

Deferred compensation and risk-taking incentives*

Florian Hoffmann[†] Roman Inderst[‡] Marcus Opp[§]

February 2016

Abstract

Our paper develops a simple principal-agent framework to analyze the equilibrium relationship between risk-taking and the timing of pay. In our setup, the agent's one-time action has persistent effects through affecting the arrival time distribution of a disaster event. While the principal receives informative signals about the agent's action over time, it is costly to rely on this information for incentive pay since the agent is relatively impatient. Optimal compensation contracts resolve the tension between impatience and information with at most two payout dates. Our framework lends itself to analyze recent regulatory interventions mandating minimum deferral periods and clawback provisions in the financial sector. It shows how such regulatory interference in the timing dimension causes the principal to adjust other dimensions of the compensation contract, which may then lead to higher risk-taking. Mandatory deferral requirements are more likely to be effective in reducing risk-taking when competition for agents is high.

Keywords: Compensation design, principal-agent models, financial regulation.

JEL Classification: D86 (Economics of Contract: Theory), G28 (Government Policy and Regulation), G21 (Banks, Depository Institutions, Mortgages).

PRELIMINARY AND INCOMPLETE

*We thank Nicolae Gârleanu, Denis Gromb, Ben Hermalin, Gustavo Manso, Robert Marquez, John Morgan, Andy Schwartz, Philipp Strack, Johan Walden, and Jeffrey Zwiebel for valuable insights. Hoffmann and Inderst gratefully acknowledge financial support from the ERC (Advanced Grant “Regulating Retail Finance”).

[†]University of Bonn. E-mail: fhoffmann@uni-bonn.de.

[‡]Johann Wolfgang Goethe University Frankfurt. E-mail: inderst@finance.uni-frankfurt.de.

[§]Haas School of Business, University of California, Berkeley. E-mail: mopp@haas.berkeley.edu.

1 Introduction

The timing of executive compensation has received widespread attention in the public debate in the aftermath of the financial crisis: short-term oriented compensation packages are blamed for having contributed to excessive risk-taking in the financial sector, in particular, since large bonuses to financial executives have been paid out before the realization of disastrous outcomes. In response, numerous regulatory initiatives around the world have introduced or proposed mandatory minimum deferral periods and clawbacks.¹

“Compensation schemes overvalued the present and heavily discounted the future, encouraging imprudent risk-taking and short-termism. (...) In the UK, we have introduced a remuneration code prescribing that payment of bonuses must be deferred for a minimum of three years and, after payment, be exposed to clawback for up to seven years.”

Mark Carney, Governor of the Bank of England, November 17, 2014.

Our paper argues that the regulator’s implicit diagnosis “Short-term pay equals high risk-taking incentives” is misguided, as the timing of pay is just one of many dimensions of the compensation contract relevant for risk-taking incentives. Indeed, what is needed is an understanding of the *equilibrium* relationship between the timing of pay and risk-taking. Such an understanding requires a theoretical framework, which we develop in a setting where the agent’s action has persistent effects. This feature allows us to capture the notion that (risk-taking) outcomes are oftentimes not immediately observable and may only be exposed in downturns. While our framework applies to many real-life settings outside the financial sector, it allows us, in particular, to meaningfully evaluate the effectiveness of regulatory intervention in the timing of pay. Such analysis has been called for by the Financial Stability Board, a quasi-regulatory body for financial institutions.

“The effectiveness of these mechanisms remains largely untested and more analysis is needed to assess whether tools such as malus and clawbacks are sufficiently developed and effectively used to deter conduct risks.”

Financial Stability Board, 2015.

Consistent with these concerns, we show that deferral regulation is indeed not a panacea but we also derive circumstances for when this policy robustly leads to less risk-taking.

¹In the EU, a new directive adopted in 2010 includes strict rules for bank executives’ bonuses. Directive 2010/76/EU, amending the Capital Requirements Directives, which took effect in January 2011. It has already been fully implemented in a number of countries, including France, Germany, and the UK and has led to mandatory deferral of bonuses for several years.

On an abstract level, regulation is (only) effective as long as the bank’s shareholders, the principal, cannot adjust the compensation contract along any margin other than the required timing dimension. Concretely, we relate the possibilities of such adjustments to the size of the agent’s outside option, which fixes the size of pay (if sufficiently high), and mandatory clawback provisions, which prescribe the contingency of pay.

Our setup aims to capture a basic economic trade-off that gives rise to endogenous payout times. Think of canonical principal-agent environments, such as Holmstrom (1979) or Harris and Raviv (1979), but suppose that the principal receives signals about the agent’s initial action over time.² Due to liquidity needs the agent is effectively impatient relative to the principal (see e.g., DeMarzo and Duffie (1999)), which makes it costly to defer compensation and rely on these more informative signals. For the case of a risk-neutral agent and principal and limited liability of the agent (such as e.g., Innes (1990)), we provide a complete characterization of the optimal timing of compensation for general information processes. To make predictions above and beyond optimal compensation design, we suppose that the agent’s action affects the arrival time distribution of a disaster event, such as the failure of a bank. Higher effort by the agent reduces the likelihood of disaster events, but comes at a personal cost to the agent. This captures the idea that generating “alpha” is costly to the agent, while taking on (down-side) risk is not. By imposing this additional structure, we can derive intuitive conditions for the validity of the first-order approach and we can relate our analysis to recent regulatory initiatives.

Optimal contracts are surprisingly simple in that they have at most two payout dates which we can explicitly characterize. When the agent’s participation constraint is slack, the optimal payout time is simply the unique payout date that minimizes the costs of delay per unit of information (where information is appropriately measured in “likelihood ratio units”). If information arrives smoothly over time, the associated first-order condition intuitively states that at the optimum payout time, the *growth rate of information* is equal to the *growth rate of discounting costs* (given the agent’s greater impatience). Interestingly, we show for various commonly used information processes that the optimal payout time may be both increasing as well as decreasing in the risk-taking action, so that longer deferral cannot generally be associated with less risk-taking. If the agent’s participation constraint binds, an additional, earlier payout may be optimal. We provide a proof with straightforward graphical intuition relying on a simple convexification argument to illustrate when this is the case. This analysis forms the backbone for our

²For related information environments, see Hopenhayn and Jarque (2010), as well as Hartman-Glaser, Piskorski, and Tchisty (2012) and Malamud, Rui, and Whinston (2013) in the context of securitization.

consideration of mandatory deferred compensation.

While the just described compensation contracts are Pareto-optimal for any given action, the induced equilibrium action may involve excessive risk-taking from society's perspective, for instance, as a bank's board or shareholders (the principal) may not fully internalize the social cost of bank failure. As the motivating quote by Mark Carney reveals, regulators seem to have implicitly asserted that restrictions on compensation design in the form of deferral and clawback clauses would effectively lower risk-taking. We show that this is indeed the case when the level of the agent's compensation package is fixed by her outside option and regulation both requires deferral and clawbacks (contingent on the absence of a disaster event). Otherwise, the principal is able to adjust other relevant margins of the compensation contract in such a way that it countervails the intent of such compensation regulation and, typically, results in higher rather than lower risk-taking. A key determinant for whether regulation "works" is thus whether the agent's outside option is sufficiently high, i.e., banks compete fiercely for talented CEOs and other risk takers.

While it is intuitive that regulation has the intended effect as long as the principal only adjusts the timing of pay (and thus uses more precise information), it is useful to provide intuition for why it may backfire when he can readjust the entire compensation package. In this case, the principal optimally induces the level of effort that equates his marginal benefit to the associated marginal compensation costs. Since regulation only affects compensation design, but not the benefit, regulation is effective at lowering risk-taking if and only if it reduces *marginal* compensation cost. However, there is a robust effect that causes mandatory deferral to pull towards an increase in marginal compensation cost. By optimality of the unconstrained payout time, any required deferral above and beyond this payout time causes discounting costs to grow faster than the associated gain in information. (Otherwise, the principal would have chosen to wait for more precise information himself). Importantly, this effect applies for every dollar that the agent receives in present value terms, robustly pulling towards an increase in marginal (and, trivially, the level of) compensation cost. In addition to this unambiguous "timing-inefficiency" effect, the total effect of mandatory deferral on marginal costs also depends on how (the growth rate of) information process changes in the action itself. For small deviations from the principal's unconstrained payout time, the former "timing-inefficiency" effect is second-order, and we demonstrate that regulation may either increase or decrease risk-taking depending on the information process. However, for sufficiently stringent regulation the induced timing inefficiency always dominates leading robustly to higher risk-taking.

Literature TBD.

Organization of paper. Our paper is organized as follows. Section 2 describes our economic and regulatory environment and sets up the full formal problem. In our main Section 3, we analyze the Principal’s Problem consisting of optimal compensation design (Section 3.1) and the equilibrium action choice (Section 3.2). Section 4 discusses the comparative statics of risk-taking with respect to the stringency of regulation and gives guidance when regulation leads to lower risk-taking and improves welfare. In Section 5, we outline that the solution to our compensation design problem applies to a larger class of principal-agent settings. Section 6 concludes.

2 Model setup

Economic environment. The modeling framework is a static moral hazard problem with a risk-neutral agent and principal (similar to Innes (1990)) in which the principal observes signals about the agent’s action over time. In our main application, the agent should be interpreted as a CEO, or key risk-taker of a financial institution and the principal as (the board of) a bank.

At time 0, an agent A with outside option v takes an unobservable action $a \in \mathcal{A} = [0, a^{\max}]$ which controls the arrival time distribution function $F(t|a)$ of a verifiable bad event affecting the principal P and society. The bad event should be interpreted as a disaster, such as the failure of a bank or a nuclear meltdown. Action a comes at a personal cost $c(a)$ to the agent, generates a discounted expected payoff $\pi(a)$ to the principal, and imposes a discounted expected negative externality $x(a)$ on society. The wedge between the principal’s payoff and society’s payoff, $x(a)$, captures the part of losses that the principal does not internalize, e.g., due to limited liability or a shorter time horizon than society.³ The rate of time preference for the agent and principal is denoted by r_A and r_P , respectively.

Assumptions. We suppose that the survival function, $S(t|a) = 1 - F(t|a)$ is twice continuously differentiable in both arguments. Throughout the text, we use subscripts as a short-hand notation for partial derivatives, so that $h(t|a) = F_t(t|a)/S(t|a)$ denotes the associated hazard rate. To capture the notion that a higher action reduces the instance of a disaster, we posit

³Alternatively, the wedge could result from a corporate governance problem, in which the board’s preferences are not perfectly aligned with long-run shareholder value.

Assumption 1 *The hazard rate $h(t|a)$ is strictly decreasing in a for all $t \in (0, \infty)$.*

$$h_a(t|a) < 0 \quad \forall t \in (0, \infty), a. \quad (1)$$

Assumption 1 implies a first-order stochastic dominance order on the (family of) survival distributions indexed by a . It is satisfied for our three leading example distributions.

Example 1 *Mixed distribution: $S(t|a) = aS_L(t) + (1-a)S_H(t)$ with $a \in [0, 1]$ and where $S_L(t)$ dominates $S_H(t)$ in the hazard rate order, i.e., $h_L(t) < h_H(t)$.*

Example 2 *Weibull distribution: $S(t|a) = e^{-\left(\frac{t}{a}\right)^\kappa}$, with $a > 0$, $\kappa > 0$.⁴*

Example 3 *Log-normal distribution: $S(t|a) = \frac{1}{2} - \frac{1}{2} \operatorname{erf} \left[\frac{\log t - a}{\sqrt{2}\sigma} \right]$, with $a \geq 0$, $\sigma > 0$.*

Assumption 1 plays a similar role as the monotone likelihood ratio property (MLRP) in static principal agent models with immediately observable outcomes, see e.g., Rogerson (1985). Section 5.1 discusses how most of our results can be extended to more general information processes, such as when information arrives at discrete points in time.

