Math Finan Econ (2007) 1:21–55
DOI 10.1007/s11579-007-0002-2

Optimal compensation with adverse selection and dynamic actions

Jakša Cvitanić · Jianfeng Zhang

Received: 1 August 2006 / Accepted: 24 January 2007 /
Published online: 16 March 2007

Abstract We consider continuous-time models in which the agent is paid at the end of the time horizon by the principal, who does not know the agent's type. The agent dynamically affects either the drift of the underlying output process, or its volatility. The principal's problem reduces to a calculus of variation problem for the agent's level of utility. The optimal ratio of marginal utilities is random, via dependence on the underlying output process. When the agent affects the drift only, in the risk-neutral case lower volatility corresponds to the more incentive optimal contract for the smaller range of agents who get rent above the reservation utility. If only the volatility is affected, the optimal contract is necessarily non-incentive, unlike in the first-best case. We also suggest a procedure for finding simple and reasonable contracts, which, however, are not necessarily optimal.

Keywords Adverse selection · Moral hazard · Principal-agent problems · Continuous-time models · Contracts · Managers compensation

---

Research supported in part by NSF grants DMS 04-03575 and 06-31298. We would like to express our gratitude to participants of the following seminars and conferences for useful comments and suggestions: UCLA (Econ Theory), Caltech (Econ Theory), Columbia (Probability), Princeton (Fin. Engineering), U. Texas at Austin (Math Finance), Banff Workshop on Optim. Problems in Fin. Econ, Kyoto U. (Economics), UC Irvine (Probability), Cornell (Fin. Engineering), Bachelier Seminar. Moreover, we are very grateful to the anonymous referee for helpful suggestions. The remaining errors are the authors' sole responsibility.

J. Cvitanić (✉)
Caltech, Humanities and Social Sciences, M/C 228-77, 1200 E. California Blvd.
Pasadena, CA 91125, USA
e-mail: cvitanic@hss.caltech.edu

J. Zhang
USC Department of Mathematics, 3620 S Vermont Ave, KAP 108,
Los Angeles, CA 90089-1113, USA
e-mail: jianfenz@usc.edu

JEL Classification C61, C73, D82, J33, M52

Mathematics Subject Classification (2000) 91B28 · 93E20

1 Introduction

We propose new continuous-time models for modeling a principal-agent relationship in the presence of adverse selection (hidden agent's type), with or without moral hazard (hidden actions). The main applications we have in mind are the compensation of executives and the compensation of portfolio managers. For executive compensation it may be satisfactory to have a model in which the agent (executive) can control the drift (return) of the underlying process (value of the firm or its stock), but the volatility is fixed. However, for portfolio management it is important to have models in which the volatility (determined by the portfolio strategy) can be affected by agent's actions. Moreover, it is important to allow for these actions (portfolio choice) to be dynamic. We consider such models in the presence of adverse selection. More precisely, the agent's type is unobservable by the principal and is represented by a parameter corresponding to the expected return of the underlying output process, when actions are fixed (at zero effort and unit volatility).

The continuous-time principal-agent literature started with the seminal paper of Holmstrom and Milgrom [17]. In that paper the agent controls only the drift, there is moral hazard but not adverse selection, the utility functions are exponential and the optimal contract is linear. That work was generalized by Schattler and Sung [28,29], Sung [30,31] and Detemple et al. [12]. See also [13,15,18,24,25]. Discrete-time adverse selection papers with applications include [3,4,9,14,19,20,23]. Articles Williams [33] and Cvitanić et al. [7] use the stochastic maximum principle and Forward–Backward Stochastic Differential Equations to characterize the optimal compensation for more general utility functions, under moral hazard. A paper in the similar spirit, but on an infinite horizon, is Sannikov [27]. Sannikov [27] and Williams [33] focus on the contracts represented as a payment at a continuous rate to the agent, as opposed to a bulk payment at the end of the time horizon, the case considered in the present paper. See also [2,11] for models of a different type. Cadenillas et al. [6], Cvitanić et al. [8] and Ou-Yang [26] consider the case when the volatility is controlled, with applications to portfolio management, but there is no loss of efficiency due to moral hazard or adverse selection.

A recent paper in continuous time that has both adverse selection and moral hazard is Sung [32]. It also contains numerous examples and references that motivate having a risk-averse agent, and being able to control both the drift and the volatility. The paper considers a risk-neutral principal, and an agent with exponential utility. Moreover, it is assumed that the principal observes only the initial and the final value of the underlying process. The optimal agent's actions in the model are constant through time and the optimal contract is again linear.

We are able to study a framework with general utility functions and with dynamic actions, in which the principal observes the underlying output process continuously and hence also observes the volatility. On the flip-side, we only consider a cost function which is quadratic in the agent's effort (the drift control) and there is no cost on the choice of volatility. If the agent only controls the drift while the volatility is fixed, we reduce the principal's problem to a deterministic calculus of variations problem of choosing the appropriate level of agent's utility. This is similar to the classical static model literature, see for example the excellent book [5]. In the static model, under the so-called single-crossing property and quasi-linear utility, the problem simplifies even further, to a calculus of variations problem, but not over the agent's level of utility, rather, over the level of compensation. In our model, we provide a more general formulation of the calculus of variations problem, for general utilities.

The generality in which we work is possible because with quadratic cost function the agent's optimal utility and the principal's problem can both be represented in a simple form which involves explicitly the contracted payoff only, and not the agent's effort process. The ratio of the marginal utilities of the principal and the agent, which is constant in the Pareto optimal first-best case of full information, is now random. The optimal contract's value at the payoff time depends also on the path of the output process, not just on its final value, unless the volatility is constant. In the case of a risk-neutral principal and agent, we solve the problem explicitly: the optimal contract is linear; there is a range of lower type agents which get no informational rent above the reservation utility; as the volatility decreases, that range gets wider, the contract becomes more incentive (sensitive), while the informational rent for higher type agents gets lower.

If only the volatility is controlled, as may be the case in delegated portfolio management, the optimal contract is a random variable which depends on the value of the underlying risky investment asset, or, equivalently, on the volatility weighted average of the output. In the first-best case, there is an optimal contract which is of benchmark type (the output value minus the benchmark value) and which is incentive in the sense that the agent implements the first-best volatility at her optimum. With adverse selection where the expected return of the portfolio manager is not known to the principal, the optimal contract is non-incentive: it is random (as it depends on the value of the underlying noise process), but independent of the manager's actions and the manager has to be told by the principal how to choose the portfolio strategy.

With adverse selection, there is a so-called "revelation principle" which says that it is sufficient to consider contracts which are "truth-telling": the principal offers a menu of contracts, one for each type (of agent), and with a truth-telling contract the agent of a certain type will choose the contract corresponding to that type. This truth-telling requirement imposes a constraint on the admissible contracts. We need to stress that our approach is the so-called "first order approach" with respect to that constraint, in the sense that we look for contracts which satisfy the first-order (first derivative equal to zero) necessary condition for this constraint. In general, it is very hard to identify under which conditions this procedure is also sufficient for producing an optimal contract. Instead, we propose a simpler way for finding reasonable contracts which are not necessarily optimal. We do this by restricting the form of the involved Lagrange multipliers in such a way that the first-order necessary condition also becomes a sufficient condition for truth-telling.

The paper is organized as follows: in Sect. 2 we consider the fixed volatility case with the agent controlling the expected return rate. We solve the risk-neutral example in Sect. 3, and consider simpler, but non-optimal contracts for risk-averse agents in Sect. 4. Section 5 deals with the control of volatility. We conclude in Sect. 6.

2 Model I: controlling the return

2.1 Weak formulation of the model

I. The state process X We take the model from Cvitanić et al. [7], henceforth CWZ [7], discussed in that paper in the context of moral hazard, without adverse selection. Let B be a standard Brownian motion under some probability space with probability measure P, and $\mathbf{F}^B = \{\mathcal{F}_t\}_{0\le t\le T}$ be the filtration generated by B up to time $T > 0$. For any $\mathbf{F}^B$-adapted process $v > 0$ such that $E\int_0^T v_t^2 dt < \infty$, let

$$X_t := x + \int_0^t v_s dB_s. \tag{2.1}$$

Note that $\mathbf{F}^X = \mathbf{F}^B$. Now for any $\mathbf{F}^B$-adapted process u, let

$$B_t^u := B_t - \int_0^t u_s ds; \quad M_t^u := \exp\left(\int_0^t u_s dB_s - \frac{1}{2}\int_0^t |u_s|^2 ds\right); \quad \frac{dP^u}{dP} := M_T^u. \tag{2.2}$$

We assume here that u satisfies the conditions required by the Girsanov Theorem (e.g., Novikov condition). Then M_t^u is a martingale and P^u is a probability measure. Moreover, B^u is a P^u-Brownian motion and

$$dX_t = v_t dB_t = u_t v_t dt + v_t dB_t^u. \tag{2.3}$$

This is a standard continuous-time "weak" formulation for principal-agent problems with moral hazard, used in [28], while used in Stochastic Control Theory at least since [10].

II. The agent's problem We consider a principal who wants to hire an agent of an unknown type $\theta \in [\theta_L, \theta_H]$, where θ_L, θ_H are known to the principal. The principal offers a menu of contract payoffs $C_T(\theta)$, and an agent θ can choose arbitrary payoff $C_T(\tilde{\theta})$, where $\tilde{\theta}$ may or may not be equal to her real type θ. We assume that the agent's problem is

$$R(\theta) := \sup_{\tilde{\theta}\in[\theta_L,\theta_H]} V(\theta,\tilde{\theta}) := \sup_{\tilde{\theta}\in[\theta_L,\theta_H]} \sup_{u\in\mathcal{A}_0} E^u[U_1(C_T(\tilde{\theta})) - G_T(\theta)], \tag{2.4}$$

where U_1 is the agent's utility function; $G_T(\theta)$ is the cost variable; E^u is the expectation under P^u; and $\mathcal{A}_0$ is the admissible set for the agent's effort u, which will be defined later in Definition 2.1.

One important assumption of this paper is that the cost G is quadratic in u. In particular, we assume

$$G_T(\theta) := \int_0^T g(u_t)dt + \xi := \frac{1}{2}\int_0^T (u_t - \theta)^2 dt + \xi, \tag{2.5}$$

where ξ is a given $\mathcal{F}_T$-measurable random variable. For example, we can take $\xi = H(X_T) + \int_0^T h(X_t)dt$ for some functions H, h.

This is equivalent to the model

$$dX_t = (u_t + \theta)v_t dt + v_t dB_t^u\,, \quad G_T = \frac{1}{2}\int_0^T u_t^2 dt + \xi. \tag{2.6}$$

The interpretation for the latter model is that θ is the return that the agent can achieve with zero effort, and can thus be interpreted as the quality of the agent. If we think of the application to the delegated portfolio management, then we can interpret the process v as related to the portfolio strategy chosen by the manager, which, given the assets in which to invest, is known and observable. In other words, the volatility process of the portfolio is fixed, and given this process, the manager can affect the mean return through her effort, for example by carefully choosing the assets in which to invest. The assumption that v is fixed can be justified by the fact that X is observed continuously, and then v is also observed as its quadratic variation process, and thus the principal can tell the agent which v to use. For example, if the principal was risk-neutral, he would tell the manager to choose the highest possible v. On the other hand, in a later section, we consider a model in which the volatility (portfolio strategy) v is not given, but it is the action to be chosen.

Remark 2.1 The cost function $G(u) = (u-\theta)^2/2$ is not monotone in effort. It is shown in CWZ [7], in the hidden action case with no adverse selection and with $\theta = 0$, that we can restrict the analysis to the case $u > 0$ without loss of generality. The argument is based on the fact that we can replace $u_t < 0$ by $u_t^* = -u_t > 0$, at the same cost to the agent, but resulting in higher output X. However, here, because of adverse selection, it is conceivable that the principal may allow the agent to apply seemingly inefficient effort in exchange for telling the truth about her type, and it is not clear that we can restrict ourselves to the case $u \geq \theta$.[1] One possible interpretation then, in the context of model (2.6), is that θ is the return the agent attains with no effort ($u = 0$), and applying negative effort $-u$ is as costly as applying positive effort u. That is to say, if the agent/manager wants to harm the principal/company and produce negative returns, that is not costless, and requires more expensive effort for more negative returns. Similar interpretation can be formulated for the case of model (2.3), in which case θ is the level of zero effort, and moving away from θ requires increasingly costly effort. Let us mention, however, that in the risk-neutral example, the only example we are able to solve explicitly, the optimal effort is, in fact, no less than θ, thus in the domain of the cost function where this function is increasing.

