In this paper we analyze the optimal contract for a portfolio manager who can exert effort to improve the quality of a private signal about future market prices. We assume complete markets over states distinguished by asset payoffs and place no restrictions on the form of the contract. We show that trading restrictions are essential because they prevent the manager from undoing the incentive effects of performance-based fees. We provide conditions under which simple benchmarking emerges as optimal compensation. Additional incentives to take risk are necessary when information can be manipulated or else the manager will understate information to offset the benchmarking. (JEL D82, G11)

The appropriate evaluation and compensation of portfolio managers is an ongoing topic of debate among practitioners and regulators. Although performance measurement and optimal managerial contracting are two sides of the same coin, the academic literature has largely considered the two questions separately. Typically, performance measurement has been studied in the context of models with realistic security returns, without consideration of the incentives created by the measure. Optimal contracting has been studied in information models with careful consideration of incentives but simplistic models of portfolio choice and security returns. This paper derives optimal contracts for portfolio managers in the tradition of agency theory but uses a rich model of security returns with full spanning of market states.

We suppose that the manager can exert effort to influence the quality of a private signal about future market prices. The investor’s problem is to find a
contract for the manager that provides incentives to exert effort and to use the signal in the investor’s interest while still sharing risk efficiently. We derive optimal contracts given a mixture assumption under which the joint density function of the manager’s signal and the market state depends affinely on the effort of the manager. Mostly we assume that both the investor and the manager have log utility, but we conduct some sensitivity analysis using power utility. In a first-best world, the manager’s effort is observable, and the optimal contract is a proportional sharing rule.

In a second-best world, the manager’s signal is observable but effort is not observable. To give the appropriate incentive to generate a high-quality signal, the optimal fee for the manager overexposes him to the optimal portfolio for the signal. This fee appears as a proportion of the managed portfolio plus a share of the excess return of the portfolio over a passive benchmark portfolio. Thus, we provide a framework in which the commonly observed practice of benchmarking emerges endogenously as part of the optimal contract.

By contrast, other authors have found that benchmarking provides no incentives for effort. For example, Stoughton (1993) examines affine and quadratic contracts in a two-asset world. He finds that affine contracts provide no incentives for effort.2 Admati and Pfleiderer (1997) present a similar result to Stoughton’s result for affine contracts.

The negative results on benchmarking in the previous literature arise from an assumption that the contracts do not restrict the portfolio choice of the manager.3 Restrictions on the manager’s portfolio choice are essential for incentive pay schemes to induce effort. Unrestricted trading may allow a manager to eliminate completely the incentive effects of the fee. For example, whatever leverage is created by the benchmarking in the fee could be undone by de-leveraging the underlying portfolio.

Our analysis shows that an optimal contract specifies not only the fee schedule for the agent but also a menu of allowable portfolio strategies.4 Actual investment guidelines include many portfolio restrictions, although not necessarily the ones predicted by the model. Common restrictions on asset allocations

---

2 Stoughton (1993) finds quadratic contracts provide some incentive but are not optimal due to their poor risk-sharing properties. The paper also claims that quadratic contracts approach the first-best in the limit as the investor becomes risk-neutral. Unfortunately, the result is uninterpretable due to an unfortunate choice of utility representation used to define convergence. Using a more reasonable sense of convergence measured by difference in certainty equivalent, the proof does not work. Intuitively, the problem with using small differences in utilities (instead of, say, small differences in certainty equivalent) is that the utility representation $U_B(W_B) = -\exp(-bW_B)$ used in that paper becomes very flat as risk aversion $b$ falls. For example, $U_B(W_B) \to -1$ and $U_B'(W_B) \to 0$ uniformly on bounded sets as $b \downarrow 0$.

3 Gómez and Sharma (2001) have shown that these nonincentive results disappear when a restriction on short-selling is imposed. Similarly, Basak, Pavlova, and Shapiro (2003) show that restricting the deviation from a benchmark can reduce the perverse incentives of an agent facing an ad hoc convex objective (motivated by performance-linked future business).

4 In Admati and Pfleiderer (1997), Proposition 5 does examine the effect of adding an affine portfolio restriction to the model. However, this restriction does not look like an optimal menu, nor does it seem similar to portfolio restrictions observed in practice.
include restrictions on the universe of assets and ranges for proportions in the various assets, while common restrictions for management within an asset class are limitations on market capitalization or style of stocks, such as growth or income, credit ratings or durations of bonds, restrictions on use of derivatives, maximum allocations to a stock or industry, and increasingly restrictions on portfolio risk measures such as duration, beta, or tracking error.\(^5\)

In a third-best world, neither the effort nor the signal is observable, and additional adjustments are necessary to induce the intended portfolio choice. With the second-best contract in a third-best world a manager could partially undo the leverage effect of the benchmarking by underreporting the signal and thus implementing a more conservative underlying portfolio position. Therefore, relative to the second-best, the third-best contract rewards the manager for reporting more extreme signals, or, in other words, for trading more aggressively on extreme information. This illustrates the limitations of the second-best contract. The failure of the second-best contract to discourage overly conservative strategies explains the concerns of practitioners about “closet indexers,” managers who collect active management fees but adopt passive strategies.

In summary, this paper makes three main contributions. First, it shows that trading restrictions are an essential part of an optimal contract because they prevent the manager from undoing the incentive effects of performance-based fees. Second, the paper delineates conditions under which simple linear benchmarking emerges endogenously as optimal compensation. Third, we show that additional incentives to take risk must be provided when information can be manipulated or else the manager will understate information, because to elicit effort, the benchmarking overexposes the manager to the optimal portfolio. In addition, by mapping out optimal contracts across a range of settings, our paper shows how the optimal compensation function depends on the nature of the information generation process. In particular, our empirical prediction is that the complexity and nonlinearity of the compensation function increases with the opacity of the information generation process.