To ensure the validity of the first-order approach, we assume that the agent cost function is a strictly convex transformation of the survival probability S , i.e.,

Assumption 2

$$\frac{c''(a)}{c'(a)} > \frac{S_{aa}(t|a)}{S_a(t|a)} \quad \forall t, a. \quad (2)$$

We impose the following standard restrictions on payoffs:

Assumption 3 *$\pi(a)$, and $-x(a)$ are strictly increasing and concave. $c(a)$ is strictly increasing, strictly convex, and satisfies $c'(0) = c''(0) = 0$ as well as $c'(a^{\max}) = \infty$.*

In addition to ensuring tractability, this assumption formalizes the idea that the principal does not bear the full social cost of low effort, leading to excessive disaster risk, as suggested by the motivating quotes.

Finally, the agent is assumed to be impatient relative to the principal:

Assumption 4 *The discount rates satisfy $r_A - r_P = \Delta_r > 0$.*

Relative impatience of the agent is a standard assumption in dynamic principal-agent models (see e.g., DeMarzo and Duffie (1999), DeMarzo and Sannikov (2006), Opp (2012),

⁴The case $\kappa = 1$ yields the exponential distribution.

or Opp and Zhu (2015)). It is typically motivated with liquidity needs on the side of the agent. In our model, relative impatience introduces a cost of obtaining more precise information.

In this environment, the principal designs a cost-minimizing compensation contract subject to the following restrictions. First, due to limited liability of the agent, a feasible cumulative compensation payout process $b(t)$ must feature positive increments, $db(t) \geq 0$. Second, it must be adapted to the filtration $\mathcal{F}_t$ generated by $Y(t)$ where $Y(t) = 1$ indicates that the disaster has occurred before time t , and $Y(t) = 0$ otherwise.

Regulatory environment. Regulation may impose additional restrictions on compensation contracts. A minimum deferral period τ mandates that the unrestricted payout to the agent (available for immediate consumption) satisfies $db(t) = 0 \quad \forall t < \tau$. A clawback clause requires the absence of a disaster at the stipulated payout time t , i.e., $db(t) = 0 \quad \forall t \geq \tau$ if $Y(t) = 1$. The latter constraint is referred to as a clawback clause, as one may implement the contractual restrictions via an escrow account subject to clawback provisions.⁵ We provide further discussion on the separate role of the two clauses in Lemma 6. Finally, we note that, since the principal would never pay the agent upon failure in an unregulated setting (see Lemma 1), our setup nests the “laissez-faire” outcome for the special case $\tau = 0$.

Formal problem. The full formal problem consists of the *Principal’s Problem* and the *Regulator’s Problem* of choosing the optimal deferral period τ . Following the approach of Grossman and Hart (1983), we decompose the analysis of the *Principal’s Problem* into the compensation design problem (Problem 1) and the optimal action choice (Problem 2).

First, consider the design of the compensation process $b(t)$ that implements a fixed action a at the lowest (present value of) wage cost $W(a|\tau)$ to the principal given the minimum deferral period τ . For any action a , the entire set of Pareto-optimal compensation contracts can be obtained by varying the agent’s outside option v .

Problem 1 (Compensation design)

$$W(a|\tau) = \min_{b(t)} \mathbb{E} \left[\int_0^\infty e^{-r_P t} db(t) \middle| a \right] \quad s.t.$$

⁵That is, the principal makes pay contributions into an escrow account even before date τ (yielding interest at rate r_P). However, the agent only has unrestricted access to the account after date τ and if the institution has not failed at the time of the stipulated payout.

$$V_A := \mathbb{E} \left[\int_0^\infty e^{-r_A t} db(t) \middle| a \right] - c(a) \geq v \quad (\text{PC})$$

$$a = \arg \max_{\tilde{a}} \mathbb{E} \left[\int_0^\infty e^{-r_A t} db(t) \middle| \tilde{a} \right] - c(\tilde{a}) \quad (\text{IC})$$

$$db(t) \geq 0 \quad \forall t \quad (\text{LL})$$

$$db(t) = 0 \quad \forall t < \tau \quad (\text{DEF})$$

$$db(t) = 0 \quad \text{if } Y(t) = 1 \quad (\text{CLAW})$$

The first constraint refers to the participation constraint (PC) of the agent. The agent value V_A , i.e., the present value of compensation (discounted at the agent's rate) net of the effort cost must exceed his outside option v . Second, incentive compatibility (IC) requires that it is optimal for the agent to choose action a given $b(t)$. Limited liability of the agent (LL) implies that $db(t) \geq 0$. Finally, (DEF) and (CLAW) refer to the additional constraints imposed by regulation. Since regulation constrains the contracting space, the principal's wage cost $W(a|\tau)$ to implement a given a must weakly increase in τ . Note further, regulation may be so restrictive, that no solution to Problem 1 exists (see Lemma 6). In this case, we define $W(a|\tau) = \infty$.

Second, given minimal wage cost $W(a|\tau)$ the induced equilibrium action, $a^*(\tau)$, maximizes the principal's present value of profits.

Problem 2 (Equilibrium action)

$$a^*(\tau) = \arg \max_a \pi(a) - W(a|\tau). \quad (3)$$

As in most of the literature on principal-agent models (see e.g., Holmstrom (1979) and Rogerson (1985)) we do not assume a participation constraint on the side of the principal.

Given the resulting equilibrium action $a^*(\tau)$, we define welfare as:

$$\omega(\tau) = \pi(a^*(\tau)) - c(a^*(\tau)) - \lambda_x x(a^*(\tau)) - \mathbb{E} \left[\int_\tau^\infty (e^{\Delta_r t} - 1) e^{-r_A t} db(t) \middle| a^*(\tau) \right], \quad (4)$$

where λ_x refers to the welfare weight that the regulator puts on the externalities and the final term measures the deadweight cost of deferred compensation resulting from relative impatience of the agent.⁶ While we are mainly interested in the *positive* implications of regulatory interference in the timing of pay, we also analyze the *Regulator's Problem* yielding the *normative*, welfare-maximizing deferral period τ^* .

⁶ We thus make the simplifying assumption that the principal and agent have the same welfare weight.

Problem 3 (Welfare)

$$\tau^* = \arg \max_{\tau} \omega(\tau). \quad (5)$$

We note that first-best welfare (absent an agency problem) does not involve any inefficient delay and the associated first-best action a^{FB} solves the first-order condition $\pi'(a^{FB}) - \lambda_x x'(a^{FB}) = c'(a^{FB})$.

3 Principal's Problem

3.1 Compensation design

This section studies Problem 1, i.e., the principal's optimal compensation design to implement a given action; both in the absence and presence of regulation. The preliminary subsection 3.1.1 derives general properties of optimal contract design which are used to simplify the subsequent analysis. For expositional reasons, the actual analysis initially studies a relaxed problem without a participation constraint of the agent (Section 3.1.2) and then continues with the full problem in Section 3.1.3.

3.1.1 Preliminaries

The goal of this section is to transform Problem 1 into a more tractable problem that clearly exposes precise measures for the costs and benefits of deferral within our setup. First, we show that (CLAW) is automatically satisfied in the solution to the unregulated program ignoring (DEF) and (CLAW). Hence, setting the deferral requirement to $\tau = 0$ in Problem 1 produces the unregulated wage cost $W(a|0)$.

Lemma 1 *Without regulation, an optimal contract never rewards the agent upon failure, i.e., $db(t) = 0$ if $Y(t) = 1$.*

Intuitively, the hazard rate condition (Assumption 1) implies that the absence of a disaster is the strongest signal of agent effort at any time t . It is now convenient to rewrite Problem 1 in terms of the process for the *expected discounted pay* of the agent, denoted as $B(t)$. Since (CLAW) applies with and without regulation, $dB(t)$ satisfies

$$dB(t) = e^{-r_A t} S(t|a) db_S(t), \quad (6)$$

where $db_S(t)$ denotes the instantaneous reward conditional on survival by time t . Moreover, since (DEF) implies that $dB(t) = 0 \forall t < \tau$, we may write (PC) concisely as:

$$V_A = \int_{\tau}^{\infty} dB(t) - c(a) \geq v. \quad (7)$$

Similarly, we can simplify (IC) using the following Lemma.

Lemma 2 *Assumption 2 implies validity of the first-order approach. (IC) becomes*

$$\int_{\tau}^{\infty} \frac{S_a(t|a)}{S(t|a)} dB(t) = c'(a). \quad (8)$$

For a fixed t , the term $\frac{S_a(t|a)}{S(t|a)}$ can be interpreted as the standard continuous-action likelihood ratio of static moral hazard environments (see e.g., Hart and Holmstrom (1987)). We define

$$L(t|a) \equiv \frac{S_a(t|a)}{S(t|a)}. \quad (9)$$

Different than in the static environment, the likelihood ratio is a function of time.

Lemma 3 *The likelihood ratio L satisfies $L(0|a) = 0$ and is strictly increasing in t with*

$$L_t(t|a) = -h_a(t|a) > 0 \quad \forall t \in (0, \infty), a. \quad (10)$$

Within our setup, $L(t|a)$ is the appropriate measure of how much “information” has been revealed by date t . Intuitively, no information is revealed by time 0, i.e., $L(0|a) = 0$. Moreover, information grows faster at a particular point in time, as measured by $L_t(t|a)$, if the hazard rate is more sensitive to effort. To see this, think of approximating the local incentive constraint with binary effort. If the hazard rate under high and low effort is identical at time t , the principal learns nothing from the absence (or occurrence) of a disaster. If instead, the hazard rate under low effort is much higher than under high effort, the principal learns a lot. Formally, this is captured by the term $-h_a(t|a)$, which is strictly positive by Assumption 1. Figure 1 plots the hazard rate and function L for specifications of Example distributions 1 to 3. We plot the functions for two effort levels. It can be inferred from the graph that the slope of L is greater the greater the difference in hazard rates.

Using the discounted expected payment process $B(t)$, the first-order approach and the definition of the likelihood ratio $L(t|a)$, we can now write Problem 1 concisely as

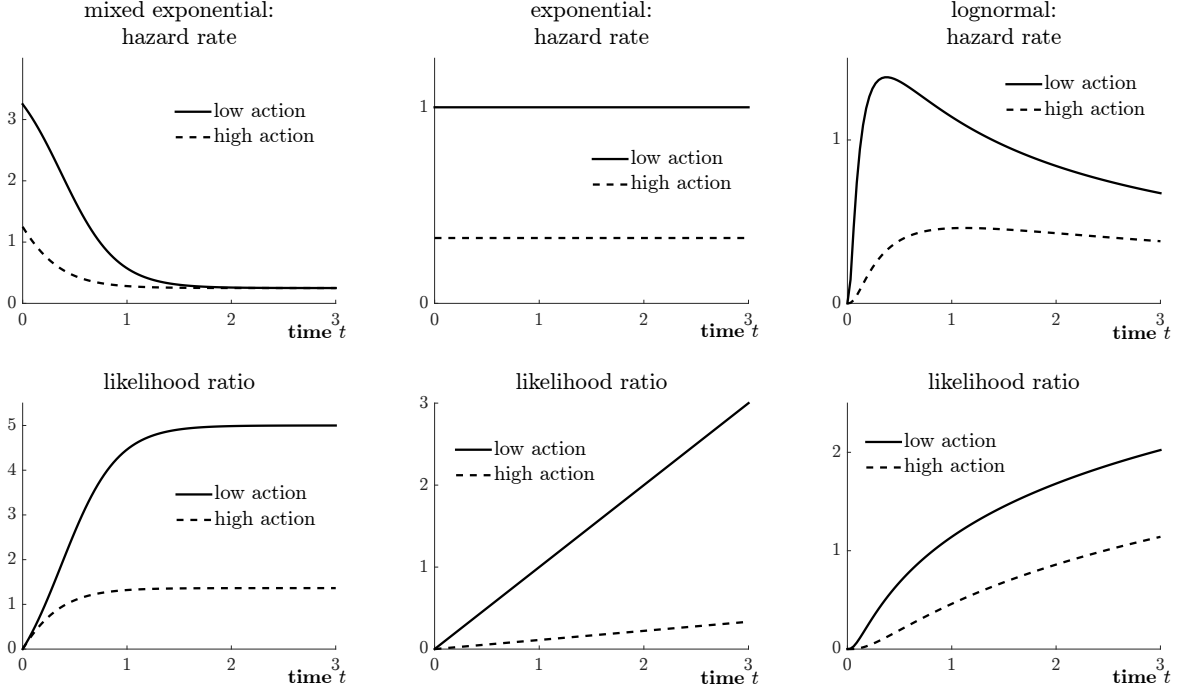


Figure 1. Example information processes. This graph plots the hazard rate (top panels) and likelihood ratio (lower panels) for different arrival time distributions falling in the class of Examples 1 to 3 (from left to right). The left panels show a mixed exponential distribution with respective parameters $\lambda_L = 1/4$ and $\lambda_H = 4$, respectively. The middle panels show the case of an exponential distribution (Weibull with $\kappa = 1$) with parameter $\lambda = 1/a$ and the right panels a lognormal distribution with parameters $\mu = a$ and $\sigma = 1$. For each distribution the plots show the respective functions for two actions corresponding to a mean arrival time of 1 (low action) and 3 (high action), respectively.