III. Constraints on the contract C_T First, we assume that the participation, or individual rationality (IR) constraint of the agent is

$$R(\theta) \geq r(\theta) \tag{2.7}$$

where $r(\theta)$ is a given function representing the reservation utility of the type θ agent. In other words, the agent θ will not work for the principal unless she can attain expected utility of at least $r(\theta)$. For example, it might be natural that $r(\theta)$ is increasing in θ, so that higher type agents require higher minimal utility. The principal offers a menu of contracts $C_T(\theta)$. Although he does not observe the type θ, he knows the function $r(\theta)$, that is, how much the agent of type θ needs to be minimally paid.

[1] We thank the referee for pointing this out.

Secondly, by standard revelation principle of the principal-agent theory, we restrict ourselves to the truth-telling contracts, that is, to such contracts for which the agent θ will choose optimally the contract $C_T(\theta)$. In other words, we have

$$R(\theta) = V(\theta,\theta), \quad \forall\theta. \tag{2.8}$$

Thirdly, we consider only implementable contracts. That is, for any θ, there exists unique optimal effort of the agent, denoted $\hat{u}(\theta) \in \mathcal{A}_0$, such that

$$R(\theta) = E^{\hat{u}(\theta)}[U_1(C_T(\theta)) - G_T(\theta)].$$

IV. The principal's problem Since θ is unobserved by the principal, we assume the principal has a prior distribution F for θ, on the interval $[\theta_L, \theta_H]$. Then the principal's optimization problem is defined as

$$\sup_{C_T \in \mathcal{A}} \int_{\theta_L}^{\theta_H} E^{\hat{u}(\theta)}[U_2(X_T - C_T(\theta))]dF(\theta), \tag{2.9}$$

where U_2 is the principal's utility function and $\mathcal{A}$ is the admissible set for contract C_T, which will be defined later, in Definition 2.3.

V. Standing assumptions First we adopt the standard assumptions for utility functions.

Assumption 2.1 *U_1, U_2 are twice differentiable such that $U_i' > 0, U_i'' \le 0, i = 1,2$.*

Throughout the paper, Assumption 2.1 will always be in force.

We now specify the technical conditions u and C_T should satisfy. Roughly speaking, we need enough integrability so that the calculations in the remainder of the paper can go through. We note that we do not aim to find the minimum set of sufficient conditions.

Definition 2.1 The set $\mathcal{A}_0$ of admissible effort processes u is the space of $\mathbf{F}^B$-adapted processes u such that

(i) $P(\int_0^T |u_t|^2 dt < \infty) = 1$;
(ii) $E\{|M_T^u|^4\} < \infty$.

We now show that for any $u \in \mathcal{A}_0$, we have the Novikov condition

$$E\Big\{e^{2\int_0^T |u_t|^2 dt}\Big\} < \infty; \tag{2.10}$$

and thus Girsanov Theorem holds for u. In fact, denote

$$\tau_n := \inf\left\{t : \int_0^t |u_s|^2 ds + |\int_0^t u_s dB_s| > n\right\}.$$

Then $\tau_n \uparrow T$. Moreover,

$$e^{\int_0^{\tau_n} u_t dB_t} = M_{\tau_n}^u e^{\frac{1}{2}\int_0^{\tau_n} |u_t|^2 dt}.$$

Squaring both sides and taking the expectation, we get

$$E\Big\{e^{2\int_0^{\tau_n} |u_t|^2 dt}\Big\} = E\Big\{|M_{\tau_n}^u|^2 e^{\int_0^{\tau_n} |u_t|^2 dt}\Big\} \le [E\{|M_{\tau_n}^u|^4\}]^{\frac{1}{2}}\Big[E\Big\{e^{2\int_0^{\tau_n} |u_t|^2 dt}\Big\}\Big]^{\frac{1}{2}}.$$

Thus

$$E\left\{e^{2\int_0^{\tau_n}|u_t|^2dt}\right\} \le E\{|M^u_{\tau_n}|^4\} \le E\{|M^u_T|^4\} < \infty.$$

Letting $n \to \infty$ we get (2.10).

The admissible set for C_T is more complicated. For any $\theta \in [\theta_L, \theta_H]$, let $B^\theta, M^\theta, P^\theta$ be defined by (2.2) with $u_t = \theta$; and denote

$$\bar{U}_1(C) := U_1(C) - \xi.$$

Definition 2.2 The set $\mathcal{A}_1$ consists of contracts C_T which satisfy:

(i) For any $\theta \in [\theta_L, \theta_H]$, $C_T(\theta)$ is $\mathcal{F}_T$-measurable.

(ii) $E\{|\bar{U}_1(C_T(\theta))|^4 + e^{5\bar{U}_1(C_T(\theta))}\} < \infty, \forall \theta \in [\theta_L, \theta_H]$.

(iii) For dF-a.s. θ, $C_T(\theta)$ is differentiable in θ and $\{e^{\bar{U}_1(C_T(\tilde{\theta}))}U_1'(C_T(\tilde{\theta}))|\partial_\theta C_T(\tilde{\theta})|\}$ is uniformly integrable under P^θ, uniformly in $\tilde{\theta}$.

(iv) $\sup_{\theta\in[\theta_L,\theta_H]} E^\theta\left\{e^{\bar{U}_1(C_T(\theta))}|U_2(X_T - C_T(\theta))|\right\} < \infty.$

Definition 2.3 The admissible set $\mathcal{A}$ of contracts C_T is the subset of $\mathcal{A}_1$ consisting of those contracts C_T which satisfy the IR constraint and the revelation principle.

We note that, as a direct consequence of Theorem 2.1 below, any $C_T \in \mathcal{A}$ is implementable. We also note that we do not impose any conditions on v and ξ here. However, depending on their properties, the set $\mathcal{A}$ may be small or even empty. So, in order to have a reasonably large set $\mathcal{A}$, we need v and ξ to have reasonably nice properties (e.g. v and ξ are bounded). We henceforth assume

$$\mathcal{A} \neq \phi.$$

2.2 Optimal solutions

I. The agent's optimal effort For fixed known θ (more precisely, for $\theta = 0$), the agent's problem is solved in CWZ [7]. We extend the result to our framework next.

Lemma 2.1 *Assume C_T satisfies* (i) *and* (ii) *of Definition* 2.2. *For any $\theta, \tilde{\theta} \in [\theta_L, \theta_H]$, the optimal effort $\hat{u}(\theta, \tilde{\theta}) \in \mathcal{A}_0$ of the agent of type θ, faced with the contract $C(\tilde{\theta})$, is obtained by solving the Backward Stochastic Differential Equation*

$$Y_t^{\theta,\tilde{\theta}} = e^{\bar{U}_1(C_T(\tilde{\theta}))} - \int_t^T (\hat{u}_s(\theta,\tilde{\theta}) - \theta) Y_s^{\theta,\tilde{\theta}} dB_s^\theta; \tag{2.11}$$

and

$$V(\theta,\tilde{\theta}) = \log E[M_T^\theta e^{\bar{U}_1(C_T(\tilde{\theta}))}]. \tag{2.12}$$

As a direct consequence, we have

Theorem 2.1 *If $C_T \in \mathcal{A}$, then the optimal effort $\hat{u}(\theta) \in \mathcal{A}_0$ for the agent is obtained by solving the Backward Stochastic Differential Equation*

$$Y_t^\theta = e^{\bar{U}_1(C_T(\theta))} - \int_t^T (\hat{u}_s(\theta) - \theta) Y_s^\theta dB_s^\theta; \tag{2.13}$$

and the agent's optimal expected utility is given by

$$R(\theta) = \log E[M_T^\theta e^{\bar{U}_1(C_T(\theta))}] = \log Y_0^\theta. \tag{2.14}$$

Remark 2.2 (i) Notice that finding optimal $\hat{u}$, in the language of option pricing theory, is mathematically equivalent to finding a replicating portfolio for the option with payoff $e^{\bar{U}_1(C_T)}$. Since that is a well studied problem, there are many ways to compute the solution (numerically, if not analytically).

(ii) A result of this type is available in CWZ [7] for other convex cost functions g, too. However, with the quadratic cost in u as here, we see that it is possible to represent the agent's utility value R in terms of the contract C_T, without dependence on u, as in (2.14). This, together with (2.19) below, will enable to represent the principal's problem in terms of C_T and R only. See also Remark 2.3 below.

Proof of Lemma 2.1 We expand here on the proof from CWZ [7]. First we show that (2.11) is well-posed and that $\hat{u}(\theta, \tilde{\theta}) \in \mathcal{A}_0$. In fact, by Definition 2.2(i) and (ii), we can solve the following linear BSDE

$$Y_t^{\theta,\tilde{\theta}} = e^{\bar{U}_1(C_T(\tilde{\theta}))} - \int_t^T Z_s^{\theta,\tilde{\theta}} dB_s^\theta.$$

Define

$$\hat{u} := \theta + \frac{Z_t^{\theta,\tilde{\theta}}}{Y_t^{\theta,\tilde{\theta}}}.$$

Then $\hat{u}(\theta, \tilde{\theta}) := \hat{u}$ satisfies (2.11). Since $Y_t^{\theta,\tilde{\theta}} > 0$ is continuous, $E^\theta\{\int_0^T |Z_t^{\theta,\tilde{\theta}}|^2 dt\} < \infty$, and P and P^θ are equivalent; we know $\hat{u}$ satisfies Definition 2.1(i). Moreover, by a straightforward calculation we have

$$e^{\bar{U}_1(C_T(\tilde{\theta}))} = Y_0^{\theta,\tilde{\theta}} e^{\int_0^T (\hat{u}_t - \theta) dB_t^\theta - \frac{1}{2}\int_0^T |\hat{u}_t - \theta|^2 dt} = Y_0^{\theta,\tilde{\theta}} M_T^{\hat{u}} [M_T^\theta]^{-1}.$$

Then

$$M_T^{\hat{u}} = [Y_0^{\theta,\tilde{\theta}}]^{-1} M_T^\theta e^{\bar{U}_1(C_T(\tilde{\theta}))}. \tag{2.15}$$

Thus

$$\begin{aligned} E\{|M_T^{\hat{u}}|^4\} &= [Y_0^{\theta,\tilde{\theta}}]^{-4} E\Big\{[M_T^\theta]^4 e^{4\bar{U}_1(C_T(\tilde{\theta}))}\Big\} \\ &\le CE\{[M_T^\theta]^{20}\} + CE\{e^{5\bar{U}_1(C_T(\tilde{\theta}))}\Big\} < \infty. \end{aligned}$$

Therefore, $\hat{u} \in \mathcal{A}_0$.

Now for any $u \in \mathcal{A}_0$, as is standard in this type of stochastic control problems, and also standard in multi-period principal-agent models, discrete or continuous, we consider the remaining utility of the agent at time t

$$Y_t^{A,u} = E_t^u \left[\bar{U}_1(C_T(\tilde{\theta})) - \frac{1}{2}\int_t^T |u_s - \theta|^2 ds \right].$$

By Definition 2.1 and (2.10), one can easily show that $Y_t^{A,u} - \frac{1}{2}\int_0^t |u_s - \theta|^2 ds$ is a square integrable P^u-martingale. We note that in general $\mathbf{F}^{B^u}$ is not the same as $\mathbf{F}^B$, so one cannot apply directly the standard martingale representation theorem. Nevertheless, one can show (see, e.g. [7]) that there exists an $\mathbf{F}^B$-adapted process $Z^{A,u}$ such that

$$Y_t^{A,u} - \frac{1}{2}\int_0^t |u_s - \theta|^2 ds = \bar{U}_1(C_T(\tilde{\theta})) - \frac{1}{2}\int_0^T |u_s - \theta|^2 ds - \int_t^T Z_s^{A,u} dB_s^u.$$

Then, switching from B^u to B^θ, we have

$$Y_t^{A,u} = \bar{U}_1(C_T(\tilde{\theta})) + \int_t^T [(u_s - \theta)Z_s^{A,u} - \frac{1}{2}|u_s - \theta|^2] ds - \int_t^T Z_s^{A,u} dB_s^\theta. \tag{2.16}$$

Note that $Y_0^{A,u} = E^u[\bar{U}_1(C_T) - \frac{1}{2}\int_0^T |u_s - \theta|^2 ds]$ is the agent's utility, given action u.

