

# Delegated Portfolio Management, Optimal Fee Contracts, and Asset Prices\*

Yuki Sato<sup>†</sup>

University of Lausanne and Swiss Finance Institute

May 1, 2016

## Abstract

This paper proposes a model of asset-market equilibrium with portfolio delegation and optimal fee contracts. Fund managers and investors strategically interact to determine funds' investment profiles, while they share portfolio risk through fee contracts. In equilibrium, their investment decisions, fee schedules, and stock price feed back into one another. The model predicts that (1) stock market's expected return and volatility increase as more investor capital is intermediated by funds, (2) fund's expense ratio is stable despite volatile market, (3) aggregate fund flow is positively (inversely) related to subsequent (past) market return, and (4) funds provide investors with a volatility hedge by adjusting market exposure counter-cyclically.

*JEL classification:* G11, G12, G23

*Key words:* portfolio delegation, optimal fee, asset prices, price volatility, fund size, fund return

---

\*I benefited from suggestions from Xavier Vives (the editor) and two anonymous referees. For helpful comments, I thank Elias Albagli, Margaret Bray, Amil Dasgupta, Bernard Dumas, Vincent Fardeau, Itay Goldstein, Amit Goyal, Denis Gromb, Stéphane Guibaud, Joel Peress, Norman Schürhoff, Dimitri Vayanos, and seminar participants at INSEAD and London School of Economics. Financial support from the Swiss Finance Institute and the Paul Woolley Centre for the Study of Capital Market Dysfunctionality at LSE are gratefully acknowledged. All errors and omissions are my own.

<sup>†</sup>Department of Finance, University of Lausanne, UNIL-Dorigny, Extranef 235, CH-1015 Lausanne, Switzerland. E-mail: yuki.sato@unil.ch.

# 1 Introduction

The fund management industry has grown significantly in the past few decades.<sup>1</sup> Reflecting that, recently more and more papers have analyzed economic implications of delegated portfolio management. While some have considered asset-pricing implications and others have considered contracting implications, few have studied both in a unified model. The goal of this paper is to develop a tractable model of such a kind and draw empirical implications from the equilibrium relations of delegation and asset prices.

Our model has two distinguishing features compared to the existing ones. First, fund fees are determined in optimal contracting problems and interact with equilibrium stock prices. The fee arrangement serves as an optimal scheme of sharing portfolio risk among fund managers and fund investors, linking the principal-agent relationship in portfolio delegation and the stock market performance. Second, our model incorporates strategic behavior of both fund managers and fund investors in an asset-pricing model. This feature creates a feedback loop between both of these agents' optimal investment decisions and the stock market performance, allowing us to derive a rich set of predictions on the time variation of stock returns, fund returns, fund fees, fund flows, and investment strategies.

Specifically, we consider a discrete-time, infinite-horizon model with a risky stock and a riskless asset. The stock yields a normally distributed dividend each period, and the price is determined by market clearing. The riskless asset has an infinitely elastic supply at an exogenous rate of return and is freely accessible to all agents. There are three types of agents: direct investors, fund investors, and fund managers. All of them are risk-averse (CARA) price takers who live for two periods, constituting overlapping generations. The direct investors can access the stock on their own, whereas the fund investors cannot. To invest in the stock, the fund investors need to delegate capital to a fund manager who manages a portfolio of the assets on their behalf. This delegation is subject to an agency problem: potentially, each manager can abscond with the fund investors' assets. To prevent it from happening, the fund investors offer to the manager a fee contract that is attractive enough for him to continue running the fund rather than abscond. The model's key endogenous variables are (1) fund size, measured by assets under management

---

<sup>1</sup>Cuoco and Kaniel (2011) report that in the U.S. as of 2004, mutual funds managed assets more than \$8 trillion, hedge funds managed about \$1 trillion and pension funds more than \$12 trillion. According to Investment Company Institute (2014), in the U.S., the number of households owning mutual funds has increased from 23.4 million in 1990 to 56.7 million in 2013. In 2012, 23% of households' financial assets are managed by registered investment companies (household holdings of ETFs, closed-end funds, unit investment trusts, and mutual funds), while that proportion was only 3% in 1980.

(AUM), (2) fund's risk exposure, measured by the number of shares of the stock held per dollar of investor capital, (3) fund's fee schedule, which we allow to be any function of the fund's investment profit (i.e., the proceeds from the fund portfolio in excess of the riskless component), and (4) the stock price.

The model's equilibrium involves a feedback loop among these endogenous variables: fund investors and managers take into account the fee schedule when choosing the capital allocation and the fund's risk exposure, while the fee schedule depends on these agents' choices, and the stock price adjusts to make certain all these actions are consistent with market clearing. Due to the CARA-normal specification, the optimal fee schedule is (not surprisingly) in an affine form. The stock price, the fund's risk exposure, and AUM are all stochastically time varying, reflecting random shocks to the stock supply. This feature allows us to study issues of significant interest such as stock price volatility, time variations of fund trading strategies and fund flows, and their interactions with stock market return. All solutions are obtained in simple closed form.

The model yields a number of testable predictions. It predicts that the expected stock market return and the stock price volatility increase with the extent of delegation as measured by the proportion of fund investors. The intuition is as follows. As more investors delegate capital to funds, the AUM per fund also increases. Consequently, each fund manager is exposed to a larger market risk because he receives a fee proportional to the fund's investment profit that increases with the AUM and varies with the stochastic stock-market performance. To compensate this risk bearing, each manager demands a larger risk premium, leading to a higher expected stock market return in equilibrium. Moreover, as the AUM increases, each fund's demand for the stock becomes less price-elastic because the manager tries to reduce the variation of the fund's stock holding to counterbalance his personal risk bearing. In equilibrium, the aggregate demand with lower price elasticity, being equated to a random supply, yields a market-clearing price with higher volatility. To see the intuition, consider the classical demand-supply diagram with price on the vertical axis and quantity on the horizontal axis. Low price elasticity means that the slope of the downward-sloping demand curve is steep. Thus, as the vertical supply curve fluctuates due to the supply shocks, the resulting market-clearing price at the intersection of these two curves exhibits large volatility.

Furthermore, the model provides explanations to the empirical observations such as (1) a positive relation between aggregate fund flows and subsequent stock returns and a negative relation between returns and subsequent flows (Warther 1995), and (2) mutual funds' counter-cyclical volatility timing, that is, they decrease their market exposure when

market volatility is high (Busse 1999; Giambona and Golec 2009). Also, we distinguish market return and fund return, identify their distributions, and offer new testable implications on their equilibrium relations: (1) the expected fund return is less volatile than the expected market return, and (2) the fund return volatility comoves negatively with the market return volatility, that is, funds provide investors with a volatility hedge.

Last, we study a model with infinitely-lived agents. While the analysis is less tractable due to the agents' hedging considerations, we show that all the qualitative results obtained in the overlapping-generations model still carry through in such a setting.

This paper relates to the growing literature that discusses equilibrium implications of delegated portfolio management (Allen and Gorton 1993; Shleifer and Vishny 1997; Vayanos 2004; Malliaris and Yan 2012; Basak and Pavlova 2012; Vayanos and Woolley 2013; Sato 2014). The following papers especially relate to this paper.

Buffa, Vayanos, and Woolley (2014) is most closely related, as they also study the joint determination of fund fees and asset prices. They focus on fund managers' contractual incentives and their impact on fund strategies, and do not study investors' endogenous capital allocation (i.e., AUM). In contrast, in our model both contractual incentives and AUM play critical roles. It is key for our results that investors' strategic choices of AUM and managers' choices of fund strategies interact with each other.<sup>2</sup>

He and Krishnamurthy (2012) study the relation between portfolio delegation and asset prices in a dynamic setting. In their model, the intermediary's capital scarcity in the "crisis" situation is central to the equilibrium characterization. In contrast, all of our results hold in non-crisis situations; the driving forces are fee contracts and strategic interactions between fund investors and managers, both of which are absent in their model.

Like our paper, Kaniel and Kondor (2013) study equilibrium asset prices and trading strategies of fund managers with concern for fund flows. While the investors' capital provision is exogenous in their paper, in ours it is one of the main endogenous variables for which we draw empirical implications. Also, their model does not discuss the optimality of fees or the volatility of asset prices, which are main focus of our paper.

Cuoco and Kaniel (2011) relates to our paper in that they also study the asset pricing implications of delegated portfolio management. While the fee structure in their model is exogenously imposed, it is optimal in ours. Also, while investors' capital provision and fee payment occur only once in their model, in ours they take place continually; thus, we can study the time variation of fund flows and fees.

---

<sup>2</sup>Moreover, while Buffa, Vayanos, and Woolley (2014) assume an affine form of fees and determine the optimal parameter values, in ours the affine form is endogenously obtained.

Ou-Yang (2003) is related to ours, as he also derives a fee contract in a portfolio delegation problem. In his model, asset prices are exogenously given, whereas they are endogenous in ours; studying the equilibrium interaction between asset prices and fees is the primary purpose of our paper. Also, his model takes the investor’s capital provision as exogenous and does not discuss the fund’s asset size, which is our focus in this paper.

The paper proceeds as follows. Section 2 presents the model. Section 3 characterizes the benchmark case without delegation. Section 4 analyzes the equilibrium with delegation. Section 5 discusses the empirical implications. Section 6 studies a model with long-lived agents. Section 7 concludes. The proofs and robustness checks are in the online appendix.

## 2 Model

Time  $t$  is discrete and runs from zero to infinity. There is a single consumption good referred to as a dollar. There is a risky stock and a riskless asset. The riskless asset has an infinitely elastic supply at an exogenous rate of return  $r > 0$  and is freely accessible to all agents.<sup>3</sup> There are three types of agents: *fund investors*, *direct investors*, and *fund managers*. All of them are risk averse and competitive (i.e., price takers). They live for two periods and constitute overlapping generations.<sup>4</sup> The fund investors can access the stock only through fund managers, who collect fees from the investors. The direct investors access the stock on their own. The fee schedule, the investors’ capital allocation, the managers’ portfolio choice, and the stock price are determined endogenously.

### 2.1 Stock

In each period  $t = 1, 2, \dots$ , the stock yields a dividend  $\delta_t$  per share, which is i.i.d. across time, normally distributed with mean  $\hat{\delta} > 0$  and variance  $\sigma^2 > 0$ . The stock is traded in the market at market-clearing price  $P_t$ . The stock’s supply  $S_t$  stochastically evolves over time as  $S_t = S_{t-1} + \epsilon_t$ , where noise  $\epsilon_t$  is normally distributed with mean 0 and variance

---

<sup>3</sup>It is standard in the literature to assume an exogenous, constant riskless rate in models with CARA-normal asset-pricing models (see, for example, Wang 1993, Spiegel 1998, Watanabe 2008, Banerjee 2011, and Vayanos and Woolley 2013). Moreover, Sundaresan (1983) provides a justification for a constant interest rate in a general equilibrium model with CARA agents. In Appendix G.1, we endogeneize the risk-free rate in the fixed supply case and show that our main result still carries through.

<sup>4</sup>The assumption of overlapping generations simplifies the analysis and allows us to obtain intuitive closed-form solutions. But the assumption is not critical to our main results. Section 6 presents a model with infinitely-lived agents and shows that all of our qualitative results still go through.

$\sigma_s^2 \geq 0$ .<sup>5</sup> Due to the stochastic nature of  $S_t$ , the price and agents' actions will be time varying in equilibrium.<sup>6</sup> We denote the stock's excess return per share of the stock by

$$R_{t+1} \equiv \delta_{t+1} + P_{t+1} - (1 + r)P_t. \quad (2.1)$$

## 2.2 Agents

In each period  $t$ , a continuum with mass one of investors and a continuum with mass  $\beta > 0$  of fund managers are born (“generation  $t$ ”). Each investor is endowed with  $W > 0$  dollars. An exogenous proportion  $\alpha \in (0, 1)$  of investors are fund investors, and the rest are direct investors. Each manager has zero endowment and runs an investment fund. Two parameters  $(\alpha, \beta)$ , representing the extent of portfolio delegation in the economy, are key comparative statics parameters in the following analysis. The letters “ $f$ ,” “ $d$ ,” and “ $m$ ” are used to represent fund investors, direct investors, and fund managers, respectively. The generation- $t$  agents make investment decisions in period  $t$  and derive utility from their terminal wealth in period  $t + 1$ . Each agent has a utility function  $u^i(W_{t+1}^i) = -\exp(-\gamma^i W_{t+1}^i)$ ,  $i \in \{f, m, d\}$ , where  $W_{t+1}^i$  is her/his terminal wealth and  $\gamma^i > 0$  is a coefficient of absolute risk aversion. When  $\alpha \rightarrow 0$  and  $\beta \rightarrow 0$ , the equilibrium of our model converges to that of Spiegel (1998) where delegation is irrelevant.

## 2.3 Portfolio management

For simplicity, we assume that fund investors are equally allocated across funds, that is, each fund has a mass  $\alpha/\beta$  of investors.<sup>7</sup> Each fund investor allocates  $X_t \geq 0$  dollars of capital to the fund in period  $t$ , where  $X_t$  is an endogenous variable. Given  $X_t$  dollars of capital, the manager purchases  $\theta_t X_t$  shares of the stock, where  $\theta_t \in (-\infty, \infty)$  is also endogenously determined. We refer to  $\theta_t$  as the fund's *risk exposure*. The rest of the

---

<sup>5</sup>The assumption that  $S_t$  is a random walk simplifies the analysis but is stronger than necessary. Any AR(1) process  $S_t = a_0 + a_1 S_{t-1} + \epsilon_t$ , where  $a_0 \geq 0$  and  $a_1 \in [0, 1]$  are constants and  $\epsilon_t$  is a normally distributed noise, would preserve all the results qualitatively. See Appendix G.2 for details.

<sup>6</sup>It is standard to study price volatility by assuming random stock supply. See Spiegel (1998), Watanabe (2008), and Banerjee (2011) for overlapping generations models, and Campbell and Kyle (1993) and Wang (1993) for infinitely lived agent models. We could alternatively assume a fixed stock supply  $k > 0$  and a stochastic demand  $(k - S_t)$  from noise traders.

<sup>7</sup>This assumption is micro-founded in Appendix G.3. We allow fund investors to choose funds to which to allocate capital, and show that they flow in and out of funds until each fund has accommodated a mass  $\alpha/\beta$  of investors. Moreover, Appendix G.3 shows that fund managers are “perfectly competitive” in attracting investors in the sense that each manager takes the mass of fund investors choosing his fund as given, as he cannot strategically influence the investors' choices of funds through his actions.

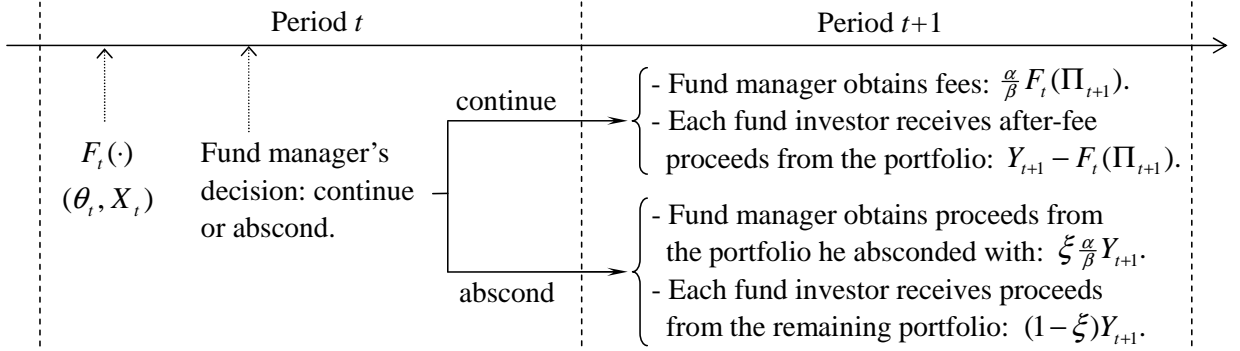


Figure 1: Timeline of portfolio management (generation- $t$  fund).

investor capital,  $(X_t - P_t \theta_t X_t)$  dollars, is invested in the riskless asset.<sup>8,9</sup> Since the fund's portfolio holding (given  $P_t$ ) is described completely by  $\theta_t$  and  $X_t$ , we refer to the pair  $(\theta_t, X_t)$  as the fund's *investment profile*.<sup>10</sup> In period  $t + 1$ , the fund receives dividend  $\delta_{t+1}$  and then sells the stock in the market (to the agents of generation  $t + 1$ ) at price  $P_{t+1}$ . Thus, the total proceeds from the fund portfolio (per investor) is

$$Y_{t+1} \equiv \Pi_{t+1} + (1 + r)X_t \quad (2.2)$$

dollars, where

$$\Pi_{t+1} \equiv R_{t+1} \theta_t X_t \quad (2.3)$$

is the proceeds in excess of the riskless component  $(1 + r)X_t$  that each fund investor could have obtained on her own without delegation.<sup>11</sup> We refer to  $\Pi_{t+1}$  as the *fund profit*.

<sup>8</sup>If  $1 > P_t \theta_t$ , the manager buys the riskless asset, while if  $1 < P_t \theta_t$ , he sells the asset (i.e., borrows funds) to finance the purchase of the stock.

<sup>9</sup>In reality, there are a number of passively managed funds, such as index funds, in which the managers do not have discretion over portfolio choice. In such funds, the managers commit to choose  $\theta_t = 1/P_t$  and do not buy the riskless asset. In Appendix G.4, we show that these passive funds would coexist with our model's funds, even if each fund investor is allowed to choose a fund to which to allocate capital.

<sup>10</sup>We implicitly assume that the managers are not allowed to invest their own capital (borrowed at rate  $r$ ) in their funds. However, as shown in Appendix G.5, even if they were allowed to do so, they would choose not to invest any capital. The reason is as follows. Suppose that a manager can invest  $Z_t \geq 0$  dollars in his fund. Then, given  $X_t$ , he has two choice variables,  $\theta_t$  and  $Z_t$ , to control the fund's stock holding,  $\theta_t(X_t + Z_t)$ . From the manager's perspective, choosing  $Z_t > 0$  entails a borrowing cost, whereas choosing  $\theta_t > 0$  is costless. Thus, he optimally sets  $Z_t = 0$ , and uses only  $\theta_t$  to control the stock holding.

<sup>11</sup> $Y_{t+1}$  is computed as follows. In period  $t$ , each fund investor invests  $X_t$  dollars in the fund. The manager buys  $\theta_t X_t$  shares of the stock (i.e., pays  $P_t \theta_t X_t$  dollars for the purchase) and invests  $(X_t - P_t \theta_t X_t)$  dollars in the riskless asset. In period  $t + 1$ , per investor, the fund receives  $\delta_{t+1} \theta_t X_t$  dollars of payoff from the stock,  $P_{t+1} \theta_t X_t$  dollars from selling that stock in the market, and  $(1 + r)(X_t - P_t \theta_t X_t)$  dollars from the riskless asset. Thus, the total proceeds from the fund portfolio per investor is  $Y_{t+1} = \delta_{t+1} \theta_t X_t + P_{t+1} \theta_t X_t + (1 + r)(X_t - P_t \theta_t X_t) = R_{t+1} \theta_t X_t + (1 + r)X_t = \Pi_{t+1} + (1 + r)X_t$ .

## 2.4 Fee contract

In period  $t$ , in each fund, the generation- $t$  fund investors collectively offer a fee contract to the manager in a take-it-or-leave-it manner. The contract specifies the fee payment schedule contingent on the fund profit  $\Pi_{t+1}$ , denoted by  $F_t(\Pi_{t+1})$ , where  $F_t(\cdot)$  can be any function. The agents determine  $F_t(\cdot)$  and  $(\theta_t, X_t)$  simultaneously (Figure 1).<sup>12</sup>

The assumption that the fee depends only on  $\Pi_{t+1}$  is stronger than necessary. In Appendix G.6, we show that all the results, including the fee schedule itself, would be identical to those of this model even if we allow for a larger contract space that includes not only  $\Pi_{t+1}$  but also  $R_{t+1}$  and  $Y_{t+1}$ .<sup>13</sup>

We model an agency friction between the fund manager and investors as follows. In period  $t$ , after  $F_t(\cdot)$  is determined and the fund portfolio is formed, the manager can choose whether to continue running the fund or abscond with a part of the fund portfolio (Figure 1). If he chooses to continue, he will receive fee  $F_t(\Pi_{t+1})$  from each of his fund's investors in period  $t+1$ . Since there is a mass  $\alpha/\beta$  of investors in each fund, his terminal wealth in this case is

$$W_{t+1}^m = \frac{\alpha}{\beta} F_t(\Pi_{t+1}). \quad (2.4)$$

Each fund investor's terminal wealth in this case is

$$W_{t+1}^f = Y_{t+1} - F_t(\Pi_{t+1}) + (1+r)(W - X_t), \quad (2.5)$$

where the first two terms on the right-hand side (RHS) are the after-fee proceeds from the fund, and the third term is the proceeds from the riskless asset she bought on her own. If the manager chooses to abscond, he can do so with a fraction  $\xi \in (0, 1)$  of the fund portfolio, where  $\xi$  measures the degree of agency friction. His terminal wealth in

---

<sup>12</sup>The timing of the choices of  $F_t(\cdot)$  and  $(\theta_t, X_t)$  is not important. When solving for the equilibrium in Section 4, we will first determine the optimal  $F_t(\Pi_{t+1})$  for an arbitrary  $(\theta_t, X_t)$ ; then, taking this  $F_t(\Pi_{t+1})$  as given, determine the optimal  $(\theta_t, X_t)$ . Hence, our solutions coincide with those of an alternative setting in which each period  $t$  has the following two stages. In the first stage, the investors determine a fee schedule  $F_t(\Pi_{t+1}, \theta_t, X_t)$ , which is contingent on not only  $\Pi_{t+1}$  but also on  $(\theta_t, X_t)$  that will be determined in the second stage. The manager decides whether to accept the schedule. If he accepts it, the agents proceed to the second stage and determine  $(\theta_t, X_t)$ .

