Is Pay-for-Performance Detrimental to Innovation?*

Florian Ederer[†] and Gustavo Manso[‡]

November 15, 2011

Abstract

Previous research in economics shows that compensation plans based on the pay-for-performance principle are effective in inducing higher levels of effort and productivity. On the other hand, research in psychology argues that performance-based financial incentives inhibit creativity and innovation. How should managerial compensation be structured if the goal is to induce managers to pursue more innovative business strategies? In a controlled laboratory setting, we provide evidence that the combination of tolerance for early failure and reward for long-term success is successful 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.

*We would like to thank Dan Ariely, Nittai Bergman, Bruno Biais, Arthur Campbell, Ernst Fehr, Bob Gibbons, Lorenz Götte, Bengt Holmstrom, Jonathan Levin, Steven Lippman, Robert Marquez, Muriel Niederle, Enrico Perotti, Thomas Philippon, Sebastien Pouget, Drazen Prelec, Panle Jia, Jason Snyder, David Robinson, Andrei Shleifer, Johannes Spinnewijn, Antoinette Schoar, Chloe Tergiman, Jean Tirole, Eric Van den Steen, Christian Zehnder and seminar participants at Berkeley Haas, Chicago GSB, Gerzensee, LSE, Stanford, Toulouse, UCLA, University of Munich, the Harvard-MIT Organizational Economics Seminar, the Kauffman Summer Legal Institute, the RFS-EFIC, 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 and the Ewing Marion Kaufmann Foundation is gratefully acknowledged.

[†]Anderson School of Management, University of California, Los Angeles, 110 Westwood Plaza Suite D515, Los Angeles, CA 90095, ederer@ucla.edu, (310) 825-7348.

[‡]MIT Sloan School of Management, 100 Main Street E62-635, Cambridge, MA 02142-1347, manso@mit.edu, (617) 253-7218.

1 Introduction

Previous research on the effectiveness of incentive pay argues that paying an 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, ranging from quasi-experiments in empirical data to economic field experiments and laboratory 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 can 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 at all to motivate agents.

In view of this seemingly conflicting evidence, can *correctly structured* performance-based financial incentives motivate innovation in creative tasks? In a controlled experimental setting in which subjects perform two tasks that involves a trade-off between exploration and exploitation, we provide evidence that incentive plans 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. We also find evidence that the threat of termination can undermine incentives for innovation while golden parachutes can alleviate these innovation-reducing effects. Finally, we show that risk-aversion is an important factor in explaining these results. Our results partially support, but also stand in contrast to previous research in psychology which suggests that financial incentives inhibit innovation (Kohn 1993, Amabile 1996). Our findings also qualify some of the predictions of 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.[1]

[1]See, for example, Rogerson (1985), Holmstrom and Milgrom (1987) and Sannikov (2008). These models focus on situations in which principals need to induce higher effort levels or to avoid diversion of

Our results have implications for various situations in which enabling entrepreneurship is an important concern. In fact, "stimulating innovation, creativity and enabling entrepreneurship" is a top priority for management and widely regarded as the "greatest human resource challenge" facing organizations according to CEO surveys.[2] These situations range from the design of compensation packages of top executives and middle managers in large corporations to structuring compensation for entrepreneurs in start-up companies. For example, some of the most innovation- and creativity-driven firms including pharmaceutical and advertising companies have adopted reward systems that tolerate or even reward failure of employees.[3] In the context of executive compensation tolerance for early failure can be achieved through commonly used practices such as managerial entrenchment, golden parachutes, or option repricing. These practices are often criticized because they protect or even reward the manager after poor performance and, therefore, can induce the manager to shirk or divert funds from the corporation. There have been several proposals to restrict the use of some of these practices.[4] Our results suggest that when combined with appropriate long-term incentives, these practices can be effective in motivating innovation. Regulations that restrict their use could thus have an adverse effect on innovation.

Innovation is often defined as 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. We conduct two economic laboratory experiments in which subjects face this tension between exploration and exploitation. In our first experiment, subjects perform a computerized decision-making task: they must decide whether to prospect gold in a well-known mountain or in a unexplored mountain. In our second experiment, which constitutes the main part of this paper, subjects control the operations of a computerized lemonade stand and must choose between fine-tuning the product choice decisions given to them by the previous manager or choosing a different location and radically altering the product mix to discover a better strategy.

To study the impact of different incentive schemes on productivity and innovation

funds by managers and do not incorporate learning from experimentation, which is a central component of the innovation process and also of the tasks used in our experiments.

[2]See, for example, "CEO Challenge 2004: Perspectives and Analysis," The Conference Board, Report 1353, which reports that the highest response (31%) among surveyed CEOs is that stimulating innovation is of the greatest concern to their company.

[3]For a discussion of this issue in the popular press see, for example, "Better Ideas Through Failure: Companies Reward Employee Mistakes To Spur Innovation, Get Back Their Edge", Wall Street Journal, September 27, 2011.

[4]See, for example, Bebchuk and Fried (2004) and "Rewards for Failure," *British DTI consultation*, June 2003.

in the above tasks, we consider three different treatment groups. The only difference between 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 a fixed percentage of the profits produced during the experiment.[5] Subjects in the third treatment group are allocated a contract that is tailored to motivate exploration. Their compensation is a fixed percentage of the profits produced during the second half of each of the two experiments.

Our main hypothesis is that subjects under the exploration contract explore more and are more likely to find a superior strategy than subjects under the fixed-wage or 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 half of the experiment while they explore new strategies. Second, the prospect of pay for performance later on encourages subjects to learn better ways of performing the task. Therefore, even though the exploration contract combines elements of the fixed-wage and pay-for-performance contracts, it is plausible that performance under the exploration contract is superior to performance under both the fixed-wage and the pay-for-performance contracts.

Our results strongly support the main hypothesis stated above. In our first experiment, the "gold-prospecting experiment", subjects given the exploration contract are significantly more likely to choose the unexplored mountain than under either the pay-for-performance or the fixed-wage contracts. In our base case specification, 24 out of 30 subjects explored the new mountain under the exploration contract while only 9 out of 30 subjects and 3 out of 30 subjects explored the new mountain under the pay-for-performance and fixed-wage contracts, respectively. Subjects under the pay-for-performance contract tended to prospect gold in the well-known mountain, while subjects under the fixed wage contract tended to shirk.