Problem 1*

$$W(a|\tau) = \min_{B(t)} \int_{\tau}^{\infty} e^{\Delta_r t} dB(t) \quad s.t. \quad (11)$$

$$\int_{\tau}^{\infty} dB(t) - c(a) \geq v \quad (PC)$$

$$\int_{\tau}^{\infty} L(t|a) dB(t) = c'(a) \quad (IC)$$

$$dB(t) \geq 0 \quad \forall t \quad (LL)$$

The transformed Problem 1* clearly reveals the costs and benefits of delay. The cost of delay is measured by $e^{\Delta_r t}$, since for every unit of discounted expected pay to the agent,

the principal incurs a cost of $e^{\Delta_r t}$. The benefit of delay is that every unit of discounted expected pay boosts incentives, i.e., relaxes IC, by $L(t|a)$. To ensure that the costs of delay are sufficiently important as to affect contract design, i.e., we impose

Assumption 5 *For any action a , the growth rate of L_t satisfies $\lim_{t \rightarrow \infty} \frac{d \log L_t}{dt} \leq 0$.*

Then, for any finite τ , all payments to the agent occur in finite time.

Before solving Problem 1*, it is useful to point out that the survival distribution $S(t|a)$ only enters the transformed problem via $L(t|a)$. Hence, it is possible to interpret $L(t|a)$ more generally as the primitive information process of the economic environment rather than an endogenous object derived from a family of survival functions. Moreover, since time maps one-to-one into information $L(t|a)$, we may then define the *cost of information* as the costs of delay viewed as a function of L rather than t :

Definition 1 *For any action a , the cost of information is given by the function $e^{\Delta_r L^{-1}(z|a)}$, mapping $z \in [0, \lim_{t \rightarrow \infty} L(t|a))$ into $\mathbb{R}^+$.*

This function will play an important role in our analysis (see, in particular, Lemma 5 and Proposition 2), and has the following properties:

Lemma 4 *For any action a , the cost of information, $e^{\Delta_r L^{-1}(z|a)}$, is strictly increasing in z for all $z \in [0, \lim_{t \rightarrow \infty} L(t|a))$. It is strictly convex if and only if*

$$\frac{d^2 e^{\Delta_r t}}{dt^2} / \frac{de^{\Delta_r t}}{dt} = \Delta_r > \frac{L_{tt}}{L_t} \quad \forall t. \quad (12)$$

In general, our assumptions on the information process do not impose any restrictions on the cost of information except for that more information is also more costly (which follows from Lemma 3). It may be strictly convex over the entire domain, such as for the exponential distribution, but this is not generically true for all information processes.⁷ Non-convexities naturally arise when information growth varies sufficiently over time, such as in Examples 1 or 3 (see Figure 1).

3.1.2 Relaxed problem

To clearly highlight the intuition for our results, we will first analyze contract design in the relaxed problem which ignores the (PC) constraint.

⁷ Assumption 5 only ensures convexity in the limit, i.e., for $L = \lim_{t \rightarrow \infty} L(t|a)$.

Proposition 1 (Relaxed problem) *For any action a , the cost-minimizing Pareto-optimal compensation contract is unique. The single payment date $\hat{T}_1$ satisfies*

$$\hat{T}_1(a|\tau) = \arg \min_{t \geq \tau} \frac{e^{\Delta_r t}}{L(t|a)}. \quad (13)$$

The agent values the associated compensation package at

$$dB(\hat{T}_1) = \frac{c'(a)}{L(\hat{T}_1|a)}. \quad (14)$$

Intuitively, due to risk-neutrality of both parties, the optimal contract shifts all payments to the date that offers the best cost-benefit trade-off. This payout date, $\hat{T}_1$, minimizes the cost of delay, $e^{\Delta_r t}$, per unit of information, $L(t|a)$. Of course, if there was no cost of delay, $\Delta_r = 0$, the optimal payout time would simply maximize information, so that $\hat{T}_1 = \infty$. Except for knife-edge cases, the global minimum of $\frac{e^{\Delta_r t}}{L(t|a)}$ is unique. If there are multiple solutions, any solution produces the same wage cost to the principal

$$W(a|\tau) = e^{\Delta_r \hat{T}_1} \frac{c'(a)}{L(\hat{T}_1|a)}.^8 \quad (15)$$

While the principal is indifferent among these solutions, the agent strictly prefers the one with the shortest payout time since $V_A = \frac{c'(a)}{L(\hat{T}_1|a)} - c(a)$. Thus, using Pareto-optimality as the selection criterion, the optimal payout date $\hat{T}_1$ is unique. Given $\hat{T}_1$, the expected discounted payout (14) is obtained from (IC).

Payout time in the absence of regulation. The optimal unregulated payout time, $\hat{T}_1(a|0)$, solves the (necessary) first-order condition:⁹

$$\left. \frac{d \log L}{dt} \right|_{t=\hat{T}_1(a|0)} = \Delta_r. \quad (16)$$

At the optimal unregulated payout time, $\hat{T}_1(a|0)$, the growth rate of information, $\frac{d \log L}{dt}$, must equal the growth rate of the costs of delay, Δ_r . For the Weibull family considered in Example 2 (and the special case of the exponential distribution, i.e., $\kappa = 1$), this

⁸ The value of $\frac{e^{\Delta_r \hat{T}_1}}{L(\hat{T}_1|a)}$ can be interpreted as the shadow value on the incentive constraint.

⁹ Existence of a solution to (16) follows from continuity of $L(t|\cdot)$ together with $\lim_{t \rightarrow 0} \frac{d \log L}{dt} = \infty$ and $\lim_{t \rightarrow \infty} \frac{d \log L}{dt} = 0$, where we have used Lemma 3 together with Assumptions 1 and 5. We note, without further assumptions, the first-order condition may not be sufficient (see Figure 2).

trade-off yields a particularly simple closed-form solution $\hat{T}_1(a|0) = \frac{\kappa}{\Delta_r}$.

The Weibull distribution example also reveals that early payout times are not necessarily reflective of poor incentives, i.e., low actions a . In general, the comparative statics of $\hat{T}_1(a|0)$ in a depend on whether the rate of information growth, $\frac{d \log L}{dt}$, increases or decreases in a . More formally, by the implicit function theorem we obtain that

$$\text{sgn} \left(\frac{d\hat{T}_1(a|0)}{da} \right) = \text{sgn} \left(\frac{\partial}{\partial a} \frac{d \log L}{dt} \Big|_{t=\hat{T}_1(a|0)} \right). \quad (17)$$

In particular, the payout time is decreasing in the action in Example 1, independent of the action in Example 2, and increasing in the action in Example 3. These findings already suggest that deferral regulation may have different effects on the equilibrium action depending on the information structure; we will come back to this in Section 4.1 (see, in particular, Proposition 3).

Contract design under deferral regulation. Regulation constrains the principal's optimal choice of payout times as soon as $\tau > \hat{T}_1(a|0)$.

Lemma 5 *In general, $\hat{T}_1(a|\tau) = \tau$ for τ sufficiently high. If the cost of information, $e^{\Delta_r L^{-1}(z|a)}$, is strictly convex for $z > L(\hat{T}_1(a|0)|a)$, then*

$$\hat{T}_1(a|\tau) = \max \left\{ \hat{T}_1(a|0), \tau \right\}. \quad (18)$$

Convexity of the cost of information is key for the effect of regulation on payout times. The first statement follows from Assumption 5 which ensures convexity in the limit. Moreover, if convexity is satisfied as soon as $\tau > \hat{T}_1(a|0)$, the constrained optimal payout time $\hat{T}_1(a|\tau)$ is simply given by the minimum deferral period τ . For Example 2, the Weibull distribution, we thus obtain that $\hat{T}_1(a|\tau) = \max \left\{ \frac{\kappa}{\Delta_r}, \tau \right\}$.

Economically, non-convexities naturally arise if information growth rates fluctuate sufficiently over time (see Condition 12). The left panel of Figure 2 plots an example information process with two phases of high growth. Hence, the cost of information plotted in the right panel exhibits a non-convex region. In this example, the optimal unconstrained payout time (pinned down by the tangent line through the origin with the minimum slope, $\frac{e^{\Delta_r t}}{L(t|a)}$) violates the regulatory constraint, i.e., $\hat{T}_1(a|0) < \tau$. However, the endogenously chosen payout time given the constraint exceeds the minimum deferral period, $\hat{T}_1(a|\tau) > \tau$, so that one may (wrongly) infer that the constraint is irrelevant.¹⁰

¹⁰ This result is reminiscent of Jewitt, Kadan, and Swinkels (2008) who show that minimum wages may harm the principal even when the agent receives more than the minimum wage in the (constrained)

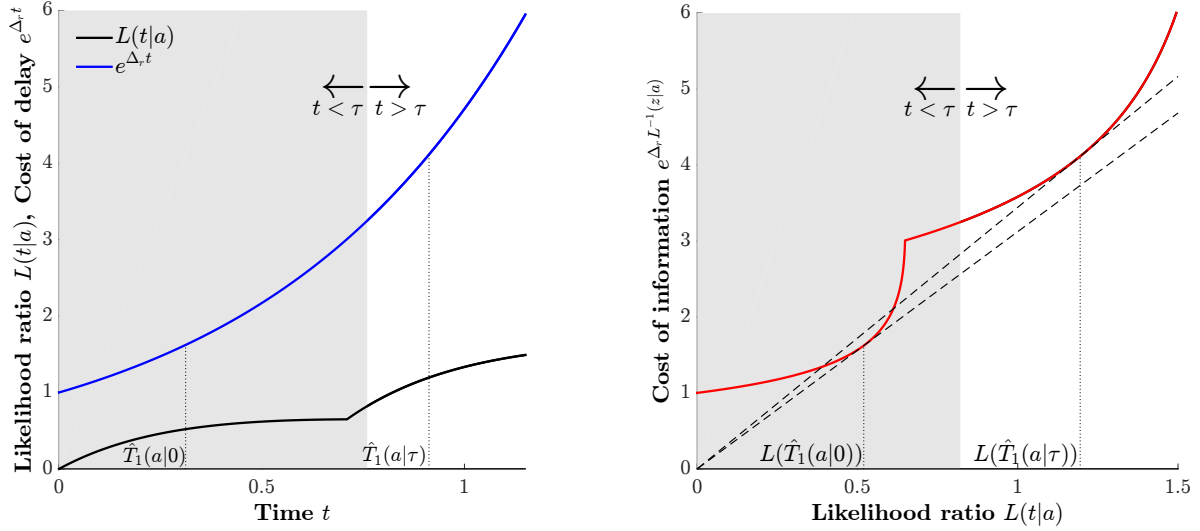


Figure 2. Payout times under deferral regulation. This graph plots the likelihood ratio and the cost of delay as a function of time (left panel) as well the cost of information (right panel) for a piecewise mixed exponential arrival time distribution: $F(t|a) = wG(t|a)$ for $t \leq s$ and $F(t|a) = F(s|a) + wG(t-s|a)$ for $t > s$, where $G(\cdot)$ is mixed exponential, i.e., $G(t|a) = aG_L(t) + (1-a)G_H(t)$ with G_L and G_H the cdfs of two exponential distributions with parameters $\lambda_L = 1/4$ and $\lambda_H = 4$ respectively, and the weight $w := 1/(1 + G(s|a))$ scaling total probability to 1. The graph shows the respective functions for $s = 0.71$, $a = 0.95$ and $\Delta_r = 1.55$.