<sup>13</sup>The intuition is simple: since  $R_{t+1}$ ,  $\Pi_{t+1}$ , and  $Y_{t+1}$  are perfectly correlated, assigning different weights on these variables in a fee arrangement would not affect risk sharing. As we shall see, an optimal fee arrangement has two properties. The first is the optimal risk sharing between the fund investors and the manager. The second is that the investors, who have full bargaining power, extract maximum surplus. Regardless of which of the three variables ( $R_{t+1}, \Pi_{t+1}, Y_{t+1}$ ) we allow the fee to depend on, as long as it satisfies these two properties, it would result in exactly the same form of contract, and consequently all the other results would be also identical.



this case equals the proceeds from the fund portfolio that he absconded with, i.e.,

$$\widetilde{W}_{t+1}^m = \xi \frac{\alpha}{\beta} Y_{t+1}. \quad (2.6)$$

Each fund investor's terminal wealth in this case is the sum of the proceeds from the remaining fund portfolio and the riskless asset:  $\widetilde{W}_{t+1}^f = (1 - \xi)Y_{t+1} + (1 + r)(W - X_t)$ .<sup>14</sup>

## 2.5 Optimization problems

Each generation- $t$  fund investor chooses capital allocation  $X_t \geq 0$  to the fund and investment in the riskless asset to maximize the expected utility derived from her terminal wealth (given that the manager will not abscond), (2.5). When offering  $F_t(\cdot)$  to the manager, the fund investors make sure that it maximizes the expected utility of each of them and that the manager is willing to continue running the fund rather than abscond, that is, it satisfies the manager's incentive compatibility constraint:

$$\mathbb{E}_t[u^m(W_{t+1}^m)] \geq \mathbb{E}_t[u^m(\widetilde{W}_{t+1}^m)]. \quad (2.7)$$

Since the manager has zero initial wealth, his participation constraint is  $\mathbb{E}_t[u^m(W_{t+1}^m)] \geq \mathbb{E}_t[u^m(0)]$ . In the following, we omit this constraint because it is implied by (2.7).

Each generation- $t$  fund manager determines the fund's risk exposure  $\theta_t \in (-\infty, \infty)$  to maximize the expected utility from his terminal wealth (given that he will not abscond), (2.4). He also decides whether to continue running the fund or abscond, though the fee schedule in equilibrium ensures that he will choose to continue.

Each generation- $t$  direct investor purchases  $K_t \in (-\infty, \infty)$  shares of the stock and invests the rest of her endowment,  $(W - P_t K_t)$  dollars, in the riskless asset to maximize the expected utility from her terminal wealth,  $W_{t+1}^d = R_{t+1} K_t + (1 + r)W$ , which is the sum of the proceeds from the stock and those from the riskless asset.

## 2.6 Definition of equilibrium

The equilibrium consists of stock price  $P_t$ , fund's fee schedule  $F_t(\cdot)$ , fund's risk exposure  $\theta_t$ , fund investor's capital allocation  $X_t$ , and direct investor's stock holding  $K_t$  such that

---

<sup>14</sup>It is not important that the investor can collect  $(1 - \xi)Y_{t+1}$  from the fund if the manager absconds. For example, we would obtain exactly the same results even if we assume that no fund assets are left for the investor if the manager absconds (i.e.,  $\widetilde{W}_{t+1}^f = (1 + r)(W - X_t)$ ). The reason is that what matters for the contracting problem is only the manager's outside option of absconding, (2.6).

the following statements hold for all  $t$ .

1. Given  $P_t$  and  $\theta_t$ ,  $X_t$  and  $F_t(\cdot)$  maximize each fund investor's expected utility.
2. Given  $P_t$ ,  $F_t(\cdot)$ , and  $X_t$ ,  $\theta_t$  maximizes each fund manager's expected utility.
3. Given  $P_t$ ,  $K_t$  maximizes each direct investor's expected utility.
4. The stock market clears.

### 3 Benchmark: Equilibrium without Delegation

Before analyzing the equilibrium with portfolio delegation, let us characterize a benchmark case without delegation, in which fund investors access the stock on their own and fund managers play no roles. This benchmark case is identical to Spiegel (1998) except that there are two types of investors (fund investors and direct investors) who have different degrees of risk aversion. We look for a linear equilibrium in which  $P_t$  is linear in  $\hat{\delta}$  and  $S_t$ .

#### 3.1 Price conjecture

We conjecture and later verify that there exists a constant  $\eta > 0$  such that, for all  $t$ , the equilibrium stock price is

$$P_t = \frac{\hat{\delta} - \eta S_t}{r}. \quad (3.1)$$

Conjecture (3.1) states that the price is the present value of expected future dividends minus a risk premium that is proportional to  $S_t$ , discounted at the riskless rate. This conjecture implies that, conditional on  $t$ ,  $R_{t+1}$  is normally distributed with mean

$$\hat{R}_t \equiv E_t[R_{t+1}] = \eta S_t \quad (3.2)$$

and variance  $\sigma_R^2 \equiv \text{Var}_t[R_{t+1}] = \sigma^2 + \eta^2 \sigma_s^2 / r^2$ , which is the sum of dividend volatility and price volatility. Note that (3.1) and (3.2) imply  $P_t = (\hat{\delta} - \hat{R}_t) / r$ .

#### 3.2 Optimization

Let  $K_t^f$  and  $K_t^d$  denote the numbers of shares of the stock held by each fund investor and direct investor, respectively. Their optimal stock holdings given  $\hat{R}_t$  have usual mean-

variance forms: for  $i \in \{f, d\}$ ,<sup>15</sup>

$$K^i(\hat{R}_t) = \frac{\hat{R}_t}{\gamma^i \sigma_R^2}. \quad (3.3)$$

### 3.3 Equilibrium

Plugging (3.3) into the stock's market clearing condition,  $\alpha K^f(\hat{R}_t) + (1 - \alpha)K^d(\hat{R}_t) = S_t$ , we obtain the equilibrium expected excess stock return:

$$\hat{R}_t = \Lambda(\eta)S_t, \quad \text{where} \quad \Lambda(\eta) \equiv \frac{\sigma_R^2}{\frac{\alpha}{\gamma^f} + \frac{1-\alpha}{\gamma^d}}. \quad (3.4)$$

Reflecting the risk premium,  $\hat{R}_t$  increases with  $\sigma_R^2$  and decreases with the total risk tolerance of the economy, measured by  $\frac{\alpha}{\gamma^f} + \frac{1-\alpha}{\gamma^d}$  (which is the sum of the total risk tolerance of fund managers  $\frac{\alpha}{\gamma^f}$  and that of direct investors  $\frac{1-\alpha}{\gamma^d}$ ). Note that  $\Lambda$  is a function of  $\eta$  since  $\sigma_R^2 \equiv \sigma^2 + \eta^2 \sigma_s^2 / r^2$  depends on  $\eta$ . Recall that we have another expression for  $\hat{R}_t$  in (3.2), which is implied by conjecture (3.1). Thus, conjecture (3.1) is correct if and only if (3.2) and (3.4) are consistent with each other. This requires

$$\eta = \Lambda(\eta). \quad (3.5)$$

That is, the equilibrium values of  $\eta$  are determined at the fixed points of  $\Lambda(\cdot)$ . Since  $\sigma_R^2$  is quadratic in  $\eta$ ,  $\Lambda$  is an increasing quadratic function of  $\eta$ . Moreover, we have  $\Lambda(0) > 0$ . So, if  $\sigma_s$  is positive and small enough, (3.5) has two positive roots of  $\eta$ :

$$\eta = \frac{r^2}{2\sigma_s^2} \left( \left( \frac{\alpha}{\gamma^f} + \frac{1-\alpha}{\gamma^d} \right) \pm \sqrt{\left( \frac{\alpha}{\gamma^f} + \frac{1-\alpha}{\gamma^d} \right)^2 - \frac{4\sigma^2 \sigma_s^2}{r^2}} \right). \quad (3.6)$$

It is well known that there are potentially multiple equilibria in overlapping generations models with infinite time horizon and random stock supply (Spiegel 1998; Watanabe 2008; Banerjee 2011).<sup>16</sup> Which of these two equilibria should one focus on? As noted by

<sup>15</sup>The derivation is straightforward. The mean and variance of each investor's terminal wealth are, for  $i \in \{f, d\}$ ,  $E_t[W_{t+1}^i] = \hat{R}_t K_t^i + (1+r)W$  and  $\text{Var}_t[W_{t+1}^i] = K_t^{i2} \sigma_R^2$ . Due to exponential utility, maximizing  $E_t[u^i(W_{t+1}^i)]$  amounts to maximizing  $E_t[W_{t+1}^i] - \frac{\gamma^i}{2} \text{Var}_t[W_{t+1}^i] = \hat{R}_t K_t^i + (1+r)W - \frac{\gamma^i}{2} K_t^{i2} \sigma_R^2$ . The first-order condition for  $K_t^i$  is  $\hat{R}_t - \gamma^i \sigma_R^2 K_t^i = 0$ , which yields (3.3).

<sup>16</sup>Each of the two values of  $\eta$  corresponds to a possible self-fulfilling belief of the agents. If they believe that  $\eta$  is large and thus the price volatility  $\text{Var}_t[P_{t+1}] = \eta^2 \sigma_s^2 / r^2$  is high, each of them tries to balance her risk bearing by submitting a demand schedule with low price elasticity. The resulting aggregate demand with low price elasticity is indeed settled by a market-clearing price with a large  $\text{Var}_t[P_{t+1}]$  (i.e., large  $\eta$ ).

Banerjee (2011), “[t]he choice of which type of equilibrium to study is somewhat arbitrary, as both types have desirable theoretical and empirical properties.” In the following, we focus on the smaller root of  $\eta$  because only the equilibrium with the smaller  $\eta$  is *stable* in the sense of the following definition, akin to that of Cespa and Vives (2015).

**Definition 1.** *An equilibrium is stable (unstable) if its corresponding value of  $\eta$  is a stable (unstable) fixed point of  $\Lambda(\cdot)$ , that is, if  $\eta$  satisfies  $|\Lambda'(\eta)| < 1$  ( $> 1$ ).*

Intuitively, a stable equilibrium is such that after a small perturbation to the agents’ belief about the value of  $\eta$ , the economy converges back to the original equilibrium through the best-reply dynamics of the agents’ beliefs and the market-clearing prices.<sup>17</sup>

The following proposition summarizes the stable equilibrium outcome. For notational clarity, we append an asterisk to the equilibrium values of the no-delegation case.

**Proposition 3.1** (No-delegation benchmark). *If fund investors could invest in the stock directly, the excess stock return  $R_{t+1}$  is, conditional on  $t$ , normally distributed with mean  $\hat{R}_t^* = \eta^* S_t$  and variance  $\sigma_R^{*2} = \sigma^2 + \eta^{*2} \sigma_s^2 / r^2$ , where  $\eta^*$  is the smaller value of (3.6); the stock price is  $P_t^* = (\hat{\delta} - \hat{R}_t^*) / r$ .*

Note that  $R_{t+1}$  in Proposition 3.1 is the *share* return (i.e., the excess return per share of the stock held) instead of the *dollar* return (i.e., the excess return per dollar invested). The next corollary derives the distribution of the dollar return on the stock market, which will be used in Section 5 where we discuss empirical implications.

**Corollary 3.1** (Market return). *Let  $\Omega_{t+1}^{market} \equiv R_{t+1} / P_t$  be the market return. Conditional on  $t$ ,  $\Omega_{t+1}^{market}$  is normal with mean  $\mu_t^{market*} = \hat{R}_t^* / P_t^*$  and standard deviation  $\sigma_t^{market*} = \sigma_R^* / P_t^*$ . The Sharpe ratio of  $\Omega_{t+1}^{market}$  is  $SR_t^{market*} = \mu_t^{market*} / \sigma_t^{market*} = (\eta^* / \sigma_R^*) S_t$ .*

## 4 Equilibrium with Delegation

This section characterizes the equilibrium with portfolio delegation. As in Section 3, we look for a linear stable equilibrium. We follow the following steps to solve the model.

---

<sup>17</sup>Moreover, the smaller- $\eta$  equilibrium is more general in the sense that it is well-defined when the stock supply is constant (i.e., when  $\sigma_s^2 = 0$ ), whereas the larger- $\eta$  equilibrium is not. As noted by Spiegel (1998), in the larger- $\eta$  equilibrium, we have  $\eta \rightarrow \infty$  (and thus  $\text{Var}_t[P_{t+1}] \rightarrow \infty$ ) when  $\sigma_s^2 \rightarrow 0$ ; if  $\sigma_s^2 = 0$ , then  $\eta$  explodes and the equilibrium is not defined. Also, as shown by Albagli (2015), the smaller- $\eta$  equilibrium is more robust in that it can be supported even when the agents’ time horizon is (infinitely) long, whereas the larger- $\eta$  equilibrium vanishes when the horizon is long enough.

1. We conjecture that the equilibrium stock price  $P_t$  is of the form (3.1), and thus  $\hat{R}_t$  is given by (3.2), but the value of  $\eta$  is different from (3.6).
2. For arbitrary investment profile  $(\theta_t, X_t)$  and  $\hat{R}_t$ , we obtain fee schedule  $F_t(\cdot)$  in the contracting problem (Section 4.1).
3. Using  $F_t(\cdot)$  of step 2, we derive each fund investor's best-response function  $X(\theta_t; \hat{R}_t)$ , i.e., her optimal choice of  $X_t$  for arbitrary  $\theta_t$  and  $\hat{R}_t$  (Section 4.2).
4. Using  $F_t(\cdot)$  of step 2, we derive each manager's best-response function  $\theta(X_t; \hat{R}_t)$ , i.e., his optimal choice of  $\theta_t$  for arbitrary  $X_t$  and  $\hat{R}_t$  (Section 4.3).
5. From the two best-response functions obtained in steps 3 and 4, we identify each fund's Nash-equilibrium investment profile given  $\hat{R}_t$ ,  $(\theta(\hat{R}_t), X(\hat{R}_t))$  (Section 4.4).
6. Using  $(\theta(\hat{R}_t), X(\hat{R}_t))$  of step 5, we compute the aggregate demand for the stock, impose market clearing, and pin down the equilibrium level of  $\hat{R}_t$ . We then verify that the resulting  $P_t$  is indeed in the linear form conjectured in step 1 (Section 4.5).

## 4.1 Fee contract

Due to the CARA-normal specification, each fund's optimal fee has a tractable affine form presented in the following lemma.

**Lemma 4.1** (Fee contract). *Each fund's fee is  $F_t(\Pi_{t+1}) = \phi\Pi_{t+1} + Q_t$ , where*

$$\phi \equiv \frac{\beta\gamma^f}{\alpha\gamma^m + \beta\gamma^f} \quad (4.1)$$

and

$$Q_t \equiv (\xi - \phi) E_t[\Pi_{t+1}] + \xi(1+r)X_t - \frac{\gamma^m \alpha}{2\beta} (\xi^2 - \phi^2) \text{Var}_t[\Pi_{t+1}]. \quad (4.2)$$

Lemma 4.1 is akin to the risk-sharing result of Wilson (1968). The optimal fee is the sum of proportional fee  $\phi\Pi_{t+1}$  and fixed fee  $Q_t$ , both functions of  $(\theta_t, X_t)$ . The optimal risk sharing, which maximizes the agents' joint surplus by adjusting the risk-return profiles of their terminal wealths, requires that the agent with relatively high risk tolerance has large exposure to the stock. Indeed, the manager's exposure measured by  $\phi$  is decreasing in  $\gamma^m$ , while each fund investor's exposure,  $1 - \phi = \alpha\gamma^m/(\alpha\gamma^m + \beta\gamma^f)$ , decreases with  $\gamma^f$ . Moreover, if the mass  $\alpha/\beta$  of investors in the fund increases,  $\phi$  decreases to counterbalance

the manager's exposure to the risky fund profit,  $(\alpha/\beta)\Pi_{t+1}$ . The fund investors set  $Q_t$  so that (2.7) binds, since otherwise they could decrease  $Q_t$  to extract more surplus.<sup>18</sup>

## 4.2 Fund investor's optimization: Capital allocation

This section characterizes each fund investor's optimal policy for her choice of capital allocation  $X_t$ . We make the following parametric assumption.

**Assumption 1:**  $\phi > \xi^2$ .

Assumption 1 ensures the investor's optimal choice of  $X_t$  to be finite by excluding an implausible situation that she increases  $X_t$  indefinitely to decrease the payment of fixed fee  $Q_t$ .<sup>19</sup> Then her optimal policy is given by the following lemma.

**Lemma 4.2** (Fund investor's best response). *Suppose the fee schedule  $F_t(\cdot)$  is as in Lemma 4.1. Given expected stock return  $\hat{R}_t$ , each fund investor's best response to the manager's choice of risk exposure  $\theta_t$  is allocating*

$$X(\theta_t; \hat{R}_t) = \max \left\{ \frac{(1-\xi)\hat{R}_t\theta_t - \xi(1+r)}{\gamma^m \frac{\alpha}{\beta} (\phi - \xi^2) \sigma_R^2 \theta_t^2}, 0 \right\} \quad (4.3)$$

dollars of capital to the fund.

To see the intuition for Lemma 4.2, recall that  $F_t(\cdot)$  in Lemma 4.1 is such that the

---

<sup>18</sup>One may wonder that if  $\xi$  is very small then the fund investors might be better off by letting the manager abscond and thus not paying fees to him. Such an action is not optimal. First, if the fund investors pay the fee to the manager, the certainty equivalent of each investor's terminal wealth is  $U_t^f \equiv E_t[W_{t+1}^f] - \frac{\gamma^f}{2} \text{Var}_t[W_{t+1}^f] = (1-\xi)\hat{R}_t\theta_t X_t - \xi(1+r)X_t + (1+r)W - \frac{\gamma^m}{2} \frac{\alpha}{\beta} (\phi - \xi^2) \theta_t^2 X_t^2 \sigma_R^2$ . Second, if the fund investors let the manager abscond, the certainty equivalent would be  $\tilde{U}_t^f \equiv E_t[\tilde{W}_{t+1}^f] - \frac{\gamma^f}{2} \text{Var}_t[\tilde{W}_{t+1}^f] = (1-\xi)\hat{R}_t\theta_t X_t - \xi(1+r)X_t + (1+r)W - \frac{\gamma^f}{2} (1-\xi)^2 \theta_t^2 X_t^2 \sigma_R^2$ . So,  $U_t^f \geq \tilde{U}_t^f \iff (\xi - \phi)^2 \geq 0$ , which means that  $U_t^f > \tilde{U}_t^f$  for any parameter values. The reason is that risk sharing in the second case is suboptimal, as the manager and each investor split  $\Pi_{t+1}$  in the proportion of  $\xi$  and  $1-\xi$ , rather than  $\phi$  and  $1-\phi$ . If  $\xi$  is very small, the investor's risk bearing in the second case is too high: she could achieve a better return-risk profile by reducing her risk exposure from  $(1-\xi)\Pi_{t+1}$  to  $(1-\phi)\Pi_{t+1}$ .

<sup>19</sup>If Assumption 1 is violated, it would be optimal for each fund investor to choose  $X_t = \infty$  for the following reason. An increase in  $X_t$  increases  $\text{Var}_t[\Pi_{t+1}]$ , which has two opposing effects on the fund investor's expected utility. On the one hand, it has a negative effect because she is risk averse. But on the other hand, it has a positive effect because it decreases her payment of  $Q_t$  (which is because a rise in  $\text{Var}_t[\Pi_{t+1}]$  increases the volatility of the manager's outside option of absconding, thereby he requires smaller  $Q_t$  to not abscond). If the mass of investors in each fund,  $\alpha/\beta$ , is very large, the second effect dominates the first one, and it would be optimal for the investor to choose an extremely large  $X_t$  to increase  $\text{Var}_t[\Pi_{t+1}]$  as much as possible. Assumption 1 eliminates such an implausible behavior by ensuring that the first effect prevails.

manager's incentive compatibility constraint (2.7) binds; that is, the manager *effectively* faces terminal wealth (2.6) that he would obtain if he absconds. With this remark, (4.3) can be viewed as a CARA-normal mean-variance solution. It states that the investor's choice of  $X_t$  is increasing in the expected *effective* after-fee fund return (the numerator), and is decreasing in the volatility of that return (the denominator).

Note that  $X(\theta_t; \hat{R}_t)$  is inverse U-shaped in  $\theta_t$ : for small  $\theta_t$  it increases with  $\theta_t$ , but for large  $\theta_t$  it decreases with  $\theta_t$  (illustrated in Figure 2 below). It makes intuitive sense that  $X(\theta_t; \hat{R}_t)$  is decreasing in  $\theta_t$  for large  $\theta_t$ . That is, the investor optimally adjusts  $X_t$  to the opposite direction of  $\theta_t$  to balance her exposure to the stock. Then, why does she *increase*  $X_t$  in response to a rise in  $\theta_t$  if  $\theta_t$  is small? The key to this is the term  $\xi(1+r)$  in the numerator of (4.3), which represents the rent that the fund investor leaves to the manager to deter him from absconding. To see this, recall that the manager can abscond with a fraction  $\xi$  of the fund's entire asset. So when delegating  $X_t$  dollars of capital to the manager, the investor effectively grants a fraction  $\xi$  of that capital—i.e.,  $\xi(1+r)X_t$  dollars in period- $(t+1)$  value—as a rent to the manager to prevent him from absconding. The term  $\xi(1+r)$  in the numerator of (4.3) is the marginal value of the rent.<sup>20</sup> Since the rent does not depend on  $\theta_t$ , it effectively serves as a fixed cost of stock investment from the investor's perspective. Thus, if the manager increases  $\theta_t$ , the rent per share of the stock held in the fund decreases, making it cheaper for the investor to get exposure to the stock. This encourages the investor to allocate more capital to the fund, creating the upward-sloping part of  $X(\theta_t; \hat{R}_t)$ . If  $\theta_t$  is very small, the rent per share of the stock is so large that the investor does not allocate any capital to the fund (i.e.,  $X(\theta_t; \hat{R}_t) = 0$ ).

