Moral hazard and the quest for linear contracts*

Marcus Opp†

First and Preliminary Draft

Abstract

This paper establishes necessary and sufficient conditions on utility functions and output distributions for linear contracts to arise in the standard static principal-agent framework of Holmström (1979). I present a complete characterization of optimal contracts for a key case: exponentially distributed output and log utility for the agent. When the principal seeks to implement effort below a certain threshold, the optimal contract consists of a fixed wage and an output (equity) share. For effort levels above this threshold, more convex compensation schemes are required, but no optimal contract exists without an additional minimum-wage constraint. When such a constraint is introduced, optimal contracts take the form of a fixed wage and a call option on output. I derive a closed-form solution for infimum compensation cost.

Keywords: Linear Contracts, Contract Theory.

JEL Classification: D86.

1 Introduction

The classical static principal agent model of Holmström (1979) asserts that the likelihood ratio of output realizations is a sufficient statistic for designing optimal compensation contracts. Yet, as successful the informativeness principle was and is for our general understanding of optimally addressing moral hazard, the model has been criticized because its predictions do not readily align with the simple, linear contracts often observed in practice. In response, subsequent research (e.g., Holmström and Milgrom (1987), Innes (1990), Hébert (2017), Yang (2019)) has departed from the classical setup to better capture these empirical regularities.

*This paper benefited from feedback by Philip Bond, Nicolae Gârleanu, Barney Hartman-Glaser, Vincent Maurin, and Moritz Meyer-ter-Vehn.

†Stockholm School of Economics and CEPR, e-mail: marcus.opp@hhs.se

This paper demonstrates that one need not abandon the Holmström (1979) framework to obtain linear contracts as the optimal resolution of the trade-off between incentives and insurance. I establish necessary and sufficient conditions on the agent's utility function and the output distribution for linear contracts to arise. In particular, linear contracts emerge generically when (a) the agent exhibits logarithmic utility and (b) output is drawn from a Gamma distribution (with the exponential distribution as a special case). Under low induced effort, the optimal contract is implemented by paying the agent a base wage along with a fixed share of output. However, for effort levels above a certain threshold, an optimal compensation contract fails to exist in the absence of a minimum wage constraint, a phenomenon analogous to the classical example by Mirrlees (1999) for a log-normal output distribution. When a minimum wage constraint is imposed, the optimal contract becomes piecewise linear, consisting of a base wage and a call option on output with a strike price that decreases with the minimum wage. I also provide a closed-form solution for the limiting compensation contract and the associated second-best compensation cost as the minimum wage converges to zero.

I derive these results within the standard setup of Holmström (1979), which features a risk-neutral principal and a risk-averse agent with additively separable preferences in consumption and effort. The agent's unobservable effort, denoted by a, generates stochastic output x. Under regularity conditions, Holmström (1979) shows that the optimal compensation schedule $c(x)$ satisfies

$$\frac{1}{u'\left(c(x)\right)} = \lambda_{PC} + \lambda_{IC}\, L\left(x|\, a\right), \tag{1}$$

where u is the agent's utility from consumption, $L\left(x|\, a\right)$ is the likelihood ratio associated with output x given effort a, and $\lambda_{PC} > 0$ and $\lambda_{IC} > 0$ are the Lagrange multipliers corresponding to the participation and incentive constraints, respectively. It follows immediately from (1) that affine contracts are optimal if both the inverse marginal utility and the likelihood ratio are affine functions. In this framework, logarithmic utility ensures the former condition, while the Gamma distribution satisfies the latter.

Affine contracts are fully characterized by two parameters: a base wage and an output share. Under such a contract, the agent's expected payoff is a strictly concave function of effort, allowing the first-order approach to hold for any convex effort cost function.[1]. Risk aversion implies that effort incentives depend on the base wage in addition to the output share. With logarithmic utility, a sufficient statistic for effort incentives is the

[1] The first-order approach holds even though the exponential distribution only satisfies the monotone likelihood ratio property, but not convexity, so the validity does not follow from the sufficient conditions established by Rogerson (1985).

ratio of the output share to the base wage. Intuitively, higher effort requires a higher ratio, i.e., a larger variable component relative to fixed pay. Maximal incentives under an affine contract are achieved when the principal relies solely on variable pay. Yet, even this pure linear scheme fails to provide sufficient incentives when the principal desires very high effort.