Results are similar in our main experiment, the "lemonade stand experiment", where a total of 379 subjects perform a real-effort task. We find that subjects given the exploration contract end the experiment in the best location 80% of the time, while subjects given 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 examine the reasons behind the relatively poor performance of subjects

[5]This exploration contract is based on Manso (2011) who shows that the combination of tolerance for early failure and reward for long-term success is optimal to motivate innovation. The exploration contract can be implemented in practice through stock grants that increase with the tenure of the manager in the firm independently of the manager's performance. Alternatively, it can be implemented via stock option grants with long vesting periods, and option repricing in case the manager performs poorly early on.

under the fixed wage and pay-for-performance contracts. Although subjects given 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; when facing 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. In both experiments, a principal would achieve higher profits under the exploration contract than under the pay-for-performance or fixed-wage contracts. The fixed-wage contract induces shirking while the pay-for-performance contract induces little exploration which in turn leads to lower profits. In the gold-prospecting experiment, if the principal were to try to induce exploration using a standard pay-for-performance contract, he would have to give 27% instead of 10% of the firm to a risk-neutral subject, reducing his profits by approximately 20% relative to the profits obtained under the exploration contract.

Different attitudes towards risk can also affect the outcome under the different contracts. In the lemonade stand experiment, 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.

To study the effects of termination on innovation and performance, we introduce two new treatment groups in the lemonade stand experiment: 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: only 45% of the subjects in the termination treatment group find the optimal location whereas 65% of the subjects in the termination with golden parachute treatment group find the optimal location.

1.1 Related Literature

Several recent papers study the effects of institutional features such as the legal environment and the degree of competition on innovation. For example, Acharya and Subramanian (2009) investigate whether debtor-friendly bankruptcy laws foster innovation and Acharya, Baghai-Wadji, and Subramanian (2009) show that stringent labor laws that restrict the dismissal of employees encourage innovation. Similarly, Sapra, Subramanian, and Subramanian (2008), Atanassov (2008), and Chemmanur and Tian (2011) study whether takeover pressure affects investment in innovation, while Liu and Wong (2011) analyze the relation between intellectual capital (as measured by patents) and leverage. Aghion, Reenen, and Zingales (2008) find that the failure-tolerant attitude of institutional investors has a positive effect on innovation that is complementary to the degree of product market competition. A burgeoning literature investigates the interplay between non-compete agreements and innovation. Gilson (1999) argues that Silicon Valley's entrepreneurial growth can be attributed to California's proscription of non-competes and high employee interfirm mobility. While Almeida and Kogut (1999) document a high mobility of California inventors and Sorenson and Stuart (2003) show that more startups appear in regions that do not enforce non-compete agreements, causal evidence for these assertions remains thin (Fallick, Fleischman, and Rebitzer 2006, Marx, Strumsky, and Fleming 2009).

Our work is most closely related to contributions that investigate the relation between explicit incentive structures and innovation.[6] Long-term incentives for the heads of research and development departments (Lerner and Wulf 2007) as well as long-term incentives and golden parachutes for CEOs (Francis, Hasan, and Sharma 2010) are associated with more (and more heavily cited) patents, while short-term incentives are unrelated to measures of innovation. Using compensation data from 237 firms in the high-technology industry, Yanadori and Marler (2006) find that greater emphasis on innovation activities is positively associated with a greater reliance on long-term incentives and longer stock option vesting period lengths. Tolerance for early failure is shown to be particularly important in entrepreneurial settings: Tian and Wang (2010) and Chemmanur, Tian, and Loutskina (2011) find that firms backed by venture capitalists who tolerate failures are significantly more innovative. Finally, Azoulay, Zivin, and Manso (2011) address whether funding policies with tolerance for early failure and long hori-

[6]Related theoretical contributions that also focus on the interplay of incentives and innovation include Aghion and Tirole (1994), Hellmann (2007) and Hellmann and Thiele (2011).

zons to evaluate results motivate creativity in scientific research. These papers provide support for the thesis that tolerance for early failure coupled with reward for long-term success motivates innovation. However, because they use naturally occurring data, they either restrict themselves to merely showing correlation patterns, or they are subject to the criticism that the variation in the incentive schemes may not be completely exogenous. In our paper, we are able to study the effects of incentives on innovation by exogenously varying compensation schemes in a controlled laboratory environment.

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, Katok and Siemsen 2011). Meyer and Shi (1995) and Banks, Olson, and Porter (1997) investigate the tension between exploitation and exploration in an experimental setting using monetary efforts and report results in broad accordance with the theoretical bandit model. Neither of these papers considers the effect of different incentive schemes on subject behavior.[7]

More recently, experimental researchers have conducted studies in both the lab and the field 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), stuffing letters into envelopes (Falk and Ichino 2006), or picking fruit (Bandiera, Barankay, and Rasul 2005, Bandiera, Barankay, and Rasul 2009). These tasks, however, are inadequate to study incentives for innovation. In this paper, we introduce two new tasks to study the incentives for exploration. In the gold-prospecting task, subjects are asked to make choices under a given cost function. In the lemonade stand experiment, subjects exert real effort and have to form beliefs about the outcomes associated with each of the available actions. While the gold-prospecting task allows us to fully control the expectations of subjects in a parametrized setting, the lemonade stand experiment allows us to address more comprehensively the effects of incentive schemes on exploration and innovation behavior.

2 Gold Prospecting Experiment

The first experimental setting allows us to measure the effects of different incentive schemes on subject choice between several payoff options. In particular, we are able to

[7] Other related experimental papers using search problems akin to our lemonade stand experiment and dealing with innovation modeled as a computationally complex problem include Gabaix, Laibson, Moloche, and Weinberg (2006) and Meloso, Copic, and Bossaerts (2009).

directly test the key predictions of Manso (2011) in a controlled laboratory setting under different parametric assumption. The results also guide our subsequent investigation of subject behavior in a real effort task setting.

2.1 Procedures and Subject Pool

The first set of experiments was programmed and conducted with the software z-Tree (Fischbacher 2007) at the California Social Science Experimental Laboratory (CASSEL) at UCLA. Participants were recruited from the CASSEL subject pool using an online recruitment system. A total of 30 subjects participated in these experiments. We employed a within-subjects design so that all subjects experienced all three treatment conditions.

After subjects completed the experiment, we elicited their degree of risk aversion and ambiguity aversion. We describe the exact procedures, which are standard, in the appendix. Subjects were then privately paid. The experimental session lasted 50 minutes.

During the experiment, experimental currency units called "gold nuggets" were used to keep track of monetary earnings. The exchange rate was set at 3 gold nuggets (gn) = \$1, and the show-up fee was \$10. On average, subjects earned \$33.

2.2 The Task

Subjects play several separate instances of a gold prospecting game that exactly mirrors the theoretical model in Manso (2011). Each gold prospecting game consists of 2 periods. In each period, each subject chooses to prospect gold in one of two mountains or to stay at home. Each of these options has an associated revenue and cost. The option "Stay at home" always yields 0gn and it costs the subject 0gn to choose this option. In contrast, the option "Mountain 1" yields 100gn with a probability of 50% and 0gn with a probability of 50% and it costs 0.25gn to choose this option. Each payoff realization of "Mountain 1" is an independent draw. Finally, the option "Mountain 2" allows for learning. With probability p, "Mountain 2" is a gold-rich mountain, in which case it yields 100gn in both period 1 and period 2 whenever this option is chosen. With a probability of $1-p$, "Mountain 2" is a gold-poor mountain, in which case it always yields 0gn whenever this option is chosen. The cost of choosing "Mountain 2" is 0.75gn. Note that each instance of the gold prospecting game is an independent draw. Thus, each subject can only learn about the quality of "Mountain 2" for that particular instance of the gold prospecting game.

