Equilibrium prices in the presence of delegated portfolio management

Domenico Cuoco,1, Ron Kanielb,c,*

A R T I C L E   I N F O

Article history:
Received 1 February 2007
Received in revised form
8 April 2010
Accepted 1 June 2010
Available online 4 March 2011

JEL classification:
G11
G12
G19
G29

Keywords:
Asset pricing
Delegation
Agency
Money management
General equilibrium

A B S T R A C T

This paper analyzes the asset pricing implications of commonly used portfolio management contracts linking the compensation of fund managers to the excess return of the managed portfolio over a benchmark portfolio. The contract parameters, the extent of delegation, and equilibrium prices are all determined endogenously within the model we consider. Symmetric (fulcrum) performance fees distort the allocation of managed portfolios in a way that induces a significant and unambiguous positive effect on the prices of the assets included in the benchmark and a negative effect on the Sharpe ratios. Asymmetric performance fees have more complex effects on equilibrium prices and Sharpe ratios, with the signs of these effects fluctuating stochastically over time in response to variations in the funds' excess performance.

© 2011 Elsevier B.V. All rights reserved.

1. Introduction

In modern economies, a significant share of financial wealth is delegated to professional portfolio managers rather than managed directly by the owners, creating an agency relationship. In the U.S., as of 2004, mutual funds managed assets in excess of $8 trillion, hedge funds managed about $1 trillion, and pension funds more than $12 trillion. In other industrialized countries, the percentage of financial assets managed through portfolio managers is even larger than in the U.S. (see, e.g., Bank for International Settlements, 2003).

While the theoretical literature on optimal compensation of portfolio managers in dynamic settings points to contracts that are likely to have very complicated path dependencies,2 the industry practice seems to favor

---

2 A distinctive feature of the agency problem arising from portfolio management is that the agent’s actions (the investment strategy and
relatively simple compensation schemes that typically include a component that depends linearly on the value of the managed assets plus a component that is linearly or non-linearly related to the excess performance of the managed portfolio over a benchmark.

In 1970, the U.S. Congress amended the Investment Advisers Act of 1940 so as to allow contracts with registered investment companies to include performance-based compensation, provided that this compensation is of the “fulcrum” type, that is, provided that it includes penalties for underperforming the chosen benchmark that are symmetric to the bonuses for exceeding it. In 1985, the SEC approved the use of performance-based fees in contracts in which the client has either at least $500,000 under management or a net worth of at least $1 million. Performance-based fees were also approved by the Department of Labor in August 1986 for ERISA-governed pension funds. As of 2004, 50% of U.S. corporate pension funds with assets above $5 billion, 35% of all U.S. pension funds, and 9% of all U.S. mutual funds used performance-based fees. Furthermore, Brown, Harlow, and Starks (1996), Chevalier and Ellison (1997), and Sirri and Tufano (1998) show that, even when mutual fund managers do not receive explicit incentive fees, an implicit non-linear performance-based compensation still arises with periodic proportional fees as a result of the fact that the net investment flow into mutual funds varies in a convex fashion as a function of recent performance.

Given the size of the portfolio management industry, studying the implications of this delegation and of the fee structures commonly used in the industry on equilibrium asset prices appears to be a critical task. The importance of models addressing the implications of agency for asset pricing was emphasized by Allen (2001): “In the standard asset pricing paradigm it is assumed investors directly invest their wealth in markets. While this was an appropriate assumption for the U.S. in the 1950 when individuals directly held over 90% of corporate equities, or even in 1970 when the figure was 68%, it has become increasingly less appropriate as time has progressed [...] For actively managed funds, the people that make the ultimate investment decisions are not the owners. If the people making the investment decisions obtain a high reward when things go well and a limited penalty if they go badly they will be willing to pay more than the discounted cash flow for an asset. This is the type of incentive scheme that many financial institutions give to investment managers.”


We complement this literature by considering a different problem. As in the literature on optimal behavior of portfolio managers, we take the parametric class of contracts as exogenously given, motivated by commonly observed fee structures. However, we carry the analysis beyond partial equilibrium by studying how the behavior of portfolio managers affects equilibrium prices when the extent of portfolio delegation and the parameters of the management contract are all determined endogenously.

A first step in studying the implications of delegated portfolio management on asset returns was made by Brennan (1993), who considered a static mean–variance economy with two types of investors: individual investors (assumed to be standard mean–variance optimizers) and “agency investors” (assumed to be concerned with the mean and the variance of the difference between the return on their portfolio and the return on a benchmark portfolio). Equilibrium expected returns were shown to be characterized by a two-factor model, with the two factors being the market and the benchmark portfolio. Closely related mean–variance models have appeared in Gómez and Zapatero (2003) and Cornell and Roll (2005).5

To our knowledge, the only general equilibrium analyses of portfolio delegation in dynamic settings are in two recent papers by Kapur and Timmermann (2005) and Arora, Ju, and Ou-Yang (2006). Kapur and Timmermann consider a restricted version of our model with mean–variance preferences, normal returns, and fulcrum performance fees, while Arora, Ju, and Ou-Yang assume CARA utilities and normal dividends and do not endogenize the extent of portfolio delegation: as a result of these assumptions, fulcrum performance fees are optimal in their

(footnote continued)
possibly the effort spent acquiring information about securities’ returns) affect both the drift and the volatility of the relevant state variable (the value of the managed portfolio), although realistically the drift and the volatility cannot be chosen independently. This makes the problem significantly more complex than the one considered in the classic paper by Holstrom and Milgrom (1987) and its extensions. With a couple of exceptions, as noted by Stracca (2006) in his recent survey of the literature on delegated portfolio management, “the literature has reached more negative rather than constructive results, and the search for an optimal contract has proved to be inconclusive even in the most simple settings.”

1 The use is concentrated in larger funds: the percentages of assets under management controlled by mutual funds charging performance fees out of funds managing assets of $0.25–1 billion, $1–5 billion, $5–10 billion, and above $10 billion were 2.8%, 4.4%, 9.2%, and 14.2%, respectively (data obtained from Greenwich Associates and the Investment Company Institute).

2 Lynch and Musto (2003) and Berk and Green (2004) provide models in which this convex relationship between flows and performance arises endogenously.

model. More importantly, both papers consider settings with a single risky asset. A key shortcoming of models with a single risky asset (or of static models) is that they are unable to capture the shifting risk incentives of portfolio managers receiving implicit or explicit performance fees (and hence the impact of these incentives on portfolio choices and equilibrium prices), as extensively described in both the theoretical and empirical literature.

In contrast to the papers mentioned above, we study the asset pricing implications of delegated portfolio management in the context of a dynamic (continuous-time) model with multiple risky assets and endogenous portfolio delegation. Specifically, we consider an economy with a continuum of three types of agents: “active investors,” “fund investors,” and “fund managers.” Active investors, who trade on their own account, choose a dynamic trading strategy so as to maximize the expected utility of the terminal value of their portfolio. Fund investors, who implicitly face higher trading or information costs, invest in equities only through mutual funds: therefore, their investment choices are limited to how much to delegate to fund managers, with the rest of their portfolio being invested in riskless assets. Fund managers, who are assumed not to have any private wealth, select a dynamic trading strategy so as to maximize the expected utility of their compensation.

The compensation contracts we consider are restricted to a parametric structure that replicates the contracts typically observed in practice, consisting of a combination of the following components: a flat fee, a proportional fee depending on the total value of the assets under management, and a performance fee depending in a piecewise-linear manner on the differential between the return of the managed portfolio and that of a benchmark portfolio.

Departing from the traditional formulation of principal-agent problems, we assume that individual fund investors are unable to make “take it or leave it” contract offers to fund managers and thus to extract the entire surplus from the agency relation: instead, we assume that the market for fund investors is competitive, so that individual investors take the fee structure as given when deciding what fraction of their wealth to delegate.

Similarly, we assume competition on the market for portfolio managers. We analyze in detail the impact on equilibrium price dynamics of commonly observed contracts, with commonly observed contract parameters. However, we also allow contract parameters to be selected in our model so that they are constrained Pareto efficient, i.e., so that there is no other contract within our parametric class that provides both fund investors and fund managers with higher welfares.

As shown in Section 6, even when fund investors and fund managers have identical preferences, the principle of preference similarity (Ross, 1973) does not apply in our setting and asymmetric performance contracts Pareto-dominate purely proportional contracts (as well as fulcrum performance contracts): intuitively, convex performance fees are a way to incentivize fund managers to select portfolio strategies having higher overall stock allocations, benefiting fund investors who have direct access to riskless investment opportunities. Because of this incentive role of performance fees, the optimal benchmark typically differs from the market portfolio.

Portfolio delegation can have a substantial impact on equilibrium prices. With fulcrum fees, the presence of a penalty for underperforming the benchmark portfolio leads risk-averse fund managers to be overinvested in the stocks included in the benchmark portfolio and underinvested in the stocks excluded from this portfolio. The bias of managed portfolios in favor of the stocks included in the benchmark portfolio results in the equilibrium expected returns and Sharpe ratios of these stocks being lower than those of comparable stocks not in the benchmark and in their price/dividend ratios being higher. At the same time, stocks in the benchmark portfolio tend to have higher equilibrium volatilities than those of comparable stocks not in the benchmark: this is due to the fact that, as the price of benchmark stocks starts to rise, the tilt of managed portfolios toward these stocks increases, lowering their equilibrium price/dividend ratios and hence moderating the price increase. Therefore, consistent with empirical evidence, our model implies that, if fund managers are mostly compensated with fulcrum fees, a change in the composition of widely used benchmark portfolios (such as the S&P 500 portfolio) should be accompanied by a permanent increase in the prices and volatilities of the stocks being added to the index and a corresponding permanent decrease in the prices and volatilities of the stocks being dropped from the index. With asymmetric performance fees, the signs of these changes become ambiguous, depending on the current average excess performance of managed portfolios relative to the benchmark.

The remainder of the paper is organized as follows. The economic setup is described in Section 2. Section 3 provides a characterization of the optimal investment strategies. Section 4 focuses on the characterization of equilibria. Section 5 provides a detailed calibrated numerical analysis of equilibrium under asymmetric and fulcrum performance fees. Section 6 discusses the

---

6 In the model of Kapur and Timmermann (2005), performance fees do not dominate fees depending only on the terminal value of the assets under management.

7 Clearly, in the presence of performance fees, tracking error volatility directly affects the reward of portfolio managers and this volatility can be dynamically controlled by varying the composition of the managed portfolio.

8 While our framework allows for both “fulcrum” and “asymmetric” performance fees, it does not allow for “high water mark” fees (occasionally used by hedge funds and discussed by Goetzmann, Ingersoll, and Ross, 2003) in which the benchmark equals the lagged maximum value of the managed portfolio.

9 As noted by Das and Sundaram (2002), the existence of regulation, such as the Investment Advisors Act of 1940, meant to protect fund investors through restrictions on the allowable compensation contracts can be viewed as tacit recognition that these investors do not dictate the form of the compensation contracts.

10 This is in contrast to the existing mean-variance equilibrium models of portfolio delegation, in which the optimal benchmark is the market portfolio and performance fees are dominated by linear contracts.
optimality of performance contracts in our model. Section 7 concludes. The Appendix contains all the proofs.

2. The economy

We consider a continuous-time economy on the finite time span [0, T], modeled as follows.

Securities: The investment opportunities are represented by a riskless bond and two risky stocks (or stock portfolios). The bond is a claim to a riskless payoff B > 0. The interest rate is normalized to zero (i.e., the bond price is normalized to B).

Stock j (j = 1, 2) is a claim to an exogenous liquidation dividend $D_j^t$ at time $t$, where

$$D_j^t = D_0^j + \int_0^t \mu_j^i(D_j^s, s) \, ds + \int_0^t \sigma_j^i(D_j^s, s) \, dw_j^s,$$

for some functions $\mu_j^i, \sigma_j^i$ satisfying appropriate Lipschitz and growth conditions and two Brownian motions $w^j$ with instantaneous correlation coefficient $\rho \in (-1, 1)$.

Since dividends are paid only at the terminal date $t$, without loss of generality we take $\mu_j^i = 0$, so that $D_t$ can be interpreted as the conditional expectation at time $t$ of stock $j$’s liquidation dividend.

We let $D_t = (B, D_1, D_2)$ denote the vector of terminal asset payoffs and denote by $S_t = (B, S_1, S_2)$ the vector of asset prices at time $t$. The aggregate supply of each asset is normalized to one share and we denote by $\bar{\theta} = (1, 1, 1)$ the aggregate supply vector.

Security trading takes place continuously. A dynamic trading strategy is a three-dimensional process $\theta$, specifying the number of shares held of each of the traded securities, such that the corresponding wealth process $W = \theta \cdot S$ satisfies the dynamic budget constraint

$$W_t = W_0 + \int_0^t \theta_s \cdot dS_s,$$

and $W_t \geq 0$ for all $t \in [0, T]$. We denote by $\mathcal{S}$ the set of dynamic trading strategies.\(^{11}\)

Agents: The economy is populated by three types of agents: active investors, fund investors, and fund managers. We assume that there is a continuum of agents of each type and denote by $\lambda \in (0, 1)$ the mass of fund investors in the economy and by $1 - \lambda$ the mass of active investors. Without loss of generality, we assume that the mass of fund managers also equals $\lambda$: this is merely a normalization, since the aggregate wealth managed by fund managers is determined endogenously by the portfolio choices of fund investors, as described later in this section.

Active investors receive an endowment of one share of each traded asset at time 0, so that their initial wealth equals $W_0^a = \bar{\theta} \cdot S_0$. They choose a dynamic trading strategy $\theta^a \in \mathcal{S}$ so as to maximize the expected utility $\mathbb{E}[u(W_T^a)]$ of the terminal value of their portfolio $W_T^a = \theta_T^a \cdot S_T$, taking equilibrium prices as given.

Fund investors also receive an endowment of one share of each asset, so that their initial wealth is $W_0^f = \bar{\theta} \cdot S_0$. However, because of higher trading or information costs (which we do not model explicitly), they do not hold stocks directly and instead delegate the choice of a dynamic trading strategy to fund managers: at time 0 they simply choose to invest an amount $\theta_0^f B \leq W_0^f$ in the riskless asset and invest the rest of their wealth in mutual funds.\(^{12}\)

Fund managers receive an initial endowment $W_0^m = W_0^f - \theta_0^f B$ from fund investors, which they then manage on the fund investors’ behalf by selecting a dynamic trading strategy $\theta^m \in \mathcal{S}$. For this, they are compensated at time $T$ with a management fee $F_T$ which is a function of the terminal value of the fund portfolio, $W_T^m = \theta_T^m \cdot S_T$, and of the terminal value of a given benchmark portfolio $W_T^g = \theta_T^g \cdot S_T$, where $\theta^b \in \mathcal{S}$.\(^{13}\) Specifically, we assume that

$$F_T = F(W_T^m, W_T^g) = \alpha + \beta W_T^m - \gamma_1 W_T^g (W_T^m - W_T^g) - \gamma_2 W_T^m (W_T^m - W_T^g),$$

$$= \alpha + \beta W_T^m - \gamma_1 W_T^g (W_T^m - \delta W_T^g) + \gamma_2 W_T^m (W_T^m - \delta W_T^g),$$

where $\alpha$, $\beta$, $\gamma_1$, and $\gamma_2$ are given parameters, $\delta = W_0^m / W_0^g$ and $x^+ = \max(0, x)$ (respectively, $x^- = \max(0, -x)$) denotes the positive part (respectively, the negative part) of the real number $x$.