For sufficiently high-effort, the issue is not implementability but implementability via an affine contract. For any given wage level, the principal can induce high effort levels by making the incentive scheme more convex, i.e., by awarding a bonus only if output exceeds a cutoff, effectively offering a call option on output. This modified contract is optimal if the principal faces an additional minimum wage constraint. In this setting, the punishment region (where the agent receives only the base wage) expands as the minimum wage increases. I provide a closed-form solution for the infimum compensation cost as the minimum wage tends to zero (with the call option strike price converging to zero). The resulting limiting contract is akin to a linear scheme from the risk-neutral principal's perspective, yet the prospect of an infinitely severe (zero-pay) punishment even if realized with zero probability provides strong incentive effects on the risk-averse agent.

2 Model

I consider the standard, static Principal-Agent model a la Holmström (1979). The principal observes output signals X according to the density function $f(x|a)$, which is parameterized by the agent's action $a \in \mathbb{R}^+$. To implement an action a, the principal designs a compensation scheme c as a function of realized output x. The principal is risk-neutral and the agent's utility has an additively separable utility function in consumption c and effort a, $U(c,a) = u(c) - k(a)$ where $u(c)$ is strictly concave and her cost of effort $k(a)$ is convex with $k'(0) < 1$.

Let $\mathbb{E}^a$ denote the expectations operator given action a, then the minimum-cost compensation contract to implement action a solves the following program:

Problem 1

$$W(a) := \min_{c(x)} \mathbb{E}^a [c(x)] \qquad s.t.$$

$$\mathbb{E}^a [u(c(x))] \geq k(a) + \underline{u}. \tag{PC}$$

$$a = \arg\max_{\tilde{a}} \mathbb{E}^{\tilde{a}} [u(c(x))] - k(\tilde{a}). \tag{IC}$$

It is customary in the literature, see Bolton and Dewatripont (2004), to assume that the agent's optimization problem in (IC) is characterized by a first-order condition. Assuming that this first-order approach holds, (IC) can be written as:

$$\mathbb{E}^a \left[L\left(x|\, a\right) u\left(c\left(x\right)\right)\right] = k'\left(a\right), \tag{IC-FOC}$$

where $L\left(x|\, a\right) := \frac{f_a(x|a)}{f(x|a)}$ denotes the continuous-action likelihood ratio. Pointwise optimization, see Holmström (1979), implies that the optimal compensation contract satisfies

$$\frac{1}{u'\left(c\right)} = \lambda_{PC} + \lambda_{IC} L\left(x|\, a\right), \tag{2}$$

where $\lambda_{PC} > 0$ & $\lambda_{IC} > 0$ are the respective Lagrange multipliers on (PC) and (IC-FOC). Economically, (2) implies that the marginal cost of transferring utility to the agent is an affine function of the likelihood ratio. As is well-known, optimality condition (2) typically generates non-linear contracts in terms of realized output.

3 Affine contracts as optimal contracts

Since linearity in output is a prevalent feature of compensation schemes observed in the real world, it is thus of special interest to determine the class of utility and probability distribution functions that generate such contracts.

Lemma 1 *The solution to (2) generates an affine compensation scheme for all values of the agent outside option $\underline{u}$ and cost functions $k\left(a\right)$, if and only if the following two conditions are satisfied.*

1. *The agent has generalized log utility $u\left(c\right) = \ln\left(c - \underline{c}\right)$ for some constant $\underline{c}$, and*

2. *the probability density function, $f\left(x|\, a\right)$, can be factorized as follows:*

$$f\left(x|\, a\right) = \kappa\left(a\right) z\left(x\right) e^{-x\delta(a)}, \tag{3}$$

for positive-valued functions $\kappa\left(a\right)$, $z\left(x\right)$, and $\delta\left(a\right)$ where $\delta'\left(a\right) > 0$ ensures that the compensation scheme is increasing in x.

These two conditions ensure that the inverse marginal utility, $\frac{1}{u'(c)}$ is affine in consumption, and the likelihood ratio is affine in output as $L\left(x|\, a\right) = \frac{\kappa'(a)}{\kappa(a)} + \delta'\left(a\right) x$. As a result, the solution to (2) is indeed affine in x.

