Optimal Contracting with Effort and Misreporting

November 2011

Agostino Capponi
Purdue University

Jakša Cvitanić
EDHEC Business School

Türkay Yolcu
Purdue University
Abstract
We propose a new continuous time contracting model, where the project value process can only be observed with noise, and there are two sources of moral hazard: effort and misreporting. Using calculus of variation techniques, we are able to find the optimal pay-per-performance sensitivity (PPS) of the contract offered to the manager, as well as optimal effort and misreporting action via a second order ordinary differential equation with time dependent coefficients. Our findings indicate that the agent will apply a higher level of effort and misreporting than if only one of those actions was present. We numerically illustrate the dependence of the agent’s actions and pay-per-performance sensitivity on the model parameters through a detailed comparative statics analysis. We find that PPS is not necessarily monotone with respect to the variance of observational noise and with respect to time. We also find that the shareholders may care more about motivating the manager to misreport than to apply effort.

Keywords: Optimal contracts, Principal-Agent problem, Variational calculus, Stochastic filtering, Moral hazard.

The research of J. Cvitanić was supported in part by NSF grants DMS 10-08219.

EDHEC is one of the top five business schools in France. Its reputation is built on the high quality of its faculty and the privileged relationship with professionals that the school has cultivated since its establishment in 1906. EDHEC Business School has decided to draw on its extensive knowledge of the professional environment and has therefore focused its research on themes that satisfy the needs of professionals.

EDHEC pursues an active research policy in the field of finance. EDHEC-Risk Institute carries out numerous research programmes in the areas of asset allocation and risk management in both the traditional and alternative investment universes.
1. Introduction

We propose a new continuous time contracting model for modeling the principal-agent relationship in which there are two possible actions that the agent can take, namely effort and misreporting. Similarly to Holmstrom and Milgrom (1987), the agent can control the drift of the project value process through effort. Differently from that paper, and similarly to Capponi, Cvitanić and Yolcu (2011), we allow the agent to misreport the accounting reports on the value of the firm. The misreporting action is not contractible, either because it is unobserved by the principal, and/or because it is not legally enforceable. Our framework extends Holmstrom and Milgrom (1987), and also generalises the contracting model proposed in Capponi, Cvitanić and Yolcu (2011), in which the output process cannot be controlled through effort by the agent, whose only available action is to misreport the accounting reports. Indeed, as we show in Section 3.4, both of these contracting models may be recovered as limiting cases of our framework. Therefore, the proposed contracting model allows analysing the interplay between effort and misreporting in a moral hazard setting.

The seminal work on continuous time contracting models by Holmstrom and Milgrom (1987) was extended along several directions, see Schättler and Sung (1993, 1997), Sung (1995, 1997), and Müller (1998). Cvitanic, Wang and Zhang (2009) use the stochastic maximum principle and Forward-Backward Stochastic Differential Equations (FBSDEs) to characterise the optimal compensation for general utility functions under moral hazard. Sannikov (2008) finds a tractable model for solving the problem with a random time of retiring the agent and with continuous payments to the agent. This model is extended by Prat and Jovanovic (2010), who consider the case where the drift is not observed. Giat et al. (2010) consider a model where the project value is observed, but its risk premium (drift term) is not, and principal and agent have different prior beliefs about it. Ou-Yang (2005) uses the principal agent model to show that the equilibrium price of the stock depends on managerial incentives. Cuoco and Kaniel (2011) study asset pricing implications of contracts in which the compensation of managers depends on the excess return of the managed portfolio over a benchmark portfolio. Goukasian and Wan (2010) consider a moral hazard framework with multiple agents who only care about their relative position and study the impact of this assumption on the optimal contract. Another recent branch of literature has considered contracting models with hidden output and misreporting. For example, De Marzo and Sannikov (2006) consider an agent who observes the actual cash flows and can divert a portion of them for personal consumption. The contract is contingent on cash flows remaining after diversion. Their model has no misreporting in equilibrium, which is not the case in this paper. In presence of asymmetric information, but assuming a model consisting of manager and shareholders who are unaware of misreporting, Benmelech, Kandel and Veronesi (2010) show that stock based compensation can induce managers to conceal bad news, besides exerting high effort.

Our contracting framework involves three entities: agent (manager), principal (shareholders) and the market (outside investors, excluding shareholders). The shareholders compensate the manager according to a contract that pays at the end of the horizon, and in their maximisation they both use the value of the firm decided by the market, which we call market value estimates. Our setting assumes that the manager and the shareholders do not observe the outcome process directly. Following evidence provided in the empirical literature that accounting reports are typically contaminated by accounting noise (see also the theoretical models of Duffie and Lando (2001) and Capponi and Cvitanić (2008)), we assume that the manager and the shareholders can only observe a white noise contaminated version of the actual output. Moreover, while managerial effort cannot be directly contracted upon by the principal, she can either observe it or compute it correctly in equilibrium. The market can also observe the realisations of the effort process over time, but it is unaware or unable to correctly estimate the reporting bias introduced by the agent. The market produces the market value estimates only on the basis of: (1) the realised effort; (2) the assumption that the uncertainty on the actual output process is governed by a Brownian motion; and (3) the knowledge that there is accounting noise in the reported output.
We assume that the utility functions of the principal and of the agent are exponential. We impose a quadratic penalty on the agent, for both the misreporting and the effort action. We consider the class of linear contracts, and assume deterministic effort and misreporting actions for tractability reasons. More precisely, a contract is a linear functional of the market value estimates, computed using a stochastic filtering approach, and under the assumption of no misreporting taking place. The latter assumption is justified if the contracts cannot legally refer to misreporting, perhaps because there will be no auditing of accounting reports.

We formulate the optimal contracting problem as a free boundary calculus of variations problem. Through a rigorous analysis, we show that the optimal pay-per-performance sensitivity, misreporting action, and effort level can be found by solving a singular second order differential equation. Moreover, unlike Capponi, Cvitanić and Yolcu (2011), the variational techniques developed in this paper can be viable to analyse moral hazard problems where multiple non-contractible actions are taken. Using numerical techniques, we analyse the dependence of those quantities on the various model parameters, including the mean-variance trade-off, the volatility of the accounting noise, the size of the penalties, and the levels of risk aversion. We find that there is non-zero misreporting in equilibrium. This occurs because of two reasons: (1) misreporting cannot be contracted upon (that is, there is no possibility of direct punishment for misreporting, at least not with 100% probability); and (2) the market values the project without being able to correctly predict the misreporting level. The level of effort applied by the agent in the proposed contracting framework is always higher than the one he would apply in the Holmstrom and Milgrom (1987) model, in which there is no misreporting. Similarly, the agent optimally applies an amount of misreporting higher than the one applied in the contracting framework proposed in Capponi, Cvitanić and Yolcu (2011), where no effort is present. We also find that the agent is compensated by the principal with a higher pay-per-performance sensitivity (PPS) relative to the one received when only one of the two actions is present. Moreover, we find that PPS exhibits an U-shaped pattern over time. Initially, it is maximal and used to induce the agent to apply full effort and zero misreporting. As time passes, the principal induces the agent to apply both actions, but providing him with PPS which is decreasing first, and then increasing again till the time of compensation. This contrasts with Capponi, Cvitanić and Yolcu (2011), where PPS always increases over time, and the agent applies an increasing level of misreporting as time progresses. Also, differently from that paper, we find that the dependence of PPS on the noise intensity is not always monotonically decreasing, but dependent on the relation between effort and misreporting penalty. This is because while the benefit of the misreporting action decreases with the level of accounting noise, the effort action can still increase the market value estimate of the output value process. This induces the principal to give the agent more incentives to apply effort relative to misreporting, especially if the former has smaller penalty than the latter. Interestingly, we find that a large part of PPS is due to providing incentives for misreporting. This implies that transferring more power to shareholders, as an improvement to corporate oversight, might be counterproductive. Instead, more frequent auditing might be a better way of improving oversight.

The rest of the paper is organised as follows. Section 2 describes the model. Section 3 introduces and analyses the optimal contracting problem in presence of non-contractible effort and misreporting. Section 4 performs a numerical study to analyse the dependence of the optimal pay-per-performance sensitivity, effort, and misreporting on the parameters of our model. Section 5 concludes the paper.

---

1 - In the classical Holmstrom-Milgrom model, the optimal effort is, in fact, deterministic and the optimal contract is linear. It is possible that also in our model these assumptions are without loss of generality, but we have not been able to prove it.
2. The Filtering Model

Our model consists of the state process and the observations process. Let \((W_t, Z_t^0)\) be a standard two-dimensional Brownian motion on a probability space \((\Omega, \mathcal{F}, \mathbb{P})\), where \((\mathcal{F}_t)_{0 \leq t \leq T}\) is the filtration generated by it, up to time \(T > 0\). Let

\[ x_t = \epsilon t + \epsilon W_t \]

We assume the “mean variance trade-off” of the output process \(\epsilon\) to be constant, and refer to Cvitanić and Zhang (2007) for the case when \(\epsilon\) can be controlled. For simplicity, we assume that \(x_0\) is observed. The state process, or output process, \(x_t\) is not directly observed by the market. Rather the agent adds bias to it, i.e., he misreports it. In addition the output is not observed directly, for example due to the noise in accounting data. Specifically, define the process \(y\) given by

\[ dy_t = x_t dt + \sigma dZ_t^0 \]  
(2.1)

We denote by \(\mathcal{F}_t^y\) the filtration generated by the observation process \(\{y_t\}\). Here, \(\sigma > 0\) is the intensity of the accounting noise.

For given deterministic bounded functions \(u_t\) and \(a_t\), let us define

\[ W_t^u \triangleq W_t - \int_0^t u_s ds, \quad Z_t^a \triangleq Z_t^0 - \int_0^t \frac{a_s}{\sigma} ds \]

\[ M_t^{u,a} \triangleq \exp \left\{ \int_0^t u_s dW_s - \frac{1}{2} \int_0^t u_s^2 ds \right\} \exp \left\{ \int_0^t \frac{a_s}{\sigma} dZ_s^0 - \frac{1}{2} \int_0^t \left( \frac{a_s}{\sigma} \right)^2 ds \right\} ; \]

\[ \frac{d\mathbb{P}^{u,a}}{d\mathbb{P}} = M_T^{u,a} \]

Then, by Girsanov theorem, \(M_t^{u,a}\) is a martingale and \(\mathbb{P}^{u,a}\) is a probability measure under which \((W_t^u, Z^a)\) is a \(\mathbb{P}^{u,a}\) Brownian motion. Moreover, we have the dynamics

\[ dx_t = u_t \epsilon dt + \epsilon dW_t^u, \]  
(2.2)

\[ dy_t = (x_t + a_t) dt + \sigma dZ_t^a \]  
(2.3)

In other words, the agent improves the return of the output by his effort \(u_t\) and adds misreporting amount \(a_t\) to the drift of the observed process \(y\), in Eq. (2.1), thus introducing bias. The change of measure approach, that is, the assumption that the agent affects the distribution of the processes \(x\) and \(y\) by her effort \(u\) and misreporting \(a\), is the standard continuous-time formulation for the principal-agent problems.

Let us introduce the notation \(\mathbb{E}^{u,a}[X] \triangleq \mathbb{E}^{P^{u,a}}[X]\), where \(X\) is a given random variable. Then, introduce the filter process

\[ \hat{x}_t^{u,a} \triangleq \mathbb{E}^{u,a}[x_t | \mathcal{F}_t^y] \]

Remark 2.1 We have that \(\hat{x}_t^{u,a}\) is the estimate of the output value at time \(t\), assuming exact knowledge of the effort process \(u_s\) and of the misreporting action as, \(0 \leq s \leq t\). As stated in the introduction, we assume that the market observes \(u_t\) but does not know misreporting process \(a_t\). More to the point, we assume that the market value estimate is \(\hat{x}_t^{u,0}\).

