Chapter 6
Special Cases and Applications

Abstract We present here well-known examples and applications of continuous-time Principal–Agent models. The seminal work of Holmström and Milgrom (Econometrica 55:303–328, 1987) is the first to use a continuous-time model, showing that doing that can, in fact, lead to simple, while realistic optimal contracts. In particular, if the principal and the agent maximize expected utility from terminal output value, and have non-separable cost of effort and exponential utilities, the optimal contract is linear in that value. With other utilities and separable cost of effort, the optimal contract is nonlinear in the terminal output value, obtained as a solution to a nonlinear equation that generalizes the first best Borch condition. In the case of the agent deriving utility from continuous contract payments on an infinite horizon, and if the principal is risk-neutral, the problem reduces to solving an ordinary differential equation for the principal's expected utility process as a function of the agent's expected utility process. That equation can then be solved numerically for various cases, including the case in which the agent can quit, or be replaced by another agent, or be trained and promoted. These cases are analyzed by studying the necessary conditions in terms of an FBSDE system for the agent's problem, and, in Markovian models, by identifying sufficient conditions in terms of the HJB differential equation for the principal's problem.

6.1 Exponential Utilities and Lump-Sum Payment

We now present a model which is an extension of the one from the seminal paper Hölmstrom and Milgrom (1987). For simplicity of notation, as we have done so far, we assume we have a one-dimensional Brownian motion.

6.1.1 The Model

We have, as usual,

$$dX_t = u_t v_t dt + v_t dB_t^u.$$

J. Cvitanić, J. Zhang, *Contract Theory in Continuous-Time Models*, Springer Finance,
DOI 10.1007/978-3-642-14200-0_6,

We assume that the agent is paid only at the final time T in the amount C_T, and the utilities are exponential: the principal maximizes

$$U_P(X_T, C_T) = U_P(X_T - C_T) := -e^{-\gamma_P(X_T - C_T)}$$

and the agent maximizes

$$U_A(C_T - G_T) := -e^{-\gamma_A(C_T - G_T)} \quad \text{with } G_t := \int_0^t \big[\mu_s X_s + g(s, u_s, v_s)\big] ds,$$

for some deterministic function of time μ_t. We consider only the participation constraint at time zero:

$$W_0^A \geq R_0.$$

We also allow the principal to choose the volatility process v.

6.1.2 Necessary Conditions Derived from the General Theory

Note to the Reader The reader not interested in the use of general theory of Chap. 5 can skip this section and go to the following section that provides a more direct approach for dealing with the above model.

In this subsection we derive the necessary conditions formally from the general theory established in the previous chapter.

Recalling (5.107), we have

$$\begin{aligned} g &= \mu x + g(t, u, v), \qquad u_A = u_P = 0, \\ U_A &= -e^{-\gamma_A(C_T - G_T)}, \qquad U_P = -e^{-\gamma_P(X_T - C_T)}. \end{aligned} \tag{6.1}$$

We first study the agent's problem. In this case, (5.110) becomes

$$\Gamma^A = 1, \qquad \bar{Y}_t^A = -\gamma_A e^{-\gamma_A(C_T - G_T)} - \int_t^T \bar{Z}_s^A dB_s^u.$$

Comparing this with (5.107), one can easily see that

$$\bar{Y}^A = \gamma_A W^A, \qquad \bar{Z}^A = \gamma_A Z^A.$$

Thus (5.111) becomes

$$Z_t^A + \gamma_A W_t^A \partial_u g(t, u_t, v_t) = 0. \tag{6.2}$$

Note that $W^A < 0$. Denote

$$\tilde{W}_t^A := -\frac{1}{\gamma_A} \ln\big[-W_t^A\big] + G_t, \qquad \tilde{Z}^A := -\frac{Z^A}{\gamma_A W^A}. \tag{6.3}$$

Then,

$$\tilde{W}_t^A = C_T - \int_t^T \left[\frac{\gamma_A}{2}\big(\tilde{Z}_s^A\big)^2 + \mu_s X_s + g(s, u_s, v_s)\right] ds - \int_t^T \tilde{Z}_s^A dB_s^u, \tag{6.4}$$

and (6.2) becomes

$$\tilde{Z}_t^A = \partial_u g(t, u_t, v_t). \tag{6.5}$$

Assume this uniquely determines a function I^A such that

$$u_t = I^A\big(t, v_t, \tilde{Z}_t^A\big). \tag{6.6}$$

Then, FBSDE (5.113) becomes

$$\begin{aligned} X_t &= x + \int_0^t v_s dB_s; \\ \tilde{W}_t^A &= C_T - \int_t^T \bigg[\frac{\gamma_A}{2}\big(\tilde{Z}_s^A\big)^2 + \mu_s X_s + g\big(s, I^A\big(s, v_s, \tilde{Z}_s^A\big), v_s\big) \\ &\qquad - \tilde{Z}_s^A I^A\big(s, v_s, \tilde{Z}_s^A\big)\bigg]ds - \int_t^T \tilde{Z}_s^A dB_s. \end{aligned}$$

This is a decoupled FBSDE that, under certain technical conditions, solves the agent's problem.

We now turn to the principal's problem. Given the principal's target action u, by (6.5) and (6.7) we have

$$\tilde{Z}_t^A = g_u(t, u_t, v_t), \qquad C_T = \tilde{W}_T^A. \tag{6.7}$$

Let w^A denote the agent's initial utility W_0^A, and

$$\tilde{R}_0 := -\frac{1}{\gamma_A}\ln[-R_0], \qquad \tilde{w}^A := -\frac{1}{\gamma_A}\ln\big[-w^A\big]. \tag{6.8}$$

Then, (5.116) becomes

$$\begin{aligned} X_t &= x + \int_0^t v_s u_s ds + \int_0^t v_s dB_s^u; \\ \tilde{W}_t^A &= \tilde{w}_A + \int_0^t \bigg[\frac{\gamma_A}{2}\big[g_u(s, u_s, v_s)\big]^2 + \mu_s X_s + g(s, u_s, v_s)\bigg]ds \\ &\qquad + \int_0^t g_u(s, u_s, v_s) dB_s^u; \\ W_t^P &= -\exp\big(-\gamma_P\big[X_T - \tilde{W}_T^A\big]\big) - \int_t^T Z_s^P dB_s^u \end{aligned} \tag{6.9}$$

and the IR constraint is $\tilde{w}^A \geq \tilde{R}_0$. It is clear that $\tilde{W}^A$ is increasing in $\tilde{w}^A$, and thus W^P is decreasing in $\tilde{w}^A$. Therefore, the principal chooses $\tilde{w}^A = \tilde{R}_0$ and then principal's problem (5.117) becomes

$$V_P := \sup_{(u,v)} W_0^P. \tag{6.10}$$

In this case, (5.118) becomes:

$$\begin{aligned}
&\Gamma^1 = 1, \qquad \Gamma^2 = 0, \qquad \bar{Y}^2 = \bar{Z}^2 = 0, \qquad \bar{Y}^3 = \bar{Z}^3 = 0,\\
&\bar{Y}_t^1 = \gamma_P \exp\big(-\gamma_P\big[X_T - \tilde{W}_T^A\big]\big) + \int_t^T \mu_s \bar{Y}_s^4 ds - \int_t^T \bar{Z}_s^1 dB_s^u, \qquad (6.11)\\
&\bar{Y}_t^4 = -\gamma_P \exp\big(-\gamma_P\big[X_T - \tilde{W}_T^A\big]\big) - \int_t^T \bar{Z}_s^4 dB_s^u,
\end{aligned}$$

and (5.119) leads to

$$\begin{aligned}
Z^P + \bar{Y}^4\left[g_u + \frac{1}{\gamma_A} g_u g_{uu} + \frac{1}{\gamma_A} g_u\right] - \frac{1}{\gamma_A}\bar{Z}^4 g_{uu} = 0;\\
\bar{Y}^1 u + \bar{Z}^1 + \bar{Y}^4\left[g_v + \frac{1}{\gamma_A} g_u g_{uv}\right] - \frac{1}{\gamma_A}\bar{Z}^4 g_{uv} = 0.
\end{aligned} \qquad (6.12)$$

It is clear that

$$\bar{Y}^4 = \gamma_P W^P, \qquad \bar{Z}^4 = \gamma_P Z^P.$$

Moreover, denote

$$\hat{Y} := \frac{\bar{Y}^1}{W^P}, \qquad \hat{Z} := \frac{\bar{Z}^1 - \hat{Y} Z^P}{W^P}.$$

We have

$$\hat{Y}_t = -\gamma_P + \int_t^T \left[\gamma_P \mu_s + \frac{Z_s^P}{W_s^P}\hat{Z}_s\right] ds - \int_t^T \hat{Z}_s dB_s^u.$$

Since μ is deterministic, we get

$$\hat{Y}_t = \gamma_P [M_T - M_t - 1], \qquad \hat{Z}_t = 0 \quad \text{where } M_t := \int_0^t \mu_s ds. \qquad (6.13)$$

Then, (6.12) becomes:

$$\begin{aligned}
Z^P + \gamma_P W^P\left[g_u + \frac{1}{\gamma_A} g_u g_{uu} + \frac{1}{\gamma_A} g_u\right] - \frac{\gamma_P}{\gamma_A} Z^P g_{uu} = 0;\\
\hat{Y} W^P u + \hat{Y} Z^P + \gamma_P W^P\left[g_v + \frac{1}{\gamma_A} g_u g_{uv}\right] - \frac{\gamma_P}{\gamma_A} Z^P g_{uv} = 0.
\end{aligned} \qquad (6.14)$$

We finally solve (6.9), (6.13), and (6.14). Denote

$$\hat{W}_t^A := \tilde{W}_t^A - M_t X_t. \qquad (6.15)$$

Then,

$$\begin{aligned}
\hat{W}_t^A = \tilde{R}_0 &+ \int_0^t \left[\frac{\gamma_A}{2}\big[g_u(s, u_s, v_s)\big]^2 + g(s, u_s, v_s) - M_s v_s u_s\right] ds\\
&+ \int_0^t \big[g_u(s, u_s, v_s) - M_s v_s\big] dB_s^u.
\end{aligned}$$