The participants in the experiment thus face the choice between an outside option of 0gn ("Stay at home"), an option with a known and constant payoff distribution

("Mountain 1") and an option with an unknown payoff distribution ("Mountain 2") that allows for learning.

We elicit subject choices using the strategy method: subjects are asked to provide a full contingent plan of action for all possible outcomes. Thus, on the input screen for each gold prospecting game subjects enter (a) their choice for period 1, (b) their choice for period 2 conditional on a payoff of 0gn in period 1, and (c) their choice for period 2 conditional a payoff of 100gn in period 1.

At the end of each gold prospecting game, subjects learn the revenue they obtained and the cost they incurred during that game. In addition, at the end of each gold prospecting game subjects are informed about their cumulative payoff balance.

2.3 Treatment Groups and Predictions

We conduct four sets of treatments which differ in terms of the probability p that Mountain 2 is a gold-rich mountain. The probability p can take on the following values: 5%, 35%, 45% or it is unknown. Note that when p is unknown to subjects we choose it such that p is equal to 50%.

In each of the four treatment sets, each subject plays the gold prospecting game under three different compensation schemes. Thus, there are a total of twelve separate gold prospecting games which each subject plays.

In particular, the three compensation schemes are:

Incentive Scheme 1 (Fixed Wage): "You will be paid a fixed wage of 1gn per period."

Incentive Scheme 2 (Pay-for-Performance): "You will be paid 10% of the gold revenue during the 1st and the 2nd period of the experiment."

Incentive Scheme 3 (Exploration): "You will be paid 10% of the gold revenue in the 2nd period of the experiment."

The first two incentive systems are motivated by previous research that studies the effect of pay-for-performance sensitivity on productivity. The third system is motivated by previous theoretical research (Manso 2011) who argues that tolerance for early failure and reward for long-term success is optimal to motivate innovation and exploration. Note that under this exploration contract, subjects who 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.

Table 1 shows the optimal choices for a risk-neutral agent in each treatment condition. Along the columns there are different values of p while the different incentive schemes are shown in the different rows.

	Optimal Strategies			
	$p = 0.05$	$p = 0.35$	$p = 0.45$	p unknown
Fixed Wage				
Period 1	Stay at Home	Stay at Home	Stay at Home	n/a
Period 2 after 100gn	Stay at Home	Stay at Home	Stay at Home	n/a
Period 2 after 0gn	Stay at Home	Stay at Home	Stay at Home	n/a
Pay-for-Performance				
Period 1	Mountain 1	Mountain 1	Mountain 2	n/a
Period 2 after 100gn	Mountain 1	Mountain 1	Mountain 2	n/a
Period 2 after 0gn	Mountain 1	Mountain 1	Mountain 1	n/a
Exploration				
Period 1	Stay at Home	Mountain 2	Mountain 2	n/a
Period 2 after 100gn	Mountain 1	Mountain 2	Mountain 2	n/a
Period 2 after 0gn	Mountain 1	Mountain 1	Mountain 1	n/a

Table 1: Optimal strategies for a risk-neutral agent in each treatment condition.

As is apparent from Table 1 as p increases the choice of "Mountain 2" becomes more attractive and hence it is the optimal choice of action in the first period for high values of p in the Pay-for-Performance and Exploration contract treatments. Our main hypothesis, however, concerns the extent to which the different payment schemes encourage the choice of "Mountain 2" in period 1 of the gold prospecting game for a given level of p. In general, due to the tolerance for early failure and reward for long-term success, subjects under the exploration contract should be more likely to choose "Mountain 2" in period 1 than subjects under the fixed wage and the pay-for-performance contracts. The comparison is most clearly visible when $p = 0.35$: in period 1 subjects should choose "Stay at Home" under the fixed wage contract, "Mountain 1" under the pay-for-performance contract and "Mountain 2" under the exploration contract.

2.4 Results

We first compare strategy choices across the three incentive treatments for the 4 different levels of the success probability p. There were 30 subjects who participated in the experiment. Each subject made choices under all the 12 different treatments which allows us to make within-subject comparisons.

We focus on the exploration behavior of subjects across the different treatment conditions. Table 2 shows the first-period choices made by subjects in the different conditions as well as the associated theoretical prediction. Consistent with the prediction that the

	Choices in Period 1			
	$p = 0.05$	$p = 0.35$	$p = 0.45$	p unknown
Fixed Wage				
Stay at Home	23	25	24	26
Mountain 1	7	2	3	3
Mountain 2	0	3	3	1
Theoretical Prediction	Stay at Home	Stay at Home	Stay at Home	n/a
Pay-for-Performance				
Stay at Home	0	1	1	1
Mountain 1	26	20	11	18
Mountain 2	4	9	18	11
Theoretical Prediction	Mountain 1	Mountain 1	Mountain 2	n/a
Exploration				
Stay at Home	17	1	2	4
Mountain 1	11	5	5	6
Mountain 2	2	24	23	20
Theoretical Prediction	Stay at Home	Mountain 2	Mountain 2	n/a

Table 2: Choices in Period 1 for each treatment condition.

exploration contract encourages exploration, more subjects choose “Mountain 2” under this incentive scheme than under the other two contracts. Moreover, subjects under the fixed wage contract overwhelmingly choose the shirking option “Stay at Home”, while subjects under the pay-for-performance contract choose to exploit “Mountain 1.”[8]

The most important comparison is for $p = 0.35$. In Tables 1 and 2 we show that for $p = 0.35$ the predicted choice of action for a risk-neutral agent in the first period is different between the pay-for-performance and the exploration contract while for $p = 0.45$ the optimal choices for a risk-neutral agent under the two incentive systems coincide. Finally, for $p = 0.05$ a risk-neutral agent under the exploration contract is expected to choose “Stay at Home” since the payoff from the exploration option “Mountain 2” is too low and subjects do not earn enough in the first period to justify the costs of choosing “Mountain 2”.