We next state a useful lemma, which will be later used to compute the optimal contract. The proof is a straightforward extension of the one given in Capponi, Cvitanić and Yolcu (2011) for the case when \(u = 0\), and omitted here.
Using the notation

\[ v_t = \epsilon \sigma \tanh \left( \frac{ct}{\sigma} \right) \]  

we have

**Lemma 2.1.**

\[ \hat{Z}_t^0 = \int_0^t \frac{\sigma}{v_s} d\hat{x}_s^{u,0} - \int_0^t \frac{\epsilon \sigma u_s}{v_s} ds \]  

and so \( \hat{Z}_t^a \) is expressed in terms of \( \hat{x}_t^{u,0} \) as

\[ \hat{Z}_t^a = \hat{Z}_t^0 - \int_0^t \frac{a_s}{\sigma} ds + \int_0^t \frac{\hat{x}_s^{u,0} - \hat{x}_s^{u,a}}{\sigma} ds \]  

Moreover, it is shown in Capponi, Cvitanic and Yolcu (2011) that

\[ \hat{x}_t^{u,0} - \hat{x}_t^{u,a} = \frac{\epsilon}{\sigma} \text{sech} \left( \frac{ct}{\sigma} \right) \int_0^t \sinh \left( \frac{es}{\sigma} \right) a_s ds \]  

and so Eq. (2.6) may be explicitly written as

\[ \hat{Z}_t^a = \int_0^t \frac{\sigma}{v_s} d\hat{x}_s^{u,0} - \int_0^t \frac{\epsilon \sigma u_s}{v_s} ds - \int_0^t \frac{a_s}{\sigma} ds + \frac{\epsilon}{\sigma} \int_0^t \left[ \text{sech} \left( \frac{es}{\sigma} \right) \left( \int_0^s \sinh \left( \frac{er}{\sigma} \right) a_r dr \right) \right] ds \]  

Using (2.5) and (2.6), we can infer that

\[ d\hat{x}_t^{u,0} = \epsilon u_t dt + \frac{v_t}{\sigma} d\hat{Z}_t^a + \frac{v_t a_t}{\sigma^2} dt - \frac{v_t}{\sigma^2} (\hat{x}_t^{u,0} - \hat{x}_t^{u,a}) dt \]

\[ = \epsilon u_t dt + \frac{v_t}{\sigma} d\hat{Z}_t^a + \frac{v_t a_t}{\sigma^2} dt - \frac{v_t}{\sigma^2} \left( \frac{\epsilon}{\sigma} \text{sech} \left( \frac{ct}{\sigma} \right) \int_0^t \sinh \left( \frac{es}{\sigma} \right) a_s ds \right) dt \]  

### 3. The Contracting Model

We consider a contracting framework in which there are two possible sources of moral hazard, one from effort, and the other from output reporting. More specifically, we consider a principal-agent model in which the principal can induce the agent to increase the market value estimate of the firm in two ways: (1) by misreporting the outcome process, (2) by applying more effort. Our framework is an extension of Holmstrom and Milgrom (1987), in which only effort is present, and of Capponi, Cvitanic and Yolcu (2011), in which only misreporting is present. The results of both of those papers can be recovered as limiting cases, as shown in Section 3.4. As mentioned in the introduction, we consider only linear contracts of the form

\[ \mathcal{G}_T = c + \int_0^T \alpha_t d\hat{x}_t^{u,0} \]

with \( \alpha \) being deterministic. We note that in Holmstrom and Milgrom (1987) the optimal contract is linear in the true final output value. In our case the latter has to be estimated using the whole history of observations. It is then not surprising that the optimal linear contract will be a functional of the whole observation history, and not just of the terminal value. The contract payment \( \mathcal{G}_T \) is an \( \mathcal{F}_T \)-measurable random variable, and should be interpreted as a payment in cash, the amount of which depends on the random outcomes of \( \hat{x}_t^{u,0} \) up to time \( T \). A justification for using \( \hat{x}_t^{u,0} \) and not \( \hat{x}_t^{u,a} \) is that for legal reasons the payments have to be offered under the assumption that there is no misreporting. In what follows, we impose the following restriction on the admissible actions.
Assumption 3.1. The set of admissible effort action misreporting action at and pay-per-performance sensitivity $\alpha_t$ is $C^1[0, T]$, the set of deterministic continuously differentiable functions on the interval $[0, T]$.

For latter use, denote

$$m_T \overset{\Delta}{=} \max_{t \in [0, T]} \{|a_t|, |u_t|, |\alpha_t|\} < +\infty \tag{3.1}$$

We will use $\alpha_t^{ME}$, $u_t^{ME}$, and $a_t^{ME}$ to denote, respectively, the optimal pay-per-performance sensitivity, effort, and misreporting action in this double moral hazard framework, where $ME$ indicates the presence of both misreporting and effort. We assume that the agent incurs a quadratic penalty for both misreporting and effort action. More specifically, we denote by $ET$ the penalty incurred for his effort action in interval $[0, T]$, given by

$$E_T = \frac{k}{2} \int_0^T u_t^2 dt$$

where $k$ is a positive constant. We denote by $G_T$ the penalty incurred for his misreporting action in interval $[0, T]$, given by

$$G_T = \frac{\xi}{2} \int_0^T a_t^2 dt.$$ 

We consider exponential utilities, assuming that

$$U_1(x) = -e^{-\gamma_1 x} \quad U_2(x) = -e^{-\gamma_2 x} \tag{3.2}$$

3.1 Moral Hazard with Effort and Misreporting

The agent minimises over $u_t$ and $a_t$ the following quantity

$$\mathbb{E}^u_\alpha[u_1(E_T - G_T - E_T)] =$$

$$\mathbb{E}^u_\alpha \left[ -\gamma_1 \left( c + \int_0^T \alpha_t u_t^2 dt - \frac{1}{2} \int_0^T (\xi a_t^2 + k u_t^2) dt \right) \right] =$$

$$-\gamma_1 \left( c + \int_0^T \alpha_t \left( u_t - \frac{1}{\alpha_t} \frac{v_t}{\sigma^2} \right) dt - \frac{\xi}{2} \int_0^T \left( \frac{e}{\sigma} \operatorname{sech} \left( \frac{e s}{\sigma} \right) \right) a_t^2 ds \right) dt \right]$$

$$\exp \left\{ \frac{\gamma_1}{2} \int_0^T (\xi a_t^2 + k u_t^2) dt \right\} \tag{3.3}$$

where the last equality follows by using (2.9) and the fact that, for a given deterministic function $\psi(t)$, we have

$$\mathbb{E}^u_\alpha \left[ \exp \left( \int_0^T \psi(t) d\hat{\mathcal{X}}_t \right) \right] = \exp \left( \frac{1}{2} \int_0^T \psi(t)^2 dt \right) \tag{3.4}$$

Equivalently, we want to maximise over $u_t$ and $a_t$ the functional given by

$$\int_0^T \mathcal{L}(t, u_t, a_t) dt \tag{3.5}$$

where

$$\mathcal{L}(t, u_t, a_t) = \alpha_t \left( u_t e - \frac{\gamma_1}{2} \frac{v_t^2}{\sigma^2} + \frac{v_t a_t}{\sigma^2} - \frac{e}{\sigma} \operatorname{sech} \left( \frac{e s}{\sigma} \right) \int_0^t \sinh \left( \frac{e s}{\sigma} \right) a_s ds \right)$$

$$- \left( \frac{k}{2} u_t^2 + \frac{\xi}{2} a_t^2 \right) \tag{3.6}$$

and $v_t$ is defined by Eq. (2.4).
3.2 Optimal Effort and Misreporting given Contract

We introduce the following function

\[ p_t = \int_0^t \sinh \left( \frac{e^s}{\sigma} \right) a_s \, ds \]  
(3.7)

which implies

\[ p_0 = 0 \quad \text{and} \quad p'_t = \sinh \left( \frac{e^t}{\sigma} \right) a_t \]  
(3.8)

Observe that \( p_t \) and \( p'_t \) are both bounded by a multiple of \( m_T \):

\[ \max_{t \in [0,T]} \{|p_t|\} \leq \max_{t \in [0,T]} \{|a_t|\} \left( \frac{\sigma}{e} \cosh \left( \frac{e^T}{\sigma} \right) - \frac{\sigma}{e} \right) < m_T \left( \frac{\sigma}{e} \cosh \left( \frac{e^T}{\sigma} \right) \right) < +\infty \]  
(3.9)

\[ \max_{t \in [0,T]} \{|p'_t|\} \leq \max_{t \in [0,T]} \max_{t \in [0,T]} \left\{ \sinh \left( \frac{e^t}{\sigma} \right) \right\} < (m_T + 1) \sinh \left( \frac{e^T}{\sigma} \right) < +\infty \]  
(3.10)

In this paper, the Lagrangians belong to a special class, called Tonelli Lagrangians, and defined as follows (see Clarke (1989))

**Definition 3.1. (Tonelli Lagrangian)** For \( d \geq 1 \), \( L : [0, T] \times \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R} \) is said to be a Tonelli Lagrangian if

1. \( L \) is a twice continuously differentiable function.
2. The Hessian matrix \( \nabla v L(t, x, v) \) of \( L \) with respect to \( v \) is positive semidefinite for all \( (t, x, v) \in [0, T] \times \mathbb{R}^d \times \mathbb{R}^d \). In other words, the map \( v \to L(t, x, v) \) is convex.
3. \( L \) is coercive. In other words, there exist constants \( m_1 > 0 \) and \( m_2 \in \mathbb{R} \) such that \( L(t, x, v) \geq m_1|v|^2 + m_2 \) for all \( (t, x, v) \in [0, T] \times \mathbb{R}^d \times \mathbb{R}^d \).

The following theorem summarises existence, uniqueness and regularity of the minimiser \( x^*_t \) of \( \int_0^T L(t, x_t, x'_t) \, dt \) that is rephrased from Tonelli’s existence theorem, Corollary 2.5 and Proposition 2.1 in Clarke (1989).

**Theorem 3.1. (Tonelli’s existence theorem)** When (T1)-(T3) hold, then \( \int_0^T L(t, x_t, x'_t) \, dt \) has a minimiser \( x^*_t \) in the space of absolutely continuous functions. Moreover,

- If \( (x, v) \to L(t, x, v) \) is convex, then the minimiser \( x^*_t \) Lipschitz continuous.
- If \( v \to L(t, x, v) \) is strictly convex, then the minimiser \( x^*_t \) is a twice continuously differentiable function on \([0, T]\).
- If \( (x, v) \to L(t, x, v) \) is strictly convex, then the minimiser \( x^*_t \) is unique.

**Remark 3.1.** The condition that \( (x, v) \to L(t, x, v) \) is strictly convex can be weakened to \( (x, v) \to L(t, x, v) \) is convex and \( v \to L(t, x, v) \) is strictly convex to get the uniqueness of the minimiser.

