Dynamic Nonmonetary Incentives

Daniel Bird (Northwestern)

7 January 2016, 11:15-12:30, bld. 72, room 465

Abstract: We consider a dynamic principal-agent environment in which investments and rewards (compensation opportunities) arrive stochastically over time, and are either taken immediately or foregone. The agent privately ob- serves whether an action is currently available, but he needs the principal's consent to take it. We show that there exists a unique optimal mechanism and analyze its qualitative properties. As the agent's promised compen- sation increases, the principal gradually becomes more selective about incentivizing investments and less selective about the rewards that she allows. Interestingly, the unique optimal use of each reward turns out to be via "time allowances"; that is, the principal allows the agent to enjoy all rewards that arrive before a certain point in time.

Dynamic Nonmonetary Incentives *

Daniel Bird Alexander Frug

November 13, 2015

Abstract

We consider a dynamic principal-agent environment in which investments and rewards (compensation opportunities) arrive stochastically over time, and are either taken immediately or foregone. The agent privately observes whether an action is currently available, but he needs the principal's consent to take it. We show that there exists a unique optimal mechanism and analyze its qualitative properties. As the agent's promised compensation increases, the principal gradually becomes more selective about incentivizing investments and less selective about the rewards that she allows. Interestingly, the unique optimal use of each reward turns out to be via "time allowances"; that is, the principal allows the agent to enjoy all rewards that arrive before a certain point in time.

1 Introduction

We study a dynamic principal-agent environment in which short-lived investment and compensation opportunities arrive stochastically over time. The agent privately observes the arrival of the investment opportunities. Pursuing investments is costly to the agent, and so the principal must use the randomly arriving short-lived compensation opportunities to compensate him for doing so. (For example, a manager can compensate her employee by allowing him to perform personal errands during the work day.) We show that in an environment with multiple types of investment projects and reward activities, where the principal

*We are grateful to Eddie Dekel, Bruno Strulovici, and Asher Wolinsky for valuable suggestions and comments. We also appreciate the helpful comments of seminar participants at Northwestern University, Tel Aviv University, and Universitat Pompeu Fabra.

perfectly observes past actions (but not which opportunities arose), there exists a unique optimal mechanism and study its qualitative properties.

Consider a situation with one type of reward activity. Even in this simple case the optimal form of compensation is not obvious. For example, the agent may be allowed to pursue a fixed number of reward activities (no matter how long it takes to do so), or enjoy all reward activities that arrive in a fixed time interval (no matter how many of them there are). The principal may begin compensation immediately ("front-loading") or she may delay compensation for as long as possible ("back-loading"). We show that the unique optimal way to compensate the agent is to allow him to enjoy all reward activities that arrive in a fixed time interval that begins at the present moment. Section 3.1 is devoted to an intuitive discussion of this result.

In the more complicated environment with multiple types of investments and compensation opportunities, the principal has to decide on how jointly to use the different reward activities and how to determine which investment opportunities are worth pursuing. Our main result shows that in such environments, there is a unique optimal mechanism under which the principal selectively incentivizes investments and the agent's compensation derives from the freedom to pursue each reward activity during the time interval specified for that type of reward. We call this mechanism the *generalized time mechanism*.

The generalized time mechanism exhibits several notable economic properties. Firstly, the set of allowed rewards depends on the magnitude of the principal's debt to the agent and thus changes over time. When the principal's debt to the agent is high, the principal permits many reward activities. As the debt decreases, the principal gradually reduces the set of rewards that she allows. Notably, the principal permits the use of rewards with a high cost of providing a util even when it is not necessary to do so. That is, even though the principal can fully compensate the agent by permitting him to enjoy only cheap reward activities, she may allow him to enjoy more expensive rewards as well.

Secondly, the principal's investment strategy changes over time. Periods with many completed investments are followed by periods in which the principal incentivizes only high-quality investments. If no such investments arrive, the principal gradually becomes less selective about the investments she is willing to incentivize. Consequently, she may incentivize an investment she previously chose to forgo. In addition, the principal eventually allows the agent to enjoy all rewards indefinitely, at which point she will be unable to incentivize further investments.

Thirdly, our analysis specifies the optimal allocation of risk between the two players. The risk associated with the uncertain arrival of reward activities is borne by both players: the agent does not know the actual value of his compensation and the principal is uncertain about the cost of providing it. The players could easily co-insure against this risk by agreeing on a discounted number of rewards that are to be enjoyed, but doing so would be inconsistent with optimality.

In this paper, we abstract away from additional adverse selection and moral hazard problems by assuming that investment projects and reward activities arrive according to independent Poisson processes. In particular, this assumption implies that the agent cannot fabricate reward activities and that both players always have the same beliefs about the availability of actions in the future.

This work is related to three strands of literature. We contribute to the dynamic principal-agent literature (e.g., Rogerson (1985), Holmström and Milgrom (1987), Spear and Srivastava (1987), and Sannikov (2008)) by analyzing an environment in which there is uncertainty over the availability of investment and compensation opportunities. This kind of uncertainty is the focus of the "trading favors" literature (e.g., Möbius (2001) and Hauser and Hopenhayn (2008)), which studies equilibrium behavior in dynamic games where each player occasionally has the ability to grant a favor to his counterpart at a cost to himself. Dynamic settings with no monetary transfers are the focus of the recent literature on dynamic delegation (e.g., Guo and Hörner (2015), Lipnowski and Ramos (2015), and Li, Matouschek, and Powell (2015)). We contribute to this literature by analyzing an environment where the principal has full commitment power and the players disagree on the desirability of any action. A detailed literature review is provided in Section 6.

The paper proceeds as follows. Section 2 introduces the model and presents some preliminary results. In Section 3 we highlight the economic intuitions behind our main result in a model with a single type of investment project and a single type of reward activity. In Section 4 we prove that there exists a unique optimal mechanism and study its qualitative properties. In Section 5 we discuss the use of nonmonetary rewards in general and consider an environment where both monetary and nonmonetary incentivization is available. Section 6 offers a review of the related literature. Section 7 concludes. All proofs are relegated to the Appendix.

2 Model

We consider an infinite-horizon continuous-time mechanism design problem in which a principal (she) incentivizes an agent (he) to implement investments via nonmonetary rewards. The following is a formal presentation of the model.

Investment projects and reward activities arrive according to independent Poisson processes. There are $I \in \mathbb{N}$ types of investment projects: investment $i \in I$ has an arrival rate μ^i, its implementation incurs a loss of l^i for the agent, and it generates a benefit of B^i for the principal. Without loss of generality we order the types of investment projects such that the relative benefit, $\frac{B^i}{l^i}$, is weakly decreasing in i. Similarly, there are $J \in \mathbb{N}$ types of reward activities: reward $j \in J$ has an arrival rate λ^j, generates a gain of g^j for the agent, and entails a cost of C^j for the principal.[1] Again, without loss of generality, we order the reward types such that the ratio $\frac{C^j}{g^j}$ is weakly increasing in j. All investments and rewards can be "scaled down" and implemented at an intensity of[2] $\alpha \in [0, 1]$.

We assume that both players discount the future using the same (strictly) positive discount factor, r. Thus, an infinite history, in which investment project

[1] All payoff parameters are non-negative.

[2] When players are expected-utility maximizers and a public randomization device is available, assuming that actions can be scaled down is equivalent to assuming that the principal can commit to approving actions probabilistically.

i (reward activity j) is performed at times t_n^i (t_n^j) at intensities α_n^i (α_n^j), induces the principal's value

$$\sum_{i \in I} \sum_{n \in \mathbb{N}} e^{-rt_n^i} B_i \alpha_n^i - \sum_{j \in J} \sum_{n \in \mathbb{N}} e^{-rt_n^j} C_j \alpha_n^j$$

and the agent's utility

$$\sum_{j \in J} \sum_{n \in \mathbb{N}} e^{-rt_n^j} g_j \alpha_n^j - \sum_{i \in I} \sum_{n \in \mathbb{N}} e^{-rt_n^i} l_i \alpha_n^i$$

We assume that both players are expected-utility maximizers, and refer throughout to the principal's value and the agent's utility as the expectation of these variables given the players' current information.

We assume that only the agent observes whether investments and rewards are available; however, at each point in time he can take only actions that are both available (by nature) and allowed by the principal.[3] To specify what actions are allowed at a given point in time we use delegation lists. Specifically, a delegation list is a vector of the form

$$D = D^{inv} \times D^{rew} \in [0,1]^I \times [0,1]^J$$

where the $k-th$ coordinate of D^{inv} (D^{rew}) is the intensity at which the $k-th$ investment project (reward activity) is allowed. When an action is permitted at intensity zero, we say that it is forbidden. Note that without loss of generality we consider delegation lists that are "tight" in the sense that there is a unique intensity at which each investment and reward activity is allowed.[4]

The public information at time t consists of the delegation list at time t and the agent's action (if any) at time t . A public history at time t is then given by the function h_t, which describes the public information for each $s \in [0, t)$. We assume that the principal has full commitment power and can choose, at the beginning of the interaction, any (measurable) delegation function that maps

[3] Our results remain unchanged if the principal observes the availability of reward activities. See Section 3.3 for a discussion on observable investment opportunities.

[4] A fully general specification would allow the agent to implement projects and rewards at multiple intensities. However, if doing that is useful, it must be the case that the agent is indifferent between the different intensities. In that case the principal may (randomly) select the intensity for the agent, and permit only a single intensity.

public histories into delegation lists. A deterministic delegation mechanism is a (measurable) delegation function. Stochastic delegation mechanisms are generated by public randomizations over deterministic delegation mechanisms.

2.1 Markovian Solution

In the stationary environment of this model it is well known that attention can be restricted to mechanisms that use the agent's expected continuation utility, u, as a state variable.[5] Observe that, for the agent, the set of continuation utilities is bounded below by 0, since he can always conceal all future investment projects, and bounded from above by his expected utility from enjoying all future rewards without carrying out any investment projects:

$$\bar{u} \equiv \int_0^\infty e^{-rt} \sum_{j \in J} g^j \lambda^j dt = \frac{\sum_{j \in J} g^j \lambda^j}{r}$$

Any continuation utility in the interval $[0, \bar{u}]$ is feasible.

To specify a Markovian delegation mechanism we need to define a delegation function specifying the delegation list for every possible continuation utility, $D(u)$, and the "law of motion," du, according to which the agent's continuation utility changes. Aside from mild technical measurability issues, we would like not to impose any restrictions on the process du. Clearly, the agent's continuation utility can be updated upon implementation of investments or rewards, and so we allow for jumps in u in such cases. Furthermore, we do not restrict ourselves to deterministic jumps in the continuation utility. In addition, even if no action has been taken we allow the principal to perform (mean-preserving) lotteries over the agent's continuation utility. Finally, we allow for a drift in the stochastic process du, which in the absence of jumps corresponds to the continuous change of u.