When $p = 0.35$ only 9 subjects chose “Mountain 2” in period 1 under the pay-for-performance contract, but 24 subjects chose to do so under the exploration contract. Due to our within-subject design we observe the choices of subjects under all the different

[8]As a result of this increased exploration the expected probability of finding 100gn in period 2 given the subjects’ strategies is 60% under the exploration contract, yet only 52.9% and 8.2% under the pay-for-performance and fixed-wage contracts.

treatment conditions. A McNemar matched samples test shows that the difference in the choice of "Mountain 2" for $p = 0.35$ under the pay-for-performance and the exploration contract is highly statistically significant (p-value 0.0007). The statistical significance of this difference is also confirmed by a subject fixed-effects logit regression where we compare the first-period choices for $p = 0.35$ under the pay-for-performance and the exploration contract. The binary dependent variable takes the value 1 if a subject chose "Mountain 2" in period 1 and value 0 otherwise and the dependent variable takes the value 1 if the treatment was the exploration contract and 0 if it was pay-for-performance contract. The associated p-value of the coefficient of the dependent variable is 0.004 indicating again that the difference is highly statistically significant. As is obvious from the raw data shown in Table 2, the difference in behavior between the exploration and the fixed wage contract is also statistically significant. Under the fixed wage contract subjects overwhelmingly choose the shirking option "Stay at Home". The relevant p-values for the McNemar test and the coefficient of the subject fixed-effects logit regression are 0.00001 and 0.001.

Although we are not able to provide precise predictions for subject behavior in the treatment conditions where p is unknown, this set of treatment conditions is of particular interest as it shares the feature of the unknown success probability with the more realistic experiment we analyze in the next section. Table 2 shows that subject behavior in the first period for unknown p is similar to when $p = 0.35$ and formal statistical tests confirm this impression. Subjects are significantly more likely to choose "Mountain 2" in the first period under exploration contract than under the pay-for-performance incentive scheme (p-value 0.032, fixed-effects logit). Thus, even when the success probability p is unknown the early failure tolerance of the exploration contract relative to the pay-for-performance contract motivates subjects to choose the exploration option "Mountain 2".

We also investigate how large the expected profit of a principal would be who implements these different incentive plans given the behavior of the subjects in the experiment. Table 3 reports the expected profits of a principal both for the case of optimal strategy choices by agents as well as for the actual strategy choices they made in the experiment.

Table 3 shows that for $p = 0.05$ the profits of the principal are highest when he offers a pay-for-performance contract. For higher expected success probabilities of the exploratory option such as $p = 0.35$ and $p = 0.45$ the principal reaps the highest profit when he offers an exploration contract. Note further that when the value of p is unknown to the subjects, the principal does best when he offers an exploration contract (under the assumption that the actual p is 50%).

How large would the principal's profits be if the principal wanted to implement exploration with a standard pay-for-performance rather than an exploration contract?

	Principal's Profits			
	$p = 0.05$	$p = 0.35$	$p = 0.45$	p unknown
Fixed Wage				
Optimal Strategy	-2	-2	-2	n/a
Actual Strategy	17.25	13.5	16.92	12.17
Pay-for-Performance				
Optimal Strategy	90	90	105.75	n/a
Actual Strategy	75.53	87.23	94.2	96
Exploration				
Optimal Strategy	45	95.75	110.25	n/a
Actual Strategy	59.47	91.23	103.13	103.33

Table 3: Expected profits of the Principal under the different treatment conditions for optimal and actual agent strategy choices.

When $p = 0.35$, for the principal to induce risk-neutral subjects to explore under a standard pay-for-performance contract, he would need to offer subjects 27% of the total gold revenue instead of 10%. This would drive down the principal's expected profits from 90 to 74.83.

3 Lemonade Stand Experiment

The insights from the gold prospecting experiments provide broad support for the hypothesis that contracts that tolerate early failure and reward long-term success are successful in inducing exploration. However, our findings may be specific to the particular task at hand that is closely related to previous theoretical work.

Naturally, this leads us to investigate whether the same conclusions hold in a more realistic setting. For this purpose we conduct experiments in which participants under different incentive schemes have to solve a real effort task in which the trade-off between exploration and exploitation again takes center stage.

3.1 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 CLER subject pool using an online recruitment system. A total of 379 subjects participated in our experiments.

After subjects completed the experiment we elicited their degree of risk aversion and ambiguity aversion. We describe the exact procedures in the appendix. Subjects were then privately paid. A session lasted, on average, 60 minutes.

Experimental currency units called francs were used. The exchange rate was set at 100 francs = \$1 and the show-up fee was \$10. Subjects on average earned \$24.

3.2 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 provided 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.[9] 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 is a letter from the previous manager which is reproduced in the appendix. The letter gives hints to the subjects about a strategy 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 with 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 manager 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

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

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.

3.3 Treatment Groups and Predictions

We initially implement three treatment conditions that are similar to the conditions analyzed in the gold-prospecting environment 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 three incentive schemes are as follows:

Incentive Scheme 1 (Fixed Wage): "You will be paid a fixed wage of 50 francs per period."

Incentive Scheme 2 (Pay-for-Performance): "You will be paid 50% of the profits you make during the 20 periods of the experiment."

Incentive Scheme 3 (Exploration): "You will be paid 50% of the profits you make during the last 10 periods of the experiment."

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 choose a strategy that is more similar to the optimal business strategy than subjects under the fixed-wage and pay-for-performance contracts.*

Two further sub-hypotheses 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.*

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.

3.4 Results

In this subsection 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.

3.4.1 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 (in terms of maximum and last-period profit) 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).[10] Finally, note that the difference in performance between the fixed wage and the pay-for-performance treatment is in line with the negative effects of performance pay found in the psychology literature.

We also examine how close subjects come to finding the optimal strategy over the course of the experiment in terms of the profits they achieve. 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 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

[10] 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).

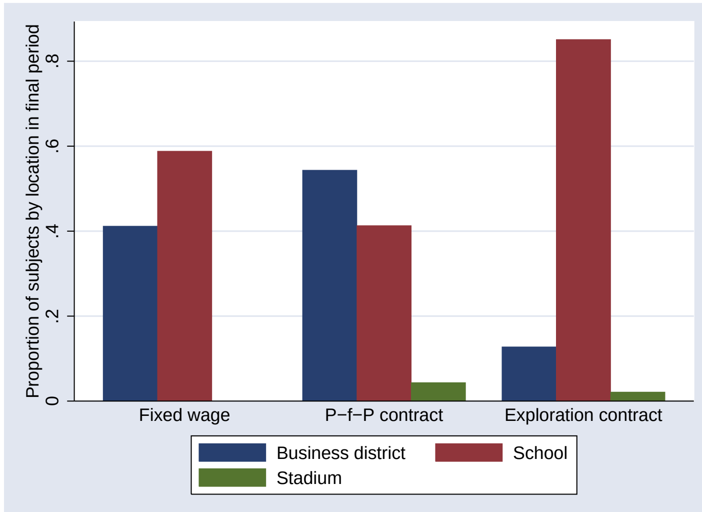

Figure 1: Proportion of subjects by location in the final period of the experiment for the fixed-wage, pay-for-performance and exploration contracts.