Thus, the fund managers’ compensation at time $T$ can consist of four components: a load fee $\alpha$ which is independent of the managers’ performance, a proportional fee $\beta W_T^m$ which depends on the terminal value of the fund portfolio, a performance bonus $\gamma_2 (W_T^m - \delta W_T^g)^+$ which depends on the performance of the managed portfolio relative to that of the benchmark portfolio, and an underperformance penalty $\gamma_1 (W_T^m - \delta W_T^g)^-$. We assume that $\alpha \geq 0$, $\beta \geq 0$, $\gamma_2 \geq \gamma_1 \geq 0$, and $\beta + \gamma_2 > 0$, so that the fund managers’ compensation $F$ is an increasing and convex function of the terminal value of the fund portfolio and a decreasing function of the terminal value of the benchmark portfolio.\(^{14}\) In addition, we assume that $\alpha + \beta > 0$, so that the fund managers always have at least

---

\(^{11}\) Implicit in the definition of $\mathcal{S}$ is the requirement that the stochastic integral in Eq. (2) is well defined.

\(^{12}\) Because fund investors in our model do not trade dynamically and are not assumed to know the return distribution of the individual assets (only knowledge of the return distribution of the fund they invest in being assumed), their behavior could be rationalized with a combination of trading and information costs. Of course, in reality there is a wide range of investors with different trading and information costs: the assumption that investors are either “active,” with full information and costless access to stock trading, or “passive,” requiring the intermediation of mutual funds to obtain exposure to risky assets, is clearly a simplifying one and is made for tractability.

\(^{13}\) Letting the benchmark be the terminal value of a dynamic (not necessarily buy-and-hold) trading strategy allows for the possibility of changes in the composition of the benchmark portfolio. The process governing the dynamics of $\theta^b$ is known to market participants.

\(^{14}\) We do not necessarily require that $\beta + \gamma_2 < 1$, i.e., that $(\partial / \partial W_T^g) F(W_T^m, W_T^g) < 1$. In order to guarantee the existence of an equilibrium, we will however later impose a condition that implies that the optimal terminal wealth of fund investors is increasing in aggregate wealth (see Eq. (23)).
one feasible investment strategy (buying the benchmark portfolio) that yields a strictly positive fee.

When $\gamma_1 = \gamma_2$, the performance-related component of the managers' compensation is linear in the excess return of the fund over the benchmark. These types of fees are known as fulcrum performance fees. As noted in the Introduction, the 1970 Amendment of the Investment Advisers Act of 1940 restricts mutual funds' performance fees to be of the fulcrum type. Hedge funds' performance fees are not subject to the same restriction, and for these funds asymmetric performance fees with $\gamma_1 = 0$ and $\gamma_2 > 0$ are the norm.

Fund managers are assumed not to have any private wealth. They therefore act so as to maximize the expected utility

$$E[u^m(F(W_m^0, W_M^T))]$$

of their management fees, while taking equilibrium prices and the investment choices of fund investors as given. Similarly, fund investors select the amount of portfolio delegation $W_0^d - \Omega^d_B$ so as to maximize the expected utility of their terminal wealth, while taking the equilibrium net-of-fees rate of return on mutual funds

$$R_T = \frac{W_T^p - F(W_T^p, W_M^T)}{W_0^p},$$

as given, subject to the constraint $W_0^d - \Omega^d_B \geq 0$.

We assume throughout that

$$u^a(W) = u^m(W) = u^e(W) = u(W) = \frac{W^{1-c}}{1-c}$$

for some $c > 0, c \neq 1$.

Equilibrium: An equilibrium for the above economy is a price process $S$ for the traded assets and a set $[\theta^a, \theta^m, \theta^e] \in \Theta \times \Theta \times \Theta$ of trading strategies such that:

1. the strategy $\theta^a$ is optimal for the active investors given the equilibrium stock prices;
2. the strategy $\theta^m$ is optimal for the fund managers given the equilibrium stock prices and the fund investors' choice of $\theta^a$;
3. the choice $\theta^e$ is optimal for the fund investors given the equilibrium stock prices and the equilibrium net-of-fees funds’ return;
4. the security markets clear:

$$1 - \lambda)\theta^a_t + \lambda \theta^m_t + \lambda(\theta^e_t, 0, 0) = \bar{S} \quad \text{for all } t \in [0, T].$$

3. Optimal investment strategies

Since active investors and fund managers face a dynamically complete market, we use the martingale approach of Cox and Huang (1989) to characterize their optimal investment strategies.

3.1. Active investors

Given the equilibrium state-price density $\pi_T$ at time $T$, the optimal investment problem for the active investors amounts to the choice of a non-negative random variable $W_T^a$ (representing the terminal value of their portfolios) solving the static problem

$$\max_{W_T^a \geq 0} E[u(W_T^a)]$$

s.t. $E[\pi_T|W_T^a] \leq W_0^a$.

This implies

$$W_T^a = g^d(\psi^a \pi_T),$$

where $g^d(y) = y^{1/c}$ denotes the inverse marginal utility function and $\psi^a$ is a Lagrangian multiplier solving

$$E[\pi_Tg^e(\psi^a \pi_T)] = W_0^a = \bar{S} \cdot S_0.$$

3.2. Fund managers

Given the equilibrium state-price density $\pi_T$ and the allocation to mutual funds by fund investors $W_m^0 = \bar{S} \cdot S_0 - \theta^d_B$, the optimal investment problem for the fund managers amounts to the choice of a non-negative random variable $W_T^m$ (representing the terminal value of the funds' portfolios) solving the static problem

$$\max_{W_T^m \geq 0} E[u(W_T^m|W_T^a)]$$

s.t. $E[\pi_T|W_T^m] \leq W_0^m$.

The added complexity in this case arises from the fact that, unless $\gamma_1 = \gamma_2$, the fund managers' indirect utility function over the terminal portfolio value, $u(F(W, W^b))$ is neither concave nor differentiable in its first argument at the point $W = \delta W^b$ (the critical value at which the performance of the fund’s portfolio equals that of the benchmark): see Fig. 1. This is a consequence of the convexity and lack of differentiability of the fee function $F(W, W^b)$. Therefore, the optimal choice $W_T^m$ is not necessarily unique and it does not necessarily satisfy the usual first-order condition. Moreover, the non-negativity constraint can be binding in this case (if $x > 0$).

On the other hand, the fact that managers have infinite marginal utility at zero wealth implies that the optimal investment strategy must guarantee that a strictly positive fee is collected at time $T$, i.e., that $W_T^m > W(W_T^a)$, where

$$W(W^b) = \inf(W \geq 0 : F(W, W^b) \geq 0)$$

$$= \left\{ \begin{array}{ll} \frac{\gamma_1 \delta W^b - \beta \bar{S}}{\beta + \gamma_1} & \text{if } \beta + \gamma_1 \neq 0, \\ 0 & \text{otherwise.} \end{array} \right.$$

Since for any $W^b > 0$ the function $u(F(\cdot, W^b))$ is piecewise concave and piecewise continuously differentiable on the interval $[W(W^b), \infty)$, we can follow Shapley and Shubik (1966), Aumann and Perles (1965), and Carpenter (2000) in constructing the concavification $v(\cdot, W^b)$ of $u(F(\cdot, W^b))$ (that is, the smallest concave function $v$ satisfying $v(W, W^b) \geq u(F(W, W^b))$ for all $W \geq W(W^b)$ and then
exist unique numbers $W_1(W^b)$ and $W_2(W^b)$ with $W(W^b) = W_1(W^b) < \delta W^b < W_2(W^b)$ such that
\[
u(F(W_1(W^b),W^b)) = u(F(W_2(W^b),W^b)) + u'(F(W_2(W^b),W^b))(\beta + \gamma_2)(W_1(W^b) - W_2(W^b))
\]
and
\[
u'(F(W_1(W^b),W^b))(\beta + \gamma_1) \leq u'(F(W_2(W^b),W^b))(\beta + \gamma_2).
\]
with equality if $W_1(W^b) > 0$. In particular, letting $\eta = ((\beta + \gamma_2)/(\beta + \gamma_1))^{1/\gamma}$,
\[
W_1(W^b) = \left(1 - \frac{\eta}{\gamma_1}\right) \frac{\Delta W^b}{\beta + \gamma_2} - \eta \left(1 - \frac{\eta}{\gamma_1}\right) \frac{\Delta W^b}{\beta + \gamma_2} + \frac{\eta - 1}{\gamma - 1} \frac{\Delta W^b}{\beta + \gamma_2}
\]
if $\beta + \gamma_1 \neq 0$, and $W_1(W^b) = 0$ otherwise. Moreover,
\[
W_2(W^b) = W_1(W^b) + \frac{1}{\gamma} \left(\frac{\Delta W^b}{\beta + \gamma_1} - \frac{\Delta W^b}{\beta + \gamma_2}\right)
\]
if $W_1(W^b) > 0$.

**Lemma 2.** Let $W_1(W^b)$ and $W_2(W^b)$ be as in Lemma 1 if $\gamma_1 \neq \gamma_2$ and let $W_1(W^b) = W_2(W^b) = \delta W^b$ otherwise. Also, let $A(W^b) = \{W(W^b), W_1(W^b)\} \cup \{W_2(W^b), \infty\}$.

Then the function
\[
\nu(W,W^b) = \begin{cases} u(F(W,W^b)) & \text{if } W^m \in A(W^b), \\ u(F(W_1(W^b),W^b)) + u'(F(W_2(W^b),W^b))(\beta + \gamma_2) \times (W - W_1(W^b)) & \text{otherwise} \end{cases}
\]
is the smallest concave function on $[W(W^b), \infty)$ satisfying $\nu(W,W^b) \geq u(F(W,W^b)) \quad \forall W \in [W, \infty)$.

Moreover, $\nu(W,W^b)$ is continuously differentiable in $W$.

The construction of the concavification $\nu(W,W^b)$ of $u(F(W,W^b))$ is illustrated in **Fig. 1**. Since $u(F(W,W^b))$ is not concave at $W = \delta W^b$ when $\gamma_1 \neq \gamma_2$, the idea is to replace $u(F(W,W^b))$ with a linear function over an interval $(W_1(W^b), W_2(W^b))$ bracketing $\delta W^b$ (a linear function is the smallest concave function between two points). The two points $W_1(W^b)$ and $W_2(W^b)$ are uniquely determined by the requirement that the resulting concavified function $\nu(W,W^b)$ be continuously differentiable and coincide with $u(F(W,W^b))$ at the endpoints of the interval. When $\beta + \gamma_1 = 0$, this requirement necessarily implies $W_1(W^b) = 0$ (as shown in Panel (a) of **Fig. 1**): this is the case studied by Carpenter (2000). On the other hand, when $\beta + \gamma_1 > 0$, both the case $W_1(W^b) = 0$ and the case $W_1(W^b) > 0$ are possible (as illustrated in Panels (b) and (c) of **Fig. 1**), depending on the sign of the expression in Eq. (13).

Letting
\[
v_W^1(y,W^b) = \{W \in [W(W^b), \infty) : \nu_W(W,W^b) = y\}
\]
be the inverse marginal utility correspondence of the concavified utility $\nu$, we then obtain the following characterization of the fund managers’ optimal investment policies.
Proposition 1. A policy \( W^m \) is optimal for the fund managers if and only if: (i) it satisfies the budget constraint in Eq. (8) as an equality and (ii) there exists a Lagrangian multiplier \( \psi > 0 \) such that

\[
W^m_0 \in \psi^{-1}(\psi^m \pi_T, W^b_T) \cap A(W^b_T) \quad \text{if} \quad W^m_0 > 0
\]

and

\[
v_W(W^m_0, W^b_T) \leq \psi^m \pi_T \quad \text{if} \quad W^m_0 = 0.
\]

To understand the characterization of optimal policies in Proposition 1, consider first the concavified problem in Eq. (10). The standard Kuhn-Tucker conditions for this problem are the same as the conditions of Proposition 1, Eq. (10). It can be easily verified from the definitions of the aggregate state-price density, any wealth level between \( Wm \) and \( Wb \) is optimal for the concavified problem in Eq. (10) and takes value \( Wm \) for any feasible policy \( W_T \), where the first inequality follows from the fact that \( W_T \) is optimal for the problem in Eq. (10), and the last equality follows from the fact that \( u(f(W, W'_T)) = v(W, W'_T) \) for \( W \in A(W^m) \).

Since managers are indifferent between selecting \( W_1(W'_T) \) or \( W_2(W'_T) \) as the terminal value of the fund's portfolio, the scaled state-price density equals \( v(W_2(W'_T), W^b_T) \), we allow them to independently randomize between \( W_1(W'_T) \) and \( W_2(W'_T) \). In this case, the optimal policy to be a lottery that selects \( W_1(W'_T) \) with some probability \( p \) and \( W_2(W'_T) \) with probability \( (1-p) \), and denote such a lottery by

\[
p \cdot W_1(W'_T) + (1-p) \cdot W_2(W'_T).
\]

Of course, \( p \) can depend on the information available to the agents at time \( T \), i.e., be a random variable. Using the expression for \( v \) in Lemma 2 and Proposition 1, we then obtain the following result.

Theorem 1. Let \( P_T \) be a random variable taking values in \([0,1]\) and let

\[
g^m(y, W^a_T, p) = \begin{cases} 
-\frac{y}{(\beta + \gamma_2)^2} & \text{if } y < v_W(W_2(W^a_T), W^b_T) \\
\frac{y}{(\beta + \gamma_2)^2} & \text{if } y > v_W(W_2(W^a_T), W^b_T) \\
0 & \text{otherwise}
\end{cases}
\]

Then the policy

\[
W^m_0 = g^m(\psi^m \pi_T, W^b_T, P_T),
\]

where \( \psi^m \) is a Lagrangian multiplier solving

\[
E[\pi_T g^m(\psi^m \pi_T, W^b_T, P_T)] = W^m_0 = T \cdot S_0 - \theta_0 B
\]

is optimal for the fund managers.

Given the existence of a continuum of mutual funds, the aggregate terminal value of the funds' portfolios equals

\[
\dot{z}^m(\psi^m \pi_T, W^b_T, P_T),
\]

where

\[
g^m(\psi^m \pi_T, W^b_T, P_T) = \begin{cases} 
-\frac{y}{(\beta + \gamma_2)^2} & \text{if } y < v_W(W_2(W^a_T), W^b_T) \\
\frac{y}{(\beta + \gamma_2)^2} & \text{if } y > v_W(W_2(W^a_T), W^b_T) \\
0 & \text{otherwise}
\end{cases}
\]

Since the fund managers' indirect utility functions are nonconvex, our assumption of a continuum of mutual funds is critical to ensure the convexity of the aggregate preferred sets and hence the existence of an equilibrium. Aumann (1966) was the first to prove that, with a continuum of agents, the existence of an equilibrium could be assured even without the usual assumption of convex preferences.
As it will become clear in the next section, this in turn implies that the randomizing probabilities \( P \) are uniquely determined in equilibrium by market clearing.