Formally, a Markovian delegation mechanism is defined by a delegation function, $D(u)$, and a stochastic process, $u_t \in [0, \bar{u}]$, with a starting value of u_0. The

[5] See Spear and Srivastava (1987).

dynamics of u is given by

$$du = \eta(u)dt + \sum_{i\in I} \varphi_i^{inv}(u)dN_i^{inv} + \sum_{j\in J} \varphi_j^{rew}(u)dN_j^{rew}$$
$$+(1 - \max_{i\in I, j\in J}\{dN_i^{inv}, dN_j^{rew}\})\varphi(u)dN^{U^*} \tag{1}$$

The drift of the process at u is given by $\eta(u)$, while all other terms are related to jumps in the process. The counting process N_i^{inv} (N_j^{rew}) counts the number of times the $i-th$ investment ($j-th$ reward) is implemented, while $\varphi_i^{inv}(u), \varphi_j^{rew}(u)$ are random variables that generate stochastic jumps in u when an action is taken in state u. The countable set in which the principal initiates lotteries independently of the agent's actions is denoted by U^*, and the distribution of these lotteries is given by the random variables $\varphi(u)$, whose support is contained in[6] $[-u, \bar{u} - u]\backslash U^*$. The counting process N^{U^*} counts the number of times u enters the set U^*. We assume that all the above random variables are independent of each other.

2.2 The Principal's Problem

The principal's objective is to choose a delegation function and a stochastic process for the agent's continuation utility that maximize her expected value at time zero:

$$\sup_{D(u),du,u_0} \mathbb{E}[\int_0^\infty e^{-rt}\Big(\sum_{i\in I} D_i^{inv}(u_t)B^i dN_{i,t}^{inv} - \sum_{j\in J} D_j^{rew}(u_t)C^j dN_{j,t}^{rew}\Big)dt] \tag{OBJ}$$

where the dynamics of the process u is given by equation (1) and is subject to the promise-keeping constraint:

$$u_s = \mathbb{E}[\int_s^\infty e^{-r(t-s)}\Big(\sum_{j\in J} D_j^{rew}(u_t)g^j dN_{j,t}^{rew} - \sum_{i\in I} D_i^{inv}(u_t)l^i dN_{i,t}^{inv}\Big)dt] \tag{PK}$$

Moreover, the chosen mechanism must be incentive compatible:

$$\mathbb{E}[\varphi_i^{inv}(u)] - D_i^{inv}(u)l^i \geq 0 \quad \forall u \in [0, \bar{u}], i \in I \tag{IC_{inv}}$$

[6] We assume that the support of these jumps does not contain any elements from U^* in order to prevent multiple instantaneous jumps. This assumption is without loss of generality, as any sequence of jumps is a compound lottery that can be reduced.

$$\mathbb{E}[\varphi_j^{rew}(u)] + D_j^{rew}(u)g^j \geq 0 \quad \forall u \in [0, \bar{u}], j \in J \qquad (IC_{rew})$$

We do not need to add an explicit IR constraint since by concealing all investments, the agent guarantees himself a non-negative expected utility. Thus, any IC mechanism is also interim-IR.

We denote the principal's value function corresponding to the solution of the above problem by $V(u)$. This value function is weakly concave due to the existence of a public randomization device.

2.3 Preliminary Observations

Since the agent has no private information at the start, in a world with money transfers she will obviously get a zero payoff under an optimal mechanism. This holds true here even in the absence of upfront transfers, since at no point does the principal need to offer expected compensation in excess of the agent's cost. Due to the lack of transfers, it is impossible to reduce the agent's utility by a uniform decrease of compensation, which, trivially, does not affect any IC constraint. However, under any mechanism in which the agent's participation utility is positive, there exist on-path public histories in which the principal can reduce the intensities of the allowed reward activities without affecting the agent's IC constraints. This is formalized by the following lemma.

Lemma 1. *In an optimal mechanism, the agent's expected utility is zero,* $u_0 = 0$.

An immediate corollary of this result is that the expected increase in the agent's continuation utility when he implements an investment project equals his cost of implementation.

Corollary 1. *Under an optimal mechanism for all* $i \in I$ *and* $u \in [0, \bar{u}]$,

$$\mathbb{E}[\varphi_i^{inv}(u)] = l^i D_i^{inv}(u).$$

A key feature of this model is that incentivization devices are both limited and perishable. This makes the principal keen to use as many of them as possible. When an investment opportunity is available, the principal can commit some of her incentivization tools in order to incentivize it. When the principal has the opportunity to incentivize an investment, she forgoes this opportunity only if she prefers to save her resources so that she can incentivize better investments in the future. Consequently, an investment project of type 1, i.e., the project with the highest relative benefit, is implemented at full intensity, or the maximal possible intensity if full intensity is not incentive compatible. This is formalized by the following lemma.

Lemma 2. *In an optimal mechanism,*

$$D_1^{inv}(u) = \min\{1, \frac{\bar{u} - u}{l^1}\}$$

3 Special Case: One Investment, One Reward

In the above model, the principal faces a complicated multi-dimensional problem. In this section we illustrate and discuss the main property of the optimal mechanism. The essentially *unique* optimal mechanism for providing compensation via a given reward activity is to allow the agent to enjoy all rewards that arrive within a given time interval, effective immediately.[7]

To illustrate this property in the most transparent way we consider the simplest possible case of our model: the case with one type of investment project and one type of reward activity. Clearly, this drastically simplifies the problem and mutes many of its dimensions; e.g., the principal need not calculate the optimal combination of rewards, nor which investment project to incentivize if it becomes available (Lemma 2). However, even in this simple case, there are multiple ways in which the principal can compensate the agent. In addition to the proposed time-allowance solution, the principal can, for example, allow the agent to pursue a fixed number of reward activities, or to enjoy many rewards in the distant future. We begin with an intuitive presentation of the *time mechanism* (henceforth TM). While this mechanism is a Markovian delegation mechanism, it is instructive to start with an alternative representation that uses

[7]By "essentially unique," we mean that there is a unique tight delegation function up to changes in $D(u), du$ for a set of utilities in which the mechanism spends a measure zero amount of time. This might include open subsets of $[0, \bar{u}]$ if these values are off the path of play. For the rest of this paper we refer to an optimal mechanism as unique, even though there are other, essentially equivalent optimal mechanisms.

the length of time in which the agent is allowed to enjoy rewards as a state variable.

To simplify the exposition we omit project indexes. Moreover, to avoid the trivial case in which the principal cannot extract a positive surplus from any (incentive compatible) mechanism, we assume that $\frac{B}{l} > \frac{C}{g}$. To incentivize the agent to implement an investment opportunity the principal must increase his continuation utility by l. The cost of providing a util is $\frac{C}{g}$, and so the expected cost of incentivizing an investment opportunity is $l\frac{C}{g}$. Therefore, it is profitable to incentivize an investment only if $B > l\frac{C}{g}$.

Time Mechanism Let $s \in [0, \infty]$ denote the amount of time, starting at the present moment, in which the agent is allowed to enjoy reward projects. Set $s_0 = 0$; that is, the initial value of the principal's promise is zero. If an investment project arrives at state s, and there is $f(s) \in \mathbb{R}_+$ that solves the indifference condition

$$\int_0^{f(s)} e^{-rt}\lambda g dt - l = \int_0^s e^{-rt}\lambda g dt$$

then the principal allows an investment project at full intensity, and sets the new promise to $f(s)$. If there is no such $f(s) \in \mathbb{R}_+$, the principal allows the investment project at the intensity α that solves

$$\int_0^{\infty} e^{-rt}\lambda g dt - \alpha l = \int_0^s e^{-rt}\lambda g dt$$

and sets the new promise at[8] $s = \infty$. Given this representation, it is easy to transform TM back into the general language of Markovian delegation mechanisms that use the agent's continuation utility as a state variable.

[8]Simple algebra shows that as long as the current promise, s, is less than $\frac{\ln(\lambda g) - \ln(lr)}{r}$, investments are implemented at full intensity and the promise is increased by $\frac{\ln\left(\frac{g\lambda}{g\lambda - lre^{rs}}\right)}{r}$. If the promise is greater than the above threshold, upon implementation of an investment it is updated to $s = \infty$, and the investment project is executed at the maximal possible intensity, $e^{-rs}\frac{g\lambda}{lr}$.

Time Mechanism as a Markovian Delegation Mechanism

$$(1)\quad D^{inv}(u) = \min\{1, \frac{\bar{u}-u}{l}\}$$

$$(2)\quad D^{rew}(u) = \begin{cases} 1 & \text{if } u > 0 \\ 0 & \text{if } u = 0 \end{cases}$$

$$(3)\quad \varphi^{inv}(u) = \min\{l, \bar{u}-u\}$$

$$(4)\quad \varphi^{rew}(u) = \varphi(u) = 0$$

$$(5)\quad \eta(u) = \begin{cases} ru - \lambda g & \text{if } u > 0 \\ 0 & \text{if } u = 0 \end{cases}$$

$$(6)\quad u_0 = 0$$

In words, conditions (1) and (2) specify the delegation function $D(u)$: investments are always incentivized at the maximal possible intensity, and rewards are allowed at full intensity whenever the principal's debt (agent's continuation utility) is positive. Condition (3) states that upon implementation of an investment, the agent's continuation utility increases by exactly the cost of implementation. The main feature of TM is reflected in (4) and (5): the agent's continuation utility does not depend on the actual number of implemented rewards but drifts down continuously (as long as no investment is implemented). The initial condition (6) trivially corresponds to $s_0 = 0$.

3.1 Optimality of the Time Mechanism

Formally, the optimality of TM follows as a special case of the main result of this paper, given in Proposition 2. The simplicity of the $J = I = 1$ case allows us to explain this result with three simple observations that apply basic economic concepts.

Our first observation relates to the nature of incentives. In a standard mechanism design problem where monetary compensation is not subject to capacity constraints, the principal is indifferent to the timing of payments with a given present value.[9] By contrast, in our environment the principal has a strict preference to compensate the agent in a timely manner since incentives are perishable. To see this, note that the *direct* cost of providing the agent u utils via reward

[9]Recall that we assume that the players share the same discount factor.

activities is $\frac{C}{g}u$. If the principal were offered a one-time opportunity to repay her debt of $u > 0$ at a cost $\frac{C}{g}u$, she would strictly benefit from accepting this offer. By paying the debt immediately, the principal would incur an identical direct cost, but, in addition, she would reclaim committed rewards, and thus increase the expected number of future investments.

The second observation is that the expected exploitation of rewards within a time interval is maximized under TM. Specifically, let $u > 0$, and denote by s the corresponding time allowance under TM. Let τ be the waiting time until the next investment project arrives (note that τ is a random variable). Obviously, there exist mechanisms for which, after certain histories, the agent's continuation utility at time τ, u_τ is lower than it would be under TM. However, TM induces the minimal value of $\mathbb{E}[u_\tau]$ over all IC mechanisms that satisfy condition (PK).

For simplicity, consider the following two cases separately. The obvious case corresponds to realizations where $\tau \geq s$. Under TM, the agent's continuation utility at time τ equals zero, which is the lower bound for any mechanism. Now consider the case where $\tau < s$. By the definition of TM, *all* reward activities arriving before time τ are allowed. On the other hand, for any mechanism under which the compensation policy is not completely independent of the agent's actions before τ, there exists a positive measure of histories in which some reward activities are forbidden before time τ. Therefore, the utility the agent receives (in expectation) from rewards before time τ is strictly higher under TM than under any other mechanism.