Conceptually, this paper is very similar to Kihlstrom (1988). However, the model in that paper has only two market states and two signal states, so it does not admit nonaffine contracts. With only two signal states, there is also no way for a manager to deviate slightly from the desired investment policy. The only choices are to take the correct position or to take the opposite position from what the signal would suggest and, consequently, when there are only two signal states, the incentive to be overly conservative does not arise. In addition, the investor in the model of Kihlstrom (1988) is risk-neutral. This would imply that no optimal contract exists except that short sales are not allowed. This leads to a corner solution.

\(^5\) Almazan et al. (2001) document the prevalence of portfolio restrictions in contracts observed in the mutual fund industry.
This paper’s model is not a model of screening managers by ability as in Bhattacharya and Pfleiderer (1985), and in fact the manager’s ability is common knowledge from the outset. Rather, there is moral hazard in information production. We are skeptical of the ability of portfolio managers to predict their own performance, and we think managers who are the most optimistic about their own performance are naive and may not be good managers (see Baranchuk and Dybvig [2009] for a development of this idea). In a similar vein, absence of any information asymmetry at the outset distinguishes this paper from Garcia (2001), whose managers already know their signals at the time of contracting.

Zender (1988) shows that the Jensen measure is the optimal affine contract in a reduced-form model of a mean-variance world. The limitations of that paper are that the mapping from effort to return properties is a black box and that it is unclear what underlying model it is a reduced form for, or indeed whether the optimal contract in the reduced form is also optimal in the underlying model. Palomino and Prat (2003) have a more complex single-period reduced-form model with some unusual assumptions; for example, there is assumed to be an internal maximum of expected return as the risk level varies. Sung (1995) and Ou-Yang (2003) analyze continuous-time models, in which both the drift and diffusion coefficient can be controlled, and affine contracts arise optimally. As in Zender (1988), the portfolio choice is a reduced form, and it is not clear whether this is the reduced form for a reasonable underlying model.

Finally, a number of economic models are not models of portfolio management but share with our model the feature of having both adverse selection and moral hazard (see Laffont and Martimort 2002, Sections 7.1 and 7.2). Very close to our model is the model of delegated expertise of Demski and Sappington (1987), which shares our structure of moral hazard followed by adverse selection. In that model, an analyst exerts costly effort to obtain information. The main differences between that paper and the current paper are that their principal is risk-neutral and the sharing rule over the output is restricted to depend only on output and not on the action taken or the signal observed by the agent. Also, a literature on the “generalized agency problem,” starting with Myerson (1979), has the reverse timing of adverse selection followed by moral hazard. Some recent papers in this literature are Faynzilberg and Kumar (1997); Faynzilberg and Kumar (2000); and Sung (2003).

The paper proceeds as follows. Section 1 describes the optimal contracting problem. Section 2 presents analytical solutions in the first-best and second-best cases, and discusses the problems that arise in the third-best case. Section 3 provides numerical examples. Section 4 discusses empirical implications and Section 5 closes the paper.

---

6 The portfolio application is mentioned in Sung (1995) and spelled out in more detail in Sung’s thesis (Sung 1991).
1. The Agency Problem

We consider the contracting problem between an investor and a portfolio manager. Our analysis takes the approach of contracting theory and looks for an optimal contract without pre-supposing that the contract conforms to known institutions or has any specific form. The optimal contract derived in this way can be compared with practice or other contracts assumed by other analyses, understanding that an equivalent contract may take a somewhat different appearance.

There are a number of different technologies that can be used to manage of information problems. For example, information problems can be managed by using various forms of information gathering, such as evaluation before-the-fact and monitoring after-the-fact and, in a multiperiod context, by the impact of reputation on future business. Our analysis considers only what can be done using contracting and communication without these other technologies. We pose this in the typical format of an agency problem (as in Ross 1973), with allowance for a direct mechanism in the signal reporting stage. Here are the assumptions of the model.

**Market returns.** Investments are made in a market that is complete over states distinguished by security prices. Let \( \omega \in \Omega \) denote such a state and let \( p(\omega) \) be the pricing density for a claim that pays a dollar in state \( \omega \). This is a single-period model in the sense that payoffs will be realized only once, but we think of market completeness as being due to dynamic trading as in a Black-Scholes world. Our agents are “small” and we assume that their trades do not affect market prices.

**Information technology.** Through the costly effort \( \varepsilon \in [0, 1] \), the manager has the ability to generate information about the future market state in the form of a private signal \( s \in S \). Given effort \( \varepsilon \),

\[
f(s, \omega; \varepsilon) = \varepsilon f^I(s, \omega) + (1 - \varepsilon) f^U(s, \omega)
\]  

is the probability density of \( s \) and \( \omega \), where the market state is \( \omega \) and the signal is \( s \). Here, \( f^I \) is an “informed” distribution and \( f^U \) is an “uninformed” distribution. We assume that \( s \) and \( \omega \) are independent in the uninformed distribution, i.e., \( f^U(s, \omega) = f^s(s) f^\omega(\omega) \), the product of the marginal distributions. These marginal distributions are assumed to be the same as for the informed distribution. For \( \omega \), this must be true or else the manager’s effort choice could influence the market return. For \( s \), this is a normalization.

One interpretation of the mixture model is that the signal observed by the manager may be informative or it may be uninformative, and the manager cannot tell which. However, the manager knows that expending more effort makes it more likely an informative signal will be generated. Using the mixture model is without loss of generality if there are only two effort levels, and it is a simple sufficient condition for the first-order approach to work in many
agency models, including our second-best problem. Perhaps most importantly, using a mixture model avoids the pathological features of the more common assumption in finance that the agent chooses the precision of a signal jointly distributed with the outcome; with this common assumption, the unbounded likelihood ratio in the tails makes it too easy to create approximately first-best incentives using a limiting “forcing solution” of Mirrlees (1974). In the mixture model, likelihood ratios are bounded and the Mirrlees forcing solution no longer approaches first-best.