It is important to highlight that when regulation constrains the timing dimension of pay (and thus increases the precision of the signal that the principal uses to incentivize the agent), it correspondingly reduces the required level of pay to incentivize a given action a , see (14). Clearly, this effect is not present when the agent's participation constraint binds. We will now explore the implications of a participation constraint in the full analysis of the compensation design problem.

3.1.3 Full problem

To highlight the additional features of optimal compensation design induced by (PC), we initially focus on optimal contracts with binding (PC). Intuitively, these contracts arise in equilibrium if the agent's outside option is sufficiently high, otherwise the contracts derived in the previous section apply (see Theorem 1).

The main difference in the case of binding (PC) is that the principal's choice of payout times no longer aims to influence the agency rent (versus (14)). Instead, for a given action a , the agent's total compensation package is fixed at $\int_{t=\tau}^{\infty} dB(t) = v + c(a)$, so that optimal compensation contracts minimize weighted average deadweight costs of delay

optimal contract.

subject to ensuring incentive compatibility. More formally, let $dw(t) = \frac{dB(t)}{v+c(a)} \geq 0$ denote the fraction of pay that the agent derives from date- t compensation, then Problem 1* can be written as

$$W(a|\tau) = (v + c(a)) \min_{w(t)} \int_{\tau}^{\infty} e^{\Delta_r t} dw(t) \quad \text{s.t.} \quad (19)$$

$$\int_{\tau}^{\infty} L(t|a) dw(t) = \frac{c'(a)}{v + c(a)} \quad (20)$$

It is immediate that if the deferral requirement τ is sufficiently stringent, so that $L(\tau|a) > \frac{c'(a)}{v+c(a)}$, no feasible contract can implement action a . Intuitively, as the entire pay, $v + c(a)$, has to be contingent on survival up to at least date τ , and, thus, on information of at least $L(\tau|a)$, there is a lower bound on incentives of $(v + c(a)) L(\tau|a)$ for any given a . Since this section focuses on compensation design, we, for now, restrict attention to the set of implementable actions $a \geq \underline{a}(\tau)$ where $\underline{a}(\tau)$ satisfies

$$L(\tau|\underline{a}) = \frac{c'(\underline{a})}{v + c(\underline{a})}. \quad (21)$$

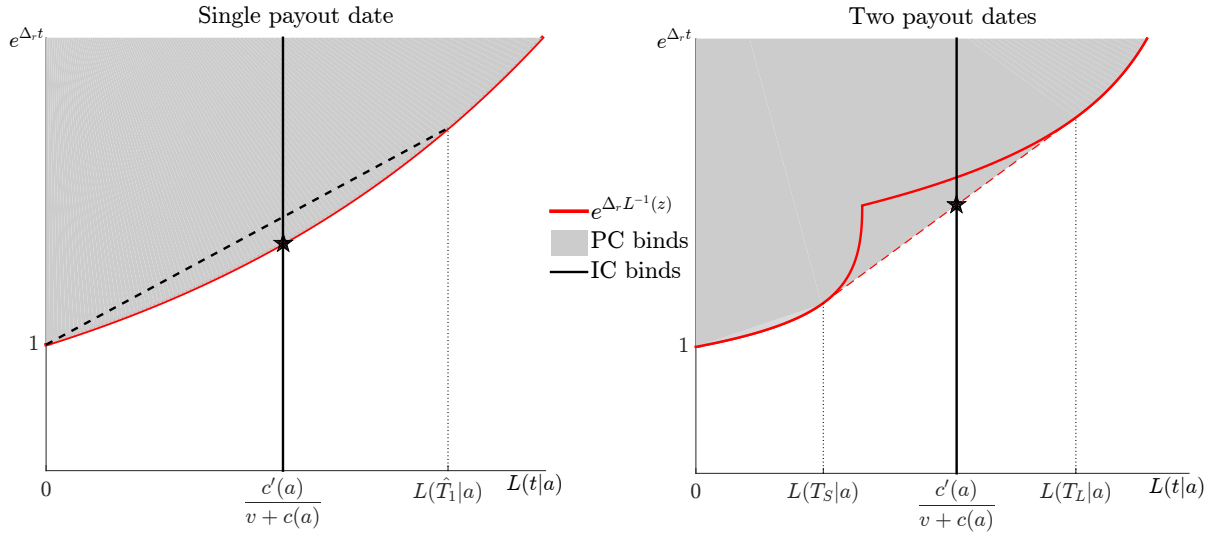


Figure 3. One vs. two payment dates. This graph plots the cost of information $e^{\Delta_r L^{-1}(z|a)}$ for the exponential distribution from Example 2 (left panel) as well as for the piecewise mixed exponential arrival time distribution as in Figure 2 (right panel). For both panels the cost function is given by $c(a) = 0.5\alpha a^2$ and parameters values are chosen as follows: $a = 0.8$, $\Delta_r = 1$, $\alpha = 1.15$ and $v = 0.632$.

Proposition 2 *In an optimal compensation contract with binding (PC), the wage costs satisfy*

$$W(a|\tau) = (v + c(a)) C \left(\frac{c'(a)}{v + c(a)} \middle| a, \tau \right),$$

where $C(z|a, \tau)$ denotes the convexification of $e^{\Delta_r L^{-1}(z|a)}$ for $z \geq L(\tau|a)$. The contract features a single payout date T_1 solving $L(T_1|a) = \frac{c'(a)}{v+c(a)}$ if and only if

$$e^{\Delta_r T_1} = C \left(\frac{c'(a)}{v + c(a)} \middle| a, \tau \right), \quad (22)$$

Otherwise, the contract features a short-term payout date $T_S(a|\tau) \in [\tau, T_1)$ and a long-term payout date $T_L(a|\tau) > T_1$ such that for $\lambda = \frac{L(T_L) - \frac{c'(a)}{v+c(a)}}{L(T_L) - L(T_S)}$

$$\lambda e^{\Delta_r T_S} + (1 - \lambda) e^{\Delta_r T_L} = C \left(\frac{c'(a)}{v + c(a)} \middle| a, \tau \right). \quad (23)$$

Intuitively, if the cost of information is strictly convex, so that $C(z|a, \tau) = e^{\Delta_r L^{-1}(z|a)}$, a single (and unique) payout date is optimal.¹¹ Any incentive-compatible combination of payments at two payout dates, such as for example at date $\hat{T}_1$ and date 0 (see left panel of Figure 3), would yield strictly higher weighted average deadweight costs of delay than paying exclusively at date T_1 .

Payments at two dates are optimal when there are non-convexities in the cost of information.¹² As an illustrative example consider the information process of Figure 2 and suppose that T_1 falls into the region where information grows very little immediately before date T_1 and grows strongly at some point after date T_1 . Put differently, the cost of information plotted in the right panel of Figure 3 grows steeply for $L < L(T_1|a) = \frac{c'(a)}{v+c(a)}$ and grows relatively little for $L > \frac{c'(a)}{v+c(a)}$. Then, it is optimal to make a payment at a long-term date $T_L > T_1$ to tap the additional information that the principal learns after date T_1 to provide incentives, and make an additional early payment at date $T_S < T_1$ to satisfy (PC) at smaller costs of delay, $e^{\Delta_r T_S} < e^{\Delta_r T_1}$. Depending on the region of non-convexities the short-term payment may be interior (as in this example) or as early as possible (at time τ) as e.g., for the following standard arrival time distributions:

¹¹ This case applies e.g., for the exponential distribution (cf. Hartman-Glaser, Piskorski, and Tchisty (2012)), where $T_1 = a^2 \frac{c'(a)}{v+c(a)}$. It does not apply for all parameter values κ of the more general Weibull distribution.

¹² In general, these two payout dates are unique (as in Figure 3). In some knife-edge cases, there may be multiple solutions for T_S and T_L . (Then, in the associated plot of discounting costs $e^{\Delta_r t}$ against information $L(t)$ all points would lie on one line). In this knife-edge case, it is without loss of generality to pick 2 payout dates so that $T_S < T_1$ and $T_L > T_1$.

Corollary 1 *In the absence of regulation, for Δ_r sufficiently low and v sufficiently high, there are two payment dates in Examples 1, 2 (for $\kappa > 1$), and 3. The short-term payment date satisfies $T_S = 0$ for all examples.*

This result follows from the fact that $\lim_{v \rightarrow \infty} T_1(a, v) = 0$ and that all of these survival distributions imply that L is locally convex at date 0. One can now consider the effect of the mandatory deferral period on contract design.

Corollary 2 *Suppose $\tau < T_1$ and $e^{\Delta_r T_1} > C\left(\frac{c'(a)}{v+c(a)} \middle| a, \tau\right)$, then if regulation constrains the short-term payment date so that $T_S(a|\tau) = \tau$, a marginal increase in τ induces the principal to lower $T_L(a|\tau)$.*

Intuitively, the principal responds to the required deferral of the short-term payment date by shifting the long-term date to an earlier point (see Figure 4 for graphical illustration).

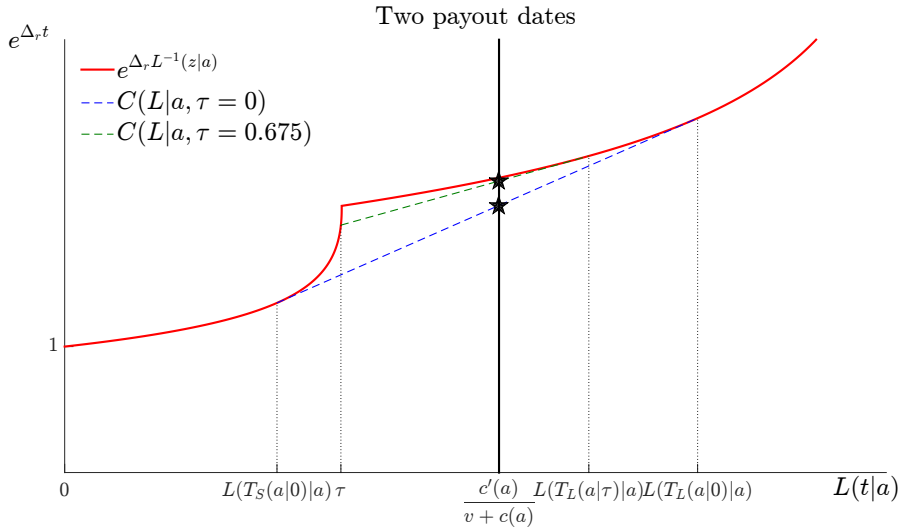


Figure 4. Adjustment of T_L under deferral regulation. This graph plots the cost of information $e^{\Delta_r L^{-1}(z|a)}$ for the piecewise mixed exponential arrival time distribution as in the right panel of Figure 3 and shows the optimal short- and long-term payout dates for both the unconstrained case as well as for $\tau = 0.675$. Functional forms and parameter values are as in Figure 3.

Solution to full problem. It is now possible to characterize the solution to the full compensation design problem for any given action a by combining the results of the relaxed compensation design problem and the results of this section.

Theorem 1 *For any action $a \geq \underline{a}(\tau)$, the optimal contract is*
1) characterized by Proposition 1 if $v \leq \frac{c'(a)}{L(\hat{T}_1(a|\tau)|a)} - c(a)$, and
2) characterized by Proposition 2 otherwise.

Intuitively, if the agency rent derived from the optimal contract in the relaxed Problem, $\frac{c'(a)}{L(\hat{T}_1(a|\tau)|a)} - c(a)$, (see Proposition 1) is sufficiently high as to satisfy (PC), then this contract is also the solution to the full problem. The proof then proceeds to show that (PC) must bind in all other cases, so that either the single-date or two-date contract applies.

Role of clawback clause. We conclude the section on compensation design by analyzing compensation contracts in the absence of the clawback clause, thereby highlighting its role.