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. 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 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

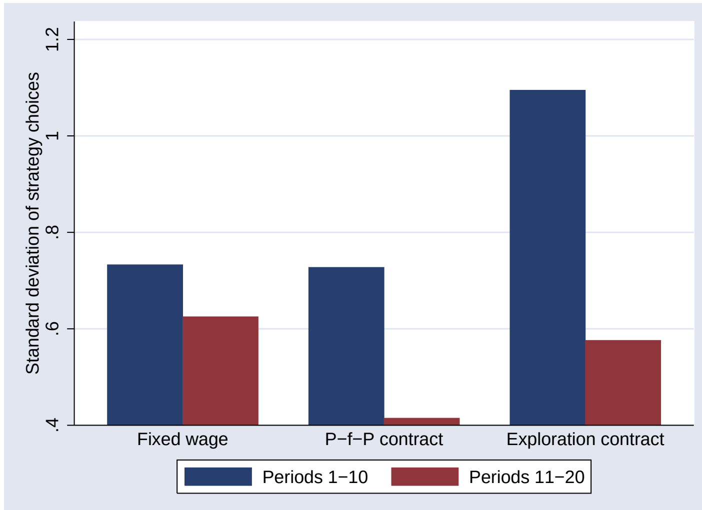

Figure 2: Average subject-specific standard deviation of strategy choices for the three continuous variables (sugar content, lemon content, price) in periods 1-10 and 11-20 of the experiment for the fixed-wage, pay-for-performance and exploration contracts.

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.

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. 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 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.[11] As can be seen from column 1 of Table 4, the hazard rate

[11] 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

	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 4: 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 ***.

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 4 shows that the estimates for the first 10 periods are qualitatively similar.

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

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.

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.

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 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 (2008) 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

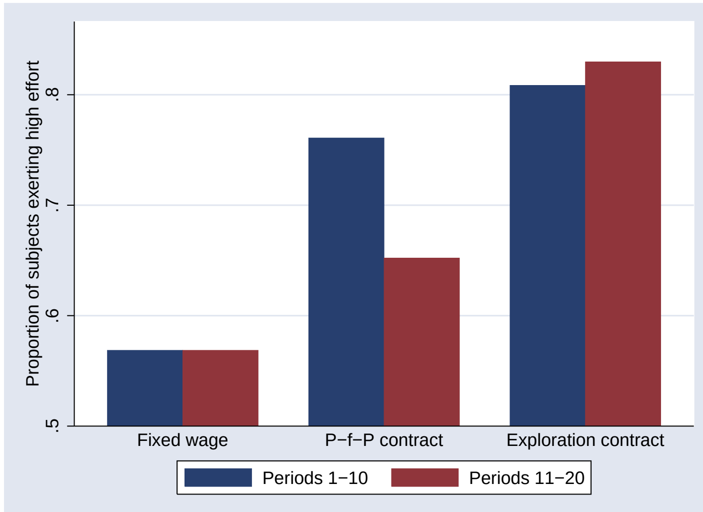

Figure 3: Proportion of subjects who complete more than half of the fields in the decision record table for the fixed-wage, pay-for-performance and exploration contracts.

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). 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 3, 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 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).[12]

3.4.2 Average Performance

Having confirmed that the innovation success, exploration behavior, time allocation and effort choice across the different contracts are in accordance with our theoretical predictions, we now 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

[12] For a study of the effect of intrinsic motivation on innovation productivity, see, for example, Sauermann and Cohen (2010).

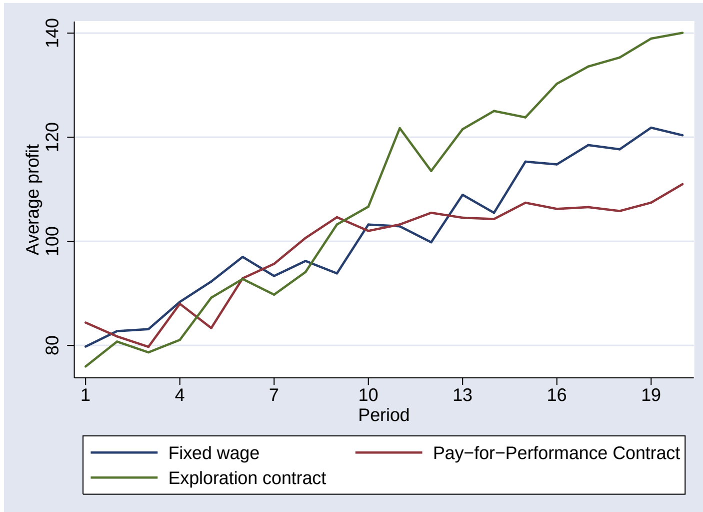

Figure 4: Average per period profits over the course of the experiment for subjects under the fixed-wage, pay-for-performance and exploration contracts.

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 2 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.

We now investigate whether attitudes toward risk can explain the differences in performance documented in Result 4.

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.*

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

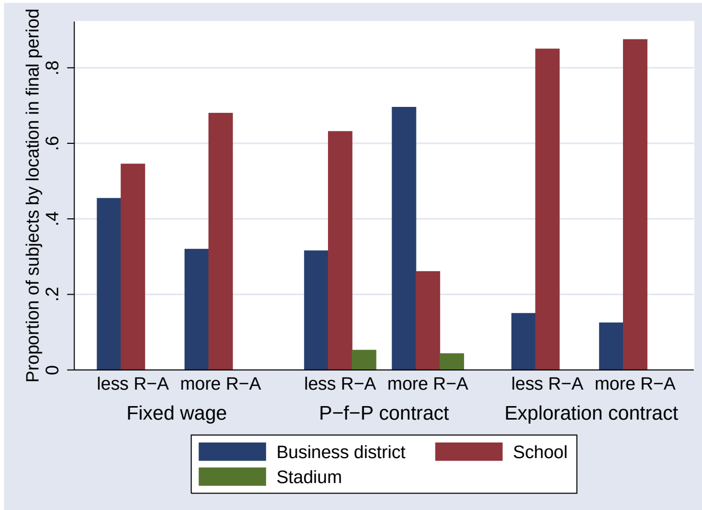

Figure 5: Proportion of subjects by location in the final period of the experiment for the fixed-wage, pay-for-performance and exploration contracts adjusting for differences in risk aversion.

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 the effect is smaller in magnitude and not statistically significant. The same is true in the fixed-wage contract where compensation entails no risk.[13]

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

[13]Qualitatively similar results hold for our 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.