Preferences. Both the investor and the manager have logarithmic von Neumann-Morgenstern utility of end-of-period consumption, and the manager also bears a utility cost of expending effort. Specifically, the manager’s (agent’s) utility is \( \log(\phi) - c(\varepsilon) \), where \( \phi \) is the manager’s fee and \( c(\varepsilon) \) is the cost of the effort \( \varepsilon \) (the hidden action). We assume that \( c(\varepsilon) \) is differentiable and convex with \( c'(0) = 0 \). We assume that all the problems we consider have optimal solutions. The investor’s (principal’s) utility is \( \log(V) \), where \( V \) is the value of what remains in the portfolio after the fee has been paid. Following Grossman and Hart (1983), we use utility levels, rather than consumption levels, as the choice variables; this choice makes most of the constraints affine. Our results extend Grossman and Hart (1983) to incorporate a risk-averse principal: this is an important extension for portfolio problems. We denote by \( u_i(s, \omega) \) the investor’s equilibrium utility level \( \log(V) \) given \( s \) and \( \omega \), and we denote by \( u_m(s, \omega) \) the manager’s equilibrium utility level for only the wealth component \( \log(\phi) \) given \( s \) and \( \omega \).

Initial wealth and reservation utility. The investor’s initial wealth is \( w_0 \), and the manager does not have any initial wealth. The agency problem is formulated as maximizing the investor’s utility subject to giving the manager a reservation utility level of \( u_0 \). We interpret the reservation utility level as the best the manager can do in alternative employment. In an alternative interpretation, the reservation utility level would be a parameter mapping out the efficient frontier in a bargaining problem between the investor and the manager. Either way, the contracting problem is the same.

Optimal contracting. The contracting problem looks at mechanisms that work in this way. First, there is a contracting phase in which the investor offers a contract to the manager. The contract specifies a menu of portfolio strategies and rules for dividing the resulting payoff between investor and manager. In particular, the contract specifies one portfolio strategy-sharing rule pair for each

---


8 For the first-best problem with positive initial wealth, we have in essence a portfolio optimization, and an optimal solution exists under growth bounds on the tail probabilities of the state-price density and the asymptotic marginal utility, as in Cox and Huang (1991) or Dybvig, Rogers, and Back (1999). For the other problems, existence can fail in more subtle ways, for example, because compensating the manager enough to induce effort never leaves enough wealth left over to meet the investor’s minimum utility level. Or, there may be a closure problem in the second-best like the forcing solution described by Mirrlees (1974) in the first-best problems.
possible signal realization. The manager either accepts or refuses the contract; in our formal analysis, this is handled as a constraint that says the investor must choose a contract that the manager will be willing to accept. Once the contract is accepted, the manager chooses effort $\varepsilon$ and receives the private signal $s$. Then the manager chooses a portfolio-strategy and sharing-rule pair from the menu. Finally, portfolio returns are realized and the manager and the investor divide the portfolio value according to the specified sharing rule.

The investor must ensure that when the manager has observed the signal and is choosing from the menu of strategies, the manager always wants to choose the strategy that is intended for that signal realization. To formalize this constraint, we model the manager’s choice of trading strategy as a report of the signal, and we require that truthful reporting be incentive compatible. Imposing truthful reporting is without loss of generality. The revelation principle ensures that any equilibrium allocation implemented with false reporting can also be implemented with truthful reporting. The intuition for this is that in any equilibrium with false reporting, the principal will still know the reporting function his contract induces and will interpret reports accordingly, and so the outcome will be the same as if truthful reporting had been induced instead. In any equilibrium, the principal will structure the menu so that for each signal realization, the manager chooses the trading strategy and sharing rule intended for that signal, and the details of how the signal was reported will not matter. In this sense, the form of the contract is invariant to the reporting function the principal chooses to induce.\(^9\)

The merit of structuring the contract as a menu of allowable portfolio strategies and sharing rules is that the search for an optimum considers all feasible allocations. In particular, this approach imposes no restriction on the form of the sharing rule. A contract such as this can implement allocations that can be implemented using the sharing rules traditionally studied in the literature. But the optimal contract may do even better. The menu structure of the contract has an economic interpretation as a set of restrictions and guidelines on investment strategies that the investor permits the manager.

The search for an optimal contract is formalized as the solution of a choice problem that makes the investor as well off as possible subject to a budget constraint, the manager’s reservation utility level, and incentive-compatibility of the choices intended for the manager. We consider three forms of the problem. The first-best assumes that the manager’s choice of action and portfolio can be dictated, and as implied by the first theorem of welfare economics it is equivalent to a competitive allocation. The first-best seems unrealistic but it is a useful ideal benchmark and may approximate reality if the agency problem

---

\(^9\) Because of the private costly effort, our model does not conform to the traditional derivation of the revelation principle, in which there is private information but no private costly effort. Nonetheless, the revelation principle still works because there are no private actions chosen after the signal is reported (the portfolio choice is reasonably assumed to be public or at least publicly verifiable). Laffont and Martimort (2002), Section 7.2, discusses how the revelation principle is still valid even with initial costly effort.
we are concerned about is, for whatever reason, not so important in practice. The second-best requires the manager to want to choose the effort optimally but assumes that the use of the information signal in constructing the portfolio can be dictated. This is consistent with an assumption that there is monitoring of the process that ensures the information will be used as intended, or with an assumption that the incentives to misuse the information are handled another way; for example, through loss of business due to a reputation for being a “closet indexer” who collects fees as an active manager but actually chooses a portfolio close to the index. The third-best problem has the most profound difficulties with incentives and requires the manager to have the incentive to select the costly effort and also the incentive to reveal truthfully the observed signal. Whether the third-best is more realistic than the second-best is an empirical question.