### 3.3. Fund investors

Fund investors select the allocation \( \theta'_{0} \) to bonds so as to maximize the expected utility of their terminal wealth, while taking the equilibrium net-of-fees rate of return on mutual funds \( R_{t} \) (defined in Eq. (4)) as given. That is, they solve

\[
\begin{align*}
\max_{\theta_{0} \in \mathbb{R}} & \quad \mathbb{E}[u'(\theta_{0}B + (W'_{0} - \theta_{0}B)R_{T})] \\
\text{s.t.} & \quad \theta_{0}B \leq W'_{0}.
\end{align*}
\]

Since this is a strictly concave static maximization problem, the optimal choice \( \theta'_{0} \) satisfies the standard Kuhn-Tucker conditions

\[
\begin{align*}
\frac{\partial}{\partial \theta_{0}} \{u'(\theta'_{0}B + (W'_{0} - \theta'_{0}B)R_{T})^{-1}(1-R_{T})\} - \psi' = 0, \\
\psi'(W'_{0} - \theta'_{0}B) = 0
\end{align*}
\]

for some Lagrangian multiplier \( \psi' \geq 0 \).

### 4. Equilibrium: characterization

In equilibrium the terminal stock prices equal the liquidation dividends, i.e., \( S_{T} = D_{T} \). Multiplying the market-clearing condition (5) at time \( T \) by \( D_{T} \) and using Eqs. (6) and (17), it then follows that the equilibrium state-price density \( \pi_{T} \) must solve

\[
(1-\lambda)\mathbb{g}^{a}(\psi_{m}\pi_{T}) + \lambda\mathbb{g}^{m}(\psi_{m}\pi_{T}, W'_{T}, P_{T}) + \lambda \theta'_{0}B = \overline{T} \cdot D_{T}.
\]

This shows that the equilibrium state-price density \( \pi_{T} \) and the randomizing probabilities \( P_{T} \) must be deterministic functions of \( D_{T} \) and \( W'_{T} \). Letting

\[
\Pi(D_{T}, W'_{T}) = \psi_{m}^{a} \pi_{T}
\]

and

\[
\varphi = (\psi_{m}/\psi_{m}^{a})^{-1/c},
\]

substituting in (19), recalling the definitions of \( \mathbb{g}^{a} \) and \( \mathbb{g}^{m} \), and rearranging gives

\[
\begin{align*}
\overline{T} \cdot D_{T} - \lambda \theta'_{0}B - (1-\lambda)\varphi \Pi(D_{T}, W'_{T})^{-1/c} &= \begin{cases} \\
\frac{\lambda}{(\beta + \gamma_{2})^{-1/c} - \frac{\lambda}{\beta + \gamma_{2}}} & \text{if } \Pi(D_{T}, W'_{T}) = \varphi W(D_{T}, W'_{T}), \\
\lambda \Psi_{W}(W_{T} + \frac{(1-\lambda)\theta'_{0}B}{\beta + \gamma_{2}}) & \text{if } \Pi(D_{T}, W'_{T}) = \varphi W(D_{T}, W'_{T}), \\
0 & \text{if } \Pi(D_{T}, W'_{T}) > \varphi W(D_{T}, W'_{T}) \text{ and } \beta + \gamma_{2} \neq 0, \\
\end{cases}
\end{align*}
\]

Solving the above equation shows that the scaled state-price density \( \Pi(D_{T}, W'_{T}) \) can take on one of four different functional forms (corresponding to the four different cases on the right-hand side of Eq. (20)),

\[
\begin{align*}
\Pi_{1}(D_{T}, W'_{T}) &= \left( \frac{(\beta + \gamma_{2})\overline{T} \cdot D_{T} - \lambda \theta'_{0}B + \lambda (x - \gamma_{2}) \delta W'_{T}}{(\beta + \gamma_{2})^{-1/c} + (1-\lambda)\phi(\beta + \gamma_{2})} \right)^{-\mathbb{c}}, \\
\Pi_{2}(D_{T}, W'_{T}) &= v_{W}(W(D_{T}, W'_{T}), W'_{T}) = F(W(D_{T}, W'_{T}), W'_{T})^{-\mathbb{c}}(\beta + \gamma_{2}), \\
\Pi_{3}(D_{T}, W'_{T}) &= \left( \frac{(\beta + \gamma_{2})\overline{T} \cdot D_{T} - \lambda \theta'_{0}B + \lambda (x - \gamma_{1}) \delta W'_{T}}{(\beta + \gamma_{1})^{-1/c} + (1-\lambda)\phi(\beta + \gamma_{1})} \right)^{-\mathbb{c}}, \\
\Pi_{4}(D_{T}, W'_{T}) &= \left( \frac{\overline{T} \cdot D_{T} - \lambda \theta'_{0}B}{(1-\lambda)\phi} \right)^{-\mathbb{c}}.
\end{align*}
\]

It then follows from the inequality conditions in Eq. (20) and the fact that Eq. (16) implies

\[
g^{m}(\Pi_{1}(D_{T}, W'_{T}), W'_{T}, P_{T}) > W_{2}(W'_{T}),
\]

\[
0 < g^{m}(\Pi_{2}(D_{T}, W'_{T}), W'_{T}, P_{T}) < W_{1}(W'_{T})
\]

that the scaled equilibrium state-price density is given by

\[
\Pi(D_{T}, W'_{T}) = \begin{cases} \\
\Pi_{1}(D_{T}, W'_{T}) & \text{if } \Pi_{1}(D_{T}, W'_{T}) < \Pi_{2}(D_{T}, W'_{T}), \\
\max(\Pi_{2}(D_{T}, W'_{T}), \Pi_{3}(D_{T}, W'_{T})), & \text{if } \Pi_{1}(D_{T}, W'_{T}) \geq \Pi_{2}(D_{T}, W'_{T}), \\
\Pi_{4}(D_{T}, W'_{T}), \Pi_{4}(D_{T}, W'_{T}) & \Pi_{4}(D_{T}, W'_{T}) > 0 \text{ and } \beta + \gamma_{1} \neq 0, \\
\max(\Pi_{2}(D_{T}, W'_{T}), \Pi_{4}(D_{T}, W'_{T})) & \text{otherwise}.
\end{cases}
\]

In addition, the equality in Eq. (20) corresponding to the case

\[
\Pi(D_{T}, W'_{T}) = v_{W}(W(D_{T}, W'_{T}), W'_{T}) = \Pi_{2}(D_{T}, W'_{T})
\]

can be solved for the market-clearing randomizing probabilities \( P_{T} \), yielding

\[
P_{T} = \Pi(D_{T}, W'_{T})^{-1/c}
\]

\[
= \left( \frac{\overline{T} W_{2}(W'_{T}) - \lambda \theta'_{0}B + (1-\lambda)\phi \Pi_{2}(D_{T}, W'_{T})}{\lambda(W_{2}(W'_{T}) - W_{1}(W'_{T}))} \right).
\]

Eq. (21) provides an explicit expression for the scaled state-price density \( \Pi(D_{T}, W'_{T}) \) in terms of three yet-undetermined constants, \( \delta, \theta'_{0}, \) and \( \varphi. \) In turn, \( \delta = W_{m}/W_{0} \) is a known function of \( \theta'_{0} \) and the initial stock prices \( S_{T}^{b} \) and \( S_{T}^{a}. \)

It follows from the shape of the state-price density that the optimal consumption policies are piecewise-linear functions of the liquidation dividends. In order to guarantee the existence of an equilibrium, we require that the coefficients of these functions be positive, i.e., that increases in aggregate consumption be shared among the agents. Assuming that \( \overline{\mathbb{c}} \geq 0 \) (i.e., that the benchmark portfolio does not include short positions), this amounts to the following parameter restriction:

\[
(\beta + \gamma_{2})\overline{\gamma_{2}} - \lambda \overline{\gamma_{2}} \overline{\theta_{0}} > 0 \quad \text{for } j = 1, 2,
\]

where \( \overline{\mathbb{d}_{j}} \) (respectively, \( \overline{\mathbb{b}_{j}} \)) denotes the number of shares of stock \( j \) in the market (respectively, in the benchmark) portfolio.
The following theorem completes the characterization of equilibria by providing necessary and sufficient conditions for existence, together with an explicit procedure to determine the unknown constants \( \theta_0^f, \phi, S_0^j, \) and \( S_0^k \).

**Theorem 2.** Assume that the condition (23) is satisfied. Then an equilibrium exists if and only if there exist constants \((\theta_0^f, \phi, S_0^j, S_0^k)\) with \( \theta_0^f \leq (B + S_0^j + S_0^k)/B \) and \( \psi^f \geq 0 \) solving the system of equations

\[
\begin{align*}
E\left(\theta_0^f \phi(B + (B + S_0^j + S_0^k) - \theta_0^f B) R_T \right) & = 0,
\psi^f (B + S_0^j + S_0^k - \theta_0^f B) = 0,
E\left(\Phi(D_T, W_T^f) \phi \Phi(D_T, W_T^f)^{-1/c} - \Phi \cdot D_T \right) & = 0,
E\left(\Phi(D_T, W_T^f) P(D_T, W_T^f) - S_0^j \right) & = 0,
E\left(\Phi(D_T, W_T^f) P(D_T, W_T^f) - S_0^k \right) & = 0,
\end{align*}
\]

where

\[ R_T = R(D_T, W_T^f) = \frac{\sigma^m(\Phi(D_T, W_T^f), W_T^f, P(D_T, W_T^f)) - \sigma^m(\Phi(D_T, W_T^f), W_T^f, P(D_T, W_T^f), W_T^j)}{B + S_0^j + S_0^k - \theta_0^f B} \]

is the equilibrium net return on mutual funds defined in (4) and \( \sigma^m, \Phi, \) and \( P \) are the functions defined in (16), (21), and (22), respectively.

Given a solution \((\theta_0^f, \phi, S_0^j, S_0^k)\), the equilibrium state-price density is given by

\[ \pi_t = E\left(\Phi(D_T, W_T^f) \phi \right)^{1/n}, \]

the equilibrium stock price processes are given by

\[ S_t^j = E\left[\pi_t D_T^j / \pi_t \right] (j = 1, 2), \]

the optimal investment policy for the fund investors is given by \( \theta_0^f \) and the optimal wealth processes for the active investors and fund managers are given by

\[ W_t^f = E\left[\pi_t g^m(\psi^f \pi_t) / \pi_t \right], \]

and

\[ W_t^m = E\left[\pi_t g^m(\Phi(D_T, W_T^f), W_T^f, P(D_T, W_T^f)) / \pi_t \right], \]

respectively, where \( \psi^g = \psi^f \phi^{-c} \) and \( \psi^m = E(\Phi(D_T, W_T^f)) \).

The five equations in (24) can be easily identified, respectively, with the first-order conditions for the fund investors in Eq. (18), the budget constraint for the active investors in Eq. (7), and the two Euler equations that define the initial stock prices.

The next corollary provides an explicit solution for the equilibrium trading strategies in the case in which the performance fees are of the fulcrum type \((\gamma_1 = \gamma_1)\) and there are no load fees.

**Corollary 1.** If \( \gamma_1 = \gamma_1 \) and \( F(0, W_b) \leq 0 \) for all \( W_b \geq 0 \), then in equilibrium the fund managers’ portfolio consists of a combination of a long buy-and-hold position in the market portfolio, a long buy-and-hold position in the benchmark portfolio, and a short buy-and-hold position in the riskless asset:

\[
\theta_0^f = \frac{(\beta + \gamma_2)^{1/c} + \phi(1 - \lambda) \gamma_2 \phi_0 + \phi(1 - \lambda) \gamma_2_0 \phi_0^{(1/\gamma_2) - 1}}{(\beta + \gamma_2)^{1/c} + \phi(1 - \lambda) \beta + \gamma_2},
\]

for all \( t \in [0, T] \), where \( \beta_1 = (1, 0, 0) \). Similarly, the active investors’ portfolio consists of a long buy-and-hold position in the market portfolio, a short buy-and-hold position in the benchmark portfolio, and a buy-and-hold position in the riskless asset (which can be either long or short):

\[
\phi_t^f = \frac{(\beta + \gamma_2)^{1/c} - \lambda \gamma_2 \phi_0 + \phi(1 - \lambda) \gamma_2_0 \phi_0^{(1/\gamma_2) - 1}}{(\beta + \gamma_2)^{1/c} + \phi(1 - \lambda) \beta + \gamma_2}.
\]

In addition, if \( \gamma_1 \neq 0 \) and \( D_1^l \) and \( D_2^l \) are identically distributed conditional on the information at time \( t \), then \( S_t^j > S_t^k \) (respectively, \( S_t^j > S_t^k \)) if the benchmark portfolio is certain to hold more (respectively, less) shares of stock 1 than of stock 2 at time \( T \).

The above corollary shows that, if performance fees are of the fulcrum type and there are no load fees (or, more generally, if \( \lambda \leq \gamma_2 \beta_1 W_b^l \)), then in equilibrium the fund managers hold more (respectively, fewer) shares of stock 1 than of stock 2 at time \( t \) if and only if they are benchmarked to a portfolio holding more (respectively, fewer) shares of stock 1 than of stock 2 at time \( t \). This tilt in the fund portfolios toward the stock more heavily weighted in the benchmark portfolio results in the equilibrium price of this stock being higher, ceteris paribus, than the equilibrium price of the other stock. Moreover, if the benchmark portfolio is buy-and-hold, then the equilibrium trading strategies are also buy-and-hold. Thus, in our model, performance fees of the fulcrum type do not increase the fund portfolios’ turnover.

Before proceeding with a detailed analysis of the equilibrium, we note that in our model fees are paid only at the terminal date, with utility derived from terminal wealth. Allowing for fees paid at discrete intervals would not qualitatively affect our results (other than reinterpreting the horizon \( T \)) as long as the fees are consumed rather than saved. If the fees are saved and the manager faces no restrictions on his portfolio choices, the standard assumption in the agency literature that the agent is unable to hedge his compensation in his private account would be violated, reducing or eliminating the incentive effects of performance fees. Allowing the manager to have outside wealth significantly complicates the analysis. The results would depend in part on what restrictions are imposed on managers when trading on their own account, as well as the wealth of managers relative to the size of the funds they manage. Examples of such restrictions include: requiring managers hold a personal portfolio that is identical to the fund portfolio, precluding managers from \( \text{\footnote{17} Clearly, this buy-and-hold result is due to the fact that fund managers and active investors are assumed to have the same utility function.} \)
investing in assets held by the fund, and disallowing managers from taking positions that are opposite of those of the fund (i.e., shorting an asset that the fund is long or going long an asset that the fund is short). Some of these issues are analyzed in a partial equilibrium setting in Kaniel, Tompaidis, and Zhou (2009). Naturally, a key state variable that determines fund portfolio allocations in such a setting is the ratio between a manager’s personal wealth and the size of the fund she manages. The smaller this ratio is, the closer her actions will be to the ones we identify in the current paper. A general equilibrium analysis for the case where managers are allowed to have outside wealth is left for future research. Equilibrium implications in such an analysis would vary depending on the trading restrictions that are imposed on managers, and in addition to the channels we identify in the current paper, will depend as well on the evolution of the ratio between managers’ personal wealth and assets under management in the funds they manage.

