Dynamic Nonmonetary Incentives *

Daniel Bird Alexander Frug

January 24, 2017

Abstract

We study a principal-agent interaction where investments and rewards arrive stochastically over time, and are privately observed by the agent. Investments (costly for the agent, beneficial for the principal) can be concealed by the agent. Rewards (beneficial for the agent, costly for the principal) can be forbidden by the principal. We ask how rewards should be used and which investments incentivized. We identify the unique optimal mechanism and analyze the dynamic investment and compensation policies. When all rewards are identical, the unique optimal way to provide incentives is by a "*carte-blanche*" to pursue all rewards arriving in a predetermined timeframe.

Keywords: Dynamic mechanism design, Uncertain action availability.
JEL Classification: D82.

*Bird: Eitan Berglas School of Economics, Tel Aviv University (e-mail: dbird@post.tau.ac.il); Frug: Department of Economics and Business, Universitat Pompeu Fabra and Barcelona GSE (e-mail: alexander.frug@upf.edu). 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 London School of Economics, Northwestern University, Tel Aviv University, The European University Institute, The Hebrew University, and Universitat Pompeu Fabra.

1 Introduction

Optimal incentivization when the interests of the involved parties are not aligned is a long-standing and intensively studied question in economics. Most of the literature focuses on issues of imperfect action observability (e.g., hidden effort) or private information (e.g., cost of effort) and assumes monetary compensation. However, in some applications, instantaneous monetary transfers are infeasible or inferior to other methods of compensation and, more importantly, there is uncertainty regarding the availability of future compensation opportunities. Examples of such situations are abundant: an employer can shorten work-days when business is slow or allow certain perks when they become available (e.g., attend exclusive events held by corporate partners or professional conferences); a start-up firm may be unable to pay its employees until it locates an investor to finance it; an anti-trust authority may incentivize welfare-increasing mergers by supporting profit-increasing mergers in the future; or the management of a large organization may incentivize its divisions by granting preferred access to the organization's capital when the need arises.

For a more concrete example consider an ongoing relationship between the city mayor and a subordinate bureaucrat whose salary is fixed by regulations. Occasionally the bureaucrat's position allows him to recognize opportunities that are either "investments" that benefit the mayor but are costly for himself (e.g., informally facilitate the application process of a desired project), or "rewards" that constitute a personal benefit for the bureaucrat and entail some cost for the mayor (e.g., appoint associates to vacated municipal positions). While investments can easily be concealed by the bureaucrat and rewards banned by the mayor, there are multiple ways in which the randomly arriving rewards can be used to incentivize investments. What is the optimal method of incentivization in this environment? Which investments should be incentivized? How does the optimal policy change over time?

To address these questions we consider a dynamic principal-agent model with a feature that, to the best of our knowledge, is novel in this literature: both investments and rewards (compensation opportunities) arrive stochastically over time, and are either taken immediately or foregone. The implementation of investments is costly for the agent and so the principal uses the randomly arriving rewards for incentivization. While investment and reward availability is privately observed by the agent, his past actions (implemented investments and rewards) are perfectly observed by the principal who can restrict the set of allowed actions based on past implementation. We identify the unique optimal mechanism which is economically appealing due to its simplicity and clear interpretation.

To understand the problem faced by the principal, consider first a situation with one type of reward activity, that is, every implemented reward generates the same (positive) payoff for the agent and entails the same cost (possibly zero) for the principal. Even in this relatively 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), 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 delay it for as long as possible ("back-loading"). We show that the *unique* optimal way to compensate the agent is via a "time-constrained *carte-blanche*," i.e., allowing all rewards that arrive in a fixed time interval that begins at the present moment. In our mayor-bureaucrat example, this implies that after the bureaucrat identifies and implements an investment for the mayor, the most efficient way to provide compensation is by granting him complete freedom to fill vacated positions in his department for a predetermined amount of time. Section 3.1 explains the economic intuition behind this result.

When there are multiple types of investments and rewards the principal has to decide on how jointly to use the different reward activities and to determine which investment opportunities are worth pursuing. Our main result shows that there is a unique optimal mechanism which we call the generalized time mechanism. Under this mechanism the principal compensates the agent by specifying a (separate) time interval for each type of reward activity within which all rewards of that type are allowed. When a desirable investment is implemented the principal increases the length of (some of) these time intervals.

The generalized time mechanism exhibits several notable economic properties that only arise when there are multiple types of investments and rewards. Firstly, the set of allowed rewards at each point in time depends on the amount of compensation owed to the agent and thus changes over time. When owed compensation is high, the principal permits multiple types of reward activities at the same time. Surprisingly, the principal permits reward activities with a high cost of providing a util to the agent even when it is not necessary to do so. In other words, even though the principal can incentivize the agent by only increasing the time interval in which cheap rewards are allowed, she allows more expensive rewards as well. As time goes by, the principal gradually reduces the set of allowed rewards, until a new investment is implemented.

Secondly, investment decisions change non-monotonically over time. Periods with many completed investments are followed by periods in which only high-quality investments are incentivized. If such investment opportunities do not arrive quickly enough, the principal gradually becomes less selective about the investments she is willing to incentivize. Consequently, the principal may incentivize an investment opportunity similar to one she previously chose to forgo. Additionally, the principal eventually allows the agent to enjoy all rewards indefinitely, at which point she will be unable to incentivize further investments. In our mayor-bureaucrat example this implies that ultimately the bureaucrat will have complete freedom to enjoy personal benefits and cease pursuing opportunities that benefit the mayor.

The only information asymmetry assumed in this paper is with regard to the availability of investments and rewards. We abstract away from additional adverse selection and moral hazard problems by assuming that the payoffs from implemented actions are perfectly observable and that investments and rewards arrive according to independent Poisson processes. In particular, this implies that the agent cannot hasten the arrival of rewards and that both players always have the same beliefs about the availability of investments and rewards in the future. We discuss the assumptions of information asymmetry and Poisson arrivals in detail in Section 7.

The paper proceeds as follows. Section 2 introduces the model and presents some preliminary analysis. In Section 3 we highlight the economic intuition and the implications of 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 multiple compensation devices in general and extend our main results to an environment where both monetary and nonmonetary incentivization is available. Section 6 offers a review of the related literature. Section 7 discusses the role of key assumptions. 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 stochastically arriving investments via stochastically arriving rewards. 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 rate of return on investments, $\frac{B^i}{l^i}$, 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. Again, without loss of generality, we order the reward types such that the cost of providing a util, $\frac{C^j}{g^j}$, is weakly increasing in j. All model parameters are non-negative. All investments and rewards can be "scaled down" and implemented at an intensity of[1] $\alpha \in [0,1]$.

