Optimal innovation time-off

Ayush Pant Federico Trombetta
Department of Economics, The University of Warwick *

September 9, 2019

PRELIMINARY AND INCOMPLETE – Do not circulate

Abstract

We look for the optimal delegation mechanism that provides an agent with time-off from his current task to pursue his creative endeavors. Driven by high intrinsic motivation, the agent would like the get time off to pursue any idea he discovers. The principal, on the other hand, would like to offer the agent the time off only if she believes he has come up with an idea that has a high potential to succeed. We show that in the optimal mechanism the principal is inefficiently harsh on an agent who was initially provided the time off but lenient to the one who wasn't given the time off. This implies that an agent with high potential idea might get only a limited chance to achieve the breakthrough. Creativity, therefore, only receives a limited opportunity.

Keywords: dynamic delegation, innovation and creativity

JEL Classification Numbers:

*Email: A.Pant@warwick.ac.uk and F.Trombetta@warwick.ac.uk. We would like to thank Motty Perry, Robert Akerlof, Dan Berhardt and Costas Cavounidis for their time and insightful comments. We would also like to thank conference participants at the Midwest Economic Theory Conference, Indiana 2019. All errors are our own.

1 Introduction

Employee-driven innovations have become critical for the growth of big organizations in recent decades. Employees are usually more capable of identifying new ideas and problems with current workstreams and finding their solutions. First-hand experience working with the technology and specific expertise gives them a distinct advantage in the innovation process.[1] For instance, almost half of Google's products have been employee-driven innovations. Paul Buchheit, credited with developing *GMail* in Google, initially sought to refine the emailing experience after identifying that it was difficult to search within specific email inboxes. He later developed *AdSense* technology to recommend advertisements based on email searches, which now makes Google around $10 billion each year in revenue.[2]

Recognizing the importance of employee-driven innovations, many firms have introduced policies of innovation time-off for their employees. For instance, Google permits its employees to take 20% time-off from their regular work to work on a project of their choice. Similarly, LinkedIn runs a yearly time-off program called the InCubator to support its employees' ideas.

However, there are apparent tradeoffs for organizations in providing employees with time off to work on their creative endeavors. On the one hand, employees would need to take time off from regular tasks to work on their projects that may lead to innovations and higher future profits for the firm. But on the other hand, providing the time off induces the agent to work on potentially unfruitful avenues, which is costly to the organizations as it diverts resources away from present tasks. In the presence of this tradeoff, we ask how a firm should optimally delegate authority to its employees to work on their creative projects.

(Currently) We build a simple two-period delegation model to answer this question. In our model, a principal (she) finances an agent's (he) work. The agent either works on a regular task, or he may request to work on his creative task. The regular task brings a normalized zero net benefit to either player. The creative task involves the agent working on his idea, which has the potential to generate a breakthrough and a higher payoff of 1 to either player. However, the type of idea that the agent possesses in any period is his private information and may be either high or low. A high-type idea has a higher ex-ante potential to succeed and generate a breakthrough than a low-type one. While the agent would like to gain authority to work on his idea independent of its type, the principal only

[1] This is especially true for tech firms in Silicon Valley that spend millions of dollars in their recruitment processes to select the best candidates who are of the highest ability.

[2] Another example of successful employee-driven innovation is Flickr. It originally a side project for the team at Ludicorp whose sole product at the time was a web-based game called Game Neverending. See Tate (2012) for more examples and anecdotes.

finds it profitable to finance a high idea. In each period, the agent may draw a new idea if he does not already have a high idea.

The principal designs an optimal delegation policy which is contingent on past observed outcomes. One may imagine several plans that the principal may use. Fixed probability of granting authority across periods, declining probability, increasing probability, or a combination of the three depending on the previous outcome are a few examples. Our main result shows that the optimal mechanism resembles a time-based screening contract in which the agent may choose to seek authority today or tomorrow. If the agent seeks authority today, then he faces a potential future punishment for lack of performance. Specifically, the principal reduces the probability of granting authority in the next period when he does not produce a breakthrough in the first period. If the agent does not seek authority today, then the principal rewards him with more authority tomorrow *independent of his type*. Thus, in the optimal policy, the principal punishes the *persistent* good type but not the bad type. Either way, creativity only gets a limited opportunity.

The reason is as follows: the principal needs to give sufficient incentives to a low-type agent to deter seeking authority in the first period. She achieves this by using the future probability of authority in each contingency, i.e., if he does not seek authority and if he does and fails to produce a breakthrough. Therefore, she rewards the agent for continuing work on the regular task *and* also punishes him for seeking authority and not providing a breakthrough. Consequently, only the high-type agent gets penalized in the optimal incentive-compatible mechanism.

There is some anecdotal evidence to support this observation. Google managers are known to clamp down on those employees that take the 20% time off but do not produce breakthroughs sufficiently quickly. In one of the Quora blog posts[3], a Google employee writes:

> "Unofficially, 20% projects are no longer encouraged. They led to many problems because it took a great deal of time away from an employee's primary team (without any measurable successes)."

While at the same time, some successful employees mention how it is important to define the objectives of their projects to be able to benefit from the time-off policy. Defining objectives allows the managers to measure success and reward (or punish) accordingly in their projects.

It is interesting to note that it is not possible to achieve the desired honesty of the low-type using just one of the tools. The principal must not only reward the low-type but also punish the high-type in the optimal policy. In doing so, she faces an intertemporal tradeoff. Punishing the agent following

[3]See https://www.quora.com/How-does-Google-company-s-Google-Innovation-Time-Off-20-time-policy-work-in-practice

failures relaxes the low-type's first-period incentive constraint but also makes the high-type's second-period constraint tighter. However, this hurts the principal as a high potential idea is persistent and taking away authority from such an agent reduces her expected payoff. The principal, therefore, tries to minimize the adverse effect on the high-type while at the same time maintaining the low-type's incentive to report honestly. How she uses the two tools optimally to balance the two effects depends on the various parameters.

