthermore, the next result also shows that the adverse effects of termination are less pronounced in the golden parachute treatment.

Result 6 (termination): *The threat of termination has adverse effects on innovation success and exploration activities, but golden parachutes alleviate these negative effects. Risk aversion further reduces innovation success, exploration activities and performance in the termination treatment.*

There were a total of 71 and 78 subjects who participated in the termination and the golden parachute treatments. Figure 7 shows final period location choices in the exploration contract, termination and golden parachute treatments where in the case of the latter two treatments we eliminated subjects that are terminated after the first

[9] Note that the prediction that termination has an adverse effect on exploration depends crucially on our choice of the termination threshold which is chosen such that it can be achieved without exploring.

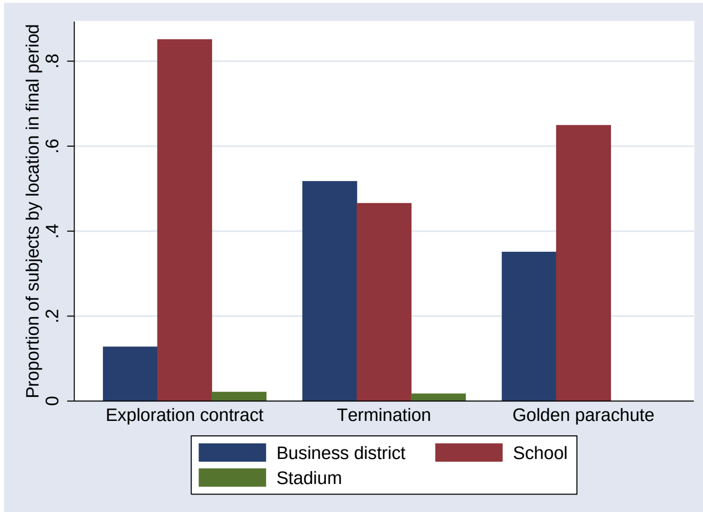

Figure 7: Proportion of subjects by location in the final period of the experiment for the exploration contract, termination and golden parachute treatments.

10 periods. The threat of termination in the pure termination and golden parachute treatment significantly reduces the probability that subjects end up choosing to sell at the best location in the final period of the experiment relative to the exploration contract treatment (p-values 0.0001 and 0.0200) while the use of golden parachutes raises the innovation success probability (p-value 0.0485) relative to the termination treatment. The same picture emerges when focusing exclusively on the final location choice after the first 10 periods using all the subjects in the termination and golden parachute treatments. As before, the threat of termination reduces the probability of finding the best location relative to the exploration treatment (p-values 0.0063 and 0.0562) and the use of reparation payments increases the innovation success probability in the golden parachute treatment relative to the termination treatment, although this effect is not large enough to be significant (p-value 0.3176).

We also analyze differences among treatments in the maximum profit and final period profit a subject achieves which serves as our other measure of innovation success. Focusing on subjects that are not terminated we again find that termination has an innovation-reducing effect since average maximum profit in the exploration contract treatment (145 francs) is significantly higher than in the termination (126 francs) and the golden parachute treatments (134 francs). The respective p-values are 0.0037 and 0.0772. Comparing the maximum profits for the termination and golden parachute treatments shows that the use of golden parachutes slightly mitigates these adverse effects, though the effect is not significant (p-value 0.1784).

The adverse effect of termination is more pronounced if we consider the full sample of subjects and only focus on the first 10 periods. The average maximum profit in the termination and the golden parachute treatments is again significantly lower than in exploration contract treatment (p-values 0.0032 and 0.0037). However, the difference between the termination and the golden parachute treatments is not statistically significant (p-value 0.7989).