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

$$\sum_i \sum_n e^{-rt_n^i} B^i \alpha_n^i - \sum_j \sum_n e^{-rt_n^j} C^j \alpha_n^j,$$

and the agent's utility

$$\sum_j \sum_n e^{-rt_n^j} g^j \alpha_n^j - \sum_i \sum_n e^{-rt_n^i} l^i \alpha_n^i.$$

We assume that both players are expected-utility maximizers, and we refer throughout to the principal's value and 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.[2] At each point in time the agent can take only actions that are both available (by nature) and allowed (by the principal). To specify what actions are allowed at a given point in time we define delegation lists. 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. 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.[3]

[1]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.

[2]The assumption of asymmetric information regarding reward availability is immaterial for our results. See Section 7 for a detailed discussion on publicly observable investments and rewards.

[3]A fully general specification would allow the agent to implement investments and rewards at multiple intensities. However, if doing that were useful, the agent must be indifferent between different intensities. In which case the principal could select the intensity for the agent and permit only that single intensity.

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 public histories into delegation lists. A deterministic delegation mechanism is a (measurable) delegation function. We allow for stochastic delegation mechanisms that are generated by the addition of an appropriately defined randomization device to the public history.

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.[4] Observe that, for the agent, the set of continuation utilities is bounded below by 0, since he can always conceal all future investment opportunities, and bounded from above by his expected utility from enjoying all future rewards without carrying out any investments:

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

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 the requisite measurability restrictions, we do not 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 (possibly stochastic) jumps in u in such cases. In addition, even if no action was taken we allow (mean-preserving) lotteries over the agent's continuation utility. Finally, we allow for a drift in the stochastic process du, which corresponds to the continuous change of u in the absence of jumps.[5]

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 dynamics of u is given

[4] See, for example, Spear and Srivastava (1987).

[5] There is no loss of generality in this formulation due to the Itô-Lévy decomposition.

by

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

where z_t is a standard Brownian motion process. The deterministic drift of the process at u is given by $\eta(u)$, the stochastic drift is given by $\sigma(u)$, and 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 (at any positive intensity), and $\varphi_i^{inv}(u), \varphi_j^{rew}(u)$ 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 aforementioned 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 that stipulates that u is indeed the agent's expected continuation utility:

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

Moreover, the chosen mechanism must incentivize the agent to implement available actions at the allowed intensity:

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

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

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

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

The principal's value function corresponding to the solution of the above problem is denoted by $V(u)$. This value function is weakly concave due to the existence of a public randomization device.[7]

2.3 Preliminary Analysis

To provide sufficient incentives for the agent to implement an investment of type i the principal must increase his continuation utility by at least l^i. The cost of providing one util by using rewards of type j is $\frac{C^j}{g^j}$, and so the expected cost of incentivizing an investment of type i via rewards of type j is $l^i \frac{C^j}{g^j}$. Therefore, incentivizing such an investment in this way can generate a positive profit only if $B^i > l^i \frac{C^j}{g^j}$.

A certain type of investment is said to be redundant if incentivizing it via the most efficient reward activity (type $j = 1$) is not profitable. Similarly, if it is not profitable to exclusively use rewards of a given type in order to incentivize the investment with the highest return (type $i = 1$), we say that this type of reward is redundant. Clearly, redundant rewards will never be used for compensation and the principal will never incentivize redundant investments. To simplify the exposition of our results we assume that all investment opportunities and reward activities are non-redundant. Formally, we make the following parametric assumptions:

$$\frac{B^I}{l^I} > \frac{C^1}{g^1},$$

and

$$\frac{C^J}{g^J} < \frac{B^1}{l^1}.$$

Since the principal observes the exact type of implemented investments at no point does she need to offer expected compensation in excess of the agent's cost.

Lemma 1. *In an optimal mechanism,* $u_0 = 0$.

[7] Were the value function to have a convex region, the principal could increase her continuation payoff by offering the agent a fair lottery over his continuation utility in this region.

An immediate corollary of this lemma is that the expected increase in the agent's continuation utility when he implements an investment equals his cost of implementation.[8]

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 compensation opportunities are both limited and perishable. This implies that the principal has a limited capacity to incentivize investments, and that this capacity is wasted if not used in time. When an investment opportunity is available, the principal can commit some of her capacity in order to incentivize it. If an investment opportunity arrives, the principal forgoes it (and risks wasting some capacity) only if she prefers to save her limited capacity so that she can incentivize better investments in the future. Consequently, an investment project of type 1, i.e., the project with the highest return, is implemented at full intensity (or at 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 A Special Case: One Investment, One Reward

In this section we illustrate and discuss a main property of the optimal mechanism: the *unique* optimal method for providing compensation by means of a certain type of reward activity is via a time allowance. That is, the principal allows the agent to enjoy all rewards that arrive within a given time interval that starts at the present moment. To illustrate this property and its implications in the most transparent way we consider the case with one type of investment project and one type of reward activity. Clearly, this dramatically simplifies the problem and mutes many of its dimensions; the principal need calculate neither the optimal combination of rewards nor which investment projects to incentivize if they become 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 allow him to enjoy many rewards in the distant future.

[8]All subsequent results hold up to changes in the delegation mechanism that effect the implementation of future actions with probability zero.

We begin with an intuitive presentation of the *time mechanism* (henceforth TM). Although this mechanism is a Markovian delegation mechanism, it is instructive to start with an alternative representation that uses the length of time in which the agent is allowed to enjoy rewards as a state variable. To simplify the exposition in this section, we omit investment and reward indexes.

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 all available reward activities. Set $s_0 = 0$; that is, the agent's initial time allowance is zero. When the state is s, and there is $f(s) \in \mathbb{R}_+$ that solves the indifference condition

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

the principal allows the investment project at full intensity, and increases the agent's time allowance to $f(s)$ if an investment is implemented. 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 increases the agent's time allowance to $s = \infty$ if an investment is implemented.[9]

The agent's expected continuation utility when his time allowance is s is

$$u(s) \equiv \int_0^s e^{-rt}\lambda g dt = \frac{1 - e^{-rs}}{r}\lambda g.$$

Therefore, it is possible to transform TM back into the general language of Markovian delegation mechanisms that use the agent's continuation utility as a state variable. A formal definition of TM as a Markovian delegation mechanism is provided in Appendix B.

3.1 Optimality and Uniqueness

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

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

[9]Simple algebra shows that as long as the current time allowance, s, is less than $\frac{\ln(\lambda g) - \ln(lr)}{r}$, investments are implemented at full intensity and the time allowance is increased by $\frac{\ln\left(\frac{g\lambda}{g\lambda - lre^{rs}}\right)}{r}$. If the time allowance is greater than the above threshold, the investment project is executed at the maximal possible intensity, $e^{-rs}\frac{g\lambda}{lr}$.

