Department of Finance

Working Paper Series

FIN-03-031

Portfolio Performance and Agency

Philip H. Dybvig, Heber K. Farnsworth and Jennifer N. Carpenter

September 2003

Portfolio Performance and Agency

Philip H. Dybvig
Washington University in Saint Louis
Heber K. Farnsworth*
Washington University in Saint Louis
Jennifer N. Carpenter
New York University

* **Corresponding Author:** Heber Farnsworth, John M. Olin School of Business, Washington University, Campus Box 1133, One Brookings Drive, St. Louis, MO 63130-4899, tel: (314) 935-4221, fax: (314) 935-6359, `mailto:farnsworth@olin.wustl.edu`

Abstract

The literature traditionally assumes that a portfolio manager who expends costly effort to generate information makes an unrestricted portfolio choice and is paid according to a sharing rule. However, the revelation principle provides a more efficient institution. If credible communication of the signal is possible, then the optimal contract restricts portfolio choice and pays the manager a fraction of a benchmark plus a bonus proportional to performance relative to the benchmark. If credible communication is not possible, an additional incentive to report extreme signals may be required to remove a possible incentive to underprovide effort and feign a neutral signal.

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 but no 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[1] but uses a rich model of security returns with full spanning of market states.

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. The manager can expend effort to influence the precision of a private signal about future market prices. The investor's problem is to find a contract for the manager that provides incentives to take costly effort and to use the signal in the investor's interest while still sharing risk reasonably efficiently. In a similar vein, absence of any information asymmetry at the outset distinguishes us from Garcia (2001), whose managers already know their signals at the time of contracting.

Negative results appear to be more common in this literature than positive ones. Stoughton (1993) examines affine (linear plus a constant) and quadratic contracts in a two-asset world. He finds that affine contracts provide no incentives for effort. Quadratic contracts provide some incentive but are not optimal due to their poor risk sharing properties.[2] Admati and Pfleiderer (1997) contains a similar result

[1] See Ross (1973). For a survey of the agency literature see Laffont and Martimort (2002) or Stole (1993).

[2] There is a claim in Stoughton (1993) that quadratic contracts approach the first-best in the limit as the investor becomes risk neutral. This claim is incorrect. The sense of convergence (small difference in utility) to the first-best as the client becomes less risk averse is not robust to different utility representations. Using a more reasonable sense of convergence measured by certainty-equivalent, one can demonstrate that the contracts do not approach first-best. Specifically, the investor's utility is $U_B(w) = -\exp(-bw)$ and therefore the certainty equivalent is the inverse of this function $CE(u) = -\log(-u)/b$. Then we can compute from (29) and (30) in Stoughton (1993) (using also (4)), that the difference in certainty equivalents for small b is approximately $-\log(1 - 2a\gamma_2/H)/(2a)$, which is a positive constant. Under that paper's normalization, the utility function converges uniformly to zero as risk aversion increases, and consequently *any*

to Stoughton's result for affine contracts. There it is shown that contracts which are affine in the excess return over a benchmark also provide no incentives to take costly effort. These negative results arise from an assumption that the contracts do not restrict the portfolio choice of the manager, as would a general contract under the revelation principle.[3] Restrictions on the manager's portfolio choice are essential for incentive pay schemes to induce effort. Unrestricted trading may allow a manager to completely undo the incentive effects of the fee. 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] This form of contract can be motivated by the revelation principle, and the revelation principle also tells us that this form of contract is general in the sense that it can replicate the equilibrium allocation of any other contract. Actual investment guidelines are full of portfolio restrictions. Common restrictions on asset allocation 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 (growth versus income) of stocks, credit ratings or durations of bonds, restrictions on use of derivatives, maximum allocations to a stock or industry, and increasingly portfolio risk measures such as duration, beta, or tracking error. Almazan, Brown, Carlson, and Chapman (2001) documents the prevalence of such restrictions in contracts observed in the mutual fund industry.

We derive optimal contracts given a mixture assumption under which the joint distribution of the manager's signal and the market state depends affinely on the effort of the manager. We also assume that both the investor and the manager have log utility. In a first-best world, where the manager's effort is contractible, the optimal contract is a proportional sharing rule. In a second-best world, in which the manager's signal is observable but his effort is not contractible, the optimal fee for the manager is a proportion of the managed portfolio plus a share of the excess return of the portfolio over a (endogenously determined) benchmark. This

bounded payoff would be considered asymptotically first-best. For example, if giving the manager 100% of the portfolio value satisfies the participation constraint, it is also asymptotically first-best in that sense. Obviously that is a meaningless sense of convergence.