First-best. Choose \( u_i(s, \omega), u_m(s, \omega), \) and \( \varepsilon \) to maximize investor’s expected utility,

\[
\int \int u_i(s, \omega)(\varepsilon f^I(\omega|s) + (1 - \varepsilon) f^{\omega}(\omega)) f^x(s) d\omega ds, \tag{2}
\]

subject to the budget constraint,

\[
(\forall s \in S) \int (\exp(u_i(s, \omega)) + \exp(u_m(s, \omega))) p(\omega) d\omega = w_0, \tag{3}
\]

and the participation constraint,

\[
\int \int u_m(s, \omega)(\varepsilon f^I(\omega|s) + (1 - \varepsilon) f^{\omega}(\omega)) f^x(s) d\omega ds - c(\varepsilon) = u_0. \tag{4}
\]

Second-best. Add the constraint for the incentive-compatibility of effort:

\[
\varepsilon = \arg \max_{\varepsilon'} \int \int u_m(s, \omega)(\varepsilon' f^I(\omega|s) + (1 - \varepsilon') f^{\omega}(\omega)) f^x(s) d\omega ds - c(\varepsilon'). \tag{5}
\]

Third-best. Instead of constraint (5), add the constraint for simultaneous incentive compatibility of effort and truthful signal reporting:

\[
\{\varepsilon, s\} = \arg \max_{\{\varepsilon', p(s)\}} \int \int u_m(\rho(s), \omega)(\varepsilon' f^I(\omega|s) + (1 - \varepsilon') f^{\omega}(\omega)) f^x(s) d\omega ds - c(\varepsilon'). \tag{6}
\]

In the choice problems, the choice variables are the effort level \( \varepsilon \) and the utility levels for investor and manager in each contingency \( (s, \omega) \). The objective function is expected utility for the investor as computed from the investor’s utility level in each contingency and the joint distribution of \( s \) and \( \omega \) given the
effort level $\epsilon$. The participation constraint says that the agent has to be treated well enough to meet the reservation utility level $u_0$ of outside opportunities.

The budget constraint computes the consumption levels for investor and manager from their utility levels and values them using the pricing rule $p(\omega)$. There is a separate budget constraint for each signal realization. This is because the contract specifies a different market-dependent portfolio payoff for each signal, and each of these payoffs must be affordable. In particular, the cost of each payoff cannot exceed the initial portfolio value $w_0$. The pricing rule $p(\omega)$ is the same for each $s$ because the signal is purely private and because we are making the “small investor” assumption that the manager does not affect pricing in security markets.

In the first-best, it is assumed that the effort and the dependence of the portfolio strategy on the signal can be dictated. In the second-best, there is an incentive-compatibility condition that says that the manager has an incentive to choose the intended effort $\epsilon$. In the third-best, the incentive compatibility condition says that the manager has incentive to choose the intended effort and to report the true state $s$, i.e., to select the intended payoff-pair.

Given a choice of effort, a realization of the signal, and a payoff for the manager, the investor’s optimal payoff is essentially just the solution to the standard investment problem. Thus, we can reduce both the number of choice variables and the number of constraints by use of the following lemma. The lemma enables us to eliminate the variables $u_i(s, \omega)$ and use as the objective the investor’s indirect utility, which equals the optimal value for the investor given the investor’s budget share, the effort level and the realization of the signal.

**Lemma 1.** In the solution to all the three forms of the investor’s problem, the expected utility conditional on $s$ for the investor is given by

$$\log \left( B_i(s) \frac{f^{\omega}(\omega)}{p(\omega)} + \epsilon(f^I(\omega|s) - f^{\omega}(\omega)) \right),$$

where

$$B_i(s) = w_0 - \int \exp(u_m(s, \omega))p(\omega)d\omega$$

is the investor’s budget share. Therefore, the indirect utility function can be substituted for the original objective in these problems.

**Proof.** Note that the choice of investor utilities $u_i(s, \omega)$ only appears in the problems in the objective function (2) and in the budget constraint (3). Therefore, the optimal solution must solve the subproblem of maximizing Equation (2) subject to Equation (3). The first-order condition of this problem is

$$[\epsilon f^I(\omega|s) + (1 - \epsilon)f^{\omega}(\omega)]f^s(s) = \lambda_B(s)p(\omega)\exp(u_i(s, \omega)),$$
where $\lambda_B(s)$ is the multiplier of the budget constraint. Integrating the above with respect to $\omega$, and rearranging, gives

$$\lambda_B(s) = \frac{f^s(s)}{B_i(s)}, \quad (10)$$

which can be substituted back into the first-order condition to give Equation (7).

Equation (7) can be taken to be an application of the usual formula for optimal consumption given log utility and complete markets (in this case conditional on $s$). The gross portfolio return

$$R^P \equiv \frac{\varepsilon f^I(\omega|s) + (1 - \varepsilon)f^\omega(\omega)}{p(\omega)} \quad (11)$$

is optimal for a log investor conditional on observing $s$.

A related gross portfolio return

$$R^B \equiv \frac{f^\omega(\omega)}{p(\omega)} \quad (12)$$

is optimal for a log investor who does not observe $s$. We will refer to this portfolio as the *benchmark* portfolio, motivated by the fact that benchmark portfolios in practice are intended to be sensible passively managed portfolios.

We can also identify the signal return

$$R^I(s) \equiv \frac{f^I(\omega|s)}{p(\omega)}, \quad (13)$$

the optimal return for maximum effort, $\varepsilon = 1$. The investor’s optimal return is then $R^P = \varepsilon R^I + (1 - \varepsilon)R^B$. Thus, the effect of better information is to tilt the portfolio more toward the signal return $R^I$, and this holds regardless of which version of the problem the investor is solving.