The simplicity of the $J = I = 1$ case enables us to explain the intuition behind this result and to highlight its main driving forces using basic economic concepts. Firstly, it is strictly desirable for the principal to provide compensation as quickly as possible in order to increase her available capacity to incentivize future investment opportunities (henceforth, we refer to the principal's capacity to incentivize investments as her capacity). It is instructive to begin by considering the extreme situation where the principal has already committed to allowing the agent to enjoy all rewards indefinitely. In this case, the principal cannot incentivize additional investments and her continuation value is the expected cost of allowing all future rewards, $\frac{C}{g}\bar{u}$. If the principal were offered a one-time opportunity to provide the agent with $\bar{u}$ utils immediately at a cost of $\frac{C}{g}\bar{u}$, she would strictly benefit from accepting this offer, because she would incur an identical *direct* cost of compensation, but also be able to incentivize profitable investments in the future. More generally, the principal has a limited capacity and hence she will eventually be forced to forgo profitable investments. Thus, if the principal has an opportunity to increase her capacity (by permitting an available reward activity), she will strictly prefer to do so.

Secondly, conditional on every possible arrival time of the next investment, τ, TM maximizes the principal's expected capacity at τ. To see this, let $u > 0$, and denote by s the corresponding time allowance under TM. Notice that if $\tau \geq s$, the agent's continuation utility at τ equals zero (which is the lower bound for any IC mechanism), and hence the principal has the maximal available capacity, $\bar{u}$. Now consider the case where $\tau < s$. By the definition of TM, regardless of the actual reward arrival process, *all* rewards arriving before time τ are allowed. Therefore, TM attains the lower bound on the agent's expected continuation utility at τ, and hence maximizes the principal's expected capacity at τ.

Thirdly, by the construction of TM, conditional on the value of τ there is no uncertainty regarding the principal's available capacity at τ. Recall, that the principal has a (weakly) concave value function over u, which implies that she also has a weakly concave value function over her capacity, $\bar{u} - u$. Thus, minimizing the risk associated with the availability of future capacity is desirable for the principal.

We have shown that TM provides the best expected outcome up to any time τ with no uncertainty for the (weakly) risk-averse principal, hence, TM must be *an* optimal mechanism. Moreover, for any mechanism that is distinct from TM, there are values of τ for which there exist histories in which some rewards are not allowed at full intensity before time τ. Thus, under such a mechanism, in expectation, the agent receives less compensation than under TM, and hence the principal's expected capacity at τ is strictly lower than under

TM. As the arrival time of the next investment opportunity is a random variable with full support, it follows that TM is the unique optimal mechanism.

3.1.1 Sub-optimality of Allowing a Fixed Number of Rewards

An important implication of Proposition 1 is that the intuitive compensation method of allowing the agent to enjoy a fixed number of rewards is not optimal. In this subsection we illustrate why such incentivization needlessly wastes compensation opportunities and is therefore sub-optimal.

Assume that the agent is currently allowed to enjoy one reward and consider two histories in which no investments arrive until time τ. However, in the first history, $hist_1$, no rewards arrive before τ and in the second history, $hist_2$, two rewards arrive. In $hist_1$, at τ the agent is still allowed to enjoy one reward and so his continuation utility is strictly positive, while in $hist_2$, by time τ the agent has already enjoyed all the rewards he was entitled to and so his continuation utility is zero.

Figure 1: The type of histories used to construct an improvement.

In $hist_2$ the agent was not allowed to enjoy the last reward that arrived. To avoid such a waste of incentives the principal can make an offer of the following gist to the agent at time[10] 0. She can promise to allow the agent to enjoy both rewards that arrive in the second history and in return decrease the agent's continuation utility at time τ if the first history is realized. The decrease is chosen such that the agent's current continuation utility will not be affected by this change. This type of offer increases the principal's capacity to incentivize investments after τ if the first history is realized without affecting the expected cost of compensating the agent or the principal's capacity after τ in the second history, and hence it is profitable.

[10] Without details, we note that the "improvement" we suggest is not well defined. This subsection aims only to provide a general intuition for why allowing a fixed number of rewards is sub-optimal. Formally, this result follows from Proposition 2.

3.2 Value Function

The use of time allowances 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 at 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 principal's value at u is determined by the cost of paying the current debt (the agent's continuation utility) and by the amount of remaining resources that will successfully be utilized in expectation. In the limiting case of always-available investment projects ($\mu = \infty$), the principal will successfully use all her resources to incentivize additional investment projects. Therefore, the value function in this limit case 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 $\mu = 0$, only the cost of compensation remains. All of the uncommitted incentivization resources will be wasted as no further investment opportunities are expected. Thus, the value function is given by

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

For any $\mu \in (0, \infty)$, the value function is strictly decreasing and lies inside the wedge created by the extreme value functions (an example is depicted in Figure 2). In Appendix C we provide an algorithm for analytically deriving the value function.

The value obtained from successfully utilizing the marginal resource is constant; however, for any $0 < \mu < \infty$, the probability of successfully utilizing the marginal resource depends on the state s. 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 opportunity arrives in s units of time. Thus, the probability of successfully utilizing the marginal resource is increasing in s. Therefore, 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. This, in turn, implies that the value function is strictly concave in u despite the risk neutrality of both players and the linearity of the environment.[11]

[11] We prove this for our general model in Lemma 7.

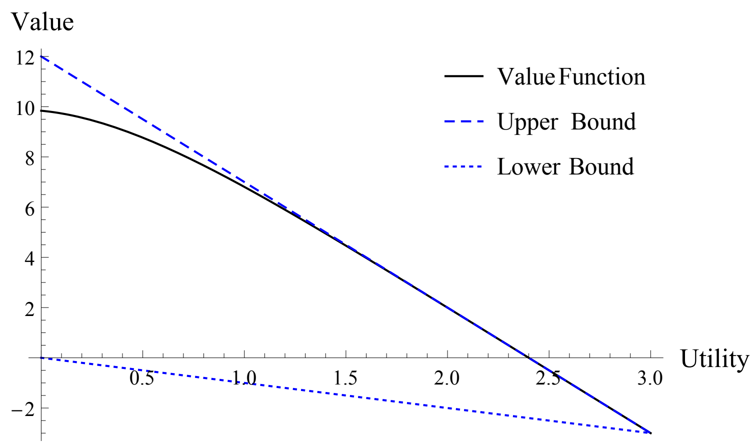

Figure 2: Value function and its bounds for $r = \frac{1}{20}, \mu = \frac{5}{20}, \lambda = \frac{3}{20}, B = 5, l = g = C = 1$.

3.3 Properties of the Time Mechanism

Reward Independence The realized arrival of rewards is not taken into account under TM. Thus, our result does not depend on the assumption that the principal observes implemented rewards. In some cases, it might be natural to assume that the principal only partially observes which rewards were implemented, possibly with some delay. In these environments, TM remains the unique optimal mechanism.