5. Analysis of equilibrium

This section contains a numerical analysis of the asset pricing implications of delegated portfolio management. Throughout most of this section, we calibrate the benchmark asset to the S&P 500 and the non-benchmark asset to the rest of the U.S. stock market. We assume that stock (portfolio) 1 is the benchmark portfolio, i.e., that \( \pi^0 = (0, 1, 0) \): we address in Section 6.1 the robustness of our results to alternative choices of the benchmark (including optimal choice).

We consider investment horizons from \( T = 1 \) to 5. Based on evidence reported by Del Guercio and Tkac (1998) that 42.6% of the pension fund sponsors who use performance-based fees to compensate managers rely on a three to five-year investment horizon to measure performance and that the median holding period among the mutual fund investors who totally redeem their shares is five years, we consider \( T = 5 \).\(^{18}\) We consider \( T = 1 \) since most performance contracts in the hedge fund industry are based on annual performance. Also, hedge fund lock-up periods are heavily clustered around one year, and exhibit little variability (see Aragon, 2007). Finally, data provided to us by the Investment Company Institute (ICI) on characteristics of fees charged by mutual funds that have a fulcrum fee in place show that performance is typically measured over a 36-month rolling window.\(^{19}\)

Our goal is to understand the asset pricing implications of commonly observed performance contracts. We consider two performance fee structures: fulcrum fees \( (\gamma_1 = \gamma_2 > 0) \) and asymmetric performance fees \( (\gamma_1 = 0, \gamma_2 > 0) \). In both cases, the performance compensation is added on top of a proportional fee \( (\beta > 0) \), as is typically done in practice.

For fulcrum fees we set \( \beta/T = 0.6\% \) and \( \gamma_1 = \gamma_2 = 2\% \); using the ICI data we find that the value-weighted average proportional component across funds charging performance fees is 60 basis points, and the typical fulcrum performance fee is 2%; both on an annual basis. For asymmetric fees we analyze the predominant two-twenty hedge fund contract \((\beta/T = 2\%, \gamma_1 = 0, \text{and } \gamma_2 = 20\%)\).

Given that U.S. financial institutions hold about 50% of financial assets (see Allen, 2001), we set the fraction \( \lambda \) of fund investors in the economy to 0.5. The investors’ relative risk aversion coefficient \( c \) is set to 10. In calibrating the economy, we need to specify the aggregate supply of the bond and the parameters for the dividend processes. The calibration is based on annual data from 1967 to 2007. For calibrating the aggregate supply of the bond, we use the Fed Flow of Funds L4 annual series. For the dividend processes, stock price data are obtained from CRSP and we use the one year constant maturity T-bill (Fed statistical series H15) as the riskfree asset for computing stock excess returns.

The processes in Eq. (1), which represent the conditional expectation of the liquidation dividend, are taken to be geometric Brownian motions. Using the time series of annual log excess returns for the S&P 500 and the rest of the market, we find the annual standard deviations and the correlation between the two series. The resulting calibrated diffusion parameters are \( \sigma^2_{D,T} = 0.153D \) and \( \sigma^2_{T,D} = 0.1911D \), and the correlation coefficient is \( \rho = 0.89 \); clearly, a lower correlation would allow for larger pricing differences across the two stocks to emerge in equilibrium as a result of benchmarking. Since dividend processes represent the conditional expectation of the liquidation dividend, we set \( \mu^1_{D,T} = \mu^2_{T,D} = 0 \).

To calibrate the initial levels of the expected liquidation dividends and the aggregate supply of the bond, we use two facts. First, the average market share of the S&P 500 during the last 40 years has been 0.69.\(^{20}\) Second, using the Fed Flow of Funds L4 annual series, we find that the average ratio between the value of Treasury securities including U.S. savings bonds and corporate equities was 0.45. Specifically, we computed the 0.45 by first computing for each year in the period 1967–2007 the average dividend processes. The calibration is based on annual data from 1967 to 2007. For calibrating the aggregate supply of the bond, we use the Fed Flow of Funds L4 annual series. For the dividend processes, stock price data are obtained from CRSP and we use the one year constant maturity T-bill (Fed statistical series H15) as the riskfree asset for computing stock excess returns.

The processes in Eq. (1), which represent the conditional expectation of the liquidation dividend, are taken to be geometric Brownian motions. Using the time series of annual log excess returns for the S&P 500 and the rest of the market, we find the annual standard deviations and the correlation between the two series. The resulting calibrated diffusion parameters are \( \sigma^2_{D,T} = 0.153D \) and \( \sigma^2_{T,D} = 0.1911D \), and the correlation coefficient is \( \rho = 0.89 \); clearly, a lower correlation would allow for larger pricing differences across the two stocks to emerge in equilibrium as a result of benchmarking. Since dividend processes represent the conditional expectation of the liquidation dividend, we set \( \mu^1_{D,T} = \mu^2_{T,D} = 0 \).

To calibrate the initial levels of the expected liquidation dividends and the aggregate supply of the bond, we use two facts. First, the average market share of the S&P 500 during the last 40 years has been 0.69.\(^{20}\) Second, using the Fed Flow of Funds L4 annual series, we find that the average ratio between the value of Treasury securities including U.S. savings bonds and corporate equities was 0.45. Specifically, we computed the 0.45 by first computing for each year in the period 1967–2007 the average dividend processes. The calibration is based on annual data from 1967 to 2007. For calibrating the aggregate supply of the bond, we use the Fed Flow of Funds L4 annual series. For the dividend processes, stock price data are obtained from CRSP and we use the one year constant maturity T-bill (Fed statistical series H15) as the riskfree asset for computing stock excess returns.

\(^{18}\) The section of the working paper comparing investment horizons in the pension fund and mutual fund industry does not appear in the published article (Del Guercio and Tkac, 2002).

\(^{19}\) We would like to sincerely thank Brian Reid, ICI’s chief economist, and Sean Collins for collecting and providing us with the ICI data. The data describe contract characteristics for those mutual funds having performance fees in place as of 2008.

\(^{20}\) The market share of the S&P 500 has been fairly stable over these 40 years, with an annual standard deviation of 0.03.
portfolio weights of the different agents. We set $r=0.06$, the mean log return of the one year constant maturity T-bill over the period 1967–2007.

All remaining figures, and all tables, are at the end of the paper. These figures and tables are for an economy horizon of $T=5$. The figures and tables for $T=1$ are, in general, qualitatively similar to the ones for $T=5$. In those cases where they differ, we elaborate on this in the text. The corresponding figures and tables for the case $T=1$ are in a separate appendix that is available upon request.

5.1. Benchmark economy and proportional fees

Before moving to economies in which fund managers receive performance fees, it is useful to review equilibrium prices in the version of our economy in which all agents have direct costless access to the equity market (i.e., $\lambda=0$ or equivalently $x=\beta=\gamma_1=\gamma_2=0$) and in the version in which purely proportional management contracts are used (i.e., $x=\gamma_1=\gamma_2=0$). For the purely proportional contract we set $\beta/T$ to 1.33%. This is the average from 1981–2006 of the annual cross-sectional value-weighted average of total fees charged by actively managed equity funds in the following investment objectives: aggressive growth, growth, and growth and income. Total fees are defined, similar to, for example, Huang, Wei, and Yan (2007), as the total expense ratio plus one-seventh of the up-front load.

We first briefly consider an economy where the two dividend processes are identically distributed, maintaining the same volatility of aggregate dividends as in the baseline calibration. Fig. 2 plots expected values of key equilibrium quantities at the midpoint ($t=T/2$) as a function of the second stock’s dividend share: $D_2^T/(D_1^T+D_2^T)$.

In the background we superimpose the density distribution of the dividend share: the visible shaded region accounts for more than 99.5% of the distribution. For comparison, the figure also plots the corresponding equilibrium values with costless access to the equity market, an economy we will refer to henceforth as the benchmark economy.

Starting from the benchmark economy with costless portfolio delegation, since in this case fund managers and active investors have identical preferences and wealths (as fund investors optimally delegate the management of their entire endowment), they must hold identical buy-and-hold portfolios in equilibrium. Hence, when a stock’s dividend share increases, resulting in an increase in the price of that stock relative to the price of the other stock, investors must be induced to allocate a higher fraction of their portfolios to that stock in order for the market to clear, as shown in the top two panels of Fig. 2. This equilibrium incentive takes the form of a higher risk premium (which compensates investors for the higher correlation between the stock’s dividend and aggregate consumption) and a higher Sharpe ratio. Since the riskfree rate is an exogenously specified constant throughout our analysis, differences in risk premiums are driven solely by differences in instantaneous expected returns. Therefore, in the rest of the paper, while the graphs plot the “risk premiums,” the discussion will focus on variations in stock expected returns. Stock volatility falls in response to an increased dividend share, which again is consistent with providing an incentive to increase the holding, although this effect is small. Finally, reflecting the increasing expected returns, price/dividend ratios are monotonically decreasing functions of the dividend shares.

The analog of Fig. 2 for the baseline calibration is shown in Fig. 3. Comparing, for example, the stock Sharpe ratio panels across the two figures shows that while with identically distributed dividend processes an increase in the second stock’s dividend share increases the Sharpe ratio of the second stock and decreases the Sharpe ratio of the first, under the baseline calibration both Sharpe ratios increase. Notice though the line depicting the Sharpe ratio of the second stock is steeper. Under the baseline calibration, the volatility of the dividends of the first stock is lower than that of the second stock. Consequently, an increase in the second stock’s dividend share increases the riskiness of the market portfolio. Since the two dividend processes are highly correlated, for markets to clear, both stocks need to become more attractive, although the Sharpe ratio of the second stock needs to increase by more due to its increased dividend share. The qualitative differences in the pattern of the expected return, volatility, and price/dividend ratio graphs across the two figures are explained by the same reasoning.

Moving from the benchmark economy to an economy with costly portfolio delegation and proportional fees, the allocation to mutual funds by fund investors decreases to less than 100% and the allocation to bonds becomes strictly positive. As a result, a lower fraction of aggregate wealth is available for investment in the equity market, and in order to restore market clearing, fund managers and active investors must be induced to increase their equity holdings relative to the benchmark economy, as shown in the top two panels of both Figs. 2 and 3.

To ensure this, expected returns and Sharpe ratios increase relative to the benchmark economy, while price/dividend

---

21 We sincerely thank Kelsey Wei for computing this for us.

22 We set $\sigma_1^2(D,t) = \sigma_2^2(D,t) = 0.165D_0$ and recalibrate $D_0 = \tilde{D}_0$ to 1.210.

23 To produce graphs as a function of the dividend share $D_t^1/(D_t^1+D_t^2) = 1/(D_t^1/D_t^2 + 1)$, we utilize the fact that if $D_t^1$ and $D_t^2$ are log-normally distributed, then the distribution of $D_t^1$ conditional on $D_t^1/D_t^2$ is log-normal. We show this in Lemma A.1 at the end of the Appendix. This allows us to efficiently integrate over all values of $D_t^1$ and $D_t^2$, conditional on a fixed level of the dividend share.

24 Our benchmark economy is similar to the two-trees Lucas economy studied in Cochrane, Longstaff, and Santa-Clara (2007), the differences being that Cochrane, Longstaff and Santa-Clara assume logarithmic utility, infinite horizon, intertemporal consumption, and bonds in zero net supply, while we assume general CRRA utilities, finite horizon, consumption at the terminal date only, and bonds in positive net supply.

25 When the correlation between the two dividend processes is below 0.65, the two Sharpe ratios move in opposite direction even when $\sigma_1^2(D,t) = 0.153D_0$ and $\sigma_2^2(D,t) = 0.191D_0$ (not shown).

26 Within each of the figures, the top two panels are identical, as expected: with CRRA preferences, portfolio choice is independent of wealth levels and is the same if utility is derived from wealth $W$ or $\beta W$. 
ratios decrease. Stock volatilities slightly rise relative to the benchmark economy. Clearly, in the absence of benchmarking, delegated portfolio management has a similar impact on the two stocks. As shown in Fig. 4, which plots unconditional expected values of key equilibrium quantities as a function of time $t$, the deviations between the benchmark economy and an economy with proportional fees persist through time for portfolio allocations, expected returns, and Sharpe ratios. Differences in volatilities and price/dividend ratios decline over time.

Although not shown, the deviations between stock expected returns, volatilities, Sharpe ratios, and price/dividend ratios in the presence of proportional fees and
the corresponding values in the benchmark economy are monotonic in the proportional fee parameter $b$. The qualitative pricing effects are identical in the presence of load fees.

Before proceeding with analyzing economies with fulcrum and asymmetric performance fees, we note that the model-generated risk premiums under our calibration are higher than those in the data: over the period 1967–2007, the average annualized excess return relative to the one year constant maturity T-bill of the S&P 500 and the rest of the market were both about 0.055. Lowering the risk aversion coefficient reduces the magnitudes of the model-generated risk premiums. For example, with a risk aversion coefficient of 3, the risk premiums at the
initial date with proportional fees are 0.055 and 0.065 for stock one and two, respectively (not shown).

5.2. Fulcrum performance fees

The 1970 Amendment of the Investment Advisers Act of 1940 restricts mutual fund performance fees to be of the fulcrum type; we examine next their impact on equilibrium. Fig. 5 plots the unconditional expected values of key equilibrium quantities as a function of time $t$. The investment horizon is set to $T=5$. The proportional fee parameter is set to $\beta/T = 1.33\%$. For comparison, the corresponding values in the benchmark economy ($\beta=0$) are also plotted. The solid (respectively, dotted) line refers to the first stock (respectively, the second stock) with proportional fees, while the dashed (respectively, dot-dashed) line refers to the first stock (respectively, the second stock) in the benchmark economy. Both the instantaneous risk premiums and instantaneous volatilities are on a per annum basis.

The qualitative results are insensitive to the specific choice of the fee parameters, or to the time horizon of the economy.
equilibrium values in an economy with a purely proportional fee of $\frac{\beta}{T} = 0.6\%$.

As shown in Corollary 1, and displayed in the top left panel of the figure, the presence of a penalty for underperforming the benchmark portfolio implicit in fulcrum fees leads fund managers to tilt their portfolios toward the benchmark stock relative to their holdings with a purely proportional contract. In addition, the overall equity allocation by fund managers is higher than with purely proportional fees. Thus, in order to ensure market

---

28 With a sufficiently large fulcrum fee component, the fund essentially behaves as an index fund, with a 100% allocation to the benchmark stock.
clearing, relative to holdings under a purely proportional contract active investors must be induced to hold portfolios that have lower equity allocations and are tilted toward the non-benchmark stock over the entire horizon, as displayed in the top right panel. The equilibrium price distortions relative to an economy with purely proportional fees have the expected signs, and naturally decline as the terminal date approaches. Specifically, expected returns and Sharpe ratios are lower than in an economy with purely proportional fees (in order to induce a lower overall equity allocation by active investors), with the difference being more pronounced for the benchmark stock (in order to induce active investors to tilt their portfolios toward the non-benchmark stock). Correspondingly, price/dividend ratios are higher than with purely proportional fees, with the difference being once again more pronounced for the benchmark stock. Interestingly, volatilities are lower than with purely proportional fees, implying that adding performance fees on top of a purely proportional contract may in fact stabilize prices.