### 4.3 Fund manager's optimization: Fund's risk exposure

This section determines the manager's optimal policy for his choice of fund's risk exposure  $\theta_t$ . There are two routes through which  $\theta_t$  affects his expected utility. First, it affects his personal exposure to the stock via the proportional fee  $\phi\Pi_{t+1}$ , because  $\Pi_{t+1}$  depends on  $\theta_t$ . Second, it affects fixed fee  $Q_t$  through  $E_t[\Pi_{t+1}]$  and  $\text{Var}_t[\Pi_{t+1}]$ , which are both increasing in  $\theta_t$ . The manager takes these effects into account when choosing  $\theta_t$ .

---

<sup>20</sup>The following argument, though not entirely rigorous, helps us grasp the economic intuition behind the numerator of (4.3). From (2.5), each fund investor's terminal wealth is  $W_{t+1}^f = Y_{t+1} - F_{t+1} + (1+r)(W - X_t) = \Pi_{t+1} - F_{t+1} + (1+r)W$ . Since the manager can abscond with  $\xi Y_{t+1}$ , the investor effectively leaves  $\xi Y_{t+1}$  to the manager to deter him from absconding. Thus, if we replace  $F_{t+1}$  with  $\xi Y_{t+1}$ , we have  $W_{t+1}^f = \Pi_{t+1} - \xi Y_{t+1} + (1+r)W = (1-\xi)\Pi_{t+1} - \xi(1+r)X_t + (1+r)W$ . So the marginal expected wealth is  $\partial E_t[W_{t+1}^f]/\partial X_t = (1-\xi)\hat{R}_t\theta_t - \xi(1+r)$ , which is identical to the numerator of (4.3).

**Lemma 4.3** (Fund manager’s best response). *Suppose the fee schedule  $F_t(\cdot)$  is as in Lemma 4.1. Given expected stock return  $\hat{R}_t$ , the manager’s best response to the investor’s choice of capital allocation  $X_t > 0$  is to choose the fund’s risk exposure*

$$\theta(X_t; \hat{R}_t) = \frac{\beta \hat{R}_t}{\gamma^m \alpha \xi \sigma_R^2 X_t}. \quad (4.4)$$

The manager’s best response (4.4) can be viewed as a CARA-normal mean-variance solution, *as if* he is facing the terminal wealth (2.6) that he would obtain if absconding. Not surprisingly, his choice of  $\theta_t$  is increasing in  $\hat{R}_t$  and decreasing in  $\sigma_R^2$  and  $\gamma^m$ . The mass of investors in his fund,  $\alpha/\beta$ , affects  $\theta_t$  negatively for the following reason. A rise in  $\alpha/\beta$  increases the “total proportional fee rate”  $\alpha\gamma^f/(\alpha\gamma^m + \beta\gamma^f)$ , i.e.,  $\phi$  for each investor multiplied by the mass  $\alpha/\beta$  of investors. This increases the manager’s personal exposure to the stock, which he tries to counterbalance by reducing  $\theta_t$ . Note that choosing  $\theta_t$  given  $X_t$  is equivalent to choosing  $\theta_t X_t$ , the number of shares of the stock held for each investor. Hence, the manager optimally decreases  $\theta_t$  in response to an increase in  $X_t$  to balance his personal exposure to the stock.

#### 4.4 Fund’s investment profile

To determine the funds’ demand for the stock, let us identify each fund’s optimal investment profile, i.e., Nash-equilibrium pair  $(\theta(\hat{R}_t), X(\hat{R}_t))$  for given  $\hat{R}_t$ . This pair is determined by solving the two best response functions, (4.3) and (4.4), for  $\theta_t$  and  $X_t$ . As illustrated in Figure 2, the nonmonotonicity of  $X(\theta_t; \hat{R}_t)$  is critical to the existence of a pure-strategy Nash equilibrium. The key to this nonmonotonicity is the rent that each fund investor leaves to the manager, as discussed already in Lemma 4.2.

To ensure that the two best response functions cross, we assume the following.

**Assumption 2:**  $\phi < \xi$ .

Assumption 2 requires that the mass  $\alpha/\beta$  of fund investors in each fund is not too small.<sup>21</sup> If  $\alpha/\beta$  increases,  $\theta(X_t; \hat{R}_t)$  shifts downward because the manager tries to rebalance his personal risk exposure in response to the increased fund size, but  $X(\theta_t; \hat{R}_t)$  shifts upward because the associated lower  $Q_t$  leads the investor to increase  $X_t$ .<sup>22</sup> Thus, the two best

<sup>21</sup>Assumptions 1 and 2 are compatible with each other. For example, the parameter values used in Figure 2 ( $\alpha = 0.4$ ,  $\beta = 0.001$ ,  $\gamma^f = 1$ ,  $\gamma^m = 1$ , and  $\xi = 0.01$ ) satisfy both of them.

<sup>22</sup> $Q_t$  decreases with  $\alpha/\beta$  for the following reason. A higher  $\alpha/\beta$  makes  $\widetilde{W}_{t+1}^m$  more volatile. With this



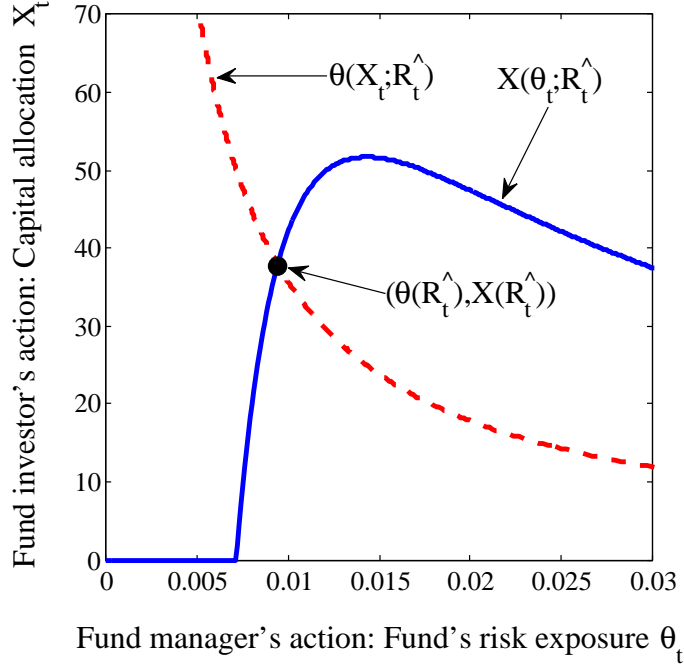


Figure 2: Determination of the fund's Nash-equilibrium investment profile given  $\hat{R}_t$ . The parameter values used in the figure are  $r = 0.01$ ,  $\gamma^f = 1$ ,  $\gamma^m = 1$ ,  $\xi = 0.01$ ,  $\alpha = 0.4$ ,  $\beta = 0.001$ , and  $\sigma_R^2 = 1$ . We set  $\hat{R}_t = 1.4286$ , which is the equilibrium level of  $\hat{R}_t$  determined later with  $S_t = 1$ .

response functions cross if  $\alpha/\beta$  is large enough. Under Assumptions 1 and 2, (4.3) and (4.4) pin down each fund's investment profile given  $\hat{R}_t$  as follows.

**Lemma 4.4** (Fund's investment profile). *Given expected stock return  $\hat{R}_t$ , there is a unique pure-strategy Nash-equilibrium investment profile  $(\theta(\hat{R}_t), X(\hat{R}_t))$  for each fund, where*

$$\theta(\hat{R}_t) = \left( \frac{(1+r)\xi^2}{\xi - \phi} \right) \frac{1}{\hat{R}_t} \quad (4.5)$$

and

$$X(\hat{R}_t) = \left( \frac{\beta(\xi - \phi)}{(1+r)\gamma^m\alpha\xi^3\sigma_R^2} \right) \hat{R}_t^2. \quad (4.6)$$

Overall, Lemma 4.4 is in line with Hugonnier and Kaniel (2010), who find (1) a positive relation between a fund's proportional fee rate and its leverage and (2) a negative relation between the fee rate and the investor's holding of the fund. Moreover, Lemma 4.4 implies that the fund's stock holding for each investor,  $\theta(\hat{R}_t)X(\hat{R}_t)$ , is independent of  $\phi$  (given  $\hat{R}_t$ ) because the positive effect of  $\phi$  on  $\theta(\cdot)$  and its negative effect on  $X(\cdot)$  just offset each other. <sup>1</sup> If the outside option is less attractive, the manager is willing to continue running the fund for a lower  $Q_t$ .

other. This feature is also consistent with the result of Hugonnier and Kaniel (2010) that the investor’s effective portfolio weights are independent of the fee rate.

## 4.5 Equilibrium

Each direct investor’s optimal policy  $K(\hat{R}_t)$  is the same as in the benchmark case, (3.3). Using (3.3), (4.5), and (4.6), the period- $t$  aggregate demand for the stock is

$$\underbrace{\alpha\theta(\hat{R}_t)X(\hat{R}_t)}_{\text{funds' demand}} + \underbrace{(1-\alpha)K(\hat{R}_t)}_{\text{direct investors' demand}} = \left(\frac{\beta}{\xi\gamma^m} + \frac{1-\alpha}{\gamma^d}\right) \frac{\hat{R}_t}{\sigma_R^2}. \quad (4.7)$$

The aggregate demand (4.7) is a decreasing affine function of  $P_t$  because  $\hat{R}_t$  is a decreasing affine function of  $P_t$ . Equating (4.7) to the aggregate supply  $S_t$  determines the equilibrium expected excess stock return:

$$\hat{R}_t = \Lambda(\eta)S_t, \quad \text{where} \quad \Lambda(\eta) \equiv \frac{\sigma_R^2}{\frac{\beta}{\xi\gamma^m} + \frac{1-\alpha}{\gamma^d}}. \quad (4.8)$$

Unlike its counterpart in the no-delegation case, (3.4),  $\hat{R}_t$  in (4.8) decreases with the effective total risk tolerance of fund managers, represented by  $\frac{\beta}{\xi\gamma^m}$ , but does not reflect the fund investors’ total risk tolerance  $\frac{\alpha}{\gamma^f}$  because the managers effectively “undo” these investors’ impact on the price by adjusting  $\theta_t$  given  $X_t$ .

The consistency between (3.2) and (4.8) requires that  $\eta$  solves the quadratic equation  $\eta = \Lambda(\eta)$  that has two positive roots:

$$\eta = \frac{r^2}{2\sigma_s^2} \left( \left( \frac{\beta}{\xi\gamma^m} + \frac{1-\alpha}{\gamma^d} \right) \pm \sqrt{\left( \frac{\beta}{\xi\gamma^m} + \frac{1-\alpha}{\gamma^d} \right)^2 - \frac{4\sigma^2\sigma_s^2}{r^2}} \right), \quad (4.9)$$

where only the smaller root is relevant for the stable equilibrium.

Viewing  $P_t$  and  $\hat{R}_t$  as functions of  $\eta$ , their functional forms are identical in the no-delegation case (Section 3) and the delegation case (Section 4). Thus, we have the following proposition on the stock price and return in the stable equilibrium. We append two asterisks to the equilibrium values of the delegation case.

**Proposition 4.1** (Stock market). *The stock return and price characterizations are similar to those without delegation (Proposition 3.1 and Corollary 3.1) with  $\eta^*$  being replaced by  $\eta^{**}$ , where  $\eta^{**}$  is the smaller value of (4.9).*

Proposition 4.1 states that, in terms of the stock price and return, the difference between the no-delegation case and the delegation case stems directly from the difference between  $\eta^*$  and  $\eta^{**}$ . More specifically, the difference is that the total risk tolerance of fund investors  $\frac{\alpha}{\gamma^f}$  in (3.6) is replaced by the effective total risk tolerance of fund managers  $\frac{\beta}{\xi\gamma^m}$  in (4.9). This implies that the presence of fund managers represents a shift in the aggregate risk bearing capacity. Thus, asset price dynamics with delegation are the same as they would be in an economy with two agents with different degrees of risk aversion, for appropriately selected levels of CARA coefficients.<sup>23</sup>

The next proposition presents the variables related to portfolio delegation in the stable equilibrium, whose counterparts do not exist in the no-delegation case.

**Proposition 4.2** (Portfolio delegation). *In the stable equilibrium,*

1. *each fund's fee schedule is*

$$F_t(\Pi_{t+1}) = \phi\Pi_{t+1} + \left( \frac{\beta(\xi - \phi)(\xi - \phi + 2)\sigma_R^{**2}}{2\alpha\xi^2\gamma^m\left(\frac{\beta}{\xi\gamma^m} + \frac{1-\alpha}{\gamma^d}\right)^2} \right) S_t^2, \quad (4.10)$$

where  $\Pi_{t+1} \equiv R_{t+1}\theta_t X_t$ ,  $\phi \equiv \beta\gamma^f/(\alpha\gamma^m + \beta\gamma^f)$ , and  $\sigma_R^{**2} = \sigma^2 + \eta^{**2}\sigma_s^2/r^2$ ;

2. *each fund's size (i.e., assets under management) is*

$$\frac{\alpha}{\beta} X_t^{**} = \left( \frac{(\xi - \phi)\sigma_R^{**2}}{\gamma^m(1+r)\xi^3\left(\frac{\beta}{\xi\gamma^m} + \frac{1-\alpha}{\gamma^d}\right)^2} \right) S_t^2; \quad (4.11)$$

3. *each fund buys  $(\alpha/\beta)\theta_t^{**} X_t^{**}$  shares of the stock, where*

$$\theta_t^{**} = \left( \frac{(1+r)\xi^2\left(\frac{\beta}{\xi\gamma^m} + \frac{1-\alpha}{\gamma^d}\right)}{(\xi - \phi)\sigma_R^{**2}} \right) \frac{1}{S_t}. \quad (4.12)$$

The equilibrium with delegation involves a feedback loop among stock price, fund's investment profile, and fee schedule. In each fund, the agents take into account the fee schedule when choosing the capital allocation and the fund's risk exposure, while the fee schedule depends on these agents' choices, and the stock price adjusts so that the agents' choices are consistent with market clearing. Since  $S_t$  is stochastic,  $P_t$  is stochastically time varying. Accordingly, the agents' optimal choices of  $F_t(\cdot)$ ,  $X_t$ ,  $\theta_t$ , and  $K_t$  are also

---

<sup>23</sup>I thank an anonymous referee for suggesting these interpretations of Proposition 4.1.

time varying with  $S_t$ .<sup>24</sup> Detailed analysis of this equilibrium is deferred to Section 5, in which we discuss the model’s empirical implications.

It is worth noting that it would be optimal for each fund investor to allocate capital to only one fund even if she was allowed to access multiple funds. That is, the equilibrium of this section is also an equilibrium for the case in which investors can allocate capital to multiple funds. The reason is simple: since there is no diversification benefit from investing in multiple funds and yet the fee schedule (4.10) involves a fixed fee, it is optimal for each investor to access only one fund to avoid duplication of fixed fees.

Note that  $S_t$  (and therefore  $\hat{R}_t$ ) can go negative because  $\epsilon_t$  is normally distributed (for example, noise traders may place a very large buy order). The model’s solution makes intuitive sense even in such a situation. The fund size (4.11) is proportional to  $S_t^2$ , meaning that fund investors provide a positive amount of capital irrespective of whether  $S_t$  is positive or negative. This is because fund managers guarantee positive expected returns on the investor capital by buying the stock in “normal times” (i.e.,  $\theta_t^{**} > 0$  when  $S_t > 0$ ) and shorting it in the reverse condition (i.e.,  $\theta_t^{**} < 0$  when  $S_t < 0$ ). The next corollary makes this point clearer. It derives the distribution of the dollar return on each fund, which we refer to as the *fund return*, and shows that the expected fund return is positive regardless of  $S_t$ . To our knowledge, this is the first paper to study the distributions of both market return and fund return explicitly in an asset-pricing model.

**Corollary 4.1** (Fund return). *Let  $\Omega_{t+1}^{fund} \equiv (\Pi_{t+1} - F_t(\Pi_{t+1}))/X_t$  be the fund return, i.e., the after-fee excess rate of return on each fund. Conditional on  $t$ ,*

1.  $\Omega_{t+1}^{fund}$  is normal with mean

$$\mu^{fund**} = \frac{(1+r)\xi\left(\phi - \frac{\xi^2 + \phi^2}{2}\right)}{\xi - \phi} \quad (4.13)$$

and standard deviation

$$\sigma_t^{fund**} = \left( \frac{(1+r)\xi^2(1-\phi)\left(\frac{\beta}{\xi\gamma^m} + \frac{1-\alpha}{\gamma^d}\right)}{(\xi - \phi)\sigma_R^{**}} \right) \frac{1}{|S_t|}; \quad (4.14)$$

---

<sup>24</sup>The funds’ aggregate stock holding,  $\alpha\theta_t^{**}X_t^{**} = \left(\frac{\beta}{\xi\gamma^m} / \left(\frac{\beta}{\xi\gamma^m} + \frac{1-\alpha}{\gamma^d}\right)\right)S_t$ , and the direct investors’ aggregate stock holding,  $(1-\alpha)K_t^{**} = \left(\frac{1-\alpha}{\gamma^d} / \left(\frac{\beta}{\xi\gamma^m} + \frac{1-\alpha}{\gamma^d}\right)\right)S_t$ , are also time varying with  $S_t$ .

2. the Sharpe ratio of  $\Omega_{t+1}^{fund}$  is

$$\text{SR}_t^{fund^{**}} \equiv \frac{\mu^{fund^{**}}}{\sigma_t^{fund^{**}}} = \left( \frac{(\phi - \frac{\xi^2 + \phi^2}{2})\sigma_R^{**}}{\xi(1-\phi)(\frac{\beta}{\xi\gamma^m} + \frac{1-\alpha}{\gamma^d})} \right) |S_t|. \quad (4.15)$$

It is readily checked from (4.15) and Assumption 2 that the funds are not able to beat the market in normal times with  $S_t > 0$ , that is,  $\text{SR}_t^{fund^{**}} < \text{SR}_t^{market^{**}}$  if  $S_t > 0$ . This result is not surprising: as long as  $S_t > 0$  and therefore  $\hat{R}_t = \eta S_t > 0$ , no fund investor should achieve a better risk-adjusted return than she would do when holding the market, because the managers access the same market and yet the investors pay fees to these managers. When  $S_t < 0$  and thus  $\hat{R}_t < 0$ , holding the market directly is obviously unprofitable because  $\text{SR}_t^{market} < 0$ . However, even in such times, investing in a fund is profitable for fund investors because the manager shorts the market (i.e.,  $\theta_t < 0$ ) to generate a positive expected return on the investor capital (i.e.,  $\text{SR}_t^{fund} > 0$ ).

## 5 Empirical Implications

This section discusses empirical implications drawn from the equilibrium outcome presented in Section 4. Unless otherwise noted, we focus on the case of  $S_t > 0$ .

### 5.1 Effects of portfolio delegation on stock prices

First, we study the impact of portfolio delegation on the stock's expected returns and price volatility. Viewing  $\mu_t^{market}$ ,  $\text{SR}_t^{market}$ , and  $\text{Var}_t[P_{t+1}]$  as functions of  $\eta$ , their functional forms are identical in the no-delegation case (Proposition 3.1 and Corollary 3.1) and the delegation case (Proposition 4.1); moreover, all of them are monotonically increasing in  $\eta$ . Thus, the equilibrium levels of these three variables in the delegation case are larger (smaller) than those in the no-delegation case if and only if  $\eta^{**} > (<)\eta^*$ .