[3] Gómez and Sharma (2001) have shown that these non-incentive 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] 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.

gives the appropriate incentive to exert effort. The form of the optimal contract suggests the use of excess return over a benchmark as a measure of portfolio performance. Use of excess returns as a performance measure is common in the portfolio management industry. Performance based fees, when they are observed, are generally tied to this measure.

In a third-best world, neither the effort nor the signal is contractible, so additional adjustments are necessary to induce correct portfolio choice. Relative to the second-best, the third-best contract rewards the manager for reporting more extreme signals, or choosing riskier portfolios. This illustrates the limitations of the second-best contract. With the second-best contract in a third-best world a manager could not fully undo the incentives but could mitigate their effect by taking more conservative portfolio positions. 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 a third-best world the manager's compensation is similar to the second-best but with explicit rewards for taking risk.

Conceptually, our 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 choice is to take the correct position or take the opposite position from what the signal would suggest. So in this context the incentive to be overly conservative is not apparent. In addition, the investor in the model of Kihlstrom (1988) is risk-neutral. This would imply that no optimal contract exists except that Kihlstrom also assumes that the manager cannot short. This leads to a corner solution.

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 efficient portfolio 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. Sung (1995)[5] and Ou-Yang (2003) analyze continuous-time models in which both the drift and diffusion coefficient can be controlled, and affine contracts arise opti-

[5]The portfolio application is mentioned in Sung (1995) and spelled out in more detail in Sung's thesis, Sung (1991).

mally. 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, although not a model of delegated portfolio management, the model of delegated expertise of Demski and Sappington (1987) bears several similarities to our model of this paper. In that paper an analyst exerts costly effort to obtain information. The main differences between that paper and the current paper are that the 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. Another model that does not consider portfolio management but is closely related to our problem is Sung (2003), which has a continuous-time model with both moral hazard and adverse selection.

The paper proceeds as follows. Section I describes the optimal contracting problem. Section II presents analytical solutions in the first-best and second-best cases and discusses the problems which arise in the third-best case. Section III concludes.

I The Agency Problem

We consider the contracting problem between an investor and a portfolio manager. This is a moral hazard (or "hidden action") problem because the investor cannot observe the level of costly effort undertaken by the manager. But, it also includes an adverse selection (or "hidden information") problem because the costly effort generates private information that the manager cannot necessarily be trusted to use in the investor's best interest. 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 minimize the impact of information problems. For example, information problems can be minimized by using various forms of information gathering before-the-fact, information gathering after-the-fact, and, in a multi-period context, the impact of reputation on future business. Our analysis considers 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 ω. The integrals used to price payoffs or compute utilities may seem most familiar if the underlying states ω (and s defined later) lie in $\Re^n$ for some n. However, the notation and derivations are also consistent with a discrete state space if the measure is a counting measure (implying that the integrals are sums). The notation and derivations are also consistent with a more complex state spaces (such as the set of paths of Brownian motion) if integrals are taken with respect to a convenient reference measure. The model is a one-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. When the state space is not discrete, there may be a technical issue of exactly what space the market is complete over, and we resolve this issue by assuming that all claims are marketed for which the integral defining the price exists and is finite.

Information Technology Through 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 ε,

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

is the probability density of s and ω where the market state is ω and the signal is s. Here, f^I is an “informed” distribution and f^U is an “uninformed” distribution. We assume that s and ω 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 ω, 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 manager receives a signal that may be informative or it may just be random, and that expending more effort makes it more likely the signal is informative. Using the mixture model would be without loss of generality if there were only two effort levels, and it is a simple sufficient condition for the first-order approach to work in many agency models,[6] 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 normally distributed signal; with this common assumption, the unbounded likelihood ratio in the tails makes it too easy to create approximately first-best incentives, as pointed out by Mirrlees (1974). In the mixture model, likelihood ratios are bounded. Besides justifying the first-order approach (which simplifies exposition), our main use of the mixture assumption is its implication that the benchmark appearing in the solution to the second-best problem is the uninformed optimum. Absent the mixture assumption, the form of the optimal compensation is the same but the benchmark loses its simple interpretation.