On the other hand, using Itô's rule and (2.11), we get

$$\log Y_t^{\theta,\tilde{\theta}} = \bar{U}_1(C_T(\tilde{\theta})) + \frac{1}{2}\int_t^T (\hat{u}_s - \theta)^2 ds - \int_t^T (\hat{u}_s - \theta) dB_s^\theta.$$

Thus, $\log Y_t^{\theta,\tilde{\theta}} = Y_t^{A,\hat{u}}$ is the agent's utility if she chooses action $\hat{u}$. Then we obtain

$$\begin{aligned} Y_0^{A,\hat{u}} - Y_0^{A,u} &= \int_0^T \Big[\frac{1}{2}[|\hat{u}_t - \theta|^2 + |u_t - \theta|^2] - (u_t - \theta)Z_t^{A,u}\Big] dt + \int_0^T [Z_t^{A,u} - (\hat{u}_t - \theta)] dB_t^\theta \\ &\geq \int_0^T \Big[(\hat{u}_t - \theta)(u_t - \theta) - (u_t - \theta)Z_t^{A,u}\Big] dt + \int_0^T [Z_t^{A,u} - (\hat{u}_t - \theta)] dB_t^\theta \\ &= \int_0^T [Z_t^{A,u} - (\hat{u}_t - \theta)] dB_t^u. \end{aligned} \tag{2.17}$$

The equality holds if and only if $u = \hat{u}$. Note that $E^u\{\int_0^T |Z_t^{A,u}|^2 dt\} < \infty$, and

$$E^u\left\{\int_0^T |\hat{u}_t|^2 dt\right\} = E\left\{M_T^u \int_0^T |\hat{u}_t|^2 dt\right\} \leq CE\Big\{|M_T^u|^2 + e^{2\int_0^T |\hat{u}_t|^2 dt}\Big\} < \infty,$$

where the last inequality is due to (2.10). Then

$$E^u\left\{\int_0^T [Z_t^{A,u} - (\hat{u}_t - \theta)]^2 dt\right\} < \infty.$$

Taking expected values under P^u in (2.17) we get $Y_0^{A,\hat{u}} \geq Y_0^{A,u}$, with equality if and only if $u = \hat{u}$. □

II. The relaxed principal's problem We now turn to the principal's problem (2.9). For $C_T \in \mathcal{A}$, the first order condition for the truth-telling constraint (2.8) is

$$E\Big\{M_T^\theta e^{\bar{U}_1(C_T(\theta))} U_1'(C_T(\theta)) \partial_\theta C_T(\theta)\Big\} = 0. \tag{2.18}$$

We now apply the standard, first-order approach of the principal-agent theory. That is, we solve the principal's problem by replacing the truth-telling constraint by its first order condition (2.18). Then, once the solution is found, it has to be checked whether it does satisfy the truth-telling constraint.

To solve the problem, we make a transformation. Recalling (2.15) and (2.14) and setting $\tilde{\theta} = \theta$, we have the following crucial observation

$$M_T^{\hat{u}(\theta)} = e^{-R(\theta)} M_T^{\theta} e^{\bar{U}_1(C_T(\theta))} \; . \tag{2.19}$$

Then we can rewrite the principal's problem as

$$\sup_{C_T} \int_{\theta_L}^{\theta^H} e^{-R(\theta)} E\left[M_T^{\theta} e^{\bar{U}_1(C_T(\theta))} U_2\left(X_T - C_T(\theta)\right)\right] dF(\theta).$$

Moreover, differentiating (2.14) with respect to θ, we get

$$E\Big\{M_T^{\theta} e^{\bar{U}_1(C_T(\theta))}\Big[[B_T - \theta T] + U_1'(C_T(\theta))\partial_\theta C_T(\theta)\Big]\Big\} = e^{R(\theta)} R'(\theta),$$

which, by (2.18), implies that

$$E\Big\{B_T M_T^{\theta} e^{\bar{U}_1(C_T(\theta))}\Big\} = e^{R(\theta)}[R'(\theta) + T\theta]. \tag{2.20}$$

Thus, the new, relaxed principal's problem is given by the following

Definition 2.4 The relaxed principal's problem is

$$\sup_{R} \sup_{C_T \in \mathcal{A}_1} \int_{\theta_L}^{\theta^H} e^{-R(\theta)} E\left[M_T^{\theta} e^{\bar{U}_1(C_T(\theta))} U_2\left(X_T - C_T(\theta)\right)\right] dF(\theta) \tag{2.21}$$

under the constraints

$$R(\theta) \geq r(\theta) \; , \quad E[M_T^{\theta} e^{\bar{U}_1(C_T(\theta))}] = e^{R(\theta)} \; , \quad E\Big\{B_T M_T^{\theta} e^{\bar{U}_1(C_T(\theta))}\Big\} = e^{R(\theta)}[R'(\theta) + T\theta]. \tag{2.22}$$

Remark 2.3 Our approach is based on the fact (2.19) for the agent's optimal choice of u. Thus, the choice of the probability measure corresponding to action $\hat{u}$ is completely determined by the choice of $e^{\bar{U}_1(C_T(\theta))}$ and by the choice of utility level $R(\theta)$ the principal is willing to offer to the agent. Therefore, the principal's objective becomes

$$e^{-R(\theta)} E\left[M_T^{\theta} e^{\bar{U}_1(C_T(\theta))} U_2\left(X_T - C_T(\theta)\right)\right]$$

which does not involve the agent's choice of u. Similarly, the IR constraint and the first order condition for the truth-telling constraint are also explicit in terms of $R(\theta)$, $R'(\theta)$ and expected values involving $C_T(\theta)$. The explicit connection between the agent's choice of the probability measure and the given contract, such as the connection (2.19), does not seem available for cost functions other than quadratic.

Remark 2.4 In the classical, single-period adverse selection problem with a continuum of types, but no moral hazard, one also has to solve a calculus of variations problem over the agents (indirect) utility. However, the problem typically reduces to a calculus of variation problem over the payment $C_T(\theta)$. Under the so-called Spence–Mirrlees condition on the agent's utility function and with a risk-neutral principal, a contract $C_T(\theta)$ is truth-telling if and only if it is a non-decreasing function of θ and the first-order truth-telling condition is satisfied. In our method, where we also have moral hazard and risk-averse principal, the calculus of variation problem cannot be reduced to the problem over $C_T(\theta)$, but remains to be a problem over the agent's utility $R(\theta)$. We are able to give a general formulation for the problem. However, unfortunately, for a general utility function U_1 of the agent, we have not been able to formulate a condition on U_1 under which we could find a necessary and sufficient conditions on $R(\theta)$ to induce truth-telling. Later below, we are able to show that the first order approach works for linear U_1 and U_2, when the hazard rate of θ is increasing, in agreement with the classical theory. For other utility functions, we suggest a way of finding reasonable contracts which are not necessarily optimal.

III. Optimal contracts for the relaxed principal's problem We proceed by fixing agent's utility $R(\theta)$, and finding first order conditions for the optimal contract $C_T(\theta)$. Introduce the Lagrange multipliers $\lambda(\theta), \mu(\theta)$ for the second and third constraint in (2.22). Denote $J_1 := U_1^{-1}$ and define a random function D:

$$D(y) := e^{U_1(y)}\Big[U_2(X_T - y) - \lambda(\theta) - \mu(\theta)B_T\Big].$$

Then, the Lagrangian is

$$E\left[\int_{\theta_L}^{\theta_H} M_T^{\theta} e^{-R(\theta)-\xi} D(C_T(\theta))dF(\theta)\right]. \tag{2.23}$$

Note that,

$$D'(y) = e^{U_1(y)}U_1'(y)\Big[G(X_T, y) - \lambda(\theta) - \mu(\theta)B_T\Big],$$

where

$$G(x, y) := U_2(x - y) - \frac{U_2'(x - y)}{U_1'(y)}.$$

The first order condition is

$$G(X_T, C_T(\theta)) = \lambda(\theta) + \mu(\theta)B_T. \tag{2.24}$$

Denote

$$\tilde{D}(y) := D(J_1(\log(y))) = y\Big[U_2(X_T - J_1(\log(y))) - \lambda(\theta) - \mu(\theta)B_T\Big], \quad y > 0.$$

Then, suppressing the arguments,

$$\tilde{D}''(y) = -U_2'\frac{J_1'}{y} + U_2''\frac{(J_1')^2}{y} - U_2'\frac{J_1''}{y} < 0.$$

So $\tilde{D}$ is concave on $(0, \infty)$ and then we can maximize it on its domain. By the relation between D and $\tilde{D}$, we may maximize inside the integral of the Lagrangian (2.23).

Since

$$\frac{d}{dy}G(x,y) = -U_2' + \frac{U_2''}{U_1'} + \frac{U_2'U_1''}{|U_1'|^2} < 0,$$

for z in the range of $G(x,\cdot)$ there exists a unique function $\tilde{H}(x,z)$ such that

$$G(x,\tilde{H}(x,z)) = z. \tag{2.25}$$

Let

$$\text{Range}(G(X_T)) := \{G(X_T,y) : y \text{ is in the domain of } U_1\}. \tag{2.26}$$

Thus, if $\lambda(\theta) + \mu(\theta)B_T \in \text{Range}(G(X_T))$, one should choose $C_T(\theta) = \tilde{H}(X_T, \lambda(\theta) + \mu(\theta)B_T)$. On the other hand, if $\lambda(\theta) + \mu(\theta)B_T \notin \text{Range}(G(X_T))$, then we should choose the smallest or the largest possible value of $C_T(\theta)$. In this case, we assume

Assumption 2.2 *If Range$(G(X_T)) \neq \mathbb{R}$, we require that the payoff $C_T(\theta)$ is bounded:*

$$L \leq C_T(\theta) \leq U$$

for some finite constants L, U. That is, we consider a smaller set $\mathcal{A}_1$ in (2.21)*. On the other hand, if Range$(G(X_T)) = \mathbb{R}$, we set*

$$L = -\infty\,, \quad U = +\infty.$$

Introduce the events

$$\begin{aligned}
A_1 &:= \{\omega : \lambda(\theta) + \mu(\theta)B_T(\omega) \leq G(X_T(\omega), U)\};\\
A_2 &:= \{\omega : G(X_T(\omega), U) < \lambda(\theta) + \mu(\theta)B_T(\omega) < G(X_T(\omega), L)\};\\
A_3 &:= \{\omega : \lambda(\theta) + \mu(\theta)B_T(\omega) \geq G(X_T(\omega), L)\}.
\end{aligned}$$

From all the above, the optimal C_T is given by

$$C_T(\theta) = U\mathbf{1}_{A_1} + \tilde{H}(X_T, \lambda(\theta) + \mu(\theta)B_T)\mathbf{1}_{A_2} + L\mathbf{1}_{A_3} =: H(\lambda(\theta), \mu(\theta)). \tag{2.27}$$

For any constant θ, λ, μ, define the following deterministic functions:

$$\begin{cases}
H_1(\theta,\lambda,\mu) := E\Big\{M_T^\theta \exp\Big(U_1(H(\lambda,\mu)) - \xi\Big)\Big\};\\
H_2(\theta,\lambda,\mu) := E\Big\{M_T^\theta \exp\Big(U_1(H(\lambda,\mu)) - \xi\Big)B_T\Big\};\\
H_3(\theta,\lambda,\mu) := E\Big\{M_T^\theta \exp\Big(U_1(H(\lambda,\mu)) - \xi\Big)U_2(X_T - H(\lambda,\mu))\Big\}.
\end{cases}$$

Assume also

Assumption 2.3 *For any $\theta \in [\theta_L, \theta_H]$, (H_1, H_2) have inverse functions (h_1, h_2).*

Then, in order to satisfy the second and third constraint in (2.22), we need to have

$$\lambda(\theta) = h_1(\theta, e^{R(\theta)}, e^{R(\theta)}[R'(\theta) + T\theta]); \quad \mu(\theta) = h_2(\theta, e^{R(\theta)}, e^{R(\theta)}[R'(\theta) + T\theta]). \tag{2.28}$$

and the principal's problem is as in the following

Theorem 2.2 *Under Assumptions* 2.2 *and* 2.3*, and assuming that C_T defined by* (2.27) *and* (2.28) *is in $\mathcal{A}_1$, the principal's relaxed problem is given by*

$$\sup_{R(\theta)\geq r(\theta)} \int_{\theta_L}^{\theta_H} e^{-R(\theta)} H_3(\theta, h_1(\theta, e^{R(\theta)}, e^{R(\theta)}[R'(\theta) + T\theta]), h_2(\theta, e^{R(\theta)}, e^{R(\theta)}[R'(\theta) + T\theta]))dF(\theta). \tag{2.29}$$

Notice that this is a deterministic calculus of variations problem.

Remark 2.5 (i) The first order condition (2.24) can be written as, using $B_T = \int_0^T dX_t/v_t$,

$$\frac{U_2'(X_T - C_T(\theta))}{U_1'(C_T(\theta))} = -\mu(\theta)\int_0^T \frac{1}{v_t}dX_t - \lambda(\theta) + U_2(X_T - C_T(\theta)). \quad (2.30)$$

This is a generalization, to the case of adverse selection, of the classical Borch condition for the first-best full information case (see (3.24) below), and the generalization of the second-best case (no adverse selection, $\mu = 0$) in CWZ [7]. In our "third-best" case of moral hazard and adverse selection, the ratio between the marginal utilities of the principal and of the agent in (2.30) becomes random, with the first term proportional to $B_T = \int_0^T \frac{1}{v_t}dX_t$, the volatility weighted average of the output process X. The optimal contract is no longer a function of only the final output value X_T, unless the volatility is constant. Instead, the optimal contract, in addition to X_T, depends on the volatility weighted average B_T of the path of the output process X, which will have high/low values depending on when the underlying random source of risk has high/low values. This term is multiplied by $\mu(\theta)$, the Lagrange multiplier for the truth-telling first-order condition. Thus, making the contract contingent on the level of the underlying source of risk, the principal is trying to get the agent to reveal her type.
Another term influencing the ratio is the utility of the principal. This makes the relationship between C_T and X_T even "more nonlinear" than in the first best case, and makes the effect of X on the marginal utilities more pronounced. This effect is present without adverse selection, too, and is due to moral hazard.