**Proposition 3.1.** Given a contract \( c^*_T \) offered by the principal to the agent, the optimal misreporting action \( a^{ME}_t \) and effort \( u^{ME}_t \) applied by the agent are given by

\[ a^{ME}_t = \frac{e}{\sigma \xi} \tanh \left( \frac{e t}{\sigma} \right) \alpha_t - \frac{e^2}{\sigma^2 \xi} \sinh \left( \frac{e t}{\sigma} \right) \int_t^T \tanh \left( \frac{e s}{\sigma} \right) \sech \left( \frac{e s}{\sigma} \right) \alpha_s \, ds \]  
(3.11)

\[ u^{ME}_t = \frac{e}{k} \alpha_t \]  
(3.12)

**Proof.** Using Eq. (3.7), we first note that the function \( t \to p_t \) is at least twice continuously differentiable with first derivative \( p'_t = \sinh \left( \frac{e^t}{\sigma} \right) a_t \) for each \( t \in [0, T] \). Therefore, we can rewrite the Lagrangian in Eq. (3.6) as

\[ \mathcal{L}(t, u_t, u'_t, p_t, p'_t) = \epsilon u_t \alpha_t - \frac{\gamma_1}{2\sigma^2} v_t^2 \alpha_t^2 + \frac{1}{2\sigma^2} \csc h \left( \frac{e t}{\sigma} \right) v_t \alpha_t p'_t - \frac{\epsilon}{\sigma^3} \sech \left( \frac{e t}{\sigma} \right) v_t \alpha_t p_t \]  

\[ - \left( \frac{k}{2} u_t^2 + \frac{\xi}{2} \left( \csc h \left( \frac{e^t}{\sigma} \right) \right)^2 (p'_t)^2 \right) \]  
(3.13)
Our aim is to maximise \( \int_0^T \mathcal{L} \) first over \( u_t \) and then over \( p_t \) satisfying Eqs.(3.7) and (3.8). Observe that \( \mathcal{L} \) is well-defined and smooth in all variables because it is a different way of writing the smooth Lagrangian in Eq. (3.6) via Eq. (3.7). Notice that the negative Lagrangian \(-\mathcal{L}\) is quadratic in \( u_t \) and there is no \( u_t' \) appearing. Since \( \mathcal{L} \) is strictly concave in \( u_t \), the maximiser \( u_t^* \) is unique and at least \( C^1[0, T] \). Moreover, it is found as the critical point. Notice that \( L \) is quadratic \( p_t' \), and linear in \( p_t \). Therefore, \(-\mathcal{L}\) is convex in \( (p_t, p_t') \). Additionally, \(-\mathcal{L}\) is strictly convex in \( p_t' \) because

\[
-\frac{\partial^2 \mathcal{L}}{(\partial p_t')^2} \geq \xi \cosh\left(\frac{eT}{\sigma}\right) > 0.
\]

It can be seen that \(-\mathcal{L}\) is coercive in \( u_t \), as it can be bounded below by \((p_t')^2\). Since \( \mathcal{L} \) is a smooth function in \( p_t' \) and \( p_t \), \(-\mathcal{L}\) becomes a Tonelli Lagrangian. By using the celebrated theorem of Tonelli (Theorem 3.1) we infer that \( \int_0^T \mathcal{L}(t, u_t, u_t', p_t, p_t') \, dt \) has a unique (at least \( C^1[0, T] \)) minimiser \( p_t^* \) that satisfy the Euler-Lagrange equation. Thus, we have the following system of equations

\[
\frac{\partial \mathcal{L}}{\partial u_t} = 0
\]
\[
\frac{d}{dt} \left( \frac{\partial \mathcal{L}}{\partial p_t'} \right) = \frac{\partial \mathcal{L}}{\partial p_t}
\]

From Eq. (3.14), we immediately obtain that the optimal \( u_t^* \) is given by Eq. (3.12). We may rewrite the second equation (Eq. (3.15)) as

\[
\frac{d}{dt} \left( \frac{1}{\sigma^2} \cosh\left(\frac{eT}{\sigma}\right) u_t \alpha_t - \xi \left( \cosh\left(\frac{eT}{\sigma}\right) \right)^2 (p_t^*)_t \right) = -\frac{\epsilon}{\sigma^3} \text{sech}\left(\frac{eT}{\sigma}\right) u_t \alpha_t
\]

Notice that \((p_t^*)_t\) is a free end point. Integrating from \( t \) to \( T \), and using the free boundary condition

\[
\left. \frac{\partial \mathcal{L}}{\partial p_t} \right|_{t=T} = 0 \iff \frac{1}{\sigma^2} \cosh\left(\frac{eT}{\sigma}\right) \alpha_T v_T = \xi \left( \sinh\left(\frac{eT}{\sigma}\right) \right)^2 (p_T^*)_T
\]

we arrive at

\[
\xi \left( \sinh\left(\frac{eT}{\sigma}\right) \right)^2 (p_t^*)_t = \frac{\epsilon}{\sigma} \text{sech}\left(\frac{eT}{\sigma}\right) \alpha_t - \frac{\epsilon^2}{\sigma^2} \int_t^T \text{tanh}\left(\frac{eS}{\sigma}\right) \text{sech}\left(\frac{es}{\sigma}\right) \alpha_s \, ds
\]

after using the explicit expression for \( v_t \) given by (2.4). From the definition of \( p_t \) in Eqs. (3.7) and (3.8), we can express Eq. (3.18) in terms of \( a_t \) leading to Eq. (3.11). This proves the proposition.

**Remark 3.2.** From Eqs. (3.11) and (3.12) we conclude that the optimal effort action \( u_t^{ME} \), the optimal misreporting action \( a_t^{ME} \) belong to \( C^1[0, T] \), as long as \( \alpha \) is \( C^1[0, T] \) as in Assumption 3.1.

### 3.3 Optimal Contract

The optimisation problem of the principal is to maximise over \( \alpha_t \in C^1[0, T] \) her utility under the reservation constraint that the agent expected utility given by Eq. (3.3) is equal to \( R_0 < 0 \). From this constraint, using Eq. (2.7), we obtain

\[
-R_0 = e^{-\gamma_1 e^{-\gamma_1 \int_0^T \alpha_t d\xi_t^u + \frac{1}{2} \gamma_1 \int_0^T \left(\alpha_t^2 - \xi_t^u \right) \, dt}} \left[ -e^{-\gamma_1 \int_0^T (\alpha_t^2 - \xi_t^u) \, dt} \right]
\]

leading to

\[
e^{-\gamma_1 e^{-\gamma_1 \int_0^T \alpha_t d\xi_t^u + \frac{1}{2} \gamma_1 \int_0^T \left(\alpha_t^2 - \xi_t^u \right) \, dt}} = -R_0 e^{-\gamma_1 \int_0^T \alpha_t d\xi_t^u + \frac{1}{2} \gamma_1 \int_0^T \left(\alpha_t^2 - \xi_t^u \right) \, dt}
\]
and therefore, solving for $c$ in the above equation, we obtain
\[
c = -\frac{1}{\gamma_1} \log(-R_0) - \int_0^T \left( \alpha_t \left[ \epsilon + u_t - \frac{1}{2} \frac{v_t^2}{\sigma^2} \gamma_1 \alpha_t + \frac{v_t a_t}{\sigma^2} - \frac{v_t}{\sigma} \left( \hat{x}_t^{u,0} - \hat{x}_t^{u,a} \right) \right] - \frac{1}{2} ku_t^2 - \frac{1}{2} \xi a_t^2 \right) dt
\]
(3.19)

Plugging it into the principal’s utility, and using again Eq. (3.4), we obtain that the principal optimises over $\alpha_t$ the following quantity
\[
U_p(\hat{x}_T^{u,0} - \mathcal{C}_T) = \mathbb{E}^{u,a} \left[ -\exp \left\{ -\gamma_2 \left( \hat{x}_T^{u,0} - c - \int_0^T \alpha_t d\hat{x}_t^{u,0} \right) \right\} \right]
\]
\[
= -\exp \left\{ -\gamma_2 \int_0^T (1 - \alpha_t) \left[ \epsilon u_t - \frac{1}{2} \gamma_2 (1 - \alpha_t) \frac{v_t^2}{\sigma^2} + \frac{v_t a_t}{\sigma^2} - \frac{v_t}{\sigma} \left( \hat{x}_t^{u,0} - \hat{x}_t^{u,a} \right) \right] dt \right\}
\]
\[
\times \exp \left\{ -\gamma_2 \int_0^T \alpha_t \left[ \epsilon u_t - \frac{1}{2} \gamma_2 \frac{v_t^2}{\sigma^2} + \frac{v_t a_t}{\sigma^2} - \frac{v_t}{\sigma} \left( \hat{x}_t^{u,0} - \hat{x}_t^{u,a} \right) \right] dt \right\}
\]
\[
\times \exp \left\{ -\frac{\gamma_2}{\gamma_1} \log(-R_0) \right\} \exp \left\{ \frac{1}{2} \gamma_2 \left( ku_t^2 + \xi a_t^2 \right) dt \right\}
\]

where we used that
\[
\hat{x}_T^{u,0} = \int_0^T d\hat{x}_t^{u,0}.
\]

Therefore, the principal wants to maximise the following quantity over $\alpha_t$:
\[
\mathcal{A}(t, \alpha_t) = \int_0^T \left[ \epsilon u_t^{ME} + \frac{v_t}{\sigma^2} a_t^{ME} - \frac{v_t}{\sigma^2} \left( \hat{x}_t^{u,0} - \hat{x}_t^{u,a} \right) - \frac{v_t^2}{2\sigma^2} (\gamma_2 (1 - \alpha_t)^2 + \gamma_1 \alpha_t^2) \right] dt
\]
\[
- \frac{1}{2} \int_0^T \left[ k \left( u_t^{ME} \right)^2 + \xi \left( a_t^{ME} \right)^2 \right] dt
\]
(3.20)

and $a_t^{ME}$ and $u_t^{ME}$ are the optimal misreporting action and effort given, respectively, by Eq. (3.11) and Eq. (3.12). Using Eq. (3.11), we can rewrite the term $\hat{x}_t^{u,0} - \hat{x}_t^{u,a}$, at in Eq. (2.7) as
\[
\hat{x}_t^{u,0} - \hat{x}_t^{u,a} = \frac{c^2}{\sigma^2 \xi} \int_0^t \left[ \sinh \left( \frac{\epsilon s}{\sigma} \right) \tanh \left( \frac{\epsilon s}{\sigma} \right) \alpha_s - \frac{\epsilon}{\sigma} \sinh^2 \left( \frac{\epsilon s}{\sigma} \right) \left( \int_s^T \tanh \left( \frac{\epsilon s}{\sigma} \right) \sech \left( \frac{\epsilon s}{\sigma} \right) \alpha_s ds \right) \right] dr
\]
(3.21)

We next introduce the following functions
\[
\beta_t \triangleq \int_0^t \sinh \left( \frac{\epsilon s}{\sigma} \right) \tanh \left( \frac{\epsilon s}{\sigma} \right) \alpha_s ds
\]
(3.22)
\[
\eta_t \triangleq \int_t^T \tanh \left( \frac{\epsilon s}{\sigma} \right) \sech \left( \frac{\epsilon s}{\sigma} \right) \alpha_s ds
\]
(3.23)
\[
\zeta_t \triangleq \int_0^t \sinh^2 \left( \frac{\epsilon s}{\sigma} \right) \eta_s ds
\]
(3.24)

so that
\[
\hat{x}_t^{u,0} - \hat{x}_t^{u,a} = \frac{c^2}{\sigma^2 \xi} \sech \left( \frac{\epsilon t}{\sigma} \right) \left( \beta_t - \frac{\epsilon}{\sigma} \zeta_t \right).
\]
(3.25)

Using the fundamental theorem of calculus, we obtain from Eq. (3.23) and (3.22) that
\[
\beta_t' = -\sinh \left( \frac{\epsilon t}{\sigma} \right) \cosh \left( \frac{\epsilon t}{\sigma} \right) \eta_t
\]
(3.26)
and from Eq. (3.24) that
\[ \zeta' = \sinh^2 \left( \frac{\epsilon t}{\sigma} \right) \eta_t. \] (3.27)

Moreover, we have that \( \beta_T, \eta_0, \) and \( \zeta_T \) are free of constraints, while \( \beta_0 = \eta_T = \zeta_0 = 0. \) Eq. (3.22), (3.23) and (3.24) turn the integral of Eq. (3.20) into
\[ \mathcal{A}^p(t, \beta_t, \eta_t, \zeta_t) = \int_0^T \mathcal{X}(t, \beta_t, \eta_t, \zeta_t, \beta'_t, \eta'_t, \zeta'_t) \, dt \] (3.28)
where
\[ \mathcal{X}(t, \beta_t, \eta_t, \zeta_t, \beta'_t, \eta'_t, \zeta'_t) = + \frac{e^2}{k} \csc \left( \frac{\epsilon t}{\sigma} \right) \coth \left( \frac{\epsilon t}{\sigma} \right) \beta'_t \]
\[ + \frac{v_t}{\sigma^2} \left( \frac{\epsilon}{\sigma \xi} \csc \left( \frac{\epsilon t}{\sigma} \right) \beta'_t - \frac{e^2}{2 \sigma^2 \xi} \sinh \left( \frac{\epsilon t}{\sigma} \right) \eta_t \right) \]
\[ - \frac{v_t}{\sigma^3} \text{sech} \left( \frac{\epsilon t}{\sigma} \right) \left( \frac{e^2}{\sigma \xi} \beta_t - \frac{e^3}{\sigma^2 \xi} \zeta_t \right) \]
\[ - \frac{v_t^2}{2 \sigma \gamma_1} \cosh^2 \left( \frac{\epsilon t}{\sigma} \right) \coth^2 \left( \frac{\epsilon t}{\sigma} \right) \left( \beta_t' \right)^2 \]
\[ - \frac{v_t^2}{2 \sigma \gamma_2} \left( \csc \left( \frac{\epsilon t}{\sigma} \right) \coth \left( \frac{\epsilon t}{\sigma} \right) \beta_t' - 1 \right)^2 \]
\[ - \frac{e^2}{2k} \csc^2 \left( \frac{\epsilon t}{\sigma} \right) \coth^2 \left( \frac{\epsilon t}{\sigma} \right) \left( \beta_t \right)^2 \]
\[ - \frac{\xi}{2} \frac{\epsilon}{\sigma \xi} \csc \left( \frac{\epsilon t}{\sigma} \right) \beta'_t - \frac{e^2}{\sigma^2 \xi} \sinh \left( \frac{\epsilon t}{\sigma} \right) \eta_t \right) \] (3.29)