Also denote

$$\tilde{W}_t^P := -\frac{1}{\gamma_P}\ln\big(-W_t^P\big), \qquad \tilde{Z}_t^P := -\frac{Z_t^P}{\gamma_P W_t^P}. \tag{6.16}$$

Then,

$$\tilde{W}_t^P = X_T - \tilde{W}_T^A - \int_t^T \frac{\gamma_P}{2}\big|\tilde{Z}_s^P\big|^2 ds - \int_t^T \tilde{Z}_s^P dB_s^u,$$

and (6.14) becomes

$$\begin{aligned} [\gamma_A - \gamma_P g_{uu}]\tilde{Z}^P &= \gamma_A g_u + g_u g_{uu} + g_u; \\ [\gamma_A \hat{Y} - \gamma_P g_{uv}]\tilde{Z}^P &= \frac{\gamma_A}{\gamma_P}\hat{Y}u + \gamma_A g_v + g_u g_{uv}. \end{aligned} \tag{6.17}$$

Note that

$$X_T - \tilde{W}_T^A = (1 - M_T)X_T - \hat{W}_T^A$$

and that M is deterministic. Denote

$$\hat{W}_t^P := \tilde{W}_t^P - (1 - M_T)X_t + \hat{W}_t^A. \tag{6.18}$$

Then,

$$\begin{aligned} \hat{W}_t^P = &-\int_t^T \Big[\frac{\gamma_P}{2}\big|\tilde{Z}_s^P\big|^2 - (1 - M_T)v_s u_s + \frac{\gamma_A}{2}\big[g_u(s, u_s, v_s)\big]^2 \\ &+ g(s, u_s, v_s) - M_s v_s u_s\Big] ds \\ &- \int_t^T \big[\tilde{Z}_s^P - (1 - M_T)v_s + g_u(s, u_s, v_s) - M_s v_s\big] dB_s^u. \end{aligned}$$

One solution for this and (6.17) is

$$\begin{aligned} \tilde{Z}_s^P &= (1 - M_T)v_s - g_u(s, u_s, v_s) + M_s v_s; \\ \hat{W}_t^P &= -\int_t^T \Big[\frac{\gamma_P}{2}\big|\tilde{Z}_s^P\big|^2 - (1 - M_T)v_s u_s + \frac{\gamma_A}{2}\big[g_u(s, u_s, v_s)\big]^2 \\ &\quad + g(s, u_s, v_s) - M_s v_s u_s\Big] ds; \end{aligned} \tag{6.19}$$

where u, v are deterministic and satisfy

$$\begin{aligned} [\gamma_A - \gamma_P g_{uu}]\big[(1 - M_T)v - g_u + Mv\big] &= \gamma_A g_u + g_u g_{uu} + g_u; \\ [\gamma_A \hat{Y} - \gamma_P g_{uv}]\big[(1 - M_T)v - g_u + Mv\big] &= \frac{\gamma_A}{\gamma_P}\hat{Y}u + \gamma_A g_v + g_u g_{uv}. \end{aligned} \tag{6.20}$$

Assume (6.20) determines uniquely deterministic functions (u^*, v^*). Solving (6.19) we obtain $\hat{W}_0^{P,*}$. Then, the principal's optimal utility is

$$W_0^{P,*} := -\exp\big(-\gamma_P \tilde{W}_0^{P,*}\big) = -\exp\big(-\gamma_P\big[\hat{W}_0^{P,*} + (1 - M_T)x - \tilde{R}_0\big]\big),$$

and the optimal contract is $C_T^* = \tilde{W}_T^{A,*}$, where the latter is defined by (6.9) with optimal control (u^*, v^*). In the following section we prove rigorously that the above solution is indeed optimal for the problem.

6.1.3 A Direct Approach

In this section we provide a direct approach for solving the problem, without using the results of Chap. 5. However, we use additional results from the BSDE theory. We start with the agent's problem. Given a pair (C_T, v), the agent's utility process is given by

$$W_t^A = U_A(C_T - G_T) - \int_t^T Z_s^A dB_s^u.$$

Recall the agent's certainty equivalent process $\tilde{W}^A$ and the corresponding process $\tilde{Z}^A$ as defined in (6.3):

$$\tilde{W}_t^A := -\frac{1}{\gamma_A}\ln\bigl[-W_t^A\bigr] + G_t, \qquad \tilde{Z}^A := -\frac{Z^A}{\gamma_A W^A}. \tag{6.21}$$

We have the following result.

Proposition 6.1.1 *Assume, for a given pair (C_T, v), that the admissible set for u is such that BSDE (6.23) below is well-posed and satisfies the BSDE comparison principle (as stated in Part V of the book). Then, the necessary and sufficient condition for the agent's optimal effort is*

$$u_t = I_A\bigl(t, v_t, \tilde{Z}^A\bigr) := \operatorname*{argmin}_u\bigl[g(t,u,v_t) - u\tilde{Z}_t^A\bigr]. \tag{6.22}$$

Proof Note that $\tilde{W}_0^A = -\frac{1}{\gamma_A}\ln(-W_0^A)$. Then, the optimization of the agent's utility W_0^A is equivalent to the optimization of $\tilde{W}_0^A$. By (6.4), or by applying Itô's rule directly, we get

$$\tilde{W}_t^A = C_T - \int_t^T \left[\frac{\gamma_A}{2}\bigl(\tilde{Z}_s^A\bigr)^2 + \mu_s X_s + g(s,u_s,v_s) - u_s\tilde{Z}_s^A\right]ds - \int_t^T \tilde{Z}_s^A dB_s. \tag{6.23}$$

By the comparison principle for BSDEs we see that the optimal u is obtained by minimizing the integrand in the first integral in the previous expression, which completes the proof. □

Remark 6.1.2 BSDE (6.23) has quadratic growth in $\tilde{Z}^A$. When C_T is bounded, we prove the well-posedness and the comparison principle for such BSDEs in Sect. 9.6. However, C_T corresponding to the optimal contract in Theorem 6.1.3 below is in general not bounded. Instead we can use the comparison theorem from Briand and Hu (2008). In order to apply that theorem, we need to assume that we only allow actions (u, v) and contracts C_T such that

$$E\bigl[e^{\lambda \sup_{0\le t\le T}|\tilde{W}_t^A|}\bigr] < \infty, \quad \forall\lambda > 0$$

and the random variable

$$\int_0^T \bigl|\mu_t X_t + g(t,u_t,v_t) - u_t\tilde{Z}_t^A\bigr| dt$$

has exponential moments of all orders.

We now turn to the principal's problem. Assume g is differentiable in u and the optimal u is in the interior of the admissible set. Then, (6.22) leads to

$$Z_t^A + \gamma_A W_t^A \partial_u g(t, u_t, v_t) = 0 \tag{6.24}$$

and thus we get, also using (6.23),

$$\tilde{Z}_t^A = g_u(t, u_t, v_t), \qquad C_T = \tilde{W}_T^A. \tag{6.25}$$

Denote

$$\tilde{R}_0 := -\frac{1}{\gamma_A} \ln[-R_0], \qquad W_0^A = w^A, \qquad \tilde{w}^A := -\frac{1}{\gamma_A} \ln\big[-w^A\big]. \tag{6.26}$$

Then, we can write

$$\begin{aligned} X_t &= x + \int_0^t v_s u_s ds + \int_0^t v_s dB_s^u; \\ \tilde{W}_t^A &= \tilde{w}_A + \int_0^t \left[\frac{\gamma_A}{2}\big[g_u(s, u_s, v_s)\big]^2 + \mu_s X_s + g(s, u_s, v_s)\right] ds \\ &\quad + \int_0^t g_u(s, u_s, v_s) dB_s^u; \\ W_t^P &= -\exp\big(-\gamma_P\big[X_T - \tilde{W}_T^A\big]\big) - \int_t^T Z_s^P dB_s^u. \end{aligned} \tag{6.27}$$

As in Sect. 5.2.3, instead of using contract payment C_T as the principal's control, we use the corresponding agent's optimal action u as the principal's control. Given a principal's "target action" u, the volatility control v, and the agent's initial utility $w^A \geq R_0$, the corresponding contract is $C_T = \tilde{W}_T^A$. Clearly, $\tilde{W}^A$ is increasing in $\tilde{w}^A$, and so W^P is decreasing in $\tilde{w}^A$. Thus, the principal chooses $\tilde{w}^A = \tilde{R}_0$ and faces the problem

$$V_P := \sup_{(u,v)} W_0^P. \tag{6.28}$$

The solution is given by the following result:

Theorem 6.1.3 *Consider the function*

$$\begin{aligned} L(t, u_t, v_t) := &\frac{\gamma_P}{2}\big[(1 - M_T + M_t)v_t - g_u(t, u_t, v_t)\big]^2 - (1 - M_T + M_t)u_t v_t \\ &+ \frac{\gamma_A}{2}\big|g_u(t, u_t, v_t)\big|^2 + g(t, u_t, v_t), \end{aligned} \tag{6.29}$$

where M is defined by

$$M_t := \int_0^t \mu_s ds.$$

Assume that, for every t, there exists a pair (u_t^*, v_t^*) minimizing this expression, and such that $\int_0^T L(t, u_t^*, v_t^*)dt$ is finite. Then, the deterministic controls (u_t^*, v_t^*) are optimal for the principal's problem. The optimal contract payoff is given by

$$C_T^* = c + \int_0^T \left[M_T - M_t + \frac{g_u(t, u_t^*, v_t^*)}{v_t^*} \right] dX_t^* \tag{6.30}$$

for a constant c chosen so that the agent's expected utility is equal to his reservation value R_0. In particular, if $g_u(t, u_t^*, v_t^*)/v_t^* + M_T - M_t$ is a constant, then contract C_T^* is linear in X_T^*.