Lemma 6 *For $a \geq \underline{a}(\tau)$, the optimal contract in the absence of (CLAW) is described in Theorem 1. For, $a < \underline{a}(\tau)$, a fraction $\gamma \in (0, 1]$ of the total expected discounted pay to the agent, $v + c(a)$, is paid out at date τ regardless of survival, where $\gamma = \frac{v+c(a)-\frac{c'(a)}{L(\tau|a)}}{v+c(a)}$. In the latter case, the associated wage costs are given by $(v + c(a)) e^{\Delta_r \tau}$.*

Thus, as long as an action is implementable with a contract that makes pay fully contingent on survival, i.e., $a \geq \underline{a}(\tau)$, the clawback clause plays no role in our compensation design analysis. The principal would endogenously make all pay conditional on survival (Lemma 1 is a special case as $\underline{a}(0) = 0$). However, without the clawback clause, the principal has the ability to design compensation contracts that implement low actions, $a < \underline{a}(\tau)$, by making unconditional payments at date τ to satisfy (PC). Thus, the clawback clause only “bites” in conjunction with the minimum deferral requirement $\tau > 0$ if it prohibits the implementation of an action altogether.

This completes the characterization of optimal contract design in the presence of deferral requirements. Given the associated minimal wage cost $W(a|\tau)$ for a given action a we now analyze the action that the principal induces in equilibrium, $a^*(\tau)$, i.e., determine the solution to Problem 2.

3.2 Equilibrium action

Following the structure of our analysis of the compensation design problem, we first study the profit-maximizing action choice in the absence of (PC), i.e.,

$$\hat{a}(\tau) = \arg \max_a \pi(a) - \hat{W}(a|\tau), \quad (24)$$

where $\hat{W}(a|\tau)$ denotes the compensation cost function of the relaxed problem given by (15). To facilitate exposition, we require that $c'(a)$ is sufficiently convex as to ensure strict convexity of $\hat{W}(a|\tau)$.¹³ Then, by concavity of π , the action $\hat{a}(\tau)$ is the unique solution to:

$$\pi'(\hat{a}(\tau)) = \hat{W}_a(\hat{a}(\tau)|\tau). \quad (25)$$

Let $\hat{v}(\tau) := \frac{c'(\hat{a}(\tau))}{L(\hat{T}_1(\hat{a}(\tau)|\tau)|a)} - c(\hat{a}(\tau))$ denote the agency rent associated with $\hat{a}(\tau)$, then

Lemma 7 *The equilibrium action satisfies*

$$a^*(\tau) = \begin{cases} \hat{a}(\tau) & v \leq \hat{v}(\tau) \\ \arg \max_{a \geq \underline{a}(\tau)} \pi(a) - [v + c(a)] C\left(\frac{c'(a)}{v+c(a)} \middle| a, \tau\right) & v > \hat{v}(\tau). \end{cases} \quad (26)$$

Thus, unless (PC) is slack in equilibrium ($v \leq \hat{v}$), the equilibrium action is not necessarily pinned down by a first-order condition, and may instead be given by the lowest implementable action $\underline{a}(\tau)$. Since $\underline{a}(0) = 0$ and by our assumptions on the cost function, such a corner case is only possible in the presence of sufficiently stringent regulation.

4 Regulator's Problem

Our analysis of the regulator's problem consists of two parts. First, we study the comparative statics of the equilibrium action $a^*(\tau)$ in the minimum deferral time τ (Section 4.1). This does not only set the stage for the analysis of the *normative* (welfare maximizing) choice of the deferral time τ^* (Section 4.2), it also allows us to develop the *positive* implications of deferral regulation. Since we restrict the set of possible regulatory interventions to a single policy, deferral regulation, our “normative” analysis cannot aim to solve for the globally optimal policy.¹⁴

4.1 Comparative statics analysis

4.1.1 Relaxed problem

We initially study the effect of mandatory deferral on the action in the absence of (PC). In this case, for any given τ , the equilibrium action is always pinned down by a first-order

¹³ As usual in static principal-agent models, the assumption that implies validity of the first-order approach, i.e., Assumption 2 in our setup, does not imply that the principal's problem is quasiconcave in the action (see Grossman and Hart (1983) or Jewitt, Kadan, and Swinkels (2008)).

¹⁴ Clearly, for deferral regulation to be part of a more general optimal regulatory policy mix, we would conjecture that there should be restrictions or costs involved in making the principal directly internalize the externalities.

condition (25). In particular, since deferral regulation only affects the cost of inducing an action, $\hat{W}(a|\tau)$, but not its benefit, $\pi(a)$, the impact of deferral on the equilibrium action is solely determined by the effect on the *marginal* cost. We note that while regulation must increase the *level* of costs, i.e., $\frac{\partial \hat{W}(a|\tau)}{\partial \tau} \geq 0$, the effect of regulation on marginal cost is a priori unclear. By the implicit function theorem,

$$\text{sgn} \left(\frac{d\hat{a}(\tau)}{d\tau} \right) = \text{sgn} \left(-\frac{\partial^2 \hat{W}(a|\tau)}{\partial a \partial \tau} \right), \quad (27)$$

where the cross-partial satisfies

$$\frac{\partial^2 \hat{W}(a|\tau)}{\partial a \partial \tau} = \underbrace{\left(\Delta_r - \frac{d \log L}{dt} \Big|_{t=\hat{T}_1(a|\tau)} \right)}_{\geq 0} \underbrace{\left[\frac{c''(a)}{c'(a)} - \frac{L_a}{L} \right]}_{> 0} - \underbrace{\frac{\partial}{\partial a} \frac{d \log L}{dt} \Big|_{t=\tau}}_{\leq 0}. \quad (28)$$

We initially consider the effect of setting τ marginally above the unconstrained payout time $\hat{T}_1(a|0)$. Then, the first term is zero by the first-order condition (16).

Proposition 3 *Let $\tau = \hat{T}_1(a|0)$, then*

$$\text{sgn} \left(\frac{d\hat{a}(\tau)}{d\tau} \right) = \text{sgn} \left(\frac{d\hat{T}_1(a|0)}{da} \right) = \text{sgn} \left(\frac{\partial}{\partial a} \frac{d \log L}{dt} \Big|_{t=\tau} \right). \quad (29)$$

Intuitively, whether the marginal effect of deferral on the equilibrium action is positive or negative is akin to the question whether the *unconstrained* timing choice, $\hat{T}_1(a|0)$, is increasing or decreasing in a , see (17). If $\frac{d\hat{T}_1(a|0)}{da} > 0$ and the principal is constrained by regulation to choose further deferral, the regulator effectively nudges the principal to induce a higher equilibrium action (lower risk-taking). In the opposite case, $\frac{d\hat{T}_1(a|0)}{da} < 0$, a large minimum deferral time induces a lower action. As shown previously, all comparative statics are possible depending on the information process. The Weibull family (Example 2), with the exponential distribution as a special case ($\kappa = 1$), represents the knife-edge case.

Corollary 3 *Let $\tau = \hat{T}_1(a|0)$, then a marginal increase in τ decreases the action in Example 1, has no effect in Example 2, and increases the action in Example 3.*

We have thus shown that *marginal* interference in the timing of pay has non-trivial effects on the equilibrium action and requires precise knowledge about the underlying information process. It is key for these comparative statics results that for small, formally

infinitesimal, perturbations the payout time still satisfies the first-order condition (17). However, as the regulator mandates further deferral and regulation binds, he induces first-order deviations from the optimality condition: The induced growth rate of the costs of delay exceeds the growth rate of information, $\Delta_r > \frac{d \log L}{dt} \Big|_{t=\hat{T}_1(a|\tau)}$, which unambiguously pulls towards an increase in marginal costs (see first term in (28)). Therefore, regardless of the underlying information process, sufficiently stringent regulation eventually reduces the action. More formally,

Proposition 4 *There exists a minimum deferral period $\check{\tau}$ such that for all $\tau > \check{\tau}$*

- 1) *the minimum deferral period binds, $\hat{T}_1(\hat{a}(\tau)|\tau) = \tau$, and*
- 2) *the action is strictly decreasing in τ with $\lim_{\tau \rightarrow \infty} \hat{a}(\tau) = 0$.*

We will now study how the principal's choices change in presence of the agent's (PC).

4.1.2 Full problem

We first consider the case of a sufficiently low outside option, $0 < v < \hat{v}(0)$, so that (PC) is slack in the absence of regulation. Then, for τ sufficiently small, the comparative statics described in Propositions 3 and 4 initially still apply. The left panel of Figure 5 illustrates the comparative statics for the example of the exponential distribution. However, as the action $\hat{a}(\tau)$ is eventually reduced (and approaching zero) according to Proposition 4, so is the associated agent's payoff $V_A = \frac{c'(\hat{a}(\tau))}{L(\tau|\hat{a}(\tau))} - c(\hat{a}(\tau))$. Therefore, starting from slack (PC) in the absence of regulation ($\tau = 0$) there exists a threshold level for the deferral period, $\hat{\tau}$, at which (PC) starts to bind.

$$\frac{c'(\hat{a}(\hat{\tau}))}{L(\hat{\tau}|\hat{a}(\hat{\tau}))} - c(\hat{a}(\hat{\tau})) = v. \quad (30)$$

At this threshold level $\hat{\tau}$, the optimal action ignoring (PC), $\hat{a}(\hat{\tau})$, is equal to the lowest implementable action i.e., $\hat{a}(\hat{\tau}) = \underline{a}(\tau)$. From then on, the comparative statics are akin to the comparative statics of τ with binding (PC) as illustrated in the right panel of Figure 5. We now characterize the new and robust result in the presence of (PC).

Proposition 5 *For any $v > 0$, a sufficiently stringent deferral period implies that*

- 1) *there is a single payment date at date τ ,*
- 2) *the agent's participation constraint binds, and*
- 3) *the equilibrium action $a^*(\tau) = \underline{a}(\tau)$ is strictly increasing in τ .*

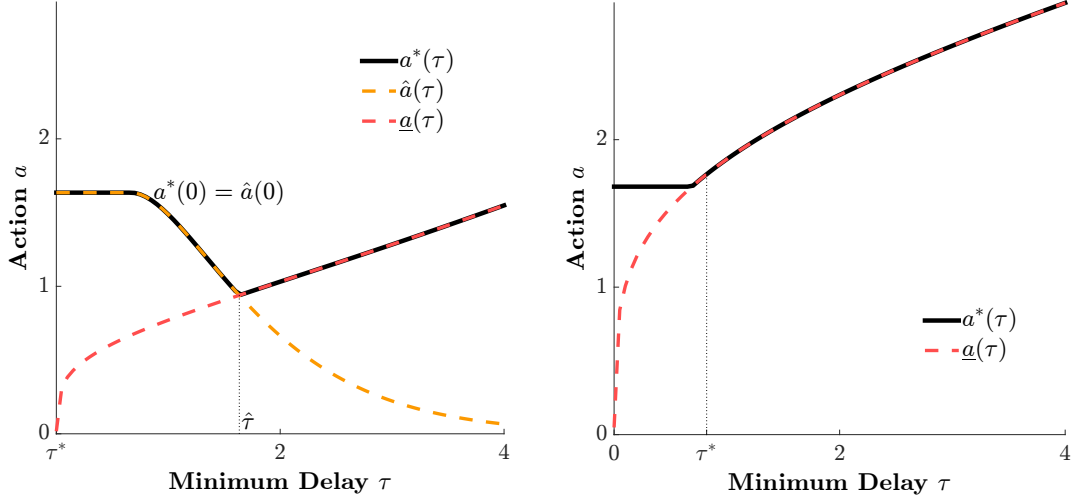


Figure 5. Comparative statics. This graph plots equilibrium effort $a^*(\tau)$ as a function of the mandatory delay of compensation τ for the case of an exponential distribution. The left panel shows a case with low $v = 0.1$, such that the participation constraint does not bind for $\tau = 0$, whereas it does bind in the right panel with $v = 5$. For both panels $\Delta_r = 1.5$, the revenue function satisfies $\pi(a) = 40a$ and effort costs are given by $c(a) = \frac{1}{6}a^3$. The welfare maximizing τ^* (cf. section 4.2) is obtained for $x(a) = \bar{x} - a$, where we normalize $\bar{x} = 0$.