Remark 3.3. (Well-posedness of the maximisation problem) First note that \( \mathcal{X} \) is defined for all \( t \in [0, T] \), as it is a different version of the Lagrangian (integrand) in Eq. (3.20). Furthermore, it is quadratic in \((\beta', \eta', \zeta')\) and smooth in all variables due to equations (3.26)-(3.27). One can easily check that \( -\mathcal{X} \) is convex in \((\beta_t, \eta_t, \zeta_t, \eta'_t, \zeta'_t)\). Indeed, letting \( \beta_t = w_1, \eta_t = w_2, \zeta_t = w_3, \beta'_t = w_4, \eta'_t = w_5, \zeta'_t = w_6 \) and \( w = (w_1, w_2, w_3, w_4, w_5, w_6) \) we have
\[ \mathcal{H}_\mathcal{X} \triangleq \frac{\partial^2 \mathcal{X}}{\partial w^2} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & a_{22} & 0 & a_{24} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & a_{42} & 0 & a_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \]

where
\[ a_{22} = -\frac{e^4}{\sigma^4 \xi} \sinh^2 \left( \frac{\epsilon t}{\sigma} \right), \quad a_{24} = a_{42} = \frac{e^3}{\sigma^3 \xi}, \]
\[ a_{44} = -\left( \frac{e^2}{\sigma^2 \xi} + e^2 (\gamma_1 + \gamma_2) + \frac{e^2}{k} \cosh^2 \left( \frac{\epsilon t}{\sigma} \right) \right) \csc^2 \left( \frac{\epsilon t}{\sigma} \right). \]

Therefore, we conclude that the eigenvalues \( \mu_i \) (1 ≤ i ≤ 3) of the Hessian matrix \( \mathcal{H}_\mathcal{X} \) given by
\[ \mu_1 = 0, \quad \mu_{2,3} = \frac{1}{2} \left[ a_{22} + a_{44} \pm \sqrt{(a_{44} - a_{22})^2 + 4(a_{44})^2} \right] \]
are all non-positive due to \( a_{24}^2 < a_{44}a_{22} \) for all \( t \geq 0. \)
Furthermore, as a consequence of Eqs. (3.26)-(3.27), it is enough to have the coercivity and strict convexity of $-X$ in only one component of $(\beta', \eta', \zeta')$, as we can rewrite $X$ to make the quadratic terms $(\beta')^2$, $(\eta')^2$, and $(\zeta')^2$ appear in $X$. Therefore, it is not difficult to see that $-X$ is coercive as well as strictly convex in $(\beta', \eta', \zeta')$. Thus, it turns out that $-X$ is a Tonelli Lagrangian and so again by Theorem 3.1 there exist unique (at least $C^2[0, T]$) maximizers $\beta, \eta, \zeta$ of $A^P$. Hence $\beta, \eta, \zeta$ satisfy

$$\max_{\alpha_t} \{ A(t, \alpha_t) : a_t^{ME} \text{ satisfies (3.11) and } u_t^{ME} \text{ satisfies (3.12)} \} =$$

$$\max_{(\beta_t, \eta_t, \zeta_t)} \{ A^P(t, \beta_t, \eta_t, \zeta_t) : \beta_t \text{ and } \eta_t \text{ satisfy (3.26), } \eta_t \text{ and } \zeta_t \text{ satisfy (3.27)} \} \tag{3.30}$$

We next give the following lemma to characterise the unique maximisers $(\beta_t, \eta_t, \zeta_t)$ of $A^P$.

**Lemma 3.2.** If $(\beta_t, \eta_t, \zeta_t)$ maximises $A^P$ in Eq. (3.28) with Eq. (3.26), Eq. (3.27) and $\beta_0, \eta_0, \zeta_0$, then the following equations are satisfied:

$$\frac{\partial X}{\partial \beta_t} - \frac{d}{dt} \left( \frac{\partial X}{\partial \beta'_t} + \lambda_2(t) \right) = 0 \tag{3.31}$$

$$\lambda_2(T) = -\left. \frac{\partial X}{\partial \beta'_t} \right|_{t=T} \tag{3.32}$$

$$\frac{\partial X}{\partial \eta_t} - \lambda_1(t) \sinh^2 \left( \frac{et}{\sigma} \right) = \frac{d}{dt} \left( \frac{\partial X}{\partial \eta'_t} + \lambda_2(t) \sinh \left( \frac{et}{\sigma} \right) \cosh \left( \frac{et}{\sigma} \right) \right) \tag{3.33}$$

$$\left[ \frac{\partial X}{\partial \eta'_t} + \lambda_2(t) \sinh \left( \frac{et}{\sigma} \right) \cosh \left( \frac{et}{\sigma} \right) \right]_{t=0} = 0 \tag{3.34}$$

$$\frac{\partial X}{\partial \zeta_t} - \frac{d}{dt} \left( \frac{\partial X}{\partial \zeta'_t} + \lambda_1(t) \right) = 0 \tag{3.35}$$

$$\lambda_1(T) = -\left. \frac{\partial X}{\partial \zeta'_t} \right|_{t=T} \tag{3.36}$$

$$\zeta'_t = \sinh^2 \left( \frac{et}{\sigma} \right) \eta_t \tag{3.37}$$

$$\beta'_t = -\sinh \left( \frac{et}{\sigma} \right) \cosh \left( \frac{et}{\sigma} \right) \eta'_t \tag{3.38}$$

**Proof.** First, we define

$$\chi^\lambda_1, \lambda_2(t, \beta_t, \eta_t, \zeta_t, \beta'_t, \eta'_t, \zeta'_t) \triangleq X(t, \beta_t, \eta_t, \zeta_t, \beta'_t, \eta'_t, \zeta'_t) + \lambda_1(t) \left( \zeta'_t - \sinh^2 \left( \frac{et}{\sigma} \right) \eta_t \right) + \lambda_2(t) \left( \beta'_t + \sinh \left( \frac{et}{\sigma} \right) \cosh \left( \frac{et}{\sigma} \right) \eta'_t \right) \tag{3.39}$$

Consider the functional

$$\Psi(\beta_t, \eta_t, \zeta_t) \triangleq \int_0^T X^\lambda_1, \lambda_2(t, \beta_t, \eta_t, \zeta_t, \beta'_t, \eta'_t, \zeta'_t) dt. \tag{3.40}$$

Suppose that $(\beta_n, \eta_n, \zeta_n)$ maximises the functional $\Psi$ defined by Eq. (3.40). Let $\tilde{\beta}_t, \tilde{\eta}_t, \tilde{\zeta}_t$ be smooth functions such that

$$\tilde{\beta}_0 = \tilde{\eta}_T = \tilde{\zeta}_0 = 0. \tag{3.41}$$

Therefore, any perturbation $(\beta_t + \delta_1 \tilde{\beta}_t, \eta_t + \delta_2 \tilde{\eta}_t, \zeta_t + \delta_3 \tilde{\zeta}_t)$ of $(\beta_n, \eta_n, \zeta_n)$ would yield

$$\Psi(\beta_t, \eta_t, \zeta_t) \geq \Psi(\beta_t + \delta_1 \tilde{\beta}_t, \eta_t + \delta_2 \tilde{\eta}_t, \zeta_t + \delta_3 \tilde{\zeta}_t).$$
for any real number $\delta_1, \delta_2, \text{ and } \delta_3$. Thus, we conclude that $(\delta_1, \delta_2, \delta_3) = (0, 0, 0)$ is the critical point of the map

$$(\delta_1, \delta_2, \delta_3) \mapsto \Psi(\beta_t + \delta_1 \tilde{\beta}_t, \eta_t + \delta_2 \tilde{\eta}_t, \zeta_t + \delta_3 \tilde{\zeta}_t).$$

Therefore, we must have

$$\frac{\partial}{\partial \delta_1} \Psi(\beta_t + \delta_1 \tilde{\beta}_t, \eta_t, \zeta_t) \bigg|_{\delta_1 = 0} = 0, \quad (3.42)$$

and

$$\frac{\partial}{\partial \delta_2} \Psi(\beta_t, \eta_t + \delta_2 \tilde{\eta}_t, \zeta_t) \bigg|_{\delta_2 = 0} = 0. \quad (3.43)$$

and

$$\frac{\partial}{\partial \delta_3} \Psi(\beta_t, \eta_t, \zeta_t + \delta_3 \tilde{\zeta}_t) \bigg|_{\delta_3 = 0} = 0. \quad (3.44)$$

By (3.42) we have

$$\int_0^T \left( \frac{\partial \mathcal{X}}{\partial \beta_t} \tilde{\beta}_t + \frac{\partial \mathcal{X}}{\partial \beta_t} \beta_t + \lambda_2(t) \tilde{\beta}_t \right) dt = 0$$

Application of the integration by parts yields

$$\int_0^T \left( \frac{\partial \mathcal{X}}{\partial \beta_t} - \frac{d}{dt} \left[ \frac{\partial \mathcal{X}}{\partial \beta_t} + \lambda_2(t) \right] \right) \tilde{\beta}_t dt + \left[ \left( \frac{\partial \mathcal{X}}{\partial \beta_t} + \lambda_2(t) \right) \tilde{\beta}_t \right]_{t=0}^{t=T} = 0.$$

This, together with the free boundary $\tilde{\beta}_T$, gives equations (3.31) and (3.32). Similarly, by (3.43) we have

$$\int_0^T \left( \frac{\partial \mathcal{X}}{\partial \eta_t} \tilde{\eta}_t + \frac{\partial \mathcal{X}}{\partial \eta_t} \eta_t - \lambda_1(t) \sinh^2 \left( \frac{\epsilon t}{\sigma} \right) \tilde{\eta}_t + \lambda_2(t) \sinh \left( \frac{\epsilon t}{\sigma} \right) \cosh \left( \frac{\epsilon t}{\sigma} \right) \tilde{\eta}_t \right) dt = 0.$$

Using the integration by parts we arrive at

$$\int_0^T \left( \frac{\partial \mathcal{X}}{\partial \eta_t} - \lambda_1(t) \sinh^2 \left( \frac{\epsilon t}{\sigma} \right) - \frac{d}{dt} \left[ \frac{\partial \mathcal{X}}{\partial \eta_t} + \lambda_2(t) \sinh \left( \frac{\epsilon t}{\sigma} \right) \cosh \left( \frac{\epsilon t}{\sigma} \right) \right] \right) \tilde{\eta}_t dt$$

$$+ \left[ \left( \frac{\partial \mathcal{X}}{\partial \eta_t} + \lambda_2(t) \sinh \left( \frac{\epsilon t}{\sigma} \right) \cosh \left( \frac{\epsilon t}{\sigma} \right) \right) \tilde{\eta}_t \right]_{t=0}^{t=T} = 0$$

from which, together with free boundary $\tilde{\eta}_0$, we obtain equations (3.33) and (3.34). By (3.44) we have

$$\int_0^T \left( \frac{\partial \mathcal{X}}{\partial \zeta_t} \tilde{\zeta}_t + \frac{\partial \mathcal{X}}{\partial \zeta_t} \zeta_t + \lambda_1(t) \tilde{\zeta}_t \right) dt = 0$$

Application of the integration by parts yields

$$\int_0^T \left( \frac{\partial \mathcal{X}}{\partial \zeta_t} - \frac{d}{dt} \left[ \frac{\partial \mathcal{X}}{\partial \zeta_t} + \lambda_1(t) \right] \right) \tilde{\zeta}_t dt + \left[ \left( \frac{\partial \mathcal{X}}{\partial \zeta_t} + \lambda_1(t) \right) \tilde{\zeta}_t \right]_{t=0}^{t=T} = 0.$$

With the aid of free boundary $\tilde{\zeta}_T$, we obtain equations (3.35) and (3.36).