Proof Doing integration by parts we get the following representation for the first part of the cost G_T:

$$\int_0^T \mu_t X_t dt = X_T M_T - \int_0^T M_t \big[u_t v_t dt + v_t dB_t^u \big]. \tag{6.31}$$

Then, by (6.27) and $\tilde{W}_0^A = \tilde{R}_0$, we see that we need to minimize

$$\begin{aligned} -W_0^P &= E^u\big[\exp\big(-\gamma_P\big[X_T - \tilde{W}_T^A\big]\big)\big] \\ &= E^u\Bigg[\exp\Bigg(-\gamma_P\Bigg[(1 - M_T)x - \tilde{R}_0 \\ &\qquad + \int_0^T \Bigg[(1 - M_T + M_t)u_t v_t - \left(\frac{\gamma_A}{2}|g_u|^2 + g\right)\Bigg]dt \\ &\qquad + \int_0^T \big[(1 - M_T + M_t)v_t - g_u\big]dB_t^u\Bigg]\Bigg)\Bigg]. \end{aligned} \tag{6.32}$$

This is a standard stochastic control problem, for which the solution, when it exists, turns out to be a pair of deterministic processes (u^*, v^*). (This can be verified, once the solution is found, by verifying the corresponding HJB equation.) Assuming that u, v are deterministic, the expectation above can be computed by using the fact that

$$E^u\left[\exp\left(\int_0^T f_s dB_s^u\right)\right] = \exp\left(\frac{1}{2}\int_0^T f_s^2 ds\right)$$

for a given square-integrable deterministic function f. Then,

$$\begin{aligned} -W_0^P &= \exp\Bigg(-\gamma_P\Bigg[(1 - M_T)x - \tilde{R}_0 + \int_0^T \Bigg[(1 - M_T + M_t)u_t v_t \\ &\qquad - \left(\frac{\gamma_A}{2}|g_u|^2 + g\right)\Bigg]dt\Bigg] + \frac{1}{2}\gamma_P^2 \int_0^T \big[(1 - M_T + M_t)v_t - g_u\big]^2 dt\Bigg) \\ &= \exp\left(-\gamma_P\big[(1 - M_T)x - \tilde{R}_0\big] + \gamma_P \int_0^T L(t, u_t, v_t)dt\right). \end{aligned}$$

Thus, the minimization can be done inside the integral in the exponent, and boils down to minimizing $L(t, u_t, v_t)$ over (u_t, v_t), which proves the first part of the theorem.

The optimal contract is found from $C_T^* = \tilde{W}_T^{A,*}$. Note that (6.31) is equivalent to

$$\int_0^T \mu_t X_t dt = \int_0^T [M_T - M_t] dX_t$$

and that

$$\int_0^T g_u\big(t, u_t^*, v_t^*\big) dB_t^{u^*} = \int_0^T \frac{g_u(t, u_t^*, v_t^*)}{v_t^*} dX_t^* - \int_0^T g_u\big(t, u_t^*, v_t^*\big) u_t^* dt.$$

Plugging these into (6.27) we obtain (6.30). □

Remark 6.1.4 Assume the functions below are smooth enough and the optimal controls (u^*, v^*) are in the interior of the admissible set. Then, the minimization of $L(t, u, v)$ leads to

$$\begin{aligned}
&-\gamma_P\big[(1 - M_T + M_t)v_t^* - g_u\big]g_{uu} - (1 - M_T + M_t)v_t^* + \gamma_A g_u g_{uu} + g_u = 0;\\
&\gamma_P\big[(1 - M_T + M_t)v_t^* - g_u\big][1 - M_T + M_t - g_{uv}] - (1 - M_T + M_t)u_t^*\\
&\quad + \gamma_A g_u g_{uv} + g_v = 0.
\end{aligned}$$

One can check straightforwardly that this is equivalent to (6.20).

6.1.4 A Solvable Special Case with Quadratic Cost

Consider now the special case of Holmström–Milgrom (1987), with

$$\mu_t \equiv 0, \qquad v_t \equiv v, \qquad g(t, x, u, v) = (uv)^2/2.$$

Then, $g_u = v^2 u$ and the expression (6.29) becomes

$$L(t, u_t) := \frac{\gamma_P}{2}\big[v - v^2 u_t\big]^2 - u_t v + \frac{\gamma_A}{2}\big|v^2 u_t\big|^2 + \frac{1}{2}|v u_t|^2. \tag{6.33}$$

Minimizing this we get constant optimal u^* of Holmström–Milgrom (1987), given by

$$u^* = \frac{\frac{1}{v} + \gamma_P v}{1 + (\gamma_A + \gamma_P)v^2}.$$

The optimal contract is linear, and given by

$$C_T^* = c + \frac{1 + \gamma_P v^2}{1 + (\gamma_A + \gamma_P)v^2} X_T,$$

where c is such that the IR constraint is satisfied,

$$c = -\frac{1}{\gamma_A}\log(-R_0) - u^* v x + \frac{|u^* v|^2 T}{2}(\gamma_A - 1). \tag{6.34}$$

In particular, one prediction is that with lower uncertainty v, the "pay-per-performance" sensitivity (the slope) of the contract is higher; in fact, it is equal to 1 when $v = 0$: the principal turns over the whole firm to the agent when there is no risk.

6.2 General Risk Preferences, Quadratic Cost, and Lump-Sum Payment

6.2.1 The Model

Consider now the setting in which the participation constraint is imposed only at time zero, there is no intermediate consumption, just the lump sum payment C_T at the end, no volatility control, and the cost is quadratic:

$$u_A = u_P = 0, \qquad g = ku^2/2 \quad \text{for some constant } k. \tag{6.35}$$

Moreover, the agent's utility is separable in effort and contract payment, so that the model becomes

$$\begin{aligned} W_t^A &= U_A(C_T) - \int_t^T \frac{ku_s^2}{2} ds - \int_t^T Z_s^A dB_s^u, \\ W_t^P &= U_P(C_T) - \int_t^T Z_s^P dB_s^u \end{aligned} \tag{6.36}$$

and the IR constraint is

$$W_0^A \geq R_0. \tag{6.37}$$

6.2.2 Necessary Conditions Derived from the General Theory

Note to the Reader The reader not interested in the use of general theory of Chap. 5 can skip this section and go to the following section that provides a more direct approach for dealing with the above model.

As usual, we start with the agent's problem. In this case, by (5.88) we have

$$I_A(z) = \frac{z}{k}, \quad \text{and the agent's optimal control satisfies } u = \frac{1}{k} Z^A. \tag{6.38}$$

Consequently, given C_T, the agent's optimal utility process satisfies

$$\begin{aligned} W_t^A &= U_A(C_T) - \int_t^T \left[\frac{ku_s^2}{2} - u_s Z_s^A \right] ds - \int_t^T Z_s^A dB_s \\ &= U_A(C_T) + \int_t^T \frac{1}{2k} |Z_s^A|^2 ds - \int_t^T Z_s^A dB_s. \end{aligned} \tag{6.39}$$

Denote

$$\tilde{W}_t^A := e^{W_t^A/k}, \qquad \tilde{Z}_t^A := \frac{1}{k} \tilde{W}^A Z^A. \tag{6.40}$$

Applying Itô's rule, we get

$$\tilde{W}_t^A = e^{U_A(C_T)/k} - \int_t^T \tilde{Z}_s^A dB_s. \tag{6.41}$$

If $E[e^{2U_A(C_T)/k}] < \infty$, the above BSDE is well-posed, with the solution $\tilde{W}_t^A = E_t[e^{U_A(C_T)/k}]$, and we obtain the agent's optimal utility and optimal control:

$$W_t^A = k\ln\big(\tilde{W}_t^A\big) = k\ln\big(E_t\big[e^{U_A(C_T)/k}\big]\big), \qquad u_t = Z_t^A/k = \frac{\tilde{Z}_t^A}{\tilde{W}_t^A}. \tag{6.42}$$

We now turn to the principal's problem. As in Sect. 5.4.2, we take two different approaches, corresponding to Sects. 5.2.2 and 5.2.4, respectively. For the first approach, we consider the relaxed principal's problem

$$V_P(\lambda) := \sup_{C_T}\big[W_0^P + \lambda W_0^A\big], \quad \text{with } u = Z^A/k \text{ in (6.36)}. \tag{6.43}$$

The first equation in (5.89) gives us the optimality condition for C_T, that translates into $C_T = I_P(D_T)$, assuming the following inverse function exists:

$$I_P := \big[-U_P'/U_A'\big]^{-1}. \tag{6.44}$$

Recall (6.38) and, as in Theorem 5.4.1,

$$\tilde{W}_t^P := W_t^P - \lambda W_0^A. \tag{6.45}$$

Then, (5.91) becomes:

$$\begin{aligned}
D_t &= \lambda + \int_0^t \frac{1}{k} Z_s^P\left[dB_s - \frac{1}{k}Z_s^A ds\right];\\
W_t^A &= U_A\big(I_P(D_T)\big) - \int_t^T \frac{1}{2k}|Z_s^A|^2 ds - \int_t^T Z_s^A\left[dB_s - \frac{1}{k}Z_s^A ds\right];\\
\tilde{W}_t^P &= U_P\big(I_P(D_T)\big) - \int_t^T Z_s^P\left[dB_s - \frac{1}{k}Z_s^A ds\right].
\end{aligned} \tag{6.46}$$

Moreover, the principal's optimal utility and the optimal contract are

$$V_P(\lambda) = W_0^P = \tilde{W}_0^P + \lambda W_0^A, \qquad C_T = I_P(D_T). \tag{6.47}$$

Comparing the equations for D_t and $\tilde{W}_t^P$ in (6.46), we see that

$$D_t = \frac{1}{k}\tilde{W}_t^P + \tilde{\lambda}$$

for some constant $\tilde{\lambda}$. In particular,

$$D_T = \frac{1}{k}U_P(C_T) + \tilde{\lambda}.$$

This means, using (6.47), that the optimal C_T can be obtained from the following generalization of Borch's rule to the hidden action case:

$$\frac{U_P'(C_T)}{U_A'(C_T)} = -\frac{1}{k}U_P(C_T) - \tilde{\lambda}. \tag{6.48}$$

Assume the above equation determines uniquely

$$C_T = \xi(\tilde{\lambda}) \quad \text{for some random variable } \xi(\tilde{\lambda}). \tag{6.49}$$