(ii) If we assume that v is constant and that the above equation can be solved for the optimal contract $C_T = C_T(X_T)$ as a function of X_T, it can be computed from the above equation, omitting the functions arguments, that

$$\frac{\partial}{\partial X_T}C_T = \frac{U_1'(U_1'U_2' - U_2'' - \mu/v)}{U_2'(U_1')^2 - U_2''U_1' - U_2'U_1''}.$$

Thus, unlike the first-best case and the second-best case (no adverse selection) in which $\mu = 0$, it is not a priori clear that the contract is a non-decreasing function of X_T. Unfortunately, the only example which we can solve is the case of linear utilities, in which case we will see that the contract is, in fact, a non-decreasing function of X_T.

Remark 2.6 Using the methods of CWZ [7], under technical conditions, it can be shown that for a more general cost function $g(u - \theta)$, the optimal solution necessarily satisfies this system of Backward SDEs:

$$Y_t^1 = U_1(C_T(\theta)) - G_T(\theta) - \int_t^T g'(u_s - \theta)dB_s^u.$$

$$Y_t^2 = U_2(X_T - C_T(\theta)) - \lambda\int_t^T g'(u_s - \theta)dt - \int_t^T Z_s^2 dB_s^u;$$

$$Y_t^3 = \frac{U_2'(X_T - C_T(\theta))}{U_1'(C_T(\theta))} - \int_t^T Z_s^3 dB_s^u.$$

$$Z_t^2 = [Z_t^3 + \lambda]g''(u_t - \theta).$$

where λ is found from the first-order condition for the truth-telling constraint, which can be written as

$$E^u \left\{ \int_0^T g'(u_t - \theta)dt \right\} = R'(\theta). \tag{2.31}$$

The principal's problem reduces then to

$$\sup_R \int_{\theta_L}^{\theta_H} E^u\{U_2(X_T - J_1(Y_T^1))\}dF(\theta),$$

under the constraint $R(\theta) \geq r(\theta)$.

It seems very hard to say anything about the existence or the nature of the solution, though, unless g is quadratic, and because of that we omit the details.

3 Risk-neutral agent and principal

The case of the risk-neutral agent and principal is the only case that we can solve explicitly, and we can also show that the first-order approach introduced in Sect. 2 works, as we do next.

3.1 Third best

Suppose that

$$U_1(x) = x, \quad U_2(x) = kx, \quad X_t = x + vB_t, \quad G_T(\theta) = \frac{1}{2}\int_0^T (u_t - \theta)^2 dt \tag{3.1}$$

for some positive constants k, x, v, and no bounds on C_T, $L = -\infty$, $U = \infty$. From (2.24) we get a linear relationship between the payoff C_T and B_T (equivalently, X_T)

$$x + vB_T - C_T(\theta) = 1 + \frac{1}{k}[\lambda(\theta) + \mu(\theta)B_T].$$

From this we can write

$$C_T(\theta) = a(\theta) + b(\theta)B_T$$

Note that

$$E[e^{(\theta+b(\theta))B_T}] = e^{\frac{T}{2}(\theta+b(\theta))^2}, \quad E[B_T e^{(\theta+b(\theta))B_T}] = (\theta + b(\theta))Te^{\frac{T}{2}(\theta+b(\theta))^2}. \tag{3.2}$$

By the last two equations in (2.22) we get

$$\begin{aligned} &e^{a(\theta)-T\theta^2/2+T(\theta+b(\theta))^2/2} = e^{R(\theta)}, \\ &(\theta + b(\theta))Te^{a(\theta)-T\theta^2/2+T(\theta+b(\theta))^2/2} = e^{R(\theta)}[R'(\theta) + T\theta]. \end{aligned} \tag{3.3}$$

Then we get, omitting the argument θ,

$$b = \frac{1}{T}R', \quad a = R - \theta R' - \frac{(R')^2}{2T}. \tag{3.4}$$

Plugging into the principal's problem, we see that he needs to maximize

$$k\int_{\theta_L}^{\theta_H} e^{a(\theta)-R(\theta)-T\theta^2/2}E[e^{(\theta+b(\theta))B_T}(x-a(\theta)+(v-b(\theta))B_T)]dF(\theta)$$

which is, using (3.2) and (3.4), equal to

$$k\int_{\theta_L}^{\theta_H}\left\{x - R(\theta) - \frac{T\theta^2}{2} + \frac{T}{2}\left(\theta + \frac{R'(\theta)}{T}\right)^2 + \left(v - \frac{R'(\theta)}{T}\right)\left(\theta + \frac{R'(\theta)}{T}\right)T\right\}dF(\theta) \tag{3.5}$$

Maximizing this is equivalent to minimizing

$$\int_{\theta_L}^{\theta_H}\left\{R(\theta) + \frac{1}{2T}(R'(\theta))^2 - vR'(\theta)\right\}dF(\theta) \tag{3.6}$$

and it has to be done under the constraint

$$R(\theta) \geq r(\theta)$$

for some given function $r(\theta)$. If this function is constant, we have the following result:

Theorem 3.1 *Assume* (3.1)*, assume that θ is uniform on $[\theta_L, \theta_H]$, and the IR lower bound is $r(\theta) \equiv r_0$. The the principal's problem* (2.9) *has a unique solution as follows. Denote $\theta^* := \max\{\theta_H - v, \theta_L\}$. The optimal choice of agent's utility R by the principal is given by*

$$R(\theta) = \begin{cases} r_0, & \theta_L \leq \theta < \theta^*; \\ r_0 + T\theta^2/2 + T(v-\theta_H)\theta - T(\theta^*)^2/2 - T(v-\theta_H)\theta^*, & \theta^* \leq \theta \leq \theta_H. \end{cases} \tag{3.7}$$

The optimal agent's effort is given by

$$\hat{u}(\theta) - \theta = \begin{cases} 0, & \theta_L \leq \theta < \theta^*; \\ v + \theta - \theta_H, & \theta^* \leq \theta \leq \theta_H. \end{cases} \tag{3.8}$$

The optimal contract is, recalling (3.4)*,*

$$\hat{C}_T(\theta) = \begin{cases} a(\theta), & \theta_L \leq \theta < \theta^*; \\ a(\theta) - xb(\theta)/v + \frac{\theta+v-\theta_H}{v}X_T, & \theta^* \leq \theta \leq \theta_H. \end{cases} \tag{3.9}$$

With fixed θ, the optimal principal's utility is

$$E^{\hat{u}(\theta)}[U_2(X_T - \hat{C}_T(\theta))] = \begin{cases} k\Big[x - r_0 + \theta v T\Big], & \theta_L \leq \theta < \theta^*; \\ k\Big[x - r_0 + T\frac{(\theta^*+v)^2}{2} + T\theta_H(2\theta - \theta^* - \frac{\theta_H}{2})\Big], & \theta^* \leq \theta \leq \theta_H. \end{cases} \tag{3.10}$$

Remark 3.1 (i) If $v < \theta_H - \theta_L$, a range of lower type agents gets no rent above the reservation value r_0, the corresponding contract is not incentive as it does not depend on X, and the effort $\hat{u} - \theta$ is zero. The higher type agents get utility $R(\theta)$ which is quadratically increasing in their type θ, while the principal's utility is linear in θ. As the volatility (noise) gets lower, the non-incentive range gets wider, and only the highest type agents get informational rent. The rent gets smaller with lower values of volatility, even though the incentives (the slope of C_T with respect to X_T) become larger.

(ii) Similar results can be obtained for general distribution F of θ, that has a density $f(\theta)$, if we notice that the solution y to the Euler equation (3.13) below is:

$$y(\theta) = \beta + vT\theta + \alpha \int_{\theta_L}^{\theta} \frac{dx}{f(x)} + T \int_{\theta_L}^{\theta} \frac{F(x)}{f(x)} dx \tag{3.11}$$

for some constants α and β.

Proof of the theorem We show here that (3.7)–(3.10) solve the relaxed principal's problem (2.21)–(2.22), and we check the truth-telling constraint in Lemma 3.1 below. First, one can prove straightforwardly that $\hat{u} \in \mathcal{A}_0$ and $\hat{C}_T \in \mathcal{A}$.

If F has density f, denote

$$\varphi(y, y') := \left[y + \frac{1}{2T}(y')^2 - vy' \right] f \tag{3.12}$$

Here y is a function on $[\theta_L, \theta_H]$ and y' is its derivative. Then, the Euler ODE for the calculus of variations problem (3.6), denoting by y the candidate solution, is (see, for example, [21])

$$\varphi_y = \frac{d}{d\theta} \varphi_{y'}$$

or, in our example,

$$y'' = T + (vT - y')\frac{f'}{f} \tag{3.13}$$

Since θ is uniformly distributed on $[\theta_L, \theta_H]$, this gives

$$y(\theta) = T\theta^2/2 + \alpha\theta + \beta$$

for some constants α, β. According to the calculus of variations, on every interval R is either of the same quadratic form as y, or is equal to r_0. One possibility is that, for some $\theta_L \le \theta^* \le \theta_H$,

$$R(\theta) = \begin{cases} r_0, & \theta_L \le \theta < \theta^*; \\ T\theta^2/2 + \alpha\theta + \beta, & \theta^* \le \theta \le \theta_H. \end{cases} \tag{3.14}$$

In this case, $R(\theta)$ is not constrained at $\theta = \theta_H$.

By standard results of calculus of variations, the free boundary condition is then, recalling notation (3.12),

$$0 = \varphi_{y'}(\theta_H) = \frac{1}{T} y'(\theta_H) - v \tag{3.15}$$

from which we get

$$\alpha = T(v - \theta_H)$$

Moreover, by the principle of smooth fit, if $\theta_L < \theta^* < \theta_H$, we need to have

$$0 = R'(\theta^*) = T\theta^* + \alpha$$

which gives

$$\theta^* = \theta_H - v$$

if $v < \theta_H - \theta_L$. If $v > \theta_H - \theta_L$ then we can take

$$\theta^* = \theta_L.$$

In either case the candidate for the optimal solution is given by (3.7).

Another possibility would be

$$R(\theta) = \begin{cases} T\theta^2/2 + \alpha\theta + \beta, & \theta_L \le \theta < \theta^*; \\ r_0, & \theta^* \le \theta \le \theta_H, \end{cases} \tag{3.16}$$

In this case the free boundary condition at $\theta = \theta_L$ would give $\alpha = Tv$, but this is incompatible with the smooth fit condition $T\theta^* + \alpha = 0$, if we assume $v > 0$.

The last possibility is that $R(\theta) = T\theta^2/2 + \alpha\theta + \beta$, everywhere. We would get again that at the optimum $\alpha = T(v - \theta_H)$, and β would be chosen so that $R(\theta^*) = r_0$ at its minimum point θ^*. Doing computations and comparing to the case (3.7), it is easily checked that (3.7) is still optimal.

Note that solving the BSDE (2.13), we get $\hat{u} = \theta + b(\theta)$, which gives (3.8). Also, (3.10) follows by computing the integrand in (3.5). □

It remains to check that the contract is truth-telling. This follows from the following lemma, which is stated for general density f.

