

# Optimal Compensation and Pay-Performance Sensitivity in a Continuous-Time Principal-Agent Model\*

Nengjiu Ju<sup>†</sup> and Xuhu Wan<sup>‡</sup>

June 19, 2011

## ABSTRACT

This paper studies the optimal contract between risk-neutral shareholders and a constant relative risk-aversion manager in a continuous-time model. Several interesting results are obtained. First, the optimal compensation is increasing but concave in output value if the manager is more risk-averse than a log-utility manager. Second, when the manager has a log-utility, a linear contract is optimal when there is no explicit lower bound on the compensation, and an option contract is optimal when there is an explicit lower bound. Third, optimal effort is stochastic (state-dependent). Fourth, consistent with empirical findings and contrary to standard agency theory predictions, the relationship between pay-performance sensitivity and firm performance and that between pay-performance sensitivity and firm risk can be non-monotonic.

JEL CLASSIFICATION: G34; C61; D86.

Keywords: Continuous-time principal-agent models, optimal concave contract, stochastic optimal effort, pay-performance sensitivity.

---

\*We appreciate comments from Gurdip Bakshi, Tao Li, Peter MacKay, Jianjun Miao, Rong Wang, and seminar participants at the Chinese University of Hong Kong, the University of Hong Kong, and the 2007 China International Conference in Finance. We are most grateful to an anonymous referee and an associate editor for their extensive and thoughtful comments and suggestions. Financial support from Hong Kong government general research fund grant GRF 620909 is acknowledged.

<sup>†</sup>Department of Finance, School of Business and Management, Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong, E-mail: nengjiu@ust.hk; and Shanghai Advanced Institute of Finance, Shanghai Jiaotong University, China 20030, email: njju@saif.sjtu.edu.cn.

<sup>‡</sup>Department of Information and Systems Management, School of Business and Management, Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong. Phone: (852) 2358-7731. E-mail: imwan@ust.hk.

# 1. Introduction

We study the optimal contract between well-diversified risk-neutral shareholders and a constant relative risk aversion (thereafter CRRA) manager in a continuous-time agency model. In our model, the manager controls the instantaneous growth rate of an output process. Cost of effort is assumed as a standard quadratic function. A number of interesting results are obtained.

First, the optimal contract is strictly increasing but *concave* in the output value if the manager is more risk-averse than a log-utility manager. Due to the wealth effect, relative to a log-utility manager, it is more cost effective to compensate a more risk-averse manager with a contract which has a relatively higher *portion* in equity-linked compensation at the lower end of the output value than at the higher end. The concavity of the optimal compensation is in sharp contrast with the convexity of compensations such as stock options whose optimality is normally not established.

If, on the other hand, the managerial compensation has a positive lower bound (which can be interpreted as a base payment or limited liability), the optimal compensation consists of a cash payment plus an equity-linked component when the output value is above a certain level. While the equity-linked component has the appearance of a stock option (kinked payoff due to the lower bound constraint), it is generally not.

The result that the constrained optimal compensation with a lower bound is option-like appears to be general and has also been obtained in, among others, Kadan and Swinkels (2008), and Jewitt, Kadan and Swinkels (2008). Even though these papers have all considered bounds on compensations, they differ in focuses. For example, while studying the effects of compensation bounds in static agency models is the focus of Jewitt, Kadan and Swinkels (2008), our paper is on the dynamics of optimal efforts and compensation in a dynamic model. On the other hand, while the focus of Kadan and Swinkels (2008) is on the effort incentives when compensation is in stock or stock option, our paper determines the optimal contract from an *unrestricted* contract space, except perhaps a lower bound constraint.

Second, when the manager has a log-utility, a linear contract is optimal when there is no explicit lower bound on the compensation, and an option contract is optimal when there is an explicit lower bound. The reason is that, besides the wealth effect, our model exhibits a scale effect. In our model, effort affects the output *scale* rather than *level*.<sup>1</sup> Same effort level has a larger impact on output value when the asset value is higher. This scale effect calls for a higher portion of equity-linked compensation which induces

---

<sup>1</sup>See Remark 1 for more detail.

higher effort. For a manager more risk-averse than a log-utility manager, the wealth effect dominates and a concave compensation becomes optimal. For a log-utility manager, the wealth effect and scale effect balance out and a linear contract becomes optimal when there is no lower bound on the compensation and an option contract becomes optimal when there is a lower bound constraint.<sup>2</sup>

Third, the optimal effort is stochastic and in stark contrast with the constant effort result in the standard exponential-normal settings of Holmstrom and Milgrom (1987) and the deterministic counterpart in Ou-Yang (2003, 2005). In models where effort affects level, its impact on final output value is independent of current asset value (state variable). Therefore, optimal effort does not depend on the state variable. By contrast, effort in our model affects scale. As a result, the optimal effort becomes state-dependent. However, the exact behavior depends on risk aversion and is our focus in section 4.

Fourth, the relationship between pay-performance sensitivity (thereafter PPS) and firm performance and that between PPS and firm risk may be non-monotonic, and hence consistent with the extant empirical evidence. One of the central features of standard agency models is that firm performance should be positively correlated with PPS. While a positive relation is reported in McConnell and Servaes (1990), and Lazear (2000), no significant relation is found in Himmelberg, Hubbard, and Palia (1999) and Palia (2001). Another central prediction of these models is that PPS should be negatively correlated with firm risk. While a negative relation is supported by the studies of Lambert and Larcker (1987), Aggarwal and Samwick (1999), Jin (2002), and Garvey and Milbourn (2003), a positive relation is found in Demsetz and Lehn (1985). Furthermore, a large body of empirical work (Garen (1994), Yermack (1995), Bushman, Indejikian, and Smith (1996), and Ittner, Larcker, and Rajan (1997)) finds that there is no significant relationship between PPS and firm risk. Prendergast (2002) provides empirical evidence, argues, and develops a theoretical model to show that the relation between PPS and risk can be positive.

The rest of the paper is organized as follows. The next section specifies the model assumptions and the maximization problems of the shareholders and the CRRA manager. Section 3 develops the major solution steps and techniques and solves the maximization problems. Section 4 discusses properties of the optimal effort. Section 5 focuses on the relations between PPS and firm performance and risk. Section 6 offers a few empirical predictions. Section 7 concludes. Appendix A provides a complete characterization of the optimal compensation. Appendix B proves some properties of the optimal effort.

---

<sup>2</sup>Studies which take stock or stock options as the compensation form without addressing their optimality include, among others, Cadenillas, Cvitanić, and Zapatero (2004), Carpenter (1998, 2000), Carpenter and Remmers (2001), Carpenter, Stanton, and Wallace (2009), Guo, and Ou-Yang (2006).

## 2. A Continuous-Time Principal-Agent Model

### 2.1. Asset value dynamics under different probability measures

**Assumption 1** *Under no managerial effort, the firm's asset value follows a geometric Brownian motion:*

$$dV_t = V_t (\mu dt + \sigma dB_t^0), \quad (1)$$

where  $\mu$  and  $\sigma$  are constants and  $B_t^0$  is a standard Wiener process under reference probability measure  $P^0$  on space  $(\Omega, \mathcal{F})$ . Let  $\mathcal{F}_t$  denote the information set generated by  $B_t^0$ .

**Assumption 2** *Through costly managerial effort  $e_t$ , the probability measure  $P^0$  is changed into a new equivalent measure  $P^e$  defined by*

$$\frac{dP^e}{dP^0} = M_T^e,$$

where  $M_T^e$  is an  $\mathcal{F}_T$ -adapted martingale under measure  $P^0$  and has the following representation:

$$M_T^e = \exp\left(-\frac{1}{2} \int_0^T \left(\frac{e_t}{\sigma}\right)^2 dt + \int_0^T \left(\frac{e_t}{\sigma}\right) dB_t^0\right). \quad (2)$$

Note that by Girsanov theorem,  $B_t^e$ , defined by

$$B_t^e = B_t^0 - \int_0^t \left(\frac{e_t}{\sigma}\right) dt, \quad (3)$$

is another standard Wiener process under the new measure  $P^e$ . Consequently, under  $P^e$  the value of the firm's assets is governed by

$$dV_t = V_t ((\mu + e_t)dt + \sigma dB_t^e). \quad (4)$$

In the following, expectations under  $P^0$  and  $P^e$  are respectively denoted as  $\mathbb{E}^0[\cdot | \mathcal{F}_t]$  and  $\mathbb{E}^e[\cdot | \mathcal{F}_t]$ .

**Remark 1** *The effect of the costly effort in (4) is to increase the (instantaneous) expected growth rate by  $e_t$  and the output level  $V_t$  grows at the rate of  $V_t e_t$ , proportional to its current level  $V_t$ , a scale effect. This is different from the setting of Holmstrom and Milgrom (1987), in which the level  $V_t$  grows at the rate of  $e_t$ .*

One standard assumption of the extant literature is that the output process is an arithmetic Brownian motion. In a recent important advance, He (2009) develops a tractable continuous-time agency model with a geometric Brownian motion output process. We too model the output process without effort in (1) as a geometric Brownian motion. However, our paper differs from He (2009) in other modeling aspects. One, the agent is risk-averse in our model while she is risk-neutral in He (2009). Two, effort is restricted to two discrete values (zero or a positive constant) in He (2009) while in our model it is optimally determined without restrictions and is stochastic. Three, managerial payment and consumption occur at the terminal date in our model while they occur intertemporally in He (2009).

## 2.2. The contract space

**Assumption 3** *At time 0, the shareholders offer a compensation,  $S_T$ , which is payable at contract horizon  $T$  to the manager.  $S_T$  is continuous and positive ( $S_T > 0$ ). Otherwise, the contract space of  $S_T$  is unrestricted except that there may be a lower bound  $K_L > 0$  such that  $S_T \geq K_L$ .*

## 2.3. The manager

Besides the arithmetic Brownian motion output process assumption, another standard one is to presume a constant absolute risk aversion (thereafter CARA) utility. While a CARA utility affords tractability in many applications, it exhibits no wealth effect. Instead, we use a CRRA utility.

**Assumption 4** *The manager has a CRRA utility of consuming  $S_T$ . Disutility of exerting costly effort is separable from utility of consumption and given by*

$$C[e_t] = \frac{\kappa}{2} e_t^2, \quad (5)$$

where  $\kappa$  is a constant.

It follows then that the objective of the manager is (for unconditional expectations, we omit the initial information set  $\mathcal{F}_0$  from the expectation operators),

$$\max_{\{e_t\}} V^M [S_T, \{e_t\}_{t=0}^T] = \mathbb{E}^e \left[ \frac{S_T^{1-\gamma}}{1-\gamma} - \int_0^T \frac{\kappa}{2} e_s^2 ds \right], \quad (6)$$

where  $\gamma \geq 1$  is the constant relative risk-aversion parameter.<sup>3</sup>

**Remark 2** *Different from the monetary effort cost in Holmstrom and Milgrom (1987) which is inseparable from the compensation, effort cost in our model is in disutility and separable from the utility from compensation (like in Sannikov (2008)). This modeling feature, coupled with the quadratic form of the cost function  $C[e_t]$ , is the key reason that our model is tractable. See subsection 3.1 and Remark 4 for details.*

## 2.4. The principal

**Assumption 5** *The shareholders are risk-neutral.*

Let  $W_0$  denote the manager's initial wealth,  $\varepsilon_0$  her reservation wage, and  $S_T$  the sum of  $W_0$  and compensation. Accordingly, the shareholders solve the following problem:

$$\max_{S_T} V^S[S_T] = \max_{S_T} \mathbb{E}^{e^*} [V_T - (S_T - W_0)] = W_0 + \max_{S_T} \mathbb{E}^{e^*} [V_T - S_T] \quad (7)$$

such that

$$\{e_t^*\}_{t=0}^T \in \arg \text{maximizes } V^M [S_T, \{e_t\}_{t=0}^T], \quad (\text{IC})$$

$$V^M [S_T, \{e_t^*\}_{t=0}^T] \geq U(W_0 + \varepsilon_0). \quad (\text{IR})$$

## 3. Optimal Strategies

### 3.1. The manager's problem: optimal effort

We follow Sannikov (2008) to derive the optimal effort. His approach relies on taking the manager's continuation value  $Y_t$  as a state variable.  $Y_t$  is the manager's expected utility at time  $t$ , given that she exerts the principal's desirable effort from  $t$  onwards. That is,  $Y_t$  is given by

$$Y_t = \mathbb{E}^e \left[ \frac{S_T^{1-\gamma}}{1-\gamma} - \int_t^T \frac{\kappa}{2} e_s^2 ds \mid \mathcal{F}_t \right], \quad Y_T = \frac{S_T^{1-\gamma}}{1-\gamma}. \quad (8)$$

---

<sup>3</sup>For an integrability condition, we require that  $\gamma \geq 1$ . See footnote 5 for detail.