Our third observation is that, by the construction of TM, for any value of τ there is no uncertainty regarding the agent's continuation utility u_τ.

To summarize, among all IC mechanisms that satisfy condition (PK), TM guarantees the maximal expected reduction in the agent's continuation utility before the next investment project arrives, and it does so with no variance. Moreover, if an investment arrives before the debt is fully repaid under TM, any other mechanism (that differs from TM before the arrival of an investment) reduces the principal's expected debt by less than TM. Thus, only a strictly risk-loving (over u) principal would consider using a mechanism other than TM. However, since the principal's value function, $V(u)$, is (weakly) concave, TM is the unique optimal mechanism. This is formalized by the following proposition.

Proposition 1. *The time mechanism is the unique optimal mechanism.*

Proof. This proposition is a special case of Proposition 2. □

3.2 Discussion of the Time Mechanism's Properties

Front-loading The principal uses the earliest available time interval to compensate the agent. By forbidding available rewards when $u > 0$, the principal is clearly acting sub-optimally as she cannot retroactively use these rewards to compensate the agent later. This property stands in contrast to the standard practice of back-loading, which is the recommended policy in many settings where transfers are allowed.[10] This compensation policy enables us to define a notion of the resources that the principal has at her disposal. If, at state s, the principal were required to marginally increase the agent's continuation utility, she would do so by allowing him to enjoy all rewards that arrive in an infinitesimal time interval starting s units of time in the future. Thus, the marginal resource at state s is the right to enjoy all reward activities in that interval.

The front-loading of compensation has a more subtle statistical manifestation in the use of time allowances to provide compensation. If a compensation scheme allows for two realizations of rewards that differ in the time period it would take to fully repay the debt, the risk of wasting resources is higher in the realization with earlier debt repayment. Thus, it is sensible to shift compensation between the two realizations such that a fraction of the compensation that should be paid in the distant future in one realization is moved to an earlier point in time in the other realization. One can shift compensation between different realizations of rewards only at the ex-ante stage. An optimum is reached by a complete disentangling of the continuation utility from the actual realization of rewards. Thus, it can be thought of as a stronger, statistical form of front-loading.

[10] See, for example, Harris and Holmström (1982), Lazear (1981), and Ray (2002).

Value Function The principal's value at u is determined by the cost of paying the current debt and by the amount of resources that will be wasted in expectation. In the limiting case of always-available investment projects ($\mu = \infty$), the principal will successfully use all her resources either to pay the existing debt or to incentivize additional investment projects. Therefore, the value function has a linear form given by

$$V_\infty(u) = \frac{B}{l}(\bar{u} - u) - \frac{\lambda C}{r}$$

In the opposite limiting case of no further investment projects ($\mu = 0$), only the cost of compensation remains. Thus, the linear value function is given by

$$V_0(u) = -\frac{C}{g}u$$

For any intermediate value of μ, the value function is strictly decreasing and lies inside the wedge created by the extreme value functions. In Appendix B we provide an algorithm for analytically deriving the value function.

The value obtained from successfully utilizing a resource is constant; however, the probability of successfully utilizing a resource depends on the state (for intermediate values of μ). This induces a strictly concave value function despite the risk neutrality of both players and the linearity of the environment.[11] At the beginning of the interaction (and whenever $s = 0$) there is no compensation until an investment project is carried out and, therefore, the marginal resource is wasted. By contrast, when $s > 0$ the marginal resource will be wasted only if no investment project arrives in s units of time. Thus, the probability of successfully utilizing the marginal resource is increasing in s, which in turn implies that the value of the marginal resource is increasing in s and that the value function is strictly concave in s. Due to discounting, the amount of time required to provide u utils via TM is convex in u, and the value function is strictly concave in u.

Risk-Bearing Compensation under TM is "tailored" to the risk neutrality (with respect to the implementation of any action) of both players. In TM the players are exposed precisely to the exogenous risk generated by the stochastic arrival of reward activities. They do not co-insure against this risk, as would happen if the principal compensated the agent by allowing him to enjoy a fixed

[11]We provide a general proof of this statement in Lemma 7.

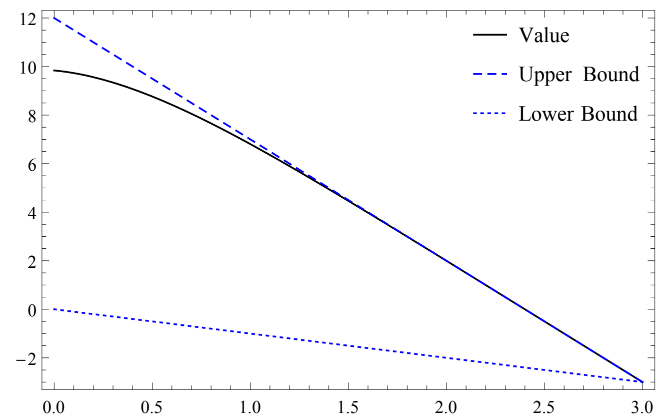

Figure 1: Example of the value function and its bounds for $r = \frac{1}{20}, \mu = \frac{5}{20}, \lambda = \frac{3}{20}, B = 5, l = g = C = 1$.

number of discounted reward activities, nor do they generate additional risk by using explicit lotteries.

Foregone Investment Opportunities Since monetary transfers are not available, our parametric assumptions do not imply that all investment projects are implemented. However, by the Borel–Cantelli lemma, the agent's continuation utility reaches $\bar{u}$ in a finite time with probability one. In this absorbing state, the agent's continuation utility never decreases and no further investments are implemented. Note that reaching this absorbing state is an ex-ante desirable outcome for the principal, since it implies that she has (efficiently) used all of her resources to incentivize investments. The time it takes to reach this state is decreasing in the arrival rate of investment projects (μ) and increasing in the principal's capacity to compensate the agent ($\bar{u}$). This suggests that both the arrival rate of investment projects and the frequency of reward activities are positively correlated with the welfare loss resulting from the lack of transfers. Thus, the combination of both parameters can be used as a (rough) indicator of the magnitude of welfare loss due to lack of transfers.[12]

[12]There is no non-trivial bound on the welfare loss due to lack of transfers. One can easily construct examples in which the percentage of implemented discounted investments approaches one or zero.

3.3 Symmetric Information: Observability of Investments

Under TM the agent implements all investment projects as long as the principal has the ability to provide compensation. Furthermore, the agent's actions are perfectly observed by the principal and he receives no information rent. This raises a question about the effect of asymmetric information on the principal's ability to extract surplus. In this section, we demonstrate that the information asymmetry is detrimental to the principal as it decreases her ability to use her limited incentivization resources effectively.

Consider an environment identical to the one described above in which the realization of projects is observed by both players.[13] In this environment, the principal must still incentivize the agent to implement investments, but she can also punish him for failing to do so. To illustrate the difference between the two cases, assume that the parameters of the model are such that the agent's expected utility from implementing all future investments and enjoying all future rewards is equal to the loss from implementing an investment opportunity, $\frac{g\lambda - l\mu}{r} = l$. If the principal applies TM, she will eventually commit all future rewards and not be able to incentivize further investments. Now consider the following mechanism: before the first investment opportunity arrives (phase 1), rewards are not allowed. When the first investment opportunity arrives, the mechanism moves to "phase 2," wherein the agent is allowed to enjoy all rewards, as long as he carries out all investment opportunities. If the agent fails to implement an available investment, future rewards are no longer allowed. By construction (and the stationariness of the environment) the agent is always indifferent whether to implement an investment opportunity or not. In other words, under the latter (incentive-compatible) mechanism all investments are implemented. Furthermore, as under TM, the agent is not overcompensated. Since more investment projects are carried out under this mechanism, it strictly outperforms TM.

The nature of compensation in this environment limits the principal's capacity to compensate the agent. However, as the previous example demonstrates, the degree to which the principal can utilize her limited capacity depends on the informational assumptions. When the principal observes which projects are

[13] This is the environment studied in Bird and Frug (2015).

available the optimal (incentive-compatible) use of the principal's limited capacity requires the principal to contract over the arrival of future investment opportunities; see Bird and Frug (2015). Such contracts allow the principal to combine histories in which investment opportunities are rare and the capacity constraint is nonbinding with other histories in which investments are abundant and the capacity constraint is binding. Under asymmetric information, i.e., when the principal cannot observe which projects are available, such contracts cannot be enforced, and thus the principal must use less efficient contracts (such as TM) that can be enforced despite her limited information. In other words, asymmetric information exacerbates the severity of the limited capacity problem.

It is worthwhile to note that if we use monetary transfers as the means of compensation in our model (instead of reward activities), the principal can reimburse the agent for the (exact) cost of implementing an investment project after observing a completed project. This compensation scheme is optimal under symmetric and asymmetric information alike. Therefore, the existence (or lack thereof) of an "informational problem" may depend on the means by which the agent is compensated.

4 The General Case

Our main result is that in the general environment presented in Section 2, there is a unique optimal mechanism under which the principal uses time allowances to compensate the agent for completed investments. Formally, we say that a Markovian delegation mechanism is a *multidimensional time mechanism* if for every agent's continuation utility u the principal's compensation can be represented in the form of J time allowances $(s^j)_{j\in J}$, such that the agent is allowed to enjoy all reward activities of type $j \in J$ that arrive in the next $s^j \geq 0$ units of time.

Proposition 2. *There is a unique optimal mechanism. Moreover, this mechanism is a multidimensional time mechanism.*

We refer to the unique optimal mechanism as the *generalized time mechanism* (henceforth GTM).

The combination of Proposition 2 and Lemma 2 pins down the optimal mechanism only for the $I = J = 1$ case. For the general case, the principal must also decide which investment opportunities to incentivize and how to combine different reward activities to compensate the agent. One particular implication of Proposition 2 is that explicit lotteries, whether they be mean-preserving gambles over different values of the agent's continuation utility or lotteries that induce a probabilistic implementation of an available reward activity, are inconsistent with optimality. This property might be surprising given the risk neutrality reflected in the linearity of the players' payoff functions. We develop this point in the next section, where we discuss several qualitative properties of the optimal mechanism.

4.1 Properties of the Generalized Time Mechanism

For ease of exposition, we now make several mild assumptions. To focus on the nontrivial part of the problem, we assume that the sets of possible investment opportunities and reward activities are non-redundant in the sense that every investment opportunity, if compensated solely via the most efficient reward activity, generates positive surplus, and that every reward activity is sufficiently efficient to generate positive surplus, if it is used to incentivize the most desirable investment project. Formally,[14]

$$\forall i \in I \quad \frac{B^i}{l^i} > \frac{C^1}{g^1}$$
$$\forall j \in J \quad \frac{C^j}{g^j} < \frac{B^1}{l^1}$$

Furthermore, we assume that the ratio $\frac{C^j}{g^j}$ is *strictly* increasing and that the rate of return on investments, $\frac{B^i}{l^i}$, is *strictly* decreasing. This assumption is made solely for ease of exposition and entails no loss of generality, as we show

[14] To incentivize the agent to implement an investment opportunity of type i the principal must increase his continuation utility by l^i. The cost of providing a util by using reward activity j is $\frac{C^j}{g^j}$, and so the expected cost of incentivizing investment opportunity i via rewards of type j is $l^i \frac{C^j}{g^j}$. Therefore, it is profitable to incentivize an investment in this way only if $B^i > l^i \frac{C^j}{g^j}$.

in Appendix C. We now present the notable properties of GTM.