As in our analysis of the three baseline treatments, we can trace the differences in innovation success back to differences in exploration behavior. To this end we again compare the number of times subjects choose to deviate from the proposed strategy and to explore a location other than the business district. To guard against potential selection effects arising from attrition we focus exclusively on choices in the first 10 periods. As expected, exploration is lower in the termination treatment where subjects shy away from exploring other locations in the first 10 periods. While the average proportion of location choices other than the default location is 82% in the exploration contract it is only 47% in the termination treatment and 59% in the golden parachute treatment. This exploration-reducing effect of the threat of termination is statistically significant (p-values 0.0001 and 0.0009). Moreover, as postulated before, golden parachutes increase exploration activities relative to the pure termination treatment and this beneficial effect is statistically significant (p-value 0.0495).

In the post-experiment questionnaire subjects argued that the threat of termination forced them to concentrate on selling in the business district and left no leeway for exploration. Further evidence for the exploration-reducing effect of the threat of termination and the exploration-increasing effect of reparation payments comes from comparing the variability of action choices in the first 10 periods for the full sample of subjects. The subject-specific standard deviation of action choices in the first 10 periods is highest in the exploration contract (standard deviation 1.09). This measure is significantly lower in the termination treatment (standard deviation 0.74, p-value 0.0014) and in the golden parachute treatment (standard deviation 0.79, p-value 0.0071). As before, the use of golden parachutes slightly increases exploration activity relative to the termination treatment, but this effect is not statistically significant (p-value 0.2821).[10]

Using the same hazard rate model as in our analysis of the baseline treatments though concentrating exclusively on the first 10 periods we can investigate how likely subjects are to persist in their exploration activities in the different treatments. Column 3 of Table 1 shows that both in the termination treatment and in the golden parachute treatment subjects are significantly more likely to stop exploring than in the exploration contract.

[10]The different proportions of subjects who are terminated in the termination and the golden parachute treatments are also in line with subjects exploring more in the latter case. While in the termination treatment 13 out of 71 subjects (18%) do not meet or exceed the termination threshold, 21 out of 78 subjects (27%) are terminated in the golden parachute treatment, but the difference is not statistically significant (p-value 0.2124).

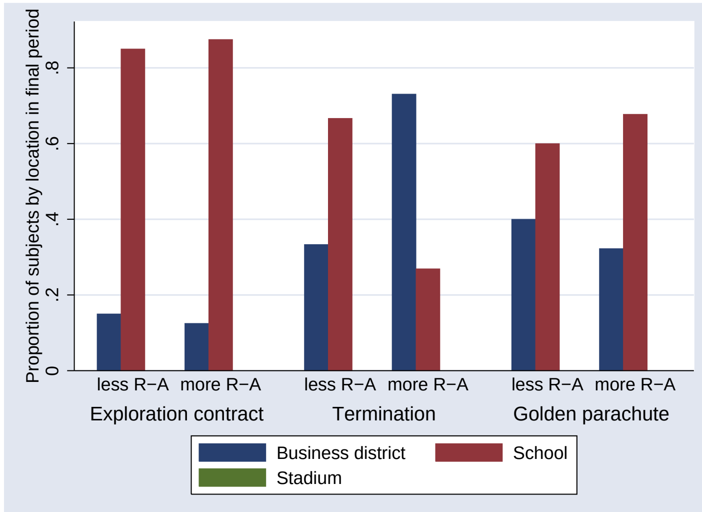

Figure 8: Proportion of subjects by location in the final period of the experiment for the exploration contract, termination and golden parachute treatment adjusting for differences in risk aversion.

Moreover, subjects in the termination treatment are also significantly more likely to stop exploring than subjects under the golden parachute treatment (p-value 0.0663). Column 4 of Table 1 reports estimates for the first 10 periods showing statistically significant differences in the hazard rate between the exploration contract treatment and the termination treatment as well as the golden parachute treatment. Note further that the difference between termination and golden parachute is also statistically significant (p-value 0.0604).