Utilising Lemma 3.2 we obtain the following result.

**Theorem 3.2.** The maximisers $\beta_t, \eta_t, \text{ and } \zeta_t$ of $A^t$ satisfy

$$A(t) \eta_t' + B(t) \eta_t + C(t) \eta_t + D(t) = 0, \quad \eta_T = 0, \quad \eta_0' = 0 \quad (3.45)$$
where
\[
A(t) = \epsilon^2 \left( \gamma_1 + \gamma_2 + \frac{1}{\sigma^2\xi} \right) \cosh^2 \left( \frac{et}{\sigma} \right) + \frac{\epsilon^2}{k} \coth^2 \left( \frac{et}{\sigma} \right) \cosh^2 \left( \frac{et}{\sigma} \right)
\]
\[
B(t) = \frac{2\epsilon^2}{\sigma} \left( \gamma_1 + \gamma_2 + \frac{1}{\sigma^2\xi} \right) \sinh \left( \frac{et}{\sigma} \right) \cosh \left( \frac{et}{\sigma} \right) + \frac{2\epsilon^3}{\sigma k} \coth \left( \frac{et}{\sigma} \right) \left( \cosh^2 \left( \frac{et}{\sigma} \right) - \coth^2 \left( \frac{et}{\sigma} \right) \right)
\]
\[
C(t) = \frac{\epsilon^4}{\sigma^4\xi} \cosh^2 \left( \frac{et}{\sigma} \right)
\]
\[
D(t) = \frac{\epsilon^3}{\sigma^3\xi} \cosh^2 \left( \frac{et}{\sigma} \right) \sech \left( \frac{et}{\sigma} \right) + \frac{\epsilon^3}{\sigma} \cosh \left( \frac{et}{\sigma} \right) \left( \gamma_2 + \frac{1}{k} \left( 1 - \cosh^2 \left( \frac{et}{\sigma} \right) \right) \right). \tag{3.49}
\]

Moreover, we infer that the optimal pay-per-performance sensitivity satisfies the initial condition \( \alpha_0 = 1 \).

**Proof.** Here, we compute equations (3.31)-(3.36) and use (3.37) and (3.38). Note that using \( v_t = \sigma \tanh \left( \frac{et}{\sigma} \right) \) we get
\[
\frac{\partial \mathcal{X}}{\partial \beta_t} = -\frac{\epsilon^2}{\sigma^2\xi} \sech \left( \frac{et}{\sigma} \right) \frac{d}{dt} \left[ \sech \left( \frac{et}{\sigma} \right) \right]. \tag{3.50}
\]
Also,
\[
\frac{\partial \mathcal{X}}{\partial \beta_t} = \frac{\epsilon^2}{k} \csch \left( \frac{et}{\sigma} \right) \coth \left( \frac{et}{\sigma} \right) \left( \gamma_1 + \gamma_2 \right) \left( \csch \left( \frac{et}{\sigma} \right) \right)^2 \beta_t' + \epsilon^2 \left( \gamma_2 \frac{1}{\sigma^2\xi} \right) \sech \left( \frac{et}{\sigma} \right)
\]
\[-\epsilon^2 \csch^2 \left( \frac{et}{\sigma} \right) \frac{1}{\sigma^2\xi} \coth^2 \left( \frac{et}{\sigma} \right) \beta_t' + \frac{\epsilon^3}{\sigma^3\xi} \eta_t \tag{3.51}
\]
Therefore, Eq. (3.31) can be written as
\[
\frac{d}{dt} \left[ \frac{\epsilon^2}{\sigma^2\xi} \sech \left( \frac{et}{\sigma} \right) - \frac{\partial \mathcal{X}}{\partial \beta_t} - \lambda_2(t) \right] = 0. \tag{3.52}
\]
Integrating Eq. (3.52) from \( t \) to \( T \) and using equations (3.51) and (3.38) together with the boundary condition given by Eq. (3.32), we get
\[
\lambda_2(t) = \left( -\frac{\epsilon^2}{k} \coth^2 \left( \frac{et}{\sigma} \right) - \epsilon^2 \left( \gamma_1 + \gamma_2 + \frac{1}{\sigma^2\xi} \right) \right) \coth \left( \frac{et}{\sigma} \right) \eta_t - \frac{\epsilon^3}{\sigma^3\xi} \eta_t
\]
\[-\epsilon^2 \frac{\sigma^2\xi^2}{\sigma^2\xi} \sech \left( \frac{et}{\sigma} \right) - \gamma_2 \epsilon^2 \sech \left( \frac{et}{\sigma} \right) - \frac{\epsilon^2}{k} \csch \left( \frac{et}{\sigma} \right) \coth \left( \frac{et}{\sigma} \right) \right) \tag{3.53}
\]
Also, using (3.38), we have
\[
\frac{\partial \mathcal{X}}{\partial \eta_t} = -\frac{\epsilon^3}{\sigma^3\xi} \sinh \left( \frac{et}{\sigma} \right) \cosh \left( \frac{et}{\sigma} \right) \eta_t - \frac{\epsilon^4}{\sigma^4\xi} \sinh^2 \left( \frac{et}{\sigma} \right) \eta_t - \frac{\epsilon^3}{\sigma^3\xi} \sinh^2 \left( \frac{et}{\sigma} \right) \sech \left( \frac{et}{\sigma} \right) \tag{3.54}
\]
Since
\[
\frac{\partial \mathcal{X}}{\partial \eta_t} = 0 \tag{3.55}
\]
Eq. (3.34) becomes
\[
\left[ \lambda_2(t) \sinh \left( \frac{et}{\sigma} \right) \cosh \left( \frac{et}{\sigma} \right) \right]_{t=0} = 0 \tag{3.56}
\]
Using the definition of \( \eta_t \) in Eq. (3.23) we have that
\[
\eta_t' = -\tanh \left( \frac{et}{\sigma} \right) \sech \left( \frac{et}{\sigma} \right) \alpha_t \tag{3.57}
\]
which together with (3.56) produce the boundary condition \( \alpha_0 = 1 \). As a result, due to Eq. (3.57), we have the boundary condition \( \eta'_0 = 0 \) in Eq. (3.45).

To use Eq. (3.33) we need to compute (3.35) and (3.36). Note that

\[
\frac{\partial \mathcal{X}}{\partial \zeta} = \frac{\alpha^3 v_t}{\sigma^3 \xi} \text{sech} \left( \frac{\sigma t}{\sigma} \right) = -\frac{\alpha^3}{\sigma^3 \xi} \frac{d}{dt} \left[ \text{sech} \left( \frac{\sigma t}{\sigma} \right) \right].
\] (3.58)

Since

\[
\frac{\partial \mathcal{X}}{\partial \zeta} = 0
\]

Eq. (3.36) becomes

\[
\lambda_1(T) = 0.
\] (3.60)

Using equations (3.58) and (3.35) we obtain

\[
\frac{d}{dt} \left[ -\frac{\alpha^3}{\sigma^3 \xi} \text{sech} \left( \frac{\sigma t}{\sigma} \right) - \lambda_1(t) \right] = 0
\] (3.61)

Integrating (3.61) from \( t \) to \( T \) and using Eq. (3.60) we get

\[
\lambda_1(t) = \frac{\alpha^3}{\sigma^3 \xi} \left[ \text{sech} \left( \frac{\sigma T}{\sigma} \right) - \text{sech} \left( \frac{\sigma t}{\sigma} \right) \right]
\] (3.62)

Now substituting equations (3.62) and (3.53) into (3.33) we obtain

\[
\frac{\partial \mathcal{X}}{\partial \eta_t} - \lambda_1(t) \sinh^2 \left( \frac{\sigma t}{\sigma} \right) \cosh \left( \frac{\sigma t}{\sigma} \right) = -\frac{\alpha^3}{\sigma^3 \xi} \sinh \left( \frac{\sigma t}{\sigma} \right) \cosh \left( \frac{\sigma t}{\sigma} \right) \eta_t - \frac{\alpha^4}{\sigma^4 \xi} \sinh^2 \left( \frac{\sigma t}{\sigma} \right) \eta_t
\]

\[
-\frac{\alpha^3}{\sigma^3 \xi} \sinh^2 \left( \frac{\sigma t}{\sigma} \right) \text{sech} \left( \frac{\sigma T}{\sigma} \right)
\]

\[
\frac{d}{dt} \left( \frac{\partial \mathcal{X}}{\partial \eta_t} + \lambda_2(t) \sinh \left( \frac{\sigma t}{\sigma} \right) \cosh \left( \frac{\sigma t}{\sigma} \right) \right) = -\frac{\alpha^3}{\sigma^3 \xi} \text{sech} \left( \frac{\sigma T}{\sigma} \right) \left[ \cosh^2 \left( \frac{\sigma t}{\sigma} \right) + \sinh^2 \left( \frac{\sigma t}{\sigma} \right) \right]
\]

\[
+\frac{\alpha^3}{\sigma k} \cosh \left( \frac{\sigma t}{\sigma} \right) \left[ \cosh^2 \left( \frac{\sigma t}{\sigma} \right) + \sinh^2 \left( \frac{\sigma t}{\sigma} \right) \right] - 1 - k\gamma_2
\]

\[
-\frac{\alpha^3}{\sigma^3 \xi} \sinh \left( \frac{\sigma t}{\sigma} \right) \cosh \left( \frac{\sigma t}{\sigma} \right) \eta_t
\]

\[
-2\frac{\alpha^3}{\sigma} \left( \gamma_1 + \gamma_2 + \frac{1}{\sigma^2 \xi} \right) \sinh \left( \frac{\sigma t}{\sigma} \right) \cosh \left( \frac{\sigma t}{\sigma} \right) \eta_t'
\]

\[
-2\frac{\alpha^3}{\sigma k} \cosh \left( \frac{\sigma t}{\sigma} \right) \left[ \cosh^2 \left( \frac{\sigma t}{\sigma} \right) - \cosh^2 \left( \frac{\sigma t}{\sigma} \right) \right] \eta_t''
\]

\[
-\frac{\alpha^3}{\sigma} \left( \gamma_1 + \gamma_2 + \frac{1}{\sigma^2 \xi} \right) \cosh \left( \frac{\sigma t}{\sigma} \right) \eta_t''
\]

\[
-\frac{\alpha^3}{\sigma k} \coth \left( \frac{\sigma t}{\sigma} \right) \cosh \left( \frac{\sigma t}{\sigma} \right) \eta_t'''
\]

(3.64)

Substitution of equations (3.63) and (3.64) into Equation (3.33) results in the differential equation given by (3.45), which completes the proof.
Remark 3.4. Finding the explicit solution of the singular ODE in Eq. (3.45) with the prescribed boundary conditions is not an easy task and accordingly we prefer to use a numerical approach to recover and analyze the solution of (3.45).

Remark 3.5. Due to Eq. (3.57) the maximizer \( \alpha_t^{ME} \) of \( A(t, \alpha_t) \) is uniquely obtained from
\[
\alpha_0^{ME} = 1, \quad \alpha_t^{ME} = - \coth \left( \frac{ct}{\sigma} \right) \cosh \left( \frac{ct}{\sigma} \right) \eta'_t \quad \text{for} \quad 0 < t \leq T
\]
and using Eq. (3.11) and Eq. (3.12) the optimal \( a_t^{ME} \) and \( u_t^{ME} \) are uniquely expressed as
\[
a_0^{ME} = 0, \quad a_t^{ME} = - \frac{c}{\sigma^2} \cosh \left( \frac{ct}{\sigma} \right) \eta'_t - \frac{c^2}{\sigma^2} \sinh \left( \frac{ct}{\sigma} \right) \eta_t \quad \text{for} \quad 0 < t \leq T
\]
\[
u_0^{ME} = \frac{c}{k}, \quad \nu_t^{ME} = - \frac{c}{k} \coth \left( \frac{ct}{\sigma} \right) \cosh \left( \frac{ct}{\sigma} \right) \eta'_t \quad \text{for} \quad 0 < t \leq T
\]
Therefore, by Remarks 3.2 and 3.3, we conclude that \( \alpha_t^{ME}, a_t^{ME} \) and \( u_t^{ME} \) are all at least \( C^1[0, T] \) functions. Notice that initially, the principal gives full incentive to the agent, who will apply the maximum level of effort. There will be no misreporting, because we are assuming that \( x_0 \) is common knowledge, and therefore there will be no utility coming from misreporting.