We show that somewhat counterintuitively, the principal may sometimes grant authority to the agent even when ex-ante expected benefit of doing so is lower than the cost. It will never happen if we restrict the mechanism to a single period. However, in two periods, the intertemporal tradeoff kicks in. In such situations, it is more costly not to give authority to high types than to fund the low-type's creativity. Thus, while the principal is forced to punish the high-type agent for failing, she does so to the minimum possible extent. In this sense, the first-period high-type agent subsidizes the low-type in the optimal mechanism.

However, if the ex-ante cost of financing creativity is too high, then there is a discontinuous fall in the probability of getting authority for both the types in the second period. In this case, the principal would rather completely take away authority from the high-type, and risk not getting a breakthrough, than grant authority to the low-type. Again, while the principal must offer some authority to the low-type, she does so to the minimum extent. In this sense, the first-period low-type taxes the first and the second-period high-type in the optimal mechanism.

Going forward, we want to provide some comparative statics and extend this model in meaningful ways (such as multiple types, moral hazard, etc.).

Related literature: Starting with Holmstrom (1982) there is a vast literature on delegation of authority in economics. Aghion and Tirole (1997) were the first to formalize the collaborative role played by employees in an organization. They showed that employee initiative could be increased by formal delegation and reducing the level of managerial effort. In contrast, Rantakari (2012) showed that employee initiative might be increased by combining formal authority (of the manager) and limited but positive involvement of the manager. This result was achieved by combining the Aghion-Tirole model with elements of costly monitoring. While ours is a model of employee initiative, we deal with the issue of dynamic delegation of authority. A number of papers have started exploring this issue recently in different contexts; Frankel (2016), Guo (2016), Datta (2017), Guo and Horner (2017), Li, Matouschek, and Powell (2017) Lipnowski and Ramos (2018) are a few.

Among these, Guo (2016) and Datta (2017) deal with the delegation in environments where the principal learns about ex-ante private information of the agent owing to experimentation. While it is common to model innovation in an experimentation-type framework, we do not explicitly need to include experimentation and learning in our model. All we require is ex-ante uncertainty over when breakthrough occurs and better information of the agent. Moreover, we are interested in the question of where ideas come from and when are they developed further. Our concern is not how a given idea is developed.

Guo and Horner (2017) is the closest in this sense to our paper. In their setting, a principal commits to an allocation policy for a perishable good (much like delegating authority) when the agent's type is persistent. While the agent would like the good either way, the principal interested in maximizing efficiency would want to grant the good only if the agent is of high type. However, there are two critical points of departure for us. First, we have a one-sided persistence of type only. In our model, only the good-type is persistent. For us, an agent who has discovered a good idea would prefer to see it through than drawing another idea, which at best is the same as the current idea. Moreover, we require that there might not be an immediate conclusion of the creative task. Second, we allow for breakthroughs to be observable. As we discuss in our analysis, these two features together produce our fundamental intertemporal tradeoff and produce our main result.

Other papers such as Li et al. (2017), Rantakari (2017) and Lipnowski and Ramos (2018) do not assume commitment by the principal and are interested in the equilibrium allocation rules that arise in similar settings using perfect public equilibrium (PPE). In Li et al. (2017), the authors show how good early choices of subordinates are rewarded with later authority. Consequently, they make more selfish decisions that end up hurting organizations in the long run. Lipnowski and Ramos (2018) build a model in which a principal sequentially delegates project choice to an agent who can assess its quality but has lower standards of acceptance. Similar to Li et al. (2017), they show that in equilibrium, the principal incentivizes the current good selection of projects by allowing future bad choices.

A common theme between this literature and our paper is the idea of linking incentives across periods or decisions. By controlling allocation or decision rights to some other (or future) units, the principal creates value to eliciting private information today. This feature appears in several different contexts, including in papers that look at relational contracts and optimal contracts.

Jackson and Sonnenschein (2007) prove that the limitation imposed by incentive constraints on attaining social efficiency can be reduced by increasing the number of copies of the decision problem and thereby linking them. In a similar spirit, Malenko (2018) develops a model of dynamic capital budgeting in which an agent with a desire to overinvest in projects privately sees the arrival and quality of investment opportunities. A principal allocates resources to the agent to be spent on these investment opportunities. The optimal mechanism involves the principal setting a dynamic spending account that gets replenished with time at a specific rate. Additionally, the mechanism outlines a threshold limit above which the agent can pass the project to the principal for auditing. Similarly, Möbius (2001) and Hauser and Hopenhayn (2008) build models of favor trading and show how the number of favors exchanged between players may be used to determine how many new favors can be exchanged.

Two related papers are Boleslavsky and Kelly (2014) and Casella (2005). In the context of environmental regulation, using a two-period model Boleslavsky and Kelly (2014) show how the regulators may vary the strength of regulations over time when the firm privately learns its compliance costs. However, again, neither do they have one-sided state persistence nor do they do they have a verifiable signal of the state (since the latter is not a concern in their setting). Similarly, Casella (2005) develops a mechanism of storable votes that allows an agent to gain more influence in future democratic decision-making by giving up the right to vote today when she expects future preferences to be strong.

The rest of the the paper is organized as follows: in Section 2 we present what our general model will look like. However, we focus currently on the simplified version of the model presented in Section 3. We discuss the model and present all the relevant results within the section itself.

2 Model

We present here the general model that we expect to solve after solving the simple two-period model.

Players, Tasks and Types: An agent (he) is in an employment relationship with a principal (she) through time $t = 0, 1, 2, ...$ The two players are denoted by $i \in \{A, P\}$ and share a common discount factor δ. While in the relationship, the agent could potentially work on one of the two tasks in any given period: regular or own. The regular task corresponds to working on assigned projects. Conducting the regular task gives a payoff of $r_i > 0$ to each player with certainty.