Corollary 1 *The Gamma-distribution family satisfies Condition (3) with* $\kappa(a) = \frac{a^{-\eta}}{\Gamma(\eta)}$*,* $z(x) = x^{\eta-1}$ *and* $\delta(a) = \frac{1}{a}$*.*

Within this family, the exponential distribution, which is obtained by setting $\eta = 1$, is the most prominent example in economics. In this special case, the agent action determines mean output. We will now fully characterize the solution to Problem 1.

Proposition 1 (Exponential Distribution) *A solution to Problem 1 exists if and only if* $a \leq \bar{a}$*,where the threshold* $\bar{a} > 0$ *solves* $ak'(a) = 1$ *The optimal compensation contract is affine in* x

$$c(x) = w + w\beta\frac{x}{a}. \tag{4}$$

Given this contract, the first-order approach is valid for any convex cost function $k(a)$*. The principal's expected wage cost is given by* $W = w(1+\beta)$*.*
Let $\mathrm{Ei}(y) = -\int_{-y}^{\infty} e^{-t}dt$ *denote the exponential integral and* $G(y) := 1 + \frac{1}{y}e^{\frac{1}{y}}\,\mathrm{Ei}\left(-\frac{1}{y}\right)$*, then the optimal contract parameters* β^* *and* w^* *to implement action* a *satisfy:*

$$G(\beta^*) = ak'(a), \tag{5}$$

$$\ln w^* = \underline{u} + k(a) - \beta^*(1 - ak'(a)). \tag{6}$$

In sum, for $a \leq \bar{a}$ we obtain a compensation contract that is affine in output x. As a result of scaling by mean output a, the principal's expected compensation costs are $w + w\beta$. The contract parameters β^* and w^* have intuitive interpretations. Here, w refers to the fixed wage and β is *ratio* of the slope $(w\beta)$ to the fixed wage w, see specification in (4). As argued in the introduction, with log utility, this ratio β is a sufficient statistic for incentives provided by a linear contracts, see incentive constraint (5). The right hand side of (5), $ak'(a)$, measures the strength of required incentives and is, hence, intuitively increasing in the implemented action a (formally due to convexity of k). The required incentives need to match the provided incentives by the contract which are captured by the strictly increasing function $G(\beta^*)$ that maps β^* into $[0, 1]$, see plot of G in the Proof of Proposition 1. Since $G(\beta)$ is bounded above by one, as $\lim_{\beta\to\infty} G(\beta) = 1$, only actions with $ak'(a) \leq 1$ can be implemented with an affine compensation scheme. The base wage, see (6), is then just set sufficiently high to ensure participation of the agent given her outside option $\underline{u}$ and cost of effort $k(a)$, where $\beta^*(1 - ak'(a)) > 0$ measures the agent's valuation of the variable pay component (in utils).

The contract parameters have the following intuitive properties, see plot in Figure 1. In this example, we set $k(a) = \frac{c^2}{2}a^2$ so that $\bar{a} = \frac{1}{c}$.[2]

[2] By making c arbitrarily small, one can make $\bar{a}$ arbitrarily large.

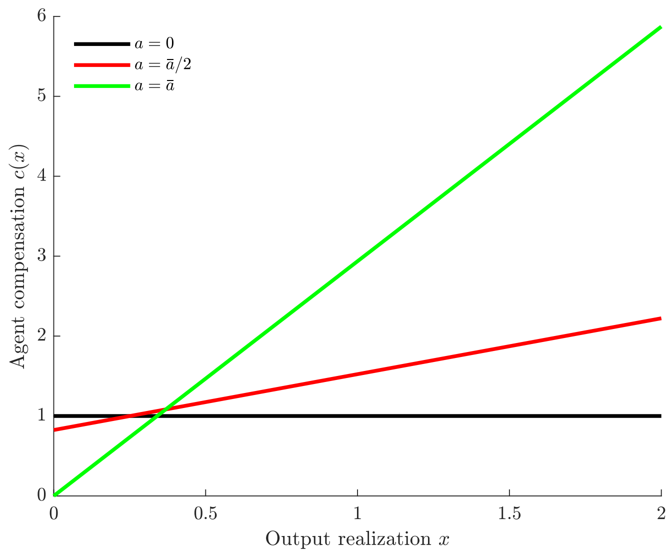

Figure 1. The graph plots the compensation scheme $c(x)$ for 3 values of a between 0 and $\bar{a}$. The agent outside option is $\underline{u} = 0$ and the cost of effort satisfies $k(a) = 0.5a^2$ so that $\bar{a} = 1$.