Foregone Investment Opportunities Under TM, after a finite amount of time the mechanism reaches an absorbing state in which no investments are implemented.[12] In contrast, if the principal could provide instantaneous compensation (e.g., via monetary transfers[13]) with an identical direct cost to the one she incurs if reward activities are used for compensation, *all* available investments would be implemented. However, reaching this absorbing state is an ex-ante desirable outcome for the principal since it implies that all of her limited resources will have been efficiently used to incentivize investments. Forgoing some investment opportunities is an unavoidable consequence of the scarcity of compensation opportunities. The expected amount of time it takes to reach this absorbing state determines the (expected) welfare loss resulting from the lack of transfers. The time it takes to reach this state is decreasing in the arrival rate of investment projects and their implementation

[12]By Lemma 2 and 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.

[13]Throughout this paper we refer to transfers as a non-negative payment the principal can deliver to the agent.

cost, and increasing in the principal's capacity to compensate the agent ($\bar{u}$).[14]

Front-loading Under TM the principal provides compensation for the investments the agent has already performed as early as possible ("front-loads compensation") in two distinct ways. Firstly, the principal front-loads compensation within a history by using the earliest available opportunities to compensate the agent. Secondly, the principal front-loads compensation across different histories by providing compensation via a time allowance. Were the time required to provide the promised continuation utility different in two histories, compensation could be front-loaded by shifting a fraction of the compensation that should be paid in the distant future in one realization to an earlier point in time in the other realization. Doing so is desirable as the risk of wasting resources is higher in the realization with the earlier payment. This property stands in contrast to the common practice of back-loading, which is the recommended policy in many settings where transfers are allowed.[15]

Renegotiation Proofness The principal's value function is strictly decreasing in the agent's continuation utility. Therefore, the principal and agent can never find a Pareto improvement by re-negotiating the continuation mechanism at some state of TM.

4 The General Case

Our main result is that in the general model presented in Section 2, there is a unique optimal mechanism under which the principal selectively incentivizes investments and uses reward-specific time allowances to compensate the agent. Formally, we say that a Markovian delegation mechanism is a *multidimensional time mechanism* if for every agent's continuation utility u his promised 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 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 optimal mechanism as the *generalized time mechanism* (henceforth GTM).

[14]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.

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

When $I = J = 1$, TM is the only (unidimensional) time mechanism that satisfies Lemmas 1 and 2, and thus Proposition 2 pins down the unique optimal mechanism. In the general case, there are many multidimensional time mechanisms that satisfy both lemmas. For example, with two types of reward activities there are many (two-dimensional) time allowances that can provide the same level of expected compensation. In addition, when there are two investments, the two-dimensional time mechanisms can differ in regard to when investments of type 2 are incentivized.

The simple case with one type of investment and reward discussed above is illuminating in many aspects; however, it is not rich enough to address two major aspects of the principal's problem: the choice of what types of rewards to permit and what types of investments to incentivize. In the next subsection we address these aspects.

4.1 Optimal Investment and Compensation Policies

For ease of exposition, we assume that the cost of providing a util, $\frac{C^j}{g^j}$, is *strictly* increasing in j and that the rate of return on investments, $\frac{B^i}{l^i}$, is *strictly* decreasing in[16] i. This assumption entails no loss of generality (see Appendix D).

4.1.1 Dynamic Compensation Structure

In Section 3 we argued that front-loading compensation reduces the waste of compensation opportunities and minimizes the *indirect* (opportunity) cost of compensation. The same reasoning suggests that the principal is inclined to permit rewards of different types simultaneously. On the other hand, different types of rewards are associated with different *direct* costs for the principal. Thus, the principal might prefer to avoid permitting expensive rewards. The optimal compensation structure balances the trade-off between minimizing these two distinct costs.

It is easy to see that if a reward activity of type $\tilde{j}$ is allowed at u, rewards with a lower direct cost ($j < \tilde{j}$) should also be allowed. 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 $u \geq \hat{u}_j^{rew}$.

[16] Recall that we have already ordered actions such that both ratios are weakly monotone.

A more interesting feature of compensation under GTM is that the above activation thresholds are 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.[17] In other words, even though the principal can incentivize an investment by increasing the agent's time allowance for cheap reward activities, at some states she will prefer to incentivize it using both cheap and expensive rewards. In particular, 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 $u = \bar{u}$. For the next proposition, recall that the agent's expected continuation utility from enjoying all future rewards of type j is $\int_0^\infty e^{-rt}\lambda^j g^j dt = \frac{\lambda^j g^j}{r}$.

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

Rephrasing this proposition into the language of time-allowances gives the following corollary.

Corollary 2.

1. *If the time allowance of reward j is positive and finite, it is strictly greater than the time allowance of rewards $\tilde{j} > j$.*

2. *The time allowance of reward j is infinite if and only if the time allowance of all rewards is infinite.*

To understand the intuition behind this result it is instructive to consider the case of two types of rewards and a single type of investment. When the agent's continuation utility is very low, the amount of time required to compensate him by using only the cheaper reward activity is very small. In such cases, with very high probability, by the time the next investment opportunity arrives the agent will have already received all his compensation even if only rewards of type 1 are used. Therefore, using the more expensive type of rewards is unlikely to increase the number of investments that will be implemented in the future, but it will definitely increase the cost of compensation. Thus, for sufficiently low levels of u, only the rewards of type 1 are allowed.

The intuition for why the principal uses the more expensive rewards before the capacity of the cheapest reward is exhausted is a bit more complicated. Suppose to the contrary

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

that she does not do so and that at present the agent is allowed to enjoy rewards of type 1 indefinitely and rewards of type 2 are forbidden. We now construct a profitable deviation. The principal reduces the agent's allowance for rewards of type 1 to a large but finite time, and to keep the agent's continuation utility unchanged grants him a small time allowance for rewards of type 2. Furthermore, the principal will incentivize future investments by first returning the time allowance for the cheap reward to infinity and then reverting to increasing the time allowance for the expensive reward.

This deviation clearly increases the cost of paying the existing debt to the agent as some fraction of it is now paid with rewards of type 2. However, it also increases the principal's continuation value in two distinct ways. Firstly, this deviation reduces the cost of incentivizing future investments as, instead of using only rewards of type 2, the principal incentivizes some of these investments with rewards of type 1. Secondly, permitting the agent to enjoy both rewards for a small amount of time increases the principal's capacity to incentivize additional investments since it decreases the agent's continuation utility when the next investment arrives compared to the original mechanism.