While working on the regular task, the agent has the ability to come up with new ideas and create own tasks. The agent costlessly comes up with an own task in each period, which could either have a low or a high potential to succeed, denoted by $\theta \in \{l, h\}$. This is the type of the own task and is agent's private information. The prior probability of drawing a high potential own task in period t is p_t. If the agent works on own task of type θ, then the probability of success or breakthrough in any given period is θ with $0 < l < h < 1$ – a high potential task is more likely to succeed in any given period. Upon achieving a breakthrough in own task, each player gets a payoff of v_i from then on in perpetuity. We will assume that success in own task is sufficiently rewarding such that $r_i < \delta v_i$. If the agent fails, then both players get 0 in that period. We denote the outcome of conducting own task by $y \in \{S, F\}$, success or failure respectively i.e. either a breakthrough is observed or not.

The higher payoff to the agent from conducting own tasks pertain to the intrinsic motivation for working on own ideas. Moreover, upon achieving success in own tasks, the principal may reward the agent with promotions and more autonomy to further conduct own task which gives the agent v_A in perpetuity. The higher payoffs to the principal from achieving success in own task pertain to the benefits of getting an innovation.

We will assume that the principal can observe the outcome y of conducting any task and also the realized payoff. This implies that the principal can perfectly distinguish between the agent conducting regular vs. own task but cannot distinguish between low potential and high potential own task.

Note that since there is nothing better than a high potential idea, the agent stops drawing new ideas after getting a high potential idea. By drawing a new idea he risks losing his current high idea. But may continue drawing ideas if she has a low potential idea, because there is only a potential benefit of doing so. Thus the high type is fully persistent.

Conflict of Preferences: While both players agree on the need for innovation, there is a conflict on which type of ideas should be pursued. We will assume

$$hv_P > r_P \text{ and } hv_A > r_A$$
$$\underline{\text{but}}\ lv_P < r_P \text{ and } lv_A > r_A$$

This implies that while the agent always prefers to conduct own tasks, the principal would rather have him conduct the regular task when she has a low potential own task. In other words, the agent's intrinsic motivation to "scratch the itch" is sufficiently high independent of the type of the task he has drawn. However, the principal would like the agent to do so if she has a high potential own task.

Policy and Payoffs: Given the realized θ_t and the policy (defined next), the agent sends a message $m_t \in \{l, h\}$. A policy (or mechanism) is a sequence of probabilities of granting authority to the agent to conduct own task for every period t denoted by a_t as a function of previous messages, outcomes and authority decisions.

$$a_t : \{l, h\}^{t+1} \times \{S, F\}^t \times \{0, 1\}^t \to \Delta(\{0, 1\})$$

Timing: The timing of the game is as follows:

1. Initially, the principal commits to a mechanism $a_t(.)$ for $t = 1, 2, ...$

2. In $t = 1$, the agent draws his type θ_1 and reports it to the principal. The principal chooses according to $a_1(\theta_1)$. Outcome y_1 is realized if $a_1 = 1$, otherwise the game moves to $t = 2$.

3. For all $t > 1$, the agent reports his θ_t to the principal and she chooses whether to grant authority using $a_t(.)$.

3 Two-period case: model and analysis

We present here a modified two-period version of the model. Consider the following changes. An agent and a principal are in an employment relationship for two periods $t = 1, 2$. The principal bears a cost of $0 < c < 1$ per period for financing the agent's work. If the agent works on the regular task, then the principal breaks even and gets 0. The agent also has no intrinsic motivation to do the regular task and gets 0 from conducting it. On the other hand, both the agent and the principal get a benefit of 1 from success in own task.

However, there is a preference conflict between the agent and the principal captured by the following assumption on parameters: $h > c > l$ while $1 > h > l > 0$. This implies that while the principal would like the agent to only pursue high potential ideas, the agent would like to have the authority to conduct own task independent of its potential to succeed. Thus, the agent has lower standards. The relationship ends after getting a success in own task or at the end of the second period, whichever happens first. All the other parameters of the problem are as above.

The principal designs a delegation mechanism which defines for every t the probability of granting authority to the agent to work on his own task a_t. Using the revelation principle, we restrict attention to those mechanisms in which the agent reports his type and the principal decides whether authority must be granted.

To begin with, let us look at the principal's first-best policy. The first-best situation is given by the situation in which the principal can perfectly observe the type of the agent's own task in each task

Proposition 1 *The principal chooses $a_t = 1$ whenever $\theta_t = h$, and $a_t = 0$ otherwise in the first-best policy.*

It is easy to see why the above is the case. As $h > c > l$, the expected payoff of the principal is always positive whenever the agent has a high-potential own task. Alternately, it is negative when the agent has a low-potential own task.[4]

[4] Note that the agent does not need to be incentivized to participate in the mechanism. He always gets 0 by being in the employment relationship with the principal.

Second, let us look at two extreme policies that the principal may adopt under the second-best. The first is to never grant any authority and the second is to always grant authority.

Lemma 1 *If the principal never grants authority, then her expected payoff is zero. If the principal always grants authority, her expected payoff is*

$$p_1h + (1-p_1)l - c + \delta\big[p_1(1-h)(h-c) + (1-p_1)(1-l)(p_2h + (1-p_2)l - c)\big].$$

Going forward, we need to determine if the principal can do better than these two mechanisms, and if so, when. Below, we will call all other non-extreme mechanisms as *interior mechanisms.*

Note first that in $t = 1$, there is no outcome to condition a_1 on. Therefore, the principal chooses a_1 as a function of m_1 which could be either h or l. With some abuse of notation, we denote this by $a_1(h)$ and $a_1(l)$ respectively. In $t = 2$, however, the principal may condition her decision on three variables – the present report, the past report and the past outcome, which could either be a failure or that no authority was granted.[5] We denote the principal's period 2 decision by $a_2(m_2; m_1, y_1)$ where $m_t \in \{h, l\}$ and $y_1 \in \{F, \varnothing\}$ is the outcome of the previous period which could be either that the authority was granted but the agent failed to produce a breakthrough (F) or that he wasn't granted authority ($\varnothing$). Observe that there are many variables that one must determine in the optimal mechanism. In the following paragraphs, we show how the problem can be reduced significantly.