#### 5.1.1 Expected returns

According to Investment Company Institute (2014), in the U.S., the number of households owning mutual funds has more than doubled in the last two decades, going from 27 million in 1993 to 56.7 million in 2013 (Figure 3, solid line). "As households have come to rely more on funds over the past decade, their demand for directly held equities has fallen" (Investment Company Institute 2014). These facts imply that  $\alpha$  has increased

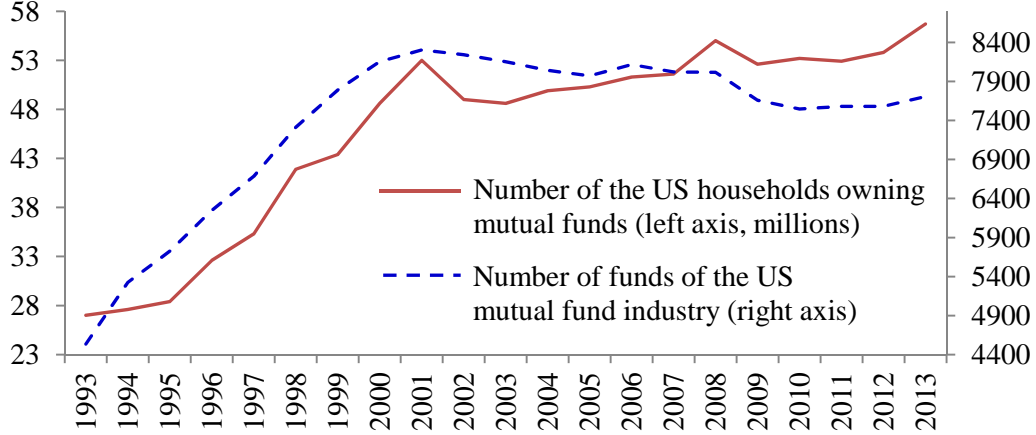


Figure 3: Number of the U.S. households owning mutual funds (solid line; left axis in millions) and the number of funds in the U.S. mutual fund industry (dashed line; right axis). Source: Investment Company Institute (2014).

dramatically in the past two decades. Meanwhile, the number of mutual funds has increased too. Investment Company Institute (2014) reports that there are 7,707 mutual funds in the U.S. in 2013, while there were 4,534 in 1993 (Figure 3, dashed line). This observation indicates that  $\beta$  has increased significantly. Despite these dramatic shifts in the portfolio management industry, however, on the surface there is no clear trend in the overall stock market return over the past two decades (see Figure 4(b) below for S&P 500 returns). The following result reconciles these three observations: (1) a rise in  $\alpha$ , (2) a rise in  $\beta$ , and (3) no trend in the stock market return.

**Result 1.** *The expected market returns, both non risk-adjusted ( $\mu_t^{market**}$ ) and risk-adjusted ( $SR_t^{market**}$ ), are*

1. *positively related to the proportion  $\alpha$  of fund investors and negatively related to the mass  $\beta$  of fund managers, and*
2. *larger or smaller than those of the no-delegation case, depending on the parameter values:  $\mu_t^{market**} > \mu_t^{market*}$  and  $SR_t^{market**} > SR_t^{market*}$  if  $\frac{\alpha}{\beta} > \frac{1}{\xi} \frac{\gamma^f}{\gamma^m}$ , but  $\mu_t^{market**} < \mu_t^{market*}$  and  $SR_t^{market**} < SR_t^{market*}$  if  $\frac{\alpha}{\beta} < \frac{1}{\xi} \frac{\gamma^f}{\gamma^m}$ .*

Statement 1 implies that both  $\alpha$  and  $\beta$  do impact the stock market returns, but those impacts may have offset each other (that is, a rise in  $\alpha$  increased the expected market returns but a rise in  $\beta$  lowered them) and therefore have not appeared clearly on the observed stock returns. The proof of statement 1 is straightforward by algebra, as  $d\eta^{**}/d\alpha > 0$  and  $d\eta^{**}/d\beta < 0$  in the stable equilibrium. Statement 2 indicates that the

stock's expected returns under delegation may exceed those of the no-delegation case if  $\alpha$  is large and  $\beta$  is small enough, i.e., if a large mass of investors allocate capital to a small mass of fund managers.<sup>25</sup> Note that Result 1 does not rely on the assumption that  $S_t$  is time-varying, as it follows from comparative statics of  $\alpha$  and  $\beta$ .

Why is the expected stock return high if  $\alpha/\beta$  is large? To see the intuition, it is useful to consider the following decomposition of the expected stock return,  $\hat{R}_t^{**} S_t$ . It is the sum of four components: (1) the total risk premium demanded by fund managers, (2) the total risk premium demanded by fund investors, (3) the total risk premium demanded by direct investors, and (4) the total rents paid by fund investors.<sup>26</sup> That is,

$$\begin{aligned} \hat{R}_t^{**} S_t = & \underbrace{\beta\gamma^m \text{Var}_t [W_{t+1}^m]}_{\text{managers' total risk premium}} + \underbrace{\alpha\gamma^f \text{Var}_t [W_{t+1}^f]}_{\text{fund investors' total risk premium}} + \underbrace{(1-\alpha)\gamma^d \text{Var}_t [W_{t+1}^d]}_{\text{direct investors' total risk premium}} \\ & + \underbrace{\alpha\xi(1+r)X_t^{**}}_{\text{rents paid by fund investors}}. \end{aligned} \quad (5.1)$$

The total proportional fee rate ultimately relevant for each manager's risk bearing is  $(\alpha/\beta)\phi = \alpha\gamma^f/(\alpha\gamma^m + \beta\gamma^f)$ , which increases with  $\alpha$  and decreases with  $\beta$ . Thus, a rise in  $\alpha$  and/or a decrease in  $\beta$  induces him to demand a larger risk premium, increasing the first term on the RHS of (5.1). Each fund investor receives a fraction  $1-\phi = \alpha\gamma^m/(\alpha\gamma^m + \beta\gamma^f)$  of  $\Pi_{t+1}$ , which also increases with  $\alpha$  and decreases with  $\beta$ . Hence, a rise in  $\alpha$  and/or a decrease in  $\beta$  also leads to a higher risk premium demanded by fund investors, increasing the second term of (5.1). Furthermore, a rise in  $\alpha$  increases the total rents too: a higher  $\alpha$  means that more investors demand extra premia on stock returns to compensate such rents, and thus the fourth term of (5.1) increases.<sup>27</sup> These effects collectively surpass the decrease in the third term of (5.1) caused by a rise in  $\alpha$ , and hence the net effect of  $\alpha$  on the expected stock return is positive.

Both  $\mu_t^{market**}$  and  $SR_t^{market**}$  increase with agency friction  $\xi$ , as  $d\eta^{**}/d\xi > 0$  in the stable equilibrium. Moreover, statement 2 of Result 1 states that these returns surpass the

---

<sup>25</sup>Both  $\frac{\alpha}{\beta} > \frac{1}{\xi} \frac{\gamma^f}{\gamma^m}$  and  $\frac{\alpha}{\beta} < \frac{1}{\xi} \frac{\gamma^f}{\gamma^m}$  can be compatible with Assumptions 1 and 2. For example, suppose  $\alpha = 0.6$ ,  $\beta = 0.2$ , and  $\gamma^f = \gamma^m = 1$ . Then,  $\xi = 0.4$  leads to  $\frac{\alpha}{\beta} > \frac{1}{\xi} \frac{\gamma^f}{\gamma^m}$  whereas  $\xi = 0.3$  leads to  $\frac{\alpha}{\beta} < \frac{1}{\xi} \frac{\gamma^f}{\gamma^m}$ , while both of them satisfy Assumptions 1 and 2.

<sup>26</sup>Each agent's risk premium is defined as her/his risk aversion  $\gamma^i$  multiplied by the conditional variance of her/his terminal wealth. Multiplying it by the mass of agents yields the total risk premium.

<sup>27</sup>Although the rent  $\xi(1+r)X_t^{**}$  is transferred to the manager (through  $Q_t$ ), it does not affect his choice of  $\theta_t$  because it is independent of  $\theta_t$ . Thus, the rent has no impact on the stock price through the managers' actions. This is the reason the rent is not offset in (5.1) in the form of a reduction in the premium demanded by the managers.

no-delegation levels if  $\xi$  is large. These results contrast with Buffa, Vayanos, and Woolley (2014), who find that the aggregate market’s expected return decreases with the agency friction. Their result arises from asymmetric cross-sectional price distortions in multiple stocks. In contrast, our result holds for a single stock, and the intuition is derived from (5.1). As  $\xi$  increases, each fund investor needs to leave a larger rent  $\xi(1+r)X_t^{**}$  to the manager. To compensate it, she demands a higher expected return on the stock market, driving up  $\hat{R}_t^{**}$  by increasing the fourth term on the RHS of (5.1).

Result 1 is in line with He and Krishnamurthy (2012), who show that the financial specialist’s presence in the economy inversely impacts the market risk premium and Sharpe ratio. However, the economic mechanisms are different. In He and Krishnamurthy (2012), the key is the specialist’s equity capital constraint: if his wealth decreases when the constraint is binding (i.e., in the “crisis” situation), the risk premium rises to reflect the capital scarcity. By contrast, our Result 1 holds in non-crisis times, and the key to the result is the risk sharing between fund investors and managers through the fee contract. Our result that  $dSR_t^{market^{**}}/d\beta < 0$  is also consistent with Basak and Pavlova (2012) who find that the Sharpe ratio decreases with the fraction of institutional investors in the economy. In their model, a large demand pressure from institutional investors—who care about their performance relative to an index—requires that, for the stock market to clear, the market becomes less attractive with a lower Sharpe ratio.

### 5.1.2 Price volatility

Ever since Shiller (1981) and LeRoy and Porter (1981) found that “stock prices move too much to be justified by changes in subsequent dividends,” the so-called excess volatility puzzle has been of interest among financial economists. Our model implies that some of the characteristics of portfolio delegation affect stock price volatility positively.

**Result 2.** *Stock price volatility  $\text{Var}_t[P_{t+1}^{**}] = \eta^{**2}\sigma_s^2/r^2$  is*

1. *positively related to the proportion  $\alpha$  of fund investors and negatively related to the mass  $\beta$  of fund managers, and*
2. *larger or smaller than that of the no-delegation case, depending on the parameter values:  $\text{Var}_t[P_{t+1}^{**}] > \text{Var}_t[P_{t+1}^*]$  if  $\frac{\alpha}{\beta} > \frac{1}{\xi} \frac{\gamma^f}{\gamma^m}$ , but  $\text{Var}_t[P_{t+1}^{**}] < \text{Var}_t[P_{t+1}^*]$  if  $\frac{\alpha}{\beta} < \frac{1}{\xi} \frac{\gamma^f}{\gamma^m}$ .*

That is, stock price tends to be highly volatile when a large fraction of investors access the market indirectly through a small number of financial experts. The result stems from



the “risk concentration effect.” As a large amount of investor capital is concentrated in a small mass of fund managers, the fee arrangement (4.10) requires that each manager bears a large market risk *ceteris paribus*. Thus, to balance his personal risk bearing by reducing the variation of stock holding, the manager submits a demand schedule with low price elasticity. In equilibrium, the aggregate demand with low price elasticity is settled by a highly volatile market-clearing price. To see this graphically, consider the classical demand-supply diagram with price on the vertical axis and quantity on the horizontal axis. If  $\alpha$  is high and/or  $\beta$  is low, the slope of the stock’s demand curve is steep because the coefficient of  $P_t$  in (4.7)—which equals  $1/\eta$  in equilibrium—is small. This makes the market-clearing price volatile for a given fluctuation of the vertical supply curve  $S_t$ .

Result 2 shares the key insight of Basak and Pavlova (2012) that “institutional investors make the stock more volatile.” Their economic mechanism for the large volatility is different from ours. They argue that a large demand pressure from institutional investors makes the market less attractive (i.e., more volatile) since otherwise the stock market would not clear. Result 2 is also consistent with Cuoco and Kaniel (2011), who show that stock volatility is higher in the economy with a proportional performance fee than in the benchmark case with costless delegation.

The signs of comparative statics of the market return volatility  $\sigma_t^{market}$  are the same as those of statement 1 of Result 2. But, in the real world, has the market return volatility changed behind the change in the fund management industry? The data say no: Campbell et al. (2001) report that the aggregate market volatility did not change during the 1962-1997 period, in which there was a dramatic shift towards institutional ownership of stocks.<sup>28</sup> A possible explanation that reconciles this empirical fact and Result 2 is that, as discussed in Section 5.1.1,  $\alpha$  has increased while  $\beta$  has also increased over time and thus offset each other’s effects on the market return volatility.

## 5.2 Comovement: Portfolio delegation and stock market

This section studies the dynamics of delegation-related variables such as fees, fund flows, and fund strategies, which are not analyzed in the no-delegation benchmark case. Our goal is to understand how and why these variables comove with the stock market.

Recall that  $S_t$  is the only source of time variation of the endogenous variables presented in Proposition 4.2 (except for  $\Pi_{t+1}$ , which varies also with  $\delta_{t+1}$ ). Thus, unlike Result 1,

---

<sup>28</sup>Although the aggregate market volatility did not change, Campbell et al. (2001) find a positive trend in the idiosyncratic volatility of individual firms. Bennett et al. (2003) and Xu and Malkiel (2003) argue that this increase in the idiosyncratic volatility is due to a rise in institutional ownership.

the results presented in this section are driven by the assumption that  $S_t$  is time-varying.<sup>29</sup>

### 5.2.1 Fund fees and stock market return

Sections 5.1.1 and 5.1.2 shed light on the role of fund fees as a risk-sharing mechanism that connects the extent of portfolio delegation and stock market performance. Then, what is the relation between fund fees *per se* and the stock market return?

**Result 3.** *Conditional on  $t$ , fund's expense ratio  $F_t(\Pi_{t+1})/X_t$  is normal with mean*

$$E_t \left[ \frac{F_t(\Pi_{t+1})}{X_t} \right] = (1+r)\xi \left( \frac{\xi^2 + \phi^2}{2(\xi - \phi)} + 1 \right), \quad (5.2)$$

*which is independent of the market return, and standard deviation*

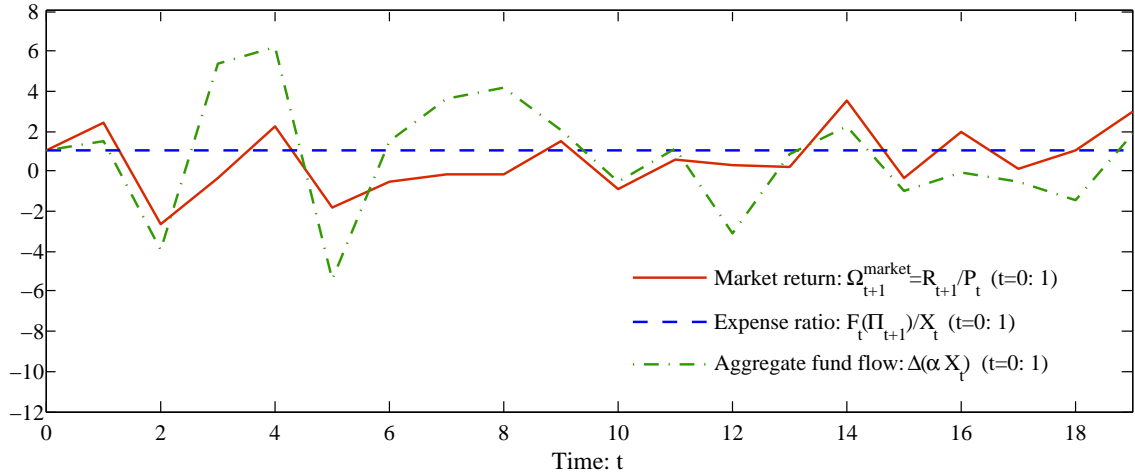
$$SD_t \left[ \frac{F_t(\Pi_{t+1})}{X_t} \right] = \left( \frac{(1+r)\xi^2 \phi \left( \frac{\beta}{\xi \gamma^m} + \frac{1-\alpha}{\gamma^d} \right)}{(\xi - \phi)\sigma_R^*} \right) \frac{1}{|S_t|}. \quad (5.3)$$

That is, regardless of  $\Omega_t^{market}$  that is time varying with  $S_t$ , funds adopt a fixed expected expense ratio every period. Moreover, the ratio's standard deviation (5.3) is much smaller than  $\sigma_t^{market}$  if  $|S_t|$  is large. So the model predicts that the time variation of the realized expense ratio can be much smaller than that of the stock market return and other variables such as fund flows. This is illustrated in panel (a) of Figure 4. It plots simulated paths of  $\Omega_{t+1}^{market} = R_{t+1}/P_t$  (solid line),  $F_t(\Pi_{t+1})/X_t$  (dashed line), and the aggregate fund flow  $\Delta(\alpha X_t)$  (dash-dotted line), with their values for  $t = 0$  being normalized to 1. The simulation appears to be in line with the real-world mutual fund industry and stock market. As shown in panel (b) of Figure 4, the average expense ratio in the U.S. mutual fund industry is quite stable compared to the stock market return and the aggregate fund flow (per fund) that exhibit large volatility on a similar scale as our model's simulation.<sup>30,31</sup>

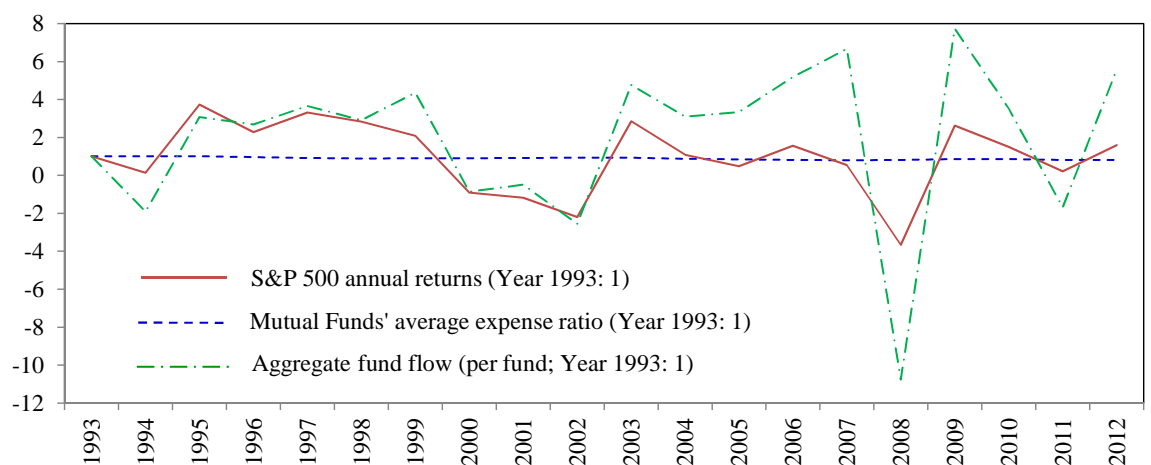
<sup>29</sup>All these results do not rely on the assumption that  $S_t$  is a random walk. As discussed in Appendix G.2, they would also go through under other assumptions that  $S_t$  is non-persistent or mean-reverting. Here, a caveat is that our attention is limited to AR(1) processes of  $S_t$ . We have not been able to check whether our results are robust to non-AR(1) specifications of  $S_t$  because we have not yet solved for equilibria under such processes. For example, if we assume  $S_t$  to be AR( $q$ ) with  $q \geq 2$ , the equilibrium price function would be different from (3.1) because  $E_t[S_{t+1}]$  would depend not only on  $S_t$  but also on  $S_{t+1-q}$ ; but we have not identified the function yet. We leave this issue for future research.

<sup>30</sup>While the expense ratio looks constant in Figure 4, it actually stochastically evolves over time in a small scale, between 0.01043 and 0.01048 in panel (a), and between 0.0078 and 0.0096 in panel (b).

<sup>31</sup>Equation (5.2) implies that  $E_t[F_t(\Pi_{t+1})/X_t]$  is a decreasing function of  $\alpha/\beta$ . Thus, although  $\alpha$  and  $\beta$  have changed significantly over the last few decades, since the ratio  $\alpha/\beta$  has been fairly stable (Figure 3), the model's prediction that the expected expense ratio is constant is consistent with the data.



(a) *Simulation.* The panel shows the time variation of our model's market return  $\Omega_{t+1}^{market} = R_{t+1}/P_t$  (solid line), expense ratio  $F_t(\Pi_{t+1})/X_t$  (dashed line), and aggregate fund flow  $\Delta(\alpha X_t)$  (dash-dotted line). For  $t = 0$ ,  $\Omega_{t+1}^{market} = 4.35\%$ ,  $F_t(\Pi_{t+1})/X_t = 0.01046$ , and  $\Delta(\alpha X_t) = 0.00041$ , which are all normalized to 1. The parameter values are  $r = 0.04$ ,  $\hat{\delta} = 0.1$ ,  $\sigma^2 = 0.01$ ,  $\sigma_s^2 = 0.001$ ,  $S_0 = 1$ ,  $\gamma^f = \gamma^m = \gamma^d = 1$ ,  $\alpha = 0.4138$ ,  $\beta = 0.000067$ , and  $\xi = 0.01$ . The value of  $\alpha$  is the percentage of U.S. households owing mutual funds (the average of 1993–2013). The value of  $\beta$  is the number of funds of the U.S. mutual fund industry divided by the number of U.S. households (the average of 1993–2013). Source: Investment Company Institute (2014).



(b) *Data.* The panel shows the time variation of S&P 500 total annual return including dividend (solid line), the U.S. mutual funds' average expense ratio (dashed line), and the aggregate fund flow per fund in the U.S. mutual fund industry (dash-dotted line). For year 1993, the S&P 500 return is 10.08%, the average expense ratio is 0.0096, and the per-fund aggregate fund flow is \$27 million, which are all normalized to 1. Source: Standard & Poor's and Investment Company Institute.

Figure 4: Market return, expense ratio, and aggregate fund flow: simulation vs. data.

The reason the expected expense ratio is independent of the market return is the following. The *level* of  $F_t(\Pi_{t+1})$  is positively related to the expected market return. This makes sense: since each manager’s outside option is high in good times when  $\hat{R}_t$  is high, the fund investors need to pay a large  $Q_t$  to let the manager continue running the fund. Meanwhile, in good times, the investors allocate a large  $X_t$  to the fund. Indeed, Figure 4 shows that the fund flow is positively related to the market return, both in simulation and data.<sup>32</sup> So when we look at the ratio  $F_t(\Pi_{t+1})/X_t$ , the effect of the market return is just offset between the numerator and the denominator, and it becomes a constant in expectation.<sup>33</sup>

### 5.2.2 Aggregate fund flows and stock market return

The model yields implications for the relation between aggregate fund flows and stock market returns. Note that studying aggregate fund flows (at macro level) is fundamentally different from studying fund flows between individual funds (at micro level).<sup>34</sup> As Warther (1995) explains, “[a]t macro level, flows between funds net out, so only aggregate flows into and out of the entire market remain. Thus, attention at the macro level centers on large-scale movements of money into and out of the market without regard to which fund it goes into or comes from.” Warther (1995) finds evidence of: (1) a positive relation between flows and subsequent returns, and (2) a negative relation between returns and subsequent flows.<sup>35</sup> The model provides theoretical support for both of these findings.

