Optimal Contracts in Portfolio Delegation: The Case of Complete Markets*

Tao Li† and Yuqing Zhou‡
The Chinese University of Hong Kong

January 31, 2006

Abstract

The optimal contracts are characterized when the underlying state variable is not contractible and the shareholders must rely on the final wealth of the portfolio to design compensation schemes for the mutual-fund managers. It is shown herein that finding the optimal contracts can be converted into solving second-order nonlinear ordinary differential equations. In general, an optimal contract is an increasing, nonlinear function of the final wealth, the shape of which depends on the risk aversions of the principal and the agent, the state price density function, the principal's initial wealth and the agent's reservation utility level. The conditions under which option-like pays are optimal are also presented. Various numerical examples are presented to show the features of the optimal contracts. In addition, the optimal contracts are compared with Pareto optimal contracts. We show that, in general, there is an efficiency loss for the optimal contracts unless the utility functions of both the principal and the agent exhibit linear risk tolerance with identical cautiousness.

*We thank Sudipto Bhattacharya, Bhagwan Chowdhry, Heber Farnsworth, Pete Kyle, James Mirrlees, Xuhu Wan, and seminar participants at the annual finance conference of India School of Business, the Chinese University of Hong Kong, City University of Hong Kong, Singapore Management University for helpful comments and suggestions. Any remaining errors are ours.

†Department of Finance, Faculty of Business Administration, The Chinese University of Hong Kong, Shatin, N.T., Hong Kong. Phone: +852 2609 7757, Fax: +852 2603 6586, Email: litao@cuhk.edu.hk.

‡Department of Finance, Faculty of Business Administration, The Chinese University of Hong Kong, Shatin, N.T., Hong Kong. Phone: +852 2609 7759, Fax: +852 2603 6586, Email: zhou@baf.msmail.cuhk.edu.hk.

Optimal Contracts in Portfolio Delegation: The Case of Complete Markets

January 31, 2006

Abstract

The optimal contracts are characterized when the underlying state variable is not contractible and the shareholders must rely on the final wealth of the portfolio to design compensation schemes for the mutual-fund managers. It is shown herein that finding the optimal contracts can be converted into solving second-order nonlinear ordinary differential equations. In general, an optimal contract is an increasing, nonlinear function of the final wealth, the shape of which depends on the risk aversions of the principal and the agent, the state price density function, the principal's initial wealth and the agent's reservation utility level. The conditions under which option-like pays are optimal are also presented. Various numerical examples are presented to show the features of the optimal contracts. In addition, the optimal contracts are compared with Pareto optimal contracts. We show that, in general, there is an efficiency loss for the optimal contracts unless the utility functions of both the principal and the agent exhibit linear risk tolerance with identical cautiousness.

1 Introduction

The recent literature documents that the number of mutual funds in the US has exceeded the number of stocks. Indeed, more than half of the US equity market is currently controlled by mutual funds, pension funds and other major financial institutions.[4] Given the fact that fund managers manage other people's money and that it is costly for investors to monitor the fund manager's trading strategies, agency costs naturally arise. Therefore, it is important to study the effect of the pay schedule on the fund manager's trading or risk-taking behavior, and, as a consequence, on the securities pricing. Recent empirical studies (see Gruber (1996), Chevalier and Ellison (1997), and Sirri and Tufano (1998) among others) offer some insights about the fund-flow market mechanism as an implicit incentive: if the fund manager performs well with respect to a benchmark index in one period, then new money will flow into the fund in the next period. A fund manager normally receives a simple percentage of initial investment, and there is asymmetry in the relation between fund performance and subsequent new money flows. Hence, the resulting implicit compensation scheme to the fund manager is typically an increasing, convex function of the realized payoff of the funds under management. In such a circumstance, the fund manager may adopt an investment policy that increases the likelihood of future fund inflows, which may be at the expense of investors.[5] In addition to the implicit incentive driven by market forces, there is also an explicit incentive induced by executive compensation such as stock options and other performance-based bonuses. There are several lines of research trying to formally model the fund manager's investment policy under different compensation schemes (See Grinblatt and Titman (1996), Carpenter (2000), Basak, Pavlova, and Shapiro (2003) and the references therein). Altogether, under either the implicit or the explicit pay schedule, the fund manager will pursue his/her own interest that is not necessarily in line with that of the investors, the current literature confirms this fact both empirically and theoretically.

However, much of research along this line has been conducted under the assumption that the pay schedule is given exogenously. Within a class of linear or quadratic contracts, Stark (1987), Stoughton (1993), Admati and Pfleiderer (1997), and Das and Sundaram (1998) explore the implications of symmetric and asymmetric incentive pay schedules. Another line of research mainly deals with the general equilibrium impacts of certain exogenously given contracts on asset

[4]At the end of 2001, there were 7177 listed stocks in the US (based on CRSP), and 8307 mutual funds (2002 Mutual Fund Fact Book).

[5]See also Elton, Gruber, and Blake (2003) for further empirical evidence.

prices: see, for example, Brennan (1993), Allen and Santomero (1997), Allen (2001), and Cuoco and Kaniel (2001) among others.

Given this, a question that arises naturally is what is the optimal pay schedule. A couple of papers have addressed such issues under very restrictive conditions. For instance, Ou-Yang (2003) studies an optimal linear pay schedule in a dynamic delegated portfolio management problem when both the investors and the fund manager have an exponential utility function. Cadenillas, Cvitanic, and Zapatero (2005) deals with a more general dynamic framework but focuses on the first-best contracts. Dybvig, Farnsworth, and Carpenter (2004) incorporates information production into the analysis and also obtains a class of linear contracts when the investors and fund manager have a log utility function.[6] However, the solution to the optimal pay schedule in a second-best world under general utility conditions remains open.

This paper addresses the optimal contracting problems in the delegated portfolio management under general preferences. To highlight the issue of risk-sharing between the investors and the fund manager in the presence of uncertainty, we analyze a type of 'pure' agency model in the spirit of Ross (1973). In other words, we believe that the most fundamental delegation problem in portfolio management is agency conflict that arises from divergence in preferences when the fund manager's portfolio choice is either unobservable (thus cannot be contracted upon) or too costly to be contracted upon even if it can be observed. The issues of moral hazards where the fund manager may shirk and other information asymmetries are thus abstracted away.[7] Under such circumstances, the optimal pay schedule will be jointly determined by the risk attitude of both the investors and the fund manager along their utility functions and other model parameters.

To be more specific, the structure of our model is similar to that of the Ross's agency model except for the fact that the agent's action space in our model is much larger than that of the Ross's agency model. In the Ross's model, the agent's action space is a set in a finite dimensional space, while in our model it is a set consisting of all random payoffs over the state of nature that satisfy a budget constraint,

[6] For more information, see Dybvig, Farnsworth, and Carpenter (2004) and the references therein.

[7] In addition to the papers mentioned earlier, in the delegated portfolio management literature there are a few papers incorporate adverse selection and moral hazard into the analysis. See, for example, Bhattacharya and Pfleiderer (1985), Dybvig and Ross (1985), Kihlstrom (1988), Kihlstrom and Matthews (1990), Heinkel and Stoughton (1994), Garcia (2001), Gómez and Sharma (2001), Gervais, Lynch, and Musto (2002), Palomino and Prat (2003), and Sung (2005) among others.

which is typically infinitely dimensional.[8] This difference in the size of the agent's action spaces at first glance seems purely technical and thus insignificant. As a consequence, our model seems not deserving a more special analysis than those already done by Ross (1973) and others. However, it is exactly the difference in the size of the agent's action space that makes our approach unique and our results significantly different.

In Ross's model, the first-order approach is applied to characterize the optimal pay schedule without adequately addressing the technical conditions it must meet. When the agent's action space is small, the principal has more leeway to design a pay schedule to work in his best interest without worrying too much about the possibility of the agent's taking advantage of the compensation scheme. As a consequence, the shape of the optimal contracts could be complicated, and may not even be monotonic. In addition, with the low dimensionality of the agent's action space, the set of pay schedules that implement any particular action is relatively large, and consequently the suboptimal pay schedule that implements a particular action contains most of the characteristics of the optimal pay schedule and the optimal action becomes less relevant. Given this, the main features of the optimal contracts can be characterized by the first-order condition with the agent's action being held constant, a standard technique used in principal-agent analysis.[9]

When the agent's action space is large, the principal will have less flexibility in designing a pay schedule to implement any particular action. Intuitively, a simple contract could be an optimal one when the agent has too many opportunities to explore. In this regard, Mirrlees and Zhou (2005a and 2005b) have developed a principal-agent model with moral hazards to capture the situation in which the agent faces a multi-dimensional action space, and have shown that the optimal contracts indeed take very simple forms under a certain class of cost functions.[10] The optimal contracts are monotonic, concave or convex, depending on the underlying model parameters.[11]

In the agency model analyzed in this paper, the agent's action space is very

[8] Focusing on the infinite dimensional action space allows us to apply our results to a model in a continuous-time setting.

[9] See Ross (1973), Mirrlees (1974, 1976 and 1999), Holmstrom (1978,1979), Grossman and Hart (1983), Rogerson (1985) and Sung (1995) among others.

[10] Indeed, optimal contracts can be linear when the players' utilities are exponential and the agent's action space is rich in a dynamic setting: see Holmstrom and Milgrom (1987) and Schättler and Sung (1993).

[11] It should be noted that, other things being equal, the principal is worse-off when faced with a large agent's action space than with a small agent's action space. Thus, from the principal's perspective, a simple contract does not mean a better one.

large in the sense that the portfolio choice space that is subject to budget constraints is infinite dimensional. Under these circumstances, the mathematical technique developed earlier in handling the case in which the agent's action space is low dimensional cannot be applied in a straightforward manner. Hence, a new approach is developed to deal with the case of the high-dimensional portfolio choice space in characterizing the optimal contracts. Via the basic techniques in the calculus of variations and the concavification technique, we can transfer the incentive constraint into a normal constraint, and restrict our analysis to a class of pay schedules that are an increasing function of final wealth. We show that, under certain smoothness conditions, finding the (local) optimal pay schedules can be converted into solving second-order nonlinear ordinary differential equations (ODEs) together with a system of algebraic equations. These second-order nonlinear ODEs allow us to characterize some features of the optimal contracts without solving them in the first place. For instance, we can show that, in general, an option-like pay schedule cannot be optimal. By examining the ODEs, we can also provide a necessary condition that a pair of utility functions must satisfy for an optimal contract to be linear. The condition is very restrictive indeed.

Our analysis reinforces the general idea in the principal-agent literature that enlarging the agent's action space could make the optimal contracts simple. In the process, we also examine the efficiency of the optimal contracts with reference to Pareto-efficient contracts, and show that in general there is an efficiency loss in the presence of agency. Indeed, Pareto efficiency contracts in our agency model can only be obtained under very limited conditions. The optimal pay schedule is Pareto-efficient if and only if the utility functions of both the principal and the agent exhibit linear risk tolerance with identical cautiousness. Furthermore, Pareto-efficient pay schedules must be linear and satisfy the similarity condition posed by Wilson (1968), and subsequently studied by Ross (1973), Ross (1974), and Dybvig and Spatt (1986), when the principal's gross pay is also a linear function of final wealth.

Various numerical examples are presented to show the features of the optimal pay schedules. Normally two offsetting forces work together in determining the shape of the optimal pay schedules. One is the risk-sharing effect, and the other is the incentive effect. The first effect demands that the marginal rate of substitution is close to a constant in order to obtain the optimal risk-sharing, and the second effect demands that the two players' utilities are similar in order to provide the best incentive. These two effects are normally mingled together and typically move in opposite directions unless the two players' utilities exhibit linear risk tolerance with identical cautiousness. As a result, the optimal pay schedules appears flatter and less sensitive compared to Pareto efficient ones in many cases. Some simple numerical results are depicted in the paper to illustrate the structures of the optimal pay schedules.

The rest of the paper is organized as follows. Section 2 sets up the basic model framework. Section 3 characterizes the optimal contracts. Section 4 discusses the efficiency of the optimal contracts. Section 5 offers some numerical results. Section 6 concludes the paper. Appendix A discusses the a variation technique employed in the paper and all proofs are provided in appendix B.

2 The Model

Consider a setting in which an investor or a fund company (the principal) wishes to hire a fund manager (the agent) to manage his/her portfolio. The portfolio return is realized over a continuum of states in a single period. The uncertainty is summarized by the probability space $(\Omega, \mathcal{F}, P)$, where Ω is the space of states that are endowed with a σ-field $\mathcal{F}$ and a probability measure P and $\omega \in \Omega$ is a state. We assume that there exist a rich set of financial securities or infinitely many trading opportunities (continuous-time models) such that the financial markets under consideration are complete. If markets are complete, then there exists a unique state price function $p(\omega)$ per unit probability over Ω. The portfolio returns are affected by the agent's actions, or individual security selections in the portfolio. Thus, the incentive scheme designed by the principal matters in order to motivate the agent to act in the best interest of the principal. If the agent's action can be observed, or if the principal can distinguish costlessly the payoff characteristics of the universe of securities in the financial markets, then the contracting problem between the principal and the agent is relatively straightforward; the contract simply specifies the exact portfolio of securities to be selected by the agent and the compensation that the principal promises to provide in return should the order be followed exactly. However, if it is too costly for the principal to distinguish the payoff characteristics of the universe of securities, then the contract can no longer effectively specify the agent's security selection. Under this circumstance, the principal must design a compensation scheme in a way that indirectly gives the agent the incentive to select the correct set of securities. As a first step, we will begin with a case in which the only way for the principal to get the agent to select a correct portfolio is to relate his/her pay to the realization of the portfolio return, which is random.

To be more specific, let $w \geq 0$ be the final wealth of the selected portfolio,[12] and $w(\omega)$ be the realization of the final wealth over Ω, which is observable. A compensation scheme specifies the agent's wage as a function of observed final wealth $y(w)$. Let the principal's utility function be $u(\cdot)$ and the agent's utility be $v(\cdot)$, where $u(\cdot)$ and $v(\cdot)$ are independent of states, smooth, increasing and concave over an interval that contains $[0, \infty)$.[13] If the final wealth is w, then the net benefit for the principal is assumed to be

$$\alpha(w) - y(w), \quad \text{where } 0 \leq \alpha(w) \leq w.$$

When $\alpha(w) = w$, our model is a typical principal-agent problem, and it can be interpreted as a large investor hiring a money manager to manage his portfolio.[14] In general, $\alpha(w)$ is used to model a situation where the principal may be an institution (e.g, a mutual fund company) or may act as an agent of a large number of shareholders himself, thus only receive a portion of the realization of the final wealth either explicitly as a management fee or implicitly through the flow of new funds. We assume that $\alpha(w)$ is piecewise smooth and increasing.

The principal will delegate initial wealth w_0 for the agent to manage, and design a fee schedule $y(w)$ to induce the agent to act in his/her best interest. Given a compensation scheme $y(w)$, the agent will select a portfolio such that his own expected utility is maximized under the budget constraint. Therefore, there is an explicit conflict of interests between the principal and the agent, and it is interesting to see how the principal and the agent share the risks and what the optimal contracts are. To formalize these ideas, let the agent's action space A be defined by

$$A = \{w(\omega) \geq 0 | \int_{\Omega} p(\omega) w(\omega) \, P(d\omega) \leq w_0\}, \tag{1}$$

where the last term in equation (1) is the budget constraint. In other words, the action space A consists of all random variables over Ω that satisfy the budget

[12] The restriction $w \geq 0$ is not necessary. All of our analysis in the paper can be carried over to the case in which $w \in (w_l, +\infty)$ as long as $w_l \leq w_0/E[p]$. One of the motivations for nonnegative wealth is the limited liability of holding equities. We do not address the limit liability issue that is associated with contracts, i.e., $y(w) \geq 0$, in this paper, but this will not be an issue for the case in which an agent has a power utility and is endowed with zero initial wealth. The latter is assumed in this paper.

[13] The differentiability is not necessary, but is used for convenience.

[14] Much work has been conducted with such a specification. See Dybvig, Farnsworth, and Carpenter (2004) and the references therein.

constraint.[15] In contrast to those one-dimensional (or low-dimensional) action spaces that are studied in the agency models in existing literature, ours is large in the sense that it is infinitely dimensional. Let the agent's reservation utility be v_0. Formally, the model is as follows:

$$\max_{y(w),w(\omega)\in A} \int_\Omega u(\alpha(w(\omega)) - y(w(\omega)))\, P(d\omega) \tag{2}$$

subject to

$$\int_\Omega v(y(w(\omega)))\, P(d\omega) \geq v_0 \tag{3}$$

and

$$w(\omega) \in \arg \max_{\hat{w}(\omega)\in A} \int_\Omega v(y(\hat{w}(\omega)))\, P(d\omega), \tag{4}$$

where equations (3) and (4) are the standard participation constraint and incentive constraint respectively.[16]

Solving the maximization problem as defined by (2)-(4) is not trivial. The classical approach to characterize an optimal solution is to replace the incentive constraint (4) with the first-order condition, and then to apply the calculus of variations. To validate this approach one has to assume that the endogenous variable $v(y(w))$ is differentiable. However, as an example by Mirrlees (1999) shows, the solution may be nondifferentiable and as a result the first-order approach can no longer be applied.