Reward Usage In Proposition 2 we established that each reward activity should be offered to the agent in the form of time allowances. This enables us to frame the principal's considerations regarding the optimal bundle of rewards as an intuitive trade-off between the *speed* and the *cost* of compensation. In the previous section, we explained that early provision of compensation is desirable and, therefore, the principal is inclined to permit rewards of different types simultaneously. On the other hand, different types of rewards are associated with different costs for the principal. Thus, to reduce the *direct* cost of compensation, the principal might prefer to avoid permitting expensive rewards.

If a reward activity of type $\tilde{j}$ is allowed at u, more efficient rewards $j < \tilde{j}$ should also be allowed.[15] Therefore, for $j < \tilde{j}$, we must have $s^j \geq s_{\tilde{j}}$. This implies that there exist weakly increasing activation thresholds $\{\hat{u}_j^{rew}\}_{j=1}^J$ such that reward activity j is permitted when the agent's continuation utility u satisfies $u \geq \hat{u}_j^{rew}$.

A more interesting feature of compensation under GTM is that the above activation thresholds are generally interior in the following sense: (1) the activation thresholds are *strictly* increasing and (2) the principal permits inefficient rewards before it is strictly necessary to do so:[16] $\hat{u}_j^{rew} < \sum_{k=1}^j \frac{\lambda^k g^k}{r}$. This implies that in some states, even though the principal can incentivize an investment by allowing the agent to enjoy a cheap activity, she will incentivize him by allowing him to enjoy a more expensive one instead. Moreover, this implies that the principal does not permit the agent to enjoy any *single* reward activity indefinitely, until she is forced to allow him to enjoy *all* rewards indefinitely at the absorbing state of $u = \bar{u}$.

The intuition behind this property is as follows. The required time allowances to provide compensation with rewards $1, \ldots, j$ become arbitrarily long as the continuation utility approaches $\sum_{k=1}^j \frac{\lambda^k g^k}{r}$. If a (sufficiently good) invest-

[15] Recall that we can compare different types of rewards in terms of the direct cost of providing one util to the agent $\frac{C_j}{g_j}$, and that this ratio is strictly increasing.

[16] Recall that the agent's expected continuation utility from enjoying all future rewards of type j is $\int_0^\infty e^{-rt}\lambda^j g^j dt = \frac{\lambda^j g^j}{r}$.

ment opportunity arrives in this period of time, the principal will incentivize it by permitting less efficient reward activities (as most of the more efficient incentives are already committed). In other words, when $u \to \sum_{k=1}^{j} \frac{\lambda^k g^k}{r}$, the probability that the principal assigns to the event that she will use less efficient rewards in the future becomes arbitrarily close to one. Thus, she would rather allow the less efficient reward project immediately and reduce the waste of her limited incentivization tools. Conversely, when the agent's continuation utility is low enough, permitting only the most efficient reward will likely repay the principal's debt before the next investment opportunity arrives. Therefore, for sufficiently low levels of debt the principal minimizes the direct cost of compensation as her choice is unlikely to affect her capacity to incentivize future investments.

Proposition 3. *Activation thresholds of reward activities, $\hat{u}_j^{rew}$, are strictly increasing in j and satisfy $\hat{u}_j^{rew} < \sum_{k=1}^{j} \frac{\lambda^k g^k}{r}$.*

An important implication of Proposition 3 is that, as long as an investment project is not implemented, the agent's continuation utility drifts down for every $u \in (0, \bar{u})$. This technical property is useful in the discussion of the following important property.

Strict Concavity The GTM induces a strictly concave value function for the principal.[17] The driving force behind this property is the combination of the linearity of the players' utility (in the implemented actions), the principal's limited capacity to compensate the agent, and her uncertainty about the usage rate of this limited capacity.

Assume that the agent's continuation utility equals u and let u_1, u_2 be such that $u = \frac{u_1+u_2}{2}$. One way the principal can deliver a promise of u is by fictitiously splitting all projects and rewards into two halves and creating two (perfectly correlated) fictitious worlds, each of which contains half of every reward and half of every investment opportunity. She can then provide u_1 utils using GTM in one fictitious world, and u_2 utils using GTM in the other. Due to the linearity of payoffs, she generates a value of $\frac{V(u_i)}{2}$ in world i.[18] Moreover,

[17] This is proven formally in Lemma 7 in Appendix A.
[18] This shows directly that the value function is weakly concave.

with positive probability at some point in time the principal will reach a debt of zero in world 1. If $u_1 < u_2$, then at this point the debt in world 2 is strictly positive. Now the principal can benefit from temporarily merging the two fictitious worlds. Specifically, instead of wasting her incentivization resources in world 1, she would rather use the most efficient reward activity in world 1 to speed up compensation in world 2. This increases the speed of compensation and (weakly) decreases the cost of compensation, and thus the principal is strictly better off.

The strict concavity of the value function immediately implies that GTM does not induce explicit lotteries over the agent's continuation utility. Thus, the increase in the agent's continuation utility induced by the implementation of an investment is deterministic, and there are no stochastic jumps in the agent's continuation utility when no action is taken.

Investment Selectivity The last notable property is concerned with the principal's selection of investment opportunities. The use of time allowances for compensation implies that the probability of an investment opportunity arriving before the marginal resource (of any type of reward) is wasted is increasing in the agent's continuation utility. Therefore, as the risk of waste decreases (and the agent's continuation utility increases), the principal becomes more selective about the quality of the investments that she is willing to incentivize. Formally, there exist thresholds $\hat{u}_i^{inv}$, such that investment project i is incentivized if and only if $u \leq \hat{u}_i^{inv}$. The strict concavity of the value function implies that $\hat{u}_i^{inv}$ is strictly decreasing in i.

Proposition 4. *There exist strictly increasing thresholds $\hat{u}_i^{inv}$ such that:*

1. *If $u \leq \hat{u}_i^{inv} - l^i$, then investment opportunities of type i are incentivized at full intensity.*
2. *If $u \in (\hat{u}_i^{inv} - l^i, \hat{u}_i^{inv})$, then investment opportunities of type i are incentivized at intensity $\frac{\hat{u}_i^{inv} - u}{l^i}$.*

This proposition, combined with the observation that the agent's continuation utility drifts down as long as investments are not implemented, provides insights into the nature of dynamic investment decisions. Specifically, it suggests a potential explanation of why an investment that was forgone in the past is implemented at present. After a large and profitable investment is carried out and the agent's continuation utility increases accordingly, the principal requires a high return on her limited resources to justify the incentivization of investment opportunities. Therefore, less profitable investments are (temporarily) forgone (despite the principal having the resources to implement them), until the agent's continuation utility decreases and the return on such an investment is deemed acceptable.

5 Modes of Compensation

The majority of principal-agent models have focused on studying how the principal should condition the agent's compensation on the outcomes she can observe. The literature generally makes the natural assumption that compensation is provided via monetary transfers that can be made at arbitrary times and be of arbitrary size. When transfers are not the exclusive compensation tool, the principal's problem becomes more complicated since, in addition to the provision of sufficient incentives, she also needs to choose the optimal compensation bundle. In our setting, we assume multiple compensation tools with an important and natural feature: the tools do not enable instantaneous compensation, and so the design of optimal compensation is an even more complex dynamic problem than when monetary transfers are available.

Our results suggest that when comparing non-instantaneous compensation devices, two (main) features should be considered. They are, first, the *cost of compensation*, i.e., the direct cost of providing one util to the agent using a given compensation tool and, second, the *speed of compensation*, i.e., the length of time it takes to provide one util to the agent.

We believe that this framework could be valuable in future research on the choice of compensation devices. Questions that could be analyzed within this framework include: What restrictions should a senior manager impose on the incentivization tools that a junior manager has at his disposal? What compensation devices should be permitted under a collective bargaining agreement? How does one determine the number and allocation of nonmonetary rewards among heterogeneous employees?

Our proposed framework has a possible application in the study of the selection of a compensation device (more generally, the subset of devices) from a given set that a principal may use. There is a natural partial ranking over compensation devices: a compensation device that is both quicker and cheaper is obviously preferred. Moreover, any set of investment projects will induce indifference curves over the space of compensation devices, providing a method to complete their ranking. In general, when investment opportunities are rare, the principal is inclined to prefer cheap compensation devices as it is unlikely that the slow speed of compensation will prevent investment projects in the near future. On the other hand, when investment opportunities are abundant she expects to utilize a high percentage of the incentives at her disposal, and thus is inclined to increase the amount of incentives at her disposal even if doing so reduces the net benefit from each investment project. This suggests a natural graphical representation of the above trade-off in $\mathbb{R}^2_+$.

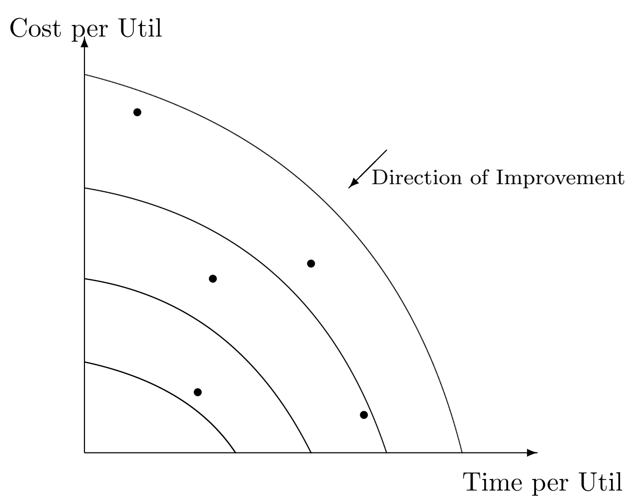

Figure 2: Possible Indifference Map for Selection between Five Devices.

The fact that there is a non-trivial trade-off between the two dimensions in the selection problem implies that the preferences regarding the choice of compensation devices are different from their activation order in our model. When choosing whether to permit a reward activity, the principal does not care about the speed of compensation as she is making a decision whether to give the agent a fixed amount of utils at a certain cost or not. However, when she is selecting which reward activity to use, the principal considers the speed of compensation, as slow compensation (even if cheap) limits the number of investment projects she can incentivize.

5.1 Adding Monetary Transfers

Our model provides insights into the structure of optimal compensation in an environment with multiple compensation tools. These insights remain valid in an environment where monetary transfers are allowed on top of other compensation devices. The point $(0, 1)$ represents transfers in the above framework.[19] The optimal mechanism in such environments is similar to GTM, with a few minor modifications.

Firstly, in contrast to GTM, the principal does not use reward activities for which the cost of providing a util to the agent is greater than one. To see this, note that the principal's ability to provide incentives via monetary transfers is unlimited and she is better off using transfers rather than more costly reward activities. Secondly, all rewards for which the cost of providing a util is strictly less than one are used and, moreover, transfers are used only when all such rewards are exhausted. By contrast, under GTM all compensation devices are used before any are exhausted. This difference reflects the unique nature of transfers as an instantaneous and non-perishable compensation device, a property that negates the need to start using transfers before it is strictly necessary to do so.