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 ϕ is the manager's fee and $c(\varepsilon)$ is the cost of the effort ε (the hidden action). We shall assume that $c(\varepsilon)$ is differentiable and convex with $c'(0) = 0$. We will assume that all the problems we consider have optimal solutions.[7] 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 will use utility levels rather than consumption levels as the choice variables, which make most of the constraints affine. We will denote by $u_i(s,\omega)$ the investor's equilibrium utility level $\log(V)$ given s and ω, and we will denote by $u_m(s,\omega)$ the manager's equilibrium utility level for only the wealth component $\log(\phi)$ given s and ω.

[6] Rogerson (1985) attributes Holmström (1984) with pointing out the appeal of the mixture model as an alternative to the more complex convexity condition of Mirrlees (1976). See also Grossman and Hart (1983) and Hart and Holmström (1987).

[7] 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 first-best problems.

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. This contract specifies a portfolio strategy for each possible signal realization and how the portfolio payoff is to be divided between the investor and the manager. 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 ε and receives the private signal s. The manager announces the signal and the portfolio associated with the signal in the contract is selected. Finally, portfolio returns are realized and the manager and the investor get the payment described in the contract.

This specification of the problem has a signal announcement that may seem somewhat artificial. This is a *direct mechanism*, which according to the *revelation principle* is guaranteed to duplicate all possible mechanisms, in effect if not in form. 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). The merit of looking at a direct mechanism is that it allows us to find contracts that implement allocations that can be implemented using the sharing rules traditionally studied in the literature as well as any alternative institutions that may do better. The more general contracts also have a nice economic interpretation in terms of restrictions on the investment strategy.

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 we are planning for the manager. We will look at several forms of the problem. 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 correctly reveal the observed signal.

Problem 1 (Third-best) *Choose* $u_i(s,\omega)$, $u_m(s,\omega)$, *and* ε *to maximize investor's expected utility,*

$$(2) \qquad \iint u_i(s,\omega)(\varepsilon f^I(\omega|s) + (1-\varepsilon)f^\omega(\omega))f^s(s)d\omega ds,$$

subject to the budget constraint,

$$(3) \qquad (\forall s \in S) \int \left(\exp(u_i(s,\omega)) + \exp(u_m(s,\omega))\right) p(\omega)d\omega = w_0,$$

the participation constraint,

$$(4) \qquad \iint u_m(s,\omega)(\varepsilon f^I(\omega|s) + (1-\varepsilon)f^\omega(\omega))f^s(s)d\omega ds - c(\varepsilon) = u_0,$$

and incentive-compatibility of effort and signal reporting,

$$\begin{aligned}(5) \qquad & (\forall \varepsilon' \in \mathcal{E} \text{ and } \rho : S \to S) \\ & \iint u_m(s,\omega)(\varepsilon f^I(\omega|s) + (1-\varepsilon)f^\omega(\omega))f^s(s)d\omega ds - c(\varepsilon) \\ & \quad \geq \iint u_m(\rho(s),\omega)(\varepsilon' f^I(\omega|s) + (1-\varepsilon')f^\omega(\omega))f^s(s)d\omega ds - c(\varepsilon').\end{aligned}$$

In the choice problem, the choice variables are the effort level ε and the utility levels for investor and manager in each contingency (s,ω). 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 ω given the effort level ε. The budget constraint computes the consumptions for investor and manager from their utility levels (the exponential function is the inverse of the logarithm) and values it using the pricing rule $p(\omega)$. There is a separate budget constraint given each signal realization because there are no opportunities to hedge consumption across signal realizations. The pricing is the same for each s because of the "small investor" assumption that the manager does not affect pricing in security markets. The participation constraint says that the agent has to be treated well enough to meet the reservation utility level u_0 of outside opportunities. Finally, the incentive-compatibility condition says that the manager wants to do as is planned and would not be better off selecting a different level of effort (ε' instead of ε) and/or misreporting states (reporting $\rho(s)$ instead of s).

The *second-best problem* assumes that signal reporting is not a problem but that incentive for effort may be, so the constraint for incentive-compatibility of effort and signal reporting is replaced by a constraint for incentive-compatibility of effort only. 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.