Lemma 2 *The optimal interior mechanism does not condition second period delegation decision on the reports in that period.*

This follows almost immediately from the fact that the game ends after the second period. The principal is not able to incentivize truth-telling by offering any future rewards or punishments in the second period. Incentive compatibility requires that she offers the same authority to the agent independent of her second period type. Thus, period 2 decisions can only be based on the past decisions and outcomes, and not on the present reports. This reduces our problem to determining the following probabilities in the second period: $a_2(l, F)$, $a_2(l, \varnothing)$, $a_2(h, F)$ and $a_2(h, \varnothing)$. These probabilities of granting authority reflect the principal's decision for when 1) the agent reported low but was given authority and failed, 2) the agent reported low and was not given authority, 3) the agent reported high and was given authority but failed, and 4) the agent reported high but was not given authority in period 1.

[5] If authority was granted in the first period and it generated a success, then there is no further need for making authority decisions as the relationship comes to an end.

We can now write the principal's optimal authority design problem as follows:

$$
\begin{aligned}
\underset{a_t(m,y)}{\text{maximize}} \quad & p_1(h-c)\big[a_1(h)+a_1(h)(1-h)\delta a_2(h,F)+(1-a_1(h))\delta a_2(h,\varnothing)\big] \\
& \quad +(1-p_1)\big[a_1(l)(l-c)+\delta(p_2h+(1-p_2)l-c)(a_1(l)(1-l)a_2(l,F)+(1-a_1(l))a_2(l,\varnothing))\big] \\
\text{subject to} \quad & a_1(h)+\delta\big[a_1(h)(1-h)a_2(h,F)+(1-a_1(h))a_2(h,\varnothing)\big] \\
& \qquad \geq a_1(l)+\delta\big[a_1(l)(1-h)a_2(l,F)+(1-a_1(l))a_2(l,\varnothing)\big], \qquad \text{(IC)}_h \\
& a_1(l)l+\delta(p_2h+(1-p_2)l)\big[a_1(l)(1-l)a_2(l,F)+(1-a_1(l))a_2(l,\varnothing)\big] \\
& \qquad \geq a_1(h)l+\delta(p_2h+(1-p_2)l)\big[a_1(h)(1-l)a_2(h,F)+(1-a_1(h))a_2(h,\varnothing)\big], \quad \text{(IC)}_l
\end{aligned}
$$

where $a_t(m,y)=\{a_1(h),a_1(l),a_2(l,F),a_2(l,\varnothing),a_2(h,F),a_2(h,\varnothing)\}$ is the set of all delegation probabilities that the principal chooses.

Lemma 3 *The low type's incentive constraint $(IC)_l$ binds in the optimal interior mechanism.*

The intuition is straightforward. Fix any incentive compatible mechanism in which both the ICs are slack, specifically (IC)_l is non-binding. Now, if the principal increases $a_1(h)$ while satisfying (IC)_l then the ex-ante expected profits increase on account of a higher probability of breakthrough for the high type without inducing any deviation by the low type. This means that optimality requires (IC)_l to bind.

Lemma 4 *Let $p_2>l$ and $\underline{\delta}<\delta<\bar{\delta}$ where $\underline{\delta}=\frac{l}{p_2h+(1-p_2)l}$ and $\bar{\delta}=\frac{p_2}{p_2h+(1-p_2)l}$. It is always optimal for the principal to grant authority following a high report and not following a low report in period 1, i.e. $a_1(h)=1$ and $a_1(l)=0$.*

The above lemma shows that the first-best optimal delegation policy is the same as the second-best optimal delegation policy in the first period. In the presence of future discounting, the principal prefers not to distort incentives in the first period. The first sufficiency condition on the discount factor, $\delta<\bar{\delta}$, ensures that future is not too valuable, which might lead to pushing all authority decisions to the second period. The second sufficiency condition $\delta>\underline{\delta}$ ensures that the future is sufficiently valuable to the agent so that the low type does not want to deviate today and inefficiently seek authority.

Using Lemma 4, it is easy to reduce the principal's interior maximization problem to

$$\begin{aligned}
\underset{a_2(h,F),\,a_2(l,\varnothing)}{\text{maximize}} \quad & p_1(h-c)\big[1+(1-h)\delta a_2(h,F)\big] \\
& + (1-p_1)\delta(p_2h+(1-p_2)l-c)a_2(l,\varnothing) \\
\text{subject to} \quad & 1+\delta(1-h)a_2(h,F) > \delta a_2(l,\varnothing) && \text{(IC)}_h \\
& a_2(l,\varnothing) = \frac{l}{\delta(p_2h+(1-p_2)l)} + (1-l)a_2(h,F) && \text{(IC)}_l
\end{aligned}$$

where we only need to determine two variables $a_2(h,F), a_2(l,\varnothing)$. Moreover, note that (IC)_h is slack at the optimum. Therefore, the optimal interior mechanism is the solution to above linear program as outlined by the following proposition.

Proposition 2 *Let $p_2 > l$. For $\underline{\delta} < \delta < \bar{\delta}$ the optimal interior mechanism is given by the following delegation probabilities:*

- $a_1^*(h) = 1$ *and* $a_1^*(l) = 0$,
- *if* $c < \bar{c}$ *where* $\bar{c} = \frac{p_1h(1-h)+(1-p_1)(1-l)(p_2h+(1-p_2)l)}{1-(p_1h+(1-p_1)l)}$,

$$a_2^*(h,F) = \frac{1}{1-l}\left(1-\frac{l}{\delta(p_2h+(1-p_2)l)}\right) < 1 \;\; \textit{and} \;\; a_2^*(l,\varnothing) = 1,$$

- *if* $c = \bar{c}$, *then any* $a_2^*(h,F)$ *and* $a_2^*(l,\varnothing)$ *that satisfies*

$$a_2^*(l,\varnothing) = \frac{l}{\delta(p_2h+(1-p_2)l)} + (1-l)a_2^*(h,F),$$

- *if* $c > \bar{c}$,

$$a_2^*(h,F) = 0 \;\; \textit{and} \;\; a_2^*(l,\varnothing) = \frac{l}{\delta(p_2h+(1-p_2)l} < 1, \;\; \textit{and}$$

- *any* $a_2^*(h,\varnothing) \in [0,1]$ *and* $a_2^*(l,F) \in [0,1]$.