To analyze the sensitivity of the unconditional distortions identified above, Fig. 6 plots the key equilibrium quantities at time $t = T/2$ as a function of the dividend share of the second (non-benchmark) stock. Given that with fulcrum fees and a fixed benchmark portfolio, equilibrium trading strategies are buy-and-hold, as shown in Corollary 1, the proportional tilt of active investors toward the non-benchmark stock must be larger when the price of the benchmark stock is large relative to that of the non-benchmark stock, i.e., when the second stock's dividend share is low. Indeed, using information from the top right panel we find that the ratio of the fraction of equity in active investors' portfolios that is invested in stock 2 with fulcrum fees to the fraction with purely proportional fees monotonically declines from 1.39 at a dividend share of 0.2–1.17 at a dividend share of 0.6. Consistent with this pattern, the equilibrium price distortions for expected returns, volatilities, Sharpe ratios, and price/dividend ratios relative to an economy with purely proportional fees are larger when the second stock's dividend share is low, having otherwise the expected signs throughout most of the dividend share space.

To relate the implications of our model to the available empirical evidence concerning the equilibrium pricing effects of benchmarking, we examine next how equilibrium quantities change in our model in response to changes in the composition of the benchmark portfolio. In particular, we consider an unanticipated change in the composition of the benchmark portfolio at time $t = T/2$ from 100% stock 1 to 100% stock 2. That is, we consider a setting where up to $t = T/2$, it was almost certain that stock 1 will be the benchmark that fund managers will be evaluated against at the terminal date $T$, and at $t = T/2$, managers learn that even though they thought they would be evaluated against stock 1 at the terminal date, they will be evaluated relative to a self-financing portfolio that was fully invested in stock 1 up to time $t = T/2$ and is fully invested in stock 2 from $t = T/2$ till $T$. To focus on the impact of the recomposition, for this analysis we reconsider the economy in which the two dividend processes are identically distributed, with the same volatility of aggregate dividends as in the baseline calibration. Table 1 displays unconditional expected values of the key equilibrium quantities immediately prior to and immediately following the announcement of the benchmark recomposition.

As shown in the Fulcrum labeled column of Panel A, the weight in the funds' portfolio of the stock that is added to (respectively, dropped from) the benchmark increases (decreases). This is consistent with the evidence reported in Pruitt and Wei (1989) that changes in the portfolios of institutional investors are positively correlated with changes in the composition of the S&P 500 index. As a result, the price of the stock that is added to the benchmark portfolio increases following the announcement, while its Sharpe ratio and expected returns decrease (Fulcrum column of Panels E, F, and C). The changes in the price, Sharpe ratio, and expected return of the stock that is dropped from the benchmark have the opposite sign. Furthermore, the effect on prices is asymmetric, with the expected absolute percentage change being larger for the stock that is added to the benchmark (1.8%) than for the stock that is dropped (1.67%).

The magnitude of the price changes are increasing in the time left ($T - t$). For example, following a recomposition at $t = 1$ (respectively, $t = 4$), the absolute percentage price change for the stock that is added to the benchmark is 2.8% (0.75%) and for the stock that is dropped from the benchmark is 2.4% (0.73%); not shown. Shleifer (1986), Harris and Gurel (1986), Beneish and Whaley (1996), Lynch and Mendenhall (1997), and Wurgler and Zhuravskaya (2002), among others, show a positive permanent price effect of about 3–5% associated with the inclusion of a stock in the S&P 500 index. Chen, Noronha, and Singal (2004) report that the price effect is asymmetric for additions and deletions, with the permanent price impact of deletions being smaller in absolute value.

Consistent with the unconditional patterns, in unreported results we find that throughout all the dividend share range, the weight in the funds' portfolio and the price of the stock that is added to (dropped from) the benchmark increase (decrease) following the announcement, while its Sharpe ratio and expected returns decrease (increase).

For volatilities, Harris (1989) shows a small positive average difference in daily return standard deviations between stocks in the S&P 500 index (the most commonly used benchmark) and a matched set of stocks not in the S&P 500 over the period 1983–1987. Since this difference was insignificant over the pre-1983 period, Harris attributed it to the growth in index derivatives trading, noticing that the contemporaneous growth in index funds was

---

29 In the $T = 1$, economy the corresponding values at $t = T/2 = 0.5$ are 1% and 0.69%.

30 Chen, Noronha, and Singal (2004) interpret the asymmetry of the price effect as evidence against the hypothesis that the effect is due to downward-sloping demand curves and in favor of the alternative hypothesis that the effect is due to increased investors’ awareness. Our analysis implies that portfolio delegation with benchmarking is an alternative possible explanation for the asymmetry of the effect.
unlikely to be a possible alternative explanation, as "it seems unlikely that volatility should increase when stock is placed under passive management." Yet, our model delivers this implication concerning volatilities: the Fulcrum column of Panel D of Table 1 shows that immediately prior to the recomposition announcement, the unconditional expected volatility is 14.99% for the benchmark stock and 14.89% for the other stock. Furthermore, a similar relation holds throughout most of the relevant dividend share range; not shown. While not shown, with identically distributed dividend processes the unconditional expected volatility of the benchmark stock is slightly higher than that of the non-benchmark stock throughout the economy horizon. This implication can also be inferred indirectly from Fig. 5: volatilities are lower with than without fulcrum performance fees, and the absolute value of the difference between the volatility with and without fulcrum performance fees is smaller for the benchmark stock than for the non-benchmark stock.

31 In the $T=1$, economy the corresponding values are 16.16% and 16.11%.
Table 1
Equilibrium quantities before and after recomposition.
The table reports unconditional expected values of equilibrium quantities at time $t=1/2$, immediately prior to and immediately following an unanticipated change in the benchmark from stock 1 to stock 2. Dividends of the two stocks are identically distributed. The contract parameters are $\beta=2\%$, $\gamma_1=0$, and $\gamma_2=20\%$ for asymmetric fees, and $\beta=0.6\%$, and $\gamma_1=\gamma_2=2\%$ for fulcrum fees.

| Panel A: Fund portfolio | Stock 1 After | 0.2903 | 0.4280 |
| | Stock 1 Before | 0.5813 | 0.4161 |
| | Stock 2 After | 0.5808 | 0.4043 |
| | Stock 2 Before | 0.2890 | 0.4182 |

| Panel B: Active investor portfolio | Stock 1 After | 0.5082 | 0.4120 |
| | Stock 1 Before | 0.2767 | 0.4206 |
| | Stock 2 After | 0.2674 | 0.4211 |
| | Stock 2 Before | 0.4997 | 0.4111 |

| Panel C: Risk premium | Stock 1 After | 0.1813 | 0.1922 |
| | Stock 1 Before | 0.1755 | 0.1906 |
| | Stock 2 After | 0.1750 | 0.1900 |
| | Stock 2 Before | 0.1813 | 0.1925 |

| Panel D: Volatility | Stock 1 After | 0.1488 | 0.1506 |
| | Stock 1 Before | 0.1499 | 0.1476 |
| | Stock 2 After | 0.1499 | 0.1472 |
| | Stock 2 Before | 0.1489 | 0.1508 |

| Panel E: Sharpe ratio | Stock 1 After | 1.2175 | 1.2747 |
| | Stock 1 Before | 1.1701 | 1.2895 |
| | Stock 2 After | 1.1665 | 1.2891 |
| | Stock 2 Before | 1.2172 | 1.2754 |

| Panel F: Price dividend ratio | Stock 1 After | 0.5185 | 0.5046 |
| | Stock 1 Before | 0.5273 | 0.5045 |
| | Stock 2 After | 0.5280 | 0.5045 |
| | Stock 2 Before | 0.5186 | 0.5048 |

5.3. Asymmetric performance fees

Moving now to equilibria in the presence of asymmetric performance fees, Fig. 7 plots the unconditional expected values of the key equilibrium quantities as a function of time $t$. As mentioned earlier, for asymmetric performance fees, we select the contract parameters $\beta/T=2\%$ and $\gamma_2=20\%$. For comparison, Fig. 7 also plots the corresponding equilibrium values in an economy with a purely proportional fee of $\beta/T=2\%$.

Comparing unconditional expected values of equilibrium quantities with asymmetric performance fees (Fig. 7) to the ones with fulcrum fees (Fig. 5) shows some interesting differences. Fund managers portfolios, relative to the ones under purely proportional fees, are no longer tilted exclusively toward the benchmark stock. While at the initial date the tilt is toward the benchmark, similar to the direction of the tilt with fulcrum fees although somewhat smaller in magnitude, beyond $t=1$ the unconditional tilt shifts to the non-benchmark stock. For markets to clear, price dynamics need to adjust to induce a reverse pattern in active investors’ portfolios. Consequently, expected returns and Sharpe ratios also exhibit similar patterns of reversals in the direction of the distortion relative to an economy with only proportional fees, as do to a lesser extent volatilities.32 Also, at the initial date, the reduction in volatilities of both stocks seems larger for asymmetric performance fees than for fulcrum fees.33

To understand the source of these differences, Fig. 8 plots key equilibrium quantities at time $t=1/2$ as a function of the second stock’s dividend share. For comparison, the figure also plots the corresponding values in an economy with purely proportional fees. It is useful to keep in mind for the following discussion that an increase in the dividend share of the second (non-benchmark) stock is associated with an increase in the difference between the price of the non-benchmark stock and the price of the benchmark stock: thus, the funds’ excess return over the benchmark portfolio, $(W_t^m-\delta S_t^1)/W_0^m$, is a monotonically increasing function of the second stock’s dividend share, with the excess return being zero at a dividend share of 40%. The 10th and 90th percentiles of the dividend share distribution are at dividend shares of 30.9% and 39.1%, with corresponding fund excess returns of $-12.2\%$ and $-1.4\%$; annualized, $-5.1\%$ and $-0.5\%$.

While fulcrum fees unambiguously induce fund managers to tilt their portfolios toward the benchmark stock relative to their portfolios under purely proportional fees, asymmetric performance fees can induce risk-averse fund managers either to select portfolios having high correlation with the benchmark in an attempt to hedge their compensation (similar to under fulcrum fees), or to select portfolios having low correlation with the benchmark in an attempt to maximize the variance of the excess return of the managed portfolio over the benchmark (and hence the expected value of their performance fees, which are a convex function of this excess return). These shifting risk incentives are evident in the top left panel of Fig. 8, which plots funds’ portfolio weights, or, even more clearly, in the top panel of Fig. 9, which plots funds’ share holdings. As shown in these figures, the first incentive dominates (inducing the fund manager to tilt the fund’s portfolio toward the benchmark stock) when the fund’s excess return is positive or slightly negative: when the second stock’s dividend share is larger than 39.5%, corresponding to an excess return above $-0.8\%$, or equivalently to an annualized excess return above $-0.32\%$. The second incentive dominates (inducing the fund manager to tilt the fund’s portfolio toward the non-benchmark stock) when the fund’s excess return is sufficiently negative (below $-0.8\%$).34 When the fund’s excess return is

---

32 It is important to keep in mind that each point on each of the graphs in this figure is obtained by integrating the respective quantity over the entire two-dimensional dividends domain. As a result, crossings between the asymmetric and proportional lines need not occur at the same position across the different subfigures.

33 Caution is warranted in comparing the magnitudes of the distortions, relative to proportional fees, between asymmetric and fulcrum fees as their proportional fees $\beta$ are different. However, later we show that assuming the same proportional fee does not change the conclusions regarding the magnitudes at the initial date for either the portfolio allocations or the volatilities.

34 Clearly, which incentive dominates at a given excess return critically depends on both the managers’ risk aversion and the time...
strongly negative, so that the probability of earning positive performance fees is negligible, the fund manager essentially behaves as if she were receiving a purely proportional fee and the difference between the holdings of the two stocks with asymmetric performance fees and with proportional fees is close to zero. As shown in the bottom panel of Fig. 9, the overall equity allocation by mutual funds exceeds the overall allocation that would have been chosen with purely proportional fees when the excess return is positive, dips below the level under purely proportional fees when the excess return is moderately negative, and converges to the level under purely proportional fees when the excess return is close to zero.
proportional fees as the excess return becomes more and more negative.

The behavior of expected returns and Sharpe ratios in Fig. 8 reflects the behavior that would be expected to ensure market clearing given the fund managers’ portfolio strategy described above. Specifically, two forces are at play in leading active investors to tilt their portfolios so that the market clears. First, a tilt in managers’ allocation toward the benchmark (non-benchmark) stock relative to the allocation under proportional fees decreases the expected return and Sharpe ratio of the benchmark (non-benchmark) stock and increases the expected return and Sharpe ratio of the non-benchmark (benchmark) stock. Second, expected returns and Sharpe ratios of both stocks decrease (increase) relative to the corresponding quantities in an economy with purely proportional fees when the overall equity allocation by fund managers is above (below) that with purely proportional fees. Price/dividend ratios also have their expected behavior given their inverse relation to expected returns.

Perhaps surprisingly, equilibrium stock volatilities in the presence of asymmetric performance fees tend to be

Fig. 8. The graph plots key equilibrium quantities at time $t = T/2 = 2.5$ with asymmetric performance fees as a function of the second stock’s dividend share. The contract parameters are $\beta/T = 2\%$, $\gamma_1 = 0$, and $\gamma_2 = 20\%$. For comparison, the corresponding values with proportional fees ($\gamma_1 = \gamma_2 = 0$) are also plotted. The solid (respectively, dotted) line refers to the first stock (respectively, the second stock) with asymmetric fees, while the dashed (respectively, dot-dashed) line refers to the first stock (respectively, the second stock) with proportional fees. Both the instantaneous risk premiums and instantaneous volatilities are on a per annum basis. The shaded region in the background is the density of the dividend share at $t = T/2$. 

proportional fees.
generally smaller than those in an economy with purely proportional fees, in spite of the higher portfolio turnover. The reduction in volatilities is particularly strong in the region in which the funds’ excess return is either moderately negative or positive: in fact, in this region the volatility of the first stock is not only lower than the corresponding volatility in an economy with purely proportional fees, but also lower than in the benchmark economy. The source of this volatility reduction can be understood by noticing that, in the presence of asymmetric performance fees, the turnover by mutual funds is concentrated in the region in which their excess return is either moderately negative or moderately positive and is associated with a portfolio reallocation away from the stock whose dividend share, and hence whose price, is rising (Fig. 9, top panel): to ensure market clearing (that is, to induce active investors to hold even more of the stock whose dividend share is rising), the price of this stock must rise less than what it would have otherwise, lowering its volatility. The small increase in volatilities over their values in an economy with purely proportional fees in the lower stock 2 dividend share ratio region can be explained by a symmetric argument, as in this region there is a small reallocation by fund managers toward the stock whose dividend share is rising.