Obviously,  $Y_t - \int_0^t \frac{\kappa}{2} e_s^2 ds$  is a  $P^e$ -martingale. By the martingale representation theorem, there exists a  $\beta_t$  process such that

$$dY_t = \frac{\kappa}{2} e_t^2 dt + \beta_t dB_t^e. \quad (9)$$

By equation (4) in Sannikov (2008), the optimal effort is determined by

$$e_t^* = \arg \min \left( \frac{1}{2} \kappa e_t^2 - \frac{\beta_t}{\sigma} e_t \right) = \beta_t / (\kappa \sigma). \quad (10)$$

Therefore, in equilibrium

$$dY_t = \frac{\kappa}{2} (e_t^*)^2 dt + \kappa \sigma e_t^* dB_t^e = \frac{\kappa}{2} (e_t^*)^2 dt + \kappa \sigma e_t^* \left( dB_t^0 - \frac{e_t^*}{\sigma} dt \right) = -\frac{\kappa}{2} (e_t^*)^2 dt + \kappa \sigma e_t^* dB_t^0. \quad (11)$$

**Remark 3** To demystify the optimal effort condition (10), we offer the following heuristic derivation.<sup>4</sup> Since  $dB_t^e = dB_t^0 - (e_t/\sigma)dt$ , we can write (9) as

$$dY_t = \left( \frac{\kappa}{2} e_t^2 - \frac{\beta_t}{\sigma} e_t \right) dt + \beta_t dB_t^0. \quad (12)$$

Therefore,

$$Y_t = Y_T - \int_t^T \left( \frac{\kappa}{2} e_s^2 - \frac{\beta_s}{\sigma} e_s \right) ds - \int_t^T \beta_s dB_s^0 \quad (13)$$

where the terminal utility  $Y_T$  is given in (8). The reason that we have transformed to  $B_s^0$  is that the measure  $P^0$  is not affected by effort  $e_t$ . Thus, for a given  $\beta_s$ ,  $t \leq s \leq T$ , we can maximize  $Y_t$  in (13) by choosing  $e_t$  to maximize the second term (i.e. minimize the integrand  $\kappa e_s^2/2 - \beta_s e_s/\sigma$  at each  $s$ ).

It is clear from (11) that  $\beta_t$  (equivalently  $e_t^*$ ) is determined by the diffusion function of  $Y_t$ . The form in (11) suggests an exponential transformation. Let

$$Y_t^{e^*} = \exp \left( \frac{Y_t}{\kappa \sigma^2} \right).$$

By Ito's lemma we have

$$dY_t^{e^*} = Y_t^{e^*} \left( \frac{e_t^*}{\sigma} \right) dB_t^0; \quad Y_T^{e^*} = \exp \left( \frac{S_T^{1-\gamma}}{(1-\gamma)\kappa\sigma^2} \right). \quad (14)$$

<sup>4</sup>For a more rigorous treatment, see Sannikov (2008).

It is clear from (14) that the diffusion of  $Y_t^{e^*}$  determines  $e_t^*$ . Furthermore,  $Y_t^{e^*}$  is a  $P^0$ -martingale and has the following conditional expectation representation:<sup>5</sup>

$$Y_t^{e^*} = \mathbb{E}^0 \left[ Y_T^{e^*} | \mathcal{F}_t \right] = \mathbb{E}^0 \left[ \exp \left( \frac{S_T^{1-\gamma}}{(1-\gamma)\kappa\sigma^2} \right) | \mathcal{F}_t \right]. \quad (15)$$

It follows then that the manager's initial optimal expected utility is given by

$$V^M [S_T, \{e_t^*\}_{t=0}^T] = Y_0 = \kappa\sigma^2 \log(Y_0^{e^*}) = \kappa\sigma^2 \log \left( \mathbb{E}^0 \left[ \exp \left( \frac{S_T^{1-\gamma}}{(1-\gamma)\kappa\sigma^2} \right) \right] \right). \quad (16)$$

Equation (15) makes it clear that under optimal managerial effort  $e_t^*$ , the manager's utility depends only on the compensation  $S_T$  and measure  $P^0$  which is independent of  $e_t^*$ . So we have eliminated the effort choice from the decision-making process. It follows that the key is to determine the optimal contract  $S_T$ . We turn our attention to that in section 3.3.

**Remark 4** *Two modeling features allow us to obtain the striking result (15). The first is that the cost function is separable in utility. This allows us to obtain (10). The second is that the cost function is quadratic. This allows us to obtain (11). If the cost function is not quadratic but still separable in utility, we can still derive  $e_t^*$  as in (10). For example, if  $C[e_t] = \kappa e_t^4/4$ ,  $e_t^* = (\beta_t/(\kappa\sigma))^1/3$ . However, we no longer have the exponential martingale form (11). Both features are required to deliver (15).*

### 3.2. Separation of optimal effort and change of measure

**Proposition 1** *Under optimal effort, the change of measure has the following representation:*

$$M_T^{e^*} = Y_T^{e^*} / Y_0^{e^*} = \exp \left( \frac{S_T^{1-\gamma}}{(1-\gamma)\kappa\sigma^2} \right) / \mathbb{E}^0 \left[ \exp \left( \frac{S_T^{1-\gamma}}{(1-\gamma)\kappa\sigma^2} \right) \right]. \quad (17)$$

**Proof:** Note that under optimal effort  $e_t^*$ , the change of measure in (2) becomes

$$dM_t^{e^*} = \frac{e_t^*}{\sigma} M_t^{e^*} dB_t^0, \quad M_0^{e^*} = 1. \quad (18)$$

<sup>5</sup>Integrability in (15) requires  $\gamma \geq 1$ . To see this, consider the case where the manager is offered a constant fraction,  $\rho$ , of the output value  $V_T$  (*non-optimal* compensation) and she exerts a constant effort  $e$  (*non-optimal* effort). It is easy to check that both the manager's net utility and the equity value approach infinity as  $e$  approaches infinity if  $0 \leq \gamma < 1$ .

On the other hand, we have just seen from (14) and (15) that the manager's utility satisfies

$$dY_t^{e^*} = \frac{e_t^*}{\sigma} Y_t^{e^*} dB_t^0, \quad Y_0^{e^*} = \mathbb{E}^0 \left[ \exp \left( \frac{S_T^{1-\gamma}}{(1-\gamma)\kappa\sigma^2} \right) \middle| \mathcal{F}_0 \right]. \quad (19)$$

Comparing (18) and (19) we see that  $M_t^{e^*} = Y_t^{e^*} / Y_0^{e^*}$ . **Q.E.D.**

**Remark 5** *In equilibrium, the change of measure,  $M_T^{e^*}$ , depends on compensation  $S_T$ , but does not depend on  $e_t^*$  explicitly. This property helps make the principal's problem much easier as we see next. For a similar representation of the change of measure, see Cvitanić, Wan, and Zhang (2009). Again, this result results from the two modeling features discussed in Remark 4.*

### 3.3. The principal's problem: optimal compensation

Given the change of measure in (17), the principal's problem becomes

$$\max_{S_T} \mathbb{E}^{e^*} [V_T - S_T] \equiv \max_{S_T} \mathbb{E}^0 [M_T^{e^*} (V_T - S_T)] = \max_{S_T} \frac{\mathbb{E}^0 \left[ \exp \left( \frac{S_T^{1-\gamma}}{(1-\gamma)\kappa\sigma^2} \right) (V_T - S_T) \right]}{\mathbb{E}^0 \left[ \exp \left( \frac{S_T^{1-\gamma}}{(1-\gamma)\kappa\sigma^2} \right) \right]}, \quad (20)$$

subject to

$$\mathbb{E}^0 \left[ \exp \left( \frac{S_T^{1-\gamma}}{(1-\gamma)\kappa\sigma^2} \right) \right] \geq \exp \left( \frac{(W_0 + \varepsilon_0)^{1-\gamma}}{(1-\gamma)\kappa\sigma^2} \right). \quad (21)$$

**Proposition 2** *For a given constant  $\lambda$ , let  $S_T^\lambda$  be the solution of the following equation*

$$S_T + \kappa\sigma^2 S_T^\gamma = V_T + \lambda \quad (22)$$

when  $V_T + \lambda > 0$  and for a positive  $K_L$  define  $V^*$  by

$$V^* = K_L + \kappa\sigma^2 K_L^\gamma - \lambda. \quad (23)$$

Then, the optimal compensation with a lower bound  $K_L$  is given by

$$S_T^*(\lambda) = K_L 1(V_T \leq V^*) + S_T^\lambda 1(V_T > V^*) \quad (24)$$

where  $1(\cdot)$  is the indicator function. In particular, (i) if reservation  $\varepsilon_0 < \varepsilon_0^L$  where  $\varepsilon_0^L$  is defined in (A.9), the reservation constraint (21) is not binding and  $\lambda$  is determined by (A.6) and negative; (ii) if, on the other hand,  $\varepsilon_0 \geq \varepsilon_0^L$ , the reservation constraint is binding and  $\lambda$  is determined by (A.11) and can be negative or positive. Moreover, if  $\varepsilon_0 > \varepsilon_0^M$  where  $\varepsilon_0^M > \varepsilon_0^L$  and is defined in (A.14),  $\lambda > 0$ . In this case, an explicit lower bound is not necessary if  $\lambda$  is large enough so that  $V^* < 0$ .<sup>6</sup>

**Remark 6** We have reduced the difficult problem of determining the optimal contract to the simple algebraic equation (22). Clearly, the optimizing  $S_T$  depends only on  $V_T$ .

We will focus on our discussions when  $\lambda > 0$ . We consider two cases separately: (1) there is no explicit lower bound on the compensation; and (2) there is an explicit positive lower bound.

### 3.4. Optimal contract without an explicit lower bound constraint

In this case,  $S_T^* = S_T^\lambda$ . Generally, an explicit solution for  $S_T^*$  is not available and the simple algebraic equation (22) is solved numerically. However, there are special cases where explicit solutions are available.

**Case 1:**  $\gamma = 1$ . The optimal contract is linear and given by

$$S_T^* = \frac{V_T + \lambda}{1 + \kappa\sigma^2}. \quad (25)$$

**Case 2:**  $\gamma = 2$ . In this case the optimal contract is nonlinear and given by

$$S_T^* = \sqrt{\frac{V_T + \lambda}{\kappa\sigma^2} + \left(\frac{1}{2\kappa\sigma^2}\right)^2} - \frac{1}{2\kappa\sigma^2}. \quad (26)$$

**Remark 7** We have obtained an optimal linear contract in a log utility–geometric Brownian motion output setting.<sup>7</sup> In contrast, the classical result of optimal linear contracts is obtained in the standard CARA utility–Brownian motion output settings of Holmstrom and Milgrom (1987), Schättler and Sung (1993).

<sup>6</sup>See Appendix A for proof.

<sup>7</sup>The output process in our model is a geometric Brownian motion *only* under the probability space  $P^0$ . It is not under the probability space  $P^e$  with managerial effort in (4) because  $e_t$  is stochastic (dependent on  $V_t$ ). In the CARA–Brownian motion models the asset value follows a Brownian motion under both probability measures because the optimal efforts are constants.

**Proposition 3** Generally, we have the following properties. (1) The optimal contract is strictly increasing; and (2) it is concave.

**Proof:** For notational purpose, let  $f[V_T] = S_T^*$ . Since the optimal compensation  $f[V_T]$  satisfies (22), we have the following first and second order derivatives:

$$f'[V_T] = \partial f[V_T]/\partial V_T = \frac{1}{1 + \gamma \kappa \sigma^2 (f[V_T])^{-(1-\gamma)}} > 0, \quad (27)$$

$$f''[V_T] = \partial^2 f[V_T]/\partial V_T^2 = \gamma(1-\gamma) \kappa \sigma^2 (f[V_T])^{\gamma-2} (f'[V_T])^3. \quad (28)$$

$f[V_T]$  is strictly increasing since  $f'[V_T] > 0$ . It is concave since  $f''[V_T] \leq 0$  for  $\gamma \geq 1$ . **Q.E.D.**

### 3.4.1. Optimal contract with an explicit positive lower bound constraint

In this case,  $V^*$  in (23) is positive. We can now rewrite (24) as

$$S_T^* = K_L 1(V_T \leq V^*) + S_T^\lambda 1(V_T > V^*) = K_L + (S_T^\lambda - K_L) 1(V_T > V^*) = K_L + (S_T^\lambda - K_L)^+. \quad (29)$$

The equity-linked compensation (29) has the appearance of stock options except that generally it depends on the asset value in complicated ways. However, when the manager has a log-utility, the equity-linked part becomes stock options. We formalize this important result in the following Corollary.

**Corollary 1** For the log-utility ( $\gamma = 1$ ) manager with an explicit positive lower bound, the optimal compensation becomes cash plus stock options and is given by<sup>8</sup>

$$S_T^* = K_L + \frac{1}{1 + \kappa \sigma^2} (V_T - ((1 + \kappa \sigma^2)K_L - \lambda))^+ \quad (30)$$

where  $\lambda$  is determined by

$$\mathbb{E}^0 \left[ (S_T^*)^{\frac{1}{\kappa \sigma^2}} \right] = (W_0 + \varepsilon_0)^{\frac{1}{\kappa \sigma^2}}. \quad (31)$$

**Remark 8** For a manager who is less efficient (i.e. higher  $\kappa$ ) or a firm with higher risk ( $\sigma$ ) or both, the shareholders should provide the manager with fewer number of stock options. It is also tempting to think

<sup>8</sup>Recall that with  $\gamma = 1$ ,  $S_T^\lambda = (V_T + \lambda)/(1 + \kappa \sigma^2)$ .

that the strike price,  $(1 + \kappa\sigma^2)K_L - \lambda$ , should also be higher. However, this is not so and in fact the strike price should be lower in order to satisfy the manager's reservation. The reason is that  $\lambda$  increases more than  $(1 + \kappa\sigma^2)K_L$  does if  $\sigma^2$  increases. The shareholders can reward and retain the more efficient (smaller  $\kappa$ ) managers by providing them a greater number of stock options but with a higher strike price.