[19] We assume that all compensation devices except for transfers are non-instantaneous.

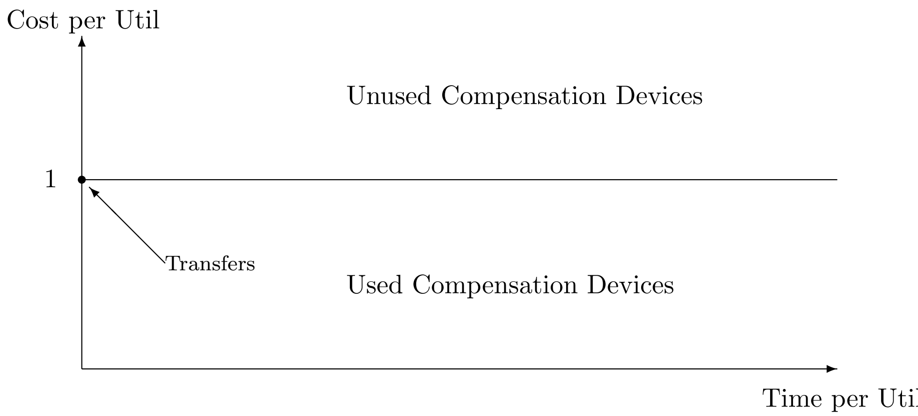

Figure 3: Used and Unused Compensation Devices when Transfers are Available.

6 Related Literature

Following the seminal work of Holmström (1977), the literature on delegation has focused on situations in which the interests of the principal and the agent are partially aligned. Initially, research focused on static settings. Melumad and Shibano (1991) and Alonso and Matouschek (2008) showed that in a static environment, sufficient alignment of interests is a necessary condition for delegation to be of value. In a more complex environment, where the agent simultaneously conducts multiple tasks on behalf of the principal, Frankel (2014) also requires (sufficient) alignment of preferences to derive the optimality of his "moment mechanism." Armstrong and Vickers (2010) consider a delegation problem where the agent has private information about the available actions, rather than information about a payoff-relevant state of nature. Their main result shows that the principal permits an action if it provides her with a high enough payoff relative to the agent's utility from performing the same action. This result demonstrates the essential role of partial preference alignment in a static version of the environment studied in this paper. In recent years, the literature on delegation has begun to analyze dynamic interactions where preferences are partially aligned. Examples of such papers include Guo and Hörner (2015), Lipnowski and Ramos (2015), and Li, Matouschek, and Powell (2015). The former two papers assume that the principal never observes the realized payoffs from previous actions, whereas the latter paper (and ours) assume that all past actions are observable. Secondly, the first paper (and ours) assume that the principal has full commitment power, whereas the last two papers assume that her commitment is limited. Several works of a more applied nature have also focused on this environment. Guo (2014) studies the delegation of experimentation when no transfers are permitted, and Nocke and Whinston (2010), Nocke and Whinston (2013) analyze the optimal dynamic merger policy for an antitrust authority.[20]

Our model is similar to that of Möbius (2001) and Hauser and Hopenhayn (2008), who study equilibria in games where each player occasionally has the ability to grant a favor to his counterpart at a cost to himself. Möbius (2001) suggests the use of a chip mechanism in which a player grants a favor if the (undiscounted) difference between the number of favors granted and received is not too large. Hauser and Hopenhayn (2008) show that the optimal perfect public equilibrium can be implemented by a modified chip mechanism in which the cost of receiving a favor depends on the current chip distribution and the number of chips held by each player reverts to the mean over time. Their modification resembles the GTM in that the principal grants the agent favors (rewards) for a limited amount of time. It differs from the GTM, however, in that, in the optimal equilibrium, granting a favor to the agent reduces his credit and the principal must grant favors to the agent even when she does not owe the agent anything.

Our work lies at the junction between the literature on delegation and the literature on the structure of compensation in dynamic principal-agent models. The latter literature addresses a plethora of economic questions in rich environments;[21] however, it does so by assuming compensation (solely) via monetary transfers. The complexity created by incentivization via nonmonetary devices forces us to consider a simple environment, indeed, so simple that a direct comparison between our model and previous dynamic principal-agent models is not insightful. However, two features of the GTM are frequently debated in this

[20]Authorization of mergers is a delegation problem since the anti-trust authority can commit to a merger approval policy but cannot give or receive payments from firms in order to incentivize or approve mergers.

[21]Seminal papers in this field include the works of Rogerson (1985), Holmström and Milgrom (1987), and Spear and Srivastava (1987).

literature, and to these we now turn.

The first debated feature of the GTM is the optimal timing of compensation. This is a long-standing question in economics. Earlier work such as Lazear (1981) and Harris and Holmström (1982) generally suggest that when the principal has full commitment power compensation should be back-loaded. Ray (2002) reaches the same conclusion when the principal does not have full commitment power. Later work on models of full commitment such as Rogerson (1985) and Sannikov (2008) shows that the optimal timing of compensation is more complicated as it relates to time-dependent effectiveness and the cost of providing incentives, in which case either back-loading or front-loading may be optimal. In the present paper, we add a novel argument to this debate by pointing out that when compensation opportunities are perishable by nature (as they often are) the principal has an unambiguous inclination to front-load compensation.

The second feature of GTM that is subject to debate is the retirement of the agent. The GTM induces the eventual retirement of the agent in line with the recommended policy of previous models (for example, Spear and Wang (2005) and Sannikov (2008)), albeit for other reasons than those hitherto given. In models with transfers the agent retires when it becomes too costly to incentivize him (or he is fired when he becomes too poor to be punished effectively), whereas in our model the agent ceases to carry out investment projects when the principal runs out of incentivization devices.

7 Concluding Remarks

Our model provides insights into patterns of long-run performance in economic environments with extensive nonmonetary compensation. Even in stationary environments, present performance is affected by past decisions, making it impossible to evaluate the efficiency of economic activity from a short-term perspective. Periods of high productivity are inevitably followed by periods when investments are seldom pursued. Periods of low productivity last longer when investment opportunities are more abundant, and efficiency loss is greater under such circumstances when nonmonetary (non-instantaneous) compensation is used.

Time is the unique optimal form of compensation in our environment as it maximizes the utilization rate of the principal's limited resources. In more complicated environments there are other contributory forces including the agent's risk aversion, different discount rates, imperfect observability of the agent's actions, and different beliefs about future events. We think that analysis combining these features with uncertainty over the availability of compensation opportunities can yield productive economic insights. Furthermore, the intriguing connection our model shows between uncertainty over compensation opportunities and information problems merits general analysis.

References

Alonso, Ricardo and Niko Matouschek (2008), "Optimal delegation", *The Review of Economic Studies* 75:1.

Armstrong, Mark and John Vickers (2010), "A model of delegated project choice", *Econometrica* 78:1.

Bird, Daniel and Alexander Frug (2015), *Long Term Authority Relationships*, Mimeo, Northwestern University.

Frankel, Alexander (2014), "Aligned delegation", *The American Economic Review* 104:1.

Guo, Yingni (2014), *Dynamic delegation of experimentation*, Mimeo, Yale University.

Guo, Yingni and Johannes Hörner (2015), *Dynamic Mechanisms Without Money*, Cowles Foundation Discussion Paper No. 1985.

Harris, Milton and Bengt Holmström (1982), "A Theory of Wage Dynamics", *The Review of Economic Studies* 49:3.

Hauser, Christine and Hugo Hopenhayn (2008), "Trading favors: Optimal exchange and forgiveness", *Collegio Carlo Alberto* 88.

Holmström, Bengt (1977), "On incentives and control in organizations", PhD thesis, Stanford University.

Holmström, Bengt and Paul Milgrom (1987), "Aggregation and Linearity in the Provision of Intertemporal Incentives", *Econometrica* 55:2.

Lazear, Edward P. (1981), “Agency, Earnings Profiles, Productivity, and Hours Restrictions”, *The American Economic Review* 71:4.

Li, Jin, Niko Matouschek, and Michael Powell (2015), *Power Dynamics in Organizations*, Mimeo, Northwestern University.

Lipnowski, Elliot and Joao Ramos (2015), *Repeated Delegation*, Mimeo, New York University.

Melumad, Nahum D. and Toshiyuki Shibano (1991), “Communication in settings with no transfers”, *The RAND Journal of Economics*.

Möbius, Markus (2001), *Trading favors*, Unpublished paper, Haravard Uvnisrstiy.

Nocke, Volker and Michael D. Whinston (2010), “Dynamic Merger Review”, *Journal of Political Economy* 118:6.

— (2013), “Merger policy with merger choice”, *The American Economic Review* 103:2.

Ray, Debraj (2002), “The Time Structure of Self-Enforcing Agreements”, *Econometrica* 70:2.

Rogerson, William P. (1985), “Repeated Moral Hazard”, *Econometrica* 53:1.

Sannikov, Yuliy (2008), “A Continuous-Time Version of the Principal-Agent Problem”, *The Review of Economic Studies* 75:3.

Spear, Stephen E. and Sanjay Srivastava (1987), “On Repeated Moral Hazard with Discounting”, *The Review of Economic Studies* 54:4.

Spear, Stephen E. and Cheng Wang (2005), “When to fire a CEO: optimal termination in dynamic contracts”, *Journal of Economic Theory* 120:2.

Appendix A

Proof of Lemma 1

We first show that reward activities are not allowed before an investment project is implemented at a positive intensity. At any point in time the agent can get zero continuation utility by hiding all investments and rewards thus the agent’s continuation utility is always non-negative. Assume that under an IC mechanism a reward is granted before an investment project is executed. Let τ be the first point in time when an investment is carried out. The IC at τ implies that removing rewards from the delegation list before time τ does not affect the agent’s future actions, but reduces the cost of compensation for the principal.

We now show that the IR is binding in an optimal mechanism. From the previous claim, the agent does not get compensated before an investment is carried out. Assume the IR is not binding. Then the IC for the first implemented investment project is not binding for a set of histories with positive measure, because otherwise the agent's utility would be the same as if he never implemented any investments. In which case, his participation utility is zero, and the IR is binding.

Consider histories in this set. There are two options: either the agent is compensated with positive probability before the next investment project arrives or he is not.

In the first case reducing the intensity of reward activities by $\epsilon > 0$ until the next investment is undertaken, while adjusting the delegation function to maintain the same continuation paths, increases the mechanism's value for the principal without affecting any ICs. Namely, there are no decisions before the first investment project is implemented, it does not affect any future ICs along the path, the adjustment of the delegation function maintains the IC of the scaled down rewards, and if ϵ is small enough the first investment project is still carried out.

In the second case, the agent undertakes another investment before compensation begins. This implies that for a set of continuation histories with positive measure, the IC for the next investment project is not binding.

Since there are a finite number of investment projects that can be incentivized before compensation begins, eventually there is an investment project for which the IC is binding and the agent is compensated with positive probability before implementing another investment project. □

Proof of Lemma 2

Assume the best project is not implemented at a promise of $u < \bar{u}$ under the optimal strategy. Consider a situation where the state is u, an investment project of type 1 is available, and construct a deviation of the following form: require the agent to implement the investment opportunity at intensity α and then permit him to enjoy all rewards for ϵ units of time. After ϵ units of time return to the original mechanism. We show that there exist $\alpha, \epsilon > 0$ such that this deviation is IC, satisfies condition (PK), and increases the value of the mechanism for the principal.