For sufficiently small changes, the benefits created by this deviation outweigh the cost. To see why, assume that the next investment arrives sufficiently fast so that increasing the agent's time allowance for the cheap reward to infinity is not sufficient to incentivize the agent to implement it. In this case, the present value of the decrease in the cost of incentivizing the next investment is equal to the increase in the cost of paying the initial debt. However, the principal is left with the benefit of an increased capacity to incentivize additional investments. Clearly, by reducing the agent's time allowance for rewards of type 1 from infinity to a finite but sufficiently large time allowance, the principal can ensure that the first investment arrives "quickly enough" with an arbitrarily high probability.

4.1.2 Dynamic Investment Policy

In GTM there are three separate forces that lead the principal to become more selective about the investments she incentivizes at higher levels of u. Firstly, as the agent's continuation utility increases, the principal's capacity to incentivize additional investments decreases. Secondly, by Proposition 3 the direct cost of incentivizing the agent is increasing in his continuation utility. Thirdly, the use of time allowances for compensation implies that the probability of a better investment opportunity arriving before the marginal resource (of any type of reward) is wasted is increasing in the agent's continuation utility. In the next proposition we show that there exist thresholds $\{\hat{u}_i^{inv}\}_{i=1}^{I}$, 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 the thresholds $\hat{u}_i^{inv}$ are in fact strictly decreasing in i.

Proposition 4. *There exist thresholds* $\bar{u} = \hat{u}_1^{inv} > \hat{u}_2^{inv} > ... > \hat{u}_J^{inv} > 0$ *such that:*

1. *If* $u \leq \hat{u}_i^{inv}$*, then investments of type* i *are incentivized at intensity* $\min\{1, \frac{\hat{u}_i^{inv} - u}{l^i}\}$.
2. *If* $u > \hat{u}_i^{inv}$*, then investments of type* i *are not incentivized.*

This proposition, and the implication of Proposition 3 that the agent's continuation utility drifts down as long as investments are not implemented and $u \in (0, \bar{u})$, provides insights into the nature of dynamic investment decisions. Specifically, it suggests a potential explanation of why an investment opportunity similar to one that was forgone in the past is implemented at present. After a large and profitable investment is carried out the principal requires a high return on her limited resources to justify the incentivization of an investment. Therefore, despite the principal having the resources, some investments opportunities are temporarily forgone, until the agent's continuation utility decreases and the return on such investments is deemed acceptable.

5 Modes of Compensation

Most principal-agent models focus 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 monetary 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.

Our results suggest that when comparing non-instantaneous compensation devices, two (main) features should be considered: (1) The *cost of compensation*, i.e., the direct cost of providing one util to the agent using a given compensation tool. (2) The *speed of compensation*, i.e., the amount of time it takes to provide one util to the agent.

This classification could be valuable in future research on 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?

In the aforementioned problems the principal must select a compensation device (more generally, a subset of devices) from a feasible set. There is a natural partial ranking: a compensation device that is both quicker and cheaper is obviously preferred. The specific characteristics of the feasible investment projects will induce indifference curves over the space of compensation devices. In general, when investment opportunities are rare, the principal is inclined to prefer cheap compensation devices as it is unlikely that the low 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 her capacity to provide incentives even if doing so reduces the net benefit from each investment project. 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 a compensation device are different from their activation order in GTM.

5.1 Adding Monetary Transfers

The main insights of our model remain valid in an environment where monetary transfers are allowed on top of other compensation devices. The optimal mechanism in such environments is similar to GTM, with a few minor modifications.

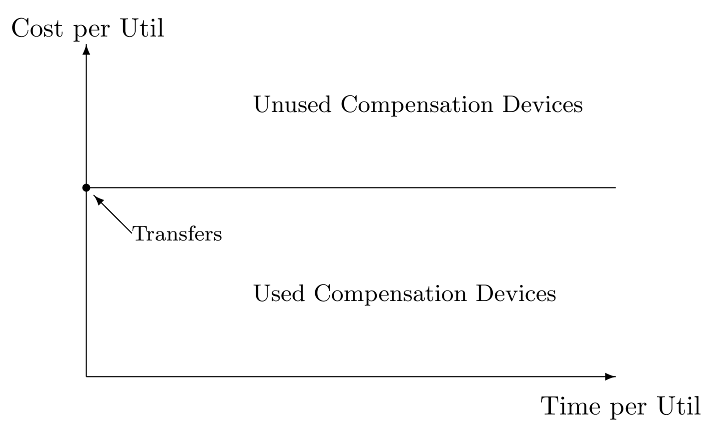

Figure 3: Used and unused compensation devices when transfers are available.

Firstly, in contrast to GTM, the principal does not use reward activities that are more expensive than monetary transfers. 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 that are cheaper than transfers 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.

6 Related Literature

Our work contributes to the literature on dynamic principal-agent interactions by providing the first analysis of a dynamic model with uncertain availability of investments and rewards. The existing literature has addressed a plethora of economic questions in rich environments;[18] however, it does so by assuming compensation (solely) via monetary transfers. We consider a model where the complexity stems from the dynamic stochasticity of action availability. On other dimensions, our environment is simple, indeed, so simple that a direct comparison between our model and previous dynamic principal-agent models is not insightful. However, our model provides new insights into two (related) features of optimal compensation that are frequently debated in the literature: the timing of compensation and the retirement of the agent.

The optimal timing of compensation is a long-standing question in economics. Earlier work such as Lazear (1981) and Harris and Holmström (1982) generally suggests 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, the principal has an unambiguous incentive to front-load compensation.

[18]Seminal papers in this field include the works of Baron and Besanko (1984), Rogerson (1985), Holmström and Milgrom (1987), and Spear and Srivastava (1987).

One interpretation of GTM is that it 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.

The type of uncertainty in our model is similar to the uncertainty in the "trading favors" literature (e.g., Möbius 2001 and Hauser and Hopenhayn 2008), which studies equilibria in games where each player occasionally has an opportunity to grant a favor to his counterpart at a cost to himself. The setting in this literature differs from ours in three important ways: (1) Both players lack commitment power and so, unlike our mechanism design model, the analysis is equilibrium based. (2) Each player privately observes when he can grant a favor to his counterpart. (3) There is a single type of favor each player can grant and receive. Möbius (2001) suggests the use of a "chip mechanism" in which a player grants a favor if the 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. As in GTM the suggested equilibria in those papers rely on a notion of credit between the two players, however this notion is different from the one in GTM because of the mechanism design approach used in this paper. Firstly, in these chip mechanisms either player can be the player who has a positive credit whereas in GTM the principal is never owed anything by the agent. Secondly, although it reverts to the mean over time, the credit in their equilibrium is adjusted *every* time a favor is granted, whereas our credit is dependent only on the timing of the investments implemented by the agent. In a more recent paper, Abdulkadirolu and Bagwell (2013) consider an environment where there is not only asymmetric information about the availability of actions but also asymmetric information with regard to the realized payoff generated by each action. They find that in such a setting a credit based mechanism must sometimes be augmented by infrequent and symmetric punishment to obtain an efficient equilibrium.