Problem 2 (Second-best) *Choose $u_i(s,\omega)$, $u_m(s,\omega)$, and ε to maximize investor's expected utility (2) subject to the budget constraint (3), the participation constraint (4), and incentive-compatibility of effort*

(6) $$(\forall \varepsilon' \in \mathcal{E})$$
$$\iint u_m(s,\omega)(\varepsilon f^I(\omega|s) + (1-\varepsilon)f^\omega(\omega))f^s(s)d\omega ds - c(\varepsilon)$$
$$\geq \iint u_m(s,\omega)(\varepsilon' f^I(\omega|s) + (1-\varepsilon')f^\omega(\omega))f^s(s)d\omega ds - c(\varepsilon').$$

And, in the first-best, there is assumed to be no incentive problem for inducing the appropriate effort and reporting, although there is still a problem of allocating scarce resources efficiently to do as well as possible for the investor while satisfying the manager's minimum requirment as given in the reservation utility constraint. The first-best is not realistic but it is a useful ideal benchmark.

Problem 3 (First-best) *Choose $u_i(s,\omega)$, $u_m(s,\omega)$, and ε to maximize investor's* *expected utility (2) subject to the budget constraint (3), and the participation constraint (4).*

We can reduce both the number of choice variables and the number of constraints by use of the following lemma, which enables us to 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 Problem 1, 2, or 3, the expected utility conditional on s for the investor is given by*

$$\log\left(B_i(s)\frac{f^\omega(\omega)+\varepsilon(f^I(\omega|s)-f^\omega(\omega))}{p(\omega)}\right) \tag{7}$$

where

$$B_i(s) = w_0 - \int \exp(u_m(s,\omega))p(\omega)d\omega \tag{8}$$

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 Problems 1, 2, and 3 in the objective function (2) and in the budget constraint (3). Therefore, the optimal solution must solve the subproblem of maximizing (2) subject to (3). The first-order condition of this problem is

$$\left[\varepsilon f^I(\omega|s) + (1-\varepsilon)f^\omega(\omega)\right] f^s(s) = \lambda_B(s)p(\omega)\exp(u_i(s,\omega)) \tag{9}$$

where $\lambda_B(s)$ is the multiplier of the budget constraint. Integrating the above with respect to ω and rearranging gives

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

which can be substituted back into the first-order condition to give (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

$$(10)\quad R^P \equiv \frac{\varepsilon f^I(\omega|s) + (1-\varepsilon) f^\omega(\omega)}{p(\omega)}$$

is optimal for a log investor conditional on observing s.

A related gross portfolio return

$$(11)\quad R^B \equiv \frac{f^\omega(\omega)}{p(\omega)}$$

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.

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

$$(12)\quad \int \log\left(w_0 - \int \exp(u_m(s,\omega))p(\omega)d\omega\right) f^s(s)ds$$
$$+ \iint \log\left(\frac{\varepsilon f^I(\omega|s) + (1-\varepsilon) f^\omega(\omega)}{p(\omega)}\right)(\varepsilon f^I(s,\omega) + (1-\varepsilon) f^\varepsilon(\omega) f^s(s))ds d\omega$$

Note that the second term depends only on effort, ε, 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 over effort levels. Note also that the first term is concave in the manager's utilities. We will assume the second term is finite for all effort levels ε to avoid some technical difficulties that are far from the main concerns of our paper.

There is no incentive-compatibility constraint in the first-best (the ideal situation in which the manager's choices can be dictated). In solving the second-best and third-best problems we desire a more convenient characterization of the incentive-compatibility constraints. We shall adopt what is known as the *first-order approach* of Holmström (1979) to solving principal-agent problems. In the first-order approach, the optimization in each incentive-compatibility condition is replaced by its first-order condition. The manager is the agent in this case and the manager's choice problem is

Problem 4 *(Manager's Problem) Choose an effort level ε' and a reporting strategy $\rho(s)$ to maximize*

$$\iint u_m(\rho(s), \omega)(\varepsilon' f^I(\omega|s) + (1-\varepsilon')f^\omega(\omega))f^s(s)d\omega ds - c(\varepsilon') \tag{13}$$

Substituting the first-order conditions of this problem, evaluated at $\varepsilon' = \varepsilon$ and $\rho(s) = s$ for the incentive-compatibility contraints, we obtain the first-order version of the third-best problem:

Problem 5 *(First-order Third-best) Choose ε and $u_m(s, \omega)$ to maximize*

$$\int \log\left(w_0 - \int \exp(u_m(s,\omega))p(\omega)d\omega\right) f^s(s)ds + K(\varepsilon) \tag{14}$$

subject to manager participation

$$\iint u_m(s, \omega)(\varepsilon f^I(\omega|s) + (1-\varepsilon)f^\omega(\omega))f^s(s)d\omega ds - c(\varepsilon) = u_0, \tag{15}$$

incentive-compatibility of effort choice

$$\iint u_m(s, \omega)(f^I(\omega|s) - f^\omega(\omega))f^s(s)dsd\omega - c'(\varepsilon) = 0, \tag{16}$$

and incentive-compatibility of truthful reporting

$$\int \frac{\partial u_m(s,\omega)}{\partial s}(\varepsilon f^I(\omega|s) + (1-\varepsilon)f^\omega(\omega))f^s(s)d\omega = 0. \tag{17}$$

($K(\varepsilon)$ is the second term of the investor's expected utility function in (12) that doesn't depend on manager utilities.)

The incentive-compatibility of truthful reporting condition (17) assumes that the support of s is a continuum of values so that a derivative is appropriate. If s is discrete, then there would be a finite difference condition instead.

In the first-order second-best problem, truthful reporting of the state is assumed not to be a problem, so we only impose incentive-compatibility of effort:

Problem 6 *(First-order Second-best) Choose ε and $u_m(s,\omega)$ to maximize the investor's indirect utility (14) subjective to manager participation (15) and incentive-compatibility of effort (16).*

The first-order version of the first-best problem is the same as the original version.

II 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

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

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

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

Since the two budget shares must sum to w_0 we have

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

from which we obtain

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

or equivalently the manager's fee is

$$\phi(s,\omega) = \frac{w_0\lambda_R}{1+\lambda_R}R^P, \tag{22}$$

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 which gives the manager a fixed proportion of the payoff of the portfolio independent of the signal. So, as expected, optimal risk sharing obtains. 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 solved for the Lagrange multiplier λ_R, but it is easy to do so by substituting the manager's fee (22) into the reservation utility contraint (15).

As mentioned above, the first-best contract assumes that moral hazard and adverse selection are not a problem, 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, a result that can be seen as consistent with the notion of preference similarity described by Ross (1974, 1979). In other words a manager who is constrained to take the first-best effort and is faced with a contract of the form (21) 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 so the contract must be incentive-compatible for effort.

Proposition 1 *The second-best contract gives the manager a payoff which 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 non-negative constants.

PROOF We will work with the first-order version of the problem. Here, the first-order condition for $u_m(s,\omega)$ is

$$\text{(23)} \quad \frac{\exp(u_m(s,\omega))p(\omega)}{B_i(s)} = \lambda_R(f^\omega(\omega) + \varepsilon(f^I(\omega|s) - f^\omega(\omega))) + \lambda_a(f^I(\omega|s) - f^\omega(\omega))$$

where λ_a 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

$$\text{(24)} \quad u_m(s,\omega) = \log\left(\frac{w_0\lambda_R}{(1+\lambda_R)}\frac{f^\omega(\omega) + (\varepsilon + \frac{\lambda_a}{\lambda_R})(f^I(\omega|s) - f^\omega(\omega))}{p(\omega)}\right)$$

or equivalently the manager's fee is

$$\text{(25)} \quad \phi(s,\omega) = B^A\left(R^P + k(R^P - R^B)\right)$$

where

$$\text{(26)} \quad k = \frac{\lambda_a}{\varepsilon\lambda_R} \geq 0.$$

∎

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. 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

$$\frac{\exp(u_m(s,\omega))p(\omega)}{B_i(s)} = \lambda_R f(\omega|s;\varepsilon) + \lambda_a f_\varepsilon(\omega|s;\varepsilon) \tag{27}$$

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

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

is a zero cost payoff. Because of complete markets this random variable is some excess return. We can take this random excess return to be the excess return of the managed portfolio over some other portfolio return defined by

$$R^O = R^P - Z.$$

The managers payoff is

$$\phi(s,\omega) = B^A \left[R^P + k'\left(R^P - R^O\right)\right]$$

where $k' = \lambda_a/\lambda_R$. 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).