From (PK) the intensity of an investment project that can be incentivized for a given ϵ is the $\alpha(\epsilon)$ that solves

$$u = -\alpha(\epsilon)l^1 + \frac{1 - e^{-r\epsilon}}{r} \sum_{j \in J} \lambda^j g^j + e^{-r\epsilon} u$$

$$\alpha(\epsilon) = \frac{1 - e^{-r\epsilon}}{rl^1} \sum_{j \in J} \lambda^j g^j - \frac{(1 - e^{-r\epsilon})u}{l^1}$$

Thus the value from using this alternative strategy is

$$\alpha(\epsilon)B^1 - \frac{1 - e^{-r\epsilon}}{r} \sum_{j \in J} \lambda^j C^j + e^{-r\epsilon} V(u)$$

Thus we need to show that

$$V(u) < \alpha(\epsilon)B^1 - \frac{1 - e^{-r\epsilon}}{r} \sum_{j \in J} \lambda^j C^j + e^{-r\epsilon} V(u)$$

Rearranging terms we get

$$V(u) + \frac{uB^1}{l^1} < \frac{B \sum_{j \in J} \lambda^j g^j}{rl^1} - \frac{\sum_{j \in J} \lambda^j C^j}{r} \tag{2}$$

Denote by $\bar{V}(u)$ the value function in an auxiliary world where the best investment project is always available. It is straightforward to show that

$$\bar{V}(u) = \left(\frac{\sum_{j \in J} \lambda^j g^j}{r} - u\right) \frac{B^1}{l^1} - \frac{\sum_{j \in J} \lambda^j C^j}{r} \tag{3}$$

Clearly, $V(u) \leq \bar{V}(u)$ and, moreover, this inequality is strict for $u < \bar{u}$ since the event that no investment projects arrive for T units of time has a positive probability for any T.

Using the upper bound (3) in equation (2) yields

$$\frac{B^1}{l^1}\frac{\sum_{j\in J}\lambda^j g^j}{r} - \frac{\sum_{j\in J}\lambda^j C^j}{r} - \frac{B^1}{l^1}u + \frac{B^1}{l^1}u \leq \frac{B^1}{l^1}\frac{\sum_{j\in J}\lambda^j g^j}{r} - \frac{\sum_{j\in J}\lambda^j C^j}{r}$$

which is vacuously true. This implies that the deviation is profitable, contradicting the assumption that it is optimal to forgo the best investment project (or any part thereof). □

Proof of Proposition 2

Lemma 3. *It is without loss to restrict attention to mechanisms with the following properties*

1. $\phi_j^{rew}(u) = 0 \quad \forall j \in J, u \in [0, \bar{u}]$.
2. $\phi(u) = 0 \quad \forall u \in [0, \bar{u}]$.
3. $D_j^{rew}(u) > 0 \implies D_k^{rew}(u) = 1 \quad \forall k < j, u \in [0, \bar{u}]$.

This lemma states that the dynamics of u is independent of the realized reward activities, that it does not depend on any lotteries, and that cheaper means of compensation are used first.

Proof. Consider an arbitrary IC mechanism and assume that the current state is $u_0 > 0$. Denote by τ the time until the next investment project arrives. For any pair u_0, τ and $t < \tau$, the above mechanism generates a distribution of the continuation utility at time t in terms of time 0 utils, $e^{-rt}\mathbb{E}[u_t]$, which is independent of τ. The expected continuation utility (measured in terms of time 0 utils) is weakly decreasing (by assumption, no investment projects are carried out in this time) and continuous. Continuity follows from rearranging condition (PK) to get

$$e^{-rt}\mathbb{E}[u_t] = u_0 - \mathbb{E}[\int_0^t e^{-rs}\sum_{j\in J} D_j^{rew}(u_s)g^j dN_{j,s}^{rew} ds] \tag{4}$$

Therefore $e^{-rt}\mathbb{E}[u_t]$ is differentiable almost everywhere; moreover, equation (4) shows that its derivative is an element of $[r\mathbb{E}[u_t] - \sum_{j\in J}\lambda^j g^j, r\mathbb{E}[u_t]]$.

Therefore, we can replicate the dynamics of $\mathbb{E}[u_t|u_0]$ by finding the unique $x \in [0, J]$ for which[22]

$$\frac{\partial \mathbb{E}[u_t|u_0]}{\partial t} = r\mathbb{E}[u_t] - \sum_{j=1}^{\lfloor x \rfloor} \lambda^j g^j - (x - \lfloor x \rfloor)\lambda^{\lfloor x \rfloor + 1} g^{\lfloor x \rfloor + 1} \tag{5}$$

Since we are focusing on a Markovian solution we suppress the time index and denote the solution to equation (5) by $x(u)$.

By construction, the reward component described above induces the minimal expected cost (to the principal) among all compensation schemes inducing the same rate of compensation as that of the chosen mechanism. Since by using this scheme there is no uncertainty about the value of u_t, it is clear that replacing the reward component of the original delegation function with the one constructed above does not harm the principal, who has a weakly concave value function.

Since the value of the mechanism is not affected by the value of D_t for any set of t's with measure zero, we can complete the delegation list arbitrarily for the points where $\mathbb{E}[u_t]$ is non-differentiable.

This implies that it is without loss of generality to assume that the delegation list does not depend on the realization of reward activities and does not use lotteries. □

□

Given the function $x(u)$ we can construct the reward component of the delegation list at u by allowing reward activities $\{1, \ldots, \lfloor x(u) \rfloor\}$ at full intensity and reward $\lfloor x(u) \rfloor + 1$ at an intensity of $x(u) - \lfloor x(u) \rfloor$. Formally, given $x(u)$ the $j-th$ reward project is allowed at the following intensity:

$$D_j^{rew}(x(u)) = \begin{cases} 1 & \text{if } j \leq \lfloor x(u) \rfloor \\ x(u) - \lfloor x(u) \rfloor & \text{if } j = \lfloor x(u) \rfloor + 1 \\ 0 & \text{if } j > \lfloor x(u) \rfloor + 1 \end{cases}$$

Given the result of Lemma 3 and with a slight abuse of notation, we now limit attention to mechanisms wherein the compensation component can be

[22] We define $\lfloor x \rfloor = max\{z \in \mathbb{Z} : z \leq x\}$.

represented by a function $x(u)$. The combination of Lemma 1, Corollary 1, and Lemma 3 enables us to write the HJB equation, corresponding to problem (OBJ), using as the control variables only $x(u)$ and the desired intensity and compensation distribution for investment project implementation, $\{D_i^{inv}(u), \varphi_i^{inv}(u)\}_{i\in I}$. This optimality condition is given by

$$
\begin{aligned}
0=&sup_{x(u),\{\alpha_i(u),\varphi_i^{inv}(u)\}_{i\in I}}\{-rV(u) + V'(u)(ru - W(x(u))) - C(x(u)) \\
&+\sum_{i\in I}\mu^i(D_i^{inv}(u)B_i + \mathbb{E}[V(u+\varphi_i^{inv}(u))] - V(u))\} \\
&s.t \quad x(u)\in[0,J], \quad D_i^{inv}(u)\in[0,\min\{1,\frac{\bar{u}-u}{l^i}\}] \\
&supp(\varphi_i^{inv}(u))\subset[-u,\bar{u}-u], \mathbb{E}[\varphi_i^{inv}(u)] = D_i^{inv}(u)l^i
\end{aligned} \tag{6}
$$

where $W(x)$ is the instantaneous compensation provided to the agent when the control is x:

$$W(x) = \sum_{j=1}^{\lfloor x\rfloor}\lambda^j g^j + (x-\lfloor x\rfloor)\lambda^{\lfloor x\rfloor+1}g^{\lfloor x\rfloor+1}$$

and $C(x)$ is the instantaneous cost of using this control:

$$C(x) = \sum_{j=1}^{\lfloor x\rfloor}\lambda^j C^j + (x-\lfloor x\rfloor)\lambda^{\lfloor x\rfloor+1}C^{\lfloor x\rfloor+1}$$

Given this representation, we can now characterize the properties of the optimal compensation scheme and prove that there is an optimal mechanism that is a multidimensional time mechanism (henceforth MDTM).

Lemma 4. *There exists an optimal $x(u)$, with an image contained in $\{0,\ldots,J\}$, that is weakly increasing.*

Proof. Since the HJB equation is locally linear in x, there exists an optimal solution that does not use partial intensities.

To see there is a non decreasing optimal solution, assume that k_1 is the solution for u. This implies that

$$-V'(u)(W(k_1)) - C(k_1) \geq -V'(u)(W(k_2)) - C(k_2) \quad \forall k_2 < k_1$$

From the weak concavity of $V(u)$, for any $\tilde{u} > u$ we have $V'(\tilde{u}) \leq V'(u)$. Therefore

$$-V'(\tilde{u})(W(k_1)) - C(k_1) \geq -V'(\tilde{u})(W(k_2)) - C(k_2) \quad \forall k_2 < k_1$$

This implies that there is an optimal solution for $x(\tilde{u})$ which is at least k_1. $\square$

$\square$

Lemma 5. *Under an optimal mechanism,* $u > \frac{\sum_{k=1}^{j} \lambda^k g^k}{r}$ *implies that* $x(u) \geq j+1$.

Proof. By the previous lemma there exists an optimal $x(u)$ that is an increasing step function into the set $\{0, 1, \ldots, J\}$.

Assume to the contrary that there exists $\tilde{u}$ such that $W(x(\tilde{u})) < r\tilde{u}$. This implies that there exists an open interval Υ such that for all $u \in \Upsilon$, $\quad x(u) = k < i+1$ and $\sum_{j=1}^{k} \lambda^j g^j < ru$.

Consider $u_0 \in \Upsilon$

- There exists $\epsilon_1 > 0$ such that $u_t \in \Upsilon$ for all $t < \epsilon_1$ if no investment project is implemented. That is,

$$u_0 - \frac{1 - e^{-\epsilon_1 r}}{r} \sum_{j=1}^{k} \lambda^j g^j < sup(\Upsilon)$$

- There exists $\epsilon_2 > 0$ such that the continuation utility after ϵ_2 units of time in which reward activities 1 to $k+1$ are allowed at full intensity is in Υ if no investment project is implemented.

$$inf(\Upsilon) < u_0 - \frac{1 - e^{-\epsilon_2 r}}{r} \sum_{j=1}^{k+1} \lambda^j g^j < sup(\Upsilon)$$

- Since $u_0 - \frac{1}{r} \sum_{j=1}^{k} \lambda^j g^j = \delta > 0$ there exists T such that by time T at least $\frac{\delta}{2}$ discounted utils are provided (in expectation) to the agent via reward activities $\{k+1, \ldots, J\}$. This implies that there exists $\epsilon_3 > 0$ such that reward $k+1$ is available to the agent in expectation for at least $\epsilon_3 > 0$ discounted units of time.