Proposition 2 (Comparative Statics) *For $a \leq \bar{a}$, the optimal percentage bonus on the wage β^* is strictly increasing in the implemented action a while the base wage w^* is strictly decreasing in a. The compensation contract for the lower and upper bound is obtained in closed form.*

- *For $a = 0$, the optimal contract consists only of a flat wage $c(x) = w^* = e^{\underline{u}}$.*
- *For $a = \bar{a}$, the optimal contract is a pure linear contract without a fixed wage*

$$c(x) = e^{\underline{u}+k(\bar{a})+\gamma}\frac{x}{\bar{a}}, \tag{7}$$

where $\gamma \approx 0.58$ denotes the Euler-Mascheroni constant.

These comparative statics are intuitive. When no incentives need to be provided, $a = 0$, the optimal second-best contract corresponds to the first-best contract consisting only of a flat wage. The higher the action that the principal wants to implement, the higher the required incentives, which translates into higher β. However, even in the limit as β approaches infinity, incentives provided from a linear contract are finite as $\lim_{\beta\to\infty} G(\beta) = 1$. At $\bar{a}$, the optimal compensation contract is obtained in closed-form. The resulting second-best wage cost $W(\bar{a}) = W^{FB}(\bar{a})\, e^{\gamma}$ are 58% higher than first-best compensation compensation costs of $W^{FB}(\bar{a}) = e^{\underline{u}+k(\bar{a})}$.

Proposition 1 states that no solution to Problem 1 exists for $a > \bar{a}$ as no affine compensation contract can generate sufficient incentives, However, it does not rule out implementability with more convex compensation schemes.

Proposition 3 (Implementability with Call Options) *Actions $a > \bar{a}$ are implementable by choosing a fixed wage $\underline{w} \in (0, \exp[\underline{u} + k(a)])$ and awarding $\frac{\underline{w}\beta}{a}$ call options on output with a strike price equal to the p quantile of the exponential distribution $Q(p,a) := -a\ln[1-p]$:*

$$c(x) = \underline{w} + \frac{\underline{w}\beta}{a}\max(x - Q(p,a), 0) \tag{8}$$

For any given $\underline{w}$, the contract parameters $p^* \in (0,1)$ and $\beta^* > 0$ satisfy

$$(1-p)(G(\beta) - (1 - G(\beta))\beta\ln(1-p)) = ak'(a) \tag{9}$$

$$\ln(\underline{w}) + (1-p)(1 - G(\beta))\beta = \underline{u} + k(a) \tag{10}$$

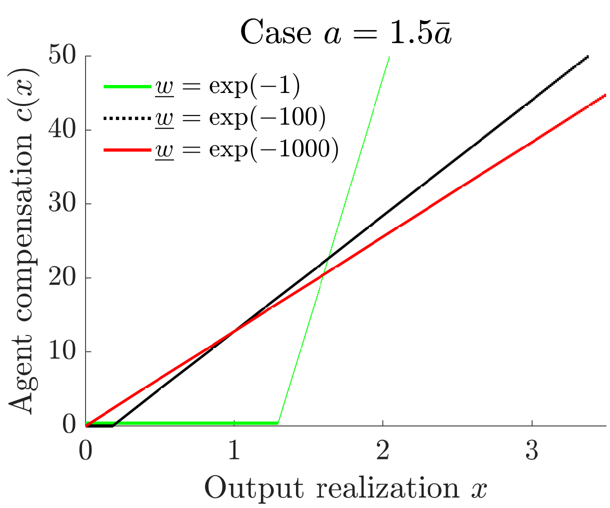

Figure 2. The graph plots optimal compensation contract for $a = 1.5\bar{a}$ for 3 different values of the lower bound on pay. The optimal contract is a call option.