We then have the BSDE system

$$\begin{aligned} W_t^A &= U_A\big(\xi(\tilde{\lambda})\big) - \int_t^T \frac{1}{2k}|Z_s^A|^2 ds - \int_t^T Z_s^A\left[dB_s - \frac{1}{k}Z_s^A ds\right]; \\ \tilde{W}_t^P &= U_P\big(\xi(\tilde{\lambda})\big) - \int_t^T Z_s^P\left[dB_s - \frac{1}{k}Z_s^A ds\right]. \end{aligned}$$

We have

$$\lambda = D_0 = \frac{1}{k}\tilde{W}_0^P + \tilde{\lambda} \quad \text{and thus } V_P(\lambda) = \tilde{W}_0^P + \left[\frac{1}{k}\tilde{W}_0^P + \tilde{\lambda}\right]W_0^A.$$

Finally, if we can find $\tilde{\lambda}^*$ such that the corresponding agent's initial wealth satisfies $W_0^{A,*} = R_0$, then we have

$$V_P = \tilde{W}_0^{P,*} + \left[\frac{1}{k}\tilde{W}_0^{P,*} + \tilde{\lambda}^*\right]R_0.$$

Remark 6.2.1 In this remark we assume that U_A is a deterministic function and

$$U_P(C_T) = \tilde{U}_P(X_T - C_T)$$

for some deterministic function $\tilde{U}_P$. Then, (6.48) is a nonlinear equation:

$$\frac{\tilde{U}_P'(X_T - C_T)}{U_A'(C_T)} = \frac{1}{k}U_P(X_T - C_T) + \tilde{\lambda} \tag{6.50}$$

and thus the optimal contract C_T is a function of the terminal value X_T only:

$$C_T = \Phi(X_T) \quad \text{for some deterministic function } \Phi. \tag{6.51}$$

For an economic discussion of this nonlinear equation see Remark 6.2.4 below.

We next study the principal's problem following the approach in Sect. 5.2.4. Let u be the principal's target action, $w^A \geq R_0$ be the agent's initial utility. Denote

$$J_A := \big(U_A'\big)^{-1} \quad \text{and} \quad \hat{U}_P := U_P(J_A). \tag{6.52}$$

Then, by (6.38) and (6.36),

$$Z^A = ku, \qquad C_T = J_A\big(W_T^A\big), \tag{6.53}$$

and thus

$$\begin{aligned} W_t^A &= w^A + \int_0^t \frac{ku_S^2}{2}ds + \int_0^t ku_s dB_s^u; \\ W_t^P &= \hat{U}_P\big(W_T^A\big) - \int_t^T Z_s^P dB_s^u. \end{aligned}$$

We assume the standard condition

$$U_A \text{ is increasing and concave,} \qquad U_P \text{ is decreasing and concave.} \tag{6.54}$$

Then, clearly W^A is increasing in w^A and W^P is decreasing in w^A. Thus, the principal would chooses $w^A = R_0$. Therefore, the principal's problem becomes

$$V_P := \sup_u W_0^P, \tag{6.55}$$

where

$$\begin{aligned} W_t^A &= R_0 + \int_0^t \frac{ku_S^2}{2} ds + \int_0^t ku_s dB_s^u; \\ W_t^P &= \hat{U}_P\big(W_T^A\big) - \int_t^T Z_s^P dB_s^u. \end{aligned} \tag{6.56}$$

In this case, (5.46) becomes

$$\begin{aligned} \Gamma_t &= 1 + \int_0^t \Gamma_s u_s dB_s; \\ \bar{Y}_t &= \hat{U}'_P\big(W_T^A\big)\Gamma_T - \int_t^T \bar{Z}_s dB_s, \end{aligned} \tag{6.57}$$

and the optimization condition (5.47) (see also (5.92)) becomes

$$\Gamma_t Z_t^P - [\bar{Y}_t u_t - \bar{Z}_t]k = 0. \tag{6.58}$$

Applying Itô's rule, we have

$$\begin{aligned} &d\big(\Gamma_t W_t^P + k\bar{Y}_t\big) \\ &\quad = -\Gamma_t Z_t^P u_t dt + \Gamma_t Z_t^P dB_t + W_t^P \Gamma_t u_t dB_t + Z^P \Gamma_t u_t dt + k\bar{Z}_t dB_t \\ &\quad = \big[\Gamma_t Z_t^P + W_t^P \Gamma_t u_t + k\bar{Z}_t\big] dB_t = \big[\Gamma_t W_t^P + k\bar{Y}_t\big] u_t dB_t, \end{aligned}$$

thanks to (6.58). This implies that

$$\Gamma_t W_t^P + \bar{Y}_t = \hat{\lambda}\Gamma_t$$

for some constant $\hat{\lambda}$. In particular,

$$\hat{\lambda}\Gamma_T = \Gamma_T W_T^P + k\bar{Y}_T = \Gamma_T \hat{U}_P\big(W_T^A\big) + k\hat{U}'_P\big(W_T^A\big)\Gamma_T.$$

Then,

$$\hat{U}_P\big(W_T^A\big) + k\hat{U}'_P\big(W_T^A\big) = \hat{\lambda}.$$

This, together with (6.52) and (6.53), leads to (6.48) again, for an appropriately chosen constant $\tilde{\lambda}$.

6.2.3 A Direct Approach

In this section we provide a direct approach for solving the problem, without using the results of Chap. 5, except for the general model described in Sect. 5.1.

We start with the agent's problem. Note that the agent's utility process satisfies

$$W_t^A = U_A(C_T) - \int_t^T \left[\frac{ku_s^2}{2} - u_s Z_s^A\right] ds - \int_t^T Z_s^A dB_s. \tag{6.59}$$

We have then immediately the following result.

Proposition 6.2.2 *Assume, for a given C_T, that the admissible set of u is given such that the BSDE (6.59) is well-posed and satisfies the comparison principle. Then, the necessary and sufficient condition for the agent's optimal effort is*

$$u_t = \frac{1}{k} Z_t^A. \tag{6.60}$$

Proof By the comparison principle for BSDEs the optimal u is obtained by minimizing the integrand $\frac{ku^2}{2} - uZ^A$ in (6.59), which implies (6.60). □

By (6.60), the agent's optimal utility W^A satisfies

$$dW_t^A = -\frac{k}{2}u_t^2 + ku_t dB_t.$$

Then,

$$de^{W_t^A/k} = e^{W_t^A/k} u_t dB_t. \tag{6.61}$$

This implies that, recalling (5.3),

$$e^{W_t^A/k} = e^{W_0^A/k} M_t^u.$$

Noting that $W_T^A = U_A(C_T)$, we get

$$M_T^u = \exp\left(\frac{1}{k}\left[U_A(C_T) - W_0^A\right]\right).$$

Moreover, under condition (6.54), as analyzed in the paragraph right after (6.54), it is optimal for the principal to offer contract C_T so that $W_0^A = R_0$. Therefore, for such contract and for the agent's optimal action, we have

$$M_T^u = e^{-R_0/k} e^{U_A(C_T)/k}. \tag{6.62}$$

This turns out to be exactly the reason why this problem is tractable: the fact that the choice of the probability measure corresponding to the optimal action u has an explicit functional relation with the promised payoff C_T.

We now turn to the principal's problem. Recall that

$$W_0^P = E^u\left[\tilde{U}_P(X_T - C_T)\right] = E\left[M_T^u \tilde{U}_P(X_T - C_T)\right].$$

Then, the principal's problem is

$$V_P := e^{-R_0/k} \sup_{C_T} E\big[e^{U_A(C_T)/k} \tilde{U}_P(X_T - C_T)\big]$$
$$\text{subject to} \quad E\big[e^{U_A(C_T)/k}\big] = e^{R_0/k}. \tag{6.63}$$

As usual, we consider the following relaxed problem with a Lagrange multiplier λ:

$$V_P(\lambda) := e^{-R_0/k} \sup_{C_T} E\big[e^{U_A(C_T)/k}\big[\tilde{U}_P(X_T - C_T) + \lambda\big]\big]. \tag{6.64}$$

The following result is then obvious:

Proposition 6.2.3 *Assume that the contract C_T is required to satisfy*

$$L \le C_T \le H$$

for some $\mathcal{F}_T$-measurable random variables L, H, which may take infinite values. If, with probability one, there exists a finite value $C_T^\lambda(\omega) \in [L(\omega), H(\omega)]$ that maximizes

$$e^{U_A(C_T)/k}\big[\tilde{U}_P(X_T - C_T) + \lambda\big] \tag{6.65}$$

and λ can be found so that

$$E\big[e^{U_A(C_T^\lambda)/k}\big] = e^{R_0/k},$$

then C_T^λ is the optimal contract.

Remark 6.2.4 Since (6.65) is considered ω by ω, we have reduced the problem to a one-variable deterministic optimization problem. In particular, if C_T is not constrained, the first order condition for optimal C_T is of the form

$$\frac{\tilde{U}_P'(X_T - C_T)}{U_A'(C_T)} = \frac{1}{k}\tilde{U}_P(X_T - C_T) + \lambda \tag{6.66}$$

and thus the optimal contract C_T is a function of the terminal value X_T only.

(i) The difference between Borch's rule (2.3) or (4.5) and condition (6.66) is the term with $\tilde{U}_P$: the ratio of marginal utilities of the agent and the principal is no longer constant, but a linear function of the utility of the principal. Increase in global utility of the principal also makes him happier at the margin, relative to the agent, and decrease in global utility makes him less happy at the margin. This will tend to make the contract "more nonlinear" than in the first best case. For example, if both utility functions are exponential, and we require $C_T \ge L > -\infty$ (for technical reasons), it is easy to check from Borch's rule that the first best contract C_T will be linear in X_T for $C_T > L$. On the other hand, as can be seen from (6.66), the second best contract will be nonlinear. Finally, we see that if cost k tends to infinity, the second best contract will tend to the first best contract.