**Result 4.** *Over time,*

1. *aggregate fund flow  $\Delta(\alpha X_t)$  is positively related to subsequent market return  $\Omega_{t+1}^{market} = R_{t+1}/P_t$ ;*
2. *aggregate fund flow  $\Delta(\alpha X_t)$  is inversely related to past market return  $\Omega_t^{market} = R_t/P_{t-1}$ .*

---

<sup>32</sup>Although  $\Omega_{t+1}^{market}$  and  $\Delta(\alpha X_t)$  are positively related, they are not *perfectly* correlated. This is because  $\Omega_{t+1}^{market}$  evolves with two random variables,  $S_t$  and  $\delta_t$ , while  $\Delta(\alpha X_t)$  depends only on  $S_t$ .

<sup>33</sup>This logic goes through even under alternative measures of the expense ratio such as  $F_t(\Pi_{t+1})/Y_{t+1}$  and  $F_t(\Pi_{t+1})/(Y_{t+1} - F_t(\Pi_{t+1}))$ . Although we do not obtain simple expressions for the mean and standard deviation of these measures, the simulation results (which are not presented but available on request) show that the realized expense ratio is still much less volatile than the market return because the variation of  $S_t$  is cancelled out between the numerator and the denominator.

<sup>34</sup>For the micro approach, see, for example, Chevalier and Ellison (1997) and Sirri and Tufano (1998).

<sup>35</sup>Fant (1999) also finds evidence of a negative relation between returns and subsequent flows.

Statement 1 holds because both  $\Delta(\alpha X_t)$  and  $\Omega_{t+1}^{market}$  comove positively with  $S_t$ . When  $S_t$  increases,  $P_t$  falls, so  $\Omega_{t+1}^{market} = \frac{\delta_{t+1} + P_{t+1}}{P_t} - (1 + r)$  tends to increase (though it *can* decrease if the realization of  $\delta_{t+1}$  is very low). Because each fund investor chooses  $X_t$  based on the expectation of  $\Omega_{t+1}^{market}$ , flow  $\Delta(\alpha X_t)$  also increases.

Statement 2 contrasts with the well-documented empirical fact at *micro* level that individual fund's flow-performance relation is positive and convex (Chevalier and Ellison 1997). Statement 2 implies that, at *macro* level, fund investors are "somewhat contrarian" (Warther 1995). The result holds because the past market return  $\Omega_t^{market}$  decreases with  $S_t$  (because  $\Omega_t^{market} = \frac{\delta_t + P_t}{P_{t-1}} - (1 + r)$  is increasing in  $P_t$  that comoves negatively with  $S_t$ ), whereas the current flow  $\Delta(\alpha X_t)$  comoves positively with  $S_t$ . The intuition is as follows. If  $S_t$  increases,  $P_t$  falls. So the investors who bought the stock in period  $t - 1$  sell them cheap in period  $t$ , and thus  $\Omega_t^{market}$  decreases. But, because low  $P_t$  implies high  $\Omega_{t+1}^{market}$  (which is relevant for the investor's choice of  $X_t$ ), the current flow  $\Delta(\alpha X_t)$  increases.

We have two remarks on the robustness of Result 4. First, both statements 1 and 2 would hold even if  $S_t$  is independent over time or mean reverting (see Appendix G.2). As seen from the economic intuitions discussed above, the three elements needed for Result 4 are  $dX_t/dS_t > 0$ ,  $d\Omega_{t+1}^{market}/dS_t > 0$ , and  $d\Omega_t^{market}/dS_t < 0$ , none of which relies on our assumption that  $S_t$  is a random walk. Second, since Warther (1995) studies monthly return patterns, one may argue that his results may not be compatible with our model where each generation lives for only two periods. However, we will show in Section 6 that Result 4 still holds in a model with infinitely-lived agents, in which the length of each time period can be interpreted more flexibly.

### 5.2.3 Fund's trading strategy and stock market return: Volatility timing

The model also has implications for fund managers' trading strategies, such as volatility timing documented by Busse (1999). He finds that more than 80% of mutual funds in his sample decrease their market exposure when market volatility is high. Giambona and Golec (2009) also find that the majority of funds time volatility counter-cyclically. The model's prediction is consistent with their findings.

**Result 5.** *Fund's volatility timing is counter-cyclical. That is, each fund manager decreases the fund's risk exposure  $\theta_t$  when the market return volatility  $\sigma_t^{market}$  increases.*

Result 5 follows from Proposition 4.1 and (4.12). Over time,  $\theta_t$  and  $\sigma_t^{market}$  move in the opposite directions because  $\theta_t$  (resp.,  $\sigma_t^{market}$ ) moves in the opposite (resp., same) direction of  $S_t$ . The intuition is the following. Suppose that  $S_t$  increases in period  $t$ , and

thus  $P_t$  falls. On the one hand, this price fall increases  $\sigma_t^{market}$ .<sup>36</sup> On the other hand, the price fall increases  $\hat{R}_t$ , encouraging fund investors to increase  $X_t$ . In response to this, each manager optimally decreases  $\theta_t$  to rebalance his personal exposure to the stock.

An alternative measure of fund’s risk exposure is the leverage ratio, defined as the amount of borrowing divided by capital, i.e.,  $L_t \equiv (P_t\theta_t X_t - X_t)/X_t = P_t\theta_t - 1$ . For a similar economic intuition as the one for Result 5, the model predicts that funds tend to use leverage also counter-cyclically, that is,  $L_t$  and  $\sigma_t^{market}$  comove negatively over time.

#### 5.2.4 Market return vs. fund return

The model allows us to distinguish market return and fund return, identify their distributions, and offer testable implications on their equilibrium relation. The following result indicates that these returns may evolve quite differently over time.

**Result 6.** *Comparing the distributions of market return and fund return,*

1. *the expected market return  $\mu_t^{market}$  is time-varying (with supply shock  $S_t$ ), but the expected fund return  $\mu^{fund}$  is constant over time;*
2. *the market return volatility  $\sigma_t^{market}$  and the fund return volatility  $\sigma_t^{fund}$  comove negatively.*

The reason  $\mu^{fund}$  is constant over time (statement 1) is that the managers undo the impact of  $S_t$  on  $\mu^{fund}$ . To see this, consider what happens if  $S_t$  decreases. This increases  $P_t$ , which decreases  $\mu_t^{market}$  because one dollar buys a smaller number of shares of the stock. However, under portfolio delegation, each manager levers up each investor’s one dollar of capital so that the expected return from investing one dollar in the fund is still the same as before the rise in  $P_t$ . More specifically, even if the stock return  $\hat{R}_t$  decreases due to a decrease in  $S_t$ , the manager increases  $\theta_t$  so that the expected rate of return on the fund (before fee) is  $\hat{R}_t\theta_t$ , independent of  $S_t$ .<sup>37</sup>

Statement 2 of Result 6 implies that funds effectively provide investors with a “volatility hedge.” The intuition follows from Result 5 that the managers decrease the funds’

---

<sup>36</sup>The price fall allows investors to purchase a larger number of shares of the stock by one dollar. Thus, the return on the market per dollar invested increases in expectation but becomes more volatile too.

<sup>37</sup>Kaniel and Kondor (2013) argue that funds are less volatile than the market on average, that is, the average fund overperforms the market in recessions and underperforms in expansions. Although the economic mechanisms are different, their result and statement 1 of Result 6 share the idea that funds provide investors with more stable returns than the market over time. This “stabilizing role” of funds is supported empirically by Kacperczyk, Van Nieuwerbuygh, and Veldkamp (2012), who find that after-fee excess fund returns are negative in expansions and positive in recessions.

risk exposure when market volatility increases. The volatility hedge is consistent with the agents' incentives as follows. Given that each fund provides a volatility hedge, each fund investor is encouraged to increase  $X_t$  when the market return volatility  $\sigma_t^{market}$  increases (i.e., when  $S_t$  increases). In response to the increased  $X_t$ , each manager decreases  $\theta_t$  to balance his personal risk exposure; this consistently translates back into the lower fund return volatility  $\sigma_t^{fund}$ , indeed providing the volatility hedge.

## 6 Long-lived Agents

This section studies a model with infinitely-lived agents and shows that our main results in the overlapping-generations model still carry through.

This section's model is much like the one of Section 2 except that the agents are infinitely lived; unless otherwise noted, the same assumptions prevail. The agents have utility function  $u^i(c_t^i) = -\exp(-\tilde{\gamma}^i c_t^i)$  for  $i \in \{f, m, d\}$ , where  $c_t^i$  is consumption and  $\tilde{\gamma}^i > 0$  is a coefficient of absolute risk aversion. Due to the model's recursive nature, it is convenient to define a "time-adjusted" CARA coefficient  $\gamma^i \equiv \frac{r}{1+r} \tilde{\gamma}^i$ . For  $i \in \{f, m, d\}$ , each agent's objective function in period  $t$  is her/his lifetime expected utility  $E_t [\sum_{\tau=0}^{\infty} \rho^\tau u^i(c_{t+\tau}^i)]$ , where  $\rho \in (0, 1)$  is a discount factor common for all agents.

In each period  $t + 1$ , each fund investor pays fee  $F_{t+1}$  to the manager. Since  $F_{t+1}$  compensates the manager for the investment he made in period  $t$  (which yields a return in period  $t + 1$ ), we allow  $F_{t+1}$  to depend not only on the fund profit  $\Pi_{t+1}$  but also on the period- $t$  investment profile  $(\theta_t, X_t)$  and the period- $t$  state variable  $S_t$ . At the beginning of  $t = 0$ , the fund investors collectively offer a fee schedule, represented by a function  $F(\Pi_{t+1}, \theta_t, X_t, S_t)$ , to the manager. We allow  $F$  to be any function. If the manager accepts  $F$  in  $t = 0$ , then for all  $t \geq 0$  the fee in that fund will be  $F_{t+1} = F(\Pi_{t+1}, \theta_t, X_t, S_t)$ . Note that, while the *functional form* of  $F$  is determined in  $t = 0$  and does not change thereafter, the *value* of the fee  $F_{t+1}$  changes every period, depending on the realizations of  $(\Pi_{t+1}, \theta_t, X_t, S_t)$ .

We assume that each manager can abscond in any period  $t \geq 0$ .<sup>38</sup> If he absconds in period  $t$ , then he acts as a direct investor from period  $t + 1$  onward.<sup>39</sup> The story behind this assumption is that if a manager ever deceives investors, his name would be mud and

---

<sup>38</sup>As in the model of Section 2, each manager who absconds in period  $t$  is not allowed to change  $\theta_t$  in that period. However, as shown in Appendix G.7, the results would not change even if we allow him to modify  $\theta_t$  when absconding.

<sup>39</sup>While acting as a direct investor, his risk aversion is still represented by  $\gamma^m$  (instead of  $\gamma^d$ ).

he can never come back to the profession; he lives the rest of his life as an individual investor, using the investors' money he absconded with as seed money. When offering  $F$  in  $t = 0$ , the fund investors make sure that the manager will never have an incentive to abscond in the future. That is, they consider incentive compatibility constraints à la (2.7) for each and every period  $t \geq 0$  (see Appendix F for details).

We look for a linear equilibrium in which  $P_t$  is linear in  $\hat{\delta}$  and  $S_t$ . The equilibrium is derived in Appendix F, using the standard dynamic programming technique. The following proposition summarizes the equilibrium outcome.

**Proposition 6.1** (Long-lived agents). *Define  $\pi \equiv \frac{\beta}{\xi\gamma^m} + \frac{1-\alpha}{\gamma^d}$ ,  $\eta \equiv \frac{\sigma^2}{\pi}$ ,  $q_1^d \equiv \frac{\sigma^2\sigma_s^2}{r\pi^2 + \sigma^2\sigma_s^2}$ ,  $q_2^d \equiv \sigma^2 \left(1 + \frac{\sigma^2\sigma_s^2}{r(r\pi^2 + \sigma^2\sigma_s^2)}\right)$ , and  $A^d \equiv \frac{\sigma^2}{r\pi^2}$ , and let  $(\zeta_X, q_1^f, q_1^m, q_2^f, q_2^m, q_3, q_4, A^f, A^m)$  be the solutions to the system of nine equations:*

$$\zeta_X = \frac{\pi^2}{(1+r)\xi^3\gamma^{m2}} \left( \gamma^m \xi \sigma^2 \left(1 - \xi + \frac{q_3}{r}\right) - \frac{\beta}{\alpha} q_4 \right), \quad (6.1)$$

$$q_1^i = \frac{A^i \sigma_s^2}{1 + A^i \sigma_s^2} \quad \text{for } i \in \{f, m\}, \quad (6.2)$$

$$q_2^i = \sigma_R^2 - \frac{\eta^2 \sigma_s^2 q_1^i}{r^2} \quad \text{for } i \in \{f, m\}, \quad (6.3)$$

$$q_3 = (1 - \phi) q_1^f + \phi q_1^m - \xi q_1^d, \quad (6.4)$$

$$q_4 = \gamma^f (1 - \phi)^2 q_2^f + \gamma^m \frac{\alpha}{\beta} (\phi^2 q_2^m - \xi^2 q_2^d), \quad (6.5)$$

$$A^f = \frac{1}{1+r} \left( \frac{\gamma^f \beta^2 q_4}{\gamma^{m2} \alpha^2 \xi^2 \pi^2} + \frac{A^f}{1 + A^f \sigma_s^2} - \frac{\gamma^f \beta}{\gamma^m \alpha} \left( \frac{A^d}{1 + A^d \sigma_s^2} - \frac{A^m}{1 + A^m \sigma_s^2} \right) \right), \quad \text{and} \quad (6.6)$$

$$A^m = A^d + 2\gamma^m \xi \zeta_X. \quad (6.7)$$

Then, in equilibrium,

1. the stock return and price characterizations are similar to those of the short-lived agents case (Proposition 4.1) with  $\eta^{**}$  being replaced by  $\eta = \sigma^2/\pi$ ,
2. each fund's fee schedule is  $F(\Pi_{t+1}, \theta_t, X_t, S_t) = \phi \Pi_{t+1} + Q(\theta_t, X_t, S_t)$ , where  $\phi \equiv \beta\gamma^f / (\alpha\gamma^m + \beta\gamma^f)$  and

$$\begin{aligned} Q(\theta_t, X_t, S_t) \equiv & \left( \xi \left(1 + \frac{q_1^d}{r}\right) - \phi \left(1 + \frac{q_1^m}{r}\right) \right) \eta S_t \theta_t X_t - \frac{\gamma^m \alpha}{2 \beta} (\xi^2 q_2^d - \phi^2 q_2^m) \theta_t^2 X_t^2 \\ & + \xi(1+r)X_t + \frac{1}{2\gamma^m} \frac{\beta}{\alpha} \left( \left( \frac{q_1^d - q_1^m}{\sigma_s^2} \right) S_t^2 + \ln \left( \frac{1 - q_1^m}{1 - q_1^d} \right) \right), \end{aligned} \quad (6.8)$$



3. each fund's size (i.e., assets under management) is  $(\alpha/\beta)X_t = \zeta_X S_t^2$ , and

4. each fund buys  $(\alpha/\beta)\theta_t X_t$  shares of the stock, where

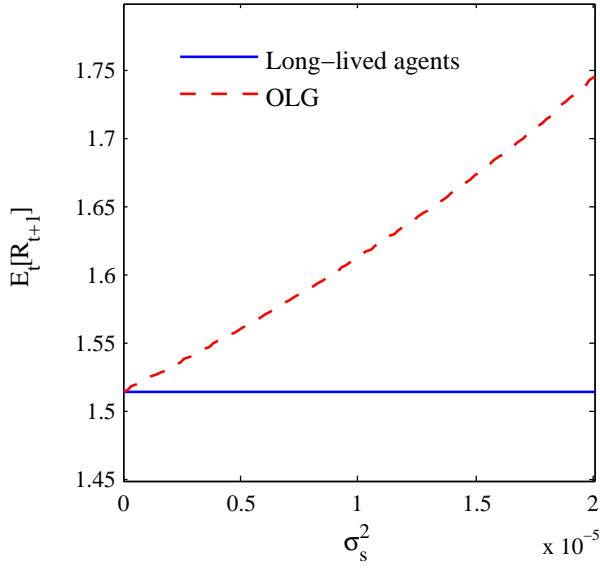
$$\theta_t = \frac{\zeta_\theta}{S_t} \quad \text{with} \quad \zeta_\theta \equiv \frac{(1+r)\xi^2\gamma^m\pi}{\gamma^m\xi\sigma^2\left(1-\xi+\frac{q_3}{r}\right) - \frac{\beta}{\alpha}q_4}. \quad (6.9)$$

The equilibrium characterization of the long-lived agents case is more complicated than that of the overlapping-generations case due to the agents' concerns for hedging against stochastic variation of  $S_t$ , which are captured by  $q_1^i$ ,  $i \in \{f, m, d\}$ . Nevertheless, the implications drawn from these two cases are similar. Indeed, *all* the qualitative results presented in Section 5 carry through also in this section's model. First, because we have  $d\eta/d\alpha > 0$  and  $d\eta/d\beta < 0$  in this section's model, both  $\mu_t^{market} = \hat{R}_t/P_t = r\eta S_t/(\hat{\delta} - \eta S_t)$  and  $SR_t^{market} = (\eta/\sigma_R)S_t$  increase with  $\alpha$  and decrease with  $\beta$  (Result 1), and the same applies to  $\text{Var}_t[P_{t+1}] = \eta^2\sigma_s^2/r^2$  (Result 2). Second, since in both models the fee  $F_{t+1}$  is an affine function of  $\Pi_{t+1}$  where  $Q_t$  is an affine function of  $S_t^2$ , the qualitative implication of Result 3 is still derived from this section's model.<sup>40</sup> Third, in both Sections 4 and 6, in equilibrium  $\hat{R}_t$ ,  $X_t$ ,  $\theta_t$ , and  $K_t$  are proportional to  $S_t$ ,  $S_t^2$ ,  $1/S_t$ , and  $S_t$ , respectively. So the two models' endogenous variables move in the same direction over time, and hence Results 4, 5, and 6 hold also in this section's model.<sup>41</sup>

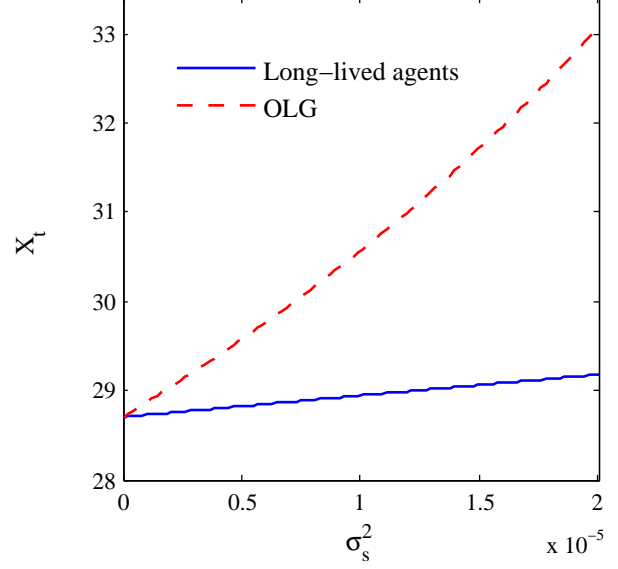
How does the assumption of long-lived agents change their behavior and stock price? Figure 5 plots the key endogenous variables of the long-lived agents model of this section (solid line) and those of the overlapping generations model of Section 2 (dashed line) as functions of the volatility  $\sigma_s^2$  of stock supply. The overlapping-generations case is straightforward. As  $\sigma_s^2$  increases, the agents demand a higher risk premium, pushing up  $\hat{R}_t$  (panel (a)). This encourages fund investors to increase  $X_t$  (panel (b)), which then

<sup>40</sup>Cuoco and Kaniel (2011) also consider a multi-period problem of portfolio delegation, but a linear fee is not optimal in their model. Why is it not optimal in theirs, while it is in ours? In Section 6 of Cuoco and Kaniel (2011), the authors argue that the critical assumption that prevents a linear fee from being optimal is that investors take the fee structure as given and do not internalize the fact that their choice of capital allocation affects the equilibrium fee. The authors note (page 289): “*If fund investors were able to choose the managers' compensation contract while committing to delegating the amount [...], this [linear] fee would indeed be optimal.*” By contrast, in our model the fund investors choose the fee structure, and each of them fully internalizes how her choice of capital allocation  $X_t$  will affect the equilibrium fee because the load component  $Q_t$  is a *function* of  $X_t$  (as opposed to Cuoco and Kaniel (2011) in which the load component is assumed to be a *constant*). Consequently, a linear fee is still optimal in our setting, consistent with the standard results such as Wilson (1968) and Ross (1973).