Thus, in the presence of (PC), sufficiently stringent regulation has an unambiguously positive effects on the action. In this case, the principal is “forced” by regulation to increase his action as it is no longer determined by a first-order condition, but instead given by the lowest possible implementable action that allows satisfying (PC), (DEF) and (CLAW), i.e., $a^*(\tau) = \underline{a}(\tau)$ (see Figure 5 for $\tau > \hat{\tau}$).

To understand the intuition, recall that regulatory interference in the timing of pay does not give the principal per se more incentives to avoid disaster events (in contrast to, for example, equity capital regulation).¹⁵ In particular, when facing regulatory constraints on one dimension of the compensation contract, the principal will, in general, try to adjust other dimensions of the contract. In our setup, the three relevant dimensions are 1) the agent’s net level of pay V_A , 2) the contingency of pay (success versus failure) and, 3), the timing of pay. If τ is sufficiently high, the principal is unable to adjust on any margin. First, binding (PC) fixes the agency rent at v . Second, (CLAW) forces the principal to only pay in the absence of the disaster, which combined with the deferral requirement τ , makes it impossible to implement a sufficiently low action.

If the principal can adjust on either of these margins, the regulatory rationale is not robust. If (PC) is slack, then the principal can adjust the size of the pay (see Section 4.1.1). If (PC) is binding but the deferral period is not sufficiently stringent and

¹⁵ See e.g., Harris, Opp, and Opp (2015).

only constrains the short-term payment T_S of a two-date contract, then the principal can both change the timing of the long-term payment and the fraction of pay made at each date. Finally, if the regulator does not impose CLAW (see Lemma 6), the principal responds by making (part of the) compensation regardless of whether there was success or failure when being confronted with sufficiently stringent deferral regulation. In all of these cases, these adjustments may lead to a lower action.

4.2 Optimal deferral periods

We now consider the regulator's optimal choice of τ . Recall our definition of welfare from (4), which can be written as

$$\omega(\tau) = \pi(a^*(\tau)) - c(a^*(\tau)) - \lambda_x x(a^*(\tau)) - \int_{\tau}^{\infty} (e^{\Delta_r t} - 1) dB(t). \quad (31)$$

By imposing a binding deferral regulation τ , the regulator, thus, affects welfare directly through the deadweight cost of delay (last term in 31) and indirectly via the effect on the equilibrium action $a^*(\tau)$. Using the comparative statics results from the previous section 4.1 together with (31) it is then straightforward to determine the welfare maximizing deferral period $\tau^* = \arg \max_{\tau} \omega(\tau)$, which satisfies the following basic properties:

Lemma 8 *For any bounded λ_x , the optimal (welfare-maximizing) deferral period τ^* is finite and satisfies $\frac{d}{d\lambda_x} \tau^* \geq 0$. If regulation binds, $\tau^* > 0$, then $a^*(\tau^*) > a^*(0)$.*

To see the intuition for the last statement in Lemma 8 note that $\omega(\tau)$ is simply the sum of the principal's and the agent's expected payoff minus the expected externality. The principal's payoff can clearly not increase with tighter restrictions on the contract space and the agent's payoff does either not depend on τ or a when (PC) binds, or, with (PC) slack, decreases in τ and increases in a . Combined with $x'(a) < 0$, this implies the intuitive result that binding deferral regulation can be welfare increasing only if it is successful in increasing the equilibrium action. However, even if $a^*(\tau) > a^*(0)$ for some τ it is still unclear whether this increase is welfare-enhancing since deferral also has a first-order effect on the deadweight loss. In fact, for sufficiently stringent deferral regulation the deadweight loss dominates, such that welfare is eventually decreasing in τ . Whether imposing a binding deferral period $\tau^* > 0$ is optimal at all, then depends on the importance of the externality in the welfare measure. It is fairly intuitive that the optimum deferral time (which increases the equilibrium action) is increasing in λ_x .

Using Lemma 8 together with the comparative statics results in Section 4.1 it is apparent that the optimality of a binding deferral period crucially depends on the arrival

time distribution as well as whether (PC) binds or not without regulation. Consider the example from Figure 5. For low values of v , (PC) is slack in the absence of regulation, so that $a^*(\tau)$ is first decreasing and increasing only for $\tau > \hat{\tau}$ (see left panel of Figure 5). Thus, relatively stringent regulation is required to increase $a^*(\tau)$ above the unregulated level. However, at that level the associated deadweight loss is so large that it dominates the reduction in the externality. Hence, $\tau^* = 0$. For higher values of v , (PC) binds also absent regulation, so that $a^*(\tau)$ is globally increasing (see right panel of Figure 5). It is now strictly optimal to set $\tau^* > 0$ by trading off the marginal reduction in the externality with the marginal increase in the deadweight loss.

For this example, it can be shown that if a binding $\tau^* > 0$ is optimal for some v' , then it is strictly optimal for any $v > v'$. When we thus stipulate that more competition for human capital raises v , our results suggest that deferral regulation is more likely to be beneficial when competition for human capital is more intense. While these results are obtained assuming the specific exponential arrival time distribution, they clearly illustrate how our positive results on the impact of deferral regulation on $a^*(\tau)$ can be used to make normative statements.

5 Extensions

In this section, we lay out that the optimal compensation contract that solves Problem 1* applies to more general information environments and principal-agent relationships. The only key assumptions of our setup are 1) risk-neutrality, 2) informative signals about the agent's action are revealed over time, and, 3) relative impatience of the agent. Risk-neutrality implies that for any payout date t , the agent is only rewarded for the outcome with the highest likelihood ratio. The second feature implies that this likelihood ratio is an increasing function of time t . Finally, impatience implies that deferring pay is costly, which introduces a trade-off between information and deadweight costs of delay.

5.1 Generalization of information processes

To see how our setup can accommodate more general information structure, let $\tilde{h}_t$ denote the history of realized signals by time t , and let H_t denote the corresponding set of all possible histories at time t . Then, for each possible history $\tilde{h}_t$, we can compute the associated likelihood ratio $\tilde{L}(\tilde{h}_t|a)$. However, for the purpose of designing incentive pay,

we only need to consider the history associated with the highest likelihood ratio, i.e.,

$$L(t|a) = \max_{\tilde{h}_t \in H_t} \tilde{L}(\tilde{h}_t|a) \quad (32)$$

In our binary information environment (disaster has occurred or not), Assumption 1 insured that a history without failure is the history with the largest likelihood-ratio at each point in time t (Lemma 1). However, in more general environments, it may be possible that early failure is an indicator of taking the intended long-run action (see e.g., Manso (2011) or Zhu (2015)). (32) formalizes how to accommodate this.

We also note that our results readily extend to environments with discrete information arrival and/ or a finite horizon. For example, one may envision situations in which information arrives continuously, as in our setting, but the principal can only observe this information at discrete points in times, say because of an annual audit. If (PC) is slack, Proposition 1 essentially applies without modification. The optimal payout time is still $\arg \min_{t \in \varsigma} \frac{e^{\Delta r t}}{L(t|a)}$; now minimized over the (possibly) finite set of information dates ς . Likewise, when (PC) binds one can readily define $C(z|a, \tau)$ and apply Proposition 2. However, for discrete information arrival, there are now generically two payout dates as there generically does not exist a single payout date that satisfies $L(t|a) = \frac{c'(a)}{v+c(a)}$.

5.2 Multi-task environment

Consider the following multi-task environment inspired by Bénabou and Tirole (2016). The agent's action consists of two tasks a and q , where only task a has persistent effects. In particular, the agent first needs to exert unobservable effort $q \in [0, 1]$ at associated effort cost $k(q)$ to create a business opportunity with probability q . If no opportunity arrives, the game ends. If a business opportunity has arrived, which is immediately observable, the agent then chooses action a at cost $c(a, q)$. We make standard convexity assumptions on $k(\cdot)$ and $c(\cdot)$, and, for simplicity, assume that the agent has a sufficiently low outside option. Since the agent will optimally never receive any compensation if no business opportunity has been created, we may write the agent's payoff as

$$V_A = q \left(\int_{\tau}^{\infty} dB(t) - c(a, q) \right) - k(q).$$

Then, the principal's compensation design problem to implement a given a and q is

$$W(a, q|\tau) = q \min_{B(t)} \int_{\tau}^{\infty} e^{\Delta r t} dB(t) \quad \text{s.t.}$$

$$\begin{aligned}\int_{\tau}^{\infty} dB(t) &= k'(q) + c(a, q) & (\text{ICq}) \\ \int_{\tau}^{\infty} L(t|a) dB(t) &= \frac{\partial c(a, q)}{\partial a} & (\text{ICa}) \\ dB(t) &\geq 0 \quad \forall t & (\text{LL})\end{aligned}$$

Since a and q are fixed, it is immediate that the solution to this compensation design problem is given by Proposition 2.

6 Conclusion

This paper characterizes optimal compensation contracts in a principal-agent framework in which signals about the agent's action are revealed over time. Optimal contracts are surprisingly simple. Regardless of the information process, there are at most two payout dates. We use this basic framework to shed light on recent regulatory initiatives to impose minimum deferral period and clawback requirements on compensation contracts in the financial sector. Facing such restrictions on the timing dimension of compensation contracts, the principal responds by restructuring the contract along other dimensions. We show that these adjustments, most notably of the level of pay, generally vary with the information environment and can, thus, have an ambiguous effect on risk-taking. However, a robust insight with respect to regulation emerges. If the agent's outside option is sufficiently high so that the level of the compensation package is determined by the outside option, a minimum deferral requirement has the intended effect of reducing risk-taking.

Our paper provides two avenues for future research. First, one of the key restrictions of our paper is that we only analyze the implications of a particular regulatory policy, namely a minimum deferral period. It would be interesting to study when deferral regulation should be part of optimal and unified financial regulation. We conjecture that there are interdependencies between various regulations that directly target the principal's risk-taking incentives, such as capital requirements, and restrictions on compensation design. Such a more general framework may also take into account general equilibrium effects of regulation. For example, by affecting the minimum deferral period for an entire industry, regulation may also affect the outside option of the agent.

Second, it may be worthwhile to generalize our paper to repeated moral hazard settings with persistent effects. There has been recent interest in these types of problems with a risk-averse agent (see e.g., Jarque (2010) or Sannikov (2014)). However, it may

be possible to exploit the simplicity of our results implied by risk-neutrality to deepen our intuitive understanding of these settings for arbitrary information processes.

Appendix A Proofs

Proof of Lemma 1. Follows from the proof of Lemma 6 below. **Q.E.D.**

Proof of Lemma 2. We will show that Assumption 2 is a sufficient condition for the agent's expected discounted pay V_A to be concave in a for any contract solving Problem 1. With V_A given by (7), given any contract, the agent's optimal choice of a must satisfy the first-order condition (8) which can be rewritten as¹⁶

$$\int_{\tau}^{\infty} \frac{L(t|a)}{c'(a)} dB(t) = 1. \quad (33)$$

The second-order condition for a maximum is then satisfied if

$$\begin{aligned} \frac{\partial^2 V_A}{\partial a^2} &= \int_{\tau}^{\infty} \frac{S_{aa}(t|a)}{S(t|a)} dB(t) - c''(a) \leq 0 \\ &\iff \int_{\tau}^{\infty} \frac{S_{aa}(t|a)}{S_a(t|a)} \frac{L(t|a)}{c'(a)} dB(t) \leq \frac{c''(a)}{c'(a)}. \end{aligned} \quad (34)$$

Now, note that the left-hand side in (34) is a weighted average of $S_{aa}(t|a)/S_a(t|a)$ over $t \in [\tau, \infty)$ with the weight for each t given by $w(t) := \frac{L(t|a)}{c'(a)} dB(t) \geq 0$ and the weights adding up to one from (33). The result then follows as for any given t it holds from Assumption 2 that $S_{aa}(t|a)/S_a(t|a) \leq c''(a)/c'(a)$. **Q.E.D.**

Proof of Lemma 3. The first result follows directly from $L(0|a) = \frac{S_a(0|a)}{S(0|a)}$ and the fact that $S(0|a) = 1$ for all $a \in \mathcal{A}$. Next, straightforward differentiation gives

$$L_t(t|a) = \frac{S_{at}(t|a) S(t|a) - S_a(t|a) S_t(t|a)}{S^2(t|a)} = -h_a(t|a),$$

and the second result follows together with Assumption 1. **Q.E.D.**

Proof of Lemma 4. This follows from the following general result. Consider two twice differentiable functions $f(t)$ and $g(t)$. Then, f is a strictly convex transformation of g if $\frac{f''(t)}{f'(t)} > \frac{g''(t)}{g'(t)}$ for all t (see e.g., Gollier (2004)). **Q.E.D.**

Proof of Proposition 1. See main text. **Q.E.D.**

¹⁶Note that from Assumption 3, the agent's optimal action choice must be interior.