Turning now to the effect of changes in the composition of the benchmark portfolio, the Asymmetric labeled

---

**Fig. 9.** The top graph plots the funds’ shares holdings and the bottom graph plots the funds’ overall equity portfolio weight with asymmetric performance fees as a function of the second stock’s dividend share, at \( t = T/2 \). The contract parameters are \( \beta/T = 2\% \), \( g_1 = 0 \), and \( g_2 = 20\% \). In the top graph, the solid (respectively, dotted) line refers to the first stock (respectively, the second stock). For comparison, the corresponding values with proportional fees \( (g_1 = g_2 = 0) \) are also plotted: with proportional fees, share holdings of the two stocks are identical and are represented by the same dashed line in the top graph. In the bottom graph, the solid line refers to the overall equity portfolio weight with asymmetric fees, and for comparison the dotted line refers to the corresponding values with proportional fees. The economy horizon is \( T = 5 \). The shaded region in the background is the density of the dividend share at \( t = T/2 \).
column of Table 1 displays unconditional expected values immediately before and after a benchmark recomposition at \( t = T/2 \) from 100% stock 1 to 100% stock 2, in an economy with two identically distributed dividend processes. Immediately prior to the recomposition, fund portfolios are on average tilted toward the non-benchmark stock, implying that on average the maximizing variance of excess returns incentive dominates the hedging incentive. Changes in manager allocations following the recomposition are as would be expected given the allocations prior to the recomposition (i.e., the allocation to the stock that is dropped from (added to) the benchmark increases (decreases)), as is the fact that the average magnitude of the change in allocations is significantly smaller than with fulcrum fees (compare Fulcrum and Asymmetric columns of Panel A). Even though the two stocks have identically distributed dividend processes, fund managers’ unconditional expected allocations to stock 1 (2) immediately before the recomposition need not be the same as the allocation to stock 2 (1) immediately after the recomposition. While in the recomposition the benchmark flips from stock 1 to 2, the dividend share remains the same and managers’ strategies are not symmetric around a dividend share of 0.5. Consequently, other equilibrium quantities need not be symmetric around the recomposition as well. The corresponding impact on volatilities, Sharp ratios, and prices reflects the expected behavior that would guarantee market clearing following the recomposition.

Interestingly, even though following the recomposition, on average, managers increase their holding of the stock that is dropped from the benchmark and decrease their holding of the stock that is added to the benchmark, on average expected returns increase as a stock is dropped from the benchmark and decrease as it is added. This stems from the fact that on average expected returns prior to the recomposition are lower for the benchmark stock, even though on average managers’ portfolios are tilted toward the non-benchmark stock. To understand the source of such an outcome, it is important to keep in mind that as shown above with asymmetric fees, the proportional tilt of managers’ portfolios depends on the level of the dividend share. The figures that correspond to Figs. 8 and 9 for the case with identically distributed dividend processes (figures not shown) show that when the dividend share of the non-benchmark stock is high (low), managers’ portfolios relative to the portfolios held under the purely proportional contract are tilted toward (away from) the benchmark stock, and expected returns of the benchmark stock are lower (higher) than those of the non-benchmark stock. However, the difference in the absolute magnitude of the variance of excess returns incentive dominates from time \( t = 2.4 \) onward, and the hedging incentive dominates up to time \( t = 2.4 \); results not shown.

5.4. Varying performance contract parameters

While in Figs. 5–9 the performance sensitivity parameters are held fixed at their calibrated level, Fig. 10 shows how key equilibrium quantities at time 0 vary as we consider different performance fee sensitivity parameters, for both asymmetric and fulcrum performance fees: for asymmetric fees we fix \( \gamma_1 = 0 \) and vary \( \gamma_2 \) and for fulcrum fees we vary \( \gamma_1 = \gamma_2 \). To simplify the comparison between the two, we consider two different fulcrum fee specifications. In the first, the proportional fee is set to the level under the calibrated fulcrum contract (\( \beta = 0.6\% \)), as used in Figs. 5 and 6. In the second, we set the proportional fee to the same level as in the asymmetric performance contract (\( \beta = 2\% \)). The levels of the different equilibrium quantities under the corresponding proportional contracts are the values at the left-hand side of the figures (i.e., at the point \( \gamma_1 = \gamma_2 = 0 \)).

For the fulcrum contracts the results are as anticipated. Increasing the performance sensitivity parameter \( \gamma_1 = \gamma_2 \) amplifies the tilt of the fund managers toward the benchmark stock. This in turn amplifies distortions of other equilibrium quantities, in general in the anticipated direction. While not shown, as the performance sensitivity parameter increases from 0% to 30%, fund investors’ allocations to mutual funds decrease when \( \beta = 0.6\% \) (\( \beta = 2\% \)) from 94% (82%) to 71% (69%) of their endowment. This decrease in fund investors’ allocations partially mitigates distortions on equilibrium quantities that are due to higher performance sensitivity parameters, including the tilt required in active investors’ portfolios for markets to clear, as the amount of assets under fund managers’ management declines.

Differences in levels of equilibrium quantities between the two fulcrum contracts are mainly driven by the fact that for a fixed performance sensitivity parameter, fund investors allocate more to funds when the proportional fee component is lower. Consequently, expected returns and Sharpe ratios are lower with the lower proportional fee component, while prices are higher. Interestingly,

\[ 36 \text{With two identically distributed dividend processes, the maximizing variance of excess returns incentive dominates from time } t = 2.4 \text{ onward, and the hedging incentive dominates up to time } t = 2.4; \text{ results not shown.} \]

\[ 37 \text{With fulcrum fees, managers’ strategies are not symmetric around a dividend share of 0.5 as well, and therefore, there too, in general, there is no exact symmetry in the unconditional expected allocations around the recomposition.} \]
volatilities are lower as well. The rate of change in the level of the different quantities as a function of the performance sensitivity component is typically higher with a lower proportional fee. This is because a smaller proportional component implies that the performance component has a larger relative impact on fund managers' compensation. Consequently, differences in convexities and concavities between the two fulcrum contracts are more pronounced at low performance sensitivity levels, since the relative impact of the proportional component decreases as the performance component increases.

For asymmetric fees, as shown in the top left panel, at time 0, when excess return is zero, fund managers are in the region in which they find it optimal to tilt their portfolios toward the benchmark stock relative to their holdings under a purely proportional contract, although the tilt is slight. Therefore, equilibrium prices must adjust so that active investors are induced to tilt their portfolios
toward the non-benchmark stock. While both stocks have equilibrium expected returns that are lower than with purely proportional fees (i.e., comparing to the level at \( \gamma_2 = 0 \)), the difference is less pronounced for the non-benchmark stock. This smaller difference is offset by a smaller difference in volatility, where both volatilities are lower than with purely proportional fees as well, resulting in the Sharpe ratio of the non-benchmark (benchmark) stock being below (above) the Sharpe ratio with purely proportional fees. Therefore, active investors’ slight tilt at time 0 toward the non-benchmark stock is not induced by the Sharpe ratio or volatility channels but by some other channel.38

Interestingly, with asymmetric fees for the \( T = 1 \) economy, the tilt in managers’ portfolio allocations is not monotone in the performance sensitivity parameter. When \( T = 1 \) and \( \gamma_2 > 10\% \), similar to above, fund managers are in the region in which they find it optimal to tilt their portfolios toward the benchmark stock, implying that at time 0, hedging considerations dominate the incentive to maximize variance of excess returns. However, when \( T = 1 \) and \( \gamma_2 < 10\% \), the incentive to maximize variance of excess returns dominates and managers tilt allocations away from the benchmark stock. Equilibrium quantities must adjust to induce active investors to tilt allocations toward the benchmark stock. In this region, however, the tilt is induced through the Sharpe ratio and volatility channels. Specifically, the increase in the Sharpe ratio of the benchmark stock, the decrease in the Sharpe ratio of the non-benchmark stock, and the larger decrease in volatility relative to when purely proportional fees are in place of the benchmark stock lead active investors to tilt their holdings toward the benchmark stock, even though the decline in expected returns is slightly more pronounced for the benchmark stock.

Finally, given that variations in fund managers’ portfolio distortions, as a function of the performance sensitivity parameter, are less pronounced with asymmetric fees compared to fulcrum fees, fund investors’ allocations to mutual funds decrease less with asymmetric fees: from 82% of their endowment when \( \gamma_2 = 0 \) to 75% of their endowment when \( \gamma_2 = 30\% \); not shown.

5.4.1. Costs of delegation

Table 2 reports certainty equivalents for fund investors, fund managers, and active investors for a variety of contract parameters.39 For the calibrated contracts (top panel): fund managers’ certainty equivalent (henceforth, CE) is lower than with a purely proportional contract with the same proportional fee (second panel), both for the fulcrum and the asymmetric performance fee contracts. The same holds for fund investors with the asymmetric performance fee contract. With fulcrum fees, fund investors’ CE is slightly higher than in the corresponding proportional fee contract. However, holding fixed fund managers’ certainty equivalent at the level obtained with the calibrated fulcrum fee parameters, a proportional contract with a slightly lower fee than in the calibrated contract yields a certainty equivalent gain to fund investors of 0.6% (not shown). While the proportional contract yields a higher CE than the calibrated asymmetric contract for both fund managers and fund investors, as we will show in Section 6, under the optimal choice of fee parameters asymmetric performance fees dominate proportional fees, while fulcrum fees are dominated.

The last two columns quantify costs of delegations. The column Delegation cost quantifies the reduction in the CE of fund investors versus active investors: \( \frac{(\text{Fund investors CE})-(\text{Active investors CE})}{\text{(Fund investors CE)}} \). This is the overall cost borne by fund investors. It accounts for direct costs of paying fees, but also for indirect costs due the nature of the agency relation. The column Distortion cost quantifies the latter. To compute this, we first compute the CE of the combined wealth of fund investors and fund managers, not shown, representing Fund investors “before fees” CE. Then, similar to the first measure, we compute \( \frac{(\text{Active investors CE})-(\text{Fund investors before fees CE})}{\text{(Active investors CE)}} \). The Distortion costs in Panel \( \gamma = 0\% \) (proportional) represent the part of the losses that is due to the fact investors delegate only part of their wealth to funds to manage, since with proportional fees managers and active investors hold the same portfolios.

In general, both costs increase as the performance sensitivity parameter increases. Comparing the two cost columns suggests that the increase in Delegation cost associated with a higher performance sensitivity parameter is driven mostly by higher Distortion costs, as opposed to higher managerial compensation. The difference between the Delegation cost column and the Distortion cost column represents the explicit cost of delegation. Alternatively, this explicit cost of delegation can be gauged by computing \( \frac{(\text{Fund managers CE})-(\text{Active investors CE})}{\text{(Active investors CE)}} \): the two measures yield very close values. Considering the costs of either two asymmetric or two fulcrum contracts with the same performance sensitivity component but different proportional fee components shows that with the higher proportional fee component (\( \beta = 2\% \)), the bulk of the cost is an explicit cost due mostly to the proportional component. However, with the lower performance fee component (\( \beta = 0.6\% \)), Distortion costs can be a significant part of the overall cost, as would be expected.

Active investors slightly benefit indirectly from a higher performance sensitivity parameter, holding fixed the proportional fee. The benefit generally increases in the performance fee sensitivity level. A higher performance sensitivity parameter is associated with a higher Delegation cost, which induces fund investors to delegate less money to funds to manage. While managers’ overall proportional allocation to equity may increase, this does not suffice to restore market clearing due to the reduced size of the portfolio they manage. To induce active investors to increase allocations to stocks, so

\[ u'((CE)') = E[u'(W_0 + \frac{1}{2} \sigma^2 + (W_0 - \frac{1}{2} \sigma^2)R_t)], \]

where \( R_t \) is the net return on mutual funds defined in Eq. (4).

38 In the corresponding economy with identically distributed dividend processes, not shown, the expected return of the non-benchmark stock exceeds that of the benchmark stock; this higher expected return is offset by a higher volatility, resulting in the Sharpe ratio of the non-benchmark stock being below that of the benchmark stock.

39 The managers’ certainty equivalent (CEm) is defined by \( u'((CE)') = E[u'(W_0^m + W_0^f)] \). Similarly, the active investors’ certainty equivalent (CEa) is defined by \( u'((CE)') = E[u'(W_0^a)] \), and the fund investors’ certainty equivalent (CEi) is defined by \( u'((CE)') = E[u'(W_0^i + W_0^f - \frac{1}{2} \sigma^2R_t)] \), where \( R_t \) is the net return on mutual funds defined in Eq. (4).
that market clearing is restored, equilibrium quantities change in their favor.

Finally, comparing the two calibrated performance fee contracts to the corresponding proportional contracts shows that differences in costs between the two calibrated contracts are mostly due to the higher proportional fee component in the asymmetric contract. While the ratio between the Delegation cost under the asymmetric calibrated contract and the fulcrum calibrated contract is 3.32 (9.60/2.89), the ratio between the corresponding proportional contracts is only slightly lower at 3.14 (9.14/2.91). Comparing for each of the calibrated contracts the magnitude of the Delegation cost and Distortion cost shows as well that the bulk of the delegation cost is due to the proportional performance component.

6. Optimality of performance contracts

While our objective in this paper is to understand the impact of commonly observed performance contracts on equilibrium returns, we briefly address in this section the rationale for performance contracts within our model.

Given that in our model investors and managers have utilities with linear risk tolerance and identical cautiousness, the principle of preference similarity of Ross (1973) would seem to imply that a linear fee should be optimal. Specifically, Ross shows that, under the stated assumption on preferences, a linear fee achieves first best. However, two distinctive features of our model are that fund investors have direct access to riskless investment opportunities and that they take the fee structure as given when formulating their investment decisions: this is in contrast to standard models of delegated portfolio management, in which the principal is assumed to delegate the management of his entire portfolio and to be able to dictate the fee structure (subject only to the managers’ participation constraint).