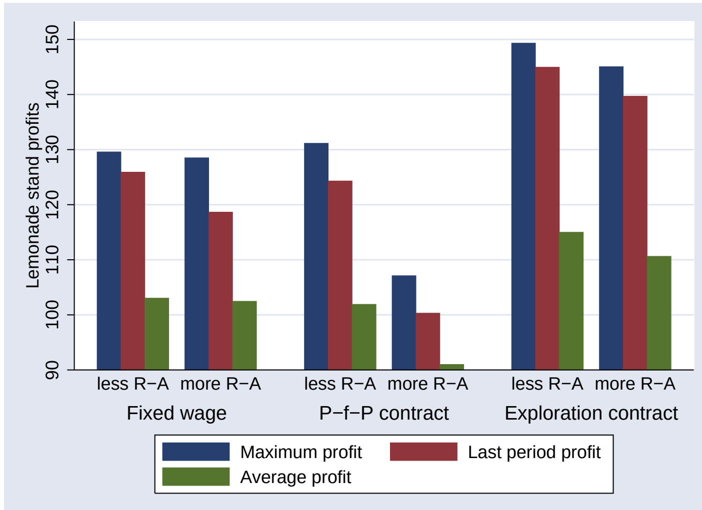

Figure 6: Maximum profit, last period profit and average per period profit of subjects under the fixed-wage, pay-for-performance and exploration contract adjusting for differences in risk aversion.

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.[14] 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.

3.5 Termination

We now turn to investigating how the threat of early termination influences exploration behavior and performance. Early termination can undermine the exploration behavior

[14]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.

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 introduce two new treatment groups that enable us to investigate the effects of termination and golden parachutes.

Incentive Scheme 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."

Incentive Scheme 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 does offer sufficient tolerance for early failure.[15] 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.

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.*

We show that the threat of termination reduces the probability that subjects successfully innovate because the threat of early termination reduces exploration activities. Furthermore, 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.*

[15]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.

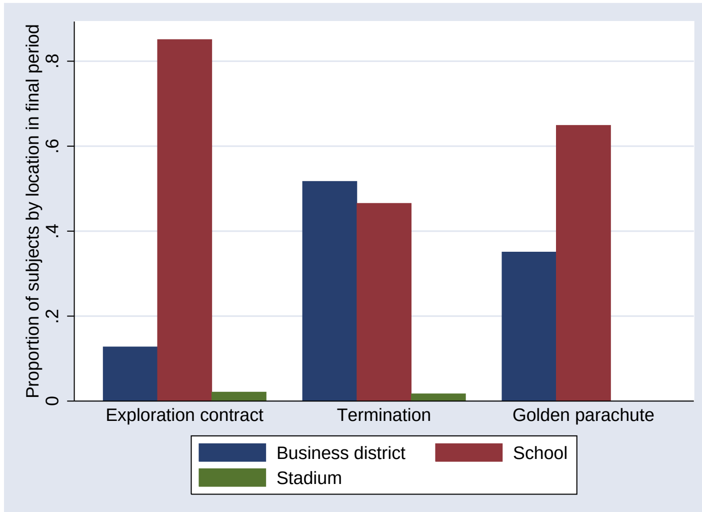

Figure 7: Proportion of subjects by location in the final period of the experiment for the exploration contract, termination and golden parachute treatments.

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 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 subject to termination 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 for the differences 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, 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).[16]

[16] 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,

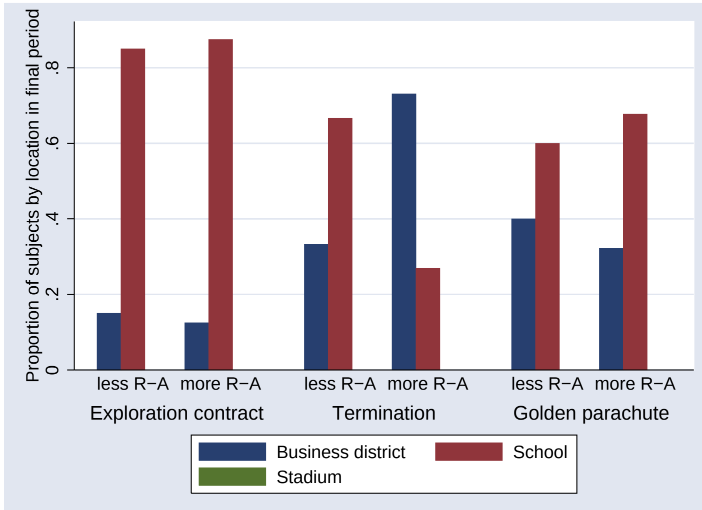

Figure 8: Proportion of subjects by location in the final period of the experiment for the exploration contract, termination and golden parachute treatment adjusting for differences in risk aversion.

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 4 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. 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 4 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),

21 out of 78 subjects (27%) are terminated in the golden parachute treatment, but the difference is not statistically significant (p-value 0.2124).

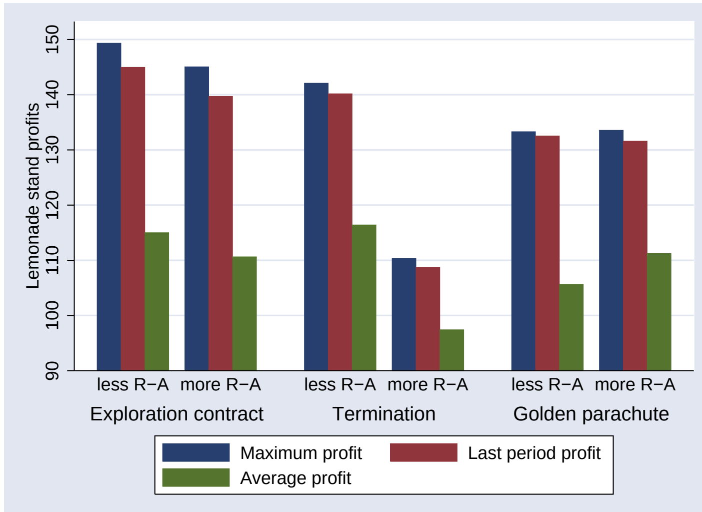

Figure 9: Maximum profit, last period profit and average per period profit of subjects under the exploration contract and termination treatment adjusting for differences in risk aversion.

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 for our finding for the exploration contract treatment, there is no statistically significant effect of risk aversion in the golden parachute treatment.

Finally, in the termination treatment a high degree of 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.

3.6 Robustness

In this subsection, we show that our results are robust to modifications in the experimental design and 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 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. 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.

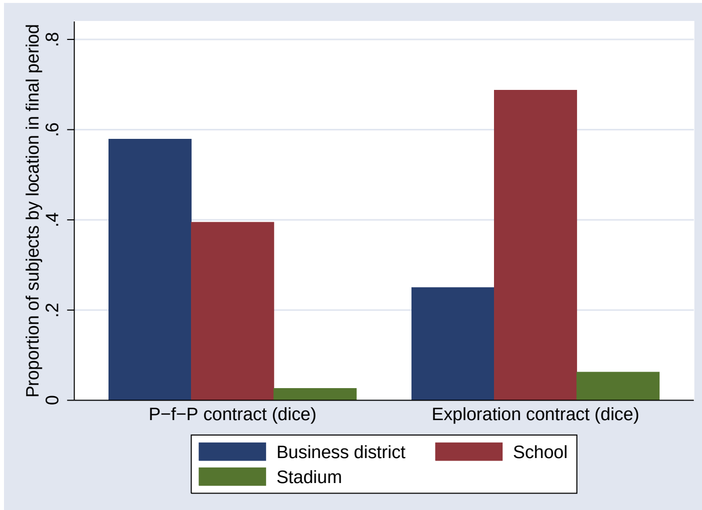

Figure 10: Proportion of subjects by location in the final period of the experiment for the pay-for-performance (dice roll) and exploration (dice roll) contracts.