Proof of Lemma 5. Note that convexity of the cost of information ensures that the problem in (13) is a convex problem. The results then follow from Proposition 1 together with $\hat{T}_1(a|0) < \infty$ and Assumption 5, which ensures convexity of the cost of information for high τ . **Q.E.D.**

Proof of Proposition 2. Consider Problem 1* with binding (PC). From (19), minimizing the wage costs is equivalent to minimizing a weighted average of the cost of information $\int_{L(\tau|a)}^{L(\infty|a)} e^{\Delta_r L^{-1}(z|a)} dw(z)$, with the weights given by $dw(z) = dB(L^{-1}(z|a)) / (v + c(a))$, which are nonnegative from (LL) and sum up to one from (PC). Any solution satisfying (LL) and (PC) must, thus, lie on the convexification of the cost of information $C(z|a, \tau)$ for $z \geq L(\tau|a)$. The result then follows as (20) pins down the weighted average of $L(\cdot)$ from $\int_{L(\tau|a)}^{L(\infty|a)} z dw(z) = \frac{c'(a)}{v+c(a)}$. **Q.E.D.**

Proof of Theorem 1. The first assertion directly follows from the fact that $v \leq \frac{c'(a)}{L(\hat{T}_1(a|\tau)|a)} - c(a)$ ensures that (PC) is satisfied at the solution to the relaxed problem as characterized in Proposition 1, such that the relaxed and the full problem are in fact equivalent. It remains to show, that in all other cases the participation constraint is binding, thus proving the second assertion in the theorem. The proof is by contradiction. So, fix τ and take any $a \in \mathcal{A}$ for which $\hat{T}_1(a|\tau) > T_1(a) \geq \tau$ such that $v > \frac{c'(a)}{L(\hat{T}_1(a|\tau)|a)} - c(a)$ and assume that the participation constraint under the optimal contract is slack.¹⁷ Then, the optimal contract would specify a single payout time $\tilde{T} = \arg \min_{\tau \leq T < T_1(a)} \frac{e^{\Delta_r T}}{L(T|a)} < \hat{T}_1(a|\tau) = \arg \min_{T \geq \tau} \frac{e^{\Delta_r T}}{L(T|a)}$, to minimize the discounting cost per unit of information for all admissible payout times with slack participation constraint. We will now show that there exists a feasible contract with lower compensation costs and binding participation constraint. To see this, consider a convex combination of the single-payment contracts with payout times $\tilde{T}$ and $\hat{T}_1(a|\tau)$ respectively, i.e., a contract with two payments $dB(\tilde{T}) = (1 - w) c'(a) / L(\tilde{T}|a)$ and $dB(\hat{T}_1(a|\tau)) = w c'(a) / L(\hat{T}_1(a|\tau)|a)$ at $\hat{T}_1(a|\tau)$, which, by construction, satisfies the incentive constraint for any weight $0 \leq w \leq 1$. Further, wage cost are clearly decreasing in w . Thus, one wants to choose w as high as possible, i.e., until the participation constraint binds or $0 < w = \left(\frac{1}{L(\tilde{T}|a)} - \frac{1}{L(T_1(a)|a)} \right) / \left(\frac{1}{L(\tilde{T}|a)} - \frac{1}{L(\hat{T}_1(a|\tau)|a)} \right) < 1$. This implies a single payment at $\tilde{T}$ is strictly dominated by the considered convex combination with binding participation constraint. **Q.E.D.**

Proof of Lemma 6. We will start by showing the first assertion in Lemma 6 that the constraint (CLAW) is irrelevant for compensation design as long as $a \geq \underline{a}(\tau)$. Note that

¹⁷ That $T_1(a) \geq \tau$ is equivalent to $a \geq \underline{a}(\tau)$ and thus ensures implementability of a .

Lemma 1 then follows as a special case as $\underline{a}(\tau) = 0$. So, consider Problem 1 without the claw-back constraint (CLAW). It is useful to introduce the following notation:

$$db(t) = \begin{cases} b_t + db_S(t) & \text{if } Y(t) = 0 \\ b_t + db_F(t) & \text{if } Y(t) = 1, \end{cases}$$

where b_t denotes the “unconditional” payment at t , $db_S(t)$ the payment in case no disaster has occurred up to t , and $db_F(t)$ the payment in case the disaster has occurred prior to or at t . As in our model no more information is revealed after the (first) bad event and we have differential discounting, it is without loss of generality to stipulate that, under the optimal contract, $b_t = 0$ for all $t > \tau$ and $db_F(t) > 0$ only if $\lim_{x \rightarrow t^-} Y(x) = 0$, i.e., the unconditional payment is made as early as possible and any payment that rewards for failure is made immediately in the event of the disaster. Then we can rewrite Problem 1 as follows,

$$W(a|\tau) = \min_{b_\tau, db_S(t), db_F(t)} \left\{ e^{-r_P \tau} b_\tau + \int_\tau^\infty e^{-r_P t} S(t|a) db_S(t) + \int_\tau^\infty e^{-r_P t} f(t|a) db_F(t) \right\} \quad \text{s.t.}$$

$$e^{-r_A \tau} b_\tau + \int_\tau^\infty e^{-r_A t} S(t|a) db_S(t) + \int_\tau^\infty e^{-r_A t} f(t|a) db_F(t) - c(a) \geq v \quad (35)$$

$$\int_\tau^\infty e^{-r_A t} S_a(t|a) db_S(t) + \int_\tau^\infty e^{-r_A t} f_a(t|a) db_F(t) - c'(a) = 0 \quad (36)$$

$$b_\tau, db_S(t), db_F(t) \geq 0 \quad \forall t, \quad (37)$$

where we have replaced (IC) by the respective first-order condition (36).¹⁸ The proof now proceeds in three steps. We first show that paying the agent after failure, $db_F(t) > 0$, cannot be optimal (Lemma A1). In a second step we then show that, further, $b_\tau = 0$ under the optimal contract implementing some $a \geq \underline{a}(\tau)$ (Lemma A2). The result then follows from the fact that action a can be implemented with a contract that only pays in the absence of a bad event if and only if $a \geq \underline{a}(\tau)$.

Lemma A1 *It is never (strictly) optimal to pay the agent after failure, i.e., $db_F(t) = 0$ for all t .*

Proof of Lemma A1. The proof is by contradiction. So, assume to the contrary that the optimal contract has $db_F(t') > 0$ for some $t' \geq \tau$. We distinguish two cases, depending

¹⁸ Note that (36) is only a necessary condition for incentive compatibility and we are, hence, considering a relaxed problem. As we will show, the solution to the relaxed problem will be from the class of contracts that do not pay after failure. For this class of contracts, however, assumption 2 ensures that (36) is a necessary and sufficient condition for incentive compatibility (cf. Lemma 2).

on whether $f_a(t'|a) \geq 0$. Assume, first, that $f_a(t'|a) \geq 0$. We will show that there exists a feasible perturbation of the initial contract resulting in lower expected compensation costs to the principal. In particular we set $db_F(t') = 0$ and adjust $db_S(t')$ and b_τ by $\Delta b_S(t')$ and Δb_τ respectively, in order to keep satisfying (36) and (35). We thus require that

$$\begin{aligned} S_a(t'|a) \Delta b_S(t') - f_a(t'|a) db_F(t') &= 0, \\ e^{-r_A \tau} \Delta b_\tau + e^{-r_A t'} [S(t'|a) \Delta b_S(t') - f(t'|a) db_F(t')] &= 0, \end{aligned}$$

implying

$$\begin{aligned} \Delta b_S(t') &= \frac{f_a(t'|a)}{S_a(t'|a)} db_F(t') \geq 0, \\ \Delta b_\tau &= -e^{-r_A(t'-\tau)} \frac{S^2(t'|a)}{S_a(t'|a)} \frac{\partial h(t'|a)}{\partial a} db_F(t') > 0, \end{aligned}$$

where we have used assumption (1) on the hazard rate, showing that the perturbation is feasible. It remains to show that it reduces expected compensation costs, but we immediately get

$$\begin{aligned} \Delta W &= e^{-r_P \tau} \Delta b_\tau + e^{-r_P t'} S(t'|a) \Delta b_S(t') - e^{-r_P t'} f(t'|a) db_F(t') \\ &= e^{-r_A t'} \left(e^{\Delta_r t'} - e^{\Delta_r \tau} \right) \frac{S^2(t'|a)}{S_a(t'|a)} \frac{\partial h(t'|a)}{\partial a} db_F(t') \leq 0 \end{aligned}$$

where we have again used assumption (1) on the hazard rate as well as $\Delta_r = r_A - r_P > 0$.¹⁹ The proof for the remaining case where $f_a(t'|a) < 0$ is immediate, as in this case setting $db_F(t') = 0$ relaxes the incentive constraint, thus allowing to reduce any incentive pay while ensuring that the participation constraint continues to hold by increasing b_τ , which minimizes discounting costs. We omit a formal argument for brevity. ■

So far we have shown that any contract that pays after failure or unconditionally at any $t > \tau$ is dominated. We will now show that also an unconditional payment at $\tau > 0$ cannot be optimal as long as the action to be implemented satisfies $a \geq \underline{a}(\tau)$, which solves (21).²⁰

Lemma A2 *Under the optimal contract implementing action $a \geq \underline{a}(\tau)$, the agent does not receive an unconditional payment, i.e., $b_t = 0$ for all t .*

Proof of Lemma A2. It remains to show that $b_\tau > 0$ cannot be part of an optimal

¹⁹ Note that the inequality is strict for $t' > \tau$.

²⁰ If no solution $\underline{a}(\tau) > 0$ to (21) exists, we set $\underline{a}(\tau) = 0$.

contract implementing $a \geq \underline{a}(\tau)$. First, it can clearly never be optimal to have $b_\tau > 0$ if (35) is slack. The proof for the case where (35) is binding is by contradiction. So assume that the solution to the principal's problem involves $b_\tau > 0$. Then, using the notation $dB(t) = e^{-rA^t} S(t|a) db_S(t)$ and $L(t|a) = S_a(t|a)/S(t|a)$, we can rewrite the principal's problem as follows:

$$W(a|\tau) = \min_{B(t) \geq 0} \left\{ \int_{\tau}^{\infty} (e^{\Delta_r t} - e^{\Delta_r \tau}) dB(t) + e^{\Delta_r \tau} [c(a) + v] \right\} \quad \text{s.t.} \\ \int_{\tau}^{\infty} L(t|a) dB(t) = c'(a), \quad (38)$$

where we have used Lemma A1 and substituted for $b_\tau > 0$ from (35). Now, note that this program is linear in $dB(t)$ and it is, hence, without loss of generality to restrict attention to a single payment date, i.e. there exists a unique $T \geq \tau$ such that $dB(T) > 0$ while $dB(t) = 0$ for all $t \neq T$. Then, substituting out $dB(T)$ using (38) we can write

$$W(a|\tau) = \min_{T \geq \tau} \left\{ (e^{\Delta_r T} - e^{\Delta_r \tau}) \frac{c'(a)}{L(T|a)} + e^{\Delta_r \tau} [c(a) + v] \right\}. \quad (39)$$

From inspection of (39) it is clearly optimal to set $T = \tau$. However, we then get from (38) that the agent's expected incentive compensation satisfies

$$dB(\tau) = \frac{c'(a)}{L(\tau|a)} \geq v + c(a),$$

where the inequality follows from our assumption that $a \geq \underline{a}(\tau)$. To see this consider $g(\tau|a) = c'(a)/L(\tau|a) - (v + c(a))$ and note that $g(\tau|\underline{a}) = 0$ as well as $g_a(\tau|a) = c'(a) \left(\frac{c''(a)}{c'(a)} - \frac{S_{aa}(\tau|a)}{S_a(\tau|a)} \right) / L(\tau|a) > 0$, where we have used Assumption 2. However, this implies together with $b_\tau > 0$ that (35) is slack leading to a contradiction. ■

It remains to show that action a can be implemented with a contract that has $b_t = db_F(t) = 0$ for all t if and only if $a \geq \underline{a}(\tau)$. But this follows directly from Theorem 1.