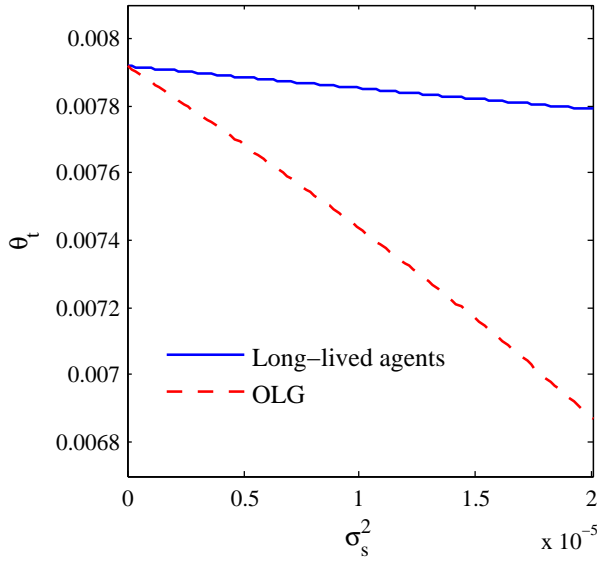
<sup>41</sup>Moreover, the funds' aggregate stock holding,  $\alpha\theta_t X_t = \frac{\beta}{\xi\gamma^m\pi}S_t$ , and the direct investors' aggregate stock holding,  $(1-\alpha)K_t = \frac{1-\alpha}{\gamma^d\pi}S_t$ , are also the same as those of Section 4 (see Footnote 24).



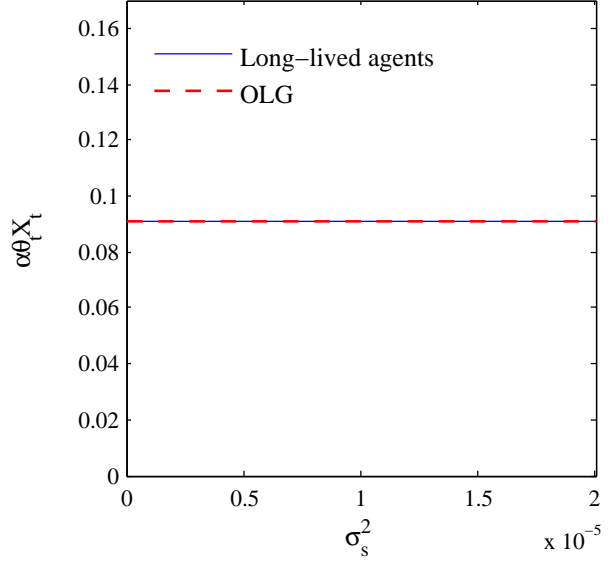
(a) Expected stock return:  $\hat{R}_t$



(b) Fund investor's capital allocation:  $X_t$



(c) Fund's risk exposure:  $\theta_t$



(d) Funds' aggregate stock holding:  $\alpha\theta_t X_t$

Figure 5: Long-lived-agents model vs. overlapping-generations (OLG) model. The graphs plot the key endogenous variables versus the volatility  $\sigma_s^2$  of stock supply. The parameter values are  $\sigma^2 = 1$ ,  $r = 0.02$ ,  $\xi = 0.01$ ,  $\gamma^f = \gamma^m = \gamma^d = 1$ ,  $\alpha = 0.4$ ,  $\beta = 0.0006$ , and  $S_t = 1$ .

leads the manager to decrease  $\theta_t$  to counterbalance his personal risk exposure (panel (c)). The funds' aggregate stock holding  $\alpha\theta_t X_t$  is independent of  $\sigma_s^2$  because the increase in  $X_t$  and decrease in  $\theta_t$  just offset each other (panel (d)). If  $\sigma_s^2 = 0$ , even the long-lived agents choose mean-variance solutions and thus the two models' solutions coincide in all four panels. However, for  $\sigma_s^2 > 0$  the long-lived agents' hedging demand creates discrepancies between the models. The expected return  $\hat{R}_t$  is independent of  $\sigma_s^2$  in the long-lived agents case (panel (a)). This is because an increase in  $\sigma_s^2$  has two opposing effects on  $\hat{R}_t$ . First, it increases risk premium, having an upward pressure on  $\hat{R}_t$ ; second, it increases hedging demand, having an upward pressure on  $P_t$ , i.e., a downward pressure on  $\hat{R}_t$ . These two effects just offset each other, leading to  $d\hat{R}_t/d\sigma_s^2 = 0$ . Despite  $d\hat{R}_t/d\sigma_s^2 = 0$ , each fund investor's  $X_t$  increases with  $\sigma_s^2$  (panel (b)) because her hedging demand increases with  $\sigma_s^2$ . The increase in  $X_t$  is moderate compared to the overlapping-generations case, in which  $\hat{R}_t$  increases with  $\sigma_s^2$  to drive up  $X_t$  more sharply. As a best response to such a moderate increase in  $X_t$ , the manager's  $\theta_t$  decreases with  $\sigma_s^2$  also moderately (panel (c)). The funds' aggregate stock holding  $\alpha\theta_t X_t$  is independent of  $\sigma_s^2$  also in the long-lived agents case because, again, the increase in  $X_t$  and decrease in  $\theta_t$  just cancel out (panel (d)).

## 7 Conclusion

This paper develops a model of asset-market equilibrium that incorporates delegated portfolio management with optimal fee contracts. Fund managers and investors strategically interact to determine funds' investment profiles, while they share the funds' portfolio risk via fee contracts. In equilibrium, their investment decisions, fee schedules, and stock prices feed back into one another. The model yields a number of empirical implications for fund size, fund fees, funds' investment strategies, stock market returns, and fund returns. It predicts that expected stock returns and price volatility tend to be high when a large fraction of investors access the market indirectly through a small number of funds (the risk concentration effect). Explanations are provided to the empirical facts documented in the literature: (1) a positive relation between aggregate fund flows and subsequent stock returns and a negative relation between returns and subsequent flows, and (2) funds' counter-cyclical volatility timing. Moreover, new testable implications on the distributions of market return and fund return are offered: (1) expected fund return is less volatile than expected market return, and (2) funds use leverage counter-cyclically to balance their risk exposure, that is, funds provide investors with a volatility hedge.

In our model, time variation of variables is caused by exogenous supply shocks (or noise trading). It will be a fruitful future research to endogenize these shocks to study the welfare implications of portfolio delegation. Given the policy debates about the effect of financial intermediaries on the real economy, such a study will be important.

A limitation of our model is that it does not allow us to analyze cross-sectional variations of funds. Indeed, in equilibrium fund returns are perfectly correlated, and all funds have the same size. These are clearly counterfactual. For future research, it will be interesting to consider heterogeneities of funds (such as investment styles or managers' skill), possibly in a multiple-stock setting, and explore the interactions of funds' investment profiles and returns, manager compensations, and asset prices.

## References

- [1] Albagli, E., 2015, "Investment Horizons and Asset Prices Under Asymmetric Information," *Journal of Economic Theory*, 158, 787–837.
- [2] Allen, F. and G. Gorton, 1993, "Churning Bubbles," *Review of Economic Studies*, 60, 813–836.
- [3] Banerjee, S., 2011, "Learning from Prices and the Dispersion in Beliefs," *Review of Financial Studies*, 24, 3025–3068.
- [4] Basak, S. and A. Pavlova, 2012, "Asset Prices and Institutional Investors," *American Economic Review*, forthcoming.
- [5] Bennett, J. A., Sias, R. W., and L. T. Starks, 2003, "Greener Pastures and the Impact of Dynamic Institutional Preferences," *Review of Financial Studies*, 16, 1203–1238.
- [6] Buffa, A., Vayanos, D., and P. Woolley, 2014, "Asset Management Contracts and Equilibrium Prices," Working Paper, Boston University and London School of Economics.
- [7] Busse, J., 1999, "Volatility Timing in Mutual Funds: Evidence from Daily Returns," *Review of Financial Studies*, 12, 1009–1041.
- [8] Campbell, J. and A. Kyle, 1993, "Smart Money, Noise Trading and Stock Price Behaviour," *Review of Economic Studies*, 60, 1–34.

- [9] Campbell, J., Lettau, M., Malkiel, B., and Y. Xu, 2001, “Have Individual Stocks Become More Volatile? An Empirical Exploration of Idiosyncratic Risk,” *Journal of Finance*, 56, 1–43.
- [10] Cespa, G. and X. Vives, 2015, “The Beauty Contest and Short-Term Trading,” *Journal of Finance*, 70, 2099–2154.
- [11] Chevalier, J. and G. Ellison, 1997, “Risk Taking by Mutual Funds as a Response to Incentives,” *Journal of Political Economy*, 105, 1167–1200.
- [12] Cuoco, D. and R. Kaniel, 2011, “Equilibrium Prices in the Presence of Delegated Portfolio Management,” *Journal of Financial Economics*, 101, 264–296.
- [13] Fant, L. F., 1999, “Investment Behavior of Mutual Fund Shareholders: The Evidence from Aggregate Fund Flows,” *Journal of Financial Markets*, 2, 391–402.
- [14] Giambona, E. and J. Golec, 2009, “Mutual fund volatility timing and management fees,” *Journal of Banking and Finance*, 33, 589–599.
- [15] He, Z. and A. Krishnamurthy, 2012, “Intermediary Asset Pricing,” *American Economic Review*, forthcoming.
- [16] Hugonnier, J. and R. Kaniel, 2010, “Mutual fund portfolio choice in the presence of dynamic flows,” *Mathematical Finance*, 20, 187–227.
- [17] Investment Company Institute, 2014, *2014 Investment Company Fact Book*, 54th Edition.
- [18] Kacperczyk, M., Van Nieuwerburgh, S., and L. Veldkamp, 2012, “Rational Attention Allocation Over the Business Cycle,” Working Paper, New York University.
- [19] Kaniel, R. and P. Kondor, 2013, “The Delegated Lucas Tree,” *Review of Financial Studies*, 26, 929–984.
- [20] LeRoy, S. F., and R. D. Porter, 1981, “The Present-Value Relation: Tests Based on Implied Variance Bounds,” *Econometrica*, 49, 555–574.
- [21] Malliaris, S. and H. Yan, 2012, “Reputation Concerns and Slow-Moving Capital,” Working Paper, Yale University.

- [22] Ou-Yang, H., 2003, “Optimal Contracts in a Continuous-Time Delegated Portfolio Management Problem,” *Review of Financial Studies*, 16, 173–208.
- [23] Ross, S., 1973, “The Economic Theory of Agency: The Principal’s Problem,” *American Economic Review*, 63, 134–139.
- [24] Sato, Y., 2014, “Opacity in Financial Markets,” *Review of Financial Studies*, 27, 3502–3546.
- [25] Shiller, R. J., 1981, “Do Stock Prices Move Too Much to be Justified by Subsequent Changes in Dividends?,” *American Economic Review*, 71, 421–436.
- [26] Shleifer, A. and R. Vishny, 1997, “The Limits of Arbitrage,” *Journal of Finance*, 52, 35–55.
- [27] Sirri, E. R., and P. Tufano, 1998, “Costly Search and Mutual Fund Flows,” *Journal of Finance*, 53, 1589–1622.
- [28] Spiegel, M., 1998, “Stock Price Volatility in a Multiple Security Overlapping Generations Model,” *Review of Financial Studies*, 11, 419–447.
- [29] Sundaresan, M., 1983, “Constant Absolute Risk Aversion Preferences and Constant Equilibrium Interest Rates,” *Journal of Finance*, 38, 205–212.
- [30] Vayanos, D., 2004, “Flight to Quality, Flight to Liquidity, and the Pricing of Risk,” Working Paper, London School of Economics.
- [31] Vayanos, D. and P. Woolley, 2013, “An Institutional Theory of Momentum and Reversal,” *Review of Financial Studies*, 26, 1087–1145.
- [32] Wang, J., 1993, “A Model of Intertemporal Asset Prices Under Asymmetric Information,” *Review of Economic Studies*, 60, 249–282.
- [33] Warther, V., 1995, “Aggregate Mutual Fund Flows and Security Returns,” *Journal of Financial Economics*, 39, 209–235.
- [34] Watanabe, M., 2008, “Price Volatility and Investor Behavior in an Overlapping Generations Model with Information Asymmetry,” *Journal of Finance*, 63, 229–272.
- [35] Wilson, R., 1968, “The Theory of Syndicates,” *Econometrica*, 36, 119–132.

- [36] Xu, Y. and B. G. Malkiel, 2003, “Investigating the Behavior of Idiosyncratic Volatility,” *Journal of Business*, 76, 613–645.

# Online Appendix: Delegated Portfolio Management, Optimal Fee Contracts, and Asset Prices

Yuki Sato

University of Lausanne and Swiss Finance Institute

## A Proof of Lemma 4.1

For an arbitrary investment profile  $(\theta_t, X_t)$ , generation- $t$  fund investors choose fee schedule  $F_t(\cdot)$  that maximizes the expected utility of each of them, subject to the manager's incentive compatibility constraint. The manager's participation constraint,  $\mathbb{E}_t[u^m(W_{t+1}^m)] \geq \mathbb{E}_t[u^m(0)]$ , is slack because it is implied by the incentive compatibility constraint. That is, the problem is

$$\begin{aligned} & \max_{F_t(\cdot)} \mathbb{E}_t \left[ u^f(W_{t+1}^f) \right] \\ \text{sub. to } & \mathbb{E}_t \left[ u^m(W_{t+1}^m) \right] \geq \mathbb{E}_t \left[ u^m(\widetilde{W}_{t+1}^m) \right]. \end{aligned} \quad (\text{A.1})$$

The Lagrangian is

$$\mathcal{L} = \int u^f(\Pi - F_t(\Pi) + (1+r)W) g_t(\Pi) d\Pi - \psi_t \left( \mathbb{E}_t \left[ u^m(\widetilde{W}_{t+1}^m) \right] - \int u^m \left( \frac{\alpha}{\beta} F_t(\Pi) \right) g_t(\Pi) d\Pi \right),$$

where  $g_t(\cdot)$  is the probability density function (conditional on  $t$ ) of the fund profit  $\Pi_{t+1}$ , and  $\psi_t$  is the Lagrange multiplier. The first-order condition (FOC) for  $F_t(\Pi_{t+1})$  is

$$\begin{aligned} & -u^{f'}(\Pi_{t+1} - F_t(\Pi_{t+1}) + (1+r)W) g_t(\Pi_{t+1}) + \psi_t u^{m'} \left( \frac{\alpha}{\beta} F_t(\Pi_{t+1}) \right) \frac{\alpha}{\beta} g_t(\Pi_{t+1}) = 0 \\ \iff & \frac{u^{m'} \left( \frac{\alpha}{\beta} F_t(\Pi_{t+1}) \right)}{u^{f'}(\Pi_{t+1} - F_t(\Pi_{t+1}) + (1+r)W)} = \frac{\beta}{\alpha} \frac{1}{\psi_t}. \end{aligned} \quad (\text{A.2})$$

Log-differentiating (A.2) with respect to  $\Pi_{t+1}$ ,

$$\begin{aligned} & \frac{u^{m''}(W_{t+1}^m)}{u^{m'}(W_{t+1}^m)} \frac{\alpha}{\beta} F_t'(\Pi_{t+1}) - \frac{u^{f''}(W_{t+1}^f)}{u^{f'}(W_{t+1}^f)} (1 - F_t'(\Pi_{t+1})) = 0 \\ \iff & -\gamma^m \frac{\alpha}{\beta} F_t'(\Pi_{t+1}) + \gamma^f (1 - F_t'(\Pi_{t+1})) = 0 \iff F_t'(\Pi_{t+1}) = \frac{\beta \gamma^f}{\alpha \gamma^m + \beta \gamma^f} \equiv \phi. \end{aligned} \quad (\text{A.3})$$

(A.3) implies that the optimal fee schedule is

$$F_t(\Pi_{t+1}) = \phi \Pi_{t+1} + Q_t$$



for some  $Q_t$  that is known as of period  $t$ . The fund investors choose  $Q_t$  such that the manager's incentive compatibility constraint (A.1) holds with equality, since otherwise they can be better off by decreasing  $Q_t$  until (A.1) binds. Thus, recalling that the manager has exponential utility,  $Q_t$  satisfies

$$\begin{aligned}
\mathbb{E}_t [u^m(W_{t+1}^m)] &= \mathbb{E}_t [u^m(\widetilde{W}_{t+1}^m)] \iff \mathbb{E}_t [W_{t+1}^m] - \frac{\gamma^m}{2} \text{Var}_t [W_{t+1}^m] = \mathbb{E}_t [\widetilde{W}_{t+1}^m] - \frac{\gamma^m}{2} \text{Var}_t [\widetilde{W}_{t+1}^m] \\
&\iff \frac{\alpha}{\beta} (\phi \mathbb{E}_t [\Pi_{t+1}] + Q_t) - \frac{\gamma^m}{2} \frac{\alpha^2}{\beta^2} \phi^2 \text{Var}_t [\Pi_{t+1}] = \frac{\alpha}{\beta} \xi \mathbb{E}_t [Y_{t+1}] - \frac{\gamma^m}{2} \frac{\alpha^2}{\beta^2} \xi^2 \text{Var}_t [Y_{t+1}] \\
&\iff \phi \mathbb{E}_t [\Pi_{t+1}] + Q_t - \frac{\gamma^m}{2} \frac{\alpha}{\beta} \phi^2 \text{Var}_t [\Pi_{t+1}] = \xi \mathbb{E}_t [\Pi_{t+1}] + \xi(1+r)X_t - \frac{\gamma^m}{2} \frac{\alpha}{\beta} \xi^2 \text{Var}_t [\Pi_{t+1}] \\
&\iff Q_t = (\xi - \phi) \mathbb{E}_t [\Pi_{t+1}] + \xi(1+r)X_t - \frac{\gamma^m}{2} \frac{\alpha}{\beta} (\xi^2 - \phi^2) \text{Var}_t [\Pi_{t+1}]. \quad \square
\end{aligned}$$

## B Proof of Lemma 4.2

Given  $F_t(\cdot)$  specified in Lemma 4.1, each fund investor's terminal wealth (2.5) is rewritten as

$$W_{t+1}^f = (1 - \phi) \Pi_{t+1} - (\xi - \phi) \mathbb{E}_t [\Pi_{t+1}] - \xi(1+r)X_t + \frac{\gamma^m}{2} \frac{\alpha}{\beta} (\xi^2 - \phi^2) \text{Var}_t [\Pi_{t+1}] + (1+r)W,$$

which implies that the mean and variance of  $W_{t+1}^f$  are

$$\mathbb{E}_t [W_{t+1}^f] = (1 - \xi) \mathbb{E}_t [\Pi_{t+1}] - \xi(1+r)X_t + \frac{\gamma^m}{2} \frac{\alpha}{\beta} (\xi^2 - \phi^2) \text{Var}_t [\Pi_{t+1}] + (1+r)W \quad \text{and} \quad (\text{B.1})$$

$$\text{Var}_t [W_{t+1}^f] = (1 - \phi)^2 \text{Var}_t [\Pi_{t+1}]. \quad (\text{B.2})$$

Since each fund investor has exponential utility, (B.1) and (B.2) imply that her maximization problem is

$$\begin{aligned}
\max_{X_t} \mathbb{E}_t [u^f(W_{t+1}^f)] &\iff \max_{X_t} \mathbb{E}_t [W_{t+1}^f] - \frac{\gamma^f}{2} \text{Var}_t [W_{t+1}^f] \\
&\iff \max_{X_t} (1 - \xi) \mathbb{E}_t [\Pi_{t+1}] - \xi(1+r)X_t + (1+r)W - \frac{\gamma^m}{2} \frac{\alpha}{\beta} \left( -\xi^2 + \phi^2 + \frac{\beta \gamma^f}{\alpha \gamma^m} (1 - \phi)^2 \right) \text{Var}_t [\Pi_{t+1}] \\
&\iff \max_{X_t} (1 - \xi) \hat{R}_t \theta_t X_t - \xi(1+r)X_t + (1+r)W - \frac{\gamma^m}{2} \frac{\alpha}{\beta} (\phi - \xi^2) \theta_t^2 X_t^2 \sigma_R^2.
\end{aligned}$$

The FOC for  $X_t$ , assuming interior solution, is

$$(1 - \xi) \hat{R}_t \theta_t - \xi(1+r) - \gamma^m \frac{\alpha}{\beta} (\phi - \xi^2) \theta_t^2 X_t \sigma_R^2 = 0.$$

This yields the fund investor's best response to  $\theta_t$  given  $\hat{R}_t$ :

$$X(\theta_t; \hat{R}_t) = \frac{(1 - \xi) \hat{R}_t \theta_t - \xi(1+r)}{\gamma^m \frac{\alpha}{\beta} (\phi - \xi^2) \sigma_R^2 \theta_t^2}.$$

Taking care of the nonnegativity condition for  $X_t$  yields the required result.  $\square$

## C Proof of Lemma 4.3

Under fee schedule  $F_t(\cdot)$  of Lemma 4.1, each manager's terminal wealth (2.4) is rewritten as

$$\begin{aligned} W_{t+1}^m &= \frac{\alpha}{\beta} (\phi \Pi_{t+1} + Q_t) \\ &= \frac{\alpha}{\beta} \left( \phi \Pi_{t+1} + (\xi - \phi) E_t[\Pi_{t+1}] + \xi(1+r)X_t - \frac{\gamma^m \alpha}{2} \frac{\alpha}{\beta} (\xi^2 - \phi^2) \text{Var}_t[\Pi_{t+1}] \right), \end{aligned}$$

which implies that the mean and variance of  $W_{t+1}^m$  are

$$E_t [W_{t+1}^m] = \frac{\alpha}{\beta} \left( \xi E_t[\Pi_{t+1}] + \xi(1+r)X_t - \frac{\gamma^m \alpha}{2} \frac{\alpha}{\beta} (\xi^2 - \phi^2) \text{Var}_t[\Pi_{t+1}] \right) \quad \text{and} \quad (\text{C.1})$$

$$\text{Var}_t [W_{t+1}^m] = \frac{\alpha^2}{\beta^2} \phi^2 \text{Var}_t [\Pi_{t+1}]. \quad (\text{C.2})$$