Risk aversion plays an important role in the termination treatment as can be seen in Figure 8, which shows final period location choice, and in Figure 9, which presents the different profit measures. More risk-averse subjects in the termination treatment are less likely to sell in the school in the final period of the experiment and they achieve lower maximum, final period and average profits. Throughout, there is a statistically significant negative effect of risk aversion in the termination treatment on the correct final period location choice (p-value 0.0041) as well as maximum profits (p-value 0.0023), final period profits (p-value 0.0041) and average profits (p-value 0.0037). This finding is in line with our previous analysis where we found similarly strong effects of risk aversion for the pay-for-performance contract which also induces individuals to achieve profits from the very beginning of the experiment instead of learning through exploration. In contrast, like our finding for the exploration contract treatment there is no statistically significant effect of risk aversion in the golden parachute treatment.

Finally, we can also confirm that in the termination treatment a high degree of

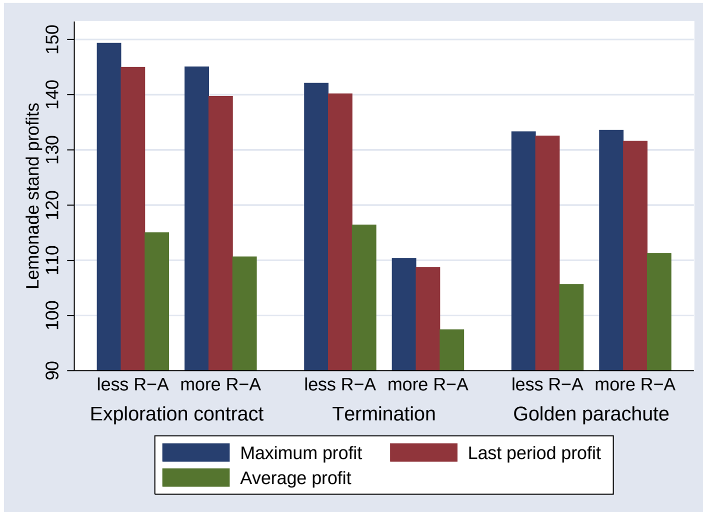

Figure 9: Maximum profit, last period profit and average per period profit of subjects under the exploration contract and termination treatment adjusting for differences in risk aversion.

risk aversion significantly decreases subjects' propensity to explore. In the termination treatment the number of times subjects choose to deviate from the proposed strategy and to explore a location other than the business district in the first 10 periods is significantly lower for subjects who are more risk-averse (p-value 0.0114). Similarly, in the termination treatment the variability of action choices in the first 10 periods is also significantly lower for more risk-averse subjects (p-value 0.0040). There are also small negative effects of risk aversion on exploration activity in the golden parachute treatment, but these effects are never statistically significant.

5 Robustness

In this section we show that our results are robust to modifications in the experimental design. In particular, we address potential signaling effects of incentive contracts. In the analysis we previously conducted each subject only ever saw one particular incentive contract. The subjects were not made aware that a variety of different incentive schemes were administered to different subjects. This means that subjects might make different inferences from the different contracts they are given about what the best strategy to play is. For example, while subjects under the pay-for-performance contract might infer that the best strategy is not to explore, subjects under the exploration contract might infer that the best strategy is to explore.

To account for these potential signaling effects we administered another treatment in

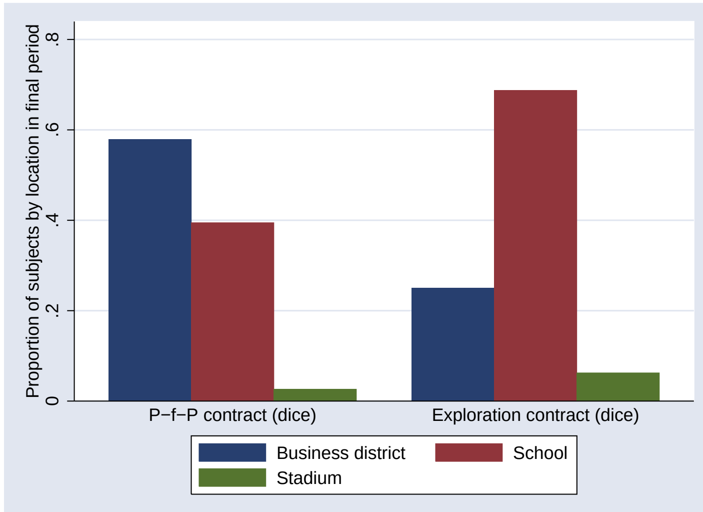

Figure 10: Proportion of subjects by location in the final period of the experiment for the pay-for-performance (dice roll) and exploration (dice roll) contracts.