(ii) By (6.66), omitting the functions arguments, we can find that

$$\frac{\partial}{\partial X_T} C_T = 1 - \frac{\tilde{U}_P' U_A''}{\tilde{U}_P'' U_A' + \tilde{U}_P' U_A'' - \frac{1}{k}\tilde{U}_P'(U_A')^2}.$$

Thus, under the standard conditions that

$$U_A \text{ and } \tilde{U}_P \text{ are increasing and concave,} \tag{6.67}$$

the contract is a non-decreasing function of X_T, and its slope with respect to X_T is not higher than one. In the first best case, Borch's rule gives us

$$\frac{\partial}{\partial X_T} C_T = 1 - \frac{\tilde{U}'_P U''_A}{\tilde{U}''_P U'_A + \tilde{U}'_P U''_A}.$$

We see that the sensitivity of the contract is higher in the second best case, partly because more incentives are needed to induce the agent to provide optimal effort when the effort is hidden. The term which causes the increase in the slope of the contract is $\frac{1}{k}\tilde{U}'_P(U'_A)^2$ in the denominator. We see that this term is dominated by the agent's marginal utility, but it also depends on the principal's marginal utility. Higher marginal utility for either party causes the slope of the contract to increase relative to the first best case. As already mentioned above, higher cost k makes it closer to the first best case.

6.2.4 Example: Risk-Neutral Principal and Log-Utility Agent

Example 6.2.5 Suppose $k = 1$ and the principal is risk-neutral while the agent is risk-averse with

$$\tilde{U}_P(C_T) = X_T - C_T, \qquad U_A(C_T) = \log C_T.$$

Since the log utility does not allow nonpositive output values, let us change the model to, with $\sigma_t > 0$ being a given process,

$$dX_t = \sigma_t X_t dB_t = \sigma_t u_t X_t dt + \sigma_t X_t dB_t^u.$$

Then, $X_t > 0$ for all t. Moreover, assume that

$$\lambda_0 := 2e^{R_0} - X_0 > 0.$$

In this case, the first order condition of (6.66) becomes

$$C_T = X_T - C_T + \lambda.$$

This gives a linear contract

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

and in order to satisfy the IR constraint in (6.63)

$$e^{R_0} = E[C_T] = \frac{1}{2}(X_0 + \lambda),$$

we need to take

$$\lambda = \lambda_0.$$

By assumption $\lambda_0 > 0$, we have $C_T > 0$, and C_T is then the optimal contract.

By (6.61), agent's optimal effort u is obtained by solving the BSDE

$$\tilde{W}_t^A = E_t[C_T] = C_T + \int_0^t \tilde{W}_s^A u_s dB_s.$$

Noting that

$$E_t[C_T] = \frac{1}{2}(X_t + \lambda_0) = e^{R_0} + \int_0^t \sigma_t X_t dB_t,$$

we get

$$\tilde{W}_t^A = \frac{1}{2}(X_t + \lambda_0), \qquad \tilde{W}_t^A u_t = \sigma_t X_t,$$

and thus

$$u_t = 2\sigma_t \frac{X_t}{X_t + \lambda_0}.$$

Since $\lambda_0 > 0$, we see that the effort goes down as the output decreases, and goes up when the output goes up. Thus, the incentive effect coming from the fact that the agent is paid an increasing function of the output at the end, translates into earlier times, so when the promise of the future payment gets higher, the agent works harder. Also notice that the effort is bounded in this example by $2\sigma_t$.

Assume now that σ is deterministic. The principal's optimal utility can be computed to be equal to

$$\begin{aligned}
V_P &= e^{-R_0} E\big[e^{U_A(C_T)} \tilde{U}_P(X_T - C_T)\big] = e^{-R_0} E\big[C_T[X_T - C_T]\big] \\
&= e^{-R_0} E\left[\frac{1}{4}[X_T + \lambda_0][X_T - \lambda_0]\right] \\
&= \frac{1}{4} e^{-R_0} E\left[X_0^2 \exp\left(2\int_0^T \sigma_t dB_t - \int_0^T \sigma_t^2 dt\right) - \left[2e^{R_0} - X_0\right]^2\right] \\
&= \frac{1}{4} e^{-R_0}\left[X_0^2 \exp\left(\int_0^T \sigma_t^2 dt\right) - 4e^{2R_0} + 4e^{R_0} X_0 - X_0^2\right] \\
&= X_0 - e^{R_0} + \frac{1}{4} e^{-R_0} X_0^2 \left[\exp\left(\int_0^T \sigma_t^2 dt\right) - 1\right].
\end{aligned}$$

The first term, $X_0 - e^{R_0}$, is what the principal can get if he pays a constant payoff C_T, in which case the agent would choose $u \equiv 0$. The second term is the extra benefit of inducing the agent to apply non-zero effort. The extra benefit increases quadratically with the initial output X_0, increases exponentially with the volatility squared, and decreases exponentially with the agent's reservation utility. While the principal would like best to have the agent with the lowest R_0, the cost of hiring expensive agents is somewhat offset when the volatility is high (which is not surprising, given that the principal is risk-neutral).

For comparison, we look now at the first best case in this example. Interestingly, we have

Proposition 6.2.6 *Assume that $\sigma_t > 0$ is deterministic and bounded. Then, the principal's first best optimal utility is infinite.*

Proof We see from Borch's rule (2.3) that, whenever the principal is risk-neutral, a candidate for an optimal contract is a constant contract C_T. With log-utility for the agent, we set

$$C_T = \lambda$$

where λ is obtained from the IR constraint, and the optimal utility of the principal is obtained from

$$\sup_u E[X_T - \lambda] = \sup_u \Big[E\Big\{X_0 e^{\int_0^T [u_t\sigma_t - \frac{1}{2}\sigma_t^2]dt + \int_0^T \sigma_t dB_t}\Big\} - e^R e^{E\{\int_0^T \frac{1}{2}u_t^2 dt\}}\Big]. \quad (6.68)$$

Under the assumption that σ is deterministic and bounded, we show now that the right-hand side of (6.68) is infinite. In fact, for any n, set

$$A_n := \left\{\int_0^{\frac{T}{2}} \sigma_t dB_t > n\right\} \in \mathcal{F}_{\frac{T}{2}}; \qquad \alpha_n := P(A_n) \to 0;$$

and

$$u_t^n(\omega) := \begin{cases} \alpha_n^{-\frac{1}{2}}, & \frac{T}{2} \le t \le T, \omega \in A_n; \\ 0, & \text{otherwise.} \end{cases} \quad (6.69)$$

Then, the cost is finite:

$$E\left\{\int_0^T \frac{1}{2}\big(u_t^n\big)^2 dt\right\} = \frac{T}{4}.$$

However, for a generic constant $c > 0$,

$$\begin{aligned}
&E\left\{x \exp\left(\int_0^T \left[u_t^n\sigma_t - \frac{1}{2}\sigma_t^2\right]dt + \int_0^T \sigma_t dB_t\right)\right\} \\
&= E\left\{x \exp\left(\alpha_n^{-\frac{1}{2}} \int_{\frac{T}{2}}^T \sigma_t dt \mathbf{1}_{A_n} - \int_0^T \frac{1}{2}\sigma_t^2 dt + \int_0^T \sigma_t dB_t\right)\right\} \\
&\ge E\left\{x \exp\left(\alpha_n^{-\frac{1}{2}} \int_{\frac{T}{2}}^T \sigma_t dt - \int_0^T \frac{1}{2}\sigma_t^2 dt + \int_0^T \sigma_t dB_t\right)\mathbf{1}_{A_n}\right\} \\
&= E\left\{x \exp\left(\alpha_n^{-\frac{1}{2}} \int_{\frac{T}{2}}^T \sigma_t dt - \int_0^{\frac{T}{2}} \frac{1}{2}\sigma_t^2 dt + \int_0^{\frac{T}{2}} \sigma_t dB_t\right)\mathbf{1}_{A_n}\right\} \\
&\ge cE\left\{x \exp\left(\alpha_n^{-\frac{1}{2}} \int_{\frac{T}{2}}^T \sigma_t dt + n\right)\mathbf{1}_{A_n}\right\} \\
&= cx \exp\left(\alpha_n^{-\frac{1}{2}} \int_{\frac{T}{2}}^T \sigma_t dt + n\right)P(A_n) \\
&= cx \exp\left(\alpha_n^{-\frac{1}{2}} \int_{\frac{T}{2}}^T \sigma_t dt + n\right)\alpha_n \ge cx\alpha_n e^{c\alpha_n^{-\frac{1}{2}}},
\end{aligned}$$

which diverges to infinity as $\alpha_n \to 0$. $\square$

We note that another completely solvable example in this special framework is the case of both the principal and the agent having linear utilities. However, in that case it is easily shown that the first best and the second best are the same, so there is no need to consider the second best.

6.3 Risk-Neutral Principal and Infinite Horizon

We now present a model which is a variation on the one in Sannikov (2008).

6.3.1 The Model

We consider a Markov model on infinite horizon with constant volatility, with risk-neutral principal and risk-averse agent with separable utility paid at rate c, both having the same discount rate r. More precisely, we have

$$T = \infty, \qquad U_A = 0, \qquad U_P = 0, \qquad v(t,x) = v,$$

and

$$\begin{aligned} &g(t,u) = re^{-rt} g(u), \qquad u_A(t,c) = re^{-rt} u_A(c), \\ &u_P(t, X_t, c) = re^{-rt}[X_t - X_0 - c]. \end{aligned} \tag{6.70}$$

Moreover, the IR constraint is only at initial time:

$$W_0^A \geq R_0. \tag{6.71}$$

Then,

$$\begin{aligned} X_t &= x + vB_t; \\ W_t^A &= \int_t^\infty re^{-rs}\big[u_A(c_s) - g(u_s)\big]ds - \int_t^\infty Z_s^A dB_s^u; \\ W_t^P &= \int_t^\infty re^{-rs}[vB_s - c_s]ds - \int_t^\infty Z_s^P dB_s^u. \end{aligned} \tag{6.72}$$

The goal of this section is to show that in this case the solution boils down to solving a differential equation in one variable, the agent's promised (remaining) utility process. Once that equation is obtained, it is possible to get economic conclusions by solving it numerically.