Intuitively, without restricting attention to affine incentives schemes, one way to satisfy (IC-FOC), $\mathbb{E}^a[L(x|a)u(c(x))] = k'(a)$, is to simply punish the agent sufficiently for negative likelihood ratios. Recall, due to log utility, paying the agent close to the subsistence level of zero, allows the principal to punish the agent sufficiently. For the compensation contract described in Proposition 3, the "punishment" corresponds to only paying a base wage $\underline{w}$ as long as output falls below the p quantile $Q(p,a)$. While this contract described in just one of many possible contracts that implements actions $a > \bar{a}$, it is of particular interest, as the following Lemma clarifies.

Lemma 2 (Minimum-wage constraint) *For any $a > \bar{a}$, the contract in Proposition 3 is the solution to Problem 1 subject to the additional constraint of a minimum wage of $\underline{w}$, i.e., $c(x) \geq \underline{w}$. The punishment probability p (and, hence, the strike price) is increasing in $\underline{w}$ and the number of options is strictly increasing in $\underline{w}$.*

Lemma 2 offers two interpretations. First, from an applied perspective, constraints on minimum compensation are not unrealistic, and, hence it offers a rationale for the wide-spread use of call options in real life. Intuitively, a higher minimum wage limits the ability to "punish," so the punishment region must optimally increase as must the number of options. Second, it allows us to shed light on the failure of existence in the original problem (without a minimum wage) by considering the limit as $\underline{w}$ becomes arbitrarily small.

Lemma 3 (An almost linear contract!) *As the minimum wage approaches zero, the option strike price approaches zero, i.e., $\lim_{\underline{w}\to 0} p = 0$ and*

$$\lim_{\underline{w}\to 0} \beta^* \underline{w} = \exp\left(\underline{u} + k(a) + \gamma + ak'(a) - 1\right). \tag{11}$$

The infimum wage costs for any action $a > \bar{a}$ are given by:

$$W(a) = e^{\underline{u}+k(a)+\gamma+ak'(a)-1} \tag{12}$$

In the limit as $\underline{w} \to 0$ the contract is akin to a linear contract from the risk-neutral principal's perspective $\frac{x}{a}e^{\underline{u}+k(a)+\gamma+ak'(a)-1}$. Yet, if the principal were to write an exact linear contract, the agent would always choose $\bar{a}$ (and PC would be slack). The "missing" incentives of $ak'(a) - 1$ are created by the prospect of an infinitely severe (zero-pay) punishment realized with zero probability just so that IC is satisfied, i.e., in the limit we obtain:

$$p\left(\left(1 - G(\beta)\right)\beta\right) = ak'(a) - 1. \tag{13}$$

References

Bolton, P. and M. Dewatripont, *Contract Theory*, MIT Press, 2004.

Holmström, Bengt, "Moral Hazard and Observability," *Bell Journal of Economics*, 1979, *10* (1), 74–91.

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

Hébert, Benjamin, "Moral Hazard and the Optimality of Debt," *The Review of Economic Studies*, 12 2017, *85* (4), 2214–2252.

Innes, Robert D, "Limited liability and incentive contracting with ex-ante action choices," *Journal of Economic Theory*, 1990, *52* (1), 45–67.

Mirrlees, J. A., "The Theory of Moral Hazard and Unobservable Behaviour: Part I," *The Review of Economic Studies*, 01 1999, *66* (1), 3–21.

Rogerson, William P., "The First-Order Approach to Principal-Agent Problems," *Econometrica*, 1985, *53* (6), 1357–1367.

Yang, Ming, "Optimality of Debt under Flexible Information Acquisition," *The Review of Economic Studies*, 07 2019, *87* (1), 487–536.

A Proofs

Proof of Lemma 1: First, we prove that an affine compensation scheme emerges from (2) if and only if $\frac{1}{u'(c)}$ is affine in c and $L(x|a)$ is affine in x. This condition is clearly sufficient. It is also necessary since $\underline{u}$ and $k(a)$ change λ_{PC} and λ_{IC}. In other cases, it is knife-edge in the sense that the solution turns out to be affine for particular values of λ_{PC} and λ_{IC}.[3] ■

Proof of Proposition 1: (4) follows from (2) due to linear inverse marginal utility $\frac{1}{u'(c)} = c$ and affine likelihood ratio $L(x|a) = \frac{x-a}{a^2}$.