Since the manager has exponential utility, (C.1) and (C.2) imply that his maximization problem is

$$\begin{aligned} \max_{\theta_t} E_t [u^m(W_{t+1}^m)] &\iff \max_{\theta_t} E_t [W_{t+1}^m] - \frac{\gamma^m}{2} \text{Var}_t [W_{t+1}^m] \\ &\iff \max_{\theta_t} \frac{\alpha}{\beta} \xi E_t[\Pi_{t+1}] + \frac{\alpha}{\beta} \xi(1+r)X_t - \frac{\gamma^m \alpha^2}{2} \frac{\alpha}{\beta^2} \xi^2 \text{Var}_t[\Pi_{t+1}] \\ &\iff \max_{\theta_t} \frac{\alpha}{\beta} \xi X_t \left( \hat{R}_t \theta_t + (1+r) - \frac{\gamma^m \alpha}{2} \frac{\alpha}{\beta} \xi \theta_t^2 X_t \sigma_R^2 \right). \end{aligned}$$

The FOC for  $\theta_t$  is

$$\hat{R}_t - \gamma^m \frac{\alpha}{\beta} \xi \theta_t X_t \sigma_R^2 = 0.$$

This yields the manager's best response to  $X_t$  given  $\hat{R}_t$ :

$$\theta(X_t; \hat{R}_t) = \frac{\beta \hat{R}_t}{\gamma^m \alpha \xi \sigma_R^2 X_t}. \quad \square$$

## D Proof of Proposition 4.2

The fixed fee  $Q_t$  is computed as follows. In equilibrium, the conditional mean and variance of  $\Pi_{t+1}$  are

$$E_t[\Pi_{t+1}] = \hat{R}_t \theta_t^* X_t^* = \left( \frac{\beta \sigma_R^2}{\alpha \xi \gamma^m \left( \frac{\beta}{\xi \gamma^m} + \frac{1-\alpha}{\gamma^d} \right)^2} \right) S_t^2 \quad \text{and} \quad (\text{D.1})$$

$$\text{Var}_t[\Pi_{t+1}] = \theta_t^{*2} X_t^{*2} \sigma_R^2 = \left( \frac{\beta^2 \sigma_R^2}{\alpha^2 \xi^2 \gamma^{m2} \left( \frac{\beta}{\xi \gamma^m} + \frac{1-\alpha}{\gamma^d} \right)^2} \right) S_t^2. \quad (\text{D.2})$$

Using (D.1) and (D.2), we have

$$\begin{aligned}
Q_t &= (\xi - \phi) \left( \frac{\beta \sigma_R^2}{\alpha \xi \gamma^m \left( \frac{\beta}{\xi \gamma^m} + \frac{1-\alpha}{\gamma^d} \right)^2} \right) S_t^2 + \xi(1+r) \left( \frac{\beta(\xi - \phi) \sigma_R^2}{\gamma^m(1+r) \alpha \xi^3 \left( \frac{\beta}{\xi \gamma^m} + \frac{1-\alpha}{\gamma^d} \right)^2} \right) S_t^2 \\
&\quad - \frac{\gamma^m \alpha}{2 \beta} (\xi^2 - \phi^2) \left( \frac{\beta^2 \sigma_R^2}{\alpha^2 \xi^2 \gamma^{m2} \left( \frac{\beta}{\xi \gamma^m} + \frac{1-\alpha}{\gamma^d} \right)^2} \right) S_t^2 \\
&= \left( \frac{\beta(\xi - \phi) \sigma_R^2}{\alpha \xi \gamma^m \left( \frac{\beta}{\xi \gamma^m} + \frac{1-\alpha}{\gamma^d} \right)^2} \right) \left( 1 + \frac{1}{\xi} - \frac{\xi + \phi}{2\xi} \right) S_t^2 = \left( \frac{\beta(\xi - \phi)(\xi - \phi + 2) \sigma_R^2}{2\alpha \xi^2 \gamma^m \left( \frac{\beta}{\xi \gamma^m} + \frac{1-\alpha}{\gamma^d} \right)^2} \right) S_t^2. \quad \square
\end{aligned}$$

## E Proof of Corollary 4.1

We have

$$\begin{aligned}
\mu_t^{fund} &= E_t \left[ \Omega_{t+1}^{fund} \right] = E_t \left[ \frac{\Pi_{t+1} - F_t(\Pi_{t+1})}{X_t^*} \right] = \frac{1}{X_t^*} ((1 - \phi) E_t[\Pi_{t+1}] - Q_t) \\
&= \left( \frac{\gamma^m(1+r) \alpha \xi^3 \left( \frac{\beta}{\xi \gamma^m} + \frac{1-\alpha}{\gamma^d} \right)^2}{\beta(\xi - \phi) \sigma_R^2} \right) \frac{1}{S_t^2} \left( \left( \frac{(1 - \phi) \beta \sigma_R^2}{\alpha \xi \gamma^m \left( \frac{\beta}{\xi \gamma^m} + \frac{1-\alpha}{\gamma^d} \right)^2} \right) S_t^2 - \left( \frac{\beta(\xi - \phi)(\xi - \phi + 2) \sigma_R^2}{2\alpha \xi^2 \gamma^m \left( \frac{\beta}{\xi \gamma^m} + \frac{1-\alpha}{\gamma^d} \right)^2} \right) S_t^2 \right) \\
&= \frac{(1+r)\xi}{\xi - \phi} \left( (1 - \phi)\xi - \frac{(\xi - \phi)(\xi - \phi + 2)}{2} \right) = \frac{(1+r)\xi(\phi - \frac{\xi^2 + \phi^2}{2})}{\xi - \phi}
\end{aligned}$$

and

$$\begin{aligned}
\sigma_t^{fund} &= \sqrt{\text{Var}_t \left[ \Omega_{t+1}^{fund} \right]} = \frac{1}{X_t^*} (1 - \phi) \sqrt{\text{Var}_t[\Pi_{t+1}]} \\
&= \left( \frac{\gamma^m(1+r) \alpha \xi^3 \left( \frac{\beta}{\xi \gamma^m} + \frac{1-\alpha}{\gamma^d} \right)^2}{\beta(\xi - \phi) \sigma_R^2} \right) \frac{1}{S_t^2} \left( \frac{(1 - \phi) \beta \sigma_R}{\alpha \xi \gamma^m \left( \frac{\beta}{\xi \gamma^m} + \frac{1-\alpha}{\gamma^d} \right)} \right) |S_t| \\
&= \left( \frac{(1+r)\xi^2(1 - \phi) \left( \frac{\beta}{\xi \gamma^m} + \frac{1-\alpha}{\gamma^d} \right)}{(\xi - \phi) \sigma_R} \right) \frac{1}{|S_t|}. \quad \square
\end{aligned}$$

## F Proof of Proposition 6.1

### F.1 Dynamic budget constraints

Each fund investor allocates her wealth  $W_t^f$  between consumption  $c_t^f$ , capital allocation  $X_t$  to the fund, and investment in the riskless asset to maximize her lifetime expected utility. Her dynamic budget constraint is

$$W_{t+1}^f = \Pi_{t+1} - F_t(\Pi_{t+1}) + (1+r)(W_t^f - c_t^f), \quad (\text{F.1})$$

which states that her next-period wealth is the sum of the after-fee fund return and the proceeds from her own riskless investment.

Each fund manager allocates his wealth  $W_t^m$  between consumption  $c_t^m$  and investment in the riskless asset, and chooses the fund's risk exposure  $\theta_t$  to maximize his lifetime expected utility. His dynamic

budget constraint if he continues his fund until period  $t + 1$  is

$$W_{t+1}^m = \frac{\alpha}{\beta} F_t(\Pi_{t+1}) + (1+r)(W_t^m - c_t^m). \quad (\text{F.2})$$

If he absconds in a certain period, his next-period wealth would be

$$\begin{aligned} \widetilde{W}_{t+1}^m &= \frac{\alpha}{\beta} \xi Y_{t+1} + (1+r)(W_t^m - c_t^m) \\ &= \frac{\alpha}{\beta} \xi (\Pi_{t+1} + (1+r)X_t) + (1+r)(W_t^m - c_t^m). \end{aligned} \quad (\text{F.3})$$

After this period onward, he acts as a direct investor; thus, his dynamic budget constraint would be

$$\widetilde{W}_{t+1}^m = R_{t+1} \widetilde{K}_t^m + (1+r)(\widetilde{W}_t^m - \widetilde{c}_t^m), \quad (\text{F.4})$$

where  $\widetilde{K}_t^m$  and  $\widetilde{c}_t^m$  are his stock purchase and consumption, respectively.

Each direct investor allocates her wealth  $W_t^d$  between consumption  $c_t^d$ , purchase of  $K_t$  shares of the stock, and investment in the riskless asset to maximize her lifetime expected utility. Her dynamic budget constraint is

$$W_{t+1}^d = R_{t+1} K_t + (1+r)(W_t^d - c_t^d). \quad (\text{F.5})$$

## F.2 Conjectures

We conjecture and later verify that each agent's value function is, for  $i \in \{f, m, d\}$ ,

$$V^i(W_t^i, S_t) = -\exp\left(-\gamma^i W_t^i - \frac{A^i}{2} S_t^2 - B^i\right), \quad (\text{F.6})$$

where  $A^i$  and  $B^i$  are constants that will be determined later. Moreover, conjecture and later verify that the value function of a manager who has absconded is

$$\widetilde{V}^m(\widetilde{W}_t^m, S_t) = -\exp\left(-\gamma^m \widetilde{W}_t^m - \frac{A^d}{2} S_t^2 - B^d\right). \quad (\text{F.7})$$

Note that (F.7) is characterized by  $A^d$  and  $B^d$  (as opposed to  $A^m$  and  $B^m$ ), reflecting that he acts as a direct investor after absconding. We also conjecture and verify later that the equilibrium stock price is

$$P_t = \frac{\hat{\delta} - \eta S_t}{r}, \quad (\text{F.8})$$

where  $\eta > 0$  is a constant that is determined later. This conjecture implies that, conditional on  $t$ ,  $R_{t+1}$  is normally distributed with mean

$$\hat{R}_t \equiv \mathbf{E}_t[R_{t+1}] = \eta S_t \quad (\text{F.9})$$

and variance  $\sigma_R^2 \equiv \text{Var}_t[R_{t+1}] = \sigma^2 + \eta^2 \sigma_s^2 / r^2$ . From (F.8) and (F.9), we have  $P_t = (\hat{\delta} - \hat{R}_t) / r$ .

### F.3 Fee contract

At the beginning of period 0, the fund investors collectively choose fee schedule  $F(\Pi_{t+1}, \theta_t, X_t, S_t)$  that maximizes the period-0 value of each of them,  $V^f(W_0^f, S_0)$ , subject to the manager's incentive compatibility constraints. Note that  $V^f$  represents the investor's value *given that the manager never absconds in the future*. So, when maximizing  $V^f$ , the investors ensure that the manager will not have an incentive to abscond at all times by considering his incentive compatibility constraint (IC) for each and every  $t \geq 0$ . That is, the optimal schedule  $F$  is the solution to the following problem.

$$\begin{aligned} & \max_F V^f(W_0^f, S_0) \\ \text{sub. to } & E_t[V^m(W_{t+1}^m, S_{t+1})] \geq E_t[\tilde{V}^m(\tilde{W}_{t+1}^m, S_{t+1})] \text{ for all } t \geq 0. \end{aligned} \quad (\text{F.10})$$

The manager's participation constraint,  $E_0[V^m(W_1^m, S_1)] \geq E_0[\tilde{V}^m(0, S_1)]$ , is omitted because it is implied by IC (F.10). Note that (F.10) is characterized by  $E_t[\cdot]$  rather than  $E_0[\cdot]$ , despite the fact that the maximization problem is solved in period 0, because (F.10) is concerning the manager's incentive to continue or abscond as of period  $t$ .

Conjecture and verify later that ICs for all  $t \geq 1$  are redundant, that is, the solution to the problem with IC for only  $t = 0$  satisfies ICs for all  $t \geq 1$ . Moreover, from the fund investor's Bellman equation

$$V^f(W_t^f, S_t) = \max_{c_t^f, X_t} \left\{ u^f(c_t^f) + \rho E_t[V^f(W_{t+1}^f, S_{t+1})] \right\}, \quad (\text{F.11})$$

maximizing  $V^f(W_0^f, S_0)$  amounts to maximizing  $E_0[V^f(W_1^f, S_1)]$ . So the original problem reduces to:

$$\max_F E_0[V^f(W_1^f, S_1)] \quad \text{sub. to } E_0[V^m(W_1^m, S_1)] \geq E_0[\tilde{V}^m(\tilde{W}_1^m, S_1)]. \quad (\text{F.12})$$

The Lagrangian is

$$\begin{aligned} \mathcal{L} = & \int \int V^f \left( \Pi_1 - F_1 + (1+r)(W_0^f - c_0^f), S_1 \right) g(\Pi_1, S_1) d\Pi_1 dS_1 \\ & - \psi \left( E_0[\tilde{V}^m(\tilde{W}_1^m, S_1)] - \int \int V^m \left( \frac{\alpha}{\beta} F_1 + (1+r)(W_0^m - c_0^m), S_1 \right) g(\Pi_1, S_1) d\Pi_1 dS_1 \right), \end{aligned}$$

where  $F_1 \equiv F(\Pi_1, \theta_0, X_0, S_0)$ ,  $g(\cdot, \cdot)$  is the probability density function of  $\Pi_1$  and  $S_1$ , and  $\psi$  is the Lagrange multiplier. Let  $V_A(A, B, C)$  denote the partial derivative of function  $V$  with respect to  $A$ . Then the FOC for  $F_1$  is

$$\begin{aligned} & -V_W^f \left( \Pi_1 - F_1 + (1+r)(W_0^f - c_0^f), S_1 \right) g(\Pi_1, S_1) \\ & + \psi V_W^m \left( \frac{\alpha}{\beta} F_1 + (1+r)(W_0^m - c_0^m), S_1 \right) \frac{\alpha}{\beta} g(\Pi_1, S_1) = 0 \\ \iff & \frac{V_W^m \left( \frac{\alpha}{\beta} F_1 + (1+r)(W_0^m - c_0^m), S_1 \right)}{V_W^f \left( \Pi_1 - F_1 + (1+r)(W_0^f - c_0^f), S_1 \right)} = \frac{\beta}{\alpha} \frac{1}{\psi}. \end{aligned} \quad (\text{F.13})$$

Log-differentiating (F.13) with respect to  $\Pi_1$ ,

$$\begin{aligned} \frac{V_{WW}^m(W_1^m, S_1)}{V_W^m(W_1^m, S_1)} \frac{\alpha}{\beta} F_{\Pi}(\Pi_1, \theta_0, X_0, S_0) - \frac{V_{WW}^f(W_1^f, S_1)}{V_W^f(W_1^f, S_1)} (1 - F_{\Pi}(\Pi_1, \theta_0, X_0, S_0)) &= 0 \\ \iff F_{\Pi}(\Pi_1, \theta_0, X_0, S_0) &= \frac{\beta\gamma^f}{\alpha\gamma^m + \beta\gamma^f} \equiv \phi. \end{aligned} \quad (\text{F.14})$$

(F.14) implies that  $F(\Pi_1, \theta_0, X_0, S_0) = \phi\Pi_1 + Q(\theta_0, X_0, S_0)$  for some function  $Q$ . Thus, for general  $t \geq 0$  we have

$$F(\Pi_{t+1}, \theta_t, X_t, S_t) = \phi\Pi_{t+1} + Q(\theta_t, X_t, S_t). \quad (\text{F.15})$$

Now, let us determine  $Q$ . In period 0, the fund investors set  $Q$  so that the manager's IC binds:

$$\mathbb{E}_0[V^m(W_1^m, S_1)] = \mathbb{E}_0[\tilde{V}^m(\tilde{W}_1^m, S_1)]. \quad (\text{F.16})$$

To solve (F.16) for  $Q$ , first we need to obtain the explicit expressions of the left-hand side (LHS) and RHS of (F.16).

First, let us compute the LHS of (F.16). Conjecture (F.6) implies that (we use an identity  $R_{t+1} = \hat{R}_t + (\delta_{t+1} - \hat{\delta}) + \frac{\eta}{r}S_t - \frac{\eta}{r}S_{t+1}$  and denote  $Q_0 \equiv Q(\theta_0, X_0, S_0)$ )

$$\begin{aligned} V^m(W_1^m, S_1) &= -\exp\left(-\gamma^m W_1^m - \frac{A^m}{2} S_1^2 - B^m\right) \\ &= -\exp\left(-\gamma^m \left(\frac{\alpha}{\beta} (\phi R_1 \theta_0 X_0 + Q_0) + (1+r)(W_0^m - c_0^m)\right) - \frac{A^m}{2} S_1^2 - B^m\right) \\ &= -\exp\left(\underbrace{-\gamma^m \frac{\alpha}{\beta} \left(\phi \hat{R}_0 \theta_0 X_0 + \phi(\delta_1 - \hat{\delta}) \theta_0 X_0 + \phi \frac{\eta}{r} S_0 \theta_0 X_0 + Q_0\right) - \gamma^m (1+r)(W_0^m - c_0^m) - B^m}_{(i)}\right) \end{aligned} \quad (\text{F.17})$$

$$\times \exp\left(\underbrace{\gamma^m \frac{\alpha}{\beta} \phi \frac{\eta}{r} \theta_0 X_0 S_1 - \frac{A^m}{2} S_1^2}_{(ii)}\right). \quad (\text{F.18})$$

In (i) of (F.18), the only random variable is  $\delta_1$ , whereas  $S_1$  is the only random variable in (ii). Since  $\delta_1$  and  $S_1$  are independent, the conditional expectation of  $V^m(W_1^m, S_1)$  is written as a product:  $\mathbb{E}_0[V^m(W_1^m, S_1)] = -\mathbb{E}_0[(i)] \mathbb{E}_0[(ii)]$ . First,  $\mathbb{E}_0[(i)]$  is

$$\mathbb{E}_0[(i)] = \exp\left(-\gamma^m \frac{\alpha}{\beta} \left(\phi \hat{R}_0 \theta_0 X_0 + \phi \frac{\eta}{r} S_0 \theta_0 X_0 + Q_0 - \frac{\gamma^m}{2} \frac{\alpha}{\beta} \phi^2 \theta_0^2 X_0^2 \sigma^2\right) - \gamma^m (1+r)(W_0^m - c_0^m) - B^m\right). \quad (\text{F.19})$$

Second,  $\mathbb{E}_0[(ii)]$  is obtained by using the following fact of normal calculus:

**Fact 1.** *If  $x$  is normally distributed with mean  $\mu$  and variance  $\sigma^2$  then*

$$\mathbb{E}\left[\exp\left(-ax - \frac{b}{2}x^2\right)\right] = \exp\left(-\frac{1}{2}\ln(1+b\sigma^2) - \frac{1}{1+b\sigma^2}\left(a\mu - \frac{a^2}{2}\sigma^2 + \frac{b}{2}\mu^2\right)\right), \quad (\text{F.20})$$

where  $a$  and  $b$  are constants.