In what follows we develop a better approach with Mirrlees's flavor and reformulate the model in a more tractable way. Given our model setup in which both the principal's utility and the agent's utility are independent of state ω, as shown below, the states only need to be distinguished by the state prices when the financial markets are complete. Let $f_p(p)$ ($F_p(p)$) and $f_w(w)$ ($F_w(w)$) denote the (cumulative) distribution functions of the random variables $p(\omega)$ and $w(\omega)$ respectively. Note that $f_p(p)$ is exogenously given, and the support of which is the half line R^+ or its subset, while $f_w(w)$ is the agent's choice variable. For sim-

[15] We will implicitly assume that $w(\omega) \in L^{\mathbf{p}}_+(\Omega, \mathcal{F}, \mathcal{P})$ where $1 \leq \mathbf{p} < \infty$, and $p(\omega) \in L^{\mathbf{q}}_+(\Omega, \mathcal{F}, \mathcal{P})$ where $\frac{1}{\mathbf{p}} + \frac{1}{\mathbf{q}} = 1$, to avoid potential arbitrage opportunities created by our model setup.

[16] At this moment we assume that $y(w)$ is any Borel-measurable function such that both $u(\alpha(w) - y(w))$ and $v(y(w))$ are well-defined over their domain. Later on we can show that we only need to focus on the class of piecewise smooth functions

plicity, we further assume that $f_p(p)$ is continuous and first-order differentiable.[17] Let

$$A_w = \{f_w(w) \geq 0 | \int f_w(w)dw = 1, E[pw] \leq w_0\},$$

where [18]

$$E[pw] = \int p(w)wf_w(w)dw = \int F_p^{-1}(1 - F_w(w))wf_w(w)\, dw, \tag{5}$$

Note that, in equation (5), $p(w) = F_p^{-1}(1 - F_w(w))$ is a nonincreasing and right-continuous function of w. Since not all distribution functions of $w(\omega) \in A$ can be written as a form in A_w. Hence A_w is a subset of A in a distribution sense.

Now we are ready to reformulate the model into a distribution form. Before doing this, we need a lemma to insure that an optimal solution is not lost in the transformation.

Lemma 1 *Suppose that $G(w)$ is upper-semicontinuous function on R^+ such that $\overline{\lim}_{w\to\infty} \frac{G(w)}{w} = 0$ and $\max_{w\in A} \int_\Omega G(w)\, P(d\omega)$ exists. Then*

$$\max_{w(\omega)\in A\cap A'} \int_\Omega G(w)\, P(d\omega) = \max_{f_w(w)\in A_w\cap A'_w} \int G(w) f_w(w)\, dw \tag{6}$$

Furthermore, we have, at optimum, $p(w) = F_p^{-1}(1 - F_w(w))$.

Proof: See appendix B. □

Lemma 1 shows that both the agent's maximization problem and participation constraint can be reformulated in terms of distribution of wealth f_w, hence, under

[17] These assumptions are not crucial to solving the contracting problems, but rather for convenience. In addition, for the incentive constraint (4) to have a solution, the state price density functions cannot be arbitrarily specified; a sufficient condition for existence, given by Cox and Huang (1991), is that the inverse of $p(\omega)$ has a certain finite moment. Since the main purpose of the paper is to characterize optimal contracts when they are known to exist, we will not discuss existence issues in details in the paper. Relevant restrictions are given explicitly when we work with specific examples.

[18] For convenience, we abbreviate $\int_0^\infty dw$ to $\int dw$ in the remaining part of the paper. We also interpret the probability density f_w in a slightly generalized sense. That is, f_w is allowed to be singular at some points as long as the number of these points is finite. The density at each of these points can be represented by a scaled Dirac delta function, $a\delta(w - t)$. See also the footnote on page 10.

a very weak condition, the principal's problem as represented by equations (2)-(4) can be reformulated as follows:

$$\max_{y(w), f_w(w)\in A_w} \int u(\alpha(w) - y(w)) f_w(w) dw, \tag{7}$$

subject to

$$\int v(y(w)) f_w(w) dw \geq v_0 \tag{8}$$

and

$$f_w \in \text{arg} \max_{\hat{f}_w(w)\in A_w} \int v(y(w)) \hat{f}_w(w) dw. \tag{9}$$

Before we get into the full principal-agent problem, let us study the agent's problem first. Clearly, given a pay schedule $y(w)$ that is upper semi-continuous, the agent's problem is the classical portfolio choice problem in complete markets, except the fact that $v(y(w))$ is not necessarily increasing and concave, or not even continuous. As a result, the traditional first-order approach, which involves the derivative of $v(y(w))$, can not be applied in a straightforward manner. However, if we work with the agent's problem in a distribution form, the smoothness condition on $v(y(w))$ can be avoided, and the resulting Lagrangian for the agent's problem is

$$\begin{aligned} \mathcal{V}(f_w, \lambda, \lambda_f) &= \int v(y(w)) f_w(w)\, dt \\ &\quad - \lambda \left[\int f_w(w) [F_p^{-1}(1 - F_w(w))] w\, dw - w_0 \right] + \int \lambda_f(w) f_w(w)\, dw, \end{aligned}$$

where λ is nonnegative constant and λ_f is nonnegative function of wealth, which is equal to zero if $f_w > 0$.

A direct calculation shows that the first-order condition (see also the proof of Lemma 2) for the agent's problem is that

$$v(y(w)) + \lambda_f(w) - \lambda \int_{\ell}^{w} F_p^{-1}(1 - F_w(t))\, dt \tag{10}$$

is invariant in w. Note that the regularity requirement on the pay schedule y is quite generous: we only require y is Borel measurable. Denote by $v \circ y$ the composite function $v(y(\cdot))$. Then we have

$$f_w(w) = -\frac{f_p(F_p^{-1}(1 - F_w(w)))}{\lambda} \left[(v \circ y)''(w) + \lambda_f''(w) \right].$$

Clearly, if $v \circ y$ is an increasing, concave function of w, then $\lambda_f \equiv 0$ for all w. In other words, $f_w(w)$ is an interior solution of the agent's problem when $v \circ y$ is increasing and concave. [19] Indeed, we have a much strong result for very general pay schedules, as shown in the following lemma

Lemma 2 *Given a Borel measurable pay schedule y, then $\{E[v \circ y] | f_w(w) \in A_w\}$ reaches its maximum at $f_w(w) \geq 0$, if and only if there exist two constants $\lambda \geq 0$ and λ_ℓ such that*

$$\tilde{x}(w) \equiv \lambda \int_\ell^w F_p^{-1}(1 - F_w(t))\, dt + \lambda_\ell \tag{11}$$

is the concavification[20] of the agent's composite utility function $v \circ y$. Furthermore, $f_w(w) \equiv 0$ for $w \in \{t | v(y(t)) \neq \tilde{x}(t)\}$.

Proof: See appendix B. □

Let f_w be the solution to the agent's problem, then the set in which x is different from its concavification $\tilde{x}$ is $\{w | f_w(w) = 0\}$. This indicates that if $v \circ y_1$ and $v \circ y_2$ share the same concavification, then the agent portfolio choice problems under these two utilities have the same solution.

Lemma 3 *If y_1 is an optimal pay schedule to the original principal-agent problem, then y_2 is also an optimal solution as long as $v \circ y_2$ has the same concavification as $v \circ y_1$.*

Proof: First note that Lemma 2 shows that $v \circ y_2$ and $v \circ y_1$ has the same solution f_w to the agent's problem. As $f_w \equiv 0$ on the set $\{w | v \circ y_2 \neq v \circ y_1\} = \{w | y_2 \neq y_1\}$, the principal's utility achieves the same level under both pay schedules. □

---

[19] Note that $v \circ y$ need not to be smooth in oder to obtain an interior solution. From the necessary condition (10), it is easy to see that a concave kink of $v \circ y$ makes F_w jump at the kink if $\lambda_f = 0$ in the neighborhood of the kink. This implies f_w is singular at the kink. However, f_w can be interpreted as a scaled Dirac delta function, i.e., $f_w(w) = \bar{f}_w(w) + a\delta(w - t)$, at the kink t such that

$$F_w(t+) - F_w(t-) = \int_{t-}^{t+} f_w(w)\, dw = \int_{t-}^{t+} \bar{f}_w(w)\, dw + a \int_{t-}^{t+} \delta(w - t)\, dw = a,$$

where $\bar{f}_w(t) < \infty$ and $a \leq 1$. In this sense, $f_w(w)$ at t can be viewed as an interior solution. Whereas, $f_w(w) = 0$ at other points can be either an interior or a corner solution depending on whether or not λ_f equals 0.

[20] The concavification of a function is the smallest concave function that dominates it.

Lemma 3 shows that we can restrict the feasible pays to those that make $v \circ y$ nondecreasing and concave. Under this restriction, the solution to the agent's problem is unique and hence the first-order approach is valid. In fact, the two lemmas above also enable us to focus on pay schedules that are piecewise smooth without loss of generality.

Lemma 4 *For any pay schedule $\tilde{y}$ that can induce an optimal choice for the agent, there exists a piecewise smooth pay schedule y such that $v \circ \tilde{y}$ and $v \circ y$ share the same concavification.*

Proof: See the proof of Lemma 2. □

Given the three lemmas above, we are now in a position to further simplify the agent's problem. Let $x = v \circ y$, and therefore, $y = v^{-1}(x) = h(x)$. As $v(\cdot)$ is increasing and concave, $h(x)$ is increasing and convex. The principal-agent problem is then reformulated as follows:

$$\max_{x,\, f_w \in A_w} \int u(\alpha(w) - h(x(w))) f_w(w) dw, \tag{12}$$

subject to

$$x(w) = \lambda \int_{\ell}^{w} F_p^{-1}(1 - F_w(t))\, dt + \lambda_{\ell}, \tag{13}$$

which is both necessary and sufficient for the agent's problem, where $\lambda \geq 0$ and λ_{ℓ} are free variables, which are constrained by the budget constraint and the participation constraint

$$\int x(w) f_w(w) dw \geq v_0. \tag{14}$$

The reformulation of the original problem leads to equations (12)-(14), in which the principal selects $x(w)$ and $f_w(w)$ to maximize his utility u subject to the first-order condition constraint (13) plus the participation and budget constraints.

Note that, equation (13) transforms the agent's incentive constraint into a normal one, thus is the most critical condition in our analysis. It could be thought of the first-order condition in distribution sense. Indeed, $x(w)$ and its concavification $\tilde{x}(w)$ are identical in a distribution sense, which is exactly what matters in our principal-agent problem. If $x(w)$ is concave and nondecreasing, then $x(w) = \tilde{x}(w)$. Equation (13) alone also reveals an important feature of the optimal contract: it can be designed in a way such that the compensation is a nondecreasing function of final wealth. Of course, to obtain additional features of the optimal contract, we need to solve the principal's maximization problem. Equation (13) tells us that the principal only needs to focus on the class of nondecreasing and concave functions in the selection of $x(w)$.

3 Characteristics of the Optimal Contracts

In this section we characterize the optimal contracts of the principal-agent problem discussed in the previous section. Given our analysis of the agent's problem in the previous section, we can investigate the principal's contracting problem (12)-(14) by studying the corresponding Lagrangian

$$\begin{aligned}\mathcal{U}(f_w, \lambda, \lambda_\ell, \lambda_w, \lambda_v, \lambda_f) \;=\; & \int u\left(\alpha(w) - h(x(w))\right) f_w(w) dw \\ & - \lambda_w \left[\int f_w(w) F_p^{-1}(1 - F_w(w)) w dw - w_0\right] \\ & + \lambda_v \left[\int x(w) f_w(w) dw - v_0\right] \\ & + \int \lambda_f(w) f_w(w)\, dw, \qquad (15)\end{aligned}$$

where $x(w)$ is a function of f_w, λ and λ_ℓ as defined by equation (13), and $\lambda > 0$, $\lambda_w > 0$, $\lambda_v > 0$ and $\lambda_f(w) \geq 0$. Similar to the case of the agent's problem, $\lambda_f(w) = 0$ if $f_w(w) > 0$. Note that in addition to the budget and the individual rationality (participation) constraints the objective function $\mathcal{U}$ is also dependent on two choice variables λ and λ_ℓ.[21] These two variables can be handled separately from the function f_w.

Lemma 5 *For any given multipliers λ_w, λ_v, $\lambda_f(w)$ and a density function f_w, the objective function $\mathcal{U}$ as the function of $(\lambda,\ \lambda_\ell)$ has a unique global maximum. Furthermore, at optimum (λ, λ_ℓ), it must satisfy the following first-order conditions:*

$$\frac{\partial \mathcal{U}}{\partial \lambda} = -\frac{1}{\lambda} \int f_w(w)[x(w) - \lambda_\ell][u'h' - \lambda_v]\, dw = 0, \qquad (16)$$

$$\frac{\partial \mathcal{U}}{\partial \lambda_\ell} = -\int f_w(w)[u'h' - \lambda_v]\, dw = 0. \qquad (17)$$

[21] We will only focus on the case $\lambda > 0$, as $\lambda = 0$ will imply that $x(w) = v_0$, which is a constant and a very trivial case.

Proof. See appendix B. $\square$

Lemma 5 conveys an important message. It shows that for any agent's action $f_w(w)$, there exists a unique optimal pay schedule $y(w)$ that implements this action. It should be noted that the set of pay schedules that implement a particular action is one dimensional. The low dimensionality is due to the large size of the agent's action space. The size of the agent's action space is assumed to be one-dimensional (or low-dimensional) in the existing literature, which makes the set of pay schedules that implement a particular action very large. Consequently, most characteristics of the optimal contract can be obtained by examining the set of pay schedules that implement a particular action alone. However, the same approach cannot be applied to our portfolio choice problem because the set of pay schedules that implements a particular f_w is one-dimensional and the resulting optimal contract is not informative. Therefore, we need to find the optimal f_w to fully characterize the optimal contract $y(w)$.

We now turn to examine the features of an optimal solution around $(f_w, \lambda, \lambda_\ell)$. Suppose that $\mathcal{U}$ reaches a maximum at $(f_w, \lambda, \lambda_\ell)$ for given multipliers, λ_w, λ_v, and λ_f. Then we can study the necessary conditions for the optimality of $\mathcal{U}$ by the perturbing the density function f_w by $\epsilon\eta$, where ϵ is an arbitrary small constant and η satisfies $\int \eta(w)\,dw = 0$. For each ϵ, Lemma 5 shows that there is a unique solution $(\lambda^\epsilon, \lambda_\ell^\epsilon)$ to the first-order conditions for given $(f_w + \epsilon\eta, \lambda_w, \lambda_v)$ in Lemma 5. As ϵ goes to zero, $(\lambda^\epsilon, \lambda_\ell^\epsilon)$ converges to (λ, λ_ℓ). Therefore, the original maximizing problem $\max_{f_w,\lambda,\lambda_\ell} \mathcal{U}$ is equivalent to

$$\max_{\epsilon} \mathcal{U}^\epsilon = \max_{\epsilon} \mathcal{U}(f_w + \epsilon\eta, \lambda^\epsilon, \lambda_\ell^\epsilon, \lambda_w, \lambda_v, \lambda_f),$$

which is maximized at $\epsilon = 0$ if $(f_w, \lambda, \lambda_\ell)$ is a solution to the original problem.Thus, the necessary conditions are derived by examining the derivatives of the perturbed $\mathcal{U}^\epsilon$ with respect to ϵ.

Proposition 1 *Suppose that $(f_w, \lambda, \lambda_\ell)$ is an optimal solution. Then the first-order necessary condition is that there exists a constant C, which is independent of w, such that $(f_w, \lambda, \lambda_\ell)$ satisfy*

$$u(\alpha(w) - h(x(w))) + \lambda_f(w) + \left[\lambda_v - \frac{\lambda_w}{\lambda}\right] x(w) + \lambda \int_\ell^w \frac{1}{f_p(F_p^{-1}(1 - F_w(s)))} \left[\int_0^s [u'h' - \lambda_v] f_w(t)\, dt\right] ds = C \quad (18)$$

almost everywhere for all ℓ such that $f_w(\ell) > 0$, where

$$C = u(\alpha(\ell) - h(x(\ell)) + \lambda_v x(\ell)$$

and $\lambda_f(w) > 0$ if $f_w(w) = 0$ and $\lambda_f(w) = 0$ if $f_w(w) > 0$.