Using lemma 1, we can compute the investor’s expected utility as

$$\int \log \left( w_0 - \int \exp(u_m(s, \omega))p(\omega)d\omega \right) f^s(s)ds$$

$$+ \int \int \log \left( \frac{\varepsilon f^I(\omega|s) + (1 - \varepsilon)f^\omega(\omega)}{p(\omega)} \right) (\varepsilon f^I(s, \omega)$$

$$+ (1 - \varepsilon)f^\omega(\omega)f^s(s))dsd\omega. \quad (14)$$

Note that the second term, which we will denote by $K(\varepsilon)$, depends only on effort, $\varepsilon$, and not on the manager’s utilities. This means we can ignore this term when solving the problem of what contract will implement a particular effort level and take it into consideration only when optimizing the effort levels. Note also that the first term is concave in the manager’s utilities. We will assume
2. Optimal Contracts

We now describe the solutions to each of the three problems stated above. We begin with the simplest problem, the first-best. Then we demonstrate the impact of the agency problems by showing how the solution changes as we add incentive compatibility constraints in the second-best and third-best.

**First-best.** In a first-best contract, we expect to find that there is optimal risk sharing between the manager and the investor. This means that the marginal utility of wealth for the manager should be proportional to the investor’s marginal utility in all states.

The first-order condition for $u_m$ is

$$\frac{\exp(u_m(s, \omega))p(\omega)}{B_i(s)} = \lambda_R (f^{\omega}(\omega) + \varepsilon(f^I(\omega|s) - f^{\omega}(\omega)))$$

where $\lambda_R$ is the Lagrange multiplier on the participation constraint. Multiplying both sides by $B_i(s)$ and integrating both sides with respect to $\omega$, we obtain

$$B_m(s) = \lambda_R B_i(s).$$

Since the two budget shares must sum to $w_0$, we have

$$B_i(s) = \frac{w_0}{1 + \lambda_R},$$

from which we obtain

$$u_m(s, \omega) = \log \left( \frac{w_0 \lambda_R f^{\omega}(\omega) + \varepsilon(f^I(\omega|s) - f^{\omega}(\omega))}{1 + \lambda_R} \right)$$

or, equivalently, the manager’s fee is

$$\phi(s, \omega) = \frac{w_0 \lambda_R R^P}{1 + \lambda_R},$$

since $u_m(s, \omega) = \log(\phi(s, \omega))$. Comparing this with Equation (7), substituting the definition of $B_i(s)$ from above, we see that the first-best contract is a sharing rule that gives the manager a fixed proportion of the payoff of the portfolio independently of the signal. So, as expected, optimal risk sharing is derived. It is worth noting that this result does not depend on the mixture distribution assumption. A proportional sharing rule would still be the first-best contract even under alternative distributional assumptions.

We have not reported the Lagrange multiplier $\lambda_R$, but it is easy to do so by substituting the manager’s fee (19) into the reservation utility constraint (4).
As mentioned above, the first-best contract assumes that moral hazard and adverse selection are not a problem, and that the effort the manager exerts and the signal the manager observes (not just what is reported) can be contracted upon. However, it turns out that even if truthful reporting of the signal cannot be verified, the manager will still report truthfully. Because we assume log utility, this result follows from Ross’s (1974, 1979) work on preference similarity. In other words, a manager who is constrained to take the first-best effort and is faced with a contract of the form (18) will choose to report the signal honestly, since the budget share does not depend on the reported signal. Misreporting will only affect $R^P$. But $R^P$ is the gross return on an optimal portfolio for a log investor. Misreporting the signal can only make the manager worse off because it is equivalent to the choice of a suboptimal portfolio.

Connecting the contract in a single-period model with the actual multiperiod economy should not be oversold. However, it is worth observing that this contract resembles the commonly observed contract paying a fixed proportion of funds under management. Of course, the implications of this contract may be a lot different in our single-period model than in a multiperiod world in which the amount of funds under management can depend on past performance.

**Second-best.** In a second-best world, effort is not observable and therefore the contract must be incentive-compatible for effort. Note that in the manager’s effort optimization problem in condition (5), the objective function is concave because his utility is affine in effort and his cost of effort is convex. Therefore, we can adopt the first-order approach of Holmström (1979) and replace the effort incentive-compatibility condition (5) with the following sufficient first-order condition for the manager’s maximization:

\[
\int \int u_m(s, \omega)(f^I(\omega|s) - f^{\omega}(\omega))f^I(s)dsd\omega - c'(\varepsilon) = 0. \tag{20}
\]

**Proposition 1.** The second-best contract gives the manager a payoff that is proportional to the investor’s payoff plus a bonus that is proportional to the excess return of the portfolio over the benchmark:

\[
\phi(s, \omega) = B_m(R^P + k(R^P - R^B)),
\]

where $B_m$ and $k$ are nonnegative constants.

**Proof.** In the first-order version of the problem, the investor’s first-order condition for $u_m(s, \omega)$ is

\[
\frac{\exp(u_m(s, \omega))p(\omega)}{B_1(s)} = \lambda_R(f^\omega(\omega) + \varepsilon(f^I(\omega|s) - f^{\omega}(\omega)))
\]

\[
+ \lambda_s(f^I(\omega|s) - f^{\omega}(\omega)), \tag{21}
\]
where $\lambda_\varepsilon$ is the Lagrange multiplier on the IC-effort constraint. Proceeding as in the derivation of the first-best case, we find that the budget shares are of the same form as in the first-best contract so that we obtain

$$u_m(s, \omega) = \log \left( \frac{u_0 \lambda_R}{1 + \lambda_R} \frac{f^0(\omega) + \lambda_\varepsilon f^1(\omega|s) - f^\omega(\omega)}{p(\omega)} \right)$$

(22)

or, equivalently, the manager’s fee is

$$\phi(s, \omega) = B_m (R_p + k(R_p - R_b)),$$

(23)

where

$$k = \frac{\lambda_\varepsilon}{\varepsilon \lambda_R} \geq 0.$$  

(24)

The difference between this contract and the first-best contract is that the second-best contract gives the manager a “bonus” that is proportional to the excess return of the fund over a benchmark in addition to a fraction of end-of-period assets under management. This suggests using excess returns over a benchmark as a measure of portfolio performance. This is intriguing since measuring portfolio performance relative to a benchmark is common practice in the portfolio management industry.