Our work also contributes to the literature on delegation. Following the seminal work of Holmström (1977), the literature on delegation has focused on situations in which a principal can delegate the choice of an action to an agent that has private in information about the payoff of the action. Moreover, the literature generally assumes that the interests of the principal and agent are partially aligned (e.g., both players prefer higher actions when the state is high). Initially, the research focused on static settings. Melumad and Shibano (1991) and Alonso and Matouschek (2008) consider the problem of delegating a single choice. They show that the optimal delegation rule is to cap the agents decisions against his bias and show that sufficient alignment of interests is a necessary condition for delegation to be of value. Armstrong and Vickers (2010) consider a single delegation problem where, like in our model, the agent has private information about the available actions rather than information about a payoff-relevant state of nature. They show 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.

Frankel (2014, 2016) studies a more complex environment where the agent simultaneously conducts multiple tasks on behalf of the principal. Frankel (2016) shows that in the multi-dimensional case, capping the weighted average of the agent's actions remains an effective mechanism. In particular, he shows that this mechanism converges to the first-best expected value as the number of decision problems goes to infinity and moreover, that this mechanism is optimal if the agent's biases are constant. Frankel (2014) considers an environment in which the principal is uncertain of the agents preferences. He shows that if there is sufficient uncertainty about the agent's preferences, the max-min optimal mechanism is "moment mechanism" under which all agent type's act as if they has the same preferences as the principal.

In recent years, the literature on delegation has begun to analyze dynamic interactions where, unlike in our model, preferences regarding single actions are partially aligned. Examples of such papers include Guo and Hörner (2015), Lipnowski and Ramos (2015), and Li, Matouschek and Powell (2016). The former two papers assume that the principal never observes the realized payoffs from previous actions, whereas the latter paper (and ours) assumes that all past actions are observable. Moreover, the first paper (and ours) assumes that the principal has full commitment power (and thus utilizes a mechanism design approach), whereas the last two papers assume that her commitment is limited (and thus utilize a repeated game approach). Several works of a more applied nature have also focused on this environment. Guo (2016) studies the delegation of experimentation when no transfers are permitted, and Nocke and Whinston (2010, 2013) analyze the optimal dynamic merger policy for an anti-trust authority.

7 Concluding Remarks

Poisson Arrival of Investments and Rewards

Even though the agent's private information is richer than the public information, the principal and agent always share the same beliefs about the availability of investments and rewards in future. The fact that we assume homogenous Poisson processes simplifies the problem since we can use the current debt as a sufficient representation of a state and specify a Markovian solution. Assuming non-homogeneous Poisson processes would not change the qualitative properties of our main results, however, the set of allowed investments and the structure of compensation would depend on the calendar time.

From the intuition provided for the optimality of our mechanism, it is easy to see that if reward availability is governed by a deterministic process GTM would remain the optimal mechanism. To incentivize the agent to implement an investment the principal must allow him to enjoy a sufficient number of discounted rewards in expectation. When compensation is provided via time allowances, this is done by calculating the expected discounted number of rewards that will be available at each point in time, and setting the time allowance to the required length. Thus, in a sense, compensation via time allowances eliminates the stochastic nature of rewards. Clearly, when rewards arrive deterministically GTM would have another intuitive description - the principal promises a fixed number of reward activities (of each type) and updates this promise (weakly increases the number of allowed rewards) upon the implementation of an investment. The uncertain arrival of investments together with the limitedness and perishability of rewards (and not their stochastic arrival) is what generates the key qualitative features of GTM (namely, the use of expensive rewards before cheap rewards are exhausted and the dynamic investment policy).

If the arrival of investments and rewards is not governed by Poisson processes, the results in our model change for the following reasons. Firstly, the players' beliefs over future action availability may no longer be the same since the agent bases his beliefs on a richer representation of a history. This might create an interesting dynamic adverse selection problem where, for example, the agent may conceal available investments if he is more pessimistic about the arrival of rewards in the near future. Secondly, information about future action availability is on its own valuable. Thus, the principal may be inclined to incentivize investments and allow rewards just for its informational value in the future. Lastly, one can think of an environment where the arrival processes are affected by past implementation. In such cases, when contemplating on the optimal investment and compensation policies the principal would need to take into account the future implications of any implementation.

Symmetric Information: Observability of Investments

In our model it is assumed that only the agent observes whether investments and rewards are available. However, in GTM, implemented rewards do not change the agent's continuation utility, and he reveals to the principal the availability of worthwhile investments at no cost (the agent does not enjoy any information rents). Moreover, past availability of investments and rewards is not informative about the future. This raises a natural question about the effect of asymmetric information in our model. Were we to assume that both players (or only the principal) observe the availability of *rewards*, GTM would remain the unique optimal mechanism. To see this, note that the agent's compensation for an investment must be delivered *after* its implementation, and that regardless of the observability of available rewards, compensation via time-allowances maximizes the utilization rate of the principals limited capacity.

However, the information asymmetry over investment opportunities is detrimental to the principal as it decreases her ability to use her limited capacity effectively. Consider an environment identical to ours in all respects but where the availability of investment opportunities is observed by both players. In this environment the principal must still incentivize the agent to implement investments. However, unlike in the baseline model she can condition the set of permissible actions on both the implementation of previous actions and on the availability of past investments. This setting in analyzed in Bird and Frug (2017), however, the effect of asymmetric information can be illustrated by the following example.

Consider the case where there is a single type of reward and a single type of investment. Moreover, 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 one investment, $\frac{g\lambda - l\mu}{r} = l$. 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. Under the latter (incentive-compatible) mechanism all investments are implemented and the agent is not overcompensated. Since more investment projects are carried out under this mechanism than under TM, it strictly outperforms TM.

The nature of compensation in our environment limits the principal's capacity to compensate the agent. However, the degree to which the principal can utilize her limited capacity depends on the informational assumptions. When the principal observes which investments are available, she can contract over the arrival of future investment opportunities. Such contracts allow the principal to combine realizations in which investment opportunities are rare and the capacity constraint is slack with other realizations in which investments are abundant and the capacity constraint is binding. Under asymmetric information regarding the availability of investment opportunities such contracts cannot be enforced, and thus asymmetric information exacerbates the severity of the limited capacity problem.

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

Appendices

Appendix A: Proofs

Lemma 1

Proof. Recall that the agent's continuation utility is non-negative. Assume to the contrary that under an IC mechanism a reward is allowed before an investment project is executed with strictly positive probability. Let τ be the first point in time when an investment is carried out in such histories. The IC at τ (and onwards) 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 agent's participation utility is zero in an optimal mechanism. The agent does not get compensated before an investment is carried out. Assume to the contrary the agent's participation utility is positive. 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, namely zero.