Lemma 3.1 *Consider the hazard function $h = f/(1-F)$, and assume that $h' > 0$. Then the contract $C_T = a(\theta) + b(\theta)B_T$, where a and b are chosen as in* (3.4)*, is truth-telling.*

Proof It is straightforward to compute

$$V(\theta, \tilde{\theta}) = \log E[M_T^\theta e^{a(\tilde{\theta}) + b(\tilde{\theta})B_T}] = R(\tilde{\theta}) + R'(\tilde{\theta})(\theta - \tilde{\theta}).$$

We have

$$\partial_{\tilde{\theta}} V(\theta, \tilde{\theta}) = R''(\tilde{\theta})(\theta - \tilde{\theta}) \tag{3.17}$$

Here, either $R(\tilde{\theta}) = r_0$ or $R(\tilde{\theta}) = y(\tilde{\theta})$ where y is the solution (3.11) to the Euler ODE. If $R(\tilde{\theta}) = r_0$ then $V(\theta, \tilde{\theta}) = r_0$, which is the lowest the agent can get. Otherwise, with $R = y$ and omitting the argument θ, note that

$$\begin{aligned} R' &= vT + \alpha/f + TF/f \\ R'' &= T - (\alpha + FT)f'/f^2. \end{aligned}$$

The free boundary condition (3.15) for $y = R$ is still the same, and gives

$$\alpha = -TF(\theta_H) = -T.$$

Table 1 Optimal contracts for various cases

	No A.S.	A.S.
Risk-neutral case, θ in the cost		
No M.H.	$C_T(\theta) = c(\theta) + X_T$	Theorem 3.1
M.H.	$C_T(\theta) = c(\theta) + X_T$	Theorem 3.1
Risk-neutral case, θ in the drift		
No M.H.	$C_T(\theta) = c(\theta) + X_T$	Contract (3.23)
M.H.	$C_T(\theta) = c(\theta) + X_T$	Theorem 3.1
General utilities, θ in the cost		
No M.H.	Borch rule (3.19)	Borch rule (3.19)
M.H.	Eq. (3.24)	Eq. (2.30)
General utilities, θ in the drift		
No M.H.	Borch rule (3.19)	Eq. (3.22)
M.H.	Eq. (3.24)	Eq. (2.30)

Notice that this implies

$$R'' = T + T\frac{f'}{f^2}(1 - F)$$

Thus, $R'' > 0$ if and only if

$$f'(1 - F) > -f^2 \tag{3.18}$$

Note that this is equivalent to $h' > 0$, which is assumed. From (3.17), we see, that under condition (3.18), that $V(\theta, \tilde{\theta})$ is increasing for $\tilde{\theta} < \theta$ and decreasing for $\tilde{\theta} > \theta$, so $\tilde{\theta} = \theta$ is the maximum.

3.2 Comparison with the first-best and the second-best

Before analyzing the first-best and the second-best cases and comparing them to the third-best, we first give tables which point out where to find the results, and a short summary thereof.

3.2.1 Tables and summary

Table 1 indicates (where to find) the main results.

Many of the results are qualitatively in the same spirit as the results from the static theory, as presented, for example, in the book by [5]. In particular, with moral hazard, the highest type agent gets the first-best contract, while the lower type agents get less incentive contracts and provide less effort. Moreover, with no adverse selection, moral hazard is not relevant when the agent is risk-neutral, because the principal "sells the whole firm" to the agent. However, a new observation here is that in the case of adverse selection without moral hazard, there is a difference whether we assume that the agents differ in their cost function, or in the return θ they can attain. In the case of both moral hazard and adverse selection, these formulations are equivalent. However, without moral hazard, and with risk-neutrality, we may use the linear contract (3.9) if the type θ represents the cost function, but we get the "bang–bang" contract (3.23) below, which depends on the random "benchmark" value B_T, if the type θ represents the return. As the referee points out, the difference is due to the fact that when θ is in the drift, then the principal gets some information about θ when observing X (or B), which is not the case when θ is in the cost function. Because of that, the truth-telling constraint is very different for the two formulations, if there is no moral hazard, and leads to the different constraints and different contracts. With moral hazard, the effort and the type are both unobservable by the principal, and equivalent up to translation (that is, we can do a change of variables from u to $u+\theta$). Hence, there is no difference in information on θ obtained from observing X in the two cases. Notice furthermore, with θ in the cost function, even if the principal does not sell the whole firm to the agent (except to the highest type), the moral hazard still is not relevant, unlike when θ is in the drift.

Another interesting observation is that the Borch rule (3.19) for the first-best risk sharing still holds in the presence of adverse selection when there is no moral hazard, if type θ is in the cost function. This is again because in this case there is no information about θ to be obtained from observing X, and the truth-telling constraint does not involve the contract payoff C_T, so that the relationship between X_T and C_T remains the same as in the first-best.

3.2.2 First-best case

We modify now the previous model to

$$dX_t = u_t v_t dt + v_t dB_t$$

with fixed positive process v. We assume that both u and θ are observed by the principal. We also assume $\xi = \int_0^T h(X_t)dt$. It follows from Cvitanić et al. [8], henceforth CWZ [8] (see also Cadenillas et al. [6], henceforth CCZ [6], that the first order conditions for the optimization problem of the principal are:

$$\frac{U_2'(X_T - C_T)}{U_1'(C_T)} = \lambda$$
$$\frac{\lambda}{v_t} g'(u_t) = U_2'(X_T - C_T) - \int_t^T \lambda h'(X_s)ds - \int_t^T Z_s^C dB_s, \tag{3.19}$$

for an appropriate adapted process Z^C. If $h \equiv 0$, the latter condition can be written as

$$\frac{\lambda}{v_t} g'(u_t) = E_t[U_2'(X_T - C_T)] = \lambda E_t[U_1'(C_T)].$$

Equation (3.19) is the standard Borch optimality condition for risk-sharing.

With linear utilities, $U_1(x) = x$, $U_2(x) = kx$, and $G_T = \int_0^T g_t dt$, in the first best case the principal has to maximize

$$E\left[X_T - R(\theta) - \int_0^T g_t dt\right] = x + E\int_0^T [v_t u_t - g_t]dt - R(\theta).$$

Thus, we have to maximize $v_t u_t - g_t$, which in our case gives

$$\hat{u}_t - \theta \equiv v_t \tag{3.20}$$

A contract which implements this is

$$\hat{C}_T(\theta) = c(\theta) + X_T$$

where $c(\theta)$ is chosen so that the participation constraint $R(\theta) = r(\theta)$ is satisfied.

The optimal effort is always larger than in the adverse selection/moral hazard case (3.8), and the contract is "more incentive", in the sense that $C_T(\theta) = c(\theta) + X_T$, while in the adverse selection/moral hazard case $C_T(\theta) = a(\theta) + \frac{b(\theta)}{v}[X_T - x]$ with $b(\theta) < v$. With the constant IR bound, $r(\theta) \equiv r_0$, the agent's rent $R(\theta) \equiv r_0$ is no larger than the adverse selection/moral hazard case (3.7).

3.2.3 Second best: adverse selection without moral hazard

Case A: Unknown cost Consider the same model as in the first-best case, but we now assume that u is observed by the principal, while θ is not. This is the case of adverse selection without moral hazard. We also also assume $\xi = 0$, $g(u_t) = (u_t - \theta)^2/2$.

The revelation principle gives

$$E\left[U_1'(C_T(\theta))C_T'(\theta) - \int_0^T g'(u_t - \theta)\partial_\theta u_t dt\right] = 0$$

which, when taking derivative of the agent's value function $R(\theta)$, implies

$$R'(\theta) = E\left[\int_0^T g'(u_t - \theta)dt\right]$$

The principal's relaxed problem is to maximize the Lagrangian

$$\int_{\theta_L}^{\theta_H} E\left[U_2(X_T - C_T(\theta)) - \lambda(\theta)\left[U_1(C_T(\theta)) - \int_0^T (g_t + \mu(\theta)g_t')dt\right]\right] dF(\theta)$$

The integrand is the same as for the case without the truth-telling constraint, but with the cost function $g + \mu g'$. Thus, as above, the first order condition for the optimization problem of the principal inside the integral is

$$\lambda(\theta)U_1'(C_T(\theta)) = U_2'(X_T - C_T(\theta))$$

We see that the optimality relation between X_T and C_T is of the same form as in the first best case, that is, the Borch rule applies. The reason for this is that, in this case in which the type θ determines only the cost function but not the output, the first-order truth-telling constraint can be written in terms of the action u, and it does not involve the principal's choice of payoff C_T, so the problem, for a fixed agent's utility $R(\theta)$, becomes equivalent to the first-best problem, but with a different cost function.

With linear utilities and constant v, we will show that the solution is the same as when u is not observed. For a given agent's utility $R(\theta)$, the principal's problem is to maximize

$$\int_{\theta_L}^{\theta_H} k\left\{x - R(\theta) + E\int_0^T [vu_t - g_t + \mu(\theta)g'(u_t - \theta)]dt\right\} dF(\theta)$$

In our case this gives

$$\hat{u}_t \equiv v + \theta + \mu(\theta)$$

The revelation principle is satisfied if

$$(v+\mu(\theta))T = R'(\theta)$$

Going back to the principal's problem, he has to maximize, over $R \geq r$,

$$\int_{\theta_L}^{\theta_H} k\{x - R(\theta) + v\theta T - \frac{(R'(\theta))^2}{2T} + vR'(\theta)\}dF(\theta)$$

This is the same problem as (3.6) in the case of hidden action u.

Case B: Unknown drift We now consider the model

$$dX_t = (u_t + \theta)v_t dt + v_t dB^\theta$$

where u is observed, but θ is not. The cost function is $g(u) = u^2/2$, so it is known, but the distribution of the output depends on θ. As before, introduce the agent's utility

$$R(\theta) = E\left[M_T^\theta\left\{U_1(C_T(\theta)) - \int_0^T u_t^2/2dt\right\}\right]$$

The truth-telling first-order condition is then

$$E\left[M_T^\theta B_T\left\{U_1(C_T(\theta)) - \int_0^T u_t^2/2dt\right\}\right] = R'(\theta) + T\theta R(\theta)$$

and the principal's Lagrangian for the relaxed problem is

$$E\left[\int_{\theta_L}^{\theta_H} M_T^\theta\left\{U_2(X_T - C_T(\theta)) - \left[U_1(C_T(\theta)) - \int_0^T u_t^2/2dt\right][\lambda(\theta) + \mu(\theta)B_T]\right\} dF(\theta)\right]. \tag{3.21}$$

We require limited liability constraint

$$C_T(\theta) \geq L.$$

We can check that the integrand as a function of C_T is decreasing in C_T if $\lambda + \mu B_T \geq 0$, and is otherwise a concave function of C_T. Thus, the first order condition in C_T is, omitting the argument θ,

$$\begin{aligned} \frac{U_2'(X_T - C_T)}{U_1'(C_T)} &= -\lambda - \mu B_T \quad \text{if } \lambda + \mu B_T < 0 \\ C_T &= L \quad\quad\quad\quad\quad \text{if } \lambda + \mu B_T \geq 0 \end{aligned} \tag{3.22}$$

Comparing to the moral hazard/adverse selection case (2.30), we see that the last term, $U_2(X_T - C_T)$ disappears, because there is no moral hazard. If the truth-telling is not binding, $\mu = 0$, then the principal pays the lowest possible payoff L, so that the IR constraint is satisfied.

In the linear utilities case, $U_1(x) = x$, $U_2(x) = kx$, looking at (3.21), we see that, in order to have a solution, we need to assume

$$L \leq C_T \leq U\,, \quad u \leq \bar{u}$$

and the contract will take only extreme values:

$$C_T = L\mathbf{1}_{\{\lambda+\mu B_T<k\}} + U\mathbf{1}_{\{\lambda+\mu B_T>k\}}. \tag{3.23}$$

So, when the truth-telling is binding, $\mu \neq 0$, in order to get the agent to reveal her type, the payoff can either take the minimal or the maximal value, depending on the level of the value B_T of the weighted average of the output process X.

The optimal action can be found to be

$$u_t = \bar{u}\mathbf{1}_{\{\lambda+\mu[B_t+\theta(T-t)]\le 0\}} + \frac{kv}{\lambda+\mu[B_t+\theta(T-t)]}\mathbf{1}_{\{\lambda+\mu[B_t+\theta(T-t)]>0\}}.$$

3.2.4 Second best: moral hazard without adverse selection

Assume now that type θ is observed, but action u is not. Then similarly as in the adverse selection/moral hazard case (see also CWZ 2005), setting $\mu \equiv 0$, we get, omitting again the dependence on θ,

$$\frac{U_2'(X_T - C_T)}{U_1'(C_T)} = -\lambda + U_2(X_T - C_T) \tag{3.24}$$

where λ is determined so that the IR constraint is satisfied with equality. The ratio of marginal utilities no longer depends on the weighted average of X, but it still increases with the principal's utility.

In the linear case we have

$$C_T = \frac{\lambda}{k} - 1 + X_T$$

and

$$\hat{u} - \theta = v,$$

the same as the first-best.

4 Suboptimal truth-telling contracts

In general, it is very hard to compute the (candidates for) third-best optimal contracts and/or check that the computed candidate contracts actually are truth-telling. We now explore the following idea to amend for that: we suggest to use contracts of the form suggested by the optimality conditions, but, instead of finding the Lagrange multipliers $\mu(\theta)$ and $\lambda(\theta)$ from the constraints (2.22) in terms of $R(\theta)$ and $R'(\theta)$, we restrict the choice of $\mu(\theta)$ and $\lambda(\theta)$ further, in order to make computations simpler, while still resulting in a truth-telling contract. This will result in effectively reducing the possible choices for $R(\theta)$ in the principal's optimization problem, leading to a contract optimal in a smaller family of functions for $R(\theta)$, hence a suboptimal contract.

Here is a result in this direction.