Proof. See appendix B. □

The necessary condition (18) is quite general: it applies to the class of utilities that are not differentiable at some particular points. For example, as option-like pay schedule can cause the principal's utility not differentiable at the strike or benchmark return. Such cases can be obtained when f_w is a delta function and, as a result, the corresponding F_w is discontinuous. The resulting optimal pay schedule $y(w)$ will have kinks at those discontinuous points of F_w.[22] Recall that, in the analysis of the agent's problem, $\lambda_f \equiv 0$ if and only $x(w)$ is concave. In other words, an interior solution f_w is obtained for the agent if and only if $x(w)$ is concave. Furthermore, the exact solution can be obtained by the concavification of the agent's utility given a fixed pay schedule. However, the first-order necessary condition for the principal's problem is much complicated, the concavification, not just over the principal utility, is over a combination of both the principal and agent's utilities plus a term caused by the incentive constraint. Because the concavification depends on the endogenous variable it is difficult to determine the conditions under which $\lambda_f = 0$ in general. However, the next proposition offers some ideas to determine λ_f.

Proposition 2 *Given the principal's problem (15), then a necessary condition for $\lambda_f(w) > 0$ is that $u(\alpha(w) - h(x(w)))$ is nonconcave over $\{t|\lambda_f(t) > 0\}$. If $\alpha(w)$ is twice differentiable, then a necessary condition for $\lambda_f(w) = 0$ is*

$$u''[\alpha' - h'x']^2 - u'h''[x']^2 + u'\alpha'' \leq 0. \tag{19}$$

Proof. See appendix B. □

Proposition 2 shows that corner solutions ($\lambda_f > 0$) can only occur in nonconcave regions of $u(\alpha(w) - h(x(w))$ and that interior solutions are obtained when $\alpha(w)$ is linear or α'^{∂} is small enough. The convex pay can be motivated by the empirical evidence of the presence of a convex fund flow.

[22]Here a corner solution refers to a case of $f_w(w) \equiv 0$ in a subset of $\{w|0 < F_w(w) < 1\}$, the measure of which is greater than zero.

Now for the region in which $\lambda_f(w) = 0$, a differential equation can be derived to further investigate the property of the optimal solutions. Taking derivatives with respect to w to the first-order condition (18) implies that

$$\frac{d[u + (\lambda_v - \lambda_w/\lambda)\, x]}{dw} + \frac{\lambda}{f_p} \int_0^w [u'h' - \lambda_v] f_w(t)\, dt = 0. \tag{20}$$

This differential-integral equation becomes an ordinary differential equation by taking derivatives one more time and using the fact that

$$f_w(w) = -\frac{1}{\lambda} f_p(x'/\lambda) x''.$$

This yields

$$\left\{ \frac{d[u + (\lambda_v - \lambda_w/\lambda)\, x]}{dw} \frac{f_p'}{\lambda f_p} - \left[2[u'h' - \lambda_v] + \frac{\lambda_w}{\lambda} \right] \right\} x'' \\ + \left[u''[\alpha' - h'x']^2 - u'h''[x']^2 + u'\alpha'' \right] = 0. \tag{21}$$

For the remaining part of the paper, unless specified otherwise, we will work with the case $\lambda_f \equiv 0$. Under such a circumstance, further simplification produces the following results.

Proposition 3 *Suppose that u,α and v are twice differentiable and that the optimal pay schedule y is twice differentiable. Then $y(w)$ must satisfy the following ordinary differential equation (ODE)*

$$(\beta v' - 2u')y'' + \frac{v''}{v'}(\beta v' - u')[y']^2 + u''[\alpha' - y']^2 + u'\alpha'' \\ + \left[v'y'' + v''[y']^2 \right] \left(u'[\alpha' - y'] + (\beta - \lambda_v)v'y' \right) \frac{f_p'}{\lambda f_p} = 0, \tag{22}$$

or $x = v \circ y$ satisfies

$$\left[\beta - 2u'h' + \left(u' \left[\alpha' - h'x' \right] + (\beta - \lambda_v)x' \right) \frac{f_p'}{\lambda f_p} \right] x'' \\ + u'' \left[\alpha' - h'x' \right]^2 - u'h''[x']^2 + u'\alpha'' = 0, \tag{23}$$

where $\beta = 2\lambda_v - \lambda_w/\lambda$ and $f_p = f_p(x'/\lambda) = f_p(v'y'/\lambda)$.

Proof. See appendix B. □

Although the two ODEs are equivalent, in practice, it might be easier to solve the ordinary differential equation of x than the equation for the pay schedule y, as x has a nicer property. Once x is solved, then the pay schedule can be determined by $y(w) = h(x) = v^{-1}(x)$. From Proposition 2, we know that a necessary condition for $\lambda_f(w) = 0$ is (19). In light of this inequality and the fact that the second derivative of x is nonpositive, equation (23) implies that the solution x must satisfy the following constraint

$$\beta - 2u'h' + (u'[\alpha' - h'x'] + (\beta - \lambda_v)x')\frac{f_p'}{\lambda f_p} < 0. \tag{24}$$

The standard theory of differential equations shows the following.

Proposition 4 *Given a set of* $(\lambda, \lambda_\ell, \lambda_w, \lambda_v)$ *and an initial condition* $(x(\ell) = \lambda_\ell, x'(\ell))$ *that satisfies constraint (24), the ordinary differential equation (23) has a unique solution that satisfies constraint (24). Furthermore,* x *is differentiable with respect to parameters* $(\lambda, \lambda_\ell, x'(\ell), \lambda_w, \lambda_v)$.

Proof. See appendix B. □

To determine $y(w)$ or $x(w)$ completely, we need equations (16) and (17), which can be rewritten as follows:

$$\int f_p(x'/\lambda)x''x(w)[u'h' - \lambda_v]\,dw = 0, \tag{25}$$

and

$$\int f_p(x'/\lambda)x''[u'h' - \lambda_v]\,dw = 0. \tag{26}$$

In addition, $(\lambda_\ell, x'(\ell))$ must satisfy the differential-integral equation (20) to ensure equation (23) holds.

Finally, we also need the budget and participation constraints

$$\int wf_p(x'/\lambda)x''x'dw = -\lambda^2 w_0 \quad \text{and} \quad \int f_p(x'/\lambda)x''xdw = -\lambda v_0. \tag{27}$$

The five parameters $(\lambda, \lambda_\ell, x'(\ell), \lambda_w, \lambda_v)$ are determined by the five equations mentioned above. When we figure out $x(w)$, the optimal pay schedule will be given by $y = h(x(w))$.

Last, one may wonder whether the necessary condition (18) is enough to insure an optimal solution or not. To do this, we need to examine the second-order condition. Surprisingly, the necessary condition (18) is also sufficient for a local optimal solution if $u(\alpha(w) - h(x)$ is globally concave in $(w, x) \in R^+ \times R$, i.e., α'' is relatively small enough. Clearly, $u(\alpha(w) - h(x))$ is concave in (w, x) for the case in which $\alpha(w)$ is linear or concave.

Proposition 5 *Suppose $u(\alpha(w)-h(x))$ is a concave function of $(w, x) \in R^+ \times R$. Given the multipliers (λ_w, λ_v), the solution $(f_w, \lambda, \lambda_\ell)$ to the first-order conditions (16), (17) and (18) is a a strong local maximizer of the principal's problem (15), i.e., the Lagrangian*

$$\mathcal{U}(f_w, \lambda, \lambda_\ell, \lambda_w, \lambda_v, \lambda_f).$$

Proof. See appendix B. □

The above analysis shows that, theoretically, there are two equivalent ways of looking for a local optimal solution $(f_w, \lambda, \lambda_\ell)$. One way is that, given (λ_w, λ_v), solving $(f_w, \lambda, \lambda_\ell)$ first, then using the budget constraint and participation constraint to fix (λ_w, λ_v). Finally, $y(w) = h(x(w))$ where $x(w)$ is uniquely determined by equation (13). The other way is to solve x by the ordinary differential equation (23) with boundary conditions directly, then f_w is also uniquely determined once x is solved. The former way is more general since it does not require the smoothness of x or y in solving the problem, while the latter way is more tractable as it converts the problem into solving nonlinear ODEs. Given this observation, the propositions (especially, 4 and 5) in this section imply the following main result of the paper.

Theorem 1 *Suppose $u(\alpha(w) - h(x))$ is a concave function of $(w, x) \in R^+ \times R$. A pay schedule $y = h(x)$ is a local optimal solution to the original principal-agent problem if and only if $(x(w), \lambda, \lambda_\ell, \lambda_w, \lambda_v, \lambda_f)$ satisfy equations (18), (16), (17) and (27); or equivalently, if $y = h(x)$ is smooth, then it is a local optimal solution if and only if $(x(w), \lambda, \lambda_\ell, x'(\ell), \lambda_w, \lambda_v)$ satisfy equations (23), (25), (26), (20) and (27).*

Proof. See appendix B. □

In general, the ordinary differential equation (22) or (23) does not admit a closed-form solution. Hence we need to develop a systematic approach of solving the ODEs in order to study the general properties of the optimal contracts.

However, the optimality of option-like pay schedules can be studied in a quite general setup. The conditions under which an option-like pay schedule is optimal are given in the following theorem.

Theorem 2 *Suppose that $\alpha(w)$ is smooth and $u(\alpha(w)-h(x))$ is a strictly concave function of $(w, x) \in R^{+} \times R$. Then the optimal pay schedule $y(w)$ contains no option-like feature and is strictly increasing.*

Proof. See appendix B. □

In the model presented here, no option-like pay schedules are optimal if the principal's pay $\alpha(w)$ is not convex shaped. When $\alpha(w)$ is a linear function of w, as is the case with most mutual fund fees, the option-like pays $y(w)$ cannot be the optimal contracts for the fund managers. An option-like pay schedule might be an optimal contract in the presence of implicit incentives that imply a convex $\alpha(w)$, as discussed in the introduction. Although, our model is static and assumes away the issues of asymmetric information, the result here offers some insights into option-like pays given that many studies that investigate the influence of option-like pays on the risk taking behaviors of managers employ the same setup.

4 First- vs Second-best Contracts

As the agent's action cannot be observed and the contract can only be written on final wealth, an efficiency issue about the optimal contract naturally arises. To address the efficiency issue we will first examine the characteristics of the first-best contract so that a benchmark case can be obtained. To be more specific, the first-best contract consists of a final wealth function $w(\omega)$, which is equivalent to a detailed instruction of portfolios, and a pay schedule $y(\omega)$ such that the pair $(w(\omega), y(\omega))$ maximizes the principal's expected utility under the budget constraint:

$$\max_{w(\omega)\in A,\, y(\omega)} \int_{\Omega} u(\alpha(w(\omega)) - y(\omega))\, P(d\omega) \tag{28}$$

subject to the participation constraints:

$$\int_{\Omega} v(y(\omega))\, P(d\omega) \geq v_0. \tag{29}$$

This maximization problem is straightforward when $u(\alpha(w) - y)$ is concave for all y, but it becomes complicated when $\alpha(w)$ is convex. Therefore, it is helpful to reformulate the problem into principal's choosing the distribution of wealth f_w instead of w. However, we cannot directly apply Lemma 1 to transform the problem into choosing f_w because of the additional term $y(\omega)$.

Lemma 6 *The first-best pay schedule $y(\omega)$, which together with $w(\omega)$ solves the maximization problem (28)-(29), can be written as $y(w(\omega))$.*

Proof. See appendix B. □

Lemma 6 shows we can replace $y(\omega)$ by $y(w(\omega))$ in the maximization problem (28)-(29). Then, applying Lemma 1 to $u+\lambda v$ implies that the first-best contracting problem is equivalent to

$$\max_{f_w(w)\in A_w,\, y(w)} \int u(\alpha(w) - y(w)) f_w(w)\, dw$$

subject to

$$\int v(y(w)) f_w(w)\, dw \geq v_0.$$

In addition, for the sake of comparison with the case of the second best, we use $y = h(x) = v^{-1}(x)$. Then the Lagrangian of the reformulated maximization problem is as follows:

$$\begin{aligned}\mathcal{L}(f_w, x, \lambda_w, \lambda_v, \lambda_f) \;&=\; \int u(\alpha(w) - h(x)) f_w(w)\, dw - \lambda_w \left(E[pw] - w_0\right) \\ &\quad + \lambda_v \left[\int x f_w(w)\, dw - v_0\right] + \int \lambda_f(w) f_w(w)\, dw.\end{aligned}$$

Using the variational method that perturbs f_w by $\epsilon_f \eta_f(w)$ and x by $\epsilon_x \eta_x(w)$, where ϵ_f and ϵ_x are two constants and

$$\int \eta_f(w)\, dw = 0 \quad \text{and} \quad \int \eta_x(w) f_w(w) < \infty,$$

we immediately have

Proposition 6 *A pair $(f_w, x(w))$ is an optimal solution for the principal's problem if and only if it satisfy the following first-order condition*

$$u(\alpha(w) - h(x(w))) + \lambda_v x(w) + \lambda_f(w) - \lambda_w \int_\ell^w F_p^{-1}(1 - F_w(t))\, dt = C \quad (30)$$

and

$$[u'(\alpha(w) - h(x))h'(x) - \lambda_v]\, f_w(w) = 0 \tag{31}$$

almost everywhere for any ℓ such that $F_w(\ell) > 0$, where C is independent of w and x.

Proof. See appendix B. □

Differentiating (30) with respect to w yields

$$u'[\alpha' - h'x'] + \lambda_v x' + \lambda_f' - \lambda_w F_p^{-1}(1 - F_w(w)) = 0. \tag{32}$$

Note that, for $f_w(w) > 0$, we have $\lambda_f \equiv 0$. Then, in such a condition, if $\alpha(w)$ is smooth, the first-order conditions become

$$u'\alpha' = \lambda_w F_p^{-1}(1 - F_w(w)) = \lambda_w p \quad \text{and} \quad u'h' = \frac{u'}{v'} = \lambda_v, \tag{33}$$

where the first term means that the principal's marginal rate of substitution across states it equal to its relative price ratio, and the second term represents the optimal risk-sharing between the principal and the agent. From equation (18) we have that if the second-best contract is efficient or $u'h' = \frac{u'}{v'} = \lambda_v$, then

$$u(\alpha(w) - y(w)) + \left[\lambda_v - \frac{\lambda_w}{\lambda}\right] v(y(w)) = C. \tag{34}$$

Note that condition (34) is called the similarity condition in the literature, and was proposed early in the theory of syndicates and agency studied by Wilson (1968) and Ross (1973).[23] Condition (34) shows that the efficiency condition implies the similarity condition. Combining these two conditions, it is easy to show that the optimal contract must take the form

$$y(w) = k_1\alpha(w) + k_2, \tag{35}$$

where $k_1 > 0$ and k_2 are some constants. In particular, if $\alpha(w)$ is linear, then $y\,(w)$ will be linear.

In the case of $\alpha(w) = w$, Ross (1973) was able to completely characterize the class of utility pairs (u, v) that satisfy both the efficiency condition and the

[23] See also Ross (1974), Dybvig and Spatt (1986)

similarity condition, and to show that in such situations the pay schedule is linear. The class is simply that of pairs (u, v) with linear risk tolerance,

$$-\frac{u'}{u''} = cw + d \quad \text{and} \quad -\frac{v'}{v''} = cw + e\,, \tag{36}$$

where c, d and e are constants. It is easy to show that in the general case $\alpha\,(w)$ the class of utility pairs that satisfy both the efficiency condition and the similarity condition is still the same. However, the pay schedule is no longer linear but takes the form of equation (35). Our discussions thus lead us to the following results.

Theorem 3 *The second-best pay schedule is Pareto efficient if and only if the utility pair (u, v) for the principal and the agent satisfies equation (36). Furthermore, in case $\alpha(w) = \alpha w$, the second-best pay schedule must be linear if it is Pareto efficient.*

Proof. Omitted. □

So far we have discussed the circumstances under which the second-best contracts are efficient. The utility pairs have to satisfy equation (36) to achieve the efficiency. In particular, if u and v are exponential utilities, then equation (36) is satisfied automatically. In general, there is an efficiency loss when the principal and the agent's risk attitudes are divergent. To see this point clearly, we rewrite equation (23) into the following form

$$\frac{u'}{v'} = \beta + \left[\frac{du}{dw} + (\beta - \lambda_v)x'\right]\frac{f_p'}{\lambda f_p} + \frac{1}{x''}\frac{d^2u}{dw^2}. \tag{37}$$

Equation (37) indicates that the marginal rate of substitution is equal to a constant plus the two terms that capture the incentive constraint imposed on the principal by the need to motivate the agent to act in the best interest of the principal. The second term captures the effect of the price density, while the last term captures the effect of the incentive constraint.