3.4 Limiting Cases

In this section, we show that the contracting framework proposed in this paper includes, as special cases, the well known Holmstrom and Milgrom (1987) model with only effort present, and the contracting model with only misreporting present. We recall the expressions for the pay-per-performance sensitivity \( \alpha_t^{NH} \) and effort \( u_t^{NH} \) in the Holmstrom-Milgrom model, given by
\[
\alpha_t^{HM} = \frac{\gamma_2 + (1/k)}{\gamma_1 + \gamma_2 + (1/k)}, \quad u_t^{HM} = \frac{c}{k} \alpha_t^{HM}
\]
and the expressions for the pay-per-performance sensitivity \( \alpha_t^{M} \) and effort \( u_t^{M} \) in the model with only misreporting derived in Capponi, Cvitanic and Yolcu (2011), given by
\[
\alpha_t^{M} = \frac{\gamma_1}{\gamma_1 + \gamma_2} + \frac{1}{\gamma_1 + \gamma_2} \frac{1}{\gamma_1 + \gamma_2} \cosh(\omega T) \left[ - \omega \sigma \sinh(\omega t) + c \cosh(\omega t) \tanh \left( \frac{ct}{\sigma} \right) \right]
\]
\[
a_t^{M} = - \frac{\gamma_1}{(1 + \gamma_2) \xi} \cosh(\omega T) \sinh(\omega t) + \frac{c^2}{\sigma^2} \cosh \left( \frac{ct}{\sigma} \right) \eta'_t
\]
where
\[
\omega = \left( \frac{\gamma_1 + \gamma_2}{\gamma_1 + \gamma_2 + (1/k)} \right)^{1/2}
\]
In case when there is no observation noise, \( \sigma \rightarrow 0 \), and no misreporting action, \( \xi \rightarrow \infty \), then the actual output is observed, and the agent is discouraged from any misreporting from the infinite penalty incurred. Then, as one would expect, our contracting framework reduces to Holmstrom and Milgrom (1987). The next proposition proves this result.

Proposition 3.2. The following limits holds:
\[
\lim_{\sigma \rightarrow 0} \left( \lim_{\xi \rightarrow \infty} \alpha_t^{ME} \right) = \alpha_t^{HM}, \quad \lim_{\sigma \rightarrow 0} \left( \lim_{\xi \rightarrow \infty} u_t^{ME} \right) = u_t^{HM}
\]

Proof. It is enough to prove that the pay-per-performance sensitivities are the same, as the same relation between \( u_t \) and \( \alpha_t \) persists in both contracting models (compare Eq. (3.12) and Eq. (3.68)).
We first compute the limiting behavior as $\xi \to \infty$ of the ODE given in (3.45):

$$\eta'' = - \tan \left( \frac{\xi t}{\sigma} \right) \sech \left( \frac{\xi t}{\sigma} \right) \alpha_t$$  (3.73)

$$\eta''' = - \tan \left( \frac{\xi t}{\sigma} \right) \sech \left( \frac{\xi t}{\sigma} \right) \alpha_t' - \frac{e}{\sigma} \sech^3 \left( \frac{\xi t}{\sigma} \right) \alpha_t + \frac{e}{\sigma} \tanh^2 \left( \frac{\xi t}{\sigma} \right) \sech \left( \frac{\xi t}{\sigma} \right) \alpha_t$$  (3.74)

$$A(t)^{(\xi \to \infty)} \eta''' = \frac{e^3}{\sigma} \left( \gamma_1 + \gamma_2 + \frac{1}{k} \coth \left( \frac{\xi t}{\sigma} \right) \right) \left( \tanh \left( \frac{\xi t}{\sigma} \right) \sinh \left( \frac{\xi t}{\sigma} \right) - \frac{\xi t}{\sigma} \right) \alpha_t$$

$$- \frac{e^2}{\sigma} \left( \gamma_1 + \gamma_2 + \frac{1}{k} \coth \left( \frac{\xi t}{\sigma} \right) \right) \sinh \left( \frac{\xi t}{\sigma} \right) \alpha_t'$$  (3.75)

$$B(t)^{(\xi \to \infty)} \eta' = - \frac{2e^3}{\sigma} \left( \gamma_1 + \gamma_2 \right) \tanh \left( \frac{\xi t}{\sigma} \right) \alpha_t$$

$$D(t)^{(\xi \to \infty)} \eta = \frac{e^3}{\sigma} \cosh \left( \frac{\xi t}{\sigma} \right) \left( \gamma_2 + \frac{1}{k} \left( 1 - \csc \left( \frac{\xi t}{\sigma} \right) \right) \right)$$  (3.78)

Using equations (3.75)-(3.78) in (3.45) and dividing everything by $\cosh \left( \frac{\xi t}{\sigma} \right)$ and multiplying by $\sigma$, we arrive at

$$0 = e^3 \left( \gamma_1 + \gamma_2 + \frac{1}{k} \coth \left( \frac{\xi t}{\sigma} \right) \right) \left( \tanh^2 \left( \frac{\xi t}{\sigma} \right) - \frac{\xi t}{\sigma} \right) \alpha_t$$

$$- \frac{e^2}{\sigma} \left( \gamma_1 + \gamma_2 + \frac{1}{k} \coth \left( \frac{\xi t}{\sigma} \right) \right) \tanh \left( \frac{\xi t}{\sigma} \right) \alpha_t'$$

$$- 2e^3 \left( \gamma_1 + \gamma_2 \right) \tanh \left( \frac{\xi t}{\sigma} \right) \alpha_t$$

$$+ \frac{2e^3}{k} \sech^2 \left( \frac{\xi t}{\sigma} \right) \left( \cosh^2 \left( \frac{\xi t}{\sigma} \right) - \coth^2 \left( \frac{\xi t}{\sigma} \right) \right) \alpha_t$$

$$+ e^3 \left( \gamma_2 + \frac{1}{k} \left( 1 - \csc \left( \frac{\xi t}{\sigma} \right) \right) \right)$$  (3.79)

Observe that

$$\lim_{\sigma \to 0} \tan \left( \frac{\xi t}{\sigma} \right) = \lim_{\sigma \to 0} \frac{\xi t}{\sigma} = 1, \quad \lim_{\sigma \to 0} \frac{\xi t}{\sigma} = \lim_{\sigma \to 0} \frac{\xi t}{\sigma} = 0.$$

Taking the limit as $\alpha \to 0$ in Equation (3.79), we get

$$- e^3 \left( \gamma_1 + \gamma_2 + \frac{1}{k} \right) \alpha_t + e^3 \left( \gamma_2 + \frac{1}{k} \right) = 0$$  (3.80)

which together with Eq. (3.68) gives Eq. (3.72).

In the case in which the penalty for effort becomes infinite, the principal knows the agent will not apply effort, and therefore only provides incentives for misreporting. Therefore, we expect the optimal $\alpha_t$ to be the same as in the model with only misreporting, developed in Capponi, Cvitanić and Yolcu (2011). This is indeed the case, and we have

**Proposition 3.3.** The following holds

$$\lim_{k \to \infty} \alpha_t^{ME} = \alpha_t^M, \quad \lim_{k \to \infty} \alpha_t^{ME} = \alpha_t^M$$  (3.81)
Proof. We start computing the limiting behavior of the ODE given by (3.45) as \( k \to \infty \). We obtain

\[
A(t)^{(k\to\infty)} = e^2 \left( \gamma_1 + \gamma_2 + \frac{1}{\sigma^2 \xi} \right) \cosh \left( \frac{et}{\sigma} \right)^2
\]  

(3.82)

\[
B(t)^{(k\to\infty)} = \frac{2e^3}{\sigma} \left( \gamma_1 + \gamma_2 + \frac{1}{\sigma^2 \xi} \right) \sinh \left( \frac{et}{\sigma} \right) \cosh \left( \frac{et}{\sigma} \right)
\]  

(3.83)

\[
C(t)^{(k\to\infty)} = \frac{e^4}{\sigma^4 \xi} \cosh \left( \frac{et}{\sigma} \right)^2
\]  

(3.84)

\[
D(t)^{(k\to\infty)} = \frac{e^3}{\sigma^2 \xi} \cosh \left( \frac{et}{\sigma} \right)^2 \operatorname{sech} \left( \frac{et}{\sigma} \right) + \frac{e^3 \gamma_2}{\sigma} \cosh \left( \frac{et}{\sigma} \right)
\]  

(3.85)

Plugging equations (3.82)-(3.85) into (3.45) and using the boundary conditions \( \eta_T = 0 \) and \( \eta'_t|_{t=0} = 0 \), we obtain that the second order ODE becomes non-singular and admits the following unique solution:

\[
\eta_t = \varrho \left[ (\gamma_1 \cosh(\omega t) + \gamma_2 \cosh(\omega T)) \operatorname{sech} \left( \frac{et}{\sigma} \right) - (\gamma_1 + \gamma_2) \cosh(\omega T) \operatorname{sech} \left( \frac{et}{\sigma} \right) \right]
\]  

(3.86)

where \( \varrho \) is as in Eq. (3.71) and

\[
\varrho = \frac{e \sigma \operatorname{sech}(\omega T)}{(\gamma_1 + \gamma_2) \left( e^2 - \omega^2 \sigma^2 \right) (1 + (\gamma_1 + \gamma_2) \xi \sigma^2)}
\]

From Eq. (3.65), we get \( a_t^\ME = a_t^\M \) after differentiating (3.86), and multiplying it by the expression \(- \coth \left( \frac{et}{\sigma} \right) \cosh \left( \frac{et}{\sigma} \right)\). Using the relation between \( a_t^\ME \) and \( a_t^\ME \) given by Eq. (3.11), we get that \( a_t^\ME = a_t^\M \), thereby proving the proposition.

We also find the following asymptotics. In case when the misreporting penalty \( \xi \) tends to zero, the principal gives full incentives to the agent to apply misreporting, and the effort level will also be maximal.

\[
\lim_{\xi \to 0} a_t^\ME = 1, \quad \lim_{\xi \to 0} a_t^\ME = \infty, \quad \lim_{\xi \to 0} u_t^\ME = \frac{e}{k}
\]  

(3.87)

The expressions for \( a_t^\ME \) and \( a_t^\ME \) correspond to the ones obtained in Capponi, Cvitanić and Yolcu (2011) in case of zero penalty for misreporting. If, instead, there is zero penalty for effort, we obtain

\[
\lim_{k \to 0} a_t^\ME = 1, \quad \lim_{k \to 0} a_t^\ME = \frac{e}{\sigma \xi} \left[ e \operatorname{sech} \left( \frac{et}{\sigma} \right) \sinh \left( \frac{et}{\sigma} \right) - (e - 1) \tanh \left( \frac{et}{\sigma} \right) \right]
\]

(3.88)

Here, the expressions for \( a_t^\ME \) and \( u_t^\ME \) agree with their counterparts in the Holmstrom-Milgrom model with zero penalty for effort. The proofs of Eq. (3.87) and (3.88) are provided in Appendix 6.1.

4. Comparative Statics

This section analyses the sensitivity of optimal \( a_t, a_t^\ME, u_t \) with respect to the various parameters of the model. Furthermore, we compare those quantities with their counterparts in the model where only misreporting or effort is present, see Capponi, Cvitanić and Yolcu (2011) and Holmstrom and Milgrom (1987) respectively. We first summarise the main findings and then separately analyse the sensitivities with respect to the different parameters. We find that, in case the agent can
apply both effort and misreporting, the pay-per-performance sensitivity is higher than in the case where he can only apply one of those actions. This is because the agent needs extra incentives from the principal to apply appropriate levels of both effort and misreporting. We find that, if the principal can induce the agent to apply both actions, then she will induce her to apply a higher level of effort, relative to what she would do in the Holmstrom-Milgrom model, in which only effort is present. Similarly, she will induce the agent to apply a higher level of misreporting relative to the contracting framework in which she can only induce her to misreport. An analysis of the joint dependence of effort and misreporting from the two penalty factors shows that the misreporting penalty mainly affects the misreporting action, while the effort action is mainly affected by the effort penalty.