## 4. Optimal effort

With the optimal compensation determined by (22), from (14) the optimal effort itself can in turn be recovered by matching the diffusion of  $dY_t^{e^*}$  with  $Y_t^{e^*} \left(\frac{e_t^*}{\sigma}\right)$ . Let  $J[t, V_t] = Y_t^{e^*}$ . Then, we have

$$J[t, V_t] \left(\frac{e_t^*}{\sigma}\right) = \text{diffusion of } dJ[t, V_t] = \sigma V_t \frac{\partial J[t, V_t]}{\partial V_t}.$$

Therefore,  $e_t^*$  is determined by

$$e_t^* = \frac{\sigma^2 V_t}{J[t, V_t]} \frac{\partial J[t, V_t]}{\partial V_t} = \sigma^2 \frac{\partial \log(J[t, V_t])}{\partial \log(V_t)}. \quad (32)$$

Recall that  $J[t, V_t]$  is obtained through (15).

### 4.1. Explicit solutions of the optimal effort

In our CRRA utility-separable disutility cost and geometric Brownian motion setting, the optimal effort is both time-dependent and state-dependent (stochastic),<sup>9</sup> and in general cannot be solved explicitly. However, when  $\gamma = 1$  (log-utility) and  $\kappa\sigma^2 = 1/n$  where  $n$  is a positive integer, we can calculate it explicitly.

To this end, note that in this case (15) becomes

$$\begin{aligned} J[t, V_t] &= \mathbb{E}^0 [S_T^n | \mathcal{F}_t] = \mathbb{E}^0 \left[ \left( \frac{V_T + \lambda}{1 + 1/n} \right)^n | \mathcal{F}_t \right] = \left( \frac{n}{1+n} \right)^n \sum_{m=0}^n C_n^m \mathbb{E}^0 [V_T^m | \mathcal{F}_t] \lambda^{n-m} \\ &= \left( \frac{n}{1+n} \right)^n \sum_{m=0}^n C_n^m V_t^m \lambda^{n-m} \exp\{(\mu + 0.5\sigma^2(m-1))m(T-t)\}, \end{aligned} \quad (33)$$

<sup>9</sup>We note that even in the CARA utility-inseparable monetary cost setting but with geometric Brownian motion output process, the optimal effort is also stochastic as in He (2011).

where  $C_n^m$  denotes the combinatory coefficient. Now by (32) straightforward calculations yield

$$e_t^* = \frac{1}{n\kappa} \frac{\sum_{m=1}^n m C_n^m V_t^m \lambda^{n-m} \exp\{(\mu + 0.5\sigma^2(m-1))m(T-t)\}}{\sum_{m=0}^n C_n^m V_t^m \lambda^{n-m} \exp\{(\mu + 0.5\sigma^2(m-1))m(T-t)\}} \quad (34)$$

which is clearly time-dependent and stochastic.

In Appendix B, we prove that  $e_t^*$  is increasing in  $V_t$  and approaches zero as  $V_t$  approaches zero and approaches the upper limit  $1/\kappa$  as  $V_t \rightarrow \infty$ . When  $n = 1$ , it is particularly simple and given by

$$e_t^* = \frac{1}{\kappa} \frac{V_t e^{\mu(T-t)}}{V_t e^{\mu(T-t)} + \lambda} = \frac{1}{\kappa \left(1 + \frac{\lambda}{V_t} e^{-\mu(T-t)}\right)} \quad (35)$$

which is obviously increasing in  $V_t$  and approaches the limit  $1/\kappa$ . For this special case,  $e_t^*$  has no *explicit* time dependence if  $\mu = 0$  and is decreasing in  $t$  if  $\mu > 0$  and increasing if  $\mu < 0$ .

## 4.2. General properties and numerical examples of the optimal effort

Except the previous special case, the optimal effort needs to be solved numerically. Before we do so, we state some general properties whose proofs are provided in Appendix B.

**Proposition 4** *The optimal effort (1) approaches zero as  $V_t$  approach zero and increases with  $V_t$  initially; and (2) approaches the upper limit  $1/\kappa$  if  $\gamma = 1$  and approaches zero if  $\gamma > 1$  as  $V_t$  approaches infinity.*

We have done extensive numerical calculations for the cases with  $\gamma = 1, 2, 3$ . Most results for different  $\gamma$ 's are similar. To save space we mostly report results for  $\gamma = 2$ .

In Figure 1 we plot the optimal effort (as a function of intermediate asset value  $V_t$ ) for a given set of model parameters for  $\gamma = 1$  and  $\gamma = 2$  without or with a lower bound on the compensation. The optimal effort is low when  $V_t$  is very low. There are two reasons. First, due to the scale effect, the effect of  $e_t^*$  on the output value is small when  $V_t$  is small. Second, the equity-linked component in the compensation will be small. This provides less incentive for exerting costly effort. However, initially, the optimal effort increases quickly as  $V_t$  increases. This is still due to the scale effect of the output process dynamics in (4). That is, the same effort level increases output value by a larger amount when  $V_t$  is higher. Therefore, there is a strong incentive to increase  $V_t$  when it is low by exerting higher effort. However, above a critical

level, effort may decline with  $V_t$  because now the wealth effect dominates. The reason is that when  $V_t$  is below the critical value, the marginal benefit (increased utility from consumption) more than offsets the marginal cost (increased dis-utility of a higher effort level). However, due to the wealth effect, when the compensation is already likely to be high enough (resulting from a high  $V_t$ ), the added benefit of a higher effort level may not offset the increased dis-utility. Therefore, the manager optimally exerts less effort.<sup>10</sup> This is consistent with the result in Sannikov (2008) that “when the agent’s consumption is high, it costs too much to compensate him for positive effort.”

By contrast, in the Brownian output model of Homstrom and Milgrom (1987), effort affects the *level* of future output but not the *scale*. Furthermore, effort cost is modeled as an inseparable monetary cost in Homstrom and Milgrom while in our model it is in separable (dis)utility. The two modeling features of level effect of effort and inseparable monetary cost of effort jointly determine that optimal effort is a constant in such a model.

In Figures 2-4 we plot the optimal efforts with various changes of the model parameters. Figure 2 indicates that there exists an inverse relation between optimal effort and  $\varepsilon_0$ . The reason is that to satisfy a higher reservation utility (wage), a larger  $\lambda$  in (22) is required. A larger  $\lambda$  has the effect of reducing the relative portion of equity-linked component in the compensation, resulting in a lower optimal effort level.

However, we caution that a manager’s reservation is likely positively correlated with her skill/capability, measured by the inverse of the cost of effort  $\kappa$ . Indeed, Figure 3 reveals that  $\kappa$  has a large impact on the optimal effort level. Managers with higher reservation because of lower  $\kappa$  can optimally exert higher effort.

Figure 4 plots the effort for different output volatility  $\sigma$ . As expected, higher  $\sigma$  leads to lower effort level because any linked-equity compensation is riskier for higher  $\sigma$ . Ideally, output volatility should be under the control of the manager. In such a model, there exists a trade-off of lower risk level and the cost of managing risk. It can also be the case that the expected growth rate of output process is (possibly positively) related to the risk level. In either case, the risk level can be endogenously determined. For tractability and our solution techniques to apply, output risk level is exogenously specified in our model.

In Figure 5 we present the optimal effort as a function of the time to terminal date  $T - t$  and intermediate asset value (state variable) in a 3D plot. When  $V_t$  is small,  $e_t^*$  is an increasing function of  $T - t$ . The reason

---

<sup>10</sup>The wealth effect argument also applies to the case for  $\gamma = 1$  (log-utility). However, given the quadratic cost function specification, the wealth effect is not enough to offset the scale effect of a higher effort level for a higher  $V_t$ . Thus, the optimal effort is an increasing function of  $V_t$  in this case.

is that in this case the scale effect dominates (the wealth effect). When  $T - t$  is larger,  $e_t$  can affect the terminal output process value through a longer period. This is unlike the arithmetic Brownian motion models where the effort affects the level and current effort has no impact on later effort. In our model, current effort affects later effort through its effect on the output process which in turn affects future effort level since effort depends on output level. In other words, effort in our model has a *propagating* effect. On the other hand, when  $V_t$  is large, the wealth effect dominates and  $e_t$  is a decreasing function of  $T - t$ . The longer is  $T - t$ , the larger is its impact on  $V_T$ . Since the compensation  $S_T$  is an increasing function of  $V_T$ ,  $T - t$  has a positive impact on  $S_T$  as well. Due to the wealth effect, the marginal benefit of  $e_t$  declines when  $S_T$  is expected to be high and higher when  $T - t$  is longer. The manager optimally reduces her effort level. Time to terminal date  $T - t$  and intermediate asset value  $V_t$  jointly affect optimal effort  $e_t^*$  non-monotonically. The implication is that a shorter maturity is called for when the equity-linked component is likely to yield high payoffs, e.g. in-the-money options. When the equity-linked component is likely to yield lower payoffs, e.g. out-of-the-money options, a longer maturity is more appropriate.

## 5. Pay-performance sensitivity, firm performance, and firm risk

One central prediction of standard agency theory is that firm performance should be positively correlated with PPS. The empirical evidence has been mixed. Another central prediction of standard agency theory is that PPS should be negatively correlated with firm risk. Again, the empirical evidence has been mixed. However, these predictions are mostly based on the CARA utility–Brownian motion settings and may not hold in other theoretical settings like ours.

For nonlinear contracts, the literature has generally used the delta ( $\Delta$ ) of compensation as a measure of PPS. For example, the Black-Scholes delta has been used as a measure of PPS for option-type compensations. It is easy to check that the Black-Scholes  $\Delta$  is the expectation of the derivative of the terminal payoff.<sup>11</sup> Similarly, for our nonlinear contracts we use  $E^{e^*}[\frac{\partial}{\partial V_T} f[V_T]] = E^0[M_T f'[V_T]]$  as a measure of PPS where  $f'[V_T]$  is given in (27).

In Figures 6 and 7 we plot the PPS, expected total output value, equity (expected residual) value as a function of  $\epsilon_0$ ,  $\sigma$ , and PPS as a function of the equity value when  $\gamma = 2$ . They indicate that, depending on the

<sup>11</sup>To see this, note that  $\frac{\partial}{\partial V_T} (V_T - K)^+ = 1(V_T > K)$  where  $K$  is the strike price. It can be easily checked that the Black-Scholes  $\Delta = E[1(V_T > K)]$  where  $\frac{dV_t}{V_t} = (r + \sigma^2)dt + \sigma dB_t$ ,  $r$  is the riskless rate,  $\sigma$  the volatility, and  $B_t$  a standard Wiener process.

structural parameter values, the relation between PPS and equity value and the relation between PPS and firm risk can be increasing or decreasing. This is consistent with the empirical evidence. In Figure 8 we plot the relation between PPS and firm performance for different compensation time-to-maturity. It indicates that performance is higher if the time-to-maturity is longer. Therefore, controlling time-to-maturity is an important consideration in empirical work.

We note that the relation between PPS and  $\varepsilon_0$  (or  $\sigma$ ) has a direct consequence for the relation between PPS and equity value. Changing  $\varepsilon_0$  corresponds to managerial heterogeneity.<sup>12</sup> We caution that it is usually less difficult to find some heterogeneous variable such as  $\varepsilon_0$  to explain the relation between two endogenous variables such as PPS and equity value.<sup>13</sup> On the other hand, explaining the relation between the endogenous PPS and exogenous  $\sigma$  is more demanding. To this end, we offer the following elaboration.

### 5.1. Decomposition of pay-performance sensitivity

To understand the behavior of PPS better and how it differs from that of a linear contract, we consider the following decomposition:

$$PPS = E^{e^*} [f'[V_T]] = E^0 [M_T f'[V_T]] = E^0 [M_T] E^0 [f'[V_T]] + cov^0 (M_T, f'[V_T]) \quad (36)$$

where  $cov^0(\cdot)$  denotes the covariance under measure  $P^0$ . Because  $E^0 [M_T] = 1$ , we have

$$PPS = E^{e^*} [f'[V_T]] = E^0 [f'[V_T]] + cov^0 (M_T, f'[V_T]). \quad (37)$$

Therefore, the PPS which measures the expected (i.e.  $E^{e^*}[\cdot]$ ) pay (i.e.  $F[V_T]$ ) change versus performance (i.e.  $V_T$ ) change (i.e.  $\partial f[V_T]/\partial V_T$ ) under measure  $P^{e^*}$ ,  $E^{e^*} [f'[V_T]]$ , can be decomposed into that under measure  $P^0$ ,  $E^0 [f'[V_T]]$ , plus the covariance between the marginal compensation  $f'[V_T]$  and  $M_T$ . For a linear contract,  $f'[V_T]$  is independent of  $V_T$  and thus covariance is zero. Generally, for a nonlinear contract, the covariance measures the difference between the PPS under the probability measure with effort ( $P^{e^*}$ ) and that without effort ( $P^0$ ). In our model with  $\gamma = 2$ ,  $f[V_T]$  is given by (26) and  $f'[V_T]$  is given

$$f'[V_T] = \frac{1}{\sqrt{1 + 2\kappa\sigma^2(V_T + \lambda)}}.$$

<sup>12</sup>We have also calculated these relations by changing  $W_0$  or  $\kappa$ . Similar results are obtained and thus not reported.