which subjects were able to see that both pay-for-performance and exploration contracts were available. In this treatment, after having observed the set of possible contracts (pay-for-performance or exploration) the incentive scheme relevant to each subject was determined by having the subject roll a dice. After having observed the outcome of the dice roll the experimenter circled the relevant compensation scheme and crossed out the irrelevant compensation scheme. A total of 70 subjects participated in this treatment of which 32 subjects rolled the dice to receive a pay-for-performance contract and 38 subjects an exploration contract.

Figure 10 confirms our results about the importance of correctly structured incentives for motivating innovation. As before, subjects who are given an exploration contract are significantly more likely (p-value 0.0152) to choose the best location in the final period of the experiment than subjects who receive a pay-for-performance contract. Subjects with an exploration contract also again achieve significantly higher maximum profits (138 francs) and higher final period profits (134 francs) than subjects under a pay-for-performance contract (120 francs, 118 francs). The respective p-values for the comparisons are 0.0372 and 0.0654.

As before this difference in innovation success is driven by the differences in exploration behavior that incentive schemes induce. In particular, the proportion of location choices other than the default location is significantly higher for subjects who obtain an exploration contract following their dice roll (p-value 0.0045) and the variability of strategy choices is also higher, although this difference is not significant (p-value 0.1343) due to the smaller sample size.

Mirroring our previous results, subjects under the pay-for-performance contract also have low average profits although this effect is not statistically significant (p-value 0.1591). Furthermore, risk aversion again has an innovation- and profit-reducing effect in the pay-for-performance treatment. In the pay-for-performance treatment there is a statistically significant negative effect of risk aversion on the correct final period location choice (p-value 0.0583) but there is no significant effect in the exploration contract treatment. The negative effect of risk aversion when subjects obtain a pay-for-performance contract is also apparent in the lower profits for more risk-averse subjects, but this effect is not statistically significant due to the small sample size.

6 Conclusion

In this paper, we argued that appropriately designed incentive schemes are effective in motivating innovation. In a task that involves innovation through experimentation, we find that subjects under an incentive scheme that tolerates early failure and rewards long-term success explore more and are more likely to discover a novel business strategy than subjects under fixed-wage or standard pay-for-performance incentive schemes. We also find that the threat of termination may undermine innovation, and that this effect is mitigated by the presence of a golden parachute.

Several important questions remain unanswered. For example, when agents work in teams, what is the optimal balance between individual and team incentives that motivate exploration? Moreover, when there are different types of agents, how do we design contracts to attract the creative types while keeping shirkers and conventional types away? We leave these questions for future research.

Appendices

A Experimental Instructions

Instructions

You are now taking part in an economic experiment. Please read the following instructions carefully. Everything that you need to know in order to participate in this experiment is explained below. Should you have any difficulties in understanding these instructions please notify us. We will answer your questions at your cubicle.

During the course of the experiment you can earn money. The amount that you earn during the experiment depends on your decisions. All the gains that you make during the course of the experiment will be exchanged into cash at the end of the experiment. The exchange rate will be:

$$100 \text{ francs} = \$1$$

The experiment is divided into 20 periods. In each period you have to make decisions, which you will enter on a computer screen. The decisions you make and the amount of money you earn will not be made known to the other participants - only you will know them.

Please note that communication between participants is strictly prohibited during the experiment. In addition we would like to point out that you may only use the computer functions which are required for the experiment. Communication between participants and unnecessary interference with computers will lead to the exclusion from the experiment. In case you have any questions don't hesitate to ask us.