6.3.2 Necessary Conditions Derived from the General Theory

Note to the Reader The reader not interested in the use of general theory of Chap. 5 can skip this section and go to the following section that provides a more direct approach for dealing with the above model.

As usual we start with the agent's problem. In this case, as in (5.24) we have that the agent's optimal control satisfies

$$-re^{-rt}g'(u_t) + Z_t^A = 0. \tag{6.73}$$

Denote

$$\hat{W}_t^A := e^{rt}W_t^A, \qquad \hat{Z}_t^A := r^{-1}e^{rt}Z_t^A. \tag{6.74}$$

Then,

$$u_t = I_A\big(\hat{Z}_t^A\big), \quad \text{where } I_A := \big(g'\big)^{-1}, \tag{6.75}$$

and, given c, the agent's optimal utility satisfies

$$\begin{aligned}\hat{W}_t^A = &\int_t^\infty re^{-r(s-t)}\big[u_A(c_s) - g\big(I_A\big(\hat{Z}_s^A\big)\big) + \hat{Z}_s^A I_A\big(\hat{Z}_s^A\big)\big]ds \\ &- \int_t^\infty re^{-r(s-t)}\hat{Z}_s^A dB_s.\end{aligned} \tag{6.76}$$

For the principal's problem, the optimization condition (5.89) becomes

$$D_t re^{-rt}u_A'(c) - re^{-rt} = 0, \tag{6.77}$$

and thus, the principal's optimal control is

$$c_t = I_P(D_t) \quad \text{where } I_P := \big(1/u_A'\big)^{-1}. \tag{6.78}$$

Then, the FBSDE (5.99) becomes

$$\begin{aligned}D_t &= \lambda + \int_0^t Z_s^P r^{-1}e^{rs}I_A'\big(\hat{Z}_s^A\big)\big[dB_s - I_A\big(\hat{Z}_s^A\big)ds\big];\\ W_t^A &= \int_t^\infty re^{-rs}\big[u_A\big(I_P(D_s)\big) - g\big(I_A\big(\hat{Z}_s^A\big)\big)\big]ds\\ &\quad - \int_t^\infty re^{-rs}\hat{Z}_s^A\big[dB_s - I_A\big(\hat{Z}_s^A\big)ds\big];\\ W_t^P &= \int_t^\infty re^{-rs}\big[vB_s - I_P(D_s)\big]ds - \int_t^\infty \tilde{Z}_s^P\big[dB_s - I_A\big(\hat{Z}_s^A\big)ds\big].\end{aligned} \tag{6.79}$$

Note that

$$\begin{aligned}\int_t^\infty re^{-rs}B_s ds &= e^{-rt}B_t + \int_t^\infty e^{-rs}dB_s\\ &= e^{-rt}B_t + \int_t^\infty e^{-rs}u_s ds + \int_t^\infty e^{-rs}dB_s^u.\end{aligned} \tag{6.80}$$

Then,

$$\begin{aligned}W_t^P = &ve^{-rt}B_t + \int_t^\infty e^{-rs}\big[vI_A\big(\hat{Z}_s^A\big) - rI_P(D_s)\big]ds\\ &- \int_t^\infty \big[Z_s^P - ve^{-rs}\big]\big[dB_s - I_A\big(\hat{Z}_s^A\big)ds\big].\end{aligned}$$

Denote

$$\hat{W}_t^P := e^{rt} W_t^P - vB_t, \qquad \hat{Z}_t^P := r^{-1}\big[e^{rt} Z_t^P - v\big]. \tag{6.81}$$

Then, (6.79) becomes

$$\begin{aligned}
D_t &= \lambda + \int_0^t \big[\hat{Z}_s^P + v/r\big] I_A'\big(\hat{Z}_s^A\big)\big[dB_s - I_A\big(\hat{Z}_s^A\big)ds\big];\\
\hat{W}_t^A &= \int_t^\infty re^{-r(s-t)}\big[u_A\big(I_P(D_s)\big) - g\big(I_A\big(\hat{Z}_s^A\big)\big)\big]ds\\
&\quad - \int_t^\infty re^{-r(s-t)}\hat{Z}_s^A\big[dB_s - I_A\big(\hat{Z}_s^A\big)ds\big];\\
\hat{W}_t^P &= \int_t^\infty e^{-r(s-t)}\big[vI_A\big(\hat{Z}_s^A\big) - rI_P(D_s)\big]ds\\
&\quad - \int_t^\infty re^{-r(s-t)}\hat{Z}_s^P\big[dB_s - I_A\big(\hat{Z}_s^A\big)ds\big].
\end{aligned} \tag{6.82}$$

The above FBSDE (6.82) is Markovian and time homogeneous. Thus, we expect to have

$$\hat{W}_t^A = \varphi_A(D_t), \qquad \hat{W}_t^P = \varphi_P(D_t), \qquad \text{for some deterministic functions } \varphi_A, \varphi_P.$$

Assume φ_A has an inverse, and denote

$$\psi := (\varphi_A)^{-1}, \qquad \hat{F}(x) := \varphi_P\big(\psi(x)\big). \tag{6.83}$$

Then, ψ and $\hat{F}$ are independent of λ, and we have

$$\hat{W}_t^P = \hat{F}\big(\hat{W}_t^A\big), \qquad D_t = \psi\big(\hat{W}_t^A\big). \tag{6.84}$$

This implies that

$$V_P = \sup_{w_A \ge R_0} \hat{F}(w_A). \tag{6.85}$$

In particular, when the function $\hat{F}$ is decreasing, then

$$V_P = \hat{F}(R_0). \tag{6.86}$$

We now formally derive the equation which the function $\hat{F}$ should satisfy. Let $(D, \hat{W}^A, \hat{Z}^A, \hat{W}^P, \hat{Z}^P)$ solve (6.82). Denote

$$u_t := I_A\big(\hat{Z}_t^A\big), \qquad c_t := I_P(D_t). \tag{6.87}$$

Then,

$$\begin{aligned}
dD_t &= \frac{\hat{Z}_t^P + v}{g''(u_t)} dB_t^u;\\
d\hat{W}_t^A &= d\big(e^{rt}W_t^A\big) = r\hat{W}_t^A dt - r\big[u_A(c_t) - g(u_t)\big]dt + rg'(u_t)dB_t^u;\\
d\hat{W}_t^P &= d\big(e^{rt}\big[W_t^P - vre^{-rt}B_t\big]\big) = r\hat{W}_t^P dt - r[vu_t - c_t]dt + r\big[\hat{Z}_t^P + v\big]dB_t^u.
\end{aligned}$$

On the other hand, applying Itô's rule, we have

$$\begin{aligned} dD_t &= d\big(\psi\big(\hat W_t^A\big)\big) = \psi'\big(\hat W_t^A\big) r\big[\hat W_t^A - u_A(c_t) + g(u_t)\big]dt + \psi'\big(\hat W_t^A\big) r \hat Z_t^A dB_t^u \\ &\quad + \frac{1}{2}\psi''\big(\hat W_t^A\big)\big|rg'(u_t)\big|^2 dt; \\ d\hat W_t^P &= d\big(\hat F\big(\hat W_t^A\big)\big) = \hat F'\big(\hat W_t^A\big) r\big[\hat W_t^A - u_A(c_t) + g(u_t)\big]dt + \hat F'\big(\hat W_t^A\big) r \hat Z_t^A dB_t^u \\ &\quad + \frac{1}{2}\hat F''\big(\hat W_t^A\big)\big|rg'(u_t)\big|^2 dt. \end{aligned}$$

Comparing the above expressions, we get

$$\begin{aligned} \psi'\big(\hat W_t^A\big)\big[\hat W_t^A - u_A(c_t) + g(u_t)\big] + \frac{r}{2}\psi''\big(\hat W_t^A\big)\big|g'(u_t)\big|^2 &= 0; \\ \psi'\big(\hat W_t^A\big) g'(u_t) &= \frac{\hat Z_t^P + v}{rg''(u_t)}; \\ \hat F'\big(\hat W_t^A\big)\big[\hat W_t^A - u_A(c_t) + g(u_t)\big] + \frac{r}{2}\hat F''\big(\hat W_t^A\big)\big|g'(u_t)\big|^2 &= \hat F\big(\hat W_t^A\big) - vu_t + c_t; \\ \hat F'\big(\hat W_t^A\big) g'(u_t) &= \hat Z_t^P + v. \end{aligned}$$

Using these, we obtain

$$vu_t - c_t - \hat F\big(\hat W_t^A\big) + \hat F'\big(\hat W_t^A\big)\big[\hat W_t^A - u_A(c_t) + g(u_t)\big] + \frac{r}{2}\hat F''\big(\hat W_t^A\big)\big|g'(u_t)\big|^2 = 0, \tag{6.88}$$

and the optimal c, u, together with the function ψ, satisfy:

$$\begin{aligned} \psi\big(\hat W_t^A\big) u_A'(c_t) &= 1; \\ r\psi'\big(\hat W_t^A\big) g''(u_t) &= \hat F'\big(\hat W_t^A\big); \\ \psi'\big(\hat W_t^A\big)\big[\hat W_t^A - u_A(c_t) + g(u_t)\big] + \frac{r}{2}\psi''\big(\hat W_t^A\big)\big|g'(u_t)\big|^2 &= 0. \end{aligned} \tag{6.89}$$

This gives us a differential equation for function $\hat F$ (and ψ), and shows how optimal u and c depend in a deterministic way on the agent utility process $\hat W_t^A$.

6.3.3 A Direct Approach

In this subsection, we solve the problem directly by using the standard approach in Stochastic Control Theory, the Hamilton–Jacobi–Bellman (HJB) equation. This approach is briefly reviewed in Sect. 5.4.4.

The usual argument implies that the agent's optimal effort satisfies

$$re^{-rt} g'(u_t) = Z_t^A.$$

(This is (6.73) in the previous section.)