<sup>13</sup>We thank the referee's urging us to be cautious on relation between two endogenous variables using a heterogeneous variable.

To satisfy a given reservation  $\varepsilon_0$ ,  $\lambda$  increases with  $\sigma$ . So  $f'[V_T]$  decreases with  $\sigma$  for a given  $V_T$ . It is intuitive then that  $E^0[f'[V_T]]$  decreases with  $\sigma$ .<sup>14</sup> We plot this term in Panel A of Figure 9. It is noted that it is indeed decreasing in volatility and conforms with the usual classical result in models with optimal linear contract such as Holmstrom and Milgrom (1987). This decreasing relation appears to be slightly convex.

On the other hand, we note that

$$M_T = \exp\left(-\frac{1}{\kappa\sigma^2 S_T}\right) / \exp\left(-\frac{1}{\kappa\sigma^2 \varepsilon_0}\right) = \exp\left(-\frac{1}{\sqrt{\kappa\sigma^2(V_T + \lambda) + 1/4} - 1/2}\right) / \exp\left(-\frac{1}{\kappa\sigma^2 \varepsilon_0}\right).$$

It is easy to see that  $f'[V_T]$  is decreasing in  $V_T$  while  $M_T$  is increasing and thus the covariance is negative. Moreover, as  $\sigma$  increases,  $f'[V_T]$  becomes less dependent on  $V_T$ . For example, as  $\sigma$  becomes very large,  $f'[V_T]$  becomes very small (close to zero) regardless of the value of  $V_T$ . Therefore, the covariance is *expected* to become smaller in *magnitude* as  $\sigma$  increases. Indeed, Panel B of Figure 9 indicates that it is negative but increasing in volatility (the magnitude becomes smaller). This increasing relation is concave.

The sum of the two panels in Figure 9 yields the PPS relation in Panel A in Figure 7. Since the concavity of Panel B is more than the convexity of Panel A in Figure 9, their sum results in a non-monotone relation. Thus, the nonlinearity which yields the covariance term plays an important role of this relation.<sup>15</sup>

## 6. Empirical implications

We propose a few more specific predictions. First, the prediction that the relationship between PPS and performance is stronger if time-to-maturity is longer can be tested by the following exercise:<sup>16</sup>

$$\text{firm performance} = a + bPPS + c(PPS \cdot \text{employment\_tenure}) + \tilde{\varepsilon} \quad (38)$$

where  $\tilde{\varepsilon}$  is a white noise and  $c$  is predicted to be positive.

<sup>14</sup>While this is intuitive, it is not proof since the distribution of  $V_T$  depends on  $\sigma$ .

<sup>15</sup>Similarly, we can define utility-performance-sensitivity (UPS) as  $E^{e*} \left[ \frac{\partial}{\partial V_T} Y[V_T] \right] = E^0 [M_T f^{-\gamma}[V_T] f'[V_T]]$  and examine its relation with risk accordingly.

<sup>16</sup>We thank the referee for suggesting the test.

Second, the prediction that the relationship between PPS and firm risk is increasing for lower risk firms and decreasing for higher risk firms can be tested using firm samples sorted into risk groups:

$$\text{PPS} = a + b \sigma d_1(\text{lowest risk group}) + c \sigma d_2(\text{highest risk group}) + \tilde{\epsilon} \quad (39)$$

where  $d_1()$  and  $d_2()$  are two dummy variables and  $b$  is predicted to be positive and  $c$  negative.

Third, based on the discussion at the end of Section 4, we propose the following test on the relationship between the moneyness and maturity of executive stock options:

$$\text{Moneyness} = a + b \text{maturity} + \tilde{\epsilon} \quad (40)$$

where  $b$  is predicted to be negative.

While options are typically issued at the money, a manager's portfolio may contain existing options with different moneyness and maturities. The overall moneyness and maturity can be constructed, for example, using value-weighted average. The average values of moneyness and maturity can be endogenously affected by choosing the amount and maturity of newly issued options even if they are issued at the money. The prediction in (40) can be tested cross-sectionally using weighted values.

## 7. Conclusion

In this paper we have studied the optimal contracting problem between risk-neutral shareholders and a CRRA manager. In our model the manager controls the expected growth rate of an asset value process by exerting costly effort. In our characterization the optimal compensation function is the solution to an algebraic equation and an increasing function of the terminal output value. Generally, neither restricted shares (linear contracts) nor stock options are optimal. The optimal compensation function is linear in the terminal output value if the manager has a log-utility and concave if he is more risk-averse.

The optimal effort in our model is stochastic and is in contrast with the results of constant or deterministic efforts in the standard CARA utility–Brownian motion output settings. We find that in our model the optimal effort is increasing in the intermediate asset value if the manager has a log-utility and is increasing and then decreasing if the manager is more risk-averse.

Numerical calculations show that the relationship between PPS and firm performance and that between PPS and firm risk can be non-monotonic. These predictions are consistent with empirical findings. We have also proposed a few more testable predictions.

## A. Characterizing the Principal's Maximization Problem

Form the Lagrangian:

$$L = \frac{\mathbb{E}^0 \left[ \exp \left( \frac{S_T^{1-\gamma}}{(1-\gamma)\kappa\sigma^2} \right) (V_T - S_T) \right]}{\mathbb{E}^0 \left[ \exp \left( \frac{S_T^{1-\gamma}}{(1-\gamma)\kappa\sigma^2} \right) \right]} + \eta \left( \mathbb{E}^0 \left[ \exp \left( \frac{S_T^{1-\gamma}}{(1-\gamma)\kappa\sigma^2} \right) \right] - \exp \left( \frac{(W_0 + \varepsilon_0)^{1-\gamma}}{(1-\gamma)\kappa\sigma^2} \right) \right). \quad (\text{A.1})$$

Let

$$N = \mathbb{E}^0 \left[ \exp \left( \frac{S_T^{1-\gamma}}{(1-\gamma)\kappa\sigma^2} \right) (V_T - S_T) \right], \quad D = \mathbb{E}^0 \left[ \exp \left( \frac{S_T^{1-\gamma}}{(1-\gamma)\kappa\sigma^2} \right) \right].$$

The derivative of  $L$  with respect to  $S_T$  is:

$$\frac{\partial L}{\partial S_T} = \frac{\mathbb{E}^0 \left[ \exp \left( \frac{S_T^{1-\gamma}}{(1-\gamma)\kappa\sigma^2} \right) \left( \frac{S_T^{-\gamma}}{\kappa\sigma^2} (V_T - S_T) - 1 \right) \right] D - \mathbb{E}^0 \left[ \exp \left( \frac{S_T^{1-\gamma}}{(1-\gamma)\kappa\sigma^2} \right) \frac{S_T^{-\gamma}}{\kappa\sigma^2} \right] N}{D^2} + \eta \mathbb{E}^0 \left[ \exp \left( \frac{S_T^{1-\gamma}}{(1-\gamma)\kappa\sigma^2} \right) \frac{S_T^{-\gamma}}{\kappa\sigma^2} \right]. \quad (\text{A.2})$$

The Kuhn-Tucker first-order conditions (state-by-state, i.e. for each  $V_T$ ) are:

$$S_T + \kappa\sigma^2 S_T^\gamma = V_T + \eta D - N/D, \quad \eta \geq 0, \quad \text{and} \quad (\text{A.3})$$

$$\eta \left( \mathbb{E}^0 \left[ \exp \left( \frac{S_T^{1-\gamma}}{(1-\gamma)\kappa\sigma^2} \right) \right] - \exp \left( \frac{(W_0 + \varepsilon_0)^{1-\gamma}}{(1-\gamma)\kappa\sigma^2} \right) \right) = 0. \quad (\text{A.4})$$

### A.1. IR constraint (21) is not binding

In this case (A.4) implies  $\eta = 0$  and we can rewrite (A.3) as

$$S_T + \kappa\sigma^2 S_T^\gamma = V_T + \lambda, \quad (\text{A.5})$$

where  $\lambda$  is determined by

$$\lambda = -\frac{N}{D} = -\frac{\mathbb{E}^0 \left[ \exp \left( \frac{S_T^{1-\gamma}}{(1-\gamma)\kappa\sigma^2} \right) (V_T - S_T) \right]}{\mathbb{E}^0 \left[ \exp \left( \frac{S_T^{1-\gamma}}{(1-\gamma)\kappa\sigma^2} \right) \right]}. \quad (\text{A.6})$$

We only consider the interesting cases where the equity value is positive. Thus,  $\lambda < 0$  in (A.6).

Let  $S_T^\lambda$  denote the solution of (A.5) when  $V_T + \lambda > 0$ . It is tempting to set  $S_T = 0$  when  $V_T + \lambda \leq 0$ . However, for  $\gamma \geq 1$ , CRRA utility is not well-defined if  $S_T = 0$  for a *positive* probability set  $Pr(V_T \leq -\lambda) > 0$ . Therefore, to have a solution, a positive lower bound  $K_L > 0$  on  $S_T$  needs to be specified. The optimal compensation with a lower positive bound can now be specified as

$$S_T^* = K_L 1(S_T^\lambda \leq K_L) + S_T^\lambda 1(S_T^\lambda > K_L). \quad (\text{A.7})$$

More conveniently, we can define

$$V^* = K_L + \kappa\sigma^2 K_L^\gamma - \lambda$$

and rewrite (A.7) as

$$S_T^* = K_L 1(V_T \leq V^*) + S_T^\lambda 1(V_T > V^*). \quad (\text{A.8})$$

The compensation (A.8) defines a minimum binding reservation level  $\varepsilon_0^L$ :

$$\mathbb{E}^0 \left[ \exp \left( \frac{(S_T^*)^{1-\gamma}}{(1-\gamma)\kappa\sigma^2} \right) \right] = \exp \left( \frac{(W_0 + \varepsilon_0^L)^{1-\gamma}}{(1-\gamma)\kappa\sigma^2} \right). \quad (\text{A.9})$$

When  $\varepsilon_0 < \varepsilon_0^L$ , IR constraint is not binding and the optimal compensation is given by (A.8). When  $\varepsilon_0 \geq \varepsilon_0^L$ , IR constraint is binding and we characterize the principal's maximization in in this case next.

## A.2. IR constraint (21) is binding

In this case  $\eta \geq 0$  and we still have

$$S_T + \kappa\sigma^2 S_T^\gamma = V_T + \lambda, \quad (\text{A.10})$$

where

$$\lambda = \eta D - \frac{N}{D} = \eta \mathbb{E}^0 \left[ \exp \left( \frac{S_T^{1-\gamma}}{(1-\gamma)\kappa\sigma^2} \right) \right] - \frac{\mathbb{E}^0 \left[ \exp \left( \frac{S_T^{1-\gamma}}{(1-\gamma)\kappa\sigma^2} \right) (V_T - S_T) \right]}{\mathbb{E}^0 \left[ \exp \left( \frac{S_T^{1-\gamma}}{(1-\gamma)\kappa\sigma^2} \right) \right]}. \quad (\text{A.11})$$

The optimal compensation with no explicit lower bound is given by  $S_T^* = S_T^\lambda$  and is given by

$$S_T^* = K_L 1(V_T \leq V^*) + S_T^\lambda 1(V_T > V^*) \quad (\text{A.12})$$

when an explicit lower bound  $K_L$  exists. In either case  $\lambda$  is determined by

$$\mathbb{E}^0 \left[ \exp \left( \frac{(S_T^*)^{1-\gamma}}{(1-\gamma)\kappa\sigma^2} \right) \right] = \exp \left( \frac{(W_0 + \varepsilon_0)^{1-\gamma}}{(1-\gamma)\kappa\sigma^2} \right). \quad (\text{A.13})$$

$\eta$  is recovered from (A.11).

By setting  $\lambda = 0$  in (A.10), (A.13) defines reservation  $\varepsilon_0^M$ :

$$\mathbb{E}^0 \left[ \exp \left( \frac{(S_T^*)^{1-\gamma}}{(1-\gamma)\kappa\sigma^2} \right) \right] = \exp \left( \frac{(W_0 + \varepsilon_0^M)^{1-\gamma}}{(1-\gamma)\kappa\sigma^2} \right). \quad (\text{A.14})$$

When  $\varepsilon_0^L < \varepsilon_0 \leq \varepsilon_0^M$ , IR constraint is binding and  $\lambda \leq 0$ . When  $\varepsilon_0 > \varepsilon_0^M$ , IR constraint is binding and  $\lambda > 0$ .