We next show that the first-order approach is valid for linear contacts. To do so, it is sufficient to show that the compensation value of a linear contract from the agent's perspective, $V(\tilde{a}) := \mathbb{E}^{\tilde{a}}\left[\ln\left(w + w\beta\frac{x}{a}\right)\right]$, given a desired action a is strictly concave in $\tilde{a}$. Integration and using the definition of $\mathrm{Ei}(y)$ yields

$$V(\tilde{a}) = \ln(w) - e^{\frac{a}{\tilde{a}}\frac{1}{\beta}}\,\mathrm{Ei}\left(-\frac{a}{\tilde{a}}\frac{1}{\beta}\right). \tag{A.1}$$

Let $z(\tilde{a}) := \tilde{a}\frac{\beta}{a}$, then the second derivative of V is negative for all $\tilde{a}$

$$V''(\tilde{a}) = -\frac{\left(\frac{1}{z(\tilde{a})} + 2\right) G(z(\tilde{a})) - 1}{\tilde{a}^2} < 0, \tag{A.2}$$

where the numerator, $\left(\frac{1}{z(\tilde{a})} + 2\right) G(z(\tilde{a})) - 1$, is between 0 and 1 and the function $G(y) := 1 + \frac{1}{y}e^{\frac{1}{y}}\,\mathrm{Ei}\left(-\frac{1}{y}\right)$ is strictly increasing in y and maps $y \in \mathbb{R}^+$ into $[0, 1]$, see plot in Figure 3.

[3] For other cases, it would be knife edge, e.g., for CRRA utility with relative risk aversion parameters γ, we would obtain that $c = (\lambda_{PC} + \lambda_{IC}L(x|a))^{\frac{1}{\gamma}}$. For $\gamma \neq 1$, this would require that $L(x|a) = \frac{(\eta+\phi x)^{\gamma} - \lambda_{PC}}{\lambda_{IC}}$, i.e., $(\lambda_{PC} + \lambda_{IC}L(x|a))^{\frac{1}{\gamma}} = \eta + \phi x$

Since $G(\beta) := 1 + \frac{1}{\beta} e^{\frac{1}{\beta}} \operatorname{Ei}\left(-\frac{1}{\beta}\right)$ is bounded above by one, see Figure 3, so a solution to (5) only exists if the right hand side of (5), $ak'(a) < 1$, is smaller or equal to 1.

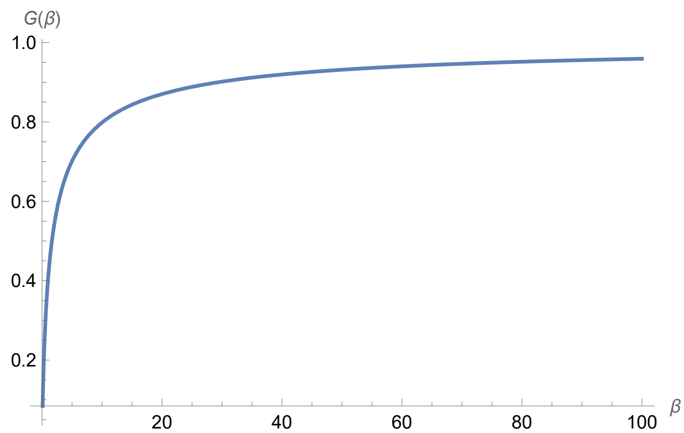

Figure 3. This graph plots the function $G(\beta) = 1 + \frac{1}{\beta} e^{\frac{1}{\beta}} \operatorname{Ei}\left(-\frac{1}{\beta}\right)$.

Due to strict concavity of V, the agent's overall objective $V(\tilde{a}) - k(\tilde{a})$, is strictly concave for any convex cost function k. This means that (IC-FOC) is both necessary and sufficient for agent optimality.

Substituting the optimal incentive scheme (4) into (IC-FOC) and integrating yields (5). Effort incentives are solely determined by β (and, hence, independent of the wage w). Since $G(\beta) \in [0, 1]$, a (unique) solution for β exists if and only if the right hand side of (5) is less than 1, i.e., $ak'(a) < 1$. Given that $ak'(a)$ is strictly increasing in a (by convexity of k) this condition is equivalent to $a < \bar{a}$. Since $k'(0) = 0$, any action $a \in [0, \bar{a})$ is implementable with a linear contract.