4 Conclusion

In this paper, we argued that appropriately designed managerial compensation is effective in enabling entrepreneurship and motivating innovation. In a decision-making task in which subjects choose between known and unexplored options we show that subjects are significantly more likely to choose the unexplored mountain than under standard pay-for-performance and fixed-wage contracts. In tasks that involve innovation through experimentation, we find that subjects under a payment scheme that tolerates early failure and rewards long-term success explore more and are thus more likely to discover superior strategies than subjects under fixed-wage or standard pay-for-performance incentive schemes. We also find that the threat of termination can undermine innovation, and that this effect is mitigated by the presence of a golden parachute.

Our results have implications for various problems in which enabling entrepreneurship and motivating innovation are important concerns, such as the design of compensation plans of top executives and middle managers in large corporations as well as of entrepreneurs in start-up companies. In the context of executive compensation, for example, implementing tolerance for failure involves the use of practices that are often criticized such as option repricing, managerial entrenchment, and golden parachutes. Our results suggest that when combined with appropriate long-term rewards, these practices are effective in motivating innovation. Therefore, regulations that restrict their use can have adverse effects on innovation.

By using a controlled laboratory experiment in which individuals are randomly assigned to different contracts, we are able to establish a causal relation between particular incentive schemes and innovation output. Our results complement other contributions to the literature on entrepreneurship and innovation which use naturally occurring data to show that tolerance for failure influences innovation activity in the contexts of bankruptcy laws, labor laws, and corporate governance.

The framework and methods proposed in this paper can be useful in studying other important problems in the entrepreneurship literature. For example, how does the choice of financing affect the entrepreneur's attitude towards innovation? How can we design compensation packages to attract creative entrepreneurs while keeping shirkers and conventional entrepreneurs away? We leave these questions for future research.

Appendices

A Gold Prospecting Experiment

A.1 Experimental Instructions

Instructions

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

At the beginning of the experiments you will receive an initial endowment of $5 in addition to your show-up fee. During the course of the experiments you can earn a further amount of money by gaining gold nuggets. The amount that you gain during the experiments depends only on your own decisions. All the gains that you make during the course of the experiments will be exchanged into cash at the end of the experiments. The exchange rate will be:

$$3 \text{ gold nuggets (gn)} = \$1$$

At the end of the series of experiments you will receive the money that you earned during the experiments.

During each experiment you have to make decisions, which you will enter on a computer screen. The decisions you make and the amount of money you gain 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 experiments. In addition we would like to point out that you may only use the computer functions which are required for the experiments. Communication between participants and unnecessary interference with computers will lead to the immediate exclusion from the experiments. In case you have any questions please don't hesitate to ask us.

Overview of the Experimental Procedures

In this series of experiments, you will be playing several instances of a gold prospecting game. In each experiment the game will be the same, but the payoffs of your choices and the way you are compensated will differ. In particular, the choices on each screen constitute an entirely new game.

Each gold prospecting game consists of 2 periods. In each period you may choose to prospect gold in one of two mountains or to stay at home. Each of these options has an associated revenue and cost:

- **"Stay at home"** always yields 0gn and it costs you 0gn to choose this option.
- **"Mountain 1"** is a well-explored mountain that is close to your home. When you choose to prospect gold at this mountain, it yields 100gn with a probability of 50% and 0gn with a probability of 50%. Thus, by choosing "Mountain 1" you cannot learn whether it yields 0gn or 100gn. It costs you 0.25gn to travel to "Mountain 1" and to prospect gold there.
- **"Mountain 2"** is an unexplored mountain that is further away. With a probability of p, "Mountain 2" is a gold-rich mountain, in which case it always yields 100gn whenever you choose this option. With a probability of $1-p$, "Mountain 2" is a gold-poor mountain, in which case it always yields 0gn whenever you choose this option. You only learn the quality of "Mountain 2" if you choose to prospect gold at that mountain. Since it is further away it costs you 0.75gn to travel to "Mountain 2" and to prospect gold there. Finally, note that each gold prospecting game you play is independent. Thus, you can only learn about the quality of "Mountain 2" for that particular instance of the gold prospecting game.

First, there will be 3 practice trials with different values of p which do not affect your compensation and will help you get acquainted with the interface.

After that, there will be a total of 4 experiments, which differ in terms of the probability p that "Mountain 2" is a gold-rich mountain. At the beginning of each experiment we will announce the probability p that will be relevant during that experiment. The probability p can take on three values: 5%, 35%, 45% or it may be unknown.

In each of the 4 experiments, you will be playing the gold prospecting game under 3 different compensation schemes. Thus, there is a total of $4 * 3 = 12$ separate gold prospecting games which are all independent. Your compensation will depend on your gold production in the two periods of each game. In particular, the 3 compensation schemes are:

1. You will be paid a fixed wage of 0gn per period plus 10% of the gold revenue during the 1st and the 2nd period of the experiment.
2. You will be paid a fixed wage of 0gn per period plus 10% of the gold revenue in the 2nd period of the experiment.
3. You will be paid a fixed wage of 1gn per period.

The different compensation schemes and experiments (12 in total after the trial phase) are completely independent. That is, your choices in each gold prospecting game have no effect on any other choices in the experiment. What you learn about the payoffs is specific to each game.

How is your income calculated?

Your total income in the two periods is calculated in the following way:

total income = fixed wage in 1st period + fixed wage in 2nd period
+ revenue in 1st period * revenue percentage in 1st period
+ revenue in 2nd period * revenue percentage in 2nd period
- costs in 1st period - costs in 2nd period

Your income is therefore higher when your fixed wage, your revenue and your revenue percentage are higher and your costs are lower.

Your fixed wage in each period depends on the compensation of the experiment you are currently in. It is either 0gn or 1gn.

Your revenue in each period depends on the choice you made and the type of mountain. The revenue percentage is 0% or 10% and depends on the compensation of the experiment you are currently in. That is to say you may receive no share of the revenue or you receive some of the revenue depending on the compensation of the current experiment. Finally, as mentioned before the cost of choosing one of the two mountains is 0.25gn or 0.75gn whereas staying at home is costless.

Example

Let's assume the following scenario. In the 1st period and the 2nd period you receive a fixed wage of 0gn and that the revenue percentage in the 1st period is 0% and 10% in the 2nd period.

Let's assume you choose "Mountain 2" in the 1st period. If the mountain revenue was 0gn in the 1st period, then in the 2nd period you choose "Mountain 1". If the mountain revenue was 100gn in the 1st period you choose "Mountain 2".