Note that in case of $\alpha\,(w) = \alpha w$, condition (36) is sufficient but not necessary for the second-best pay schedule to be linear. In addition, a linear optimal pay schedule does not have to be Pareto efficient. To see this, we can focus on the special case where the price density is uniformly distributed. Under such a condition, equation (37) can be reduced to

$$\frac{u'}{v'} = \beta + \frac{1}{x''}\frac{d^2u}{dw^2}. \tag{38}$$

In what follows we derive a necessary condition for a utility pair (u, v) to have a linear second-best pay schedule. Suppose a second-best pay schedule $y(w) = k_1 w$. Then, equation (38) implies that

$$\frac{u'}{v'} = \beta + \frac{\alpha - k_1}{k_1}\left(\frac{du'}{dw} / \frac{dv'}{dw}\right). \tag{39}$$

Let $m = \frac{\alpha - k_1}{k_1}$. It is easy to check that the following satisfies equation (39)

$$u' = \frac{C}{1-m}[v']^{\frac{1}{m}} + \frac{\beta}{1-m}v', \tag{40}$$

where $C \geq 0$ is a constant. Simplifying equation (40) further and we have

$$u(w) = \frac{\beta m}{1-m} v\left(\frac{w}{m}\right) + \frac{Cm}{1-m}\int_0^{\frac{w}{m}} (v'(s))^{\frac{1}{m}} ds + C_0. \tag{41}$$

Equation (41) can be used to generate examples for a pair (u, v) to have a linear second-best pay schedule (with an appropriate adjustment of other parameters such as v_0 and w_0, of course). In particular, when the agent's utility has a power form $v = \frac{1}{1-\gamma}w^{1-\gamma}$, the principal's utility u can be calculated explicitly as follows

$$u(w) = \frac{\beta m}{1-m}\frac{1}{1-\gamma}\left(\frac{w}{m}\right)^{1-\gamma} + \frac{Cm}{1-m}\frac{1}{1-\frac{\gamma}{m}}\left(\frac{w}{m}\right)^{1-\frac{\gamma}{m}} + C_0. \tag{42}$$

By setting $\beta = 0$ (via adjusting other parameters properly) and $C_0 = 0$, equation (42) implies that for any power utility pair (u, v) given by $v = \frac{1}{1-\gamma}w^{1-\gamma}$ and equation (42), there exists a linear second-best pay schedule for a certain set of initial parameters.

5 Numerical Examples

As discussed so far, the optimal contracts do not admit closed-form solutions in general. To further understand the properties of the optimal contracts we solve several examples numerically. As we do not intend to develop full-fledged numerical methods to solve the contracting problems, to ease the programming we will ignore the effect of price density on the optimal contracts and assume that the state price p is uniformly distributed over $[p_l, p_h]$. In addition, we also assume that the gross benefit to the principal is linear. That is, $\alpha(w) = \alpha w$, where α is a constant. Under such a condition, we have $f_p(p) = \frac{1}{p_h - p_l}$, $F_p(p) = \frac{p - p_l}{p_h - p_l}$ and $f_p' = 0$. Thus, equation (22) can be further reduced to

$$(\beta v' - 2u')y'' + \frac{v''}{v'}(\beta v' - u')(y')^2 + u''(\alpha - y')^2 = 0, \tag{43}$$

and equation (23) becomes

$$[\beta - 2u'h']\, x'' + u''\, [\alpha' - h'x']^2 - u'h''[x']^2 = 0. \tag{44}$$

Either of the two ODEs above can be used to solve the contracting problem with the constraint

$$\beta - 2u'h' < 0.$$

This constraint is needed to make sure that $x'' < 0$ (see equation (24)).

For the case of uniform distribution of state prices, in addition to the four integral constraints (25) - (27), we also need the following equations to determine the boundary of the distribution of the final wealth

$$x'(w_l) = \lambda p_h \quad \text{and} \quad x'(w_h) = \lambda p_l. \tag{45}$$

The uniform distribution assumption also makes the differential-integral equation (20) straightforward: we can fix the constraint at a particular point. A natural choice is one of the two boundary points, i.e., at w_l, the constraint becomes

$$\left[u'[\alpha - h'x'] + \left(\lambda_v - \frac{\lambda_w}{\lambda} \right) x' \right] \Bigg|_{w = w_l} = 0, \tag{46}$$

because $f_w(w) \equiv 0$ for all $w \notin [w_l, w_h]$.

For preferences, we assume both the principal and agent have a power utility in the form

$$\frac{1}{1-\gamma} w^{1-\gamma},$$

where γ is the relative risk aversion coefficient. Throughout the remaining part of this section we will use γ_p and γ_a to represent the risk aversion coefficients of the principal and agent, respectively.

For ease of exposition, we use a parameter κ, a fraction of the initial wealth, to measure the agent's reservation utility. Parameter κ is a fraction of the initial wealth such that

$$\frac{1}{1-\gamma_a} (\kappa w_0)^{1-\gamma_a} = v_0.$$

From our discussions in the previous section, we know that the first best contracts do not need to consider the incentive effect, and only the risk sharing effect should be considered. The optimal condition for the first best contracts implies that the marginal rate of substitution is a constant. Consequently, the risk is allocated between the principal and the agent optimally. However, the second best contracts would have different concerns. In addition to this risk sharing effect, the second best contracts also need to consider the incentive effect. These two effects — risk sharing and incentive — mingle together and cannot be separated.

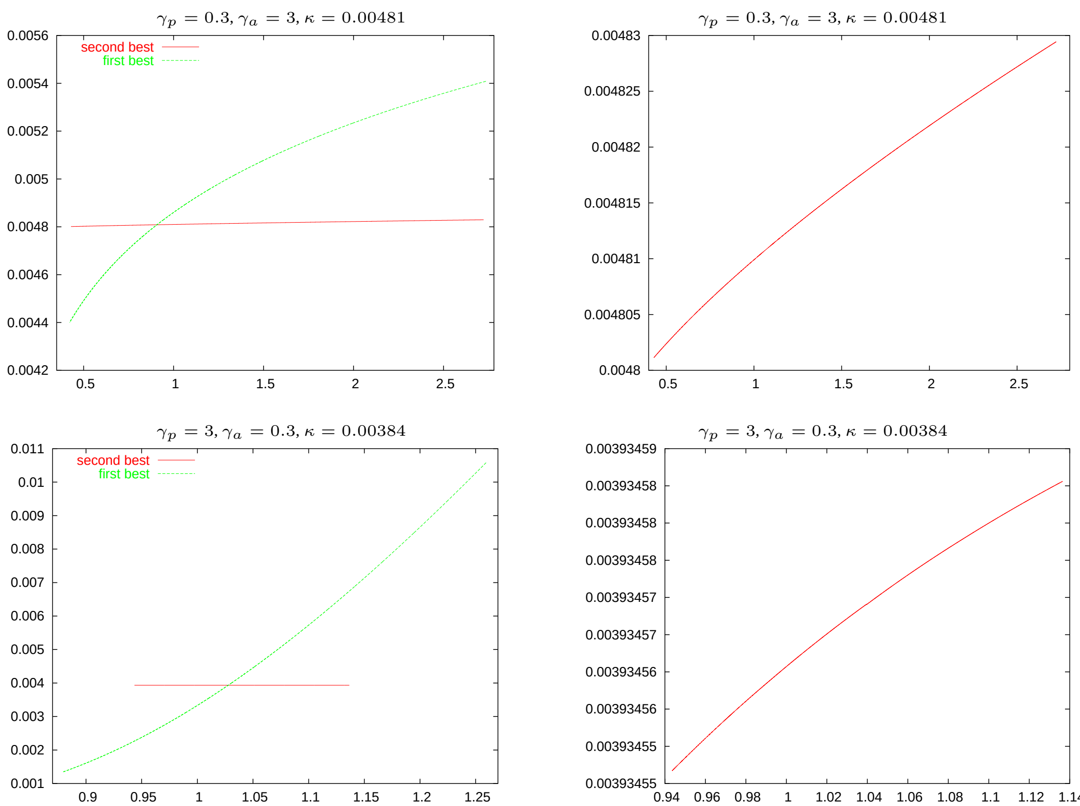

Figure 1: The first and second best contracts. The right graphs plot the second best contract only. Preferences are power utilities. The gross benefit of the principal $\alpha(w) = 0.05w$. State price follows uniform distribution with bounds: $p_l = 0.7$ and $p_h = 1.3$. Both the pay schedule and the wealth are normalized by the initial wealth w_0.

In the first numerical example, we choose the risk aversion coefficients as 0.3

and 3, such that the incentive effect could be relatively strong. Figures 1 and 2 plot the first best and the second best contracts for these two risk parameters. These two figures show that the shape of the second best contracts can not be solely determined by the relative risk aversion. Figure 1 shows that no matter who is more risk averse, the second best contract is concave and flat. In contrast, the first best contract is concave for a less risk-averse principal and convex for a more risk-averse principal.

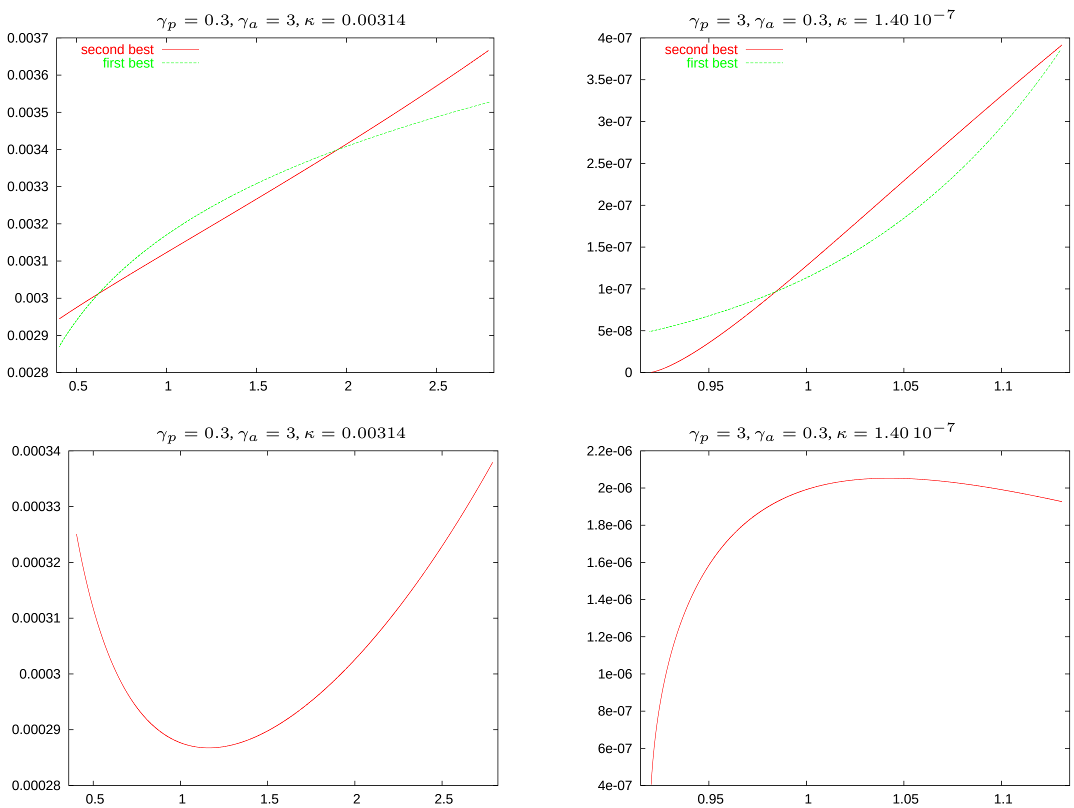

Figure 2: The first and second best contracts and the derivative of the second best contract. Preferences are power utilities. The gross benefit of the principal $\alpha(w) = 0.05w$. State price follows uniform distribution with bounds: $p_l = 0.7$ and $p_h = 1.3$. Both the pay schedule and the wealth are normalized by the initial wealth w_0.

Note that the risk-sharing effect demands that the marginal rate of substitution is constant, while the incentive effect demands that the pay schedule is designed in a way such that the composite utility functions of both the principal and the agent are similar. These two roles are normally conflicting. The shape of the second best contracts are the net effects of these two opposite forces.

A comparison of the plots in Figure 1 with the ones in Figure 2 shows the reservation utility plays a very important role in determining the shape of the second best contracts. However, the shape of the first best contracts is more or less only determined by the relative risk aversion, independent of the reservation utility level. This is due to the nature of the power utilities, and the fact that the first best contracts are characterized by a constant marginal rate of substitution.

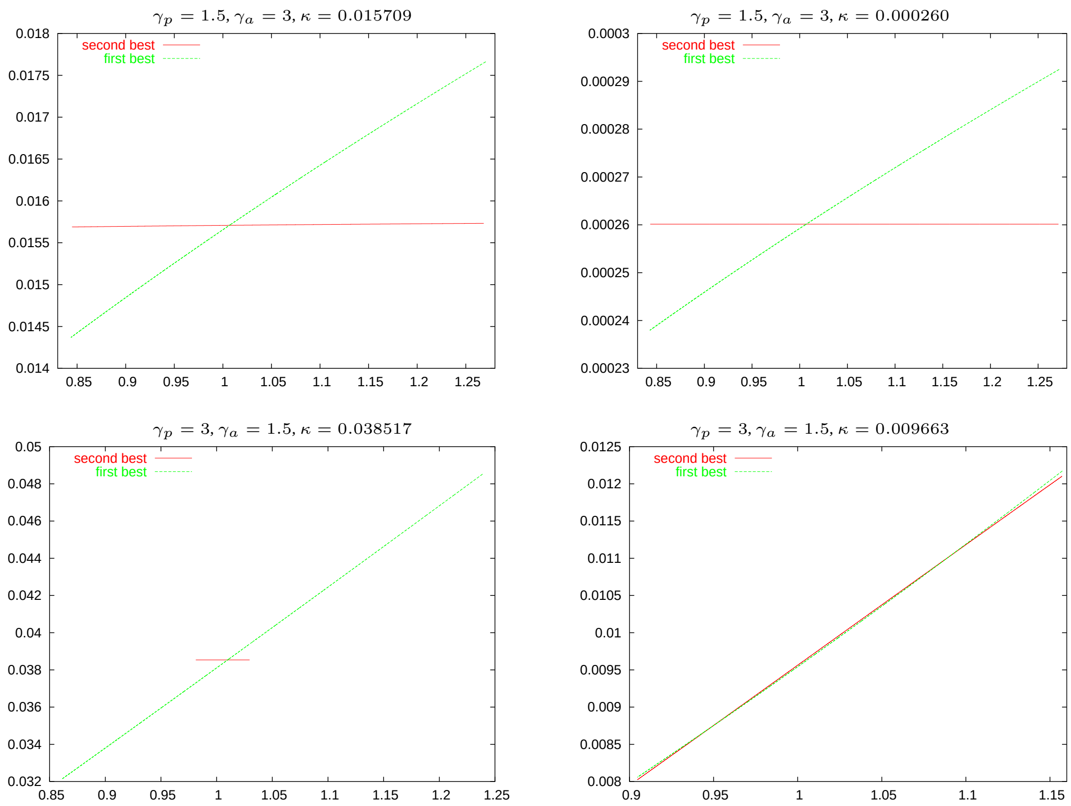

Figure 3: The first and second best contracts. Preferences are power utilities. The gross benefit of the principal $\alpha(w) = 0.05w$. State price follows uniform distribution with bounds: $p_l = 0.7$ and $p_h = 1.3$. Both the pay schedule and the wealth are normalized by the initial wealth w_0.

When the reservation utility is relatively low, the second best contracts contain both concave and convex regions. Figure 2 shows that the second best contract is concave in the low wealth region and convex at the high wealth region when the principal is less risk averse. The concave and convex regions reverse when the principal is more risk averse.

In the case in which the principal and agent's risk aversion coefficients are close the second best contracts are still quite different from the first best contracts. Figure 3 plots some examples in this situation. When the principal is more risk averse, the second best contracts are flat and concave even for a relatively low reservation utility level. When the agent is more risk averse, the second best contracts can be close to the first best contracts for low reservation utility levels.