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, this change does not affect any future IC constraints, 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 not binding and the agent is compensated with positive probability before implementing another investment project. □

Lemma 2

Proof. Assume that there exists a set of agent continuation utilities U in which the mechanism spends a strictly positive expected amount of time and that investment project 1 is not incentivized at the maximal possible intensity. For each $u \in U$ construct a deviation of the following form: require the agent to increase the intensity at which investment 1 is implemented by α_u and upon its implementation permit him to enjoy all rewards for ϵ_u units of time. After ϵ_u units of time return to the original mechanism at state u. We show that there exist $\alpha_u, \epsilon_u > 0$ such that this deviation is IC, satisfies condition (PK), and increases the value of the mechanism for the principal. We now consider a specific value of u and assume that an investment opportunity of type 1 is currently available. Moreover, we drop the subscripts of α and ϵ.

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 following this deviation is

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

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

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

Denote by $\bar{V}(u)$ the value function in an auxiliary world where the investment project 1 is always available. In this world the principal uses all of her capacity to incentivize investment 1 and thus

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

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.

Substituting $V(u)$ with its strict upper bound $\bar{V}(u)$ in the LHS of equation (2) yields the RHS of (2). This implies that the deviation is profitable. □

Proof of Proposition 2

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

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

These conditions state that the dynamics of u is independent of the realized reward activities, that it does not depend on any lotteries, and that the 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, $e^{-rt}u_t|u_0$, and this distribution 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

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

Therefore $e^{-rt}\mathbb{E}[u_t|u_0]$ is differentiable almost everywhere (in t); moreover, equation (4) shows that its derivative is an element of $[r\mathbb{E}[u_t|u_0] - \sum_{j\in J}\lambda^j g^j, r\mathbb{E}[u_t|u_0]]$.

Therefore, we can replicate the dynamics of $\mathbb{E}[u_t|u_0]$ by finding the unique $x \in [0, J]$ for which[19]

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

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 (for the principal) out of all the compensation schemes that induce the same expected rate of compensation as that of the chosen mechanism. Since there is no uncertainty about the value of u_t under this scheme, it is clear that replacing the reward component of the original mechanism 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 $\mathbb{E}[u_t|u_0]$ is non-differentiable. □

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 is allowed at the following

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

intensity:

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

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 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 $x(u)$ and $\{D_i^{inv}(u), \varphi_i^{inv}(u)\}_{i\in I}$ (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 that there is a non-decreasing optimal solution, assume that k_1 is the solution for u. This implies that

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

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 $x(\tilde{u})$ that is at least k_1. □

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

If the assertion in the lemma is false, we can construct a profitable deviation. We do so by providing the agent with a small amount of utils via the cheapest reward that will not be used under the original mechanism (until additional investment are implemented) in the near future, and then offsetting this change by revoking his right to enjoy more expensive rewards afterward. Every investment opportunity that would have been incentivized under the original mechanism is also incentivized under the new mechanism. The new mechanism weakly reduces the cost of providing the promised continuation utility and with positive probability also increases the number of investments that can be incentivized.

Proof. Assume to the contrary that there exists $\tilde{u}$ such that $W(x(\tilde{u})) < r\tilde{u}$. First, define the reward to be added and for how many units of time it is to be added. Given Lemma 4, our assumption implies that there exists an open interval Υ such that for all $u \in \Upsilon, \quad x(u) = k < j+1$ and $\sum_{i=1}^{k} \lambda^i g^i < ru$.

Consider $u_0 \in \Upsilon$.

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

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

- There exists $\epsilon_2 > 0$ such that if no investment is implemented by time ϵ_2, 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.

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

- Since $\delta \equiv u_0 - \frac{1}{r} \sum_{i=1}^{k} \lambda^i g^i > 0$ 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.

Let $\epsilon \in (0, \min\{\epsilon_1, \epsilon_2, \epsilon_3\})$. Add reward activity $k+1$ to the delegation list for ϵ units of time, unless an investment is implemented beforehand, in which case treat its 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.

Now, define the rewards that are to be revoked. Construct a process that measures the discounted expected utils the agent receives from rewards $\{k+1, \ldots, J\}$ after ϵ units of time under the original mechanism:

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

While $y_t < \omega(\epsilon)$, the compensation under the new mechanism is given by $\tilde{x}_t = \min\{k, x_t\}$, where $\omega(\epsilon)$ is the unique solution to

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

Once $y_t \geq \omega(\epsilon)$, return to the original mechanism. The new mechanism provides the same expected continuation utility as the original one since the first change provides $\xi(\epsilon)$ discounted utils to the agent and the second change counterbalances this increase.

The new mechanism weakly reduces the cost of compensation as $\xi(\epsilon)$ discounted utils are provided via reward activity $k+1$, as opposed to some mixture of (weakly) more expensive rewards under the original mechanism. Moreover, with positive probability, $u_t = \bar{u}$ and $y_t < \xi(\epsilon)$, in which case additional investments are implemented under the new mechanism. □

By Lemma 3 there exists an optimal mechanism in which the set of allowed rewards is fully determined by the arrival of past investments. By Lemmas 4 and 5 the principal front-loads compensation via each reward to which she has committed. Together, this shows that there exists an optimal mechanism that is a MDTM.

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, $x(\tilde{u}(j)) \geq j+1$.*

The main part of this proof is to show that $ur < W(x(u))$ for all $u \in (0, \bar{u})$. The result then trivially follows from the monotonicity of $W(x(u))$ and 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. In a MDTM the condition $ur = W(x(u))$ implies that the agent is allowed to enjoy rewards $\{1, \ldots, j-1\}$ indefinitely and is not allowed to enjoy the more expensive rewards. We show that by providing the agent with a small time allowance for reward j the principal increases her capacity to incentivize investments via rewards $\{j,\dots,J\}$ and with an arbitrarily high probability she still manages to use all rewards of types $\{1,\dots,j-1\}$. This implies, that providing the agent with a small time allowance for reward j increases the principal's continuation payoff.

Proof. 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 $T+k$ and 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 using the original compensation strategy.

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

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

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 $e^{-\sum_{i:D_i^{inv}(u)>0}\mu^i T}$. The increase in the cost of compensation is at most[20]

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

For this deviation to create a profit, it must increase the discounted amount of time in which reward project j is allowed.[21] The value generated from using reward activity j for the next d units of time is at least

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

[20]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.