### A.3. Checking second-order condition

First, for a given  $V_T$ , note that the first-order condition (A.5) or (A.10) has a unique solution when  $V_T + \lambda > 0$  since the left hand side is an increasing function of  $S_T$ . Rearranging terms we can rewrite (A.2) as

$$\frac{\partial L}{\partial S_T} = \frac{\mathbb{E}^0 \left[ \exp \left( \frac{S_T^{1-\gamma}}{(1-\gamma)\kappa\sigma^2} \right) \frac{S_T^{-\gamma}}{\kappa\sigma^2} (V_T + \lambda - S_T - \kappa\sigma^2 S_T^\gamma) \right]}{D} \quad (\text{A.15})$$

where  $\lambda = \eta D - N/D$  is determined by (A.6) for the nonbinding case ( $\eta = 0$ ) and by (A.13) for the binding case ( $\eta \geq 0$ ). The second-order derivative is given by

$$\begin{aligned} \frac{\partial^2 L}{\partial S_T^2} = & - \frac{\mathbb{E}^0 \left[ \exp \left( \frac{S_T^{1-\gamma}}{(1-\gamma)\kappa\sigma^2} \right) \frac{S_T^{-\gamma}}{\kappa\sigma^2} (V_T + \lambda - S_T - \kappa\sigma^2 S_T^\gamma) \right] \frac{\partial D}{\partial S_T}}{D^2} + \frac{\mathbb{E}^0 \left[ \exp \left( \frac{S_T^{1-\gamma}}{(1-\gamma)\kappa\sigma^2} \right) \left( \frac{S_T^{-\gamma}}{\kappa\sigma^2} \right)^2 (V_T + \lambda - S_T - \kappa\sigma^2 S_T^\gamma) \right]}{D} + \\ & \frac{\mathbb{E}^0 \left[ \exp \left( \frac{S_T^{1-\gamma}}{(1-\gamma)\kappa\sigma^2} \right) (-\gamma) \frac{S_T^{-\gamma-1}}{\kappa\sigma^2} (V_T + \lambda - S_T - \kappa\sigma^2 S_T^\gamma) \right]}{D} + \frac{\mathbb{E}^0 \left[ \exp \left( \frac{S_T^{1-\gamma}}{(1-\gamma)\kappa\sigma^2} \right) \frac{S_T^{-\gamma}}{\kappa\sigma^2} (-1 - \gamma\kappa\sigma^2 S_T^{\gamma-1}) \right]}{D}. \end{aligned} \quad (\text{A.16})$$

The first three terms are zero at the first-order condition (A.5) or (A.10) and the last term is clearly negative. It follows that the first-order conditions determine the unique maximum for each  $V_T$  when a solution exists.

## B. Proof of Proposition 4

### B.1. For $\gamma > 1$ , $e_t^*$ initially increases with $V_t$ and then drops to zero as $V_t$ approaches infinity

To prove this result, recall that

$$J[t, V_t] = E^0 \left[ \exp \left( \frac{S_T^{1-\gamma}}{(1-\gamma)\kappa\sigma^2} \right) \middle| \mathcal{F}_t \right]. \quad (\text{B.1})$$

Therefore, by (32) we have

$$e_t^* = \frac{\sigma^2 V_t}{J[t, V_t]} \frac{\partial J[t, V_t]}{\partial V_t} = \frac{1}{\kappa E^0 \left[ \exp \left( \frac{S_T^{1-\gamma}}{(1-\gamma)\kappa\sigma^2} \right) \middle| \mathcal{F}_t \right]} E^0 \left[ \exp \left( \frac{S_T^{1-\gamma}}{(1-\gamma)\kappa\sigma^2} \right) S_T^{-\gamma} \frac{\partial S_T}{\partial V_T} V_T \middle| \mathcal{F}_t \right] \quad (\text{B.2})$$

where  $V_t \frac{\partial V_T}{\partial V_t} = V_T$  has been used since

$$V_T = V_t e^{(\mu - \sigma^2/2)(T-t) + \sigma(B_T^0 - B_t^0)}$$

under measure  $P^0$ . Using (27) we can rewrite (B.2) as

$$e_t^* = \frac{1}{\kappa} E^{e^*} \left[ \frac{V_T S_T^{-\gamma}}{1 + \gamma\kappa\sigma^2 S_T^{-(1-\gamma)}} \middle| \mathcal{F}_t \right] \quad (\text{B.3})$$

Now, as  $V_t$  approaches 0, it is clear from (B.3) that  $e_t^*$  approaches zero. Since  $e_t^*$  is obviously positive for  $V_t > 0$ , it follows that  $e_t^*$  is increasing in  $V_t$  initially. To see that it approaches zero (again) as  $V_t$  approaches infinity, we note that from the first order condition (22) we have (for  $\gamma > 1$  as  $V_T \rightarrow \infty$ )

$$S_T \sim \left( \frac{V_T}{\kappa\sigma^2} \right)^{\frac{1}{\gamma}}.$$

Therefore,

$$e_t^* \rightarrow \sigma^2 E e^* \left[ \frac{1}{1 + \kappa\sigma^2(\gamma) \left( \frac{V_T}{\kappa\sigma^2} \right)^{\frac{-(1-\gamma)}{\gamma}} | \mathcal{F}_t} \right] \rightarrow 0. \quad \mathbf{Q.E.D.}$$

### B.2. For $\gamma = 1$ , $e_t^*$ initially increases with $V_t$ and approaches $1/\kappa$ as $V_t$ approaches infinity

In this case ( $\gamma = 1$ ), (B.2) becomes

$$e_t^* = \frac{E^0 \left[ (V_T + \lambda)^{\frac{1}{\kappa\sigma^2} - 1} V_T | \mathcal{F}_t \right]}{\kappa E^0 \left[ (V_T + \lambda)^{\frac{1}{\kappa\sigma^2}} | \mathcal{F}_t \right]} = \frac{1}{k} \left( 1 - \frac{\lambda E^0 \left[ (V_T + \lambda)^{\frac{1}{\kappa\sigma^2} - 1} | \mathcal{F}_t \right]}{E^0 \left[ (V_T + \lambda)^{\frac{1}{\kappa\sigma^2}} | \mathcal{F}_t \right]} \right), \quad (\text{B.4})$$

which approaches 0 as  $V_t \rightarrow 0$  and approaches  $1/\kappa$  as  $V_t \rightarrow \infty$ .  $\mathbf{Q.E.D.}$

Our calculations indicate that  $e_t^*$  is monotonically increasing in  $V_t$  when  $\gamma = 1$ , but we are unable to prove this result. However, when  $\kappa\sigma^2 = 1/n$  where  $n$  is a positive integer, this is true as we show next.

### B.3. For $\gamma = 1$ , $e_t^*$ is an increasing function of $V_t$ if $\kappa\sigma^2 = 1/n$ where $n$ is a positive integer

Note that when  $\kappa\sigma^2 = 1/n$ , (B.4) reduces to (34). Taking derivative of (34) with respect to  $V_t$  and letting  $b[V_t]$  denote the resulting numerator, we have

$$\begin{aligned} V_t b[V_t] &= \sum_{m=1}^n m^2 C_n^m V_t^m \lambda^{n-m} a(m) \sum_{k=0}^n C_n^k V_t^k \lambda^{n-k} a(k) - \sum_{m=1}^n m C_n^m V_t^m \lambda^{n-m} a(m) \sum_{k=1}^n k C_n^k V_t^k \lambda^{n-k} a(k) \\ &= C_n^0 \lambda^n a(0) \sum_{m=1}^n m^2 C_n^m V_t^m \lambda^{n-m} a(m) + \sum_{m=1}^n \sum_{k=1}^n m(m-k) C_n^m C_n^k V_t^{m+k} \lambda^{2n-m-k} a(m) a(k), \end{aligned}$$

where  $a(i)$  is defined as

$$a(i) = \exp\{(\mu + 0.5\sigma^2(i-1))i(T-t)\}.$$

To prove  $e_t^*$  is increasing in  $V_t$ , we only need to prove  $b[V_t] > 0$ . The first term is positive. Now consider the second term which can be rewritten as

$$\begin{aligned} & \sum_{m>k}^n m(m-k)C_n^m C_n^k V_t^{m+k} \lambda^{2n-m-k} a(m)a(k) + \sum_{m<k}^n m(m-k)C_n^m C_n^k V_t^{m+k} \lambda^{2n-m-k} a(m)a(k) = \\ & \sum_{m>k}^n m(m-k)C_n^m C_n^k V_t^{m+k} \lambda^{2n-m-k} a(m)a(k) + \sum_{k<m}^n k(k-m)C_n^m C_n^k V_t^{m+k} \lambda^{2n-m-k} a(m)a(k), \end{aligned}$$

where we have switched the dummy indices  $m$  and  $k$  in the summation in the second term. Thus,

$$\sum_{m=1}^n \sum_{k=1}^n m(m-k)C_n^m C_n^k V_t^{m+k} \lambda^{2n-m-k} a(m)a(k) = \sum_{m>k}^n (m-k)^2 C_n^m C_n^k V_t^{m+k} \lambda^{2n-m-k} a(m)a(k) > 0.$$

Therefore,  $b[V_t] > 0$  and  $e_t^*$  is increasing in  $V_t$ . **QED**

## References

- Aggarwal, R., and A. Samwick, 1999, The other side of the trade-off: The impact of risk on executive compensation, *Journal of Political Economy* 107, 65-105.
- Bushman, R., R. Indejikian, A. Smith, 1996, CEO compensation: The role of individual performance evaluation, *Journal Accounting Economics* 21, 161-193.
- Cadenillas, A., J. Cvitanić, and F. Zapatero, 2004, Leverage decision and manager compensation with choice of effort and volatility, *Journal of Financial Economics* 73, 71-92.
- Carpenter, J., 1998, The exercise and valuation of executive stock options, *Journal of Financial Economics* 48, 127-158.
- Carpenter, J., 2000, Does option compensation increase managerial risk appetite?, *Journal of Finance* 55, 2311-2331.
- Carpenter, J., and B. Remmers, 2001, Executive stock option exercises and inside information, *Journal of Business* 74, 513-534.
- Carpenter, J., R. Stanton, and N. Wallace, 2009, Estimation of employee stock option exercise rates and firm cost: methodology, working paper, New York University and University of California, Berkeley.

- Cvitanović, J., X. Wan and J. Zhang, 2009, Continuous-time principal-agent problems with hidden action and lump-sum payment, *Applied Mathematics and Optimization* 59, 99-146.
- Demsetz, H., and K. Lehn, 1985, The structure of corporate ownership: Causes and consequences, *Journal of Political Economy* 93, 1155-1177.
- Garvey, G., and T. Milbourn, 2003, Incentive compensation when executive can hedge the market: Evidence of relative performance evaluation in the cross section, *Journal of Finance* 58, 1557-1581.
- Garen, J. 1994, Executive compensation and principalagent theory, *Journal of Political Economy* 102, 1175-1199.
- Guo, M., and H. Ou-Yang, 2006, Incentives and performance in the presence of wealth effects and endogenous risk, *Journal of Economic Theory* 129, 150-191.
- He, Z., 2009, Optimal executive compensation when firm size follows geometric Brownian motion, *Review of Financial Studies* 22, 859-892.
- He, Z., 2011, A model of dynamic compensation and capital structure, *Journal of Financial Economics* 100, 351-366.
- Himmelberg, C., R. Hubbard, and D. Palia, 1999, Understanding the determinants of managerial ownership and the links between ownership and firm performance, *Journal of Financial Economics* 53, 353-384.
- Holmstrom, B., and P. Milgrom, 1987, Aggregation and linearity in the provision of intertemporal incentives, *Econometrica* 55, 303-328.
- Ittner, C., D. Larcker, and M. Rajan, 1997, The choice of performance measures in annual bonus contracts, *Accounting Review* 73, 231-255.
- Jewitt, I., O. Kadan, and J. Swinkels, 2008, Moral Hazard with Bounded Payments, *Journal of Economic Theory* 143, 59-82.
- Jin, L., 2002, CEO compensation, diversification, and incentives, *Journal of Financial Economics* 66, 29-63.
- Kadan, O., and J. Swinkels, 2008, Stocks or Options? Moral Hazard, Firm Viability, and the Design of Compensation Contracts, *Review of Financial Studies* 21, 451-482.
- Lambert, R., and D. Larker, 1987, An analysis of the use of accounting and market measures of performance in executive compensation contracts, *Journal of Accounting Research* 25, 85-125.

- Lazear, E., 2000, Performance pay and productivity, *American Economics Review* 90, 1346-1361.
- McConnell, J., and H. Servaes, 1990, Additional evidence on equity ownership and corporate value, *Journal of Financial Economics* 27, 595-612.
- Ou-Yang, H., 2003, Optimal contracts in a continuous-time delegated portfolio management problem, *Review of Financial Studies* 16, 173-208.
- Ou-Yang, H., 2005, An equilibrium model of asset pricing and moral hazard, *Review of Financial Studies* 18, 1353-1303.
- Palia, D., 2001, The endogeneity of managerial compensation in firm valuation: A solution. *Review of Financial Studies* 14, 735-764.
- Prendergast, C., 2002, The tenuous tradeoff between risk and incentives, *Journal of Political Economy* 110, 1071-1102.
- Sannikov, Y., 2008, A continuous-time version of the principal-agent problem, *Review of Economic Studies* 75, 957-984.
- Schättler, H., and J. Sung, 1993, The first-order approach to the continuous-time principal-agent problem with exponential utility, *Journal of Economic Theory* 61, 331-371.
- Yermack, D. 1995, Do corporations award CEO stock options effectively?, *Journal of Financial Economics* 39, 237-270.