Finally, we will consider optimal compensation design in the absence of (CLAW) for $a < \underline{a}(\tau)$. With binding (35), the agent's expected discounted payoff is fixed at $v + c(a)$ such that we have the following lower bound on compensation costs

$$W(a|\tau) \geq e^{\Delta_r \tau} (v + c(a)).$$

We will show that there exists a contract as characterized in the Lemma which achieves this bound with equality. To see this consider the contract which specifies $b_\tau = e^{rA\tau} \gamma (v + c(a)) \geq$

0 and $dB(\tau) = e^{-r_A\tau} S(\tau|a) db_S(\tau) = \frac{c'(a)}{L(\tau|a)} \geq 0$ with payments at all other dates and contingencies equal to zero. Clearly, this contract satisfies (36) and it satisfies (35) for $\gamma = \left(v + c(a) - \frac{c'(a)}{L(\tau|a)}\right) / (v + c(a)) > 0$, where the inequality follows from $a < \underline{a}(\tau)$. But clearly it also implies wage costs of $W(a|\tau) = e^{-r_P\tau} b_\tau + e^{\Delta_r\tau} dB(\tau) = e^{\Delta_r\tau} (v + c(a))$. **Q.E.D.**

Proof of Lemma 7. Follows from Theorem 1 together with the fact that $\hat{a}(\tau)$ does not depend on v . **Q.E.D.**

Proof of Proposition 3. See main text. **Q.E.D.**

Proof of Proposition 4. The first assertion follows directly from Lemma 5. As for the second assertion, note that for $\tau \rightarrow \infty$ marginal costs $\partial \hat{W}(a|\tau) / \partial a \rightarrow \infty$ for all $a > 0$. To see this, consider marginal costs which from (15) can be written

$$\frac{\partial \hat{W}(a|\tau)}{\partial a} = \left(\frac{1}{L} \left[\frac{c''(a)}{c'(a)} - \frac{S_{aa}}{S_a} \right] + 1 \right) e^{\Delta_r \hat{T}_1(a|\tau)} c'(a),$$

where, from $L(\hat{T}_1(a|\tau)|a) > 0$ and Assumption 2, the term in square brackets is strictly positive. Hence, $\hat{T}_1(a|\tau) \geq \tau$ directly implies that the expression goes to infinity as $\tau \rightarrow \infty$ for any $a > 0$. It then directly follows from (25) and Assumption 3 that $\lim_{\tau \rightarrow \infty} \hat{a}(\tau) = 0$. It remains to show monotonicity of the action for τ sufficiently large. Using that $\hat{T}_1(a|\tau) = \tau$ for all $a > 0$ and τ sufficiently large, we have from (28) that

$$\begin{aligned} & \lim_{\tau \rightarrow \infty} \frac{\partial^2 \hat{W}(a|\tau)}{\partial a \partial \tau} \\ &= \lim_{\tau \rightarrow \infty} \hat{W}(a|\tau) \lim_{\tau \rightarrow \infty} \left(\left(\Delta_r - \frac{d \log L}{dt} \Big|_{t=\tau} \right) \left[\frac{c''(a)}{c'(a)} - \frac{S_{aa}}{S_a} + L \right] - \frac{\partial}{\partial a} \frac{d \log L}{dt} \Big|_{t=\tau} \right) \\ &= \Delta_r \lim_{\tau \rightarrow \infty} \hat{W}(a|\tau) \lim_{\tau \rightarrow \infty} \left[\frac{c''(a)}{c'(a)} - \frac{S_{aa}}{S_a} + L \right] = \infty, \end{aligned}$$

where we have used Assumption 5. The result then follows again from (25) and strict concavity of $\pi(a)$ (cf. Assumption 3). **Q.E.D.**

Proof of Proposition 5. Assertion two is shown in the main text. With binding participation constraint for $\tau \geq \hat{\tau}$, it then follows from Proposition 2 together with the convexity of the cost of information for τ sufficiently large (cf. Assumption 5 and Lemma 4), that there is only a single payment date. It remains to show that $a^*(\tau)$ is equal to $\underline{a}(\tau)$ for τ sufficiently large. But, as for any $a > 0$ the wage costs with a single payment date and binding participation constraint are given by $W(a|\tau) = (v + c(a)) e^{\Delta_r T_1(a|\tau)}$,

we have

$$W_a(a|\tau) = \left(c'(a) + (v + c(a)) \Delta_r \frac{dT_1(a|\tau)}{da} \right) e^{\Delta_r T_1(a|\tau)},$$

where $\frac{dT_1(a|\tau)}{da} = \frac{L}{L_t} \left(\frac{c''(a)}{c'(a)} - \frac{S_{aa}}{S_a} \right) > 0$ from Assumption 2. Hence, it follows from $T_1(a|\tau) \geq \tau$ that $W_a(a|\tau) \rightarrow \infty$ for $\tau \rightarrow \infty$, such that the principal, optimally chooses the lowest possible value $a^*(\tau) = \underline{a}(\tau)$, implying $T_1(a|\tau) = \tau$. Finally, the comparative statics of $a^*(\tau)$ in τ then follow from $\frac{da(\tau)}{d\tau} = L_t / \left(L \left(\frac{c''(a)}{c'(a)} - \frac{S_{aa}}{S} \right) \right) > 0$ where we have again used Assumption 2. **Q.E.D.**

Proof of Lemma 8. It follows from Proposition 5 that for τ sufficiently large (PC) is binding with a single payment at τ and $a^*(\tau) = \underline{a}(\tau)$. Hence, the deadweight loss of delay can be written as

$$D(\tau|a^*(\tau)) = (e^{\Delta_r \tau} - 1) (v + c(\underline{a}(\tau))),$$

which goes to infinity exponentially. That τ^* must be finite then follows from (31) and the existence of a finite first-best action a^{FB} solving $\pi'(a^{FB}) - \lambda_x x'(a^{FB}) = c'(a^{FB})$, with associated bounded first-best value $\pi(a^{FB}) - c(a^{FB}) - \lambda_x x(a^{FB})$, where we have used Assumption 3.

As for the comparative statics of τ^* in λ_x , note that, by optimality of τ^* , it must hold that $a^*(\tau) < a^*(\tau^*)$ for all $\tau < \tau^*$. Hence, if $\tau^*(\lambda_x)$ is optimal for λ_x we must have $\tau^*(\lambda'_x) \geq \tau^*(\lambda_x)$ for all $\lambda'_x > \lambda_x$ as $x'(a) > 0$ from Assumption 3 and $a^*(\tau)$ does not depend on λ_x .

To prove the last assertion rewrite (31) as follows

$$\omega(\tau) = \Pi(a^*(\tau)|\tau) + V_A(a^*(\tau)|\tau) - \lambda_x x(a^*(\tau)),$$

where $\Pi(a|\tau) := \pi(a) - W(a|\tau)$ denotes the highest profit the principal can attain net of compensation costs and subject to the regulatory constraint. That $\frac{d}{d\tau} \Pi \leq 0$ then trivially follows from optimality. As $x'(a) < 0$ from Assumption 3 it remains to show that V_A cannot strictly increase with τ unless $a^*(\tau)$ increases. To see this note first that V_A as given by (7) is equal to v and thus constant if (PC) binds. When (PC) is slack, we have $V_A(a|\tau) = \frac{c'(a)}{L(\hat{T}_1|a)} - c(a)$ which is decreasing in τ , where we have used that $\hat{T}_1$ is (weakly) increasing in τ and $L_t > 0$ from Lemma 3, and increasing in a

$$\frac{\partial V_A(a|\tau)}{\partial a} = c'(a) \frac{\frac{c''(a)}{c'(a)} - \frac{S_{aa}(\hat{T}_1|a)}{S_a(\hat{T}_1|a)}}{L(\hat{T}_1|a)} > 0,$$

where we have used Assumption 2. **Q.E.D.**

References

- BÉNABOU, R., AND J. TIROLE (2016): “Bonus Culture: Competitive Pay, Screening, and Multitasking,” *Journal of Political Economy*.
- DEMARZO, P., AND D. DUFFIE (1999): “A Liquidity-based Model of Security Design,” *Econometrica*, 67(1), 65–99.
- DEMARZO, P. M., AND Y. SANNIKOV (2006): “Optimal Security Design and Dynamic Capital Structure in a Continuous-Time Agency Model,” *Journal of Finance*, 61(6), 2681–2724.
- GOLLIER, C. (2004): *The Economics of Risk and Time*, vol. 1 of *MIT Press Books*. The MIT Press.
- GROSSMAN, S. J., AND O. D. HART (1983): “An Analysis of the Principal-Agent Problem,” *Econometrica*, 51(1), 7–45.
- HARRIS, M., C. OPP, AND M. OPP (2015): “The System-Wide Effects of Bank Capital Regulation on Credit Supply and Risk-Taking,” Unpublished working paper, University of California, Berkeley.
- HARRIS, M., AND A. RAVIV (1979): “Optimal incentive contracts with imperfect information,” *Journal of Economic Theory*, 20(2), 231–259.
- HART, O., AND B. HOLMSTROM (1987): “The Theory of Contracts,” in *Advances in Economic Theory, Fifth World Congress*, ed. by T. Bewley, pp. 71–155. Cambridge University Press.
- HARTMAN-GLASER, B., T. PISKORSKI, AND A. TCHISTYI (2012): “Optimal securitization with moral hazard,” *Journal of Financial Economics*, 104(1), 186 – 202.
- HOFFMANN, F., R. Inderst, AND M. M. OPP (2015): “Regulating deferred incentive pay,” Discussion paper, University of Bonn, University of Frankfurt, University of California, Berkeley.
- HOLMSTROM, B. (1979): “Moral Hazard and Observability,” *Bell Journal of Economics*, 10(1), 74–91.

- HOPENHAYN, H., AND A. JARQUE (2010): “Unobservable persistent productivity and long term contracts,” *Review of Economic Dynamics*, 13(2), 333 – 349.
- INNES, R. D. (1990): “Limited liability and incentive contracting with ex-ante action choices,” *Journal of Economic Theory*, 52(1), 45–67.
- JARQUE, A. (2010): “Repeated moral hazard with effort persistence,” *Journal of Economic Theory*, 145(6), 2412–2423.
- JEWITT, I., O. KADAN, AND J. M. SWINKELS (2008): “Moral hazard with bounded payments,” *Journal of Economic Theory*, 143(1), 59 – 82.
- MALAMUD, S., H. RUI, AND A. WHINSTON (2013): “Optimal incentives and securitization of defaultable assets,” *Journal of Financial Economics*, 107(1), 111 – 135.
- MANSO, G. (2011): “Motivating Innovation,” *The Journal of Finance*, 66(5), 1823–1860.
- OPP, C. C., M. M. OPP, AND M. HARRIS (2013): “Rating agencies in the face of regulation,” *Journal of Financial Economics*, 108(1), 46–61.
- OPP, M. M. (2012): “Expropriation risk and technology,” *Journal of Financial Economics*, 103, 113 – 129.
- OPP, M. M., AND J. Y. ZHU (2015): “Impatience versus Incentives,” *Econometrica*, 83(4), 1601–1617.
- ROGERSON, W. P. (1985): “The First-Order Approach to Principal-Agent Problems,” *Econometrica*, 53(6), 1357–67.
- SANNIKOV, Y. (2014): “Moral Hazard and long-run incentives,” Unpublished working paper, Princeton University.
- ZHU, J. (2015): “Myopic Agency,” Unpublished working paper, University of Pennsylvania.