Theorem 4.1 *Assume $\xi = 0$ for simplicity. Also assume that*

$$\left[|U_1(x)|^4 + U_1'(x)U_1'''(x) - 3(U_1''(x))^2\right]U_2'(y) - 3U_1'(x)U_1''(x)U_2''(y) - (U_1'(x))^2U_2'''(y) \le 0, \tag{4.1}$$

for any x, y in the domains of U_1, U_2, respectively. Assume further that λ is convex and μ is linear, and there is a random variable $\eta \in \mathcal{F}_T$, which may take value $-\infty$ and is independent of θ, such that for any $\theta \in [\theta_L, \theta_H]$,

$$W(\theta) := \max(\lambda(\theta) + \mu(\theta)B_T, \eta) \in \text{Range}(G(X_T)), \quad P-a.s., \tag{4.2}$$

where $Range(G(X_T))$ is defined by (2.26). Consider a smaller $\mathcal{A}_1$ when $\eta \neq -\infty$, as defined by Assumption 2.2, extended to the current framework. Then the first order condition is sufficient for truth telling.

Remark 4.1 (i) One sufficient condition for (4.1) is

$$|U_1(x)|^4 + U_1'(x)U_1'''(x) - 3(U_1''(x))^2 \leq 0; \quad U_2''' \geq 0. \tag{4.3}$$

(ii) The following examples satisfy (4.3):

$$U_1(x) = \log(x),\ x > 0; \quad \text{or} \quad U_1(x) = -e^{-\gamma x},\ x \geq -\frac{\log(2)}{2\gamma};$$

and

$$U_2(x) = x; \quad \text{or} \quad U_2(x) = -e^{-\gamma x}; \quad \text{or} \quad U_2(x) = \log(x); \quad \text{or}$$
$$U_2(x) = x^{\gamma}, \quad 0 < \gamma < 1.$$

(iii) The following example satisfies (4.1) but not (4.3):

$$U_1(x) = kx; \quad U_2(x) = -e^{-\gamma x}; \quad k \leq \gamma.$$

(iv) The example when both U_1 and U_2 are linear does not satisfy (4.1). However, see Example 4.1 for an illustration how good the present approach is in finding a suboptimal contract for this case.

The condition (4.2) is much more difficult to satisfy than (4.1). In (4.2) we truncate $\lambda + \mu B_T$ from below. If we could truncate it from above, we would be able to apply the theorem to all the examples in Remark 4.1. However, in the proof we need $W(\theta)$ to be convex in θ, which is true for the truncation from below, but not true for the truncation from above.

Proof In Appendix.

4.1 Examples

We now look at a couple of examples and a further simplification.

Example 4.1 We assume the set-up of Theorem 3.1. We already know what the optimal solution is in this case, but let's pretend we don't, and use the present approach of looking only at linear μ and convex λ. It can be easily computed that with $\mu(\theta) = \alpha + \beta\theta$, the first order condition (2.20) implies

$$\lambda(\theta) = \text{const} - \beta T\left[\frac{1}{2}\theta^2(1-\beta) + \theta(v-\alpha)\right]$$

and that the difference between the agent's utility when lying and not lying is

$$V(\theta, \tilde{\theta}) - V(\theta, \theta) = \frac{1}{2}\beta(\tilde{\theta} - \theta)^2$$

so the truth-telling contracts are those with $\beta \leq 0$. It can also be verified straightforwardly that the principal's average utility is of the form

$$-\frac{1}{2}c_1\beta^2 + \beta\left[c_1 - \alpha c_2 - \frac{1}{2}\theta_L^2\theta_H + \frac{1}{2}\theta_L^3\right] - \frac{1}{2}\alpha^2 c_3 + \alpha[c_2 - \theta_L\theta_H + \theta_L^2]$$

where

$$c_1 = T\int_{\theta_L}^{\theta_H}\theta^2 d\theta, \quad c_2 = T\int_{\theta_L}^{\theta_H}\theta d\theta, \quad c_3 = T\int_{\theta_L}^{\theta_H} d\theta$$

The optimization gives

$$\alpha = \theta_H,\ \beta = -1\ ,\ \text{thus}\ \mu(\theta) = \theta_H - \theta$$

If we look back to Sect. 3.1, we can see that the unrestricted $\mu = \nu - b = \nu - R'/T$, and, in fact, we also have $\mu(\theta) = \theta_H - \theta$, but only for $\theta \geq \theta^* = \max[\theta_H - \nu, \theta_L]$. Thus, the suboptimal contract is actually the same as the optimal contract if the volatility is high enough, $\nu \geq \theta_H - \theta_L$. However, in general our suboptimal contract ignores the distinction that the optimal contract makes between the higher type agents and the lower type agents, and is of the same form as the optimal contract only for the higher type agents.

Example 4.2 Assume

$$U_1(x) = \log(x); \quad U_2(x) = x.$$

Then (4.1) holds. We consider only those λ, μ such that (4.2) holds true with $\eta = -\infty$. Note that the first order condition (2.24) for C_T gives

$$C_T(\theta) = \frac{1}{2}[X_T - \lambda(\theta) - \mu(\theta)B_T]. \tag{4.4}$$

Therefore,

$$e^{V(\theta,\tilde{\theta})} = E^\theta\{C_T(\tilde{\theta})\} = \frac{1}{2}[E^\theta\{X_T\} - \lambda(\tilde{\theta}) - \mu(\tilde{\theta})\theta T].$$

Since here we obtain $V(\theta,\tilde{\theta})$ explicitly, we may study the truth telling directly without assuming λ is convex and μ is linear. From the previous equation, the first order condition for truth-telling is

$$\lambda'(\theta) + T\theta\mu'(\theta) = 0.$$

Then, for some constant a,

$$\lambda(\tilde{\theta}) = a - T\tilde{\theta}\mu(\tilde{\theta}) + T\int_{\theta_L}^{\tilde{\theta}}\mu(\tau)d\tau. \tag{4.5}$$

Thus

$$e^{V(\theta,\tilde{\theta})} - e^{V(\theta,\theta)} = \frac{T}{2}\int_{\theta}^{\tilde{\theta}}[\mu(\tilde{\theta}) - \mu(\tau)]d\tau.$$

We find that the contract is truth telling if and only if μ is decreasing. (Note: this may not be true for λ, μ which do not satisfy (4.2), so what we obtain here is still suboptimal contracts.) It remains to see when (4.2) holds true. Since $C_T > 0$, we need

$$X_T - \lambda(\theta) - \mu(\theta)B_T > 0, \quad a.s. \tag{4.6}$$

This obviously depends on X_T (or v_t). We discuss three cases.

Case 1 $X_T = x_0 + v_0 B_T$. In this case we must have $\mu(\theta) = v_0, \forall\theta \in [\theta_L, \theta_H]$. Then, by (4.5), λ is a constant and $\lambda < x_0$. Thus $C_T(\theta) = \frac{1}{2}[x_0 - \lambda]$ and $e^{R(\theta,\tilde{\theta})} = \frac{1}{2}[x_0 - \lambda]$. To satisfy the IR constraint $R(\theta) \geq r_0$, we need $\lambda \leq x_0 - 2r_0$.

Case 2 $X_T = x_0 + \frac{1}{2}B_T^2$. Then

$$X_T - \lambda(\theta) - \mu(\theta)B_T \geq x_0 - \lambda(\theta) - \frac{1}{2}\mu(\theta)^2.$$

Thus, we should consider those λ, μ such that

$$\lambda(\theta) + \frac{1}{2}\mu(\theta)^2 \leq x_0.$$

We note that we allow equality above, because the probability is zero that B_T is such that the equality holds.

Case 3 $X_T = x_0 e^{\sigma B_T}$ with $x_0 > 0, \sigma > 0$. If $\mu < 0$, we have

$$\lim_{y \to -\infty}[x_0 e^{\sigma y} - \mu y] = -\infty.$$

Hence, in order to ensure (4.6) we need $\mu \geq 0$. Then

$$\inf_y [x_0 e^{\sigma y} - \mu y] = \frac{\mu}{\sigma}\left[1 - \log\left(\frac{\mu}{x_0\sigma}\right)\right].$$

Thus

$$X_T - \lambda(\theta) - \mu(\theta)B_T \geq \frac{\mu(\theta)}{\sigma}\left[1 - \log\left(\frac{\mu(\theta)}{x_0\sigma}\right)\right] - \lambda(\theta).$$

Consequently, we need $\mu \geq 0$ decreasing such that

$$\frac{\mu(\theta)}{\sigma}\left[1 - \log\left(\frac{\mu(\theta)}{x_0\sigma}\right)\right] - \lambda(\theta) \geq 0.$$

We can compute

$$e^{R(\theta)} = \frac{1}{2}E^\theta\left\{X_T - \lambda(\theta) - \mu(\theta)B_T\right\} = \frac{1}{2}\left[x_0 e^{\frac{1}{2}\sigma^2 T + \sigma T\theta} - \lambda(\theta) - T\theta\mu(\theta)\right];$$

$$E^\theta\left\{e^{U_1(C_T(\theta))}U_2(X_T - C_T(\theta))\right\} = \frac{1}{4}\left[x_0^2 e^{2\sigma^2 T + 2\sigma T\theta} - [\lambda(\theta) + T\theta\mu(\theta)]^2 - T\mu(\theta)^2\right].$$

Denote $\bar{\lambda}(\theta) := \lambda(\theta) + T\theta\mu(\theta)$. Then the suboptimal principal's problem is a deterministic calculus of variations problem given by

$$\max_{\bar{\lambda},\mu} \int_{\theta_L}^{\theta_H} \frac{x_0^2 e^{2\sigma^2 T + 2\sigma T\theta} - \bar{\lambda}(\theta)^2 - T\mu(\theta)^2}{x_0 e^{\frac{1}{2}\sigma^2 T + \sigma T\theta} - \bar{\lambda}(\theta)} dF(\theta)$$

under the constraints:

$$\bar{\lambda}'(\theta) = T\mu(\theta); \quad \mu \geq 0; \quad \mu' \leq 0;$$
$$\bar{\lambda}(\theta) + \frac{\mu(\theta)}{\sigma}[\log(\mu(\theta)) - T\sigma\theta - 1 - \log(x_0\sigma)] \leq 0;$$
$$\frac{1}{2}\left[x_0 e^{\frac{1}{2}\sigma^2 T + \sigma T\theta} - \bar{\lambda}(\theta)\right] \geq e^{r(\theta)}.$$

This is still a hard problem. A further simplification would be to set $\mu = \mu_0$, a constant, and to consider only those $\lambda(\theta)$ for which there is an admissible solution $C_T = C(T, X_T, B_T, \mu_0, \lambda(\theta))$ to the first order condition (2.30). The first order condition for truth-telling is

$$\lambda'(\theta)E\left[M_T^\theta e^{U_1(C_T)}U_1'(C_T)\frac{\partial}{\partial\lambda}C(T, X_T, B_T, \mu_0, \lambda(\theta))\right] = 0$$

In general, this will be satisfied only if $\lambda = \lambda_0$ is a constant independent of θ. Thus, we reduce a calculus of variations problem to a regular calculus problem of finding optimal λ_0 (and μ_0, if we don't fix it). We no longer have a menu of contracts, but the same contract for each type.

Assume now $\theta_L \geq 0$. Moreover, set $\mu \equiv 0$, and assume that

$$\tilde{\lambda} := xe^{\sigma\theta_L T} - 2e^{r_0} < 0$$

The first order condition (2.30) with $\mu = 0$ gives

$$C_T = \frac{1}{2}(X_T - \lambda)$$

and in order to satisfy the IR constraint

$$e^{r_0} = E^{\theta_L}[C_T] = \frac{1}{2}(xe^{\sigma\theta_L T} - \lambda)$$

we need to take $\lambda = \tilde{\lambda}$. By the assumptions, we have $C_T > 0$, and C_T is then the optimal contract among those for which $\mu = 0$, and it is linear, and of the same form as the second best contract. The corresponding u is obtained by solving the BSDE

$$\bar{Y}_t = E_t^\theta[C_T] = \bar{Y}_0 + \int_0^t \bar{Y}_t(u_t - \theta)dB_t^\theta$$

Since

$$E_t^\theta[C_T] = \frac{1}{2}(X_t e^{\sigma\theta(T-t)} - \lambda) = \bar{Y}_0 + \frac{1}{2}\int_0^t e^{\sigma\theta(T-t)}\sigma X_t dB_t^\theta$$

we get

$$u_t - \theta = \frac{e^{\sigma\theta(T-t)}\sigma X_t}{e^{\sigma\theta(T-t)}X_t - \lambda} = \sigma + \frac{\sigma\lambda}{e^{\sigma\theta(T-t)}X_t - \lambda}.$$

Recall that $\lambda < 0$. We see that the effort is increasing in the value of the output so when the promise of the future payment gets higher, the agent works harder. Moreover, the agent of higher type applies more effort, with very high types getting close to the effort's upper bound σ.