To see why these features negate the optimality of a linear fee, recall from Ross (1973) or Cadenillas, Cvitanic, and Zapatero (2007) that the first-order condition for a
linear fee $F(W) = x + \beta W$ to achieve first best is

$$u'(\theta_0 B + W_T^m - F(W_T^m)) = \omega e^{-u'(F(W_T^m))},$$

(30)

where $\omega$ is a positive constant that depends on the managers’ reservation utility: Eq. (30) states that the fee should make the marginal utility for the principal proportional to that of the manager, ensuring Pareto-optimal risk sharing. This condition is satisfied if and only if $x = \theta_0 B / (1 + \omega) \geq 0$ and $\beta = 1 / (1 + \omega) > 0$. Thus, in order for a linear fee to achieve first best, the load component $x$ must depend on the portfolio allocation chosen by fund investors (in particular, $x = 0$ if the fund investors delegate their entire portfolios). If fund investors were able to choose the managers’ compensation contract while committing to delegating the amount $W_0' - \theta_0' B$, this fee would indeed be optimal. However, since in our model the fund investors choose their portfolio allocation taking the fee as given, Eq. (30) being satisfied becomes equivalent to the investors choosing $\theta_0 B = x / \beta$ ex post when confronted with a fee $F(W) = x + \beta W$ with $x \geq 0$ and $\beta > 0$. It is immediate to see that this would not be the case when $x = 0$, as having $\theta_0' = 0$ (that is, delegating the entire portfolio) is clearly suboptimal if $\beta > 0$. More generally, whenever there is some portfolio delegation (i.e., $\theta_0' B < W_0'$), $\theta_0'$ satisfies the first-order condition in Eq. (18), which can in this case be written as

$$E[u'(\theta_0' B + W_T^m - F(W_T^m))(W_T^m - F(W_T^m) - W_0^m)] = 0.$$  

(31)

Eqs. (30) and (31) imply that the fund investors choosing the level of delegation that ensures that a linear fee achieves first best is equivalent to having

$$E[u'(F(W_T^m))(W_T^m - F(W_T^m) - W_0^m)] = 0.$$  

(32)

However, since $u(W, W^b) = u(F(W))$ and $A(W^b) = (0, \infty)$ with linear fees, it follows from Eq. (14) that $\beta u'(F(W_T^m)) = \psi \pi_t$. Substituting in Eq. (32) and using the fact that $\pi_t W_t^m$ and $\pi_t$ are martingales gives

$$E[u'(F(W_T^m))(W_T^m - F(W_T^m) - W_0^m)] = -\psi \beta (x + \beta W_0^m) < 0.$$

Hence, given a linear fee, individual fund investors would not choose ex post the level of delegation that ensures that a linear fee achieves first best: since individual investors do not internalize the fact that fees will have to increase if they “underinvest” in mutual funds in order to continue to guarantee a given certainty equivalent to fund managers, they will always invest less than the amount needed to achieve efficient risk sharing.40

What types of compensation contracts could dominate linear fees? Since investors do not pay management fees on bonds they hold directly, but do pay fees on bonds they hold indirectly through mutual funds, it is in their best interest to select compensation contracts that induce portfolio managers to hold portfolios with high equity exposures. Contracts with convex payoffs provide a possible way to incentivize managers to increase the overall equity exposure, although, as noted by Ross (2004), this incentive does not necessarily hold over the entire state space.41

Fig. 11 plots the Pareto frontiers for purely proportional contracts and for contracts including asymmetric or load fees: for asymmetric contracts we set $\gamma_1 = x = 0$ varying $\beta$ and $\gamma_2$. For the load contract we set $\gamma_1 = \gamma_2 = 0$ varying $\beta$ and $x$. Fund managers’ certainty equivalent is on the horizontal axis. Fund investors’ (respectively, active investors’) certainty equivalent is on the vertical axis. In the top and bottom left panels correspond to the calibrated economy we have analyzed throughout the paper: $\sigma^2_0(D,T) = 0.153D$ and $\sigma^2_2(D,t) = 0.191D$. In the top and bottom right panels, $\sigma^2_0(D,t) = 0.1D$ and $\sigma^2_2(D,t) = 0.3D$. The middle panels are similar to the top panels but with the aggregate bond supply set to $B = 0.7$ (including in the riskless asset also government agency securities), instead of $B = 0.45$. The reason we consider these two middle panels is that for the $B = 0.45$ specifications, the load contract frontier is visually indistinguishable from the proportional frontier, under the resolution of the figures.42 For low managers’ reservation utilities neither asymmetric performance fees nor load fees generate significant Pareto improvements over purely proportional contracts: this is consistent with the above discussion, as when fees are low fund investors optimally choose to delegate almost their entire portfolio and the “underinvestment” relative to the amount ensuring efficient risk sharing with proportional fees is minimal. However, for high managers’ reservation utilities both load fees and asymmetric performance fees Pareto-dominate purely proportional contracts, with performance fees in turn dominating load fees. This is again what could be expected from the above discussion: when fund investors hold a significant amount of bonds in their private accounts, linear contracts with positive load fees dominate purely proportional contracts. On the other hand, any positive load fee $x$ is, given the ex post allocation chosen by fund investors, always higher than what would be needed to ensure optimal risk sharing, resulting in significant utility losses for fund investors in states in which their terminal wealth is low. This creates the potential for performance fees to dominate load fees, although, as we have shown in Section 5, performance fees have a negative impact in terms of portfolio diversification.43,44 Not all benchmark specifications lead

41 Simply restricting fund managers to trade only equity is suboptimal if the allocation to mutual funds cannot be continuously rebalanced, as it leads to significant variations over time in fund investors’ effective portfolio mix of bonds and equity.

42 In panels where the parameters differ from the calibrated parameters, we rescale $D_0^1$ and $D_0^2$, maintaining the assumption that in an economy in which all agents have direct costless access to the equity market, the equity share of stock one is 0.69. In the top and bottom right panels $D_0^1 = 1.153$ and $D_0^2 = 1.568$, in the left middle panel $D_0^1 = 1.402$ and $D_0^2 = 0.733$, in the right middle panel $D_1^1 = 1.070$ and $D_0^2 = 1.249$.

43 Some empirical support for the existence of welfare gains associated with the use of performance fees is provided by Coles, Suay, and Woodbury (2000), who find that closed-end funds that use performance fees tend to command a premium that is about 8% larger than similar funds that do not use these fees.

44 The Pareto benefits of using performance fees with the $T = 1$ horizon economy, while present, are very small.
to a dominance of asymmetric performance fees over load fees (for example, compare in the left panel of the middle row the frontier with load fees to the frontier with stock one as the benchmark). For each of the economies considered in the figure there exists a benchmark specification for which the corresponding asymmetric performance fee frontier dominates the frontier with load fees, however, this may not be a general result. In Fig. 11 the different benchmark portfolios are exogenously fixed: thus, selecting the benchmark optimally would lead to an even stronger dominance of performance fees over proportional fees. We will return to this point in Section 6.1.

Fig. 11. The graph plots the Pareto frontier at time $t=0$ with asymmetric performance fees and three different benchmark portfolios: stock 1 (solid line), stock 2 (dotted line), and the market portfolio of risky assets (dashed line). Also plotted is the Pareto frontier with proportional fees (dot-dashed line), and in the middle two panels with load fees (long-dashed line). Managers' certainty equivalent is on the horizontal axis. Fund investors' certainty equivalent is on the vertical axis in the top four panels, and active investors' certainty equivalent is on the vertical axis in the bottom two panels. In each panel, we calibrate $D_0^1$ and $D_0^2$ so that in an economy in which all agents have direct costless access to the equity market, the equity share of stock one is 0.69. Plot ranges are restricted so that the difference between the different frontiers is clearly visible. Investment horizon is set to $T=5$. 
Moving along the efficient frontier with asymmetric performance contracts and an exogenously fixed benchmark, increases in fund managers’ certainty equivalent are associated with simultaneous increases in both the proportional fee component and the asymmetric performance fee component. Focusing on the calibrated economy: the optimal proportional fee parameter $\beta/T$ increases from 0% to 8.58%, while the optimal performance sensitivity parameter $\gamma_2$ increases from 0% to 11.41%. In particular, the model implies a positive correlation between managers’ overall compensation and contract performance sensitivity, which appears to be consistent with anecdotal empirical evidence. The optimal performance sensitivity parameter that corresponds to the calibrated proportional fee of $\beta/T = 2\%$ is only $\gamma_2 = 0.75\%$, considerably lower than the 20% observed in practice. The model also implies a positive correlation between the performance sensitivity parameter and both expected returns and Sharpe ratios (not shown). Fund investors’ endogenously chosen allocation to mutual funds decreases along the frontier from 100% to 42.6% of their endowment. As a result of this quickly decreasing allocation to mutual funds, the number of shares of both stocks held by mutual funds is a monotonically decreasing function of $\gamma_2$, in spite of the increasing proportional allocation to stocks by mutual funds. Hence, the active investors’ proportional allocation to both stocks must be an increasing function of $\gamma_2$. This is obtained by expected returns and Sharpe ratios increasing along the frontier. Similar patterns exist for the other frontiers.

Within our model, for reported results as well as for all other parameter specifications we have considered, fulcrum fees never generate a Pareto improvement over purely proportional fees: along the fulcrum frontier $\gamma_1 = \gamma_2 = 0$, coinciding with the proportional frontier.\(^{45}\) Adding a fulcrum fee to a proportional contract increases the welfare of fund investors (since fulcrum fees also induce an increase in the equity allocation chosen by fund managers) but strongly decreases the welfare of fund managers, due to the utility losses in states in which the excess return of the managed portfolio over the benchmark is negative.

In the asymmetric contracts considered in Fig. 11, no penalty is imposed on managers for underperforming the benchmark. However, performance contracts can include a penalty, while remaining convex ($0 < \gamma_1 < \gamma_2$). This can potentially enhance welfare as it partially mitigates the distortions in fund managers’ allocations that are induced in the region where the incentive to maximize the variance of the excess return of the managed portfolio over the benchmark dominates. Fig. 12 considers adding and varying the performance sensitivity penalty at points along the efficient frontier when stock 1 is the benchmark. For high fund manager certainty equivalents there is a benefit, although fairly small: the figure is restricted to a very narrow range at the very top of the viable manager certainty equivalent range. For lower manager certainty equivalents there is no benefit (i.e., it is best to set $\gamma_1 = 0$). Focusing on the calibrated economy, the optimal penalty performance sensitivity parameter is zero below a certainty equivalent of 0.36 and increases along the upper part of the frontier up to 2.8%. Comparing the two panels shows that the qualitative impact on fund investors’ certainty equivalent of adding a penalty can differ depending on economic fundamentals. The differential impact is because while in the calibrated economy (top panel) managers’ incentives to maximize the variance of the excess return of the managed portfolio over the benchmark dominate, in the lower panel hedging incentives dominate so that an introduction of a penalty exacerbates fund managers’ portfolio distortions. The impact on managers is similar across the two panels, as in both cases the allocation by fund investors increases with the introduction of the penalty.

### 6.1. Optimal benchmarking

In order to understand the impact of delegated portfolio management on equilibrium prices, in an empirically relevant setting, our numerical analysis has focused on a calibrated example. In this calibrated example the two liquidation dividends have fairly similar distributions and the benchmark portfolio is exogenously specified to be the stock with the lower volatility and higher market share. However, in equilibrium, the composition of the benchmark portfolio should be determined endogenously and so we briefly address the robustness of our results to optimal benchmarking in the context of asymmetric performance fees.

When optimally determining the composition of the benchmark, fund investors seek to maximize the overall equity allocation by fund managers while at the same time minimizing the investment distortions induced by benchmarking. In general, the portfolio held by fund managers is a (dynamically rebalanced) combination of the market portfolio, the benchmark portfolio, and the riskless asset. Therefore, when the benchmark portfolio coincides with the market portfolio of risky assets, fund managers’ equilibrium trading strategy consists of a combination of this portfolio and the riskless asset, eliminating the distortions in the allocation to risky assets induced by benchmarking and benefiting fund investors. On the other hand, because it is suboptimal in this case for fund managers to increase the tracking error volatility of their portfolio (and hence the expected value of their performance fees) by holding a portfolio of risky assets that deviates from the benchmark portfolio, they optimally choose to reduce the risk of their compensation by holding a portfolio characterized by a lower overall equity allocation. Clearly, this latter effect is detrimental to fund investors.

When the two liquidation dividends are identically distributed and the aggregate supplies of both stocks are

---

\(^{45}\) Das and Sundaram (2002) find that asymmetric performance fees also dominate fulcrum fees in a signaling model with fund managers of different skills, although this dominance arises in their model for a completely different reason: the ability of fund managers to more easily signal their skill, and thus to extract a higher surplus, in the presence of fulcrum fees. Ou-Yang (2003) provides a model in which investors are assumed to delegate the management of their entire portfolio and fulcrum fees are optimal.
the same, it is easy to see that the optimal benchmark portfolio must coincide with the market portfolio of risky assets, as the equilibrium must then be fully symmetric in the two stocks. As a result, all cross-sectional price distortions are eliminated in this case, although delegated portfolio management still has an impact on equilibrium risk premiums and volatilities.46

This, however, is a knife-edge case: if the two liquidation dividends are not identically distributed, the market portfolio of risky assets is not the optimal benchmark. This is highlighted in the right top and middle panels of Fig. 11, for the case $\sigma_1^2(D_t) = 0.1D$ and $\sigma_2^2(D_t) = 0.3D$. Interestingly, in this case, benchmarking fund managers to either asset results in welfare gains over benchmarking them to the market portfolio of risky assets.47 In the left top and middle panels, the frontier with the market portfolio as the benchmark Pareto-dominates the frontiers corresponding to the benchmark consisting of either of the individual assets. However, it does not coincide with the optimal benchmark which is tilted toward the higher volatility asset. In general, the optimal benchmark tilt accounts for the induced tilt in fund managers’ portfolio allocations, so that it increases fund volatility, and hence the systematic risk and the expected return of the fund’s portfolio.

Naturally, the percentage gain in certainty equivalent relative to using a proportional fee increases along the frontier. For the calibrated economy (top left panel): with stock 1 (respectively, the market portfolio) as the benchmark the maximal percentage gain of fund investors’ certainty equivalent is 1.1% (2.3)% and of fund managers’ certainty equivalent is 0.1% (0.6%). For an optimal contract (not shown) the corresponding values are 2.5% and 0.7%.48

With larger differences in the volatilities of the two dividend processes, gains are larger. For example, in the

---

46 In an earlier version of the paper we have confirmed this numerically.

47 A similar pattern emerges if we maintain the assumption that $\sigma_1^2(D_t) = 0.1D$ and $\sigma_2^2(D_t) = 0.3D$, but recalibrate $D_1^0$ and $D_2^0$ so that in an economy in which all agents have direct costless access to the equity market, the equity share of stock 1 is 0.5.

48 The optimal benchmark at the high end of managers’ certainty equivalent is 34% stock 1 and 64% stock 2. There is a slight benefit to allowing a performance penalty, but this benefit is even smaller than the one in the top panel of Fig. 12. When managers’ certainty equivalent is 0.18, percentage gain increase is 0.01%.
top right panel with stock 1 as the benchmark the maximal percentage gain to fund investors is 4.1%, and to fund managers is 2.5%. The bottom two panels show that similar to the pattern observed in Table 2, along the frontier active investors indirectly benefit as the cost to fund investors of acquiring delegation services increases. However, the benefit is smaller when asymmetric performance fees are used to enhance fund investors’ welfare, compared to when proportional fees are used. This indirect benefit is the lowest when the optimal contract is used.

Finally, as it should be clear, the optimal benchmark not being the market portfolio of risky assets is all that is needed for the qualitative results described in Section 5 to hold.

7. Conclusion

We have examined the impact of delegated portfolio management on equilibrium prices within a dynamic general-equilibrium setting in which the parameters of the management contract, the extent of delegated portfolio management, and the returns of both benchmark and non-benchmark securities are all determined endogenously.