Experimental Procedures

In this experiment, you will take on the role of an individual running a lemonade stand. There will be 20 periods in which you will have to make decisions on how to run the business. These decisions will involve the location of the stand, the sugar and lemon content and the lemonade color and price. The decisions you make in one period, will be the default choices for the next period.

At the end of each period, you will learn what profits you made during that period. You will also hear some customer reactions that may help you with your choices in the following periods.

Previous Manager Guidelines

Dear X,

I have enclosed the following guidelines that you may find helpful in running your lemonade stand. These guidelines are based on my previous experience running this stand.

When running my business, I followed these basic guidelines:

Location:	Business District
Sugar Content:	3%
Lemon Content:	7%
Lemonade Color:	Green
Price:	8.2 francs

With these choices, I was able to make an average profit of 85 francs per period.

I have experimented with alternative choices of sugar and lemon content, as well as lemonade color and price. The above choices were the ones I found to be the best. I have not experimented with alternative choices of location though. They may require very different strategies.

Regards,

Previous Manager

Compensation

(The following paragraph is used in the instructions for subjects in the treatment with the fixed wage contract.) You will get paid a fixed wage of 50 francs per period during the 20 periods of the experiment. Your final compensation does not depend on your profits from the lemonade stand.

(The following paragraph is used in the instructions for subjects in the treatment with the pay-for-performance contract.) Your compensation will be based on the profits you make with your lemonade stand. You will get paid 50% of your own total lemonade stand profits during the 20 periods of the experiment.

(The following paragraph is used in the instructions for subjects in the treatment with the exploration contract.) Your compensation will be based on the profits you make with your lemonade stand. You will get paid 50% of your own lemonade stand profits in the last 10 periods of the experiment.

(The following paragraph is used in the instructions for subjects in the treatment with the termination contract.) Your compensation will be based on the profits you make with your lemonade stand. You will get paid 50% of the profits you make during the last 10 periods of the experiment. However, if the profits you make during the first 10 periods of the experiment are below 800 francs, the experiment will end early.

(The following paragraph is used in the instructions for subjects in the treatment with the golden parachute contract.) You will get paid 50% of the profits you make during the last 10 periods of the experiment. If the profits you make during the first 10 periods of the experiment are below 800 francs, the experiment will end early and you will receive a payment of 250 francs.

B Experimental Design

B.1 Parameters of the Business Game

The subjects were able to make the following parameter choices:

- Location = {Business District, School, Stadium}
- Sugar Content = $\{0, 0.1, 0.2, ..., 9.9, 10\}$
- Lemon Content = $\{0, 0.1, 0.2, ..., 9.9, 10\}$
- Lemonade Color = {Green, Pink}
- Price = $\{0, 0.1, 0.2, ..., 9.9, 10\}$

The table below shows the optimal product mix in each location.

	Business District	School	Stadium
Sugar	1.5%	9.5%	5.5%
Lemon	7.5%	1.5%	5.5%
Lemonade Color	Green	Pink	Green
Price	7.5	2.5	7.5
Maximum Profit	100	200	60

In order to calculate the profits in each location when the choices are different from the optimal choices above, we implemented a linear penalty function. In each location, the penalty factors associated with a deviation of one unit for each of the variables are given by the next table.

	Business District	School	Stadium
Sugar	5	6	0.5
Lemon	5	6	0.5
Lemonade Color	20	60	0.5
Price	5	6	0.5

B.2 Eliciting Risk Aversion

I measured the subjects' risk aversion by observing choices under uncertainty in an experiment that took place after the business game experiment. As part of this study, the subjects participated in a series of lotteries of the following form.

Lottery A: Win $10 with probability 1/2, or win $2 with probability 1/2. If subjects reject lottery A they receive $7.

Lottery B: Win $10 with probability 1/2, or win $2 with probability 1/2. If subjects reject lottery B they receive $6.