The principal's expected utility is found to be

$$\int_{\theta_L}^{\theta_H} e^{-R(\theta)} E^{\theta}[C_T(X_T - C_T)]dF(\theta)$$

$$= \int_{\theta_L}^{\theta_H} e^{-R(\theta)}[xe^{r_0+\sigma\theta_L T} - e^{2r_0}]dF(\theta) + \int_{\theta_L}^{\theta_H} e^{-R(\theta)}\left[\frac{x^2}{4}[e^{(2\sigma\theta+\sigma^2)T} - e^{2\sigma\theta_L T}]\right]dF(\theta)$$

The first integral is what the principal can get if he pays a constant payoff C_T, in which case the agent would choose $u - \theta \equiv 0$. The additional benefit of providing incentives to the agent to apply non-zero effort $u - \theta$ is represented by the second integral. This increases quadratically with the initial value of the output, increases exponentially with the volatility squared, and decreases exponentially with the agent's reservation utility (because $e^{R(\theta)} = e^{r_0} + \frac{x}{2}[e^{\sigma\theta T} - e^{\sigma\theta_L T}]$) . Since the principal is risk-neutral, he likes high volatility.

Let us mention that it is shown in CWZ [8] that in this example of Case 3, the first-best value for the principal is actually infinite, while the second-best is finite. Let us also mention that we don't know how far from optimal the above contract is, in our third-best world.

5 Model II: control of the volatility-return trade-off

Consider the model

$$dX_t = \theta v_t dt + v_t dB_t^{\theta} = v_t dB_t$$

where v_t is controlled, with no cost function. We assume that v is $\mathbf{F}^B$-adapted process such that $E\int_0^T v_t^2 dt < \infty$, so that X is a martingale process under P. We will follow similar steps as in the previous sections, but without specifying exact technical assumptions.

One important example that corresponds to this model is the example of a portfolio manager whose portfolio value is given by the process X, and who produces expected return rate (above the risk-free rate) θv_t. In other words, θ is the Sharpe-ratio that the manager is able to achieve.

Remark 5.1 As an illustrative example, suppose that an investment fund manager dynamically re-balances the money between a risk-free asset and, for simplicity, only one risky asset, but the choice of the risky asset may change over time. In the sequel, we will see that the optimal contract will depend heavily on $B_T = B_T^{\theta} + \theta T$. If the manager invests in the same risky asset all the time, the Sharpe-ratio of the risky asset is the same as the Sharpe-ratio of the fund. In that case, B_T can be obtained from a weighted average of the risky asset's log-prices. In particular, if the volatility of the risky asset is constant, B_T is a function of the time T-value of the risky asset. In this case, the contracts below use the underlying risky asset as a benchmark, indirectly, via B_T. However, this interpretation is no longer valid when the manager keeps changing the choice of the risky asset, in which case B_T depends on the Sharpe-ratio specific to the manager, and is not an exogenously given benchmark.[2]

[2] We thank the referee for pointing this out.

We now assume that v, X are observed by the principal, but θ, B^θ are not. This is consistent with the above application, in the sense that it is well known that it is much harder for the principal to estimate what level of expected return a portfolio manager can achieve, than to estimate the volatility of her portfolio. Actually, in our model, instead of estimation, the principal has a prior distribution for θ, maybe based on historical estimation. On the other hand, we assume somewhat unrealistically, but in agreement with existing models, that the manager knows with certainty the mean Sharpe-ratio θ she can achieve, and she does not have to estimate it.

As before, let

$$M_T^\theta = \exp\left(\theta B_T - \frac{1}{2}\theta^2 T\right).$$

The agent's utility is

$$R(\theta) := E\Big\{M_T^\theta U_1(C_T(\theta))\Big\}, \tag{5.1}$$

and the IR constraint and the first order truth-telling constraint are

$$R(\theta) \geq r(\theta); \quad E\Big\{M_T^\theta U_1'(C_T(\theta))\partial_\theta C_T(\theta)\Big\} = 0.$$

Note that, by differentiating (5.1) with respect to θ, we have

$$E\Big\{M_T^\theta U_1(C_T(\theta))[B_T - \theta T] + M_T^\theta U_1'(C_T(\theta))\partial_\theta C_T(\theta)\Big\} = R'(\theta),$$

which implies that

$$E\Big\{B_T M_T^\theta U_1(C_T(\theta))\Big\} = [R'(\theta) + T\theta R(\theta)]. \tag{5.2}$$

There is also a constraint on X_T, which is the martingale property, or "budget constraint"

$$E[X_T] = x.$$

It is sufficient to have this constraint for the choice of X_T, because we are in a "complete market" framework. More precisely, for any $\mathcal{F}_T$-measurable random variable Y_T that satisfies $E[Y_T] = x$, there exists an admissible volatility process v such that $X_T = X_T^v = Y_T$, by the martingale representation theorem (as is well known in the standard theory of option pricing in complete markets). This constraint is conveniently independent of θ.

If we denote by ν the Lagrange multiplier corresponding to that constraint, the Lagrangian relaxed problem for the principal is then to maximize, over X_T, C_T,

$$E\left[\int_{\theta_L}^{\theta_H} \{M_T^\theta U_2(X_T - C_T(\theta)) - \nu(\theta)X_T - M_T^\theta U_1(C_T(\theta))[\lambda(\theta) + \mu(\theta)B_T]\}dF(\theta)\right] \tag{5.3}$$

If we take derivatives with respect to X_T and disregard the expectation, we get that the optimal X_T is obtained from

$$M_T^\theta U_2'(X_T - C_T(\theta)) = \nu(\theta) \tag{5.4}$$

or, denoting

$$I_i(x) = (U_i')^{-1}(x),$$
$$X_T = C_T(\theta) + I_2\left(\frac{\nu(\theta)}{M_T(\theta)}\right) \tag{5.5}$$

Substituting this back into the principal's problem, and noticing that

$$W_T(\theta) := X_T - C_T(\theta),$$

is fixed by (5.4), we see that we need to maximize over $C_T(\theta)$ the expression

$$E\left[\int_{\theta_L}^{\theta_H} \{-\nu(\theta)[W_T(\theta) + C_T(\theta)] - M_T^\theta U_1(C_T(\theta))[\lambda(\theta) + \mu(\theta)B_T]\} dF(\theta)\right]$$

If $\lambda(\theta) + \mu(\theta)B_T < 0$, the integrand is maximized at $C_T = I_1(\frac{-\nu}{M_T^\theta(\lambda+\mu B_T)})$ where I_1 is defined in (5.5). However, if $\lambda(\theta) + \mu(\theta)B_T \geq 0$, the maximum is attained at the smallest possible value of C_T. Therefore, in order to have a solution, we assume that we have a lower bound on $C_T(\theta)$,

$$C_T(\theta) \geq L$$

for some constant L. Also, to avoid trivialities, we then assume

$$E^\theta[U_1(L)] \geq r(\theta)\ ,\ \ \theta \in [\theta_L, \theta_H]$$

Thus, the optimal $C_T(\theta)$ is given by

$$\hat{C}_T(\theta) = L \vee I_1\left(\frac{-\nu(\theta)}{M_T^\theta(\lambda(\theta) + \mu(\theta)B_T)}\right)\mathbf{1}_{\{\lambda(\theta)+\mu(\theta)B_T<0\}} + L\mathbf{1}_{\{\lambda(\theta)+\mu(\theta)B_T\geq 0\}}. \tag{5.6}$$

Remark 5.2 (i) Notice from (5.4) that the optimal terminal output is given by

$$\hat{X}_T = I_2\left(\frac{\nu(\theta)}{M_T^\theta}\right) + \hat{C}_T(\theta).$$

Hence, the problem of computing the optimal volatility $\hat{v}$ is mathematically equivalent to finding a replicating portfolio for this payoff $\hat{X}_T$, which is a function of B_T (an "option" written on B_T).

(ii) Note that the optimal contract does not depend on the agent's action process v_t or the output X, but only on her type θ and the underlying (fixed) noise B_T. Thus, the agent is indifferent between different choices of action v given this contract. We discuss this issue in more detail in (iii) below, and the next section.

(iii) In the first part of the paper, with controlled drift, the optimal contract was only a function of the final values X_T, B_T, and this was true because of the assumption of the quadratic cost. Here, with volatility control, it can be shown, similarly as in [6,26], that even if there was a general cost function on the volatility, the optimal payoff $\hat{C}_T$ would still be a function of B_T only. However, unlike in those two papers, the optimal contract, in addition to specifying the payoff $\hat{C}_T$, also has to specify the whole path of v, which the principal has to ask the agent to follow (more on this in the next section). In other words, the optimal contract is a pair $(\hat{C}_T, v)$. Notice that we can write

$$B_T = \theta T + B_T^\theta = \int_0^T \frac{dX_t}{v_t}$$

so that the optimal payoff $\hat{C}_T$ is a function of the volatility weighted average of the accumulated portfolio value. This is because that value is a sufficient statistic for the unknown parameter θ, as first pointed out in the moral hazard context by Holmstrom [16].[3]

(iv) Our framework is a pure adverse selection framework without moral hazard, in the sense that v is observable, as a quadratic variation of the observable process X. In [26] it is assumed that only X_T is observable (in addition to the underlying stock prices), but, nevertheless, the first-best efficiency is attained, since the first-best optimal contract depends only on X_T in [26] anyway.

Example 5.1 Consider the case $U_1(x) = \log x$. If $L = 0$, for example, we get that, omitting dependence on θ,

$$C_T = -\frac{1}{v} M_T^\theta(\lambda + \mu B_T)\mathbf{1}_{\{\lambda+\mu B_T<0\}}.$$

Notice that this contract is not linear in X_T, B_T, unlike the case of controlled drift, in which the optimal contract is given by (4.4) (if the principal is risk-neutral). The nonlinearity has two causes: $\mu(\theta)$ may be different from zero (binding truth-telling constraint), and M_T^θ may be a nonlinear function of X_T (this depends on the principal's utility).

Next, we would have to compute $E^\theta[U_1(C_T(\theta))]$ and $E^\theta[B_T U_1(C_T(\theta))]$ in terms of the normal distribution function, in order to get a system of two nonlinear equations in λ and μ. However, this is hard. Suppose now that we have the case in which the agent/manager re-balances money between a risk-free asset with interest rate zero, and a single risky asset with price $S_t = e^{-t\sigma^2/2+\sigma B_t}$ (the Black–Scholes model). Even though we are unable to compute λ and μ, we can see what the shape of the optimal contract is, as a function of the underlying risky asset S. It can easily be checked that

$$B_T = \frac{1}{\sigma}\log(S_T) + \frac{1}{2}\sigma^2 T$$
$$M_T^\theta = e^{-\frac{1}{2}\theta^2 T + \frac{1}{2}\sigma^2 T}(S_T)^{\frac{\theta}{\sigma}}$$

This gives, omitting dependence on θ,

$$C_T = \frac{1}{v}(S_T)^{\frac{\theta}{\sigma}}\left[-\frac{\mu}{\sigma}\log(S_T) - \lambda - \frac{\mu}{2}\sigma^2 T\right]e^{-\frac{1}{2}\theta^2 T+\frac{1}{2}\sigma^2 T}\mathbf{1}_{\{\lambda+\frac{\mu}{\sigma}\log(S_T)+\frac{\mu}{2}\sigma^2 T<0\}}$$

It can be verified that for high types, that is, high values of the asset's risk premium θ, this is an increasing convex function of S_T, and it is concave for low types. This is not with the intention to penalize the low types, since a similar functional form holds true

[3] We thank the referee for pointing out this intuition.

even when the agent invests for herself, not for the principal. In fact, in that case we would simply have $\mu = 0$ and (see, for example, [22])

$$C_T = -\frac{\lambda}{\nu} M_T^\theta$$

It can then be checked that there is no $\log(S_T)$ term, which means that the functional form of C_T is concave if $\theta/\sigma < 1$, while here it is concave if $\theta/\sigma < c^*$, where $c^* < 1$, so that fewer agents are considered low type. Notice also that the payoff is zero for large values of S_T, unlike in the first-best case. This is necessary in order to induce truth-telling, otherwise, because of concavity vs. convexity of the payoff, the lower type agents would pretend they were high type.

5.1 Comparison with the first-best

We now consider the pure adverse selection model

$$dX_t = \theta v_t dt + v_t dB_t$$

where everything is observable. This model was considered in CCZ [6]; also in Ou-Yang [26], but with only X_T observable (see Remark 5.2, though) . We recall some results from those papers.