The mixture-model assumption plays two roles in this analysis. First, as noted in the literature, it implies that any first-order solution is a solution of the underlying agency model since the first-order conditions for the manager are necessary and sufficient. Second, the mixture model assumption implies that the benchmark in the solution can be chosen to be the uninformed optimum.

Absent the mixture model assumption, the optimal contract will include a bonus that is proportional to the excess return over a benchmark but, in general, this benchmark will not be the uninformed optimum and it may depend on the reported signal. Let $f(\omega|s; \varepsilon)$ be the conditional distribution of the market state given the signal. If this distribution is differentiable in effort and the first-order approach is still valid, then the first-order condition for $u_m(s, \omega)$ is

$$\exp(u_m(s, \omega)) p(\omega) B_i(s) = \lambda_R f(\omega|s; \varepsilon) + \lambda_\varepsilon f_\varepsilon(\omega|s; \varepsilon),$$

(25)

where the subscript indicates partial derivative. When we multiply both sides by state prices and integrate with respect to market states, the term involving $\lambda_\varepsilon$ drops out because $f(\omega|s; \varepsilon)$ integrates to one for all $s$ and we can interchange the order of integration and differentiation. Therefore, the budget share is constant and is of the same form as in the first-best case. The random variable

$$Z \equiv \lambda_\varepsilon \frac{f_\varepsilon(\omega|s; \varepsilon)}{p(\omega)}$$

(26)
is a zero-cost payoff. Because of complete markets, this random variable is some excess return. We can interpret this random excess return as the excess return of the managed portfolio over some other portfolio return defined by

\[ R^O = R^P - Z. \quad (27) \]

The manager’s payoff is

\[ \phi(s, \omega) = B_m[R^P + k'(R^P - R^O)], \quad (28) \]

where \( k' = \lambda_e/\lambda_R \). In general, of course, this “benchmark” \( R^O \) will not be the uninformed optimum because it will be some function of \( s \), the reported signal (which is okay since the signal is observed in the second-best, but not consistent with the usual choice of a benchmark in practice as an uninformed portfolio).

Even if the first-order approach fails and there are non-locally-binding incentive compatibility constraints, a similar expression can be derived. In this general case, the manager still receives a proportion of the portfolio payoff plus a constant times excess return relative a benchmark; however, the benchmark loses the simple interpretation as the uninformed log-optimal portfolio.

If we relax the assumption of log utility, we lose the strict linearity of the contract. An Appendix available from the authors shows that when the investor and manager have power utility, with the risk aversion coefficients \( \gamma_i \) and \( \gamma_m \), respectively, the manager’s second-best payoff is a power function of a linear combination of the returns \( R^P \) and \( R^B \) above, and the budget share depends on the signal. In particular, the manager’s fee is of the form

\[ \phi(s, \omega) = \frac{B_m(s)}{Q(s)} [R^P + k(R^P - R^B)]^{1/\gamma_m}, \quad (29) \]

where the manager’s budget \( B_m(s) \) and the function \( Q(s) \) are defined by

\[ Q(s) = \int [R^P + k(R^P - R^B)]^{1/\gamma_m} p(\omega) d\omega, \quad (30) \]

\[ Q_1(s) = \left[ \int (R^P)^{1/\gamma_i} p(\omega) d\omega \right]^{\gamma_i}, \quad (31) \]

\[ B_m(s) = \left( \frac{B_i(s)^{\gamma_i} \lambda_e}{Q_1(s)} \right)^{1/\gamma_m} Q(s), \quad (32) \]

\[ B_i(s) = w_0 - B_m(s). \quad (33) \]

It is clear from Equation (29) that, for \( \gamma_m \) not equal to unity, the manager’s optimal fee is no longer linear in \( R^P \) and \( R^B \). To see how well the manager’s optimal fee is approximated by a linear contract, we simulate values of the returns \( R^P \) and \( R^B \) and the corresponding values of the manager’s optimal fee, and then calculate the \( R^2 \)s of numerical regressions of the manager’s fee on these returns. In an Appendix available from the authors, we find that these \( R^2 \)s
are quite high over a range of risk aversion levels, indicating that the optimal contracts are nearly linear.

**Third-best.** In the third-best, truthful reporting of the signal must also be incentive compatible. However, in the manager’s optimization problem in condition (6), the maximum across reporting strategies of the double integral is the maximum of affine functions and is, therefore, convex. The curvature in the cost function may or may not overcome the curvature in the optimized double integral. If not, the objective function fails to be concave and the first-order conditions may fail to characterize the incentive-compatibility constraint. Moreover, since the utility levels in the double integral are endogenous to the investor’s choice problem, we do not know how to specify a priori a level of convexity in the cost function $c(\cdot)$ large enough to ensure the manager’s objective function is concave.\(^{10}\)

However, if the cost function is sufficiently convex, we can apply the first-order approach and replace the joint effort and reporting incentive compatibility constraint (6) with the manager’s first-order condition for effort, Equation (20), together with the following first-order conditions for report choice, evaluated at $\rho(s) = s$:

\[
(\forall s \in S) \int \frac{\partial u_m(s, \omega)}{\partial s} (\varepsilon f^I(\omega|s) + (1 - \varepsilon) f^o(\omega)) f^s(s) d\omega = 0. \tag{34}
\]