If the first-order approach fails and there are non-locally-binding incentive compatibility constraints then a similar expression can be derived. The general first-order condition for the principal's problem may put weights on both local and non-local changes. Combining the weighted average of the corresponding density changes from the optimum and dividing by $p(\omega)$ gives the appropriate change from the optimum to the benchmark.

In one of the numerical examples, the solution is almost the same

Third-best The third-best contract solves Problem 5. The first-order condition for u_m is

$$
\begin{aligned}
(28)\quad \frac{\exp(u_m(s,\omega))p(\omega)}{B_i(s)} \quad = \quad & \lambda_R(f^\omega(\omega) + \varepsilon(f^I(\omega|s) - f^\omega(\omega))) \\
& + \lambda_a(f^I(\omega|s) - f^\omega(\omega)) - \varepsilon\lambda_s(s)\frac{\partial f^I(\omega|s)}{\partial s} \\
& - \lambda_s'(s)(\varepsilon f^I(\omega|s) + (1-\varepsilon)f^\omega(\omega))
\end{aligned}
$$

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

$$
(29)\quad B_i(s) = \frac{w_0}{1 + \lambda_R - \frac{\lambda_s'(s)}{f^s(s)}}
$$

and

$$
(30)\quad B_m(s) = \frac{w_0(\lambda_R - \frac{\lambda_s'(s)}{f^s(s)})}{1 + \lambda_R - \frac{\lambda_s'(s)}{f^s(s)}}.
$$

It does not seem possible to solve for $\lambda_s(s)$ (or the fee $\phi(s,\omega) = \exp(u_m(s,\omega))$) analytically. Nor indeed is it even clear that a solution which satisfies this first-order condition would be solution to the problem because the manager's objective is not concave in general so the the first-order approach is not valid. To see this note that for a fixed reporting strategy (for example, when the correct signal is known in the second-best), the double integral in the manager's objective function (13) is affine (linear plus a constant), so that convexity of the cost function implies that the overall objective is concave. However, in the third-best the manager can vary the reporting strategy. In this case, the maximum across reporting strategies of the double integral is the maximum of affine functions and is therefore convex. In this case, 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.[8] For example, in the limiting case of a proportional cost function, the objective is convex in effort (once we have optimized 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. Numerical examples presented in an earlier draft confirm that there are cases in which the first-order approach works and cases in which the first-order approach fails.

What does seem clear from the first-order condition above is that none of the simple contractual forms we are familiar with from theory or from real-world contracts would solve this problem. Numerical results (included in some detail in an early unpublished version) suggest that compared to the second-best contract the third-best contract must provide extra rewards for reporting more extreme signals. The intuition for this is straightforward. In order to induce effort the manager is overexposed to the risk of the signal as in the bonus of the second-best contract. However a manager who may misreport will tend to try and report a signal which is too conservative in order to try and reduce this risk exposure. 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.

Conceptually, the third-best model seems more compelling than the first- and second-best models because effort is probably not contractible and the manager's beliefs after exerting effort are not publicly verifiable. Nonetheless, explicit contracts observed in practice seem to look like our first- and second-best solutions:

[8] Matters are actually somewhat more subtle than what is described in the text since the utility levels in the double integral are endogenous to the investor's choice problem. Therefore, we cannot specify a priori a level of convexity in the cost function $c(\cdot)$ large enough to ensure the manager's objective function is concave, since concavity still depends on the investor's specification of utility payoffs.

fees based on a proportion of assets under management (as in the first-best), with or without additional compensation based on performance relative a benchmark (as in the second-best), are common, while explicit contracts that compensate directly for taking more extreme positions (as in our third-best solution) are not. There are several possible explanations for this apparent inconsistency between theory and practice. All of these explanations are outside the scope of our model, and we do not have strong views about which explanation is most accurate.

One possible explanation is that the first- and second-best problems are better representations of the underlying economic problem, perhaps because there is a mechanism outside the model for handling the problem of truthful reporting or closet indexing. For example, perhaps site visits to the manager and examination of the records assure the investor that the manager is investing as intended. Or, it could be a reputation effect: we do observe monitoring for closet indexing in the hiring of managers, and this monitoring may produce a reputation-based incentive for taking the requisite risk. In general, there seems to be no reason to expect that alternative mechanisms will generate the correct incentives.