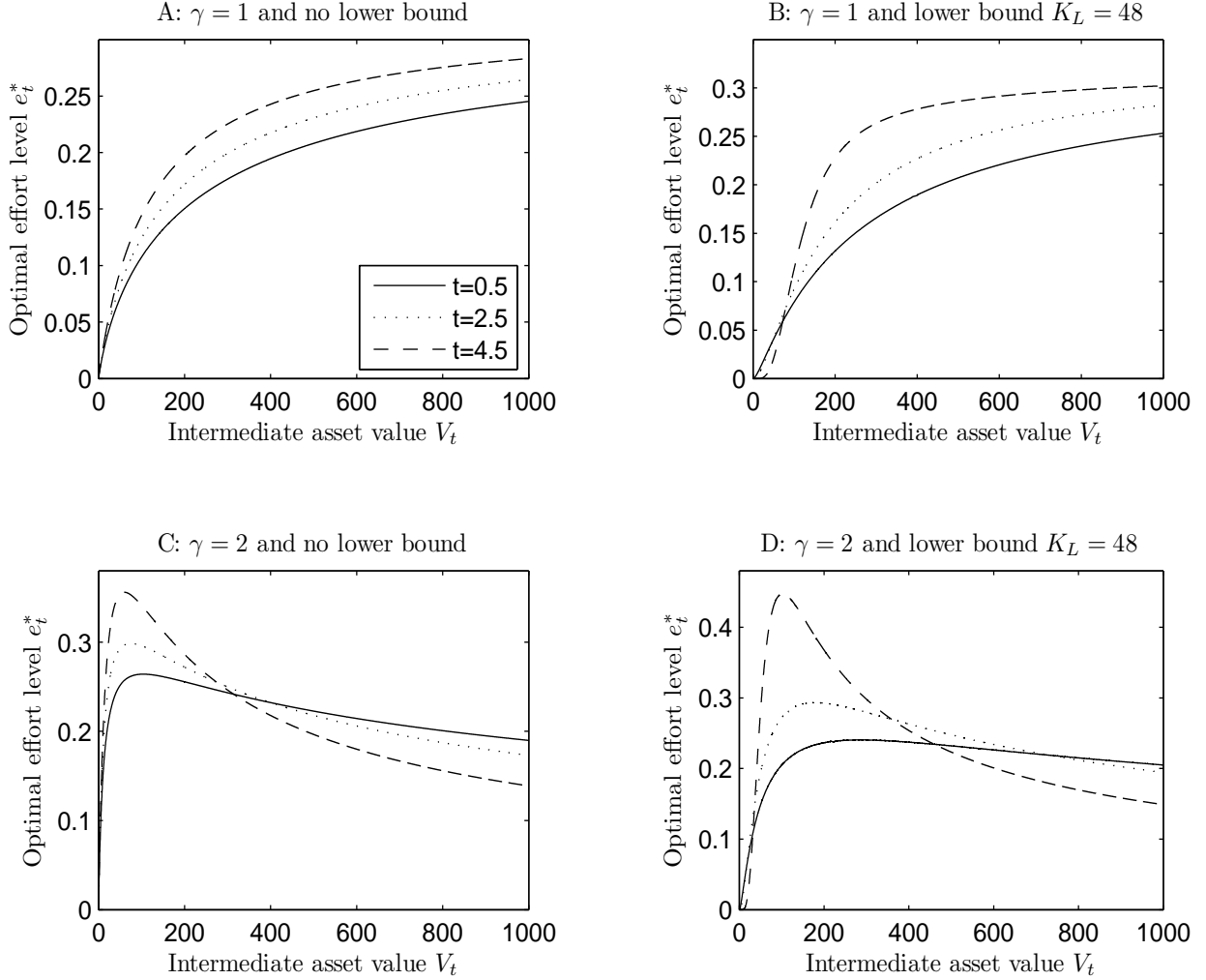


Figure 1: Optimal effort versus intermediate asset value. Panels A and B plot the optimal effort levels under utility  $U(W) = \log(W)$  at  $t = 0.5, 2.5, 4.5$  as a function of the prevailing asset value without and with a lower bound on the compensation, respectively. The parameter values are:  $T = 5, V_0 = 100, \mu = 0, \sigma = 0.8, \kappa = 2/\sigma^2, W_0 = 0, \varepsilon_0 = 60, K_L = 48$ . Panels C and D do the same under  $U(W) = W^{1-\gamma}/(1-\gamma)$  with  $\gamma = 2$ . The parameter values are:  $T = 5, V_0 = 100, \mu = 0, \sigma = 0.8, \kappa = 10^{-2}/\sigma^2, W_0 = 0, \varepsilon_0 = 60, K_L = 48$ .

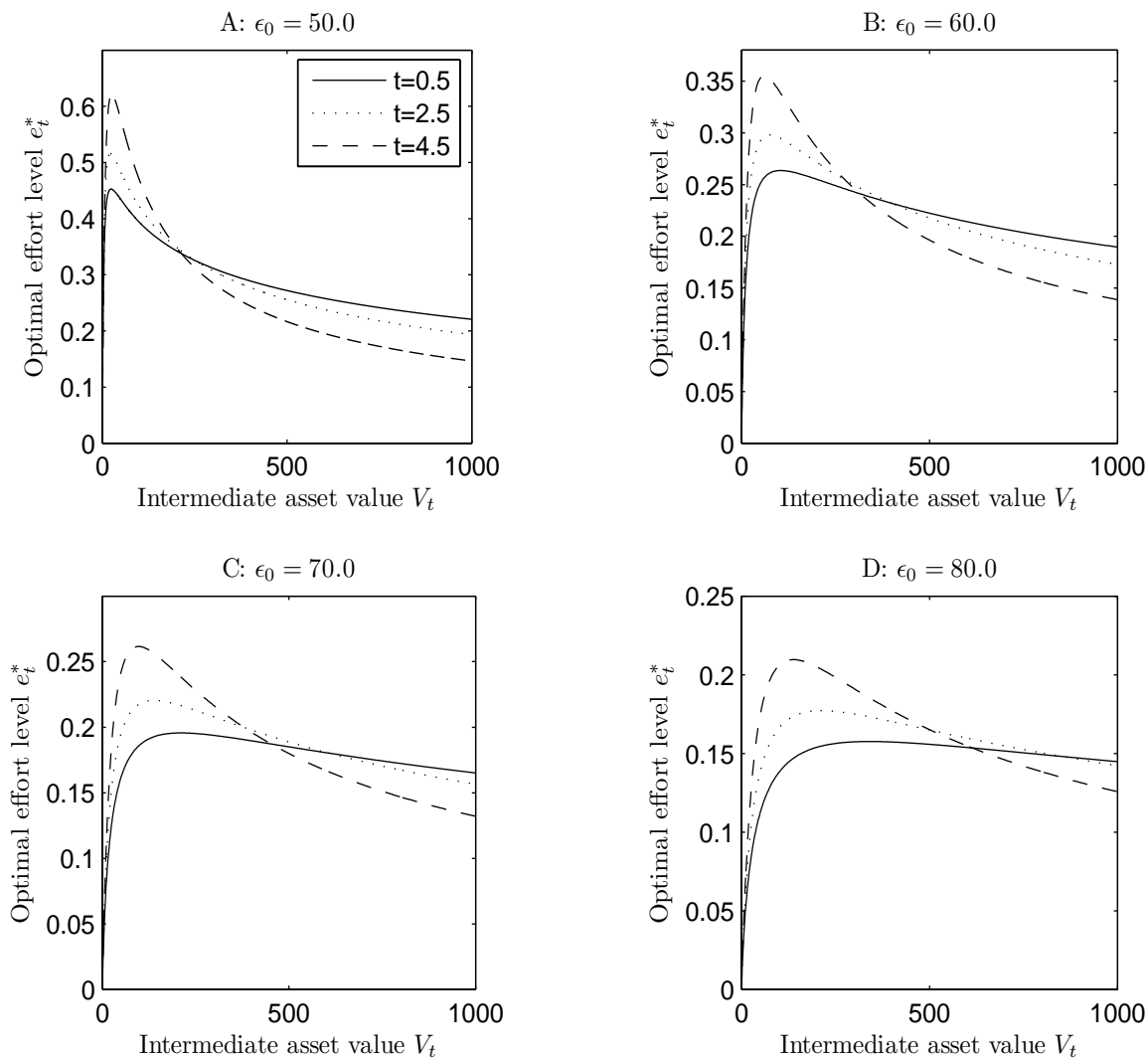


Figure 2: Effects of reservation wage  $\epsilon_0$  on optimal effort under utility  $U(W) = W^{1-\gamma}/(1-\gamma)$  with  $\gamma = 2$ . Panels A, B, C, and D plot the optimal effort levels at  $t = 0.5, 2.5, 4.5$  as a function of the prevailing asset value, respectively. The parameter values are:  $T = 5, V_0 = 100, \mu = 0, \sigma = 0.8, \kappa = 10^{-2}/\sigma^2$ .

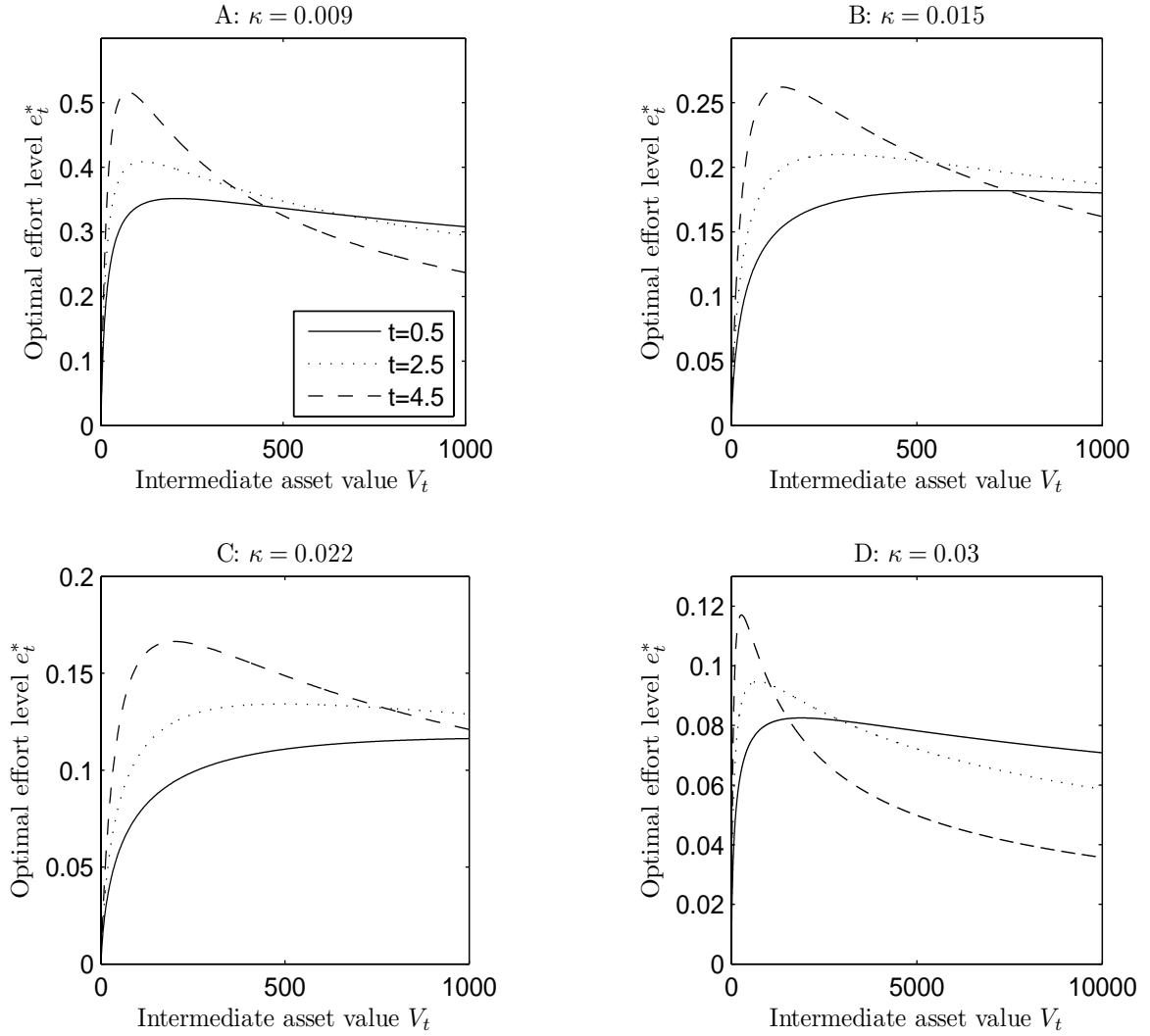


Figure 3: Effects of cost of effort coefficient  $\kappa$  on optimal effort under utility  $U(W) = W^{1-\gamma}/(1-\gamma)$  with  $\gamma = 2$ . Panels A, B, C, and D plot the optimal effort levels at  $t = 0.5, 2.5, 4.5$  as a function of the prevailing asset value, respectively. The parameter values are:  $T = 5, V_0 = 100, \mu = 0, \sigma = 1.0, \varepsilon_0 = 60$ .