With this substitution, the investor’s first-order condition for $u_m$ is

\[
\frac{\exp(u_m(s, \omega)) p(\omega)}{B_i(s)} = \lambda_R (f^o(\omega) + \varepsilon(f^I(\omega|s) - f^o(\omega)))
\]

\[
+ \lambda_s(f^I(\omega|s) - f^o(\omega))
\]

\[
- \lambda'_s(s)(f^o(\omega) + \varepsilon(f^I(\omega|s) - f^o(\omega)))
\]

\[
- \varepsilon \lambda_s(s) \frac{\partial f^I(\omega|s)}{\partial s},
\]

where $\lambda_s(s)$ is the Lagrange multiplier on the truthful reporting constraint. In this case, we have

\[
B_i(s) = \frac{w_0}{1 + \lambda_R - \frac{\lambda'_s(s)}{f^s(s)}}, \tag{36}
\]

and

\[
B_m(s) = \frac{w_0 \left( \lambda_R - \frac{\lambda'_s(s)}{f^s(s)} \right)}{1 + \lambda_R - \frac{\lambda'_s(s)}{f^s(s)}}. \tag{37}
\]

\(^{10}\) In the limiting case of a proportional cost function, the objective is convex in effort, once we have optimized the over-reporting strategy, and the manager will never choose an interior effort level. In this case, any binding incentive-compatibility constraint will compare full effort with no effort, and will not be the same as the local condition.
Equation (37) indicates that the manager’s share of the budget in the third-best is no longer constant, as it was in the second-best, but will generally depend on the signal. This is part of the mechanism for inducing truthful reporting. In addition, the final term in Equation (35) indicates that even conditional on the signal, the manager’s fee is no longer a linear combination of the returns $R_P$ and $R_B$, but contains an additional payoff, which can be shown to have zero net present value. To develop intuition for these results, we examine numerical examples of these payoffs in the next section.

3. Numerical Results

Now we turn to numerical results that compare the first-best, second-best, and third-best. We assume conditional joint normality of $\omega$ and $s$, with correlation $\rho > 0$ under the “informed” distribution and $\rho = 0$ under the “uninformed” distribution. The marginal densities of $\omega$ and $s$ are the same under the informed as under the uninformed. We think of this as a model of market timing, with $\omega$ representing the demeaned log one-year market return in the usual lognormal model. Let $n(\cdot; \cdot, \cdot)$ be the normal density parameterized by mean and variance. Then $f^i(s) = n(s; 0, \sigma^2)$ is the density of $s$ in either case, $f^{\omega}(\omega) = n(\omega; 0, \sigma^2)$ is both the unconditional density and the conditional density of the market state $\omega$ in the uninformed case, and $f^I(\omega|s) = n(\omega; \rho s, \sigma^2(1 - \rho^2))$ is the conditional density of the market state $\omega$ given $s$ in the informed case. Thus, with maximum effort, i.e., $\varepsilon = 1$, observing the signal is like observing the market for a fraction $\rho^2$ of the year and seeing that its excess log return is up by $\rho s$.

State prices are consistent with Black-Scholes and can be computed as the discount factor times the risk-neutral probabilities as $p(\omega) = e^{-r}n(\omega; r - \mu, \sigma^2)$. In these expressions, $r$ is the risk-free rate, $\mu$ is the mean return on the market, and $\sigma$ is the standard deviation of the market return. Without loss of generality, the signal $s$ has mean 0 and the same variance as the log of the market return. The parameter values used are $\mu = 0.10$, $\sigma = 0.2$, $\rho = 0.5$, $r = 0.05$, and $w_0 = 100$.

To facilitate the comparison of the cases, we vary the cost function to make the same effort level optimal in the first-, second-, and third-best. This removes the obvious distinction among the contracts that higher equilibrium effort implies a more informative signal and therefore more aggressive portfolios for both agents. By fixing effort exogenously, we isolate the differences among the contracts due solely to the addition of the IC constraints.\footnote{Another question of interest is how optimal effort and principal utility vary as we add constraints, holding the effort cost function constant. To illustrate the cost of these constraints in terms of investor welfare and suboptimal effort, we fix the effort cost function to $c(\varepsilon) = (\varepsilon + 0.3)^{10} - 0.3^{10}$, which ensures interior optimal effort levels, and examine the three different solutions. In the first-best, the optimal managerial effort level is 0.69 and the principal’s uninformed equivalent wealth, i.e., the amount of wealth that would make the principal equally happy if he invested alone and without information, is 104.2. The first-best fixed fee is 2.15%. In the second best, the optimal effort level falls to 0.5, the principal’s uninformed equivalent wealth falls to 101.5, the fixed component of the fee is 1.77%, and the benchmark coefficient is $k = 1$. In the third-best, the optimal effort level falls to 0.35,} We adjust the
cost of effort function so that $u_0 - c(\varepsilon) = 0.955$ and $c'(\varepsilon) = 0.2$ at the chosen equilibrium effort level, $\varepsilon = 0.5$.

We work with discretized versions of $f^I(\omega|s)$, $f^U(\omega)$, and $p(\omega)$ with $N$ market states and $M$ signal states. In order to circumvent the difficulty imposed by the presence of $\lambda'_s(s)$ in this first-order condition of the third-best problem, we work with a discrete version in which the reporting constraint is replaced by two sets of reporting constraints. The first set imposes the restriction that reporting the state just higher than the true state is not optimal and the other does the same for reporting the state just lower than the true state. Together this makes $2(M - 1)$ constraints. As the discretization becomes very fine, this problem approximates the continuous state case.

The manager’s utilities from the first-best problem are plotted in Figure 1. Not surprisingly, the figure shows that for higher signal states, the optimal payoff is a longer position in the market.