When fund managers receive performance fees of the fulcrum type, they optimally choose to tilt their portfolios toward stocks that are part of the benchmark: in equilibrium, this results in a significant positive price effect associated with the addition of a stock to the benchmark and a smaller negative price effect associated with a deletion. These implications are consistent with empirical evidence regarding changes in the composition of the S&P 500 index (the most widely used benchmark portfolio). Everything else being the same, we also find that benchmark stocks have lower expected returns, lower Sharpe ratios, and higher volatilities than similar non-benchmark stocks.

With asymmetric performance fees, the composition of the portfolios selected by fund managers depends critically on the funds’ excess return. As a result, cross-sectional differences between benchmark and non-benchmark stocks can have either sign, depending on the funds’ performance relative to the benchmark. Interestingly, the presence of portfolio managers receiving asymmetric performance fees can stabilize prices by decreasing the equilibrium stock volatilities of both benchmark and non-benchmark stocks, although portfolio turnovers are higher with asymmetric fees.

Previous literature has typically taken asset returns and the level of delegated portfolio management as given, analyzing managers’ portfolio choice decisions and deriving optimal contracts. We have complemented this literature by analyzing the asset pricing implications of a prevalent parametric class of existing contracts, when the level of delegated portfolio management is determined endogenously. We have also demonstrated that when full delegation is not exogenously imposed, performance contracts may be welfare-improving, even when there is no asymmetric information and investors and fund managers have CRRA preferences with the same risk aversion coefficient.

Appendix A

To simplify the notation, throughout this appendix we frequently suppress the explicit dependence of quantities such as \( W, W_1, \) and \( W_2 \) on \( W^b \).

Proof of Lemma 1. Suppose first that \( \beta + \gamma_1 \neq 0 \) and consider the system of equations

\[
\begin{align*}
\{ & \frac{1}{\beta + \gamma_1} z = -\gamma_1 \delta W^b + u'(z - \gamma_1 \delta W^b)(\beta + \gamma_1), \\
& \frac{1}{\beta + \gamma_2} z = -\gamma_2 \delta W^b + u'(z - \gamma_2 \delta W^b)(\beta + \gamma_2),
\}
\]

in the unknowns \( (W_1, W_2) \). Direct computation shows that the system has the unique solution

\[
\begin{align*}
W_1 &= \frac{1}{\eta - 1} \left( \frac{z - \gamma_1 \delta W^b}{\beta + \gamma_1} - \frac{z - \gamma_2 \delta W^b}{\beta + \gamma_2} \right), \\
W_2 &= W_1 + \frac{1}{\beta + \gamma_1} \left( \frac{z - \gamma_1 \delta W^b}{\beta + \gamma_1} - \frac{z - \gamma_2 \delta W^b}{\beta + \gamma_2} \right),
\end{align*}
\]

where \( \eta = ((\beta + \gamma_2)/(\beta + \gamma_1))^{1-1/c} \). Moreover,

\[
\gamma_1 \delta W^b - z < W_1 < \delta W^b < W_2.
\]

Letting \( W(W^b) = (\gamma_1 \delta W^b - z)/(\beta + \gamma_1) \) (as in Eq. (9)) and \( W_1(W^b) = (W_1)^+ \) (as in Eq. (13)), it follows from the above inequality that \( W \leq W_1 < \delta W^b \).

If \( W_1(W^b) > 0 \), it then immediately follows from the system (33) that Eqs. (11) and (12) are satisfied, with equality in Eq. (12) and \( W_2(W^b) = W_2 \), establishing the lemma for this case.

If, on the other hand, \( W_1(W^b) = 0 \), then \( W_1 \leq 0 \) and it follows from Eq. (35) that \( z - \gamma_1 \delta W^b > 0 \). Therefore, the function

\[
f(W) = u(z - \gamma_1 \delta W^b) - u(\beta + \gamma_2(\beta - \delta W^b)) + u'(z - \gamma_2 \delta W^b)(\beta + \gamma_2) W
\]

is well-defined for \( W \geq W^b \) and the existence of a solution \( W_2(W^b) > \delta W^b \) to Eq. (11) is equivalent to the existence of a zero of \( f \) in \( \delta W^b, \infty \). Clearly, \( f \) is continuous and strictly decreasing on \( [\delta W^b, \infty) \). Moreover,

\[
f(\delta W^b) = u(z - \gamma_1 \delta W^b) - u(\beta + \gamma_2(\beta - \delta W^b)) + u'(z - \gamma_2 \delta W^b)(\beta + \gamma_2) \delta W^b
\]

is well-defined for \( W \geq W^b \). The existence of a solution \( W_2(W^b) > \delta W^b \) to Eq. (11) is equivalent to the existence of a zero of \( f \) in \( \delta W^b, \infty \). Clearly, \( f \) is continuous and strictly decreasing on \( [\delta W^b, \infty) \). Moreover,

\[
f(\delta W^b) = u(z - \gamma_1 \delta W^b) - u(\beta + \gamma_2(\beta - \delta W^b)) + u'(z - \gamma_2 \delta W^b)(\beta + \gamma_2) \delta W^b
\]

where the second (in)equality follows from the fact that \( u' \) is decreasing and \( W^b > \delta W^b \), the third from the first equation in the system (33), the fourth from the fact that \( u' \) is decreasing and \( W_1 < 0 \), and the last from the fact that
the function \( u(x + \beta W + \gamma_1(W - \delta W))^2 \) is concave in \( W \). Similarly,
\[
\begin{align*}
 f(W_2) &= u(x + \gamma_1 \delta W - u(x + \beta W + \gamma_2 (W_2 - \delta W))^2) \\
 &= u((x + \beta W_2^* + \gamma_1 (W_2^* - \delta W))^2) + u(x + \gamma_2 (W_2^* - \delta W))^2 \\
 &\leq 0,
\end{align*}
\]
where the second equality follows from the second equation in the system (33), while the inequality follows from the concavity of the function \( u(x + \beta W + \gamma_1 (W - \delta W))^2 \). Therefore, there exists a unique \( W_2^* \) in \( \delta W_2^* \) such that \( f(W_2^*) = 0 \), thus establishing the existence of a unique solution to Eq. (11). The inequality in Eq. (12) follows from the fact that in this case, we have
\[
 u'((x + \beta W_2^* + \gamma_1 (W_2^* - \delta W))^2) + u(x + \gamma_2 (W_2^* - \delta W))^2 < 0,
\]
where the first (in)equality follows from the fact that \( u' \) is decreasing and \( W_2^* \geq W_1^* \), the second from the first equation in the system (33), and the last from the fact that \( u' \) is decreasing and \( W_2^* < W_2 \).

Finally, suppose that \( \beta + \gamma_1 = 0 \) and let \( W_2^* = 0 \) (as in Eq. (9)) and \( W_1^* = 0 \). With \( f(W) \) defined as in Eq. (36), we then have
\[
 f(\delta W) = u'(x) \gamma_1 \delta W > 0.
\]
Since
\[
 2(x + \gamma_2 (W - \delta W))^2 \gamma_2 (W - \delta W) < 0 \quad \text{as} \quad W \to +\infty
\]
and \(-Wu'(W)/u(W) = c\) for all \( W \), there exists a \( W_3^* > \delta W \) such that
\[
 2(x + \gamma_2 (W_3^* - \delta W))^2 \gamma_2 (W_3^* - \delta W) < c
\]
or
\[
 u(x + \gamma_2 (W_3^* - \delta W))^2 \gamma_2 (W_3^* - \delta W) < 0.
\]
Moreover, since \( u'' > 0 \), it follows from a second-order Taylor expansion of the function \( u(x + \gamma_2 (W - \delta W))^2 \) around the point \( W_3^* \) that
\[
 u(x) \leq u(x + \gamma_2 (W_3^* - \delta W)^2) - u(x + \gamma_2 (W_3^* - \delta W))^2 \gamma_2 (W_3^* - \delta W) \\
 + \frac{1}{2} u'(x + \gamma_2 (W_3^* - \delta W))^2 \gamma_2^2 (W_3^* - \delta W)^2. \tag{37}
\]
Hence,
\[
 f(W_3^*) = u(x) - u(x + \gamma_2 (W_3^* - \delta W)^2) + u'(x + \gamma_2 (W_3^* - \delta W)^2) \gamma_2 (W_3^* - \delta W) \\
 < u'(x + \gamma_2 (W_3^* - \delta W))^2 \gamma_2 (W_3^* - \delta W).
\]

where the first inequality follows from Eq. (38), while the second follows from Eq. (37). Therefore, \( f \) has a zero on \( (\delta W_2, W_2) \), thus establishing the existence of a solution to Eq. (11) in this case. The inequality in Eq. (12) is trivially satisfied, since the left-hand side equals zero in this case. \( \square \)

**Proof of Lemma 2.** Eq. (11) shows that \( v_1(W, \delta W) \) is continuous at \( W_2(W, \delta W) \), while Eq. (12) shows that \( v_2(W, \delta W) \) is continuously differentiable at \( W_1(W, \delta W) \). If \( W_1(W, \delta W) > 0 \) (that is, if \( W_1(W, \delta W) > W_2(W, \delta W) \)), then it immediately follows from the definition of \( v \) that \( v_1(W, \delta W) \) is continuously differentiable and concave on \( [W, W_2(W, \delta W), \infty) \). Moreover, since \( v_1(W_1(W, \delta W), \delta W) = u(F(W_1(W, \delta W)), \delta W) \) and \( u(F(W_1(W, \delta W)), \delta W) \) is strictly concave on the interval \( (W_1(W, \delta W), \delta W) \), \( v_1(W, \delta W) \) is linear, it follows that \( v_1(W, \delta W) > u(F(W_1(W, \delta W)), \delta W) \) on that interval. A similar argument implies that \( v_2(W_2(W, \delta W), \delta W) \) is concave and \( v_2(W_2(W, \delta W), \delta W) \) is linear on that interval. Therefore, \( v \) is the smallest concave function \( v(W, \delta W) \) with \( v(W_1(W, \delta W), \delta W) > u(F(W_2(W, \delta W)), \delta W) \). \( \square \)

**Proof of Proposition 1.** Standard optimization theory implies that a policy \( W_r \) is optimal in the concavified problem (10) if and only if \( W_r \) satisfies the conditions of Proposition 1 with Eq. (14) replaced by the weaker condition
\[
 W_r^m \in v(W_r^m)^{-1}(\psi^m \pi_T, \delta W_r^m) \quad \text{if} \quad W_r^m > 0.
\]
Thus, a policy \( W_r^m \) satisfying the conditions of Proposition 1 is optimal in the concavified problem (14). The fact that \( W_r^m \) takes values in \( A(W_r^m) \) implies \( u(F(W_r^m, \delta W_r^m)) = v(W_r^m, \delta W_r^m) \). Since \( u(F(W_2(W, \delta W), \delta W)) \leq v(W_2(W, \delta W), \delta W) \), optimality of \( W_r^m \) in the concavified problem (10) then implies optimality in the original problem (8). This proves that the conditions in Proposition 1 are sufficient for optimality.

Necessity follows immediately by noticing that if \( W_r \) is any other feasible policy, then
\[
 E[u(F(W_r^m, W_r^m))] \leq E[u(F(W_2(W, \delta W), \delta W))] \\
 \leq E[u(W_r^m, W_r^m)] \\
 = E[u(W_r^m, W_r^m)],
\]
with the first inequality being strict if \( W_r \not\in A(W_r^m) \), and the second inequality being strict otherwise. \( \square \)

**Proof of Theorem 1.** It can be easily verified that \( g^m(y, W_r, \delta W) = (v_r^m(y, W_r))/y \) for \( y \neq v_r(W_r, \delta W) \). The claim then immediately follows from Proposition 1. \( \square \)

**Proof of Theorem 2.** The proof of Theorem 2 is straightforward and is thus omitted. \( \square \)
Proof of Corollary 1. Under the stated assumptions, $\gamma_1 = \gamma_2$ and $x_1 - \gamma_1 \delta W^T \leq 0$,
so that $H_1(D_T, W^T) = H_3(D_T, W^T) \geq H_4(D_T, W^T)$.

It then follows from Eq. (21), using the fact that $\beta + \gamma_1 = \beta + \gamma_2 > 0$, that $\mathbb{P}(\mathcal{I}(D_T, W^T) > \mathcal{I}(D_T, W^T))$.

In addition, since by Lemma 2 $W^T = W^T_{2}$ when $\gamma_1 = \gamma_2$, it follows from (16) and (17) that
\[
W^T = \mathcal{G}(\mathcal{P}, T, P_T)
\]
\[
\mathcal{G}(\mathcal{I}(D_T, W^T)_{1/(\beta + \gamma_2)}, \mathcal{I}(D_T, W^T)) = \theta_T^m \cdot D_T,
\]
where
\[
\theta_T^m = \frac{\phi(1-\lambda)^2}{1-\lambda^2} + \beta + \gamma_2\lambda^2 + \theta_0^m + \phi(1-\lambda)^2 - \theta_0^m + \phi(1-\lambda)^2.
\]

It then immediately follows that the optimal trading strategy for the fund managers is as given in Eq. (28) for all $t \in [0, T]$. The optimal trading strategy for the active investors follows from the market-clearing conditions.

Next, Eq. (25) implies that $\int_{S} - S^2 \geq 0$ if and only if $E[\mathcal{I}(D_T, W^T)] \geq 0$.

Let $\mathcal{D}_T$ be a two-dimensional random variable obtained from $D_T$ by exchanging the values of $D_1$ and $D_2$. Since $D_1$ and $D_2$ are identically distributed conditional on the information at time $t$, we have
\[
E[\mathcal{I}(D_T, W^T)] = E[\mathcal{I}(D_T, W^T)] - E[\mathcal{I}(D_T, W^T)]_{1} + E[\mathcal{I}(D_T, W^T)]_{1} - E[\mathcal{I}(D_T, W^T)]_{1} + E[\mathcal{I}(D_T, W^T)]_{1} - E[\mathcal{I}(D_T, W^T)]_{1}.
\]

The claim regarding the equilibrium stock prices then easily follows from the fact that, when $\gamma_2 \neq 0$ and $D_1^2 > D_2^2$, $H_4(D_T, \theta^m) > H_1(D_T, \theta^m)$ (respectively, $H_4(D_T, \theta_1^m < H_1(D_T, \theta_1^m)$ if and only if the benchmark portfolio $\theta^m$ holds more (respectively, less) of stock 1 than of stock 2).

Lemma A.1. If $D_1^2$ and $D_2^2$ are log-normally distributed, then $\frac{D_1^2}{D_2^2}$ is log-normally distributed.

Proof of Lemma A.1. Let
\[
D_1^2 = D_1^2 \cdot \omega_1, \quad D_2^2 = D_2^2 \cdot \omega_2,
\]
where $m_1 = (\mu_1 - \frac{1}{2} \sigma_1^2) \cdot t$, $m_2 = (\mu_2 - \frac{1}{2} \sigma_2^2) \cdot t$, and $\omega_1$ and $\omega_2$ are normal standard variables with correlation coefficient $\rho$.

Define
\[
s^2 = \sigma_1^2 - 2 \rho \sigma_1 \sigma_2, \quad s_{12} = \sigma_1^2 - \rho \sigma_1 \sigma_2 \cdot t.
\]

References