We also find that PPS is not necessarily monotone in the noise intensity \( \sigma \), or in time \( t \). Moreover, a large part of PPS is aimed at inducing the agent to apply misreporting.

In all graphs below we fix the time horizon at \( T = 1 \), and we investigate the behaviour of the optimal \( a_n, u_t, \) and \( \alpha_t \) at time \( t = 0.5 \). In the legends, we use "ME" to indicate the proposed contracting framework with misreporting and effort. We use "M" to denote the contracting framework with only misreporting, and "HM" to denote the setting in which only effort is present.

From Eq. (3.12), we deduce that \( u_t \) is obtained scaling \( \alpha_t \) by a factor \( \frac{c}{k} \). For this reason, we do not report the graphs for effort.

4.1 Sensitivity Analysis with respect to time
We study the behaviour of optimal pay-per-performance sensitivity, optimal misreporting and optimal effort with respect to time. Figure 1 shows that the amount of misreporting increases over time, both for the proposed model and the model in which only misreporting is allowed. Interestingly, the PPS in the model in which both effort and misreporting are present is U-shaped, and contrasts with the one obtained in the contracting model with only misreporting, where it is always increasing. In both cases the optimal amount of misreporting is negligible initially, because the actual output value is assumed to be known at time zero, and thus there is no benefit from misreporting. Initially, the principal gives full incentives to the agent, aimed primarily at inducing him to apply effort, rather than misreporting. As time goes by, the principal wants the agent to apply a higher level of misreporting, while not as much effort is needed, and the pay-per-performance sensitivity is decreased. However, later on, it is optimal to have the agent apply increasingly higher effort. To this purpose, the principal increases again the pay-per-performance sensitivity to keep the levels of both misreporting and effort increasing.

Figure 1: The parameters common to all graphs are \( \sigma = 0.2, \epsilon = 1, \xi = 0.9, k = 1.2, \gamma_2 = 0.9, \gamma_1 = 0.3, \gamma_2 = 0.1 \). The left graph refers to the pay-per-performance sensitivity over time, while the right graph to the misreporting amount over time.
Moreover, we can see that the PPS in contracting models in which misreporting is allowed is always higher than the (constant) PPS obtained in the Holmstrom-Milgrom model. This shows that higher benefit is extracted from imperfect observations and misreporting, relative to the situation of full transparency.

4.2 Sensitivity Analysis with respect to \( \sigma \)
In this part, we study the behaviour of optimal pay-per-performance sensitivity, optimal misreporting and optimal effort with respect to \( \sigma \).

4.2.1 Optimal pay-per-performance sensitivity versus \( \sigma \)
Figure 2 shows that while the optimal pay-per-performance sensitivity is decreasing in \( \sigma \) in absence of effort, it may exhibit different patterns when effort is present. For a fixed misreporting level \( a_0 \), as \( \sigma \) increases the market value estimate \( \hat{x}^{u_0} \) gets closer to \( \hat{x}^{u_0} \), i.e., the estimate which would be produced by the market, were it aware of the exact misreporting level. Indeed, this can be easily checked from Eq. (2.7). This means that the misreporting action yields smaller benefits, and thus explains why in the model with only misreporting, PPS \( \alpha_t \) always decreases with \( \sigma \). However, even for large values of noise volatility \( \sigma \), principal could still benefit from agent's effort. This is because, for a given level of \( \sigma \), if the agent applies effort \( u \), then he would increase the drift of the market value estimate by \( \varepsilon u \). The precise benefit that this gives to the principal depends on the relation between misreporting and effort penalty. If the misreporting penalty is high and the effort penalty small, the agent is more resilient to applying misreporting and prefers to increase the output value by applying effort.

Figure 2: The parameters common to all graphs are \( \gamma_1 = 0.3, \gamma_2 = 0.1 \). Each graph plots \( \alpha_t \) versus \( \sigma \), for a specific triple \( (k, \xi, \varepsilon) \) of effort and misreporting penalty. Top left graph: \( (k, \xi, \varepsilon) = (0.2, 10, 0.1) \). Top right graph: \( (k, \xi, \varepsilon) = (10, 0.2, 0.1) \). Bottom left graph: \( (k, \xi, \varepsilon) = (0.4, 0.4, 0.1) \). Bottom right graph: \( (k, \xi, \varepsilon) = (10, 0.2, 0.4) \).

Additionally, the higher \( \sigma \), the higher the motivation for the principal in giving the agent more incentives to apply effort relative to misreporting, because, as mentioned above, the benefit of misreporting decreases with \( \sigma \). This parametric case is illustrated in the top left graph of Figure 2, where we can see that \( \alpha_t \) monotonically increases with \( \sigma \). On the other hand, when effort has significantly higher penalty than misreporting, the principal knows that it is costly for the agent to apply effort, and therefore she offers her smaller incentives to do so, unless \( \sigma \) is large and he
can only extract benefits from effort. A direct comparison of the top right and bottom left graph in Figure 2 shows that the threshold value $\sigma$, after which the principal increases again the agent incentives in misreporting, depends heavily on the effort penalty. If the latter is high enough, then the principal, knowing that the agent is less inclined to apply effort, increases $\alpha_t$ for larger values of $\sigma$, for which the effect of misreporting become less and less effective.

4.2.2 Optimal misreporting versus $\sigma$

Figure 3 shows that the principal always induces the agent to apply a higher misreporting action when effort is present than when it is not present. Moreover, when $\sigma$ gets larger, the applied misreporting action generally becomes smaller, because for high levels of noise volatility $\sigma$, the misreporting action does not affect much the market value estimate $\hat{X}_t^{u,0}$, as mentioned above. However, if the mean-variance trade-off $\varepsilon$ is large enough, and the penalty for misreporting is significantly smaller than the effort penalty, then the principal may have incentives to induce the agent to apply higher misreporting for larger values of $\sigma$, as long as $\sigma$ does not become too large. In other words, the relation between $\alpha_t$ and $\sigma$ may not be monotone, as seen in the case of the bottom right graph of Figure 3. We also remark that Figure 2 shows that $\alpha_t$ is not very sensitive to $\sigma$, while Figure 3 shows that $\alpha_t$ is quite sensitive to $\sigma$.

4.3 Sensitivity Analysis with respect to $\varepsilon$

Here, we investigate the behaviour of optimal pay-per-performance sensitivity, optimal misreporting and optimal effort with respect to $\varepsilon$.

4.3.1 Optimal pay-per-performance sensitivity versus $\varepsilon$

We see from Figure 4 that the pay-per-performance sensitivity $\alpha_t$ is decreasing in $\varepsilon$, regardless of whether the agent can apply both effort and misreporting or only misreporting. This contrasts with Figure 2, where the principal typically increases the incentives in the ME case when the noise volatility $\sigma$ is large enough. This may be explained observing that, for a contract consisting of cash and of performance related compensation depending linearly on $\hat{X}_t^{u,0}$, increasing $\varepsilon$ is equivalent to increasing portfolio holdings of the agent in the risky asset $\hat{X}_t^{u,0}$ relative to cash. This interpretation helps understanding why high $\alpha$ and high $\varepsilon$ have the same effect on the agent, and result in the agent having higher exposure to the risky part of the compensation. Therefore, if $\varepsilon$ goes up, then $\alpha_t$ should go down to compensate for that.

Figure 3: The parameters common to all graphs are $\gamma_1 = 0.3$, $\gamma_2 = 0.1$. Each graph plots the misreporting action $a_t$ versus $\sigma$, for a specific triple $(k, \xi, \varepsilon)$. Top left graph: $(k, \xi, \varepsilon)=(0.2,10,0.1)$. Top right graph: $(k, \xi, \varepsilon)=(10,0.2,0.1)$. Bottom left graph: $(k, \xi, \varepsilon)=(0.4,0.4,0.1)$. Bottom right graph: $(k, \xi, \varepsilon)=(10,0.2,0.4)$.
Figure 4: The parameters common to all graphs are $\gamma_1 = 0.3$, $\gamma_2 = 0.3$. Each graph plots $\alpha_t$ versus $\epsilon$, for a specific triple $(\kappa, \xi, \sigma)$. Top left graph: $(\kappa, \xi, \sigma) = (0.2, 10, 0.2)$. Top right graph: $(\kappa, \xi, \sigma) = (10, 0.2, 0.1)$. Bottom left graph: $(\kappa, \xi, \sigma) = (0.4, 0.4, 0.1)$. Bottom right graph: $(\kappa, \xi, \sigma) = (0.4, 0.4, 1)$.

Figure 5: The parameters common to all graphs are $\gamma_1 = 0.3$, $\gamma_2 = 0.3$. Each graph plots $\alpha_t$ versus $\epsilon$, for a specific triple $(\kappa, \xi, \sigma)$. Top left graph: $(\kappa, \xi, \sigma) = (0.2, 10, 0.2)$. Top right graph: $(\kappa, \xi, \sigma) = (10, 0.2, 0.1)$. Bottom left graph: $(\kappa, \xi, \sigma) = (0.4, 0.4, 0.1)$. Bottom right graph: $(\kappa, \xi, \sigma) = (0.4, 0.4, 1)$.

4.3.2 Optimal misreporting action versus $\epsilon$

Figure 5 shows that the principal does not necessarily induce the agent to apply more misreporting when $\epsilon$ increases. The dependence of $\alpha_t$ on $\epsilon$ is generally inverted U-shaped, except when $\sigma$ is very large as in the right bottom graph, where it becomes increasing.
4.3.3 Optimal effort versus $\epsilon$
From Figure 6 it is seen that the principal always induces the agent to apply more effort if $\epsilon$ increases. This is because larger $\epsilon$ results in larger benefit of increasing values of the drift $u_t \epsilon$.
Moreover, we see that when the agent can apply both effort and misreporting, he will always apply more effort than in the pure effort Holmstrom-Milgrom scenario.

Figure 6: The parameters used are $\gamma_1 = 0.3$, $\gamma_2 = 0.3$, $\sigma = 0.1$, $\xi = 0.4$. Left graph: $k = 0.4$. Right graph: $k = 10$.

4.4 Sensitivity Analysis with respect to effort and misreporting penalty
In this part, we analyse the dependence of optimal pay-per-performance sensitivity, optimal misreporting and optimal effort on the size of the penalties incurred when taking the actions.

4.4.1 Optimal pay-per-performance sensitivity versus $k$ and $\xi$
From the top graphs of Figure 7 we note the usual pattern, the optimal pay-per-performance sensitivity is always higher than in the cases in which only one of the two actions is allowed. Moreover, the two top graphs show that, in the ME model, PPS has similar sensitivity with respect to the increases in misreporting penalty (for a fixed effort penalty $k$), and to the increases in effort penalty (for a fixed misreporting penalty $\xi$). The graphs also indicate that the principal provides higher incentives in the model in which only misreporting is present, relative to the model in which only effort can be applied. This is perhaps due to the different manner in which $a_t$ and $u_t$ affect the drift of the market value estimate – effort contributes $\epsilon u_t d_t$, while $a_t$ contributes via a more complex term. It may also be specific to the parameters choice in the numerical examples.

The bottom graph, which analyses the joint dependence of $\alpha_t$ on misreporting and effort reinforces previous statements, showing that for any given level of effort (misreporting) penalty, the pay-per-performance sensitivity increases to one as the misreporting (effort) penalty decreases to zero. In other words, the whole firm is transferred to the agent if he is willing to apply an infinite amount of effort or misreporting.

4.4.2 Optimal misreporting action versus $k$ and $\xi$
Figure 8 shows that the level of misreporting decreases as the misreporting penalty $\xi$ gets larger. Moreover, the misreporting is mainly driven by the size of $\xi$, that is, for a given level of $\xi$, it exhibits negligible variations with respect to the effort penalty $k$. 

23
Figure 7: The parameters used for the top graphs are $\zeta = 0.2$, $\sigma = 0.1$, $\gamma_1 = 1$, $\gamma_2 = 0.3$. Top Left graph: $\alpha_t$ versus $\xi$ for a level of effort penalty $k = 0.2$. Top Right graph: $\alpha_t$ versus $k$ for a level of misreporting penalty $k = 0.2$. The bottom graph plots the pay-per-performance sensitivity $\alpha_p$, with respect to both $k$ and $\xi$. The parameters used for the bottom graph are $\zeta = 0.2$, $\sigma = 0.3$, $\gamma_1 = 0.8$, $\gamma_2 = 0.4$.