Lottery C: Win $10 with probability 1/2, or win $2 with probability 1/2. If subjects reject lottery C they receive $5.

Lottery D: Win $10 with probability 1/2, or win $2 with probability 1/2. If subjects reject lottery D they receive $4.

Lottery E: Win $10 with probability 1/2, or win $2 with probability 1/2. If subjects reject lottery E they receive $3.

After subjects had made their choices one lottery was chosen at random and each subject was compensated according to his or her choice. The above lotteries enable us to construct individual measures of risk aversion.

We then used the median risk aversion measure to split the sample into a more risk-averse group and a less risk-averse group.

B.3 Eliciting Ambiguity Aversion

We also measured the subjects' ambiguity aversion by observing choices under uncertainty in another experiment that took place after the business game experiment and the risk aversion experiment. As part of this study, we presented the subjects with the opportunity to participate in a series of lotteries of the following form.

If a red ball is chosen you will win $7, if a blue ball is chosen you will win $2.

Case A: Choose Urn 1 containing 20 balls that are either red or blue OR choose Urn 2 containing 16 red balls and 4 blue balls.

Case B: Choose Urn 1 containing 20 balls that are either red or blue OR choose Urn 2 containing 14 red balls and 6 blue balls.

Case C: Choose Urn 1 containing 20 balls that are either red or blue OR choose Urn 2 containing 12 red balls and 8 blue balls.

Case D: Choose Urn 1 containing 20 balls that are either red or blue OR choose Urn 2 containing 10 red balls and 10 blue balls.

Case E: Choose Urn 1 containing 20 balls that are either red or blue OR choose Urn 2 containing 8 red balls and 12 blue balls.

Case F: Choose Urn 1 containing 20 balls that are either red or blue OR choose Urn 2 containing 6 red balls and 14 blue balls.

Case G: Choose Urn 1 containing 20 balls that are either red or blue OR choose Urn 2 containing 4 red balls and 16 blue balls.

After subjects had made their choices one case was chosen at random and the subject was compensated according to his choice. The above lotteries enable us to construct individual measures of ambiguity aversion.

We then used the median ambiguity aversion measure to split the sample into a more ambiguity-averse group and a less ambiguity-averse group.

References

Acharya, Viral V. and Krishnamurthy Subramanian, "Bankruptcy Codes and Innovation," *SSRN eLibrary*, 2007.

Aghion, Philippe, John Van Reenen, and Luigi Zingales, "Innovation and Institutional Ownership," *Working Paper*, 2008.

Amabile, Teresa M., *Creativity in Context*, Boulder, CO: Westview Press, 1996.

Arrow, Kenneth J., "Classificatory Notes on the Production and Transmission of Technological Knowledge," *The American Economic Review*, May 1969, *59* (2), 29–35.

Atanassov, Julian, "Quiet Life or Managerial Myopia: Is the Threat of Hostile Takeovers Beneficial for Technological Innovation?," *SSRN eLibrary*, 2007.

Azoulay, Pierre, Joshua Graff Zivin, and Gustavo Manso, "Incentives and Creativity: Evidence From the Howard Hughes Medical Institute Investigator Program," *Working Paper*, 2007.

Banks, Jeffrey, Mark Olson, and David Porter, "An experimental analysis of the bandit problem," *Economic Theory*, Jun. 1997, *10* (1), 55–77.

Bull, Clive, Andrew Schotter, and Keith Weigelt, "Tournaments and Piece Rates: An Experimental Study," *The Journal of Political Economy*, Feb. 1987, *95* (1), 1–33.

Dickinson, David L., "An Experimental Examination of Labor Supply and Work Intensities," *Journal of Labor Economics*, 10 1999, *17* (4), 638–670.

Fahr, Rene and Bernd Irlenbusch, "Fairness as a constraint on trust in reciprocity: earned property rights in a reciprocal exchange experiment," *Economics Letters*, 3 2000, *66* (3), 275–282.