The above proposition characterizes the optimal mechanism for different cost ranges. Observe how the probabilities of delegation switch for the different types between first and the second period. An agent who is of high type in the first period must suffer some punishment in the second period for failing to achieve a breakthrough. This is despite the high type being fully persistent. At the same time, a low type must be rewarded for waiting to get a high idea when he presently doesn't have one. This implies that the first period low type must get a higher probability of delegation in the second period independent his future type. Thus, the principal buys period 1 efficiency from the low type agent by inefficiently offering rewards to the low type and punishments to the high type in period 2.[6]

We note here that the flipping of delegation probabilities for the two types is unique to our environment and arises from the two differentiating features of our model – observable outcomes and persistent high type. These two features together generate a tradeoff for the principal. On the one hand, the possibility of conditioning future delegation on the past performance relaxes the low type's constraint by making deviations less likely today. But on the other hand, since the high type is fully persistent, it hurts his incentives tomorrow. This in turn reduces the expected payoffs of the principal as she must knowingly take away authority from someone who brings in positive expected profits. Such an intertemporal tradeoff is resolved by the principal by switching delegation probabilities across periods for the two types in the optimal mechanism.

Why must the principal use both the tools? This is best understood by looking at $(\text{IC})_l$ when $a_1(h) = 1$ and $a_1(l) = 0$;

$$a_2(l, \varnothing) = \frac{l}{\delta(p_2 h + (1 - p_2)l)} + (1 - l)a_2(h, F)$$

It is easy to see that two delegation probabilities work complementarily. An increase (decrease) in one is accompanied by an increase (decrease) in the other in order to maintain incentive compatibility of the low type. However, given the incentive compatibility of the high type, the principal would like to increase $a_2(h, F)$ to the maximum possible extent *independent* of $a_2(l, \varnothing)$. This might not always be doable or in the interest of the principal. The two cost cases depict this.

To begin with, note that given the linearity of the principal's objective function and the IC constraints, the solution must be an extreme point. Now, when the cost of delegating is low enough, i.e. lower than $\bar{c}$, then the principal prefers to still offer authority to the high type with a positive probability and simultaneously with probability 1 to the low type. However, when the cost is high, then financing the high type for a second period while simultaneously delegating to an ex-ante unknown type becomes too costly. In this situation, the principal would rather give no authority to the high type and restrict the delegation probability to the minimum for the low type. We explain this point further using the corollary below. For this purpose, let $p_1 = p_2 = p$.

Corollary 1 *Let $p_1 = p_2 = p$. For $\tilde{c} \leq c < \bar{c}$ where $\tilde{c} = ph + (1 - p)l$ is the ex-ante expected benefit of*

[6] Since it never happens that a low type is offered authority and the high type is not offered authority in the first period, $a_2(h, \varnothing)$ and $a_2(l, F)$ do not matter in anyone's decision-making, and we may assign any probabilities to them.

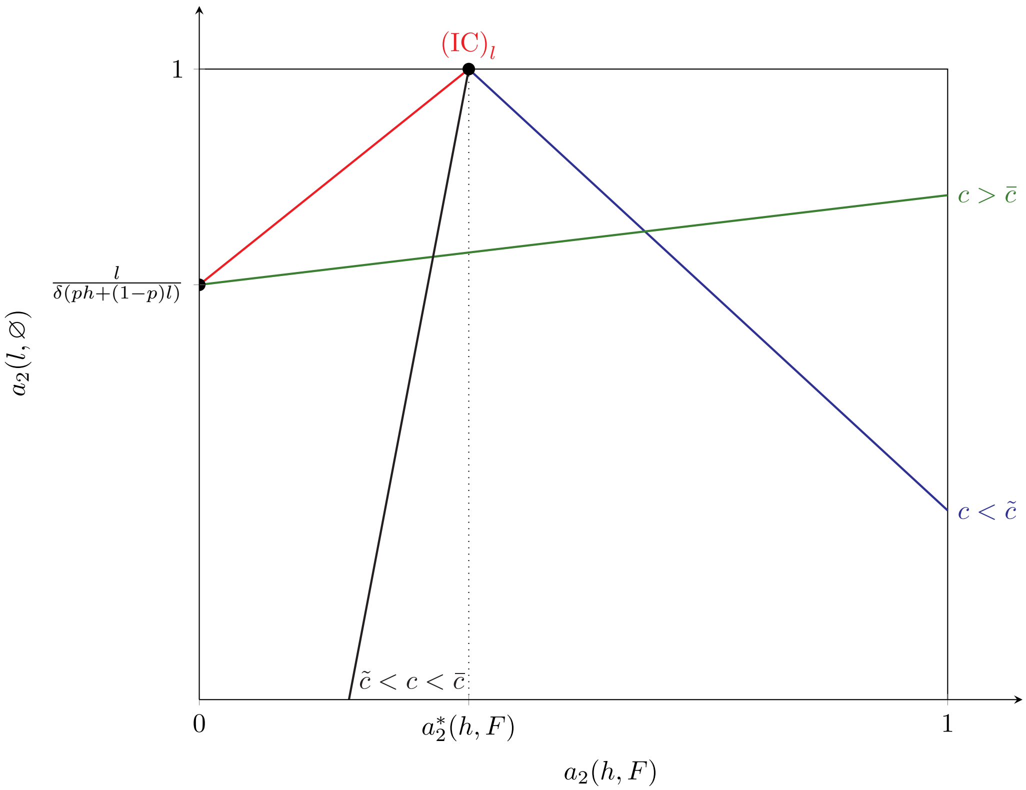

Figure 1: Optimal delegation mechanism for different c ranges

delegation, the optimal mechanism involves

$$a_2(h,F) = \frac{1}{1-l}\left(1 - \frac{l}{\delta(ph+(1-p)l)}\right) < 1 \;\; \textit{and} \;\; a_2(l,\varnothing) = 1.$$

The above corollary shows that even if the ex-ante expected benefit of granting authority is lower than the cost of doing so, the principal optimally delegates authority with probability 1 to such an agent some times. Observe that this would not be the case if this was a single period mechanism. In a single period, the principal would never offer authority to an agent who is too costly to fund. The difference arises because of how the principal resolves the aforementioned intertemporal tradeoff. See Figure 1 in relation to the explanation below.