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

Note that the ratio of size of the gain to the size of the loss $\frac{g^1\left(g^j \frac{B^{i^*}}{l^{i^*}} - C^j\right)}{g^1 C^j - g^j C^1}$ 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 the 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 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 could not 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 probability of this event in histories in which the first investment project to arrive is of type 1. In this case,

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

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. Assume that the agent's continuation utility equals u and let $u_1 < u_2$ such that $u_1, u_2 \in [0, \bar{u}]$ and $u = \frac{u_1+u_2}{2}$. One (non-natural) way the principal can deliver a promise of u is by fictitiously splitting all investments and rewards into two halves and creating two (perfectly correlated) fictitious worlds, each of which contains half of each reward and half of each investment opportunity. She can then provide $\frac{u_1}{2}$ utils using GTM in fictitious world 1, and $\frac{u_2}{2}$ utils using GTM in fictitious world 2.

Notice that our environment is insensitive to scaling in the following sense. Refer to the original environment as the unscaled world; however, if all payoff parameters are multiplied by a constant, refer to it as a scaled world. For any realization of actions, an action is implemented in the unscaled world if and only if it is implemented in the scaled world. Therefore, in the scaled world at every stage of the interaction, the agent's continuation utility and the principal's expected value are simply multiplied by the chosen scale factor. By the above compensation method, all payoff parameters are scaled down by $(\frac{1}{2})$; accordingly, the dynamics of GTM in fictitious world i, in terms of implemented actions, is identical to the dynamics of GTM in the unscaled world with an initial state u_i. Therefore, in fictitious world i the principal generates a value of $\frac{V(u_i)}{2}$ and, in total, such a compensation method generates a total value of $\frac{V(u_1)+V(u_2)}{2}$, which shows directly that the value function is weakly concave.

We now show that $V(u) > \frac{V(u_1)+V(u_2)}{2}$ by offering an improvement on this method of compensation. Notice that with positive probability, at some point in time the agent's continuation utility in world 1 reaches zero. Moreover, when it happens for the first time, the agent's continuation utility in world 2 is strictly positive. Now the principal can benefit from temporarily merging the two fictitious worlds. Specifically, instead of wasting her capacity in world 1 (where all rewards are forbidden before the next investment opportunity arrives), 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. □

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

Proof. The strict concavity of $V(u)$ implies that lotteries are not used to incentivize investment projects. I.e, in any optimal mechanism

$$\varphi_i^{inv}(u) = D_i^{inv}(u)l^i \quad \forall u \in [0, \bar{u}], i \in I$$

Moreover, given the strict concavity of $V(u)$, any optimal mechanism must satisfy the properties described in Lemmas 3 and 4.

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.

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 it thus has a unique maximizer for any u.

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

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

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, and, by Lemma 1, $x(0) = 0$ is the unique optimal solution.

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. □

Proposition 3

By Lemma 6 we know that $\hat{u}_j^{rew} < \sum_{k=1}^{j-1} \frac{\lambda^k g^k}{r}$. Since the optimal mechanism satisfies the conditions specified in Lemma 3 we know that $\hat{u}_j^{rew}$ is weakly increasing. Thus, all we need to show is the strict monotonicity. We do so by formalizing the intuition provided in the text.

Proof. 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 that rewards j and k together give the agent $\epsilon \leq \delta$ discounted 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 committed) 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 while the gain is of order t_1. Clearly, t_1 converges to zero with δ, which concludes the proof. $\square$

Appendix B: TM as a Markovian Delegation Mechanism

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

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

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

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

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

$$\text{(6)} \quad u_0 = 0$$

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 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 enjoyed rewards but drifts down continuously and deterministically (as long as no investment is implemented). The initial condition (6) trivially corresponds to $s_0 = 0$.

Appendix C: Derivation of the Value Function

From the discussion in the paper it is clear that 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 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:

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

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

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

This 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

$$0 = -(\frac{\mu}{r}+1)h_{k+1}(x) + h'_{k+1}(x)(x-\bar{u}) + \frac{\mu}{r}h_k(x+l)$$
$$h_{k+1}(\bar{u}-kl) = h_k(\bar{u}-kl)$$

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 D: Weak Order on Actions

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 that 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 $V(u)$.

References

Abdulkadirolu, Atila, and Kyle Bagwell. 2013. "Trust, reciprocity, and favors in cooperative Relationships." *American economic journal: Microeconomics*, 5(2): 213–259.

Alonso, Ricardo, and Niko Matouschek. 2008. "Optimal delegation." *Review of Economic Studies*, 75(1): 259–293.

Armstrong, Mark, and John Vickers. 2010. "A model of delegated project choice." *Econometrica*, 78(1): 213–244.

Baron, David P., and David Besanko. 1984. "Regulation and information in a continuing relationship." *Information Economics and Policy*, 1(3): 267–302.

Bird, Daniel, and Alexander Frug. 2017. "Long-term authority relationships." Tel-Aviv University Mimeo.

Frankel, Alexander. 2014. "Aligned delegation." *American Economic Review*, 104(1): 66–83.

Frankel, Alexander. 2016. "Delegating multiple decisions." *American economic journal: Microeconomics*, 8(4): 16–53.

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

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.” *Review of Economic Studies*, 49(3): 315–333.

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 diss. Stanford University.

Holmström, Bengt, and Paul Milgrom. 1987. “Aggregation and Linearity in the Provision of Intertemporal Incentives.” *Econometrica*, 55(2): 303–328.

Lazear, Edward P. 1981. “Agency, earnings profiles, productivity, and hours restrictions.” *The American Economic Review*, 71(4): 606–620.

Li, Jin, Niko Matouschek, and Michael Powell. 2016. “Power dynamics in organizations.” *American economic journal: Microeconomics*, Forthcoming.

Lipnowski, Elliot, and Joao Ramos. 2015. “Repeated delegation.” New York University Mimeo.

Melumad, Nahum D., and Toshiyuki Shibano. 1991. “Communication in settings with no transfers.” *RAND Journal of Economics*, 173–198.

Möbius, Markus. 2001. “Trading Favors.” Harvard Unifversity Mimeo.

Nocke, Volker, and Michael D. Whinston. 2010. “Dynamic merger review.” *Journal of Political Economy*, 118(6): 1201–1251.

Nocke, Volker, and Michael D. Whinston. 2013. “Merger policy with merger choice.” *American Economic Review*, 103(2): 1006–1033.

Ray, Debraj. 2002. “The time structure of self-enforcing agreements.” *Econometrica*, 70(2): 547–582.

Rogerson, William P. 1985. “Repeated moral hazard.” *Econometrica*, 53(1): 69–76.

Sannikov, Yuliy. 2008. “A continuous-time version of the principal-agent problem.” *The Review of Economic Studies*, 75(3): 957–984.

Spear, Stephen E., and Cheng Wang. 2005. “When to fire a CEO: Optimal termination in dynamic contracts.” *Journal of Economic Theory*, 120(2): 239–256.

Spear, Stephen E., and Sanjay Srivastava. 1987. “On repeated moral hazard with discounting.” *Review of Economic Studies*, 54(4): 599–617.