A visual inspection of the solution to the second- and third-best contracts at these parameter values is not very instructive. However, we can gain insight by examining the incremental changes in the contract when we move from first-best to second-best to third-best. Figure 2 plots the manager’s utilities

---

the principal’s uninformed equivalent wealth falls to 100.6, and the manager’s budget share, which corresponds roughly to a fixed fee, is about 1% on average across signal states.
Figure 2
Manager's utilities: Second-best minus first-best levels

in the second-best minus the manager's utilities in the first-best. When signal and market are both high (or both low), \( f^{I}(\omega|s) > f^{o}(\omega) \), so the manager is rewarded in those states. In the other corners of the distribution, the manager has less utility than in the first-best case. This provides the incentive to exert effort.

Figure 3 plots the manager's utilities in the third-best minus the manager's utilities in the second-best. The difference between these two contracts is that the third-best provides incentives to report truthfully. From the figure, we can see that compared to the second-best, the third-best contract provides extra rewards for reporting more extreme signals.

The intuition for this is straightforward. In order to induce an effort to generate a quality signal, the second-best contract overexposes the manager to the signal's optimal portfolio. A manager who could misreport the signal would therefore have incentive to report a more conservative signal in order to partially undo this overexposure. This may be related to plan sponsors' common concern that managers might be "closet indexers" who mimic the index but collect fees more appropriate for active managers. To address this problem, the third-best contract must provide an additional component of compensation that rewards the manager for reporting more extreme signals.

Another way to understand the difference between the third- and second-best contracts is to look at the differences in payoff, which are plotted in Figure 4.
Notice that, in dollar terms, the difference between the two contracts is very subtle in low-signal states of the world. When the signal is low the manager’s pay is also low in the second-best contract and so marginal utility is high. This means that only a small increase in pay is required to induce the manager to report the correct state. However, when the signal state is high the manager’s pay is also high and so a very large bonus is required to induce truthful reporting. Thus, the way to truthful reporting may be more of a “carrot” approach than a “stick.” Note, finally, that in the high-signal states, which are associated with high expected market returns, the incremental payments in the third-best look like those of a long position, perhaps conditional a call option position, in the market. Thus, the manager’s extra reward is invested efficiently, given the signal.

Finally, Figure 5 illustrates the difference between third- and second-best manager’s fees in the case when the investor is less risk averse than the manager. In particular, the investor’s coefficient of relative risk aversion is 0.75. The results are quite similar to the case of log utility. In the Appendix available from the authors, we also examine the case when the investor has higher risk aversion, and the results are still qualitatively the same. In particular, the effect of the third-best contract is still to encourage more risk-taking in the form of reports of more extreme states.
Figure 4
Manager’s payoff: Third-best minus second-best levels

Figure 5
Manager’s payoff: Third-best minus second-best levels when investor has RRA = 1.0 and manager has RRA = 0.75
4. Interpretation and Implications

By mapping out optimal contracts across a range of contracting settings, we show how the optimal compensation varies with the transparency of the information generation process. At one extreme, when the collection and analysis of information is verifiable and easy to interpret, a simple proportional sharing rule is best. This might be the case in settings where there are standardized protocols for what data to gather and what analysis to perform, and where the results of the analysis are easily understood.

However, if adequate information generation is hard to verify, though the results are still straightforward to interpret or difficult to manipulate, then a simple linear benchmark contract can effectively give a manager incentives to generate information when used in conjunction with portfolio restrictions that prevent his undoing those incentives. This situation might arise, for example, if the misreporting problem is mitigated by strong governance mechanisms, such as boards of directors who closely monitor and question the manager, or with reputation effects that make it costly for the manager to be a “closet indexer.”

By overexposing the manager to his own strategy, the benchmarking gives the manager incentive to understate his information. Thus, at the other extreme, if both effort and information are easy to conceal, and there are no mechanisms outside the contract that discourage the manager from misreporting the information, then a third-best compensation component is needed to induce the manager to act aggressively on extreme information.

Explicit contracts used in practice may look something like these. We see simple percentage fees for most mutual fund managers, linear benchmarking in some cases, and option-like contracts in the hedge fund industry. It is tempting to conclude, and perhaps plausible, that these contracts must coincide with increasingly complex information generation processes. However, in the real, multiperiod world, the total compensation function is also affected by the fund’s performance-flow relation. Our paper predicts that the complexity and nonlinearity of the total compensation function will increase with the opacity of the information generation process.

5. Conclusion

We have proposed a new model of optimal contracting in the agency problem in delegated portfolio management. We have shown that in a first-best world with log utility, the optimal contract is a proportional sharing rule over the portfolio payoff. In a second-best world, the optimal contract (if it exists) is a proportional sharing rule plus a bonus proportional to the excess return over a benchmark to give incentives to the manager to work hard. In a third-best world, such excess return strategies will provide incentives to work but will tend to make the manager overly conservative. These results have been demonstrated in the context of a realistic return model and the derived performance
measurement criterion looks more like the simple benchmark comparisons used by practitioners than more elaborate measures such as the Jensen measure, Sharpe measure, or various marginal-utility weighted measures. In addition, the optimal contract includes restrictions on the set of permitted strategies. These institutional features are more similar to practice than other existing agency models in finance.

Many interesting extensions are possible. For example, it could be illuminating to add more securities to extend the numerical results from a model of market timing to a model of security selection. A more challenging extension is to extend the model to consider quality of trade execution, which will require some modeling of trading opportunities and probably requires many changes in the model. This seems especially promising because studying execution could take advantage of the rich trade-by-trade data available in reports from custodians.

A different kind of extension would include explicitly the two levels of portfolio management we see in practice, with the separation of responsibilities for asset allocation across asset classes and management of sub-portfolios in each asset class. The ultimate beneficiaries have to create incentives for the overall manager to hire and compensate the asset class managers, and this could be modeled as a hierarchy of agency contracts.

References