For any given w_A, we restrict our control (c, u) to be an element of the set $\mathcal{A}(w_A)$ of all controls that, besides the standard measurability and integrability conditions, satisfy

$$\lim_{t\to\infty} W_t^A = 0,$$
$$\text{where } W_t^A = w_A - \int_0^t re^{-rs}\big[u_A(c_s) - g(u_s)\big]ds + \int_0^t re^{-rs} g'(u_S)dB_s^u. \quad (6.90)$$

Here, the limit is in L^2 sense.

Denote

$$\tilde{W}_t^P := W_t^P - vre^{-rt}B_t, \qquad \tilde{Z}_t^P := Z_t^P - vre^{-rt}. \quad (6.91)$$

Using

$$\int_t^\infty re^{-rs}B_s ds = e^{-rt}B_t + \int_t^\infty e^{-rs}dB_s = e^{-rt}B_t + \int_t^\infty e^{-rs}u_s ds + \int_t^\infty e^{-rs}dB_s^u, \quad (6.92)$$

we get

$$\tilde{W}_t^P = \int_t^\infty re^{-rs}[vu_s - c_s]ds - \int_t^\infty \tilde{Z}_s^P dB_s^u. \quad (6.93)$$

Note that $W_0^P = \tilde{W}_0^P$. Assume, for each $(c, u) \in \mathcal{A}(w_A)$, that c is implementable using effort u that is optimal for the agent. Then,

$$V_P = \sup_{w_A \ge R_0} F(w_A) \quad \text{where } F(w_A) := \sup_{(c,u)\in\mathcal{A}(w_A)} \tilde{W}_0^P. \quad (6.94)$$

We now derive formally the HJB equation that F should satisfy. For any $(c, u) \in \mathcal{A}(w_A)$ and $t \ge \delta > 0$, note that

$$\begin{aligned} e^{r\delta}W_t^A &= e^{r\delta}W_\delta^A - \int_\delta^t re^{-r(s-\delta)}\big[u_A(c_s) - g(u_s)\big]ds + \int_\delta^t re^{-r(s-\delta)}g'(u_s)dB_s^u, \\ e^{r\delta}\tilde{W}_t^P &= \int_t^\infty re^{-r(s-\delta)}[vu_s - c_s]ds - \int_t^\infty e^{r\delta}\tilde{Z}_s^P dB_s^u. \end{aligned} \quad (6.95)$$

Following the standard arguments in Stochastic Control Theory, we have the following Dynamic Programming Principle:

Proposition 6.3.1 *Assume that function F above is continuous. Then, for any $\delta > 0$,*

$$F(w_A) = \sup_{c,u\in\mathcal{A}(w_A)} E^u\left[\int_0^\delta re^{-rs}[vu_s - c_s]ds + e^{-r\delta}F\big(\hat{W}_\delta^A\big)\right] \quad (6.96)$$

where

$$\begin{aligned} &\hat{W}_t^A := e^{rt}W_t^A \quad \text{and} \\ &W_\delta^A = w_A - \int_0^\delta re^{-rs}\big[u_A(c_s) - g(u_s)\big]ds + \int_0^\delta re^{-rs}g'(u_s)dB_s^u. \end{aligned} \quad (6.97)$$

Assume now that F is sufficiently smooth. Applying Itô's rule, we have

$$d(\hat{W}_t^A) = d(e^{rt} W_t^A) = r\hat{W}_t^A dt - r[u_A(c_t) - g(u_t)]dt + rg'(u_t)dB_t^u,$$

and thus

$$\begin{aligned} d(e^{-rt}F(\hat{W}_t^A)) = -re^{-rt}F(\hat{W}_t^A)dt + \frac{1}{2}e^{-rt}F''(\hat{W}_t^A)|rg'(u_t)|^2dt \\ + e^{-rt}F'(\hat{W}_t^A)[r\hat{W}_t^A dt - r[u_A(c_t) - g(u_t)]dt \\ + rg'(u_t)dB_t^u]. \end{aligned} \tag{6.98}$$

Plugging this into (6.96), dividing both sides by δ, and then sending $\delta \to 0$, we get

$$\sup_{c,u}\Big[vu - c - F(w_A) + \frac{r}{2}F''(w_A)\big|g'(u)\big|^2 + F'(w_A)[w_A - u_A(c) + g(u)]\Big] = 0. \tag{6.99}$$

Furthermore, if the supremum is attained by a control couple $(c, u) \in \mathcal{A}(w_A)$ such that

$$\begin{aligned} c_t &= \operatorname*{argmin}_c [c + F'(\hat{W}_t^A)u_A(c)], \\ u_t &= \operatorname*{argmax}_u \Big[vu + F'(\hat{W}_t^A)g(u) + \frac{r}{2}F''(\hat{W}_t^A)[g'(u)]^2 \Big], \end{aligned} \tag{6.100}$$

then they are optimal.

We now prove a verification result, under quite strong conditions. More general results can be obtained following the viscosity solution approach, as in Fleming and Soner (2006) and Yong and Zhou (1999).

Proposition 6.3.2 *Assume that the HJB equation* (6.99) *has a classical solution* $\tilde{F}$ *that has linear growth. Then,* $F \le \tilde{F}$*. Moreover, if there exists a pair* $(c, u) \in \mathcal{A}(w_A)$ *such that* (6.100) *holds, then* $F = \tilde{F}$*, and* c *and* u *are optimal.*

Proof For any $(c, u) \in \mathcal{A}(w_A)$, by (6.98) and (6.99) we have

$$\begin{aligned} d(e^{-rt}\tilde{F}(\hat{W}_t^A)) &= re^{-rt}\Big[-\tilde{F}(\hat{W}_t^A) + \frac{r}{2}\tilde{F}''(\hat{W}_t^A)\big|g'(u_t)\big|^2 \\ &\quad + \tilde{F}'(\hat{W}_t^A)[\hat{W}_t^A - u_A(c_t) + g(u_t)]\Big]dt \\ &\quad + re^{-rt}\tilde{F}'(\hat{W}_t^A)g'(u_t)dB_t^u \\ &\le re^{-rt}[c_t - vu_t]dt + re^{-rt}\tilde{F}'(\hat{W}_t^A)g'(u_t)dB_t^u. \end{aligned} \tag{6.101}$$

Then, (6.93) leads to

$$d(e^{-rt}\tilde{F}(\hat{W}_t^A) - \tilde{W}_t^P) \le [re^{-rt}\tilde{F}'(\hat{W}_t^A)g'(u_t) - \tilde{Z}_t^P]dB_t^u. \tag{6.102}$$

Note that, by the linear growth of F and (6.90),

$$\big|e^{-rt}\tilde{F}(\hat{W}_t^A)\big| \le Ce^{-rt}[1 + |\hat{W}_t^A|] \le C[e^{-rt} + |W_t^A|] \to 0, \quad \text{as } t \to \infty.$$

Since we are assuming enough integrability, we obtain from (6.102) that

$$\tilde{F}\big(w^A\big) - \tilde{W}_0^P = \tilde{F}\big(\hat{W}_0^A\big) - \tilde{W}_0^P \geq 0.$$

This, together with the arbitrariness of (c,u), implies $\tilde{F}(w^A) \geq F(w_A)$.

On the other hand, if $(c,u) \in \mathcal{A}(w_A)$ satisfies (6.100), then the inequality in (6.101) becomes an equality. Consequently, (6.102) becomes an equation, and thus $\tilde{F}(w^A) = \tilde{W}_0^P$ for this (c,u). Then clearly $\tilde{F}(w^A) = F(w_A)$ and (c,u) is an optimal control. □

Remark 6.3.3 For the readers who are familiar with the previous section, we remark that function F here is in general different from function $\hat{F}$ of that section. Indeed, let

$$\mathcal{A} := \bigcup_{w_A \geq R_0} \mathcal{A}(w_A). \tag{6.103}$$

Then, the system (6.79) is obtained by solving the optimization problem

$$V_P(\lambda) := \sup_{(c,u)\in\mathcal{A}} \big[W_0^{P,c,u} + \lambda W_0^{A,c,u}\big], \tag{6.104}$$

and

$$W_0^P = \hat{W}_0^P = \varphi_P(\lambda) = V_P(\lambda) - \lambda\varphi_A(\lambda). \tag{6.105}$$

For given $w_A \geq R_0$, if we choose $\lambda := \psi(w_A)$, then

$$\begin{aligned}
\hat{F}(w_A) = \hat{W}_0^P &= V_P\big(\psi(w_A)\big) - w_A\psi(w_A) \\
&= \sup_{(c,u)\in\mathcal{A}} \big[W_0^{P,c,u} + \psi(w_A)W_0^{A,c,u}\big] - w_A\psi(w_A) \\
&\geq \sup_{(c,u)\in\mathcal{A}(w_A)} \big[W_0^{P,c,u} + \psi(w_A)W_0^{A,c,u}\big] - w_A\psi(w_A) \\
&= \sup_{(c,u)\in\mathcal{A}(w_A)} W_0^{P,c,u} = F(w_A).
\end{aligned} \tag{6.106}$$

However, we note that

$$\sup_{w_A\geq R_0} \hat{F}(w_A) = V_P = \sup_{w_A\geq R_0} F(w_A). \tag{6.107}$$

6.3.4 Interpretation and Discussion

(i) From (6.100) we see that the principal faces a tradeoff between minimizing the payment c and maximizing the agent's utility $u_A(c)$, but weighted by the marginal change $F'(\hat{W}_t^A)$ in the principal's utility relative to the agent's utility.

(ii) From (6.100), we see that the optimally induced effort faces a tradeoff between maximizing the drift of the output, minimizing the cost of the effort, and minimizing the risk to which the agent is exposed. The latter risk is represented by the term

$$-F''\big(\hat{W}_t^A\big)\big[re^{-rt}g'(u_t)\big]^2$$

thus, equal to a (minus) product of the marginal change in sensitivity of the principal's utility with respect to the agent's promised utility, with the squared volatility of the agent's promised utility.

(iii) If c_t is unconstrained, it follows from (6.100) that the first order condition for optimality in c_t is

$$c_t = I_P\big(F'\big(\hat{W}_t^A\big)\big)$$

in the notation of (6.78). On the other hand, by that equation the Stochastic Maximum Principle approach gives us $c_t = I_P(D_t)$ which implies

$$D_t = F'\big(\hat{W}_t^A\big)$$

and provides an economic meaning to the process D as the marginal change in the principal's utility with respect to the agent's utility.