Choose $\epsilon \in (0, \min\{\epsilon_1, \epsilon_2, \epsilon_3\})$ and construct the following process.

$$z_t = \int_{\epsilon}^{t} e^{-rs} \sum_{j=k+1}^{J} \lambda^j g^j D_j^{rew}(u_s) ds$$

That is, construct a process that measures the discounted expected utils the agent receives from rewards $\{k+1,\dots,J\}$ between in the time interval $[\epsilon, t]$.

Construct the following deviation. Implement an investment project if it would have been implemented under $x(u)$ (we make sure that this is possible by constructing a deviation that weakly reduces the agent's continuation utility at any point in time). Add reward activity $k+1$ to the delegation list for ϵ units of time (unless an investment project arrives beforehand, in which case treat this arrival time as ϵ).

This change provides $\xi(\epsilon) = \frac{1-e^{-\epsilon r}}{r}\lambda^{k+1}g^{k+1}$ utils to the agent. Define $\omega(\epsilon)$ as the unique solution to[23]

$$\xi(\epsilon) = \mathbb{E}[\min\{\omega(\epsilon), z_\infty\}]$$

While $z_t < \omega(\epsilon)$, change the the delegation list to $\tilde{x}_t = \min\{k, x_t\}$. Once $z_t \geq \omega(\epsilon)$, return to the original delegation function. This change satisfies condition (PK) since the first change provides $\xi(\epsilon)$ discounted utils to the agent and the second change counterbalances this increase.

This deviation weakly reduces the cost of compensation as $\xi(\epsilon)$ discounted utils are provided by reward activity $k+1$, as opposed to some mixture of (weakly) more expensive rewards without this change. Moreover, with positive probability (for example, if $u_t = \bar{u}$ and $z_t < \xi(\epsilon)$), this reduces the discounted time in which the least efficient reward is used, and thus implies that the deviation is strictly profitable. □

□

Lemmas 3–5 show that there exists an optimal mechanism which is MDTM. We now use the structure of MDTM to derive the uniqueness of this optimal mechanism.

Lemma 6. *For all $j < J$, there exists $\tilde{u}(j) < \sum_{k=1}^{j} \frac{\lambda^k g^k}{r}$ for which under the optimal mechanism $x(\tilde{u}(j)) \geq j+1$.*

Proof. The main part of this proof is to show that $ur < W(x(u))$ for all $u \in (0, \bar{u})$. The result trivially follows from the monotonicity of $W(x(u))$ and

[23] If lotteries are used to incentivize investment projects, it is possible that $\xi(\epsilon) < \omega(\epsilon)$.

the fact that $x(0) = 0$.

From Lemma 5 we know that $ur \leq W(x(u))$; thus it is enough to show a profitable deviation if $ur = W(x(u))$. Furthermore, as we know the value function can be attained by a MDTM, we construct a profitable deviation from this implementation.

Consider such a u, denote $j = x(u) + 1$ and $i^* = argmax_i\{D_i^{inv}(u) : D_i^{inv}(u) > 0\}$ (the worst investment incentivized at u), and construct a deviation as follows. Remove reward activity 1 from the delegation list between periods T and $T + k$, add activity j to the delegation list for the next d units of time (or until a permitted investment project arrives, in which case consider this arrival time to be d). Incentivize the next investment project by first returning reward activity 1 to the delegation list between T and $T + k$ and then reverting to use the original compensation strategy.

For the initial change to satisfy condition (PK) , it must be the case that

$$d = \frac{\log\left(\frac{1}{1-\frac{g^1\lambda^1\left(e^{kr}-1\right)e^{-r(T+k)}}{g^j\lambda^j}}\right)}{r}$$

Moreover, we choose k such that the next investment cannot be incentivized only by reward activity 1:

$$\min_{i:D_i^{inv}(u)>0} l^i D_i^{inv}(u) > \lambda^1 g^1 \frac{1-e^{-rk}}{r}$$

We show that there exists T^* such that for all $T > T^*$ this deviation is profitable. Therefore, it is without loss of generality to assume that no investment project arrives before d.

The deviation is costly if no allowed investment arrives until time T, which happens with probability $e^{-\sum_{i:D_i^{inv}(u)>0}\mu^i T}$. In this case the increase in the cost of compensation is at most[24]

$$\frac{1-e^{-rd}}{r}\lambda^j C^j - e^{-rT}\frac{1-e^{-rk}}{r}\lambda^1 C^1 = e^{-rT}\frac{\lambda^1\left(1-e^{-kr}\right)\left(g^1C^j - g^jC^1\right)}{g^j r}$$

[24]If the first investment project arrives after $T+k$ this bound is exact, but if the first project arrives between T and $T + k$ the actual cost is slightly less.

For this deviation to create a profit, it must increase the discounted amount of time in which reward project j is allowed.[25] The value generated from using reward activity j for the next d units of time is at least

$$\frac{1-e^{-rd}}{r}(\lambda^j g^j \frac{B^{i^*}}{l^{i^*}} - \lambda^j C^j) = e^{-rT}\frac{g^1\lambda^1\left(1-e^{-kr}\right)}{r}(\frac{B^{i^*}}{l^{i^*}} - \frac{C^j}{g^j})$$

Note that the ratio of size of the gain to the size of the loss $\frac{g^1\left(g^j\frac{B^{i^*}}{l^{i^*}}-C^j\right)}{g^1C^j-g^jC^1}$ is bounded away from zero, and thus it is enough to show that the ratio of the probability of loss to the probability of gain converges to zero.

The probability of gain is the probability that reward activity j will always be allowed after the next investment project arrives, as this implies better usage of the most efficient available reward project. Denote by $p_j(u)$ the probability that reward project j is used in full given an initial promise of u.

Denote by $u_t(T)$ the agent's continuation utility at time t conditional on the an initial choice of T and no allowed investment project arriving before time t.

We are interested in $\mathbb{E}^\tau[p_j(u_\tau(T))]$, where τ is the arrival time of the first allowed investment project. Since k was chosen so that the first investment cannot be incentivized solely by reward 1, we have that $p_j(u_\tau(T)) > 0$. Moreover, $p_j(\cdot)$ is an increasing function because the amount of time for which reward activity j is allowed in a MDTM is increasing in u.

The previous expectation is bounded from below by $Pr[(u_\tau(T) > x)]p_j(x)$ for any $x \in (u - \lambda^1 g^1 \frac{1-e^{-rk}}{r}, u)$. As the second term is a strictly positive constant that does not depend on T, it is sufficient to derive $Pr[(u_\tau(T) > x]$. This probability is bounded from below by the portability of this event in histories in which the first investment project to arrive is of type 1. In this case,

$$u_\tau(T) = \begin{cases} u - \lambda^1 g^1 \frac{1-e^{-rk}}{r} e^{-r(T-\tau)} & \text{if } d < \tau \leq T \\ u - \lambda^1 g^1 \frac{1-e^{-r(T+k-\tau)}}{r} & \text{if } T \leq \tau < T+k \\ u & \text{if } \tau > T+k \end{cases}$$

[25]By Lemmas 3 and 4 and the assumption that $ur \leq W(x(u))$, rewards $1, \ldots, j-1$ are permitted until the next implemented investment project arrives.

Therefore, clearly

$$\lim_{T\to\infty} Pr[u_\tau(T) > x] = 1$$

This implies that the probability of gain (loss) converges to 1 (0) with T, and thus the deviation is profitable for a large enough T.

□

Lemma 7. *$V(u)$ is strictly concave.*

Proof. For any $u \in (0, \bar{u})$ take any two values $u^1 < u^2$ in $[0, \bar{u}]$ such that $u = \frac{u^1+u^2}{2}$. We now present an unorthodox way to provide the agent with a continuation utility of u, and use it to derive the strict concavity of $V(u)$.

Consider an auxiliary world in which the maximal intensity of all projects and activities is $\frac{1}{2}$ instead of one. Note that two such (perfectly correlated) auxiliary worlds can be embedded simultaneously in our original environment.

A promise of $\tilde{u}^k = \frac{u^k}{2}$ in auxiliary world $k \in \{1, 2\}$ gives the agent a total utility of u. It is easy to see that the value function in each auxiliary world $\tilde{V}(u)$ is related to the original value function by the following equation:

$$\tilde{V}(\tilde{u}) = \frac{V(u)}{2}$$

since the auxiliary world is a world in which all payoff parameters (for both players) are normalized by $\frac{1}{2}$.

This implies that if the principal were to use an optimal MDTM in each auxiliary world (independently of the other auxiliary world) she would get a payoff of $\frac{V(u_1)+V(u_2)}{2}$. We will now construct an improvement on this strategy and conclude the strict concavity of $V(u)$.

Consider a strategy in which the principal uses an optimal MDTM in each auxiliary world, until the first time the agent's continuation utility in world 1 reaches zero. At this point, and until either the agent's continuation utility in world 2 reaches zero or an investment project arrives, the principal allocates all reward activities of type 1 to world 2. After the first of these two events happens, she reverts to treating each auxiliary world separately. By construction, it is feasible to allow rewards at an intensity of one in auxiliary world 2 in the specified time frame as no reward activities are permitted in auxiliary world 1.

In auxiliary world 2, the principal uses the same delegation function $\tilde{x}(\cdot)$ as she would have done if all projects had been allowed at intensity one-half, with the exception of allowing reward activities of type 1 at full intensity. Note, that this changes increases the downward drift of the agent's continuation utility in world 2 and weakly decreases the cost of compensation. Clearly, this strategy is no worse than the strategy that treats each auxiliary world as separate.

Define $\tilde{u}^k_t$ as the agent's continuation utility in world k if no investment project arrives in t periods of time (when the initial promise is $\tilde{u}^k$ and all reward activities are limited to intensity one-half). Define $T = argmin\{t \in \mathbb{R}^+ : \tilde{u}^1_T = 0\}$. By Lemma 6, $\tilde{u}^1_T = 0$ for all large enough T. Furthermore, as $\tilde{u}^k_t$ has a continuous downward drift, a minimal such T exists and $\tilde{u}^2_T > 0$. Choose $\epsilon, \delta > 0$ such that $\tilde{u}^2_{T+\epsilon+\delta} > 0$. Now consider the histories (with a positive measure) in which no investment project arrives for $T + \epsilon$ units of time, after which $k > \frac{\bar{u}}{l^1} + \frac{1-e^{-r\delta}}{r} W(J)$ investments arrive in the next δ units of time. By Lemma 2 investment projects of type 1 are implemented at the maximal possible intensity. By the choice of k not all investments can be implemented in either auxiliary world. However, as the suggested deviation reduces the agent's continuation utility in auxiliary world 2 at time $T + \epsilon$, this deviation increases the number of implemented investments in world 2. Since this deviation does not increase the cost of compensation, it is a strict improvement over the full separation strategy.

□

Lemma 8. *There is essentially a unique optimal IC mechanism.*

Proof. By the separability of the HJB equation (6) in the different controls, and by the strict concavity of the value function $V(u)$, there is a unique optimal delegation list for all but a measure zero set of $u's$.

The strict concavity of $V(u)$ implies that lotteries are not used to incentivize investment projects, $\varphi^{inv}_i(u) = D^{inv}_i(u) l^i$ for all u, i.