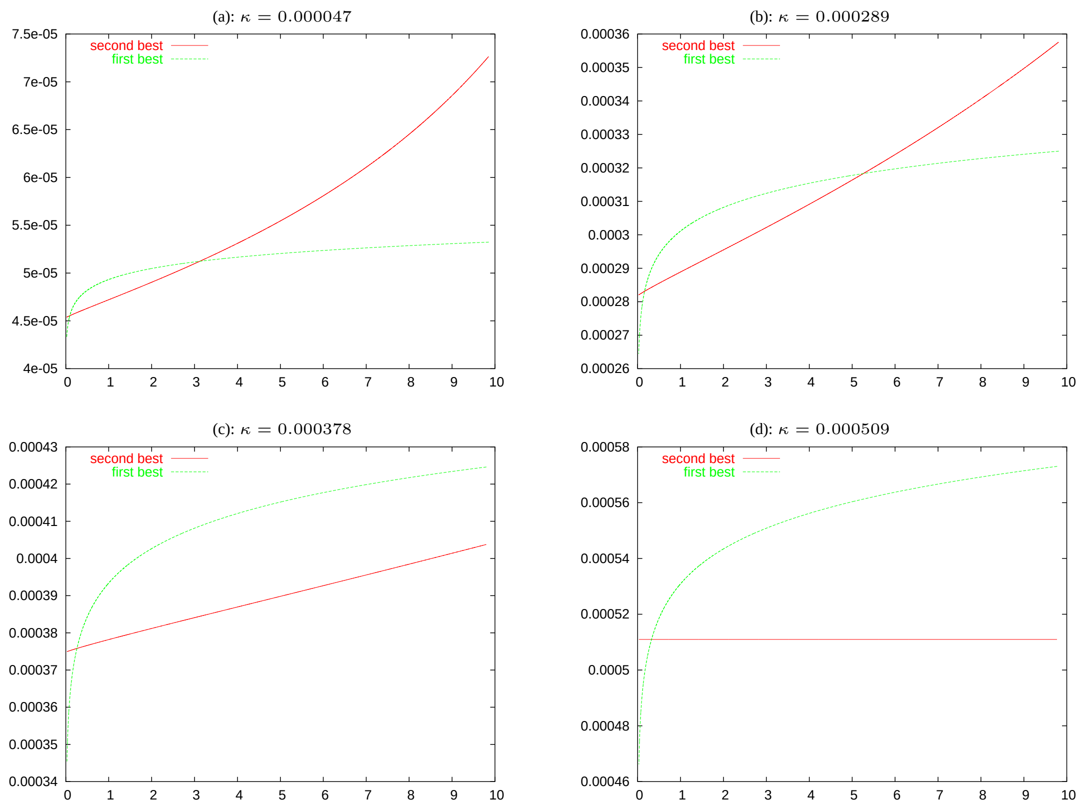

Figure 4: The first and second best contracts. Preferences are power utilities with risk aversion $\gamma_p = 0.1$ and $\gamma_a = 3$. The gross benefit of the principal $\alpha(w) = 0.05w$. State price follows uniform distribution with bounds: $p_l = 0.7$ and $p_h = 1.3$. Both the pay schedule and the wealth are normalized by the initial wealth w_0.

In Figure 4, the risk aversion parameters take values of 0.1 and 3 hence the incentive effects become larger than what we had before. It further supports the idea that the reservation utility level plays an important role in determining the shapes of the second best contracts. The reservation level increases for the four plots from (a) to (d). The shapes of these contracts change from the convex, to mix of concave and convex (left and bottom), and to the concave.

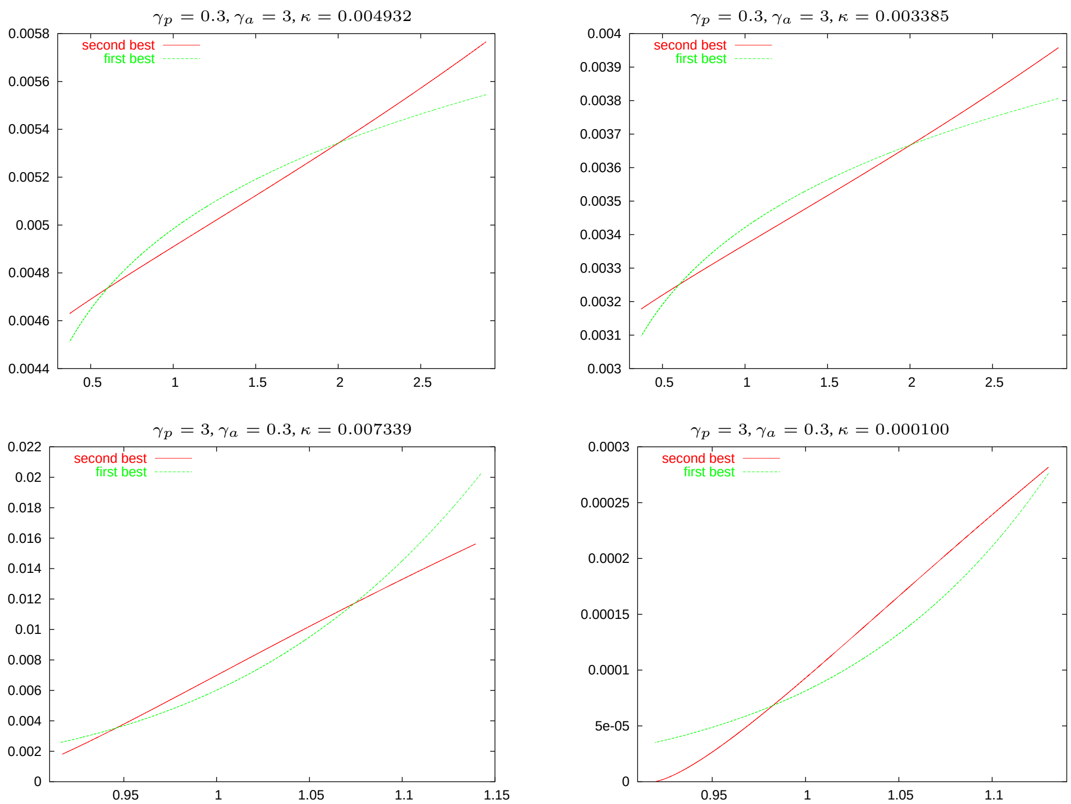

Figure 5: The first and second best contracts. Preferences are power utilities. The gross benefit of the principal $\alpha(w) = w$. State price follows uniform distribution with bounds: $p_l = 0.7$ and $p_h = 1.3$. Both the pay schedule and the wealth are normalized by the initial wealth w_0.

Another interesting observation we can draw from Figure 1-4 is that most of the second-best pay schedules are relatively "flat" for the case when $\alpha(w) = 0.05w$, even when the risk aversions of the principal and agent are quite different. This suggests that the flat pays should be common in the mutual fund industry, in which a fund company typically charges a small fixed percentage fee. This result is consistent with the empirical findings of Elton, Gruber, and Blake (2003) that fund managers' pays are usually flat in the US mutual fund industry.

Finally, Figure 5 plots some examples when the principals are themselves investors, hence $\alpha(w) = w$. In this case, the shape of the second best contracts become less sensitive to the reservation utility level due to the low "costs" of solving the incentive problems. The second best contracts are much steeper than those for smaller α in the previous examples.

Overall, these numerical examples show that at least in the class of power utilities, there is no determinant relationships between the shapes of the second-best contracts and the differences of the risk aversion coefficients between the principal and agent. Given the risk aversion coefficients, the shape of the second best contracts strongly depends on the agent's reservation utility level. In addition, it is unusual for a second best-contract to have a strong concave or convex shape. Second best contracts are more "linear." As shown by Theorem 2, option-like pay schedules cannot be optimal without implicit incentives, and the numerical examples seem to suggest that the option-like pays cannot even be approximations to the optimal contracts. Option-like pays are only possible when there are convex implicit incentives such as the observed fund flows.

6 Conclusion

This paper has addressed a fundamental problem in the portfolio delegation problem: what is the optimal pay schedule when the investors delegate their portfolio choice decision to a fund manager? Much of the existing literature addresses this issue under very limited conditions. Indeed, only optimal linear contracts under restricted preferences are fully analyzed for technical reasons. In this paper, we develop a new approach to deal with those difficult issues arising from characterizing the agent's optimal action. By reformulating the principal-agent problem into a distribution form, we are able to show that an optimal indirect utility function $x(w)$(or, equivalently, an optimal pay schedule $y(w)$) can be found within the class of the agent's indirect utility functions that are increasing and concave. As a result, the traditional first-order approach can be applied to transform the incentive constraint into a normal one and the principal-agent problem becomes a standard constrained maximization problem. Consequently, the standard calculus of variation techniques are applied to convert the principal-agent problem into one that solves a second-order nonlinear ODE. Various numerical examples are solved to illustrate the shape of the (local) optimal pay schedules, as compared to the Pareto-efficient benchmark case. An important result is that the shape of the second-best contracts is smoother and flatter than that of the first-best ones. This is due to the fact that two offsetting forces, the risk-sharing effect and the incentive effect, jointly determine the shape of the second-best contracts, while the risk-sharing effect alone determines the shape of the first-best contracts completely.

Some issues about the nonlinear ODEs deserve further discussion. One important issue is existence problem. In the paper, we use the nonlinear ODEs to characterize the optimal contracts when they are known to exist. Surprisingly, the nonlinear ODEs are both necessary and sufficient for local optimality. As a result, an natural research topic in the next step is to provide conditions on utilities and the price density function such that the nonlinear ODEs have a solution. We have solved a simple case in which both the principal and the agent have a linear risk tolerance with identical cautiousness. However, developing a systematic approach to solve the nonlinear ODEs under very general conditions, at least numerically, is a worthwhile undertaking.

Our model can be extended along several lines. First, one may wonder whether or not the approach that we use in this paper can be carried over to the case of incomplete securities markets. Note that the concavification of the agent's composite utility function in Lemma 2 relies on the complete market assumption. In the presence of market incompleteness, the price density function can not summarize all relevant information. As a result, the optimal contracts will naturally be related to the players' risk attitude as well as the securities payoff structure. This is an interesting area for future research.

Second, we can also incorporate moral hazards and adverse selection into the analysis. For instance, the agent's efforts on information production are a typical moral hazard issue in the mutual fund industry. Dybvig, Farnsworth, and Carpenter (2004) and Sung (2005) have done some work along this line. However, their model places strong restrictions on the players' utility functions, and assumes that the states of nature can be contracted upon. Consequently, their model has limited applications. Exploring the implications of an principal-agent model with moral hazards and adverse selection in the context of delegated portfolio management would be another interesting research topic.

Third, although our agency model is static, it also captures the structure of the optimal pay schedule under the worst scenario in a dynamic setting. In other words, if the final wealth is the only variable that can be contracted upon, and there is no intertemporal consumption, then the optimal contract in this simple dynamic setting is identical to that in our static model. However, in the general case in which the agent makes optimal consumption and portfolio decisions over time, and the principal uses securities price information in the design of the pay schedules, the optimal pay schedule in the presence of signals remains an interesting topic, which we are currently working on (see Li and Zhou (2005)).

Appendix A: An Auxiliary Lemma

To solve the principal-agent problem (7)-(9), the method we will use here is the basic technique in the calculus of variations. The idea is as follows. Suppose that f_w is an optimal solution, then perturb f_w by $\epsilon\eta$, where ϵ is a small constant and η is an arbitrary integrable function that satisfies $\int \eta(w)\, dw = 0$. Such a restriction makes sure that $\int [f_w + \epsilon\eta]\, dw = 1$.[24] The fact that f_w is an optimal solution implies that all directional derivatives at f_w are zero. A direct calculation of all the directional derivatives, although quite involved, gives a necessary and sufficient condition for the local optimal solutions. The technique described above is the basic one, which is used to derive the Euler equations or first-order conditions, in the calculus of variations. There are at least two major reasons for us not directly use the Euler equations, though there are several ways of doing it. First is that our problems are not typical, standard results on the second-order necessary conditions cannot be applied. On the other hand, using the basic technique, the second-order conditions work out quite nicely. Secondly, corner solutions can be handled easily by the basic variational technique. Such corner solutions are important in economics because contracts with corner solutions contain option payoff structures, which are widely used in practice.

Since the variations of the expected wealth are used several times, we present them formally in the following lemma.

Lemma A.1 *For any given a distribution function f_w, let $E^\epsilon[pw]$ be the budget under a perturbed density function $f_w^\epsilon = f_w + \epsilon\eta$, where ϵ is a constant and η satisfies $\int \eta(t)\, dt = 0$. Then,*

$$\frac{dE^\epsilon[pw]}{d\epsilon} = \int_0^\infty \eta(w) \int_\ell^w F_p^{-1}(1 - F_w^\epsilon(t))\, dt\, dw$$

[24] This is analogous to the finite dimensional case, in which η is a directional vector.

and

$$\frac{d^2E^\epsilon[pw]}{d\epsilon^2} = \int_0^\infty \frac{1}{f_p(F_p^{-1}(1-F_w^\epsilon(w)))}\left(\int_0^w \eta(t)\,dt\right)^2 dw > 0,$$

where ℓ is an arbitrary nonnegative number such that $f_w(\ell) > 0$ or $\ell F_p^{-1}(1 - F_w(\ell))$ is finite.

Proof of Lemma A.1

We perturb the density function f_w by $\epsilon\eta$, where ϵ is a small constant and η is an arbitrary piece wise smooth function that satisfies

$$\int_0^\infty \eta(w)\,dw = 0.$$

Let F_w^ϵ denote the perturbed cumulative distribution function. Note that

$$1 - F_w^\epsilon(w) = 1 - \int_0^w [f_w(s) + \epsilon\eta(s)]\,ds = \int_w^\infty [f_w(s) + \epsilon\eta(s)]\,ds.$$

We use $f_p(w)$ as a shortcut for $f_p(F_p^{-1}(1-F_w^\epsilon(w)))$ in the remaining of the proof. Since

$$\begin{aligned}
&\int_0^\infty w f_w^\epsilon(w)\frac{dF_p^{-1}(1-F_w^\epsilon(w))}{d\epsilon}\,dw \\
= \; & -\int_0^\infty w f_w^\epsilon(w)\left[\frac{1}{f_p(w)}\int_0^w \eta(t)\,dt\right]dw \\
= \; & -\int_\ell^w t f_w^\epsilon(t)\frac{1}{f_p(t)}\,dt \int_0^w \eta(t)\,dt\Big|_0^\infty + \int_0^\infty \eta(w)\int_\ell^w t\frac{f_w^\epsilon(t)}{f_p(t)}\,dt\,dw \\
= \; & \int_0^\infty \eta(w)\Big[\, wF_p^{-1}(1-F_w^\epsilon(w)) - \ell F_p^{-1}(1-F_w^\epsilon(\ell)) \\
& \qquad\qquad - \int_\ell^w F_p^{-1}(1-F_w^\epsilon(t))\,dt\Big]\,dw,
\end{aligned}$$

where ℓ is a positive number, the first derivative of the perturbed budget then is

$$\frac{dE^\epsilon[pw]}{d\epsilon}$$

$$
\begin{aligned}
&= \int_0^\infty \eta(w) w F_p^{-1}(1 - F_w^\epsilon(w))\, dw - \int_0^\infty w f_w^\epsilon(w) \frac{dF_p^{-1}(1 - F_w^\epsilon(w))}{d\epsilon}\, dw \\
&= \int_0^\infty \eta(w) \left[\int_\ell^w F_p^{-1}(1 - F_w^\epsilon(t))\, dt + \ell F_p^{-1}(1 - F_w^\epsilon(\ell)) \right] dw \\
&= \int_0^\infty \eta(w) \left[\int_\ell^w F_p^{-1}(1 - F_w^\epsilon(t))\, dt \right] dw,
\end{aligned}
$$

where the last equality is due to the fact that $\int_0^\infty \eta(w)\, dw = 0$.

Then, taking derivatives with respect to ϵ to the above equation again and integrating by parts shows that

$$
\begin{aligned}
\frac{d^2 E^\epsilon[pw]}{d\epsilon^2} &= -\int_0^\infty \eta(w) \left[\int_\ell^w \frac{1}{f_p(w)} \int_0^t \eta(s)\, ds\, dt \right] dw \\
&= \int_0^\infty \frac{1}{f_p(w)} \left(\int_0^w \eta(t)\, dt \right)^2 dw > 0.
\end{aligned}
$$

The second derivative is positive because for all $w \in R^+$ when ϵ is zero or small enough such that F_w^ϵ is a valid distribution function.