Falk, Armin and Andrea Ichino, "Clean Evidence on Peer Effects," *Journal of Labor Economics*, 2006, *24* (1), 39–57.

Fehr, Ernst and Bettina Rockenbach, "Detrimental effects of sanctions on human altruism," *Nature*, 2003, *422*, 137–140.

———**, Simon Gachter, and Georg Kirchsteiger**, "Reciprocity as a Contract Enforcement Device: Experimental Evidence," *Econometrica*, Jul. 1997, *65* (4), 833–860.

Fischbacher, Urs, "z-Tree: Zurich toolbox for ready-made economic experiments," *Experimental Economics*, 2007, *10* (2), 171–178.

Gneezy, Uri and Aldo Rustichini, "Pay Enough or Don't Pay at All," *The Quarterly Journal of Economics*, Aug. 2000, *115* (3), 791–810.

Holmstrom, Bengt and Paul Milgrom, "Aggregation and Linearity in the Provision of Intertemporal Incentives," *Econometrica*, Mar. 1987, *55* (2), 303–328.

Kohn, Alfie, *Punished by rewards: the trouble with gold stars, incentive plans, A's, praise, and other bribes*, Boston: Houghton Mifflin Co., 1993.

Lazear, Edward P., "Performance Pay and Productivity," *The American Economic Review*, Dec. 2000, *90* (5), 1346–1361.

Manso, Gustavo, "Motivating Innovation," *Working Paper*, 2007.

March, James G., "Exploration and Exploitation in Organizational Learning," *Organization Science*, 1991, *2* (1, Special Issue: Organizational Learning: Papers in Honor of (and by) James G. March), 71–87.

McCullers, John C., *Issues in Learning and Motivation* The Hidden Costs of Reward, Hillsdale, NJ: Lawrence Erlbaum Associates, 1978.

McGraw, Kenneth O., *The Detrimental Effects of Reward on Performance: A Literature Review and a Prediction Model* The Hidden Costs of Reward, Hillsdale, NJ: Lawrence Erlbaum Associates, 1978.

Merlo, Antonio and Andrew Schotter, "A Surprise-Quiz View of Learning in Economic Experiments," *Games and Economic Behavior*, 7 1999, *28* (1), 25–54.

Meyer, Robert J. and Yong Shi, "Sequential Choice under Ambiguity: Intuitive Solutions to the Armed-Bandit Problem," *Management Science*, May 1995, *41* (5), 817–834.

Nalbantian, Haig R. and Andrew Schotter, "Productivity Under Group Incentives: An Experimental Study," *The American Economic Review*, Jun. 1997, *87* (3), 314–341.

Rogerson, William P., "Repeated Moral Hazard," *Econometrica*, Jan. 1985, *53* (1), 69–76.

Sannikov, Yuliy, "A Continuous-Time Version of the Principal-Agent Problem," *The Review of Economic Studies*, forthcoming.

Sapra, Haresh, Ajay Subramanian, and Krishnamurthy Subramanian, "Corporate Governance and Innovation: Theory and Evidence," *SSRN eLibrary*, 2008.

Sauermann, Henry and Wesley M. Cohen, "What makes them tick? Employee Motives and Industrial Innovation," *Working Paper*, 2008.

Shearer, Bruce, "Piece Rates, Fixed Wages and Incentives: Evidence from a Field Experiment," *The Review of Economic Studies*, Apr. 2004, *71* (2), 513–534.

Sillamaa, Mary-Anne, "How work effort responds to wage taxation: An experimental test of a zero top marginal tax rate," *Journal of Public Economics*, 7 1999, *73* (1), 125–134.

van Dijk, Frans, Joep Sonnemans, and Frans van Winden, "Incentive systems in a real effort experiment," *European Economic Review*, 2 2001, *45* (2), 187–214.

Weitzman, Martin L., "Optimal Search for the Best Alternative," *Econometrica*, May 1979, *47* (3), 641–654.