When $c < \tilde{c}$, i.e. the cost of delegating is lower than the the expected benefit of delegating, then the preferences of the principal and period 1 low type agent are more aligned. On an average, she expects to gain by granting authority to period 1 low type in period 2. At the same time, she would like to continue granting authority in period 2 to period 1 high type. This is reflected in her isoprofit lines. The slope of the isoprofit line in $(a_2(h,F), a_2(l,\varnothing))$ space is negative. To remain at a given profit level, the principal must increase $a_2(l,\varnothing)$ following a decrease in $a_2(h,F)$. This leads to an easy resolution of the tradeoff in the optimal mechanism; the principal happily gives authority to the low type while holding the punishment to the high type to a minimum.

But when $c > \tilde{c}$, i.e the cost of delegating authority to period 1 low type in period 2 is higher than the expected benefit, then the preferences are less aligned. This means that the principal would like to take away authority from the low type i.e. reduce $a_2(l, \varnothing)$. At the same time, the desire to not punish the high type agent in period 2 remain. Thus, to maintain a given profit level a reduction in $a_2(h, F)$ must be accompanied by a reduction in $a_2(l, \varnothing)$. The intertemporal tradeoff now has a bite - the principal must decide between using $a_2(h, F)$ more intensely or less intensely.[7] By choosing a lower $a_2(h, F)$, she can minimize the punishment to the high type but it comes at the cost of inefficiently giving authority to the low type in period 2. Alternately, she can reduce the inefficiency by increasing the high type's punishment tomorrow, i.e. a higher $a_2(h, F)$. How this tradeoff is resolved now depends on a second cost threshold $\bar{c}$.

$\bar{c}$ reflects the cost threshold at which the principal is indifferent between how the tradeoff is resolved. This implies that for $\tilde{c} \leq c < \bar{c}$ the principal still prefers to grant authority to period 1 low type in period 2. This is so because in this range of cost parameters the principal's cost of taking away authority from a high type is higher than the cost of inefficiently granting authority to the low type. Thus, the high type agent 'subsidizes' the low type in the optimal mechanism. But when $c > \bar{c}$, the cost of granting authority to the low type is higher than the cost of taking away authority from the high type. Thus, the principal prefers to set $a_2(h, F) = 0$ and minimize $a_2(l, \varnothing)$ in the process. The low type, in this case, 'taxes' the high type.

Finally, the optimal mechanism is evaluated by comparing the principal's payoffs from the optimal interior mechanism with the extreme mechanisms outlined in Lemma 1.

Proposition 3 *Let $l < p_2 < \frac{l}{1-(h-l)}$. For $\underline{\delta} < \delta < \bar{\delta}$ the optimal mechanism is given by cost thresholds $\underline{c}$, $\bar{c}$ and $\bar{\bar{c}}$ where $\underline{c} < \bar{c} < \bar{\bar{c}}$ such that*

- *if $c < \underline{c}$, the principal always grants authority in the optimal mechanism;*
- *if $\underline{c} \leq c < \bar{c}$, the principal grants authority as follows in the optimal mechanism*

$$a_1^*(h) = 1 \, , \, a_1^*(l) = 0 \, , \, a_2^*(h, F) = \frac{1}{1-l}\left(1 - \frac{l}{\delta(p_2 h + (1-p_2)l)}\right) < 1 \;\; \textit{and} \;\; a_2^*(l, \varnothing) = 1,$$

and any $a_2^*(h, \varnothing) \in [0, 1]$ and $a_2^*(l, F) \in [0, 1]$;

[7] Note that not using $a_2(h, F)$, i.e. setting $a_2(h, F) = 1$ is not an option for the principal because $(IC)_l$ must be respected at all times. Setting $a_2(h, F) = 1$ breaks $(IC)_l$ as it requires an $a_2(l, \varnothing) > 1$ to match it.

- *if $\bar{c} \leq c < \bar{\bar{c}}$, the principal grants authority as follows in the optimal mechanism*

$$a_1^*(h) = 1\,,\; a_1^*(l) = 0\,,\; a_2^*(h, F) = 0 \;\; \textit{and} \;\; a_2^*(l, \varnothing) = \frac{l}{\delta(p_2 h + (1 - p_2) l)} < 1,$$

and any $a_2^*(h, \varnothing) \in [0, 1]$ and $a_2^*(l, F) \in [0, 1]$;

- *if $c \geq \bar{\bar{c}}$, the principal never grants authority.*

The above proposition outlines the overall optimal mechanism for different cost ranges. As expected, we show that when the cost of granting authority is too low or too high, the principal prefers to implement the extreme mechanism instead of the interior one. When the cost is low, the principal always grants authority; when the cost is too large, the principal never grants authority. In between, the principal prefers the optimal interior mechanism.

4 Conclusion

We showed using a simple two-period model how organizations may limit the creativity of their employees. Organizations end up being inefficiently harsh on employees who are capable of achieving breakthroughs in own ideas to limit the behavior of those employees who do not have good ideas. Such mechanisms are, therefore, not likely to be successful in promoting creativity among the employees. Our model suggests that organizations must look for alternative ways of fostering creativity.

A few points are in order about the model that we built in this paper. First, there are some obvious issues in extending our model to more than two periods. Since we need that the agent's outcome is observable, when we expand our model to include many periods, the problem blows up immediately. Second, this is not a model of experimentation even though we are attempting to model innovation. By assuming that θ is perfectly known to the agent, we are essentially killing any learning, a standard of experimentation models. We believe this will only complicate things further without adding any value. Finally, there are no monetary transfers in the model. This is so because we are interested in employee-driven innovation where the employer-employee are already in a relationship and the employee wants to conduct innovation driven by intrinsic motivation. Our model, therefore, shows further scope of research.