Denote

$$Z_t^\theta = e^{-t\theta^2/2 - \theta B_t}.$$

We have the budget constraint $E[Z_t^\theta X_t] = x$. Similarly as in CCZ [6], it follows that the first order conditions are

$$X_T - C_T(\theta) = I_2(\nu(\theta) Z_T^\theta)$$
$$C_T(\theta) = L \vee I_1(\mu(\theta) Z_T^\theta)$$

where $\nu(\theta)$ and $\mu(\theta)$ are determined so that $E[Z_T^\theta X_T] = x$ and the IR constraint is satisfied.

We see that the contract is of the similar form as the one we obtain for the relaxed problem in the adverse selection case, except that in the latter case there is an additional randomness in determining when the contract is above its lowest possible level L; see (5.6). With adverse selection, for the contract to be above L, we need, in addition to the first-best case requirements, that $\lambda + \mu B_T$ is small enough, which is the same as small values of $\lambda + \mu \int_0^T \frac{1}{v_t} dX_t$ or the small values of $\lambda + \mu(B_T^\theta + \theta T)$. Thus, the contract depends on average values of X normalized by volatility, equivalently on return plus noise.

In the first best case the ratio of marginal utilities U_2'/U_1' is constant, if $C_T > L$. In the adverse selection relaxed problem, we have, omitting dependence on θ,

$$\frac{U_2'(X_T - \hat{C}_T)}{U_1'(\hat{C}_T)} = -\mathbf{1}_{\{\hat{C}_T > L\}}[\lambda + \mu B_T] + \mathbf{1}_{\{\hat{C}_T = L\}} \frac{\nu}{U_1'(L) M_T^\theta}$$

where $\hat{C}_T$ is given in (5.6). We see that this ratio may be random, as in the case of controlled drift, but it is no longer linear in X (or B_T).

In the first best case (see CCZ [6]), the optimal contract is not unique, and it is also optimal to offer the contract

$$C_T = X_T - I_2(\nu Z_T^\theta). \tag{5.7}$$

Not only that, but this contract is incentive in the sense that it will force the agent to implement the first-best action process v, without the principal telling her what to do. This is *not* the case with adverse selection, in which the agent is given the contract payoff (5.6). As noted above, payoff $\hat{C}_T$ does not depend on the output X, but only on the exogenous random variable B_T and the agent's type θ, so that the payoff does not depend on the agent's performance, hence it is not incentive in the sense that the agent is indifferent what action v_t she will use. Thus, the principal should tell the agent which v_t to use, or alternatively, we can assume that the agent will follow the principal's best interest.

If we tried a different contract which would be incentive, similar to the first-best contract (5.7), for example if we assumed that the agent is given a symmetric benchmark contract, as is typical in portfolio management,

$$C_T = \alpha X_T - \beta_T$$

where β_T is a benchmark random variable, the following would happen: assuming the agent can choose v freely, it can be shown that at the optimum for the principal we have $\alpha \to 0$, and $\beta_T \to \hat{C}_T$, where $\hat{C}_T$ is given by (5.6).

To recap, our analysis indicates that in the "portfolio management" model of this section, in which the portfolio strategy v_t of the manager is observed, but the expected return θ on the managed portfolio is unobserved by the principal, (while known by the manager), a non-incentive, non-benchmark (but random) payoff $\hat{C}_T$ is optimal (providing that the solution of the relaxed problem is indeed the solution to the original problem). That payoff depends on the underlying source of risk B_T.

6 Conclusions

We consider several models of adverse selection with dynamic actions, with control of the return, and with control of the volatility. The problem can be transformed into a calculus of variations problem on choosing the optimal expected utility for the agent. When only the drift is controlled and the cost on the control of the return is quadratic, the optimal contract is a function of the final output value (typically nonlinear). When the volatility is controlled, the optimal contract is a non-incentive random payoff. The article Admati and Pfleiderer [1] argues against the use of benchmarks when rewarding portfolio managers, in favor of using contracts which depend only on the value of the output process. While our optimal contracts are not of a typical benchmark type, the payment, in addition to being the function of the underlying output, also depends on whether the driving noise process happened to have a high or a low value. In specific simple examples in our model, this is equivalent to the underlying risky assets attaining a high or a low value, and thus, there is a role for a "benchmark" in compensation, even though not by comparing the managed output to the benchmark output. Comparing to CWZ [8] and CCZ [6], we see that this extra randomness comes from the adverse selection effect.

We do not model here the possibility of a continuous payment or a payment at a random time chosen by the agent, or the possibility of renegotiating the contract in the future. Moreover, we do not have general results on how far away from optimal are specific truth-telling contracts that we identify. These and other timing issues would be of significant interest for future research.

7 Appendix

Proof of Theorem 4.1. By the assumed extended version of Assumption 2.2 and by (4.2), we can write the optimal contract $C_T(\theta) = \tilde{H}(X_T, W(\theta))$, where $\tilde{H}$ is a deterministic function defined by (2.25). For notational simplicity at below we denote $H := \tilde{H}$. Then

$$e^{V(\theta,\tilde{\theta})} = E\{M_T^\theta \varphi(X_T, W(\tilde{\theta}))\}; \quad \varphi(x,z) := e^{U_1(H(x,z))}.$$

If λ is convex and μ is linear, then $W(\theta)$ is convex as a function of θ. We claim that

$$\varphi_z < 0; \quad \varphi_{zz} < 0. \tag{7.8}$$

If so, then for any $\theta_1, \theta_2 \in [\theta_L, \theta_H]$ and any $\varphi \in [0,1]$,

$$\begin{aligned}\varphi(X_T, W(\alpha\theta_1 + (1-\alpha)\theta_2)) &\geq \varphi(X_T, \alpha W(\theta_1) + (1-\alpha)W(\theta_2)) \\ &\geq \alpha\varphi(X_T, W(\theta_1)) + (1-\alpha)\varphi(X_T, W(\theta_2)).\end{aligned}$$

That implies that $e^{V(\theta,\tilde{\theta})}$ is concave in $\tilde{\theta}$. By the first order condition, we get

$$e^{V(\theta,\theta)} = \max_{\tilde{\theta}} e^{V(\theta,\tilde{\theta})}.$$

Therefore, we have truth-telling.

It remains to prove (7.8). Note that

$$\varphi_z = \varphi U_1' H_z; \quad \varphi_{zz} = \varphi\Big[|U_1' H_z|^2 + U_1''|H_z|^2 + U_1' H_{zz}\Big];$$

Recall that

$$U_2(x - H(x,z)) - \frac{U_2'(x - H(x,z))}{U_1'(H(x,z))} = z.$$

Then

$$H_z = \frac{|U_1'|^2}{-|U_1'|^2 U_2' + U_1' U_2'' + U_1'' U_2'} < 0.$$

Thus, $\varphi_z < 0$.

Moreover,

$$H_{zz} = I_z^2\left[\frac{2U_1''}{U_1'} + \frac{-2U_1'U_1''U_2' + |U_1'|^2 U_2'' - U_1'U_2''' + U_1'''U_2'}{|U_1'|^2 U_2' - U_1'U_2'' - U_1''U_2'}\right].$$

So

$$\begin{aligned}\varphi_{zz} &= \varphi H_z^2\left[|U_1'|^2 + U_1'' + 2U_1'' + U_1'\frac{-2U_1'U_1''U_2' + |U_1'|^2 U_2'' - U_1'U_2''' + U_1'''U_2'}{|U_1'|^2 U_2' - U_1'U_2'' - U_1''U_2'}\right] \\ &= -\frac{\varphi H_z^3}{|U_1'|^2}\Big[[|U_1'|^2 + 3U_1''][|U_1'|^2 U_2' - U_1'U_2'' - U_1''U_2'] \\ &\qquad + U_1'[-2U_1'U_1''U_2' + |U_1'|^2 U_2'' - U_1'U_2''' + U_1'''U_2']\Big] \\ &= -\frac{\varphi H_z^3}{|U_1'|^2}\Big[[|U_1'|^4 + U_1'U_1''' - 3|U_1''|^2]U_2' - 3U_1'U_1''U_2'' - |U_1'|^2 U_2'''\Big].\end{aligned}$$

By (4.1) we get $\varphi_{zz} \leq 0$. □

References

1. Admati, A.R., Pfleiderer, P.: Does it all add up? Benchmarks and the compensation of active portfolio managers. J. Bus. **70**, 323–350 (1997)
2. Arora, N., Ou-Yang, H.: A Continuous-time model of explicit and implicit incentives. Working paper, University of North Carolina (2000)
3. Baron, D., Besanko, D.: Monitoring, moral hazard, asymmetric information, and risk sharing in procurement contracting. Rand J. Econ. **18**, 509–532 (1987)
4. Baron, D., Holmstrom, B.: The investment banking contract for new issues under asymmetric information: delegation and the incentive problem. J. Finance **35**, 1115–1138 (1980)
5. Bolton, P., Dewatripont, M.: Contract theory. MIT, Cambridge (2005)
6. Cadenillas, A., Cvitanić, J., Zapatero, F.: Optimal risk-sharing with effort and project choice. J. Econ. Theory **133**, 403–440 (2007)
7. Cvitanić, J., Wan, X., Zhang, J.: Continuous-time principal- agent problems with hidden action and lump-sum payment. Working paper (2005)
8. Cvitanić, J., Wan, X., Zhang, J.: Optimal contracts in continuous-time models. J. Appl. Math. Stoch. Anal. (2006) Article ID 95203
9. Darrough, M.N., Stoughton, N.M.: Moral hazard and adverse selection: the question of financial structure. J. Finance XLI **2**, 501–513 (1986)
10. Davis, M.H.A., Varaiya, P.P.: Dynamic programming conditions for partially-observable stochastic systems. SIAM J. Control **11**, 226–261 (1973)
11. DeMarzo, P., Sannikov, Y.: A Continuous-time agency model of optimal contracting and capital structure. Working paper (2004)
12. Detemple, J., Govindaraj, S., Loewenstein, M.: Hidden actions, agents with non-separable utility and wealth-robust intertemporal incentive contracts. Working paper, Boston University (2001)
13. Dybvig, P., Farnsworth, H., Carpenter, J.: Portfolio performance and agency. Working paper, Washington University in St. Louis (2001)
14. Heinkel, R., Stoughton, N.M.: The dynamics of portfolio management contracts. Rev. Financ. Stud. **7**(2), 351–387 (1994)
15. Hellwig, M., Schmidt, K.M.: Discrete-time approximations of Holmstrom- Milgrom Brownian-motion model of intertemporal incentive provision. Working paper, University of Mannheim (1998)
16. Holmstrom, B.: Moral hazard and observability. Bell J. Econ. **10**, 74–91 (1979)
17. Holmstrom, B., Milgrom, P.: Aggregation and linearity in the provision of intertemporal incentives. Econometrica **55**, 303–328 (1987)
18. Hugonnier, J., Kaniel, R.: Mutual fund portfolio choice in the presence of dynamic flows. Working paper, University of Lausanne (2001)
19. Kadan, O., Swinkels, J.: Moral Hazard with bounded payments. Working paper (2005a)
20. Kadan, O., Swinkels, J.: Stocks or Options? Moral Hazard, Firm Viability and the Design of Compensation Contracts. Working paper (2005b)
21. Kamien, M.I., Schwartz, N.L.: Dynamic optimization. Elsevier, Amsterdam (1991)
22. Karatzas, I., Shreve, S.E.: Methods of mathematical finance. Springer, New York (1998)
23. McAfee, R.P., McMillan, J.: Bidding for contracts: a principal-agent analysis. Rand J. Econ. **17**, 326–338 (1986)
24. Muller, H.: The first-best sharing rule in the continuous-time principal-agent problem with exponential utility. J. Econ. Theory **79**, 276–280 (1998)
25. Muller, H.: Asymptotic efficiency in dynamic principal-agent problems. J. Econ. Theory **91**, 292–301 (2000)
26. Ou-Yang, H.: Optimal contracts in a continuous-time delegated portfolio management problem. Rev. Financ. Stud. **16**, 173–208 (2003)
27. Sannikov, Y.: A Continuous-time version of the principal-agent problem. working paper, UC Berkeley (2004)
28. Schattler, H., Sung, J.: The first-order approach to continuous-time principal-agent problem with exponential utility. J. Econ. Theory **61**, 331–371 (1993)
29. Schattler, H., Sung, J.: On optimal sharing rules in discrete- and continuous- times principal-agent problems with exponential utility. J. Econ. Dyn. Control **21**, 551–574 (1997)

30. Sung, J.: Linearity with project selection and controllable diffusion rate in continuous-time principal-agent problems. Rand J. Econ. **26**, 720–743 (1995)
31. Sung, J.: Corporate insurance and managerial incentives. J. Econ. Theory **74**, 297–332 (1997)
32. Sung, J.: Optimal contracts under adverse selection and moral hazard: a continuous-time approach. Rev. Financ. Stud. **18**, 1021–1073 (2005)
33. Williams, N.: On dynamic principal-agent problems in continuous time. Working paper, Princeton University (2004)