Appendix B: Proofs

Proof of Lemma 1

Aumann and Perles (1965) show that G has a concavification $\tilde{G}$. Then, for each nonconcave region, there exists an open interval $B = (w_1, w_2) \in R^+$ such that

$$
L_B(w) = G(w_1) + \frac{G(w_2) - G(w_1)}{w_2 - w_1}(w - w_1) > G(w) \tag{47}
$$

for all $w \in B$. The the concavification is

$$
\tilde{G}(w) = \begin{cases} L_B(w) & \text{if } w \in \cup B \\ G(w) & \text{otherwise.} \end{cases}
$$

Lemma B.1 *Suppose that $G(w)$ satisfies the conditions in Lemma 1. Then*

$$
\max_{w(\omega) \in A} \int_\Omega G(w(\omega))\, P(d\omega) = \max_{w(\omega) \in A} \int_\Omega \tilde{G}(w(\omega)\, P(d\omega).
$$

If w^* solves the the maximization problem with G, it also solves the maximization problem with $\tilde{G}$.

Proof. Without loss of generality, suppose there is only one B. Let $\Omega' = \{\omega | \tilde{w}(\omega) \in B\}$. The lemma is true because for every choice $\tilde{w}$ under $\tilde{G}$ there is an equivalent choice of w under G

$$w(\omega) = \begin{cases} \tilde{w}(\omega) & \text{if } \omega \notin \Omega' \\ w_1 \text{ with prob } a(\tilde{w})P(d\omega) & \text{if } \omega \in \Omega' \\ w_2 \text{ with prob } (1 - a(\tilde{w}))P(d\omega) & \text{if } \omega \in \Omega', \end{cases}$$

where $a(\tilde{w}) = \frac{w_2 - \tilde{w}}{w_2 - w_1}$, such that

$$\begin{aligned} \int_\Omega G(w(\omega))\, P(d\omega) &= \int_{\Omega \setminus \Omega'} G(\tilde{w}(\omega)\, P(d\omega) \\ &\quad + \int_{\Omega'} \{a(\tilde{w})G(w_1) + [1 - a(\tilde{w})]G(w_2)\}\, P(d\omega) \\ &= \int_\Omega \tilde{G}(\tilde{w}(\omega))\, P(d\omega) \end{aligned}$$

and

$$\int_\Omega p(w)w(\omega)\, P(d\omega) = \int_\Omega p(\omega)\tilde{w}(\omega)\, P(d\omega).$$

The last equation is true because of the identity $\tilde{w} = a(\tilde{w})w_1 + [1 - a(\tilde{w})]w_2$. $\square$

Lemma B.1 shows that it is sufficient to prove Lemma 1 to show that it is true for concave G. The following lemma further shows we can assume G is also nondecreasing.

Lemma B.2 *Let $w_m = \inf\{w | G(w) = \max_{\hat{w}} G(\hat{w})\}$. Suppose that G is concave and bounded on all compact set in R^+. Then G is either nondecreasing or $P(\{\omega | w^*(\omega) \in (w_m, \infty)\}) = 0$ and $G(w)$ is nondecreasing on $[0, w_m]$.*

Proof. Since G is bounded and concave, there exists a maximum if it is not increasing. As w_m dominates all $w > w_m$ and $pw_m < pw$, we have $P(\{\omega | w^*(\omega) \in (w_m, \infty)\}) = 0$. $\square$

Lemmas B.1 and B.2 show that it suffices to show the lemma is true for G that is nondecreasing and concave. Since markets are complete, there is a unique p corresponding to each $\omega \in \Omega$. This implies

$$\int_\Omega G(w^*(\omega))\, P(d\omega) = \int F_p(dp) \int_{\{\omega \in \Omega | p(\omega) = p\}} G(w^*(\omega))\, P(d\omega | p), \tag{48}$$

where F_p is the induced probability measure through $p(\omega)$ and $P(\omega|p)$ is the probability measure conditional on p. In addition, for the budget, we also have

$$\begin{aligned}\int_\Omega p(\omega)w^*(\omega)\,P(d\omega) &= \int p\,F_p(dp)\int_{\{\omega|p(\omega)=p\}} w^*(\omega)\,P(d\omega|p)\\ &= \int pE[w^*|p]\,F_p(dp). \qquad (49)\end{aligned}$$

As G is concave, applying Jensen's inequality to the righthand side of equation (48) then yields

$$\int_\Omega G(w^*(\omega))\,P(d\omega) = \int E[G(w^*)|p]\,F_p(dp) \le \int G(E[w^*|p])\,F_p(dp).$$

The inequality becomes equality if $w^*(\omega) = w^*(p)$ is the same on $\{\omega \in \Omega | p(\omega) = p\}$. This choice satisfies the budget constraint in light of equations (49). Therefore,

$$\max_{w\in A}\int_\Omega G(w(\omega))\,P(d\omega) = \max_{w\in A_p}\int G(w(p))\,F_p(dp), \qquad (50)$$

where

$$A_p = \{w > 0 | \int pw(p)\,F_p(dp) \le w_0\}.$$

On the other hand, for any feasible choice of $w(\omega)$, we have (see, i.e., Theorem 7 in §6, Chapter II of Shiryaev (1995))

$$\int_\Omega G(w(\omega))\,P(d\omega) = \int G(w)\,F_w(dw),$$

where F_w is the $w(\omega)$ induced probability measure. Combining this with equation (50) implies that

$$\max_{w\in A_p}\int G(w(p))\,F_p(dp) = \max_{w\in A_p}\int G(w)\,F_w(dw). \qquad (51)$$

To prove equation (6), it is sufficient to establish that the inverse of $w(p)$ is monotonically decreasing with w at maximum. However, this is true by the first-order necessary and sufficient condition $G'(w) \le \lambda p$ for the righthand maximization problem of (50), where $\lambda > 0$. This shows that the positive relation between p and w cannot be optimal. Therefore, for the maximization problems in equation

(51), we can restrict the feasible set A_p such that p is decreasing (non-increasing) with w, which can be represented in terms of distribution function of w by

$$F_p(p) = 1 - F_w(w) = \int_w^\infty f_w(w)\, dw, \tag{52}$$

where $\infty \geq f_w \geq 0$. Note that the only requirement on F_w is that it is non-decreasing, or equivalently, $f_w \geq 0$. This restriction is also equivalent to $p = F_p^{-1}(1 - F_w(w))$. Hence we can replace A_p by A_w in the righthand side of equation (51). This establishes the second equality of equation (6).

Proof of Lemma 2

Let $x(w) = v(y(w))$. In the form of distribution with the budget constraint (5), the Lagrangian for the agent's problem is

$$\begin{aligned}\mathcal{V} &= \int_0^\infty x(w) f_w(w)\, dt - \lambda \left(\int_0^\infty w F_p^{-1}(1 - F_w(w)) f_w(w)\, dw - w_0 \right) \\ &\qquad + \int_0^\infty \lambda_f(w) f_w(w)\, dw,\end{aligned}$$

where λ is a positive constant and $\lambda_f(w)$ is a nonnegative function, which equals zero when $f_w > 0$ and is nonnegative when $f_w = 0$. Then, following the procedure as described in the proof of Lemma A.1, let $f_w^\epsilon(w) = f_w(w) + \epsilon\eta(w)$, where $\int_0^\infty \eta(w)\, dw = 0$ and f_w is the optimal solution. Then define the perturbed expected utility as

$$\mathcal{V}^\epsilon = \int_0^\infty [x(w) + \lambda_f(w)]\, f_w^\epsilon(w)\, dw - \lambda \left(\int_0^\infty w[F_p^{-1}(1 - F_w^\epsilon(w)) f_w^\epsilon(w)\, dw \right).$$

Using Lemma A.1, the first two derivatives of the Lagrangian with respect to ϵ are

$$\frac{d\mathcal{V}^\epsilon}{d\epsilon} = \int_0^\infty \eta(w) \left[x(w) + \lambda_f(w) - \lambda \int_\ell^w F_p^{-1}(1 - F_w^\epsilon(t))\, dt \right] dw,$$

and

$$\frac{d^2\mathcal{V}^\epsilon}{d\epsilon^2} = -\lambda \int_0^\infty \frac{1}{f_p} \left(\int_0^w \eta(t)\, dt \right)^2 dw < 0.$$

This shows that the Lagrangian $\mathcal{V}$ always has an interior maximum and the first-order condition $\frac{d\mathcal{V}^\epsilon}{d\epsilon}|_{\epsilon=0} = 0$ is both sufficient and necessary when we restrict $f_w \geq 0$. Yet this is true if and only if there exists a constant C such that

$$x(w) - \lambda \int_\ell^w F_p^{-1}(1 - F_w(t))\, dt + \lambda_f(w) \equiv C \tag{53}$$

holds almost everywhere. As x is bounded above by $v(\alpha(w))$ there exists a function F_w or f_w such that equation (53) holds.

This equality and Lemma A.1 imply $C = x(\ell)$ for all $\ell \in \{t|\infty > f_w(t) > 0\}$.

Let $\lambda_\ell = x(\ell)$ and define

$$\tilde{x}(w) = \lambda \int_\ell^w F_p^{-1}(1 - F_w(t))\, dt + \lambda_\ell.$$

As $\lambda_f(w) \geq 0$, we have $x(w) \leq \tilde{x}(w)$ and $\tilde{x}(w)$ is nondecreasing and concave, $\tilde{x}(w)$ is the concavification of $x(w)$. Otherwise, it contradicts the first-order condition. Because $\tilde{x}(w) > x(w)$ implies that $\lambda_f(w) > 0$ and hence $f_w(w) = 0$. Therefore, by the fact that $\tilde{x} \geq x$, we have $f_w(w) = 0$ for all $w \in \{t|\tilde{x}(t) \neq x(t)\}$

Proof of Lemma 5

The first-order derivatives of $\mathcal{U}$ with respect to λ and λ_ℓ are straightforward by noting that

$$x(w) = \lambda \int_\ell^w F_p^{-1}(1 - F_w(t))\, dt + \lambda_\ell.$$

Then the second order derivatives are given by

$$\begin{aligned}
\frac{d^2\mathcal{U}}{d\lambda^2} &= -\frac{1}{\lambda^2}\int_0^\infty f_w(w)[x(w) - \lambda_\ell]^2 \frac{\partial[u'h']}{\partial x} dw < 0; \\
\frac{d^2\mathcal{U}}{d\lambda d\lambda_\ell} &= -\frac{1}{\lambda}\int_0^\infty f_w(w)[x(w) - \lambda_\ell] \frac{\partial[u'h']}{\partial x} dw; \\
\frac{d^2\mathcal{U}}{d\lambda_\ell^2} &= -\int_0^\infty f_w(w) \frac{\partial[u'h']}{\partial x} dw < 0,
\end{aligned}$$

where the negativeness of the second derivatives is due to the fact that

$$\frac{\partial[u'h']}{\partial x} = -u''[h']^2 + u'h'' > 0,$$

because both u and v are concave functions. A direct calculation shows

$$\frac{d^2\mathcal{U}}{d\lambda^2}\frac{d^2\mathcal{U}}{d\lambda_\ell^2} - \left[\frac{d^2\mathcal{U}}{d\lambda d\lambda_\ell}\right]^2 > 0.$$

Since this holds for any (λ, λ_ℓ), the solution to the first-order condition maximizes the objective globally and is unique.

Proof of Proposition 1

Let $\mathcal{U}^\epsilon$ denote the perturbed $\mathcal{U}$, that is

$$\mathcal{U}^\epsilon = \int_0^\infty [u + \lambda_v x^\epsilon + \lambda_f(w)] f_w^\epsilon \, dw - \lambda_w E^\epsilon[pw] + \lambda_w w_0 - \lambda_v v_0.$$

Then the first-order derivative of $\mathcal{U}^\epsilon$ is

$$\begin{aligned}\frac{d\mathcal{U}^\epsilon}{d\epsilon} &= \int_0^\infty \eta(w)[u + \lambda_v x^\epsilon + \lambda_f(w)] \, dw \\ &\quad + \int_0^\infty f_w^\epsilon \frac{d[u + \lambda_v x^\epsilon]}{d\epsilon} \, dw - \lambda_w \frac{dE^\epsilon[pw]}{d\epsilon}. \qquad (54)\end{aligned}$$

First note that, using (13) or Lemma 2, we have

$$\frac{dx^\epsilon(w)}{d\epsilon} = -\lambda^\epsilon \int_\ell^w \frac{1}{f_p} \int_0^t \eta(s) \, ds \, dt + \frac{x^\epsilon(w) - \lambda_\ell^\epsilon}{\lambda^\epsilon} \frac{d\lambda^\epsilon}{d\epsilon} + \frac{d\lambda_\ell^\epsilon}{d\epsilon}.$$

Then,

$$\begin{aligned}\int_0^\infty f_w^\epsilon(w) \frac{d[u + \lambda_v x^\epsilon]}{d\epsilon} \, dw \;&=\; -\int_0^\infty [u'h' - \lambda_v] f_w^\epsilon \frac{dx^\epsilon(w)}{d\epsilon} \, dw \\ = \;\; \lambda^\epsilon \int_0^\infty [u'h' - \lambda_v] f_w^\epsilon(w) &\int_\ell^w \frac{1}{f_p} \int_0^t \eta(s) \, ds \, dt \, dw \\ - \frac{d\lambda^\epsilon}{d\epsilon} \int_0^\infty [u'h' - \lambda_v] f_w^\epsilon &\frac{x - \lambda_\ell^\epsilon}{\lambda^\epsilon} dw - \frac{d\lambda_\ell^\epsilon}{d\epsilon} \int_0^\infty [u'h' - \lambda_v] f_w^\epsilon \, dw,\end{aligned}$$

where we use the fact that the derivatives of λ^ϵ and λ_ℓ^ϵ do not depend on w. Note that the last two terms are equal to zero due to the first-order conditions of Lemma

5. Integration by parts shows the remaining integral is

$$
\begin{aligned}
&\int_0^\infty [u'h' - \lambda_v] f_w^\epsilon(w) \int_\ell^w \frac{1}{f_p} \int_0^t \eta(s)\, ds\, dt\, dw \\
&\quad = \quad \int_0^w [u'h' - \lambda_v] f_w^\epsilon(t)\, dt \int_\ell^w \frac{1}{f_p} \int_0^t \eta(s)\, ds\, dt \Big|_0^\infty \\
&\qquad\qquad\qquad - \int_0^\infty \frac{1}{f_p} \int_0^w \eta(t)\, dt \left[\int_0^w [u'h' - \lambda_v] f_w^\epsilon\, dt \right] dw \\
&\quad = \quad \int_0^\infty \eta(w) \int_0^w \frac{1}{f_p} \left[\int_0^t [u'h' - \lambda_v] f_w^\epsilon\, ds\, dt \right] dw,
\end{aligned}
$$

where we have use the first-order condition for λ_ℓ^ϵ, equation (17). Using this equation and Lemma A.1, the first derivative of $\mathcal{U}^\epsilon$ with respect to ϵ is:

$$
\begin{aligned}
\frac{d\mathcal{U}^\epsilon}{d\epsilon} = &\int_0^\infty \eta(w) \left[u + \lambda_v x^\epsilon + \lambda_f + \lambda^\epsilon \int_0^w \frac{1}{f_p} \int_0^t [u'h' - \lambda_v] f_w^\epsilon\, ds\, dt \right] dw \\
&- \lambda_w \int_0^\infty \eta(w) \int_\ell^w F_p^{-1}(1 - F_w^\epsilon(t))\, dt\, dw,
\end{aligned} \tag{55}
$$

where we have used Lemma A.1. However, when $\epsilon = 0$, we have $\lambda^\epsilon = \lambda$, $\lambda_\ell^\epsilon = \lambda_\ell$, and hence

$$
\begin{aligned}
\frac{d\mathcal{U}^\epsilon}{d\epsilon}\Big|_{\epsilon=0} = &\int_0^\infty \eta(w) \left[u + \lambda_v x + \lambda_f + \lambda \int_\ell^w \frac{1}{f_p} \int_0^t [u'h' - \lambda_v] f_w\, ds\, dt \right] dw \\
&- \lambda_w \int_0^\infty \eta(w) \int_\ell^w F_p^{-1}(1 - F_w(t))\, dt\, dw \\
= &\int_0^\infty \eta(w) \left[u + \lambda_v x + \lambda_f + \lambda \int_\ell^w \frac{1}{f_p} \int_0^t [u'h' - \lambda_v] f_w\, ds\, dt \right. \\
&\left. - \frac{\lambda_w}{\lambda} x + \frac{\lambda_w}{\lambda} \lambda_\ell \right] dw = 0,
\end{aligned}
$$

where the second last equality is obtained by using the constraints (13) or Lemma 2. The above integral equals zero for any arbitrary η such that $\int_0^\infty \eta(w)\, dw = 0$ if and only if there exists a constant C such that

$$
\begin{aligned}
&u(\alpha(w) - h(x(w))) + \lambda_v x(w) - \frac{\lambda_w}{\lambda} x(w) + \lambda_f(w) \\
&\quad + \lambda \int_\ell^w \frac{1}{f_p} \int_0^t [u'h' - \lambda_v] f_w\, ds\, dt - \frac{\lambda_w}{\lambda} x + \frac{\lambda_w}{\lambda} \lambda_\ell = C
\end{aligned} \tag{56}
$$

holds almost everywhere. However, $f_w(\ell) > 0$ by assumption and $\lambda_\ell = x(\ell)$, equation (56) then implies that $C = u(\alpha(\ell) - h(x(\ell)) + \lambda_v x(\ell)$.