Case A: Revenue of "Mountain 2" is 0gn in 1st period

Your income in the 1st period therefore is

0 + 0*0 - 0.75 = -0.75gn

and in the 2nd period you choose "Mountain 1" which may yield 0gn or 100gn (and you know for sure that "Mountain 2" yields a revenue of 0gn in the 2nd period). Let's assume that "Mountain 1" yields a revenue of 100gn in the 2nd period. Hence your income in the 2nd period is

$$0 + 100*0.1 - 0.25 = 9.75\text{gn}$$

Case B: Revenue of "Mountain 2" is 100gn in 1st period
Your income in the 1st period therefore is

$$0 + 100*0 - 0.75 = -0.75\text{gn}$$

and in the 2nd period you choose "Mountain 2". Since "Mountain 2" had a revenue of 100gn in the 1st period you know for sure that "Mountain 2" yields a revenue of 100gn in the 2nd period. Hence your income in the 2nd period is

$$0 + 100*0.1 - 0.75 = 9.25\text{gn}$$

B Lemonade Stand Experiment

B.1 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 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 about 90 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.2 Experimental Design

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 with a floor set 0 such that losses in absolute terms for the subject were impossible. 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

C Eliciting Risk Aversion

We 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.

D 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, V. V., R. Baghai-Wadji, and K. Subramanian, 2009, "Labor Laws and Innovation," *SSRN eLibrary*.

Acharya, V. V., and K. Subramanian, 2009, "Bankruptcy Codes and Innovation," *Review of Financial Studies*, 22, 4949–4988.

Aghion, P., J. V. Reenen, and L. Zingales, 2008, "Innovation and Institutional Ownership," *Working Paper*.

Aghion, P., and J. Tirole, 1994, "The Management of Innovation," *Quarterly Journal of Economics*, 109, 1185–1209.

Almeida, P., and B. Kogut, 1999, "Localization of Knowledge and the Mobility of Engineers in Regional Networks," *Management Science*, 45, 905–917.

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

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

Atanassov, J., 2008, "Quiet Life or Managerial Myopia: Is the Threat of Hostile Takeovers Beneficial for Technological Innovation?," *SSRN eLibrary*.

Azoulay, P., J. G. Zivin, and G. Manso, 2011, "Incentives and Creativity: Evidence From the Academic Life Sciences," *Rand Journal of Economics*, 42, 527–554.

Bandiera, O., I. Barankay, and I. Rasul, 2005, "Social preferences and the response to incentives: Evidence from personnel data," *Quarterly Journal of Economics*, 120(3), 917.

——— , 2009, "Social connections and incentives in the workplace: Evidence from personnel data," *Econometrica*, 77(4), 1047–1094.

Banks, J., M. Olson, and D. Porter, 1997, "An experimental analysis of the bandit problem," *Economic Theory*, 10(1), 55–77.

Bebchuk, L., and J. Fried, 2004, *Pay Without Performance*. Harvard University Press, Cambridge, MA.

Bull, C., A. Schotter, and K. Weigelt, 1987, "Tournaments and Piece Rates: An Experimental Study," *The Journal of Political Economy*, 95(1), 1–33.

Chemmanur, T., and X. Tian, 2011, "Do Antitakeover Provisions Spur Corporate Innovation?," *Working Paper*.

Chemmanur, T., X. Tian, and E. Loutskina, 2011, "Corporate Venture Capital, Value Creation, and Innovation," *Working Paper*.

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

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

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

Fallick, B., C. A. Fleischman, and J. B. Rebitzer, 2006, "Job-Hopping in Silicon Valley: Some Evidence Concerning the Microfoundations of a High-Technology Cluster," *The Review of Economics and Statistics*, 88(3), 472–481.

Fehr, E., S. Gachter, and G. Kirchsteiger, 1997, "Reciprocity as a Contract Enforcement Device: Experimental Evidence," *Econometrica*, 65(4), 833–860.

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

Francis, B., I. Hasan, and Z. Sharma, 2010, "Do Incentives Create Innovation? Evidence from CEO Compensation Contracts," *Working Paper*.

Gabaix, X., D. Laibson, G. Moloche, and S. Weinberg, 2006, "Costly Information Acquisition: Experimental Analysis of a Boundedly Rational Model," *American Economic Review*, 96(4), 1043–1068.

Gilson, R. J., 1999, "The Legal Infrastructure of High Technology Industrial Districts: Silicon Valley, Route 128, and Covenants Not to Compete," *New York University Law Review*, 74, 575–629.

Hellmann, T., 2007, "When do employees become entrepreneurs?," *Management Science*, 53, 919–933.

Hellmann, T., and V. Thiele, 2011, "Incentives and Innovation : A Multi-tasking Approach," *American Economic Journal: Microeconomics*, 3, 78–128.

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

Katok, E., and E. Siemsen, 2011, "Why Genius Leads to Adversity: Experimental Evidence on the Reputational Effects of Task Difficulty Choices," *Management Science*, 57, 1042–1054.

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

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

Lerner, J., and J. Wulf, 2007, "Innnovation and Incentives: Evidence from Corporate R&D," *Review of Economics and Statistics*, 89, 634–644.

Liu, Q., and K. P. Wong, 2011, "Intellectual Capital and Financing Decisions: Evidence from the U.S. Patent Data," *Management Science*, 57(10), 1861–1878.

Manso, G., 2011, "Motivating Innovation," *Journal of Finance*, 66, 1823–1860.

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

Marx, M., D. Strumsky, and L. Fleming, 2009, "Mobility, Skills, and the Michigan Non-Compete Experiment," *Management Science*, 55(6), 875–889.

McCullers, J. C., 1978, *Issues in Learning and Motivation*Lawrence Erlbaum Associates, Hillsdale, NJ.

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

Meloso, D., J. Copic, and P. Bossaerts, 2009, "Promoting Intellectual Discovery: Patents Versus Markets," *Science*, 323(5919), 1335–1339.

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

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

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

Sannikov, Y., 2008, "A Continuous-Time Version of the Principal-Agent Problem," *The Review of Economic Studies*, 75, 957–984.

Sapra, H., A. Subramanian, and K. Subramanian, 2008, "Corporate Governance and Innovation: Theory and Evidence," *SSRN eLibrary*.

Sauermann, H., and W. M. Cohen, 2010, "What makes them tick? Employee Motives and Industrial Innovation," *Management Science*, 56, 2134–2153.

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

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

Sorenson, O., and T. E. Stuart, 2003, "Liquidity events and the geographic distribution of entrepreneurial activity," *Administrative Science Quarterly*, 48(175-201).

Tian, X., and T. Wang, 2010, "Tolerance for Failure and Corporate Innovation," *Working Paper*.

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

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

Yanadori, Y., and J. H. Marler, 2006, "Compensation strategy: does business strategy influence compensation in high-technology firms?," *Strategic Management Journal*, 27(6), 559–570.