(iv) In order to solve the problem using HJB equation (6.99), we need boundary conditions. These will depend on the specifics of the model. For example, suppose that we allow only non-negative effort, $u_t \geq 0$. Then, at the minimum possible value for the agent's utility, denoted w_L, the agent will apply minimal effort zero, and also be paid minimal possible consumption, denoted c_L, which will make the principal's utility equal to $-r\int_0^\infty e^{-rt}c_L dt = -c_L$. Thus, the boundary condition at the bottom range is $F(w_L) = -c_L$. If there is no upper bound on W^A and no lower bound on $W^{P,\lambda}$, then we also have $F(+\infty) = -\infty$. Sannikov (2008) works with more realistic assumptions, discussed next.

6.3.5 Further Economic Conclusions and Extensions

Having obtained the HJB differential equation, Sannikov (2008) is able to do numerical computations and discuss economic consequences of the model. We list some of them here. First, there are some additional conditions in his model: The consumption c_t and effort u_t are restricted to be non-negative, as is the utility process W_t^A. It is assumed that $g(u) \geq \varepsilon u$ for all $u \geq 0$, and $g(0) = 0$. The utility function u_A is bounded below, with $u_A(0) = 0$. Thus, the boundary conditions for the PDE are, first:

$$F(0) = 0.$$

That is, once the agent's remaining utility hits zero, the principal retires him with zero payment. The second boundary condition for the solution is

$$F(w_{gp}) = -u_A^{-1}(w_{gp}), \qquad F'(w_{gp}) = -\big[u_A^{-1}\big]'(w_{gp})$$

where w_{gp} is an unknown point ("gp" stands for "golden parachute", an expression for the retirement payment). The interpretation of this condition is that when the agent's promised utility reaches too high a point w_{gp}, the principal retires the agent and continues paying him constant consumption, which then has to be equal to $u_A^{-1}(w_{gp})$. Hence the expected value of output is zero after that time, and the principal's remaining utility is equal to $[-u_A^{-1}(w_{gp})]$. For the use below, denote the principal's retirement profit

$$F_0(w) = -u_A^{-1}(w).$$

Another condition on F is

$$F(w) \geq F_0(w) \quad \text{for all } w \geq 0.$$

That is, the principal's profit is no less than the value obtained by retiring the agent.

In this model where the consumption c is constrained from below, it may happen that the agent is paid more than his reservation value $R_0 \geq 0$. This is because the function F, under the above conditions and restrictions, is not necessarily decreasing in the area where $c \equiv 0$. The principal gives the agent the value $W_0^A = w_0$ which maximizes $F(w)$ on $[R_0, w_{gp}]$, if $F(w_0) > 0$. Otherwise, if $F(w) \leq 0$ for all w on $[R_0, w_{gp}]$ the principal does not hire the agent.

Using the above machinery, it is possible to show the following general principle for this model: a change in boundary conditions that makes the principal's utility $F(w)$ uniformly higher, increases the agent's optimal effort $u = u(w)$ for all wage levels w.

Changing boundary conditions allows us to consider the following extensions:

1. **The agent can quit at any time and take an outside job** ("outside option") with expected utility $\tilde{R}_0 < R_0$. $\tilde{R}_0$ is interpreted as the value of new employment minus the search costs. In this case, the solution $\tilde{F}$ is obtained by solving the same HJB equation, except that the boundary conditions change: now $\tilde{F}(\tilde{R}_0) = 0$, since $w^A = \tilde{R}_0$ is the low retirement point, not $w^A = 0$. The boundary conditions for the new high retirement point $\tilde{w}_{gp}$ are the same as before, except we have a constraint $\tilde{w}_{gp} > \tilde{R}_0$. Numerical computations and/or analytical results in Sannikov (2008) show that what happens is:
   (i) $\tilde{F} \leq F$: the principal's profit is lower;
   (ii) $\tilde{w}_{gp} < w_{gp}$: the high retirement point occurs sooner, as the principal's profit is lower;
   (iii) the agent works less hard;
   (iv) the consumption payment c is lower: the payments are "backloaded" when the principal is trying to tie the agent more closely to the firm;
   (v) the agent's promised utility W_0^A is at least as large as without the outside option.
2. **The principal can replace the agent with another agent** of the same reservation value R_0, at a fixed cost C. In this case the principal's retirement profit will be higher, $\tilde{F}_0(w) = F_0(w) + D$, for D of the form $D = F(w_0) - C$, where w_0 has to be determined. The boundary conditions are $F(0) = \tilde{F}_0(0) = D$,

$F(w_{gp}) = \tilde{F}_0(w_{gp})$ and $F'(w_{gp}) = \tilde{F}_0'(w_{gp})$. Then, $w_0 = w_0(D)$ has to be chosen so that F is maximized on the interval $[R_0, w_{gp}]$. Since we don't know D in advance, numerically we start with an arbitrary value of D. If, doing the above procedure, we get $D = F(w_0) - C$, we are done. Otherwise, we have to adjust the value of D up or down, and repeat the procedure.

What happens is:

(i) the principal's profit is higher;
(ii) the agent works harder;
(iii) w_{gp} is increasing in C.
(iv) The principal's utility may be higher than the first best utility with only one agent.

3. **The principal can train and promote the agent** at a cost $K \geq 0$, instead of retiring him. When promotion happens, it increases the drift from u to θu for $\theta > 1$, but also increases the agent's outside option from zero to $\tilde{R} > 0$. If we denote by F_1 the principal's profit without promotion and with F_2 her profit with promotion, we will have a boundary condition $F_1(w_p) = F_2(w_p) - K$, $F_1'(w_p) = F_2'(w_p)$, where w_p is the point of promotion. What happens is:
   (i) the principal's profit is higher;
   (ii) the agent works harder until promotion;
   (iii) the consumption payment c is lower;
   (iv) the agent's promised utility W_0^A is at least as large as without promotion.
   (v) The promotion is not offered right away for the following reasons: (a) the agent first has to show good performance; (b) the agent's outside option may increase with promotion, making him more likely to leave; (c) training for promotion is costly.
4. **The principal cannot commit to the payments and to not replacing the agent**. What happens is:
   (i) the principal's profit is lower;
   (ii) the agent works less hard;
   (iii) the consumption payment c is higher;
   (iv) the agent's promised utility W_0^A is smaller;
   (v) with lack of commitment, the principal's profit with replacement or promotion options can actually be lower than without those options.

6.4 Further Reading

Section 6.1 summarizes some of the main results from Hölmstrom and Milgrom (1987). The example of Sect. 6.2 is a generalization of a case studied in Cvitanić et al. (2006). The setting and the results of Sect. 6.3 are taken from Sannikov (2008). An interesting application of Sannikov (2008) can be found in Fong (2009).

There is a growing literature extending methods of this chapter to various new applications involving moral hazard in continuous-time. These include: (i) processes driven by jumps and not by Brownian motion, see Zhang (2009) and Biais et al. (2010); (ii) imperfect information and learning, see Adrian and Westerfield (2009), DeMarzo and Sannikov (2011), Giat and Subramanian (2009), Prat and Jovanovic (2010), He et al. (2010), and Giat et al. (2011); (iii) asset pricing, see Ou-Yang (2005); (iv) executive compensation, see He (2009); (v) stochastic interest rates and mortgage contracts, see Piskorski and Tchistyi (2010). Additional references can be found in a nice survey paper, Sannikov (2012).

References

Adrian, T., Westerfield, M.: Disagreement and learning in a dynamic contracting model. Rev. Financ. Stud. **22**, 3839–3871 (2009)

Biais, B., Mariotti, T., Rochet, J.-C., Villeneuve, S.: Large risks, limited liability, and dynamic moral hazard. Econometrica **78**, 73–118 (2010)

Cvitanić, J., Wan, X., Zhang, J.: Optimal contracts in continuous-time models. J. Appl. Math. Stoch. Anal. **2006**, 1–27 (2006)

DeMarzo, P.M., Sannikov, Y.: Learning, termination and payout policy in dynamic incentive contracts. Working paper, Princeton University (2011)

Fong, K.G.: Evaluating skilled experts: optimal scoring rules for surgeons. Working paper, Stanford University (2009)

Giat, Y., Subramanian, A.: Dynamic contracting under imperfect public information and asymmetric beliefs. Working paper, Georgia State University (2009)

Giat, Y., Hackman, S.T., Subramanian, A.: Investment under uncertainty, heterogeneous beliefs and agency conflicts. Rev. Financ. Stud. **23**(4), 1360–1404 (2011)

He, Z.: Optimal executive compensation when firm size follows geometric Brownian motion. Rev. Financ. Stud. **22**, 859–892 (2009)

He, Z., Wei, B., Yu, J.: Permanent risk and dynamic incentives. Working paper, Baruch College (2010)

Holmström, B., Milgrom, P.: Aggregation and linearity in the provision of intertemporal incentives. Econometrica **55**, 303–328 (1987)

Ou-Yang, H.: An equilibrium model of asset pricing and moral hazard. Rev. Financ. Stud. **18**, 1219–1251 (2005)

Piskorski, T., Tchistyi, A.: Optimal mortgage design. Rev. Financ. Stud. **23**, 3098–3140 (2010)

Prat, J., Jovanovic, B.: Dynamic incentive contracts under parameter uncertainty. Working paper, NYU (2010)

Sannikov, Y.: A continuous-time version of the principal-agent problem. Rev. Econ. Stud. **75**, 957–984 (2008)

Sannikov, Y.: Contracts: the theory of dynamic principal-agent relationships and the continuous-time approach. Working paper, Princeton University (2012)

Zhang, Y.: Dynamic contracting with persistent shocks. J. Econ. Theory **144**, 635–675 (2009)

Springer

http://www.springer.com/978-3-642-14199-7

Contract Theory in Continuous-Time Models
Cvitanic, J.; Zhang, J.
2013, XII, 256 p., Hardcover
ISBN: 978-3-642-14199-7