Proof of Proposition 2

First note that $x(w)$ is continuous on R^+. Let

$$L(w) \equiv u(\alpha(w) - h(x)) + \left[\lambda_v - \frac{\lambda_w}{\lambda}\right] x + \lambda \int_\ell^w \frac{1}{f_p} \int_0^t [u'h' - \lambda_v] f_w \, ds \, dt,$$

which is also continuous. By the necessary condition (56) $\lambda_f(w) + L(w) \equiv C$ we then have $\lambda_f(w)$ is continuous. Therefore $\{w|\lambda_f(w) > 0\}$ is open intervals. Let $S = (w_1, w_2)$ be such an open interval. Because $\lambda_f(w_1) = \lambda_f(w_2) = 0$ we have $L(w_1) = L(w_2) = C$. Since $f_w \equiv 0$ in S, we have $F_p^{-1}(1 - F_w(w)) = F_p^{-1}(1 - F_w(w_1)) = p_1$ hence

$$L(w) = u(\alpha(w) - h(x(w))) + \left[\lambda_v - \frac{\lambda_w}{\lambda}\right] [\lambda p_1(w - \ell) + \lambda_\ell] + \lambda A(w - w_1),$$

where

$$A = \frac{1}{f_p(F_p^{-1}(1 - F_w(w_1)))} \int_0^{w_1} [u'h' - \lambda_v] f_w(w) \, dw.$$

Rearranging the linear terms, we have

$$L(w) = u(\alpha(w) - h(x(w))) + a(w - w_1) + b,$$

where a and b are constants. Then by $L(w_1) = L(w_2) = C$, we have

$$b = C - u(\alpha(w_1) - h(x(w_1)))$$

and

$$a = -\frac{u(\alpha(w_2) - h(x(w_2))) - u(\alpha(w_1) - h(x(w_1)))}{w_2 - w_1}.$$

Therefore, we have by $L(w) + \lambda_f(w) = C$ and $\lambda_f(w) > 0$

$$\begin{aligned} u(\alpha(w) - h(x(w))) \quad < \quad & \frac{w_2 - w}{w_2 - w_1} u(\alpha(w_1) - h(x(w_1))) \\ & + \frac{w - w_1}{w_2 - w_1} u(\alpha(w_2) - h(x(w_2))) \end{aligned}$$

for all $w \in (w_1, w_2)$.

If $\alpha(w)$ is twice differentiable and $f_w(w) = 0$, then $x(w)$ and $L(w)$ is twice differentiable. Therefore, $\lambda_f(w)$ is twice differentiable on S, and its second derivative is given by

$$\lambda_f''(w) = -u''[\alpha' - h'x']^2 + u'h''[x']^2 - u'\alpha'',$$

where we have used $x''(w) = f_w(w) = 0$ for all $w \in S$. However, the first part of the proposition states that $\lambda_f(w) > 0$ can only occur in a nonconcave region of $u(\alpha(w) - h(x(w)))$ hence $L(w)$ is concave on a subset of S. Then that $\lambda_f(w) + L(w) = C$ implies that $\lambda_f(w)$ is convex on this subset, that is $\lambda_f(w) > 0$. Hence (19) is a necessary condition for $\lambda_f(w) > 0$

Proof of Proposition 3

Taking derivatives with respect to w to the first-order condition (18) implies that

$$\left[u'\alpha' - \left(u'h' - \lambda_v + \frac{\lambda_w}{\lambda}\right)x'\right]f_p + \lambda\int_0^w [u'h' - \lambda_v]f_w(t)\,dt = 0.$$

This equation can be simplified further by taking the derivative with respect to w one more time and rearranging terms on the two sides. Using the fact that

$$f_w(w) = -\frac{1}{\lambda}f_p x'' \text{ and } \frac{df_p}{dw} = -\frac{f_p'}{f_p}f_w = \frac{1}{\lambda}f_p' x'',$$

we have

$$\frac{d}{dw}\left[u'\alpha' - \left(u'h' - \lambda_v + \frac{\lambda_w}{\lambda}\right)x'\right]f_p + \left[u'\alpha' - \left(u'h' - \lambda_v + \frac{\lambda_w}{\lambda}\right)x'\right]\frac{x''}{\lambda}f_p' - [u'h' - \lambda_v]x''f_p = 0,$$

or

$$\left\{\left[u'\alpha' - \left(u'h' - \lambda_v + \frac{\lambda_w}{\lambda}\right)x'\right]\frac{f_p'}{\lambda f_p} - \left[2[u'h' - \lambda_v] + \frac{\lambda_w}{\lambda}\right]\right\}x'' + \left[\frac{d[u'\alpha']}{dw} - x'\frac{d[u'h']}{dw}\right] = 0. \qquad (57)$$

$$\left[\beta - \frac{u'}{v'}\right]x'' + u''(\alpha - y')^2 - u'y'' + \frac{x''}{\lambda}\left[u'(\alpha - y') + (\beta - \lambda_v)x'\right]\frac{f_p'}{f_p} = 0, \qquad (58)$$

where $\beta = -\lambda_w/\lambda + 2\lambda_v$. Note that $y = h(x)$ or $x = v(y)$, thus

$$y' = h'(x)x' \text{ and } y'' = h'(x)x'' + h''(x)[x']^2,$$

or put another way,

$$x' = v'(y)y' \text{ and } x'' = v'(y)y'' + v''(y)[y']^2.$$

Thus, equation (58) can be written as a second-order nonlinear ODE with respect to $y\,(w)$. That is

$$\begin{aligned}[t] &[\beta v' - 2u']y'' + \frac{v''}{v'}(\beta v' - u')[y']^2 + u''[\alpha' - y']^2 + u'\alpha'' \\ &\qquad + \left[v'y'' + v''(y')^2\right][u'[\alpha' - y'] + (\beta - \lambda_v)v'y']\frac{f_p'}{\lambda f_p} = 0. \end{aligned}$$

This equation also implies that y is twice differentiable if u, α and v are twice differentiable.

Proof of Proposition 4

This is a direct application of standard results: see Chapter 3 of Walter (1998). Specifically, given the constraint (24), the ODE (23) can be rewritten as

$$x'' = -\frac{[u''[h']^2 + u'h''][x']^2 - 2u''\alpha' h' x' + u''[\alpha']^2 + u'\alpha''}{\beta - 2u'h' + (u'[\alpha' - h'x'] + (\beta - \lambda_v)x')\frac{f_p'}{\lambda f_p}}.$$

Then, letting $z = x'$ and $x'' = z'$ transforms this ODE into a system of two first-order ODEs, which have a unique solution given initial conditions that satisfy (24) by the fact that all of the coefficients are continuous. Or in the case when $u(\alpha(w) - h(x(w)))$ is piece-wise smooth (countable kinks), multiple sets of initial conditions can derived by using equation (18).

Proof of Proposition 5

Using the first-order derivative (54), the second derivative of $\mathcal{U}^\epsilon$ with respect to ϵ, then, is given by

$$\frac{d^2\mathcal{U}^\epsilon}{d\epsilon^2} = 2\int_0^\infty \eta \frac{d[u + \lambda_v x^\epsilon]}{d\epsilon}\,dw + \int_0^\infty f_w^\epsilon \frac{d^2[u + \lambda_v x^\epsilon]}{d\epsilon^2}\,dw - \lambda_w \frac{d^2 E^\epsilon[pw]}{d\epsilon^2}.$$

To simplify the notation, define

$$B(w) \equiv \int_0^w \eta(t)\,dt$$

in the remaining proof. Direct calculations show that

$$\frac{dx^\epsilon(w)}{d\epsilon} = -\lambda^\epsilon \int_\ell^w \frac{B}{f_p}\,dt + \frac{x^\epsilon(w) - \lambda_\ell^\epsilon}{\lambda^\epsilon}\frac{d\lambda^\epsilon}{d\epsilon} + \frac{d\lambda_\ell^\epsilon}{d\epsilon}, \tag{59}$$

$$\frac{d^2x^\epsilon(w)}{d\epsilon^2} = -\lambda^\epsilon \int_\ell^w \frac{f_p'}{f_p^3}B^2\,dt - 2\frac{d\lambda^\epsilon}{d\epsilon}\int_\ell^w \frac{B}{f_p}\,dt + \frac{x^\epsilon - \lambda_\ell^\epsilon}{\lambda^\epsilon}\frac{d^2\lambda^\epsilon}{d\epsilon^2} + \frac{d^2\lambda_\ell^\epsilon}{d\epsilon^2}, \tag{60}$$

where the derivatives of λ^ϵ and λ_ℓ^ϵ do not depend on w. Differentiating the first-order conditions for λ^ϵ and λ_ℓ^ϵ, equations (16) and (17), with respect to ϵ, yield

$$\begin{aligned} &\int_0^\infty \eta[u'h' - \lambda_v]\frac{x^\epsilon - \lambda_\ell^\epsilon}{\lambda^\epsilon}\,dw \\ = \; &\int_0^\infty f_w^\epsilon[u'h' - \lambda_v]\int_\ell^w \frac{B}{f_p}dt\,dw - \int_0^\infty f_w^\epsilon \frac{x^\epsilon - \lambda_\ell^\epsilon}{\lambda^\epsilon}\frac{\partial[u'h']}{\partial x}\,\frac{dx^\epsilon}{d\epsilon}\,dw, \end{aligned} \tag{61}$$

and

$$\int_0^\infty \eta[u'h' - \lambda_v]\,dw = -\int_0^\infty f_w^\epsilon \frac{\partial[u'h']}{\partial x}\frac{dx^\epsilon}{d\epsilon}\,dw, \tag{62}$$

respectively. These two equations further imply the following results, which are crucial to proving the second-order necessary condition.

Lemma B.3 *The first-order derivatives of λ^ϵ and λ_ℓ^ϵ with respect to ϵ have to vanish at $\epsilon = 0$.*

Proof. Substituting (59) into equation (62) yields

$$\begin{aligned} &\int_0^\infty \eta[u'h' - \lambda_v]\,dw - \lambda^\epsilon \int_0^\infty f_w^\epsilon \frac{\partial[u'h']}{\partial x}\int_\ell^w \frac{1}{f_p}\int_0^t \eta(s)\,ds\,dt\,dw \\ &\quad = \frac{d\lambda^\epsilon}{d\epsilon}\int_0^\infty f_w^\epsilon \frac{\partial[u'h']}{\partial x}\,\frac{x^\epsilon(w) - \lambda_\ell^\epsilon}{\lambda^\epsilon}dw + \frac{d\lambda_\ell^\epsilon}{d\epsilon}\int_0^\infty f_w^\epsilon \frac{\partial[u'h']}{\partial x}dw. \end{aligned}$$

As $\epsilon \to 0$, the righthand side of the equation above does not depend on η but the lefthand side does. Both sides of the equation have to be the same constant that is independent of the choice of η. However, this constant has to equal zero because, as $\eta(w) \to 0$ for all w, the lefthand side also goes to zero. Thus, we have

$$\left.\frac{d\lambda^\epsilon}{d\epsilon}\right|_{\epsilon=0} \int_0^\infty f_w \frac{\partial[u'h']}{\partial x} \frac{x-\lambda_\ell}{\lambda} dw + \left.\frac{d\lambda_\ell^\epsilon}{d\epsilon}\right|_{\epsilon=0} \int_0^\infty f_w \frac{\partial[u'h']}{\partial x} dw = 0.$$

The same argument also applies to equation (61) and yields

$$\left.\frac{d\lambda^\epsilon}{d\epsilon}\right|_{\epsilon=0} \int_0^\infty f_w \frac{\partial[u'h']}{\partial x} \left[\frac{x-\lambda_\ell}{\lambda}\right]^2 dw + \left.\frac{d\lambda_\ell^\epsilon}{d\epsilon}\right|_{\epsilon=0} \int_0^\infty f_w \frac{\partial[u'h']}{\partial x} \frac{x-\lambda_\ell}{\lambda} dw = 0.$$

This system of two equations with two unknowns has a unique solution, that is

$$\left.\frac{d\lambda^\epsilon}{d\epsilon}\right|_{\epsilon=0} = \left.\frac{d\lambda_\ell^\epsilon}{d\epsilon}\right|_{\epsilon=0} = 0,$$

because the determent of the matrix that is formed by the coefficients of the system of equations is not equal to zero. This completes the proof of the lemma. $\square$

Define

$$C(w) \equiv \frac{x^\epsilon(w) - \lambda_\ell^\epsilon}{\lambda^\epsilon} \frac{d\lambda^\epsilon}{d\epsilon}.$$

An application of the integration by parts, using (59) and the assumption that $B(0) = B(\infty) = 0$, shows

$$\begin{aligned}
\int_0^\infty \eta \frac{d[u + \lambda_v x^\epsilon]}{d\epsilon}\, dw \;&=\; -\int_0^\infty \eta[u'h' - \lambda_v] \frac{dx^\epsilon}{d\epsilon}\, dw \\
&= \int_0^\infty B(w) \frac{d[u'h']}{dw} \left[\frac{dx^\epsilon}{d\epsilon} - C(w)\right] dw - \lambda^\epsilon \int_0^\infty [u'h' - \lambda_v] \frac{B^2(w)}{f_p}\, dw \\
&\quad - \frac{d\lambda^\epsilon}{d\epsilon} \int_0^\infty \eta[u'h' - \lambda_v] \frac{x^\epsilon - \lambda_\ell^\epsilon}{\lambda^\epsilon}\, dw \\
&= \int_0^\infty B(w) \frac{d[u'h']}{dw} \left[\frac{dx^\epsilon}{d\epsilon} - C(w)\right] dw - \lambda^\epsilon \int_0^\infty [u'h' - \lambda_v] \frac{B^2(w)}{f_p}\, dw \\
&\quad + \int_0^\infty f_w^\epsilon C(w) \frac{\partial[u'h']}{\partial x} \frac{dx^\epsilon}{d\epsilon}\, dw - \frac{d\lambda^\epsilon}{d\epsilon} \int_0^\infty f_w^\epsilon [u'h' - \lambda_v] \int_\ell^w \frac{B}{f_p} dt\, dw,
\end{aligned}$$

where the last equality is obtained by equation (61).