References

Aghion, P., & Tirole, J. (1997). Formal and real authority in organizations. *Journal of political economy*, *105*(1), 1–29.

Boleslavsky, R., & Kelly, D. L. (2014). Dynamic regulation design without payments: The importance of timing. *Journal of Public Economics*, *120*, 169–180.

Casella, A. (2005). Storable votes. *Games and Economic Behavior*, *51*(2), 391–419.

Datta, B. (2017). Delegation and learning. *Working paper*.

Frankel, A. (2016). Discounted quotas. *Journal of Economic Theory*, *166*, 396–444.

Guo, Y. (2016). Dynamic delegation of experimentation. *American Economic Review*, *106*(8), 1969–2008.

Guo, Y., & Horner, J. (2017). Dynamic mechanisms without money.

Hauser, C., & Hopenhayn, H. (2008). Trading favors: Optimal exchange and forgiveness. *Working paper*.

Holmstrom, B. (1982). *On the theory of delegation.* Northwestern University.

Jackson, M. O., & Sonnenschein, H. F. (2007). Overcoming incentive constraints by linking decisions. *Econometrica*, *75*(1), 241–257.

Li, J., Matouschek, N., & Powell, M. (2017). Power dynamics in organizations. *American Economic Journal: Microeconomics*, *9*(1), 217–41.

Lipnowski, E., & Ramos, J. (2018). Repeated delegation. *Working paper*.

Malenko, A. (2018). Optimal dynamic capital budgeting. *Working paper*.

Möbius, M. (2001). Trading favors. *Working paper*.

Rantakari, H. (2012). Employee initiative and managerial control. *American Economic Journal: Microeconomics*, *4*(3), 171–211.

Rantakari, H. (2017). Relational influence. *Working paper*.

Tate, R. (2012). *The 20% doctrine: How tinkering, goofing off, and breaking the rules at work drive success in business.* HarperCollins.

Appendices

A Proofs from the main text

Proof of Lemma 4

Proof. We know from Lemma 3 that $(\text{IC})_l$ binds in the optimal mechanism. Rearranging $(\text{IC})_l$ and substituting in the principal's expected profit function we get:

$$\begin{aligned}
&p_1(h-c)\big[a_1(h)+a_1(h)(1-h)\delta a_2(h,F)+(1-a_1(h))\delta a_2(h,\varnothing)\big]\\
&+(1-p_1)\big[a_1(h)l+\delta(p_2h+(1-p_2)l)(a_1(h)(1-l)a_2(h,F)+(1-a_1(h))a_2(h,\varnothing))\big]\\
&-(1-p)c\big[a_1(l)+a_1(l)(1-l)\delta a_2(l,F)+(1-a_1(l))\delta a_2(l,\varnothing)\big]
\end{aligned}$$

It is now easy to verify that

$$\frac{\partial \pi_P}{\partial a_1(l)} = -(1-p_1)c\big[1+\delta(1-l)a_2(l,F)-\delta a_2(l,\varnothing)\big] < 0$$

since $1 > \delta a_2(l,\varnothing)$. Now, given the linearity of the profit function in $a_1(l)$, it is immediate that the principal should set $a_1^*(l) = 0$ in the optimal mechanism.

Now, substitute $a_1^*(l) = 0$ and $(1-a_1(h))a_2(h,\varnothing) = a_2(l,\varnothing) - a_1(h)\frac{l}{\delta(p_2h+(1-p_2)l)} - a_1(h)(1-l)a_2(h,F)$ from $(\text{IC})_l$ in the original profit function and maximize with respect to $a_2(h)$:

$$\frac{\partial \pi_P}{\partial a_1(h)} = p_1(h-c)\left[1-\delta(h-l)a_2(h,F)-\frac{l}{p_2h+(1-p_2)l}\right]$$

which is always positive for $\delta < \bar{\delta} = \frac{p_2}{p_2h+(1-p_2)l}$. Thus, $a_1^*(h) = 1$.

Moreover, since $a_2(h,F), a_2(l,\varnothing)$ are numbers between 0 and 1, from $(\text{IC})_l$ we get the second sufficiency condition

$$\frac{l}{\delta(p_2h+(1-p_2)l)} < 1 \implies \delta > \underline{\delta} = \frac{l}{p_2h+(1-p_2)l}.$$

■

Proof of Proposition 2

Proof. Substitute $a_1^*(h) = 1$ and $a_1^*(l) = 0$ in the IC conditions and the principal's profit function. This reduces the principal's optimization problem to

$$\begin{aligned}
\underset{a_2(h,F),\, a_2(l,\varnothing)}{\text{maximize}} \quad & p_1(h-c)\big[1 + (1-h)\delta a_2(h,F)\big] \\
& + (1-p_1)\delta(p_2 h + (1-p_2)l - c)a_2(l,\varnothing) \\
\text{subject to} \quad & 1 + \delta(1-h)a_2(h,F) > \delta a_2(l,\varnothing) && (\text{IC})_h \\
& a_2(l,\varnothing) = \frac{l}{\delta(p_2 h + (1-p_2)l)} + (1-l)a_2(h,F) && (\text{IC})_l
\end{aligned}$$

First, note that $(\text{IC})_h$ must be slack at the optimum. Second, by the fact that we have already assumed $\delta < \bar{\delta}$, it is easy to verify that $a_2(h,F) < 1$.

Now, given the linearity of the profit function and the IC constraints it is easy to derive the optimal mechanism by comparing the slopes of the isoprofit lines and $(\text{IC})_l$. In the $(a_2(h,F), a_2(l,\varnothing))$ space, the slope of the isporofit line is given by

$$\frac{-p_1(h-c)(1-h)}{(1-p_1)(p_2 h + (1-p_2)l - c)}$$

and the slope of $(\text{IC})_l$ is $1-l$.