Moreover, when selecting the optimal intensity for investment project i the principal is maximizing

$$\max_{D_i^{inv}(u)\in[0,\min\{1,\frac{\bar{u}-u}{l^i}\}]} \mu^i(D_i^{inv}(u)B^i + V(u + D_i^{inv}(u)l^i) - V(u))$$

which is a strictly concave function in $D_i^{inv}(u)$, and thus has a unique maximizer for any u.

Similarly, when deciding which reward projects to allow, the principal is maximizing

$$\max_{x\in\{0,\ldots,J\}} V'(u)(ru - w(x)) - C(x)$$

which has a unique solution unless, $V'(u) = \frac{g^j}{C^j}$ for some $j \in J$, a condition that can hold for at most a finite set of $u's$ due to strict monotonicity of $V'(u)$. By condition (PK) $x(\bar{u}) = J$ is the unique optimal solution when $u = \bar{u}$. Similarly, by Lemma 1 the unique optimal solution for $u = 0$ is $x(0) = 0$.

By Lemma 6 under the optimal mechanism the measure of time for which $u = \tilde{u}$ for any $\tilde{u} \in (0, \bar{u})$ is zero, and thus there is an essentially unique optimal mechanism.

□

Proof of Proposition 3

Assume to the contrary that two rewards with $\frac{C_k}{g_k} < \frac{C_j}{g_j}$ have the same threshold and consider a promise of $u = \hat{u}_k^{rew} + \delta$. This implies rewards j and k together give the agent $\epsilon \leq \delta$ utils and are used for t_2 units of time.

$$(\lambda^k g^k + \lambda^j g^j)\frac{1 - e^{-rt_2}}{r} = \epsilon$$

If, instead, only reward k is used to provide the ϵ utils that should have been provided by reward activities j and k, it would need to be used for t_1 units of time such that

$$(\lambda^k g^k)\frac{1 - e^{-rt_1}}{r} = \epsilon$$

This gives

$$t_2 = \frac{\log\left(\frac{\lambda^k g^k + \lambda^j g^j}{\lambda^j g^j + \lambda^k g^k e^{-rt_1}}\right)}{r}$$

The expected cost of allowing both rewards for t_2 units of time is

$$(\lambda^k C^k + \lambda^j C^j)\frac{1-e^{-rt_2}}{r}$$

while the expected cost of allowing reward k for t_1 units of time is

$$(\lambda^k C^k)\frac{1-e^{-rt_1}}{r}$$

Thus the direct savings in compensation costs by using reward k for t_1 units of time instead of using k and j for t_2 units of time is

$$\frac{\lambda^k \lambda^j \left(C^j g^k - C^k g^j\right) e^{-rt_1} \left(e^{rt_1} - 1\right)}{r\left(\lambda^k g^k + \lambda^j g^j\right)}$$

Taking a Taylor expansion around $t_1 = 0$, we get

$$\frac{\lambda^k \lambda^j \left(C^j g^k - C^i g^j\right)}{\lambda^k g^k + \lambda^j g^j} t_1 + O(t_1^2)$$

Thus the savings in costs is in the order of t_1.

The probability of a loss of value to the principal due to this change (an investment project arriving before t_1 when both rewards activities are not commited) is

$$(1 - e^{-\sum_{i\in I}\mu^i t_1})$$

Taking a Taylor expansion around $t_1 = 0$, we get

$$\sum_{i\in I}\mu^i t_1 + O(t_1^2)$$

The maximal loss from a misallocation of reward activity k in a period of length t_1 (which occurs if an investment project of type 1 arrives immediately) is

$$g^k\lambda^k\left(\frac{B^1}{l^1} - \frac{C^k}{g^k}\right)\frac{1-e^{-rt_1}}{r}$$

Taking a Taylor expansion around $t_1 = 0$, we get

$$g^k\lambda^k\left(\frac{B^1}{l^1} - \frac{C^k}{g^k}\right)t_1 + O(t_1^2)$$

The maximal expected loss is thus in the order of t_1^2, and when t_1 is small enough it is optimal to use only the reward activity with the better ratio of transfer. Clearly, t_1 converges to zero with δ.

Appendix B

From the discussion in the paper it is clear the principal's value is related to the discounted amount of time the agent is permitted to enjoy reward activities. Therefore, once we know the expected discount factor at the first time $u_t = 0$ for every initial value of u_0, we can calculate the expected discounted measure of time in which rewards are allowed, from which, in turn, we can derive the principal's expected value.

By the definition of TM an agent whose initial continuation utility is u_0 and who carried out no investment projects for t periods has a time t continuation utility of

$$u(t, u_0) = \frac{e^{rt}(ru_0 - \lambda g) + \lambda g}{r}$$

Denote by $\tau(u)$ the first hitting time of $u_t = 0$ given $u_0 = u$.

$$\tau = \min_{t \in R_+ \cup \{\infty\}} \{u_t = 0 : u_0 = u\}$$

Define the expected discount factor at time $\tau(u)$ to be

$$h(u) = \mathbb{E}[e^{-r\tau(u)}]$$

Consider an initial promise of x and a short interval of time ϵ in which two (or more) investment projects are are unlikely to occur. Then $h(x)$ must satisfy the recursion

$$h(x) = e^{-\mu\epsilon}e^{-r\epsilon}h(u(\epsilon, x)) + e^{-r\epsilon}\mu\int_0^{\epsilon} e^{-\mu t}h(u(\epsilon - t, u(t, x) + \min\{l, \frac{\bar{u} - u(t, x)}{l}\}))dt + e^{-r\epsilon}O(\epsilon^2)$$

Rearranging gives

$$e^{r\epsilon}h(x) = e^{-\mu\epsilon}h(u(\epsilon, x)) + \mu\int_0^{\epsilon} e^{-\mu t}h(u(\epsilon - t, u(u(t, x) + \min\{l, \frac{\bar{u} - u(t, x)}{l}\})))dt + O(\epsilon^2)$$

Since $h(x)$ is monotone decreasing and continuous it is differential a.e. and thus we can differentiate the last equality with regard to ϵ:

$$re^{r\epsilon}h(x) = -\mu e^{-\mu\epsilon}h(u(\epsilon, x)) + e^{-\mu\epsilon}h'(u(\epsilon, x))(e^{\epsilon r}(rx - \lambda g)) + \mu e^{-\mu\epsilon}h(u(\epsilon, x) + \min\{l, \frac{\bar{u} - u(t, x)}{l}\})$$
$$+ \int_0^{\epsilon} e^{-\mu t}h'(u(\epsilon - t, u(t, x) + \min\{l, \frac{\bar{u} - u(t, x)}{l}\}))(e^{r(\epsilon - t)}\left(e^{rt}(rx - \lambda g) + r\right))dt + O(\epsilon^2)$$

Taking the limit of $\epsilon \to 0$ and rearranging, we get

$$0 = -(\frac{\mu}{r} + 1)h(x) + h'(x)(x - \frac{\lambda g}{r}) + \frac{\mu}{r}h(x + \min\{l, \frac{\bar{u} - u(t,x)}{l}\}) \tag{7}$$

Clearly, we have the boundary conditions

$$h(0) = 1, \;\; h(\bar{u}) = 0 \tag{8}$$

Therefore, $h(x)$ is the solution to a differential difference equation with suitable boundary conditions.

Consider the range $x \in [\bar{u} - l, \bar{u}]$. For this range we know that $h(x + \min\{l, \frac{\bar{u}-x}{l}\}) = h(\bar{u}) = 0$, and thus for this interval equation (7) becomes

$$0 = -(\frac{\mu}{r} + 1)h_1(x) + h_1'(x)(x - \bar{u})$$

Which is an equation that is a simple ODE whose solution is

$$h_1(x) = (\bar{u} - x)^{\frac{\mu}{r}+1}\alpha$$

for some scalar α.

This, in turn, implies that in the interval $[\bar{u} - 2l, \bar{u} - l]$ equation (7) becomes

$$0 = -(\frac{\mu}{r} + 1)h_2(x) + h_2'(x)(x - \bar{u}) + \frac{\mu}{r}h_1(x + l)$$

with the boundary condition $h_2(\bar{u} - l) = h_1(\bar{u} - l)$, which is again an ODE in $h_2(x)$. This ODE can also be solved as $h_1(x + l)$ is already known (up to α).

We can continue in an iterative fashion until in the interval $[\bar{u}-(k+1)l, \bar{u}-kl]$ the solution to equation (7) is the solution to the ODE with h_{k+1} given by

$$\begin{aligned} &0 = -(\frac{\mu}{r} + 1)h_{k+1}(x) + h_{k+1}'(x)(x - \bar{u}) + \frac{\mu}{r}h_k(x + l) \\ &h_{k+1}(\bar{u} - kl) = h_k(\bar{u} - kl) \end{aligned}$$

Since $\bar{u}$ is finite and $l > 0$ there are a finite number of iterations before we reach $k \geq \frac{\bar{u}}{l}$, at which point we can use the boundary condition $h(0) = 1$ to solve for α in h_1.

When $u = 0$ the expected discount factor at the arrival of the first investment project is $\frac{\mu}{r+\mu}$, after which time there are $\frac{1-h(l)}{r}$ expected discounted units of time in which reward activities are allowed before $u = 0$ is hit again. Thus the discounted amount of time in which the agent is expected to be allowed to benefit from rewards when the current promise is $u = 0$, $W(0)$, solves

$$W(0) = \frac{\mu}{r+\mu}\frac{1-h(l)}{r} + \frac{\mu}{r+\mu}h(l)W(0)$$

or

$$W(0) = \frac{1}{r} - \frac{1}{\mu(1-h(l))+r}$$

Clearly for any $u > 0$ we have

$$W(u) = \frac{1-h(u)}{r} + h(u)W(0) = \frac{1}{r} + h(u)(W(0) - \frac{1}{r})$$

Therefore, the principal's value function is given by

$$V(u) = \frac{B}{l}(W(u)g\lambda - u) - W(u)\lambda C$$

Appendix C

At the beginning of Section 4 we assumed that $\frac{B_i}{l_i}$ is strictly decreasing and $\frac{C^j}{g^j}$ is strictly increasing in order to simplify the exposition of our results. The proof of our main result makes it clear that if there are two reward activities, j_1, j_2, with the same rate of transfer, then the principal treats them identically. We could just as well merge the two rewards and create one reward activity with the same rate of transfer and with an expected benefit (to the agent) per unit of time of $\lambda^{j_1}g^{j_1} + \lambda^{j_2}g^{j_2}$. Since the implementation of rewards has no effect on the continuation path of GTM, the perfect correlation created by merging the two projects does not matter. Conversely, splitting one reward activity into several smaller ones does not affect the optimal mechanism in this environment.

The same holds for investment projects. If the relative benefit of investment to the principal is the same for two projects, then the principal incentivizes both until the agent's continuation utility reaches the appropriate threshold. However, in contrast to the case of reward activities, the implementation of investment opportunities *does* change the continuation path of the mechanism.

Two investment projects, therefore, cannot be merged. To illustrate this, consider the case where $J = 1$. Furthermore, imagine the principal can split the investment project into two independent projects. By doing so she reduces the expected discounted amount of time for which $u = 0$, which suggests that she is using her resources more efficiently and generating a higher value from the mechanism. In the general case, splitting investment opportunities not only increases the efficiency of resource usage but also changes the activation thresholds of projects due to the change in the dynamics of u and the increase in $V(u)$.