Figure 8: The left graph plots $\alpha_t$ versus $\xi$. The parameters used are $\zeta = 0.2$, $k = 0.2$, $\sigma = 0.3$, $\gamma_2 = 0.6$, and $T = 1$. The right graph plots the misreporting $\alpha$, jointly with respect to $k$ and $\xi$ under the parameter choice $\zeta = 0.2$, $\sigma = 0.3$, $\gamma_1 = 0.8$, $\gamma_2 = 0.4$.

4.4.3 Optimal effort versus $k$ and $\xi$

Figure 9 shows that the amount of effort decreases as the penalty $k$ for effort increases. Moreover, the effort choice is mainly driven by the size $k$, that is, for a given level $k$, it exhibits negligible variations with respect to $\xi$.

4.5 Sensitivity Analysis with respect to the level of risk aversions

In this part, we analyse the dependence of optimal pay-per-performance sensitivity, optimal misreporting and optimal effort on the level of risk aversions of the principal and of the agent.

4.5.1 Optimal pay-per-performance sensitivity versus $\gamma_1$ and $\gamma_2$

Figure 10 presents a comparison across all three contracting models. Not surprisingly, we see that the pay-per-performance sensitivity is decreasing in the risk aversion level $\gamma_1$ of the agent. This is because when the agent does not like risk the incentives are less efficient. Moreover, as already noted in graphs above, we see that the optimal compensation has higher pay-per-performance sensitivity in the model in which only misreporting can be applied than in the model in which
only effort can be applied. This means that a significant part of $\alpha^{ME}_t$ is aimed at providing incentives to the agent to apply misreporting. From a regulator point of view, this suggests that transferring more power to the shareholders may even be counterproductive, because they might be more concerned about inducing the manager to misreport rather than to apply effort. Potential remedies to this might be to impose misreporting penalties more severe than effort penalties, or impose frequent auditing to discourage misreporting.

Finally, we also notice that pay-per-performance sensitivity is increasing in the risk aversion level of the principal in all the cases. This is because with higher $\gamma_2$ the principal dislikes risk more, and wants to transfer it to the agent.

Figure 9: The left graph plots $u_t$ versus $k$. The parameters used are $\xi = 0.2$, $\epsilon = 0.2$, $\sigma = 0.3$, $\gamma_1 = 1$, $\gamma_2 = 0.6$. The right graph plots the effort $u_t$ jointly with respect to $k$ and $\xi$ under the parameter choice $\epsilon = 0.2$, $\sigma = 0.3$, $\gamma_1 = 0.8$, $\gamma_2 = 0.4$.

Figure 10: The parameters common to both graphs are $\epsilon = 1$, $\xi = 0.4$, $\sigma = 0.4$, $k = 0.4$, $T = 1$. The left graph plots $\alpha_t$ versus $\gamma_1$ for a given level of principal’s risk aversion $\gamma_2 = 0.3$. The right graph plots $\alpha_t$ versus $\gamma_2$ for a given level of agent risk aversion $\gamma_1 = 0.3$.

Figure 11: The parameters common to both graphs are $\epsilon = 1$, $\xi = 0.4$, $\sigma = 0.4$, $k = 0.4$, $T = 1$. The left graph plots $a_t$ versus $\gamma_1$ for a given level of principal’s risk aversion $\gamma_2 = 0.3$. The right graph plots $a_t$ versus $\gamma_2$ for a given level of agent risk aversion $\gamma_1 = 0.3$. 
4.5.2 Optimal misreporting action versus $\gamma_1$ and $\gamma_2$

We see from Figure 11 that the amount of misreporting decreases in $\gamma_1$. This is because a more risk averse agent gets a lower percentage of the firm's output, and is less inclined to misreport. Similarly, since everything is driven by the relative size of risk aversions, the amount of misreporting increases if the principal is more risk averse. More precisely, the higher the risk aversion of the principal, the higher percentage of the firm she gives to the agent, who then has higher incentives to misreport.

4.5.3 Optimal effort versus $\gamma_1$ and $\gamma_2$

The behaviour of the effort action follows a similar pattern as the misreporting action, and for the same reasons. More precisely, from Figure 11 we see that the amount of effort decreases with $\gamma_1$. Similarly, the higher the risk aversion of the principal, the higher the agent's effort. This is in agreement with all the standard contracting models a la Holmstrom-Milgrom.

5. Conclusions

This paper provides a general contracting framework in which there are two possible sources of moral hazard, namely effort and misreporting. The proposed framework has been shown to be a generalisation of the classical Holmstrom-Milgrom model of hidden effort, and of the contracting framework proposed by Capponi, Cvitanić and Yolcu (2011), in which only misreporting may be applied. The agent applies effort to affect the drift of the true (but unobserved) underlying outcome process, and applies misreporting to manipulate the drift of the observed accounting reports process. While the market observes the realisations of the effort process over time, it cannot correctly estimate the misreporting level introduced by the agent. Therefore, the market produces project estimates which do not factor in misreporting, but only account for effort, accounting noise, and uncertainty governing the underlying actual output process. We consider contracts which are linear functionals of the market value estimates of the output process. We formulate the contracting problem as a calculus of variation optimisation problem, and reduce the problem of finding the optimal pay-per-performance sensitivity, effort, and misreporting action to solving a singular second order differential equation.

Figure 12: The parameters common to both graphs are $\epsilon = 1$, $\xi = 0.4$, $\sigma = 0.4$, $k = 0.4$, $T = 1$. The left graph plots $u_t$ versus $\gamma_1$ for a given level of principal's risk aversion $\gamma_2 = 0.3$. The right graph plots the effort $u_t$ versus $\gamma_2$ for a given level of agent risk aversion $\gamma_1 = 0.3$.

We develop a thorough comparative statics analysis to analyse the interplay between effort and misreporting in our model, as well as the dependence of the actions and of the optimal contract on the various model parameters. We find that both the effort and the misreporting actions are time dependent. Moreover, it is always in the interest of the principal to have non-zero misreporting. Indeed, we find that the optimal compensation exhibits higher pay-per-performance sensitivity relative to a framework in which the agent is only compensated for effort, or only compensated for misreporting. The fact that shareholders pay a significant fraction of the compensation to induce the manager to misreport indicates that transferring more power to shareholders may not alleviate the concerns about corporate oversight. We leave for future research a construction...
of a model in which auditing, or other oversight measures would be incorporated and whose advantages and disadvantages would be studied.

6. Proofs of Lemmas and Propositions

6.1 Proofs of Comparative Statics

6.1.1 Limiting cases of pay-per-performance sensitivity

Proof of Eq. (3.87). Multiplying the ODE (and the boundary conditions) by $\xi$, and taking the limit as $\xi \to 0$, we obtain that the ordinary differential equation reduces to

$$A^{\xi \to 0}(t)\eta''_t + B^{\xi \to 0}(t)\eta'_t + C^{\xi \to 0}(t)\eta_t + D^{\xi \to 0}(t) = 0$$

(6.1)

with boundary conditions given by $\eta_T = 0$ and $\eta'_0 = 0$. Here

$$A^{\xi \to 0}(t) = \frac{\epsilon^2}{\sigma^2} \cosh^2 \left(\frac{\epsilon t}{\sigma}\right)$$

$$B^{\xi \to 0}(t) = 2\frac{\epsilon^3}{\sigma^3} \sinh \left(\frac{\epsilon t}{\sigma}\right) \cosh \left(\frac{\epsilon t}{\sigma}\right)$$

$$C^{\xi \to 0}(t) = \frac{\epsilon^4}{\sigma^4} \cosh^2 \left(\frac{\epsilon t}{\sigma}\right)$$

$$D^{\xi \to 0}(t) = \frac{\epsilon^3}{\sigma^3} \cosh^2 \left(\frac{\epsilon t}{\sigma}\right) \sech \left(\frac{\epsilon T}{\sigma}\right)$$

It can be checked that the solution of the ODE in Eq. (6.1) is given by

$$\eta^{\xi \to 0}_t = \frac{\sigma}{\epsilon} \left( \sech \left(\frac{\epsilon t}{\sigma}\right) - \sech \left(\frac{\epsilon T}{\sigma}\right) \right)$$

and, using Eq. (3.65), we obtain $\lim_{\xi \to 0} \alpha^{ME}_t = 1$. Using relations (3.67), and (3.66), we then obtain Eq. (3.87).

6.1.2 Limiting cases $k \to 0$

Multiplying the ODE (and the boundary conditions) by $k$, and taking the limit as $k \to 0$, we obtain that the ordinary differential equation reduces to

$$A^{k \to 0}(t)\eta''_t + B^{k \to 0}(t)\eta'_t + C^{k \to 0}(t)\eta_t + D^{k \to 0}(t) = 0$$

(6.2)

with boundary conditions given by $\eta_T = 0$ and $\eta'_0 = 0$. Here

$$A^{k \to 0}(t) = \epsilon^2 \cosh \left(\frac{\epsilon t}{\sigma}\right)^4 \sinh \left(\frac{\epsilon t}{\sigma}\right)$$

$$B^{k \to 0}(t) = 2\frac{\epsilon^3}{\sigma} \cosh \left(\frac{\epsilon t}{\sigma}\right)^3 \sinh \left(\frac{\epsilon t}{\sigma}\right) \sinh \left(\frac{\epsilon t}{\sigma}\right) - 2\frac{\epsilon^3}{\sigma} \cosh \left(\frac{\epsilon t}{\sigma}\right)^3$$

$$C^{k \to 0}(t) = 0$$

$$D^{k \to 0}(t) = \frac{\epsilon^3}{\sigma} \cosh \left(\frac{\epsilon t}{\sigma}\right) \sinh \left(\frac{\epsilon t}{\sigma}\right) \left( \sinh \left(\frac{\epsilon t}{\sigma}\right)^2 - 1 \right)$$

It can be easily checked that the solution $\eta^{k \to 0}_t$ is given by

$$\eta^{k \to 0}_t = \frac{\sigma}{\epsilon} \left( \sech \left(\frac{\epsilon t}{\sigma}\right) - \sech \left(\frac{\epsilon T}{\sigma}\right) \right)$$

and, using Eq. (3.65), we obtain $\lim_{\xi \to 0} \alpha^{ME}_t = 1$. Using relations (3.67), and (3.66), we then obtain Eq. (3.88).
References


• Ke, B. (2001): Why do CEOs of publicly traded firms prefer reporting small increases in earnings and long duration of consecutive earnings increases? Working paper, Penn State University.


EDHEC-Risk Institute is part of EDHEC Business School, one of Europe’s leading business schools and a member of the select group of academic institutions worldwide to have earned the triple crown of international accreditations (AACSB, EQUIS, Association of MBAs). Established in 2001, EDHEC-Risk Institute has become the premier European centre for applied financial research.

In partnership with large financial institutions, its team of 85 permanent professors, engineers and support staff implements six research programmes and ten research chairs focusing on asset allocation and risk management in the traditional and alternative investment universes. The results of the research programmes and chairs are disseminated through the three EDHEC-Risk Institute locations in London, Nice, and Singapore.

EDHEC-Risk Institute validates the academic quality of its output through publications in leading scholarly journals, implements a multifaceted communications policy to inform investors and asset managers on state-of-the-art concepts and techniques, and forms business partnerships to launch innovative products. Its executive education arm helps professionals to upgrade their skills with advanced risk and investment management seminars and degree courses, including the EDHEC-Risk Institute PhD in Finance.

Copyright © 2012 EDHEC-Risk Institute

For more information, please contact:
Carolyn Essid on +33 493 187 824
or by e-mail to: carolyn.essid@edhec-risk.com

EDHEC-Risk Institute
393-400 promenade des Anglais
BP 3116
06202 Nice Cedex 3 - France

EDHEC Risk Institute—Europe
10 Fleet Place - Ludgate
London EC4M 7RB - United Kingdom

EDHEC Risk Institute—Asia
1 George Street - #07-02
Singapore 049145

www.edhec-risk.com