Then, using Lemma A.1, we have

$$
\begin{aligned}
& 2\int_0^\infty \eta(w)\frac{d[u+\lambda_v x^\epsilon]}{d\epsilon}\,dw - \lambda_w \frac{d^2 E^\epsilon[pw]}{d\epsilon^2} \qquad (63)\\
= \; & 2\int_0^\infty \frac{d[u'h']}{dw} B(w)\left[\frac{dx^\epsilon}{d\epsilon} - C(w)\right] dw - \int_0^\infty \left[2\lambda^\epsilon[u'h'-\lambda_v] + \lambda_w\right]\frac{B^2}{f_p}\,dw \\
& + 2\int_0^\infty f_w^\epsilon C(w)\frac{\partial[u'h']}{\partial x}\frac{dx^\epsilon}{d\epsilon}\,dw - 2\frac{d\lambda^\epsilon}{d\epsilon}\int_0^\infty f_w^\epsilon[u'h'-\lambda_v]\int_\ell^w \frac{B(t)}{f_p}dt\,dw.
\end{aligned}
$$

For the second order term, because

$$
\frac{d^2[u+\lambda_v x^\epsilon]}{d\epsilon^2} = -\frac{\partial[u'h']}{\partial x}\left[\frac{dx^\epsilon}{d\epsilon}\right]^2 - [u'h'-\lambda_v]\frac{d^2x^\epsilon}{d\epsilon^2},
$$

therefore,

$$
\begin{aligned}
\int_0^\infty f_w^\epsilon \frac{d^2[u+\lambda_v x^\epsilon]}{d\epsilon^2}\,dw \; &= \; 2\frac{d\lambda^\epsilon}{d\epsilon}\int_0^\infty f_w^\epsilon[u'h'-\lambda_v]\int_\ell^w \frac{B}{f_p}\,dt\,dw \\
& - \int_0^\infty f_w^\epsilon\frac{\partial[u'h']}{\partial x}\left[\frac{dx^\epsilon}{d\epsilon}\right]^2 dw + \lambda^\epsilon\int_0^\infty f_w^\epsilon[u'h'-\lambda_v]\int_\ell^w \frac{f_p'}{f_p^3}B^2\,dt\,dw,
\end{aligned}
$$

where the two terms with the second derivatives of λ^ϵ and λ_ℓ^ϵ, which are independent of w, vanish because of the first-order conditions (16) and (17). Integration by parts shows the last term of the equation above is

$$
\int_0^\infty f_w^\epsilon[u'h'-\lambda_v]\int_\ell^w \frac{f_p'}{f_p^3}B^2\,dt\,dw = -\int_0^\infty \frac{f_p'}{f_p^3}B^2\int_0^w f_w^\epsilon[u'h'-\lambda_v]\,dt\,dw.
$$

Therefore,

$$
\begin{aligned}
\int_0^\infty f_w^\epsilon \frac{d^2[u+\lambda_v x]}{d\epsilon^2}\,dw \; &= \; 2\frac{d\lambda^\epsilon}{d\epsilon}\int_0^\infty f_w^\epsilon[u'h'-\lambda_v]\int_\ell^w \frac{B}{f_p}\,dt\,dw \qquad (64)\\
& - \int_0^\infty f_w^\epsilon\frac{\partial[u'h']}{\partial x}\left[\frac{dx^\epsilon}{d\epsilon}\right]^2 dw - \int_0^\infty \frac{f_p'}{f_p^3}B^2\int_0^w f_w^\epsilon[u'h'-\lambda_v]\,dt\,dw.
\end{aligned}
$$

Let

$$
A(w) = \frac{dx^\epsilon}{d\epsilon} - C(w).
$$

Combining (63) and (64) yields

$$\begin{aligned}\frac{d^2\mathcal{U}^\epsilon}{d\epsilon^2} &= -\int_0^\infty f_w^\epsilon \frac{\partial[u'h']}{\partial x}\left[A^2 - C^2\right]\,dw + 2\int_0^\infty \frac{d[u'h']}{dw}AB\,dw \\ &\quad - \int_0^\infty \left[2\lambda^\epsilon[u'h' - \lambda_v] + \lambda_w\right]\frac{B^2}{f_p}\,dw - \int_0^\infty \frac{f_p'}{f_p^3}B^2\int_0^w f_w^\epsilon[u'h' - \lambda_v]\,dt\,dw.\end{aligned}$$

By Lemma B.3, $\frac{d\lambda^\epsilon}{d\epsilon}|_{\epsilon=0} = 0$, hence $C(w) \equiv 0$ at $\epsilon = 0$. Moreover, at $\epsilon = 0$

$$-\lambda\int_0^\infty B^2\frac{f_p'}{f_p^3}\int_0^w f_w^\epsilon[u'h' - \lambda_v]\,dt\,dw = \int_0^\infty \frac{f_p'}{f_p^2}\frac{d[u + (\lambda_v - \lambda_w/\lambda)x]}{dw}B^2\,dw,$$

where we have used equation (20), and

$$\begin{aligned}&-\int_0^\infty \left[2\lambda[u'h' - \lambda_v] + \lambda_w\right]\frac{1}{f_p}B^2(w)\,dw \\ &\quad = -\int_0^\infty \frac{f_p'}{f_p^2}\frac{d[u + (\lambda_v - \lambda_w/\lambda)x]}{dw}B^2(w)\,dw \\ &\qquad\qquad - \int_0^\infty \left[u''[\alpha - h'x']^2 - u'h''[x']^2 + u'\alpha''\right]\frac{\lambda}{x''f_p}B^2(w)\,dw,\end{aligned}$$

where we have used (21) or (23). Combining these results yields

$$\begin{aligned}\left.\frac{d^2\mathcal{U}^\epsilon}{d\epsilon^2}\right|_{\epsilon=0} = -\int_0^\infty \Bigg[& f_w\frac{\partial[u'h']}{\partial x}A^2(w) + 2\frac{d[u'h']}{dw}A(w)B(w) \\ &- \left[u''[\alpha' - h'x']^2 - u'h''[x']^2 + u'\alpha''\right]\frac{1}{f_w}B^2(w)\Bigg]\,dw,\end{aligned} \tag{65}$$

where we have used the fact that $f_w = -f_p x''/\lambda$. This second derivative is negative if the integrand is always positive. As $f_w > 0$, this is true for the following reasons. First

$$\frac{\partial[u'h']}{\partial x} = -u''[h']^2 + u'h'' > 0$$

is true because u is concave and h is convex. Secondly, because $u(\alpha(w) - h(x))$ is concave in (w, x), we have

$$-\frac{\partial[u'h']}{\partial x}\frac{\partial[u'\alpha']}{\partial w} - [\alpha'u''h']^2 > 0,$$

which implies

$$-u'\alpha''\frac{\partial[u'h']}{\partial x} > [\alpha']^2u'u''h''.$$

This inequality shows that

$$\begin{aligned}
&-\left[u''[\alpha'-h'x']^2 - u'h''[x']^2 + u'\alpha''\right]\frac{\partial[u'h']}{\partial x} \\
&\quad > -\left[\alpha'[\alpha'-2h'x']u'' + \left[u''[h']^2 - u'h''\right][x']^2\right]\frac{\partial[u'h']}{\partial x} + [\alpha']^2u'u''h'' \\
&\quad = [\alpha'u''h']^2 + 2\alpha'h'x'u''\frac{\partial[u'h']}{\partial x} + \left[\frac{\partial[u'h']}{\partial x}\right]^2[x']^2 \\
&\quad = \left[\alpha'u''h' + \frac{\partial[u'h']}{\partial x}x'\right]^2 > 0.
\end{aligned}$$

This shows that the coefficient of B^2 in equation (65) is positive and further implies,

$$\begin{aligned}
&\left[\frac{d[u'h']}{dw}\right]^2 + \frac{\partial[u'h']}{\partial x}\left[u''[\alpha'-h'x']^2 - u'h''[x']^2 + u'\alpha''\right] \\
&\qquad < [u''[\alpha'-h'x']h' + u'h''x']^2 - \left[\alpha'u''h' + \frac{\partial[u'h']}{\partial x}x'\right]^2 \\
&\qquad = \left[\alpha'u''h' + \frac{\partial[u'h']}{\partial x}x'\right]^2 - \left[\alpha'u''h' + \frac{\partial[u'h']}{\partial x}x'\right]^2 \\
&\qquad = 0.
\end{aligned}$$

These facts imply that the second derivative of $\mathcal{U}^\epsilon$ as defined in equation (65) is negative. This shows that the solutions which satisfy the the first-order condition is local maximum.

Proof of Theorem 1

This is the direct implication of Propositions 1-5 and the fact that the differential-integral equation (18) is equivalent to the set of equations (20) and (23).

Proof of Theorem 2

First note that $x(w) = v(y(w))$ contains a linear segment if a pay schedule y contains an option-like structure or is nondecreasing. Thus, to examine whether the optimal pay schedule contains an option-like structure is to check whether there exists a region of w such that $f_w(w) \equiv 0$ on this region. However, as shown in the proof of Proposition 5, the necessary condition (19) for $\lambda_f \equiv 0$ is always satisfied when $u(\alpha(w) - h(x(w))$ is strictly concave in (w, x). This and smoothness of $\alpha(w)$ imply that $f_w(w) = 0$ can only occur on a zero measure set. Therefore, the optimal pay contains no option-lie features for this case. The monotonicity of the pay schedule follows by the observation that a decreasing part of the optimal pay schedule can only occur on $\{w|\lambda_f(w) > 0\}$.

Proof of Lemma 6

Let $F(w, y)$ is the joint cumulative distribution function for given $w(\omega) \in A$ and $y(\omega)$, we then have

$$\begin{aligned}
\int_\Omega u(\alpha(w(\omega)) - y(\omega))\, P(d\omega) &= \int u(\alpha(w) - y)\, dF(w, y) \\
&= \int dF_w(w) \int u(\alpha(w) - y)\, dF(w, y|w) \\
&\leq \int u(\alpha(w)) - E[y|w])\, dF_w(w),
\end{aligned}$$

the last inequality is due to Jensen's inequality and becomes equality if

$$y(\omega) = y(w) \text{ for all } \omega \in \{\omega' |(w(\omega') = w\}.$$

This means the optimal pay takes the form of $y(w)$.

Proof of Proposition 6

Because x or ($y(w)$) is a free choice function, the maximization problem is quite straightforward. It is similar to the case of agent's problem for the f_w. The second order is automatically satisfied by the budget constraint. And for the case of x, the second order is due to that fact that $u(\alpha(w) - h(x))$ is a concave function of x for any fixed w. We skip the details of the calculations.

References

Admati, A. R., and P. Pfleiderer, 1997, "Does It All Add Up? Benchmarks and the Compensation of Active Portfolio Managers," *Journal of Business*, 70(3), 323–350.

Allen, F., 2001, "Do Financial Institutions Matter?" *Journal of Finance*, 56, 1165–1175.

Allen, F., and A. M. Santomero, 1997, "The Theory of Financial Intermediation," *Journal of Banking and Finance*, 21, 1461–1485.

Aumann, R. J., and M. Perles, 1965, "A Variational Problem Arising in Economics," *Journal of Mathematical Analysis and Applications*, 11, 488–503.

Basak, S., A. Pavlova, and A. Shapiro, 2003, "Offsetting the Incentives: Risk Shifting and Benefits of Benchmarking in Money Management," Working paper 4304-03, MIT Sloan School of Management.

Bhattacharya, S., and P. Pfleiderer, 1985, "Delegated Portfolio Management," *Journal of Economic Theory*, 36(1), 1–25.

Brennan, M. J., 1993, "Agency and Asset Pricing," Working paper, UCLA.

Cadenillas, A., J. Cvitanic, and F. Zapatero, 2005, "Dynamic Principal-Agent Problems with Perfect Information," Working paper, University of Southern California.

Carpenter, J., 2000, "Does Option Compensation Increase Managerial Risk Appetite?" *Journal of Finance*, 55, 2311–2331.

Chevalier, J., and G. Ellison, 1997, "Risk Taking by Mutual Funds as a Response to Incentives," *Journal of Political Economy*, 105, 1167–1200.

Cox, J. C., and C. Huang, 1991, "A Variational Problem Arising in Financial Economics," *Journal of Mathematical Economics*, 20, 465–487.

Cuoco, D., and R. Kaniel, 2001, "Equilibrium Prices in the Presence of Delegated Portfolio Management," Working paper, University of Texas at Austin.

Das, S., and R. Sundaram, 1998, "On the Regulation of Mutual Fund Fee Structure," Working paper, Harvard University and New York University.

Dybvig, P. H., H. K. Farnsworth, and J. Carpenter, 2004, "Portfolio Performance and Agency," Working paper, Washington University in St. Louis.

Dybvig, P. H., and S. A. Ross, 1985, "Differential Information and Performance Measurement Using a Security Market Line," *Journal of Finance*, 40(2), 383–399.

Dybvig, P. H., and C. Spatt, 1986, "Agency and the Market for Mutual Fund Managers: The Principle of Preference Similarity," Working paper, Graduate School of Industrial Administration, Carnegie Mellon University.

Elton, E., M. Gruber, and C. Blake, 2003, "Incentive Fees and Mutual Funds," *Journal of Finance,* 58, 779–804.

Ewing, G., 1985, *Calculus of Variations with Applications*. Dover Publications, New York.

Garcia, D., 2001, "Optimal Menus of Linear Contracts in Portfolio Selection Models," Working paper, Tuck School of Business Administration.

Gervais, S., A. W. Lynch, and D. K. Musto, 2002, "Delegated Monitoring of Fund Managers: An Economic Rationale," Working paper, University of Pennsylvania.

Gómez, J.-P., and T. Sharma, 2001, "Portfolio Delegation under Short-selling Constraints," Working paper, Norwegian School of Management.

Grinblatt, M., and S. Titman, 1996, "Adverse Risk Incentives and the Design of Performance-Based Contracts," *Management Science*, 35, 807–822.

Grossman, S., and O. Hart, 1983, "An Analysis of the Principal-Agent Problem," *Econometrica,* 51, 7–45.

Gruber, M., 1996, "Another Puzzle: The Growth in Actively Managed Mutual Funds," *Journal of Finance,* 51, 783–810.

Heinkel, R., and N. M. Stoughton, 1994, "The Dynamics of Portfolio Management Contracts," *Review of Financial Studies,* 7(2), 351–387.

Holmström, B., 1978, "On Incentives and Control in Organizations," Ph.D. thesis, Stanford University.

——— , 1979, “Moral Hazard and Observability,” *Bell Journal of Economics*, 10, 74–91.

Holmström, B., and P. Milgrom, 1987, “Aggregation and Linearity in the Provision of Intertemporal Incentives,” *Econometrica*, 55(2), 303–28.

Kihlstrom, R. E., 1988, “Optimal Contracts for Security Analysts and Portfolio Managers,” *Studies in Banking and Finance*, 5, 291–325.

Kihlstrom, R. E., and S. A. Matthews, 1990, “Managerial Incentives in an Entrepeneurial Stock Market Model,” *Journal of Financial Intermediation*, 1, 57–79.

Li, T., and Y. Zhou, 2005, “Optimal Financial Contracts with Imperfect Information,” Working paper, The Chinese University of Hong Kong.

Meyers, D., and E. M. Rice, 1979, “Measuring Portfolio Performance and the Empirical Content of Asset Pricing Models,” *Journal of Financial Economics*, 7, 3–28.

Mirrlees, J. A., 1974, “Notes on Welfare Economics, Information, and Uncertainty,” in *Essays on Economic Behavior Under Uncertainty*, ed. by Balch, McFadden, and Wu. North Holland, Amsterdam.

——— , 1976, “The Optimal Structure of Incentives and Authority Within an Organization,” *Bell Journal of Economics*, 7, 105–131.

——— , 1999, “The Theory of Moral Hazard and Unobsrvable Behaviour: Part I,” *Review of Economic Studies*, 66, 3–21.

Mirrlees, J. A., and Y. Zhou, 2005a, “Principal-Agent Models Revisited Part I: Static Models,” Working paper, The Chinese University of Hong Kong, Hong Kong.

——— , 2005b, “Principal-Agent Models Revisited Part II: Dynamic Models,” Working paper, The Chinese University of Hong Kong, Hong Kong.

Ou-Yang, H., 2003, “Optimal Contracts in a Continuous-Time Delegated Portfolio Management Problem,” *Review of Financial Studies*, 16(1), 173–208.

Palomino, F., and A. Prat, 2003, “Risk Taking and Optimal Contracts for Money Managers,” *RAND Journal of Economics*, 34(1), 113–137.

Rogerson, W., 1985, "The First-Order Approach to Principal-Agent Problems," *Econometrica*, 53, 1357–1367.

Ross, S. A., 1973, "The Economic Theory of Agency: The Principal's Problem," *American Economic Review*, 63, 134–139.

——— , 1974, "On the Economic Theory of Agency and the Principal of Preference Similarity," in *Essays on Economic Behavior Under Uncertainty*, ed. by Balch, McFadden, and Wu. North Holland, Amsterdam.

——— , 1979, "Equilibrium and Agency – Inadmissable Agents in the Public Agency Problem," *American Economic Review*, 69, 308–312.

——— , 2004, "Compensation, Incentives, and the Duality of Risk Aversion and Riskiness," *Journal of Finance*, 59, 207–225.

Schättler, H., and J. Sung, 1993, "The First-Order Approach to the Continuous-Time Principal-Agent Problem with Expotential Utility," *Journal of Economic Theory*, 61, 331–371.

Shiryaev, A., 1995, *Probability*. Springer-verlag, New York, 2nd edn.

Sirri, E., and P. Tufano, 1998, "Costly Search and Mutual Fund Flows," *Journal of Finance*, 53, 1589–1622.

Stark, L., 1987, "Performance Incentive Fees: An Agency Theoretic Approach," *Journal of Financial and Quantitative Analysis*, 22, 17–32.

Stoughton, N. M., 1993, "Moral Hazard and the Portfolio Management Problem," *Journal of Finance*, 48(5), 2009–2028.

Sung, J., 1995, "Linearity with Project Selection and Controllable Diffusion Rate in Continuous-Time Principal-Agent Problems," *Rand Journal of Economics*, 26(4), 720–43.

——— , 2005, "Optimal Contracts under Moral Hazard and Adverse Selection: A Continuous-Time Approach," *Review of Financial Studies*, 18, 1021–1073.

Walter, W., 1998, *Ordinary Differential Equations*. Springer-Verlag, New York.

Wilson, R., 1968, "On the Theory of Syndicates," *Econometrica*, 36, 119–132.