A second possible explanation is that there is some psychological reason that people do not behave as in the third-best. For example, perhaps investors and managers do not realize there is an incentive problem. Or, managers may like the feeling of honestly mapping the information into the signal while promises of effort are necessarily vague and underexertion of effort may be easier to rationalize. These explanations are not very useful as theory, since having this kind of explanation does not seem to put any restriction on behavior.

A third possible explanation (not inconsistent with the other two) is that perhaps the incentive to misreport in practice is small. In our numerical results, several choices of parameter values consistent with the first-order approach generated third-order solutions that were close numerically to the second-best, especially for parameter values for which the manager's information is not very informative about the market state (consistent with our priors that the market is not too inefficient).

At this point, we do not have a favorite explanation (or even a comprehensive list of explanations) for the apparent inconsistency between our theoretical sensibilities (which suggest the third-best is the "correct" analysis) and practice (which looks a lot like our first- and second-best solutions). We hope that future models or empirical work improve our understanding of this issue.

III 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 bonuse 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.

We have only just started to tap the potential of this framework to tell us about agency problems in portfolio management. Although some of the general results extend to stock selection models as well as the market timing examples given in this paper, it would be interesting to see the exact form of the contracts for stock-pickers. Analyzing career concerns would be an interesting variant: in this case, the current client has to take as given the manager's incentives to demonstrate superior performance this period in order to attract new clients or achieve a larger wage next period. In this case, there is probably a limit to the extent to which the client can neutralize the impact of career concerns. It would also be interesting to consider problems in which the manager's utility function (as well as consumption) is bounded below, given that the actual economy has restrictions on indentured servitude. Rajan and Srivastava (2000) considers a simple model of delegated portfolio management with limited punishment. It would be interesting to see what limited punishment or career concerns would imply in our model.

In the model, we have obtained a lot of mileage from the transparent and frictionless markets assumption that allows us to look at an equivalent formulation in which the manager simply reports information and does not actually manage the money. However, there are aspects of performance (such as quality of execution) that are not handled adequately in this way. While institutions receive complete reports of which trades were made (and mutual fund performance reports can depend in this information in any necessary way as computed credibly by the custodian or consultant), even the full trade record combined with full quote and trade histories of each stock would not necessarily tell us what trading opportunities were available at each point in time. It would be useful to have a fuller exploration of when the reporting formulation is equivalent and of what happens otherwise. Another 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

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

Almazan, A., K. C. Brown, M. Carlson, and D. A. Chapman, 2001, "Why Constrain Your Mutual Fund Manager?," Working paper, University of Texas at Austin.

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, 1–25.

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

Demski, J. S., and D. E. M. Sappington, 1987, “Delegated Expertise,” *Journal of Accounting Research*, 25, 68–89.

Dybvig, P. H., L. C. G. Rogers, and K. Back, 1999, “Portfolio Turnpikes,” *The Review of Financial Studies*, 12, 165–195.

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

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

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

Hart, O., and B. Holmström, 1987, “The Theory of Contracts,” in Truman F. Bewley (ed.), *Advances in Economic Theory – Fifth World Congress* . chap. 3, pp. 71–155, Cambridge University Press.

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

Holmström, B., 1984, “A Note on the First-Order Approach in Principal Agent Models,” mimeo.

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

Laffont, J.-J., and D. Martimort, 2002, *The Theory of Incentives: The Principal-Agent Model*, Princeton University Press, Princeton, NJ.

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

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

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

Rajan, U., and S. Srivastava, 2000, "Portfolio Delegation with Limited Liability," Working paper, GSIA, Carnegie Mellon University.

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.

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

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

Stole, L. A., 1993, "Lectures on the Theory of Contracts and Organizations, Chapters 1-2," Monograph in preparation. Available at http://gsblas.uchicago.edu/teaching/papers/lectures.pdf.

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

Sung, J., 1991, "Managerial compensation and corporate management under moral hazard," Ph.D. thesis, Washington University in Saint Louis.

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

Sung, J., 2003, "Optimal Contracts under Adverse Selection and Moral Hazard: a Continuous-Time Approach," Working paper, University of Illinois at Chicago.

Zender, J. F., 1988, "Asymmetric Information in Financial Markets," Ph.D. thesis, Yale.