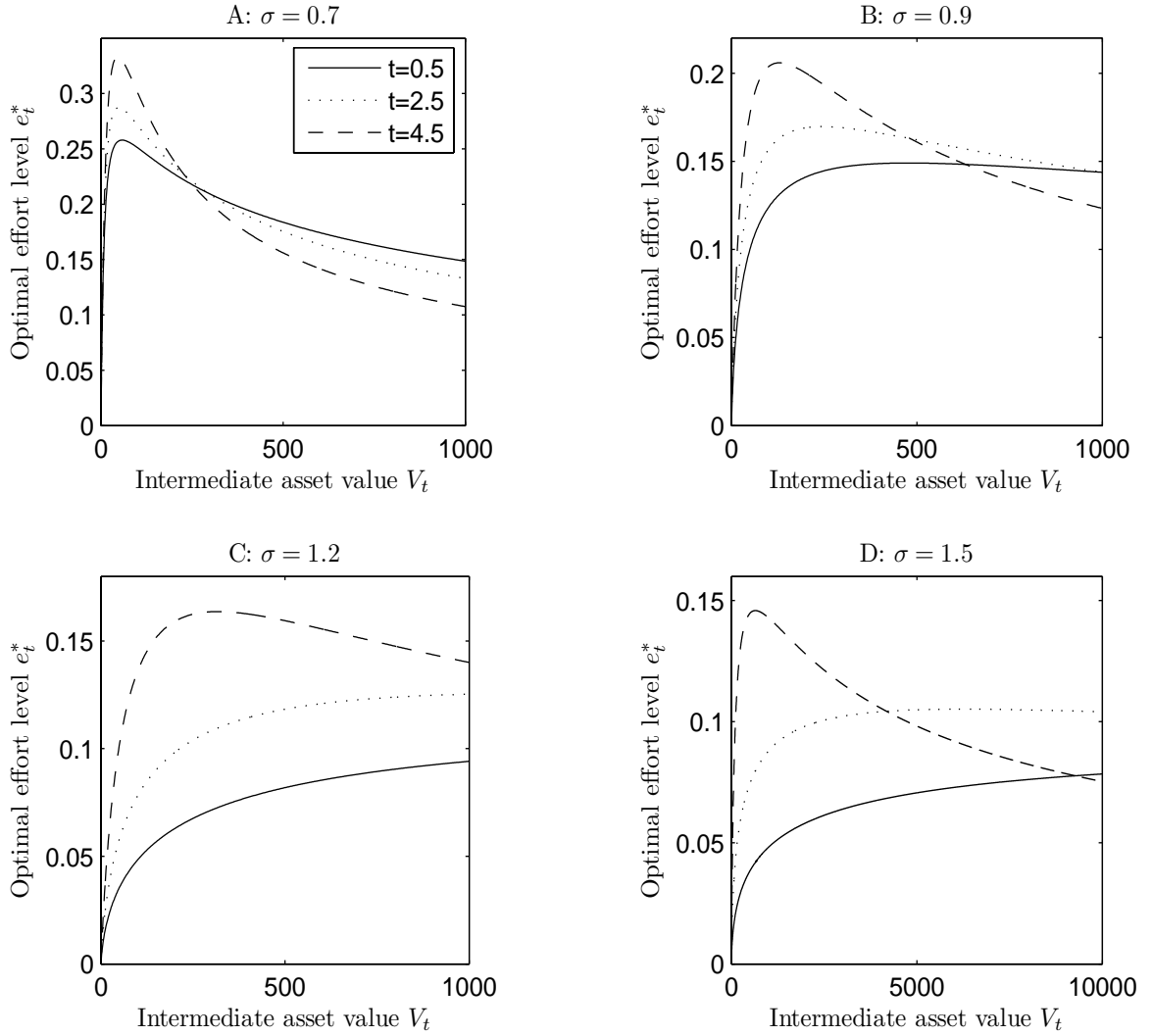


Figure 4: Effects of output volatility  $\sigma$  on optimal effort under utility  $U(W) = W^{1-\gamma}/(1-\gamma)$  with  $\gamma = 2$ . Panels A, B, C, and D plot the optimal effort levels at  $t = 0.5, 2.5, 4.5$  as a function of the prevailing asset value, respectively. The parameter values are:  $T = 5, V_0 = 100, \mu = 0, \kappa = 0.02, \varepsilon_0 = 60$ .

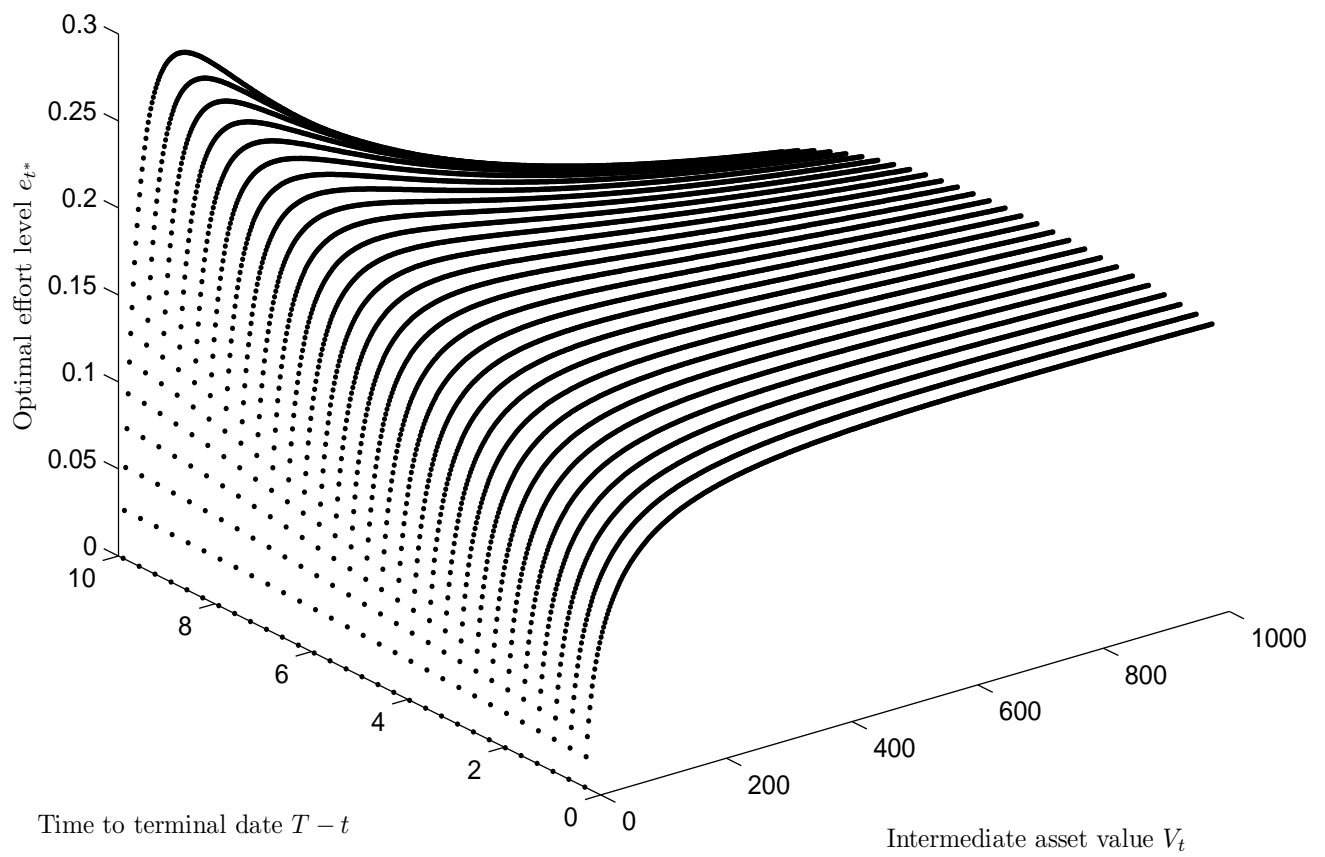


Figure 5: Optimal effort plotted against time to terminal date  $T-t$  and intermediate asset value  $V_t$  under utility  $U(W) = W^{1-\gamma}/(1-\gamma)$  with  $\gamma = 2$ . The parameter values are:  $T = 5$ ,  $V_0 = 100$ ,  $\mu = 0$ ,  $\sigma = 0.8$ ,  $\kappa = 10^{-2}/\sigma^2$ ,  $\varepsilon_0 = 60$ .

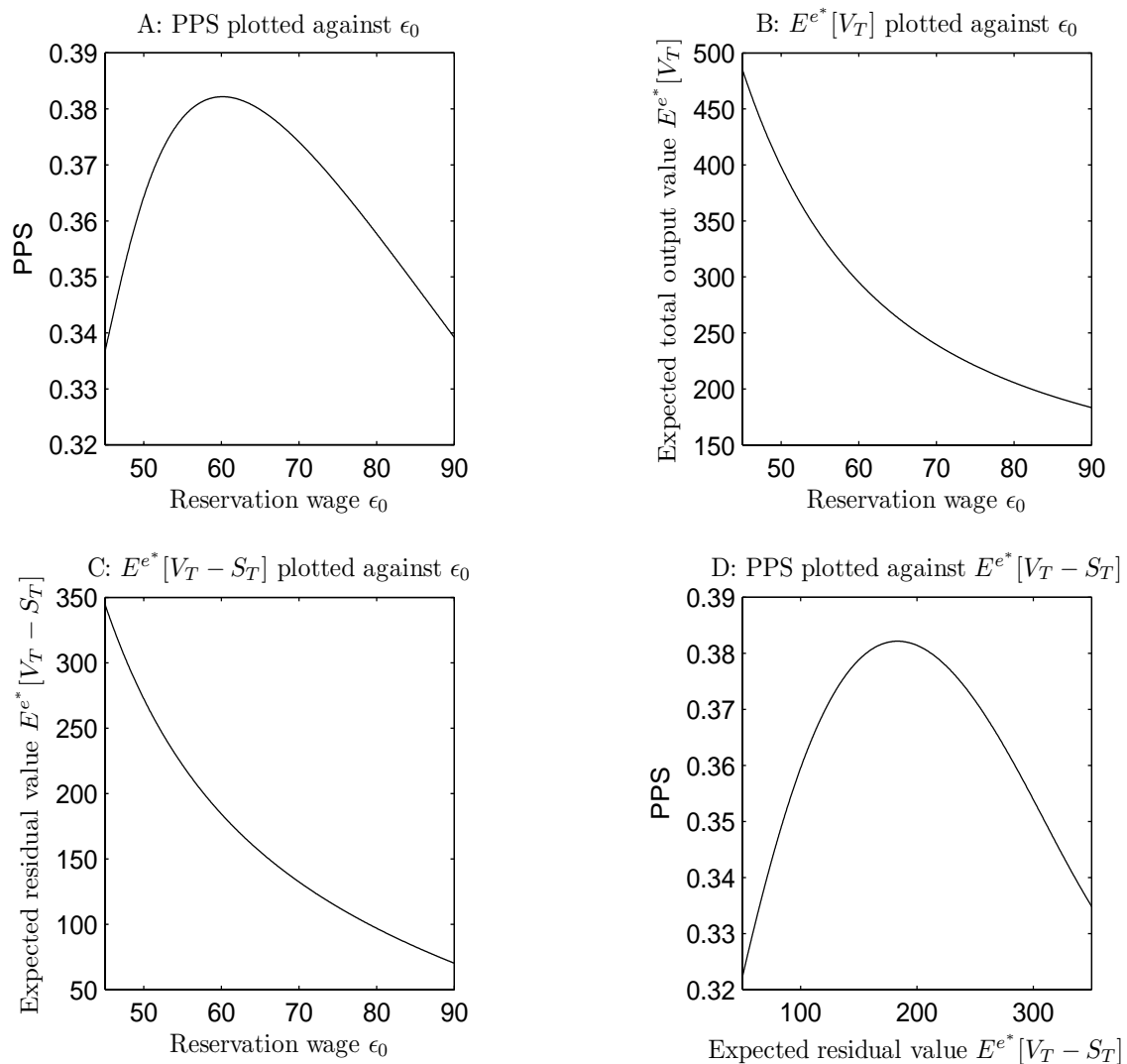


Figure 6: Effects of reservation wage  $\epsilon_0$  under utility  $U(W) = W^{1-\gamma}/(1-\gamma)$  with  $\gamma = 2$  for PPS and performance. Panels A, B, and C plot the PPS, expected terminal output, and expected terminal residual value as a function of reservation wage, respectively. Panel D plots the PPS as a function of expected terminal residual value. The parameter values are:  $T = 5$ ,  $V_0 = 100$ ,  $\mu = 0$ ,  $\sigma = 0.8$ ,  $\kappa = 10^{-2}/\sigma^2$ ,  $W_0 = 0$ .