Given the solution for β^* from (5), we obtain the optimum base wage w^* from binding (PC), i.e., the agent's compensation value given the optimal action a, $V(a) = \ln(w) - e^{\frac{1}{\beta^*}} \operatorname{Ei}\left(-\frac{1}{\beta^*}\right)$, see (A.1), matches the agent's outside option $\underline{u} + k(a)$, so that

$$\ln(w^*) = \underline{u} + k(a) + e^{\frac{1}{\beta^*}} \operatorname{Ei}\left(-\frac{1}{\beta^*}\right). \tag{A.3}$$

Using the optimality condition (5) for β^*, we obtain (6). ■

Proof of Proposition 2: The first statement follows directly from the properties of $G(\beta)$, see Proof of Proposition 1. For $a = 0$, the compensation contract is simply the (flat) first-best compensation contract. Next consider the limit as a approaches $\bar{a}$, in which case $\lim_{a \to \bar{a}} \beta^*(a) = \infty$. Since $\lim_{\beta \to \infty} e^{\frac{1}{\beta}} \operatorname{Ei}\left(-\frac{1}{\beta}\right) = -\infty$, (A.3) implies that the log wage does so as well. The limiting compensation contract can, thus, be written expressed as

$$\lim_{a \to \bar{a}} b(x) = \lim_{a \to \bar{a}} w + w\beta \frac{x}{a} = \frac{x}{\bar{a}} \lim_{a \to \bar{a}} w\beta.$$

To determine $\lim_{a\to\bar{a}} \beta w$, it is useful to consider the limit of its logarithm:

$$\begin{aligned}\lim_{a\to\bar{a}} \ln \beta w &= \lim_{a\to\bar{a}} \ln \beta + \ln w^* \\ &= \underline{u} + k(\bar{a}) + \lim_{\beta\to\infty} \ln \beta + e^{\frac{1}{\beta^*}} \operatorname{Ei}\left(-\frac{1}{\beta^*}\right) \\ &= \underline{u} + k(\bar{a}) + \gamma,\end{aligned}$$

where the second line follows from (A.3) as well as $\lim_{a\to\bar{a}} \beta^*(a) = \infty$ and the third line follows from the fact that $\lim_{\beta\to\infty} \ln \beta + e^{\frac{1}{\beta^*}} \operatorname{Ei}\left(-\frac{1}{\beta^*}\right) = \gamma$, the Euler-Mascheroni constant. ■

Proof of Proposition 3: Given this contract, the payment to the agent is given by:

$$c(x) = \begin{cases} \underline{w} & \text{if } x \leq -a\ln[1-p] \\ \underline{w} + \frac{w\gamma}{a}(a\ln[1-p] + x) & \text{else} \end{cases}.$$

Since $\mathbb{P}^a[x \leq -a\ln[1-p]] = p$, the agent, thus, get $\underline{w}$ with probability p and

$$\mathbb{E}^a[u(b(x))] = u(\underline{c}) + (1-p)(1-G(\beta))\beta \tag{A.4}$$

$$\mathbb{E}^a[L(x|a)u(b(x))]a = (1-p)(G(\beta) - (1-G(\beta))\beta\ln(1-p)) \tag{A.5}$$

For any interior $p \in (0,1)$, you can set β sufficiently high so that IF-FOC is satisfied because $\lim_{\beta\to\infty}(1-G(\beta))\beta = \infty$. Binding PC requires that

$$1 - p = \frac{\underline{u} + k(a) - u(\underline{c})}{(1-G(\beta))\beta}$$

which is only possible if $\underline{u} + k(a) > u(\underline{c})$ (i.e., punishment region involves pay below reservation wage). This gives us:

$$\frac{G(\beta)}{(1-G(\beta))\beta} + \ln((1-G(\beta))\beta) = \frac{ak'(a)}{\underline{u} + k(a) - u(\underline{c})} + \ln(\underline{u} + k(a) - u(\underline{c}))$$

The function on the left hand side can take on any value in $\mathbb{R} \Longrightarrow \beta^*$. ■