- If $p_2 h + (1-p_2)l > c$, then the slope of isoprofit line is negative. Moreover, a higher profit is a shift of the isoprofit line to the right. This implies that in the optimal mechanism $a_2^*(l,\varnothing) = 1$ and $a_2^*(h,F) = \frac{1}{1-l}\left(1 - \frac{l}{\delta(p_2 h + (1-p_2)l)}\right)$.

- If $ph + (1-p)l < c$, then the slope of the isoprofit line is positive and a higher profit is a shit of the line to the right and down. Two cases are possible depending on the comparison of slopes

$$\frac{p_1(h-c)(1-h)}{(1-p_1)(c - p_2 h + (1-p_2)l)} \lesseqgtr 1-l$$
$$\Longrightarrow c \lesseqgtr \bar{c} = \frac{p_1 h(1-h) + (1-p_1)(1-l)(p_2 h + (1-p_2)l)}{1 - (p_1 h + (1-p_1)l)}.$$

- If $c < \bar{c}$, then the isoprofit line is steeper than $(\text{IC})_l$. The optimal mechanism as in the case above, i.e. $a_2^*(l,\varnothing) = 1$ and $a_2^*(h,F) = \frac{1}{1-l}\left(1 - \frac{l}{\delta(p_2 h + (1-p_2)l)}\right)$.

- If $c > \bar{c}$, then the isporofit line is flatter than $(\text{IC})_l$. The optimal mechanism now is $a_2^*(l,\varnothing) = \frac{l}{\delta(p_2 h + (1-p_2)l)}$ and $a_2^*(h,F) = 0$.

This completes the proof of Proposition 2 and Corollary 1. ■

Proof of Proposition 3

Proof. Begin by rewriting the value to the principal of always granting authority to the agent. From Lemma 1 we know that it is equal to

$$p_1 h + (1-p_1)l - c + \delta\big[p_1(1-h)(h-c) + (1-p_1)(1-l)(p_2 h + (1-p_2)l - c)\big].$$

Using the fact that $\bar{c} = \frac{p_1 h(1-h)+(1-p_1)(1-l)(p_2 h+(1-p_2)l)}{1-(p_1 h+(1-p_1)l)}$, we can simplify the above as

$$p_1 h + (1-p_1)l - c + \delta[1-(p_1 h + (1-p_1)l)](\bar{c} - c). \tag{A.1}$$

Consider first the situation of $c = \bar{c}$. From Proposition 2, we know that

$$a_1^*(h) = 1\,,\; a_1^*(l) = 0\,,\; a_2^*(h, F) = 0 \;\text{ and }\; a_2^*(l, \varnothing) = \frac{l}{\delta(p_2 h + (1-p_2)l)} < 1,$$

and any $a_2^*(h, \varnothing) \in [0,1]$ and $a_2^*(l, F) \in [0,1]$ is an optimal interior mechanism. At this optimal mechanism, the value to the principal is

$$p_1(h-c) + (1-p_1)l\left(1 - \frac{c}{p_2 h + (1-p_2)l}\right). \tag{A.2}$$

Comparing the value to the principal from (A.1) with (A.2), it is easy to see that the latter is better. Therefore, at $c = \bar{c}$, the principal prefers the optimal interior mechanism.

It also now immediately follows, that for the case $c > \bar{c}$, either the interior mechanism is optimal or the extreme where no one is granted authority is optimal. The extreme mechanism where the agent always gets authority can not be optimal anymore since such a mechanism performs worse than the extreme one outlined above. To check whether it is optimal to grant no authority we need to check if

$$p_1(h-c) + (1-p_1)l\left(1 - \frac{c}{p_2 h + (1-p_2)l}\right) > 0.$$

Simplifying the above gives the condition,

$$c < \frac{(p_1 h + (1-p_1)l)(p_2 h + (1-p_2)l)}{p_1 p_2 h + (1-p_1 p_2)l} := \bar{\bar{c}}.$$

Finally, we now need to determine what happens when $c < \bar{c}$. Observe that at $c = 0$, the extreme mechanism where the agent always gets authority is better than the interior mechanism. Reminding ourselves that the optimal mechanism from Proposition 2 is given by

$$a_1^*(h) = 1\,,\; a_1^*(l) = 0\,,\; a_2^*(h,F) = \frac{1}{1-l}\left(1 - \frac{l}{\delta(p_2 h + (1-p_2)l)}\right) < 1 \;\text{ and }\; a_2^*(l,\varnothing) = 1,$$

and any $a_2^*(h,\varnothing) \in [0,1]$ and $a_2^*(l,F) \in [0,1]$. The value to the principal is

$$p_1 h\left(1 - \frac{1-h}{1-l}\frac{l}{\tilde{c}_2}\right) + \delta\left(p_1 h \frac{1-h}{1-l} + (1-p_1)\tilde{c}_2\right) \tag{A.3}$$

where $\tilde{c}_2 = p_2 h + (1-p_2)l$. Also, at $c = 0$, the value to the principal from the extreme mechanism is given by

$$p_1 h + (1-p_1)l + \delta(p_1 h(1-h) + (1-p_1)\tilde{c}_2(1-l)). \tag{A.4}$$

Comparing (A.4) with (A.3) we get that the extreme mechanism gives a higher value if

$$1 - p_1 + p_1 h \frac{1-h}{1-l}\frac{1}{\tilde{c}_2} > \delta\left((1-p_1)\tilde{c}_2 + p_1 h \frac{1-h}{1-l}\right),$$

which is always verified because $\tilde{c}_2 < 1$ and under the assumption $p_2 < \frac{l}{1-(h-l)}$, we have that $\delta < 1$.

Combining the above with the fact that the interior optimal mechanism is better for the principal at $c = \bar{c}$ and that the value is a linear (decreasing) function of c, it must be that there exists a $\underline{c} \in (0, \bar{c})$ where

- for $c < \underline{c}$, the principal prefers the extreme mechanism where the agent always gets the authority, and
- for $c \geq \underline{c}$, the principal prefers the optimal interior mechanism.

This completes the proof. ■