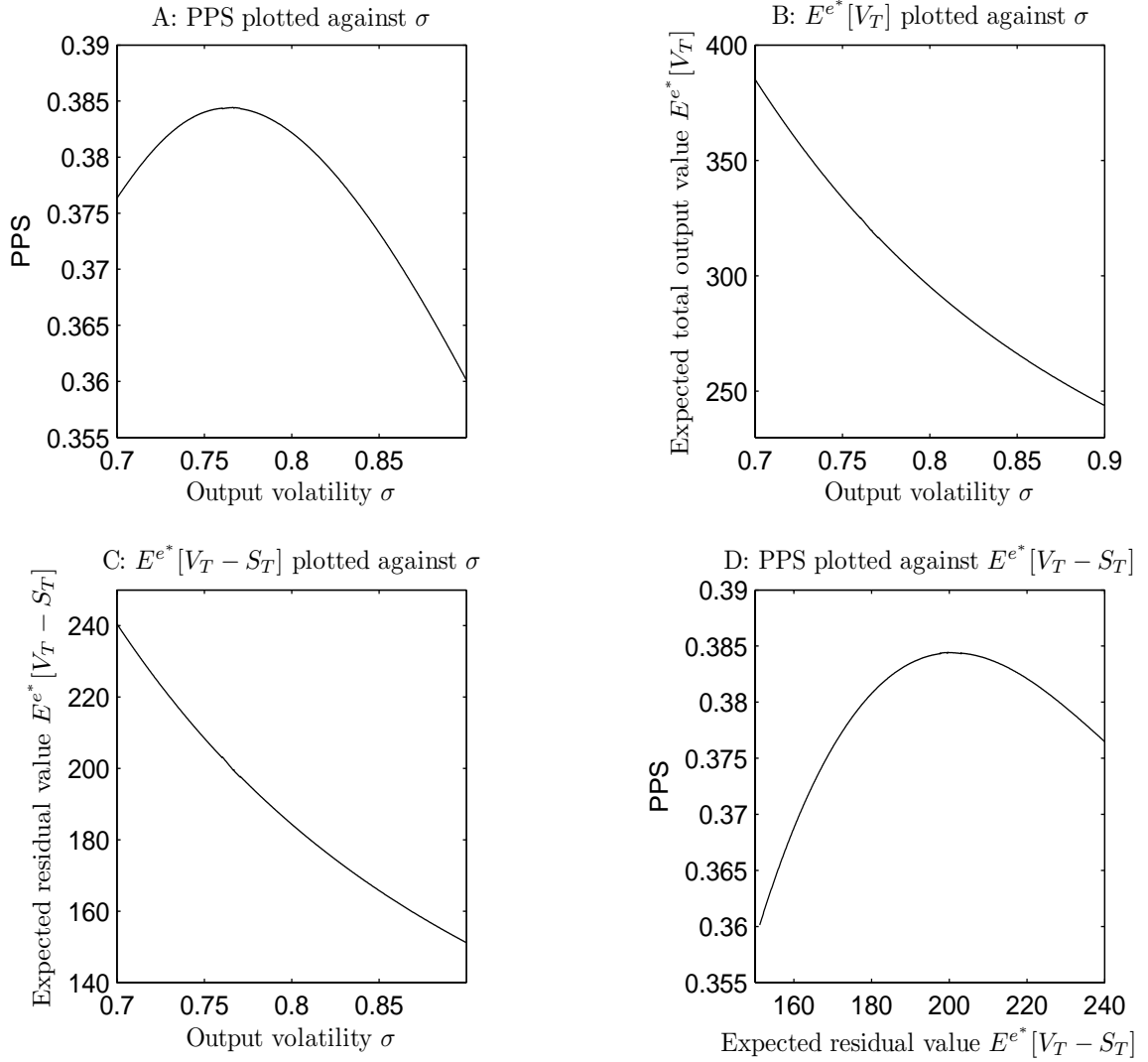


Figure 7: Effects of output volatility  $\sigma$  under utility  $U(W) = W^{1-\gamma}/(1-\gamma)$  with  $\gamma = 2$  for PPS and performance. Panels A, B, and C plot the PPS, expected terminal output, and expected terminal residual value as a function of volatility, respectively. Panel D plots the PPS as a function of expected terminal residual value. The parameter values are:  $T = 5$ ,  $V_0 = 100$ ,  $\mu = 0$ ,  $\kappa = 10^{-2}/0.8^2$ ,  $W_0 = 0$ ,  $\varepsilon_0 = 60$ .

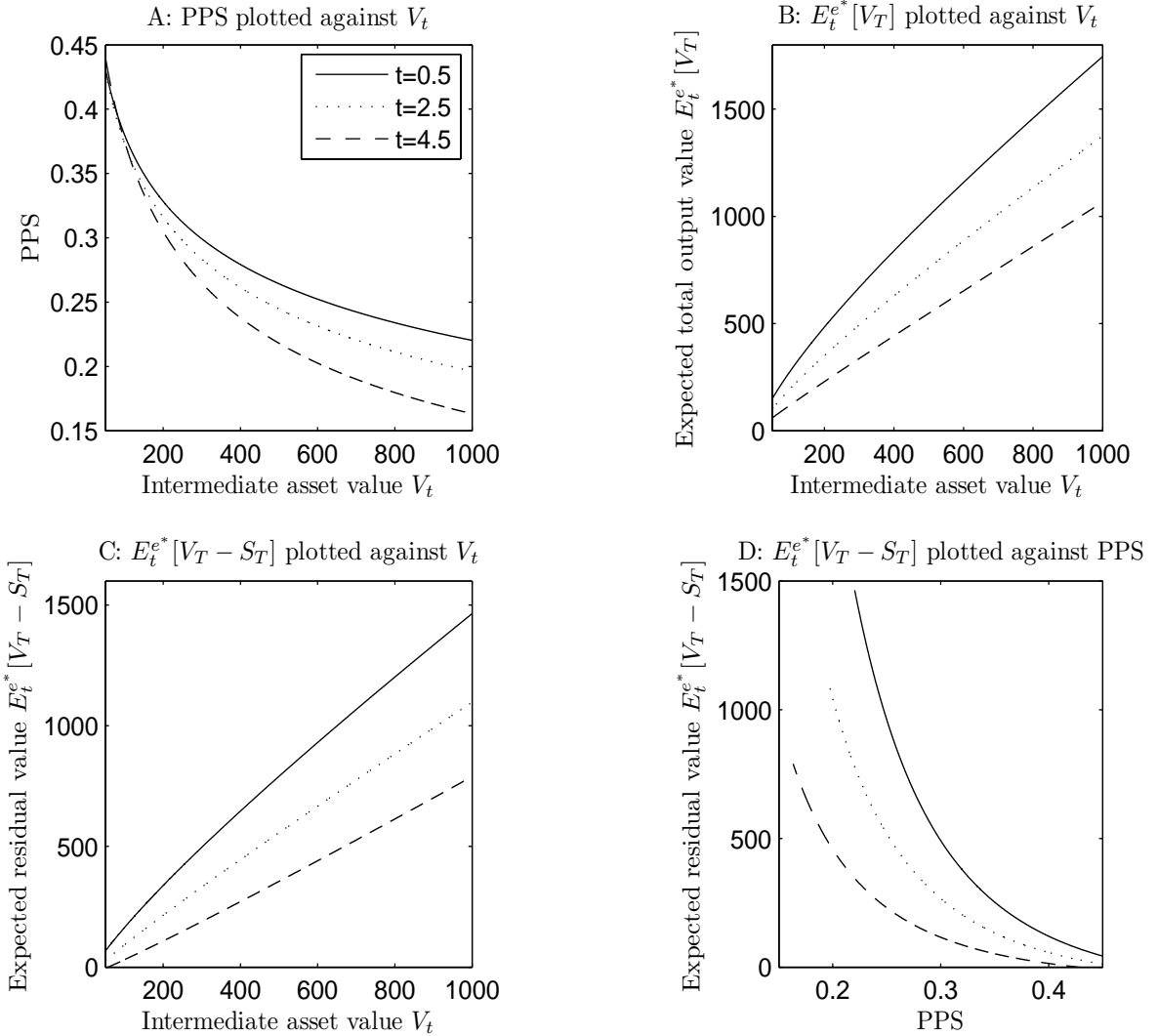


Figure 8: Effects of intermediate asset value  $V_t$  under utility  $U(W) = W^{1-\gamma}/(1-\gamma)$  with  $\gamma = 2$  on PPS and performance. Panels A, B, and C plot the PPS, expected terminal output, and expected terminal residual value as a function of  $V_t$ , respectively. Panel D plots the expected terminal residual value as a function of PPS. The parameter values are:  $T = 5$ ,  $\sigma = 0.8$ ,  $\mu = 0$ ,  $\kappa = 10^{-2}/\sigma^2$ ,  $W_0 = 0$ ,  $\varepsilon_0 = 60$ .  $t$  is the calendar time, i.e.  $T - t$  is time to terminal date.

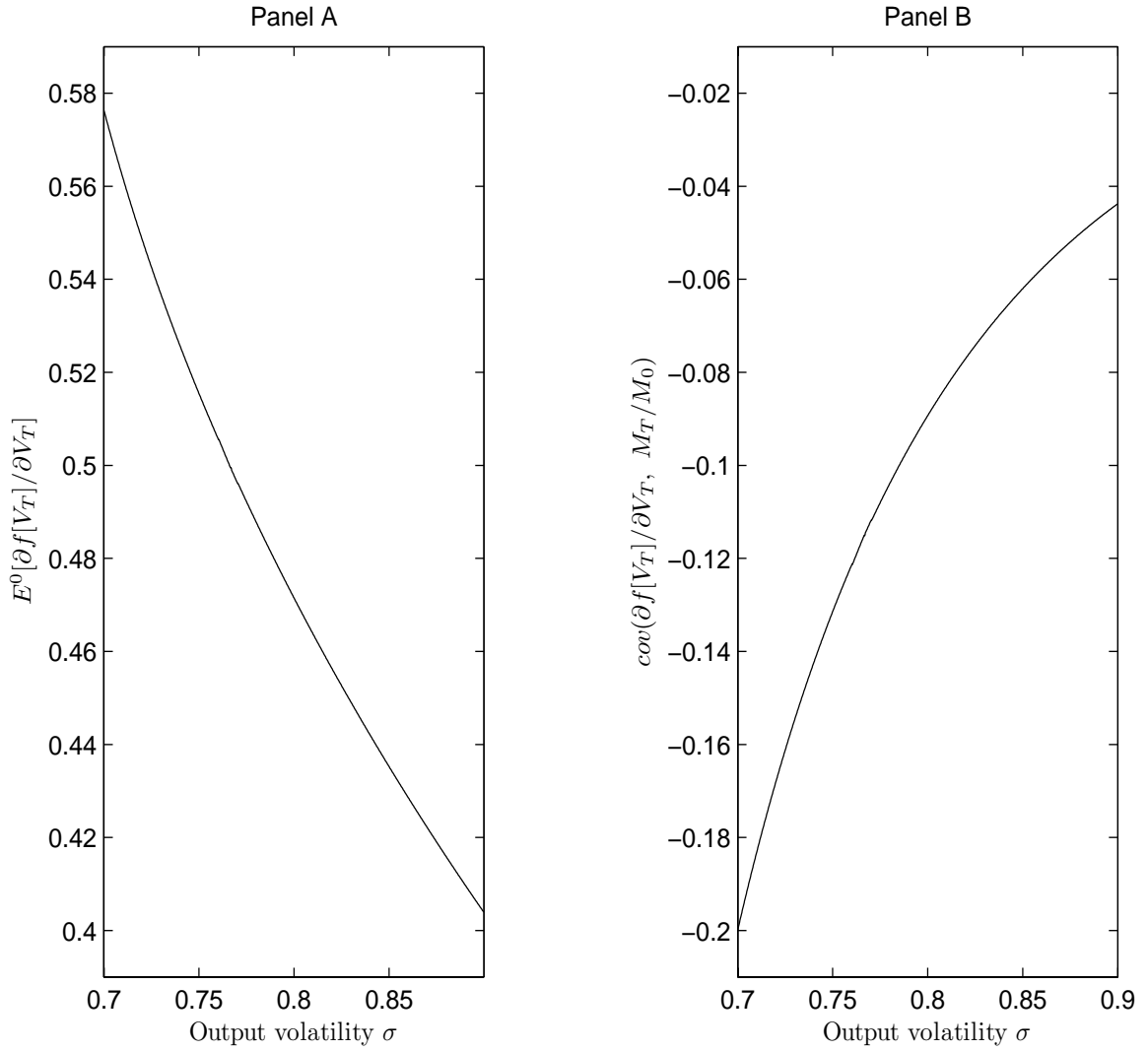


Figure 9: Effects of output volatility  $\sigma$  under utility  $U(W) = W^{1-\gamma}/(1-\gamma)$  with  $\gamma = 2$  for PPS. Panel A plots the PPS under the probability measure without effort. Panel B plots the difference between the PPS under the probability measure with and without effort. The parameter values are:  $T = 5$ ,  $\sigma = 0.8$ ,  $\mu = 0$ ,  $\kappa = 10^{-2}/\sigma^2$ ,  $W_0 = 0$ ,  $\varepsilon_0 = 60$ .