Setting  $x = S_1$ ,  $a = -\gamma^m \frac{\alpha}{\beta} \phi \frac{\eta}{r} \theta_0 X_0$ ,  $b = A^m$ ,  $\mu = S_0$  and  $\sigma^2 = \sigma_s^2$  in Fact 1, we have

$$\mathbb{E}_0[(ii)] = \exp\left(-\frac{1}{2} \ln(1 + A^m \sigma_s^2) - \frac{1}{1 + A^m \sigma_s^2} \left(-\gamma^m \frac{\alpha}{\beta} \phi \frac{\eta}{r} \theta_0 X_0 S_0 - \frac{\gamma^{m2}}{2} \frac{\alpha^2}{\beta^2} \phi^2 \frac{\eta^2}{r^2} \theta_0^2 X_0^2 \sigma_s^2 + \frac{A^m}{2} S_0^2\right)\right). \quad (\text{F.21})$$

Thus, from (F.19) and (F.21) we have

$$\begin{aligned} \mathbb{E}_0[V^m(W_1^m, S_1)] &= -\mathbb{E}_0[(i)] \mathbb{E}_0[(ii)] \\ &= -\exp\left(\begin{aligned} &-\gamma^m \frac{\alpha}{\beta} \phi \left(\left(\hat{R}_0 + \frac{\eta}{r} q_1^m S_0\right) \theta_0 X_0 - \frac{\gamma^m}{2} \frac{\alpha}{\beta} \phi q_2^m \theta_0^2 X_0^2\right) - \gamma^m \frac{\alpha}{\beta} Q_0 \\ &+ \frac{1}{2} \ln(1 - q_1^m) - \frac{1}{2} \frac{q_1^m}{\sigma_s^2} S_0^2 - \gamma^m(1+r)(W_0^m - c_0^m) - B^m \end{aligned}\right), \end{aligned} \quad (\text{F.22})$$

where

$$q_1^i \equiv \frac{A^i \sigma_s^2}{1 + A^i \sigma_s^2} \quad \text{for } i \in \{f, m, d\} \quad (\text{F.23})$$

and

$$q_2^i \equiv \sigma^2 + \frac{\frac{\eta^2 \sigma_s^2}{r^2}}{1 + A^i \sigma_s^2} = \sigma^2 + \frac{\eta^2 \sigma_s^2}{r^2} - \frac{\eta^2 \sigma_s^2}{r^2} \frac{A^i \sigma_s^2}{1 + A^i \sigma_s^2} = \sigma_R^2 - \frac{\eta^2 \sigma_s^2 q_1^i}{r^2} \quad \text{for } i \in \{f, m, d\}. \quad (\text{F.24})$$

Next, let us compute the RHS of (F.16),  $\mathbb{E}_0[\tilde{V}^m(\tilde{W}_1^m, S_1)]$ . Conjecture (F.7) implies that

$$\begin{aligned} \tilde{V}^m(\tilde{W}_1^m, S_1) &= -\exp\left(-\gamma^m \tilde{W}_1^m - \frac{A^d}{2} S_1^2 - B^d\right) \\ &= -\exp\left(-\gamma^m \left(\frac{\alpha}{\beta} \xi (R_1 \theta_0 X_0 + (1+r) X_0) + (1+r)(W_0^m - c_0^m)\right) - \frac{A^d}{2} S_1^2 - B^d\right). \end{aligned} \quad (\text{F.25})$$

Following the steps from (F.17) to (F.22) again, we have

$$\begin{aligned} \mathbb{E}_0[\tilde{V}^m(\tilde{W}_1^m, S_1)] &= -\exp\left(\begin{aligned} &-\gamma^m \frac{\alpha}{\beta} \xi \left(\left(\hat{R}_0 + \frac{\eta}{r} q_1^d S_0\right) \theta_0 X_0 - \frac{\gamma^m}{2} \frac{\alpha}{\beta} \xi q_2^d \theta_0^2 X_0^2\right) - \gamma^m \frac{\alpha}{\beta} \xi (1+r) X_0 \\ &+ \frac{1}{2} \ln(1 - q_1^d) - \frac{1}{2} \frac{q_1^d}{\sigma_s^2} S_0^2 - \gamma^m(1+r)(W_0^m - c_0^m) - B^d \end{aligned}\right). \end{aligned} \quad (\text{F.26})$$

Now, plugging (F.22) and (F.26) into (F.16), we have

$$\begin{aligned} &-\gamma^m \frac{\alpha}{\beta} \phi \left(\left(\hat{R}_0 + \frac{\eta}{r} q_1^m S_0\right) \theta_0 X_0 - \frac{\gamma^m}{2} \frac{\alpha}{\beta} \phi q_2^m \theta_0^2 X_0^2\right) - \gamma^m \frac{\alpha}{\beta} Q_0 \\ &+ \frac{1}{2} \ln(1 - q_1^m) - \frac{1}{2} \frac{q_1^m}{\sigma_s^2} S_0^2 - \gamma^m(1+r)(W_0^m - c_0^m) - B^m \\ &= -\gamma^m \frac{\alpha}{\beta} \xi \left(\left(\hat{R}_0 + \frac{\eta}{r} q_1^d S_0\right) \theta_0 X_0 - \frac{\gamma^m}{2} \frac{\alpha}{\beta} \xi q_2^d \theta_0^2 X_0^2\right) - \gamma^m \frac{\alpha}{\beta} \xi (1+r) X_0 \\ &+ \frac{1}{2} \ln(1 - q_1^d) - \frac{1}{2} \frac{q_1^d}{\sigma_s^2} S_0^2 - \gamma^m(1+r)(W_0^m - c_0^m) - B^d. \end{aligned} \quad (\text{F.27})$$

Solving (F.27) for  $Q_0$  and noting that  $\hat{R}_0 = \eta S_0$  (see (F.9)), we have

$$Q_0 = \left( \xi \left( 1 + \frac{q_1^d}{r} \right) - \phi \left( 1 + \frac{q_1^m}{r} \right) \right) \eta S_0 \theta_0 X_0 - \frac{\gamma^m \alpha}{2 \beta} (\xi^2 q_2^d - \phi^2 q_2^m) \theta_0^2 X_0^2 + \xi(1+r)X_0 + \frac{1}{2\gamma^m} \frac{\beta}{\alpha} \left( \left( \frac{q_1^d - q_1^m}{\sigma_s^2} \right) S_0^2 + \ln \left( \frac{1 - q_1^m}{1 - q_1^d} \right) + 2(B^d - B^m) \right). \quad (\text{F.28})$$

Thus, for general  $t \geq 0$  we have

$$Q(\theta_t, X_t, S_t) = \left( \xi \left( 1 + \frac{q_1^d}{r} \right) - \phi \left( 1 + \frac{q_1^m}{r} \right) \right) \eta S_t \theta_t X_t - \frac{\gamma^m \alpha}{2 \beta} (\xi^2 q_2^d - \phi^2 q_2^m) \theta_t^2 X_t^2 + \xi(1+r)X_t + \frac{1}{2\gamma^m} \frac{\beta}{\alpha} \left( \left( \frac{q_1^d - q_1^m}{\sigma_s^2} \right) S_t^2 + \ln \left( \frac{1 - q_1^m}{1 - q_1^d} \right) + 2(B^d - B^m) \right). \quad (\text{F.29})$$

As shown later, we have  $B^m = B^d$  in equilibrium (see (F.60) in Section F.9.1 and (F.67) in Section F.9.3). Thus, the term  $2(B^d - B^m)$  in (F.29) is zero, and we obtain (6.8) of Proposition 6.1.

Last, we can readily verify our conjecture that the fee schedule (F.15) satisfies ICs for all  $t \geq 1$  (with equalities) by plugging (F.15) and (F.29) into those ICs. This is not surprising: since we obtained  $Q_0$  from the binding IC for  $t = 0$ , if we replace  $t = 0$  with an arbitrary  $t$ ,  $Q_t$  should satisfy IC for that period with equality too.

## F.4 Fund manager's optimization

Each fund manager's Bellman equation is

$$V^m(W_t^m, S_t) = \max_{c_t^m, \theta_t} \left\{ u^m(c_t^m) + \rho E_t [V^m(W_{t+1}^m, S_{t+1})] \right\}. \quad (\text{F.30})$$

Plugging (F.15) and (F.29) into the period- $t$  version of (F.22) and rearranging, we have

$$E_t [V^m(W_{t+1}^m, S_{t+1})] = - \exp \left( \begin{array}{l} -\gamma^m \frac{\alpha}{\beta} \xi \left( \left( \hat{R}_t + \frac{\eta}{r} q_1^d S_t \right) \theta_t X_t + (1+r)X_t - \frac{\gamma^m \alpha}{2 \beta} \xi q_2^d \theta_t^2 X_t^2 \right) \\ + \frac{1}{2} \ln(1 - q_1^d) - \frac{1}{2} \frac{q_1^d}{\sigma_s^2} S_t^2 - \gamma^m (1+r)(W_t^m - c_t^m) - B^d \end{array} \right). \quad (\text{F.31})$$

The FOC for  $\theta_t$  is  $\partial E_t [V^m(W_{t+1}^m, S_{t+1})] / \partial \theta_t = 0$ , which is equivalent to

$$\left( \hat{R}_t + \frac{\eta}{r} q_1^d S_t \right) X_t - \gamma^m \frac{\alpha}{\beta} \xi q_2^d \theta_t X_t^2 = 0. \quad (\text{F.32})$$

Therefore, given that there is a mass  $\alpha/\beta$  of fund investors in his fund, each of whom invests  $X_t > 0$  dollars of capital, each manager's optimal choice of risk exposure given  $\hat{R}_t$  is

$$\theta_t(X_t; \hat{R}_t) = \frac{\hat{R}_t + \frac{\eta}{r} q_1^d S_t}{\gamma^m \frac{\alpha}{\beta} \xi q_2^d X_t}. \quad (\text{F.33})$$



## F.5 Fund investor's optimization

Conjecture (F.6) implies that each fund investor's next-period value function is

$$\begin{aligned} V^f(W_{t+1}^f, S_{t+1}) &= -\exp\left(-\gamma^f W_{t+1}^f - \frac{A^f}{2} S_{t+1}^2 - B^f\right) \\ &= -\exp\left(-\gamma^f \left((1-\phi)R_{t+1}\theta_t X_t - Q_t + (1+r)(W_t^f - c_t^f)\right) - \frac{A^f}{2} S_{t+1}^2 - B^f\right). \end{aligned} \quad (\text{F.34})$$

Repeating the steps from (F.17) to (F.22) (for an arbitrary period  $t$  rather than period 0), we have

$$\mathbb{E}_t[V^f(W_{t+1}^f, S_{t+1})] = -\exp\left(\begin{array}{c} -\gamma^f(1-\phi)\left(\left(\hat{R}_t + \frac{\eta}{r}q_1^f S_t\right)\theta_t X_t - \frac{\gamma^f}{2}(1-\phi)q_2^f \theta_t^2 X_t^2\right) + \gamma^f Q_t \\ + \frac{1}{2}\ln(1-q_1^f) - \frac{1}{2}\frac{q_1^f}{\sigma_s^2} S_t^2 - \gamma^f(1+r)(W_t^f - c_t^f) - B^f \end{array}\right). \quad (\text{F.35})$$

Plugging (F.29) into (F.35) and rearranging, we have

$$\mathbb{E}_t[V^f(W_{t+1}^f, S_{t+1})] = -\exp\left(\begin{array}{c} -\gamma^f\left((1-\xi)\hat{R}_t\theta_t X_t + \frac{\eta}{r}q_3 S_t\theta_t X_t - \xi(1+r)X_t - \frac{1}{2}q_4\theta_t^2 X_t^2\right) \\ - \frac{1}{2}z_1 S_t^2 - z_2 - \gamma^f(1+r)(W_t^f - c_t^f) \end{array}\right), \quad (\text{F.36})$$

$$\text{where } q_3 \equiv (1-\phi)q_1^f + \phi q_1^m - \xi q_1^d, \quad (\text{F.37})$$

$$q_4 \equiv \gamma^f(1-\phi)^2 q_2^f + \gamma^m \frac{\alpha}{\beta} (\phi^2 q_2^m - \xi^2 q_2^d), \quad (\text{F.38})$$

$$z_1 \equiv \frac{q_1^f}{\sigma_s^2} - \frac{\gamma^f}{\gamma^m} \frac{\beta}{\alpha} \left(\frac{q_1^d - q_1^m}{\sigma_s^2}\right), \quad \text{and} \quad (\text{F.39})$$

$$z_2 \equiv B^f - \frac{1}{2}\ln(1-q_1^f) - \frac{\gamma^f}{\gamma^m} \frac{\beta}{\alpha} \left(\frac{1}{2}\ln\left(\frac{1-q_1^m}{1-q_1^d}\right) + B^d - B^m\right). \quad (\text{F.40})$$

Assuming positive solution, the FOC for  $X_t$  is  $\partial \mathbb{E}_t[V^f(W_{t+1}^f, S_{t+1})]/\partial X_t = 0$ , i.e.,

$$(1-\xi)\hat{R}_t\theta_t + \frac{\eta}{r}q_3 S_t\theta_t - \xi(1+r) - q_4\theta_t^2 X_t = 0. \quad (\text{F.41})$$

Thus, given  $\theta_t$  and  $\hat{R}_t$ , each fund investor's optimal capital allocation is

$$X_t(\theta_t; \hat{R}_t) = \max\left\{\frac{(1-\xi)\hat{R}_t\theta_t + \frac{\eta}{r}q_3 S_t\theta_t - \xi(1+r)}{q_4\theta_t^2}, 0\right\}. \quad (\text{F.42})$$

## F.6 Direct investor's optimization

The Bellman equation of each direct investor is

$$V^d(W_t^d, S_t) = \max_{c_t^d, K_t} \{u^d(c_t^d) + \rho \mathbb{E}_t[V^d(W_{t+1}^d, S_{t+1})]\}. \quad (\text{F.43})$$

Conjecture (F.6) implies that the next-period value function of each direct investor is

$$\begin{aligned} V^d(W_{t+1}^d, S_{t+1}) &= -\exp\left(-\gamma^d W_{t+1}^d - \frac{A^d}{2} S_{t+1}^2 - B^d\right) \\ &= -\exp\left(-\gamma^d (R_{t+1} K_t + (1+r)(W_t^d - c_t^d)) - \frac{A^d}{2} S_{t+1}^2 - B^d\right). \end{aligned} \quad (\text{F.44})$$

Following the steps from (F.17) to (F.22), we have

$$\mathbb{E}_t [V^d(W_{t+1}^d, S_{t+1})] = -\exp\left(\begin{array}{c} -\gamma^d \left(\left(\hat{R}_t + \frac{\eta}{r} q_1^d S_t\right) K_t - \frac{\gamma^d}{2} q_2^d K_t^2\right) \\ + \frac{1}{2} \ln(1 - q_1^d) - \frac{1}{2} \frac{q_1^d}{\sigma_2^2} S_t^2 - \gamma^d (1+r)(W_t^d - c_t^d) - B^d \end{array}\right). \quad (\text{F.45})$$

The FOC for  $K_t$ ,  $\partial \mathbb{E}_t [V^d(W_{t+1}^d, S_{t+1})] / \partial K_t = 0$ , yields her optimal stock purchase:

$$\hat{R}_t + \frac{\eta}{r} q_1^d S_t - \gamma^d q_2^d K_t = 0 \iff K_t(\hat{R}_t) = \frac{\hat{R}_t + \frac{\eta}{r} q_1^d S_t}{\gamma^d q_2^d}. \quad (\text{F.46})$$

## F.7 Fund's investment profile given $\hat{R}_t$

As in Lemma 4.4, the unique pure-strategy Nash-equilibrium investment profile given  $\hat{R}_t$ ,  $(\theta_t(\hat{R}_t), X_t(\hat{R}_t))$ , is obtained by solving the two best response functions, (F.33) and (F.42), for  $\theta_t$  and  $X_t$ :

$$\theta_t(\hat{R}_t) = \frac{(1+r)\xi^2 \gamma^m \frac{\alpha}{\beta} q_2^d}{\left(\xi(1-\xi)\gamma^m \frac{\alpha}{\beta} q_2^d - q_4\right) \hat{R}_t + \frac{\eta}{r} \left(\gamma^m \frac{\alpha}{\beta} \xi q_2^d q_3 - q_1^d q_4\right) S_t} \quad (\text{F.47})$$

and

$$X_t(\hat{R}_t) = \frac{\left(\hat{R}_t + \frac{\eta}{r} q_1^d S_t\right) \left(\left(\xi(1-\xi)\gamma^m \frac{\alpha}{\beta} q_2^d - q_4\right) \hat{R}_t + \frac{\eta}{r} \left(\gamma^m \frac{\alpha}{\beta} \xi q_2^d q_3 - q_1^d q_4\right) S_t\right)}{(1+r)\xi^3 \gamma^{m2} \frac{\alpha^2}{\beta^2} q_2^{d2}}. \quad (\text{F.48})$$

## F.8 Market clearing

Plugging (F.46), (F.47), and (F.48) into the market clearing condition and solving for  $\hat{R}_t$ , we have

$$\hat{R}_t = \left(\frac{q_2^d}{\frac{\beta}{\gamma^m \xi} + \frac{1-\alpha}{\gamma^d}} - \frac{\eta}{r} q_1^d\right) S_t. \quad (\text{F.49})$$

For (F.9) and (F.49) to be consistent with each other, i.e., for conjecture (F.8) to be correct, we need

$$\eta = \frac{q_2^d}{\frac{\beta}{\gamma^m \xi} + \frac{1-\alpha}{\gamma^d}} - \frac{\eta}{r} q_1^d \iff \eta = \frac{q_2^d}{\left(\frac{\beta}{\gamma^m \xi} + \frac{1-\alpha}{\gamma^d}\right) \left(1 + \frac{q_1^d}{r}\right)}. \quad (\text{F.50})$$

Plugging  $\hat{R}_t = \eta S_t$  into (F.47), (F.48), and (F.46) yields the equilibrium levels of  $\theta_t$ ,  $X_t$ , and  $K_t$ :

$$\theta_t = \left(\frac{(1+r)\xi^2 \gamma^m \frac{\alpha}{\beta} q_2^d}{\left(\gamma^m \frac{\alpha}{\beta} \xi q_2^d \left(1 - \xi + \frac{q_3}{r}\right) - \left(1 + \frac{q_1^d}{r}\right) q_4\right) \eta}\right) \frac{1}{S_t} \equiv \zeta_\theta \frac{1}{S_t}, \quad (\text{F.51})$$

$$X_t = \left( \frac{\left(1 + \frac{q_1^d}{r}\right) \left(\gamma^m \frac{\alpha}{\beta} \xi q_2^d \left(1 - \xi + \frac{q_3}{r}\right) - \left(1 + \frac{q_1^d}{r}\right) q_4\right) \eta^2}{(1+r) \xi^3 \gamma^{m2} \frac{\alpha^2}{\beta^2} q_2^{d2}} \right) S_t^2 \equiv \zeta_X \frac{\beta}{\alpha} S_t^2, \quad \text{and} \quad (\text{F.52})$$

$$K_t = \frac{\eta}{\gamma^d q_2^d} \left(1 + \frac{q_1^d}{r}\right) S_t. \quad (\text{F.53})$$

## F.9 Value functions and consumptions

This section verifies value functions (F.6) and (F.7), and determine the agents' consumption policies.

### F.9.1 Fund managers

Plugging (F.51) and (F.52) into (F.31) and rearranging, we have

$$E_t [V^m(W_{t+1}^m, S_{t+1})] = -\exp \left( \begin{array}{l} -\frac{1}{2} \left( \frac{\eta^2}{q_2^d} \left(1 + \frac{q_1^d}{r}\right)^2 + 2\gamma^m \xi (1+r) \zeta_X + \frac{q_1^d}{\sigma_s^2} \right) S_t^2 \\ + \frac{1}{2} \ln(1 - q_1^d) - \gamma^m (1+r) (W_t^m - c_t^m) - B^d \end{array} \right). \quad (\text{F.54})$$

For notational convenience, define

$$\varphi_t^m \equiv -\ln \rho + \frac{1}{2} \left( \frac{\eta^2}{q_2^d} \left(1 + \frac{q_1^d}{r}\right)^2 + 2\gamma^m \xi (1+r) \zeta_X + \frac{q_1^d}{\sigma_s^2} \right) S_t^2 - \frac{1}{2} \ln(1 - q_1^d) + B^d. \quad (\text{F.55})$$

Then, recalling that  $\gamma^m \equiv \frac{r}{1+r} \tilde{\gamma}^m$ , the Bellman equation is

$$V^m(W_t^m, S_t) = \max_{c_t^m} \left\{ -\exp \left( -\left( \frac{1+r}{r} \right) \gamma^m c_t^m \right) - \exp(-\varphi_t^m - \gamma^m (1+r) (W_t^m - c_t^m)) \right\}. \quad (\text{F.56})$$

The FOC for  $c_t^m$  is

$$\left( \frac{1+r}{r} \right) \gamma^m \exp \left( -\left( \frac{1+r}{r} \right) \gamma^m c_t^m \right) - \exp(-\varphi_t^m) \gamma^m (1+r) \exp(-\gamma^m (1+r) (W_t^m - c_t^m)) = 0,$$

which implies that his optimal consumption is an affine function of wealth  $W_t^m$ :

$$c_t^m = \left( \frac{r}{1+r} \right) W_t^m + \frac{r}{\gamma^m (1+r)^2} (\varphi_t^m - \ln r). \quad (\text{F.57})$$

Plugging (F.57) into (F.56) yields

$$\begin{aligned} & -\exp \left( -\gamma^m W_t^m - \frac{A^m}{2} S_t^2 - B^m \right) = -\exp \left( -\gamma^m W_t^m - \frac{\varphi_t^m}{1+r} \right) r^{-\frac{r}{1+r}} (1+r) \\ \Leftrightarrow & \frac{1}{2} (1+r) A^m S_t^2 + (1+r) B^m \\ & = \frac{1}{2} \left( \frac{\eta^2}{q_2^d} \left(1 + \frac{q_1^d}{r}\right)^2 + 2\gamma^m \xi (1+r) \zeta_X + \frac{q_1^d}{\sigma_s^2} \right) S_t^2 - \ln \frac{\rho (1+r)^{1+r}}{r^r} - \frac{\ln(1 - q_1^d)}{2} + B^d. \end{aligned} \quad (\text{F.58})$$

Comparing the coefficients of  $S_t^2$  and the constant terms in (F.58), we have

$$A^m = \frac{1}{1+r} \left( \frac{\eta^2}{q_2^d} \left( 1 + \frac{q_1^d}{r} \right)^2 + 2\gamma^m \xi (1+r) \zeta_X + \frac{q_1^d}{\sigma_s^2} \right) \quad \text{and} \quad (\text{F.59})$$

$$B^m = \frac{1}{1+r} \left( B^d - \ln \frac{\rho(1+r)^{1+r}}{r^r} - \frac{\ln(1-q_1^d)}{2} \right). \quad (\text{F.60})$$

Once we obtain  $A^d$  ((F.66) in Section F.9.3), we see that (F.59) is (6.7) in the main text. Also, once we get  $B^d$  ((F.67) in Section F.9.3), we see from (F.60) that  $B^m = B^d$ .

## F.9.2 Fund investors

Plugging (F.51) and (F.52) into (F.36) and rearranging, we have

$$\mathbb{E}_t[V^f(W_{t+1}^f, S_{t+1})] = -\exp \left( -\frac{1}{2} \left( \frac{\gamma^f \beta^2 q_4 \eta^2}{\gamma^{m2} \alpha^2 \xi^2 q_2^{d2}} \left( 1 + \frac{q_1^d}{r} \right)^2 + z_1 \right) S_t^2 - z_2 - \gamma^f (1+r)(W_t^f - c_t^f) \right). \quad (\text{F.61})$$

Following the steps from (F.54) to (F.60), we obtain the fund investor's optimal consumption policy as (F.57) with  $m$  replaced with  $f$ , and the equations for  $A^f$  and  $B^f$  as

$$A^f = \frac{1}{1+r} \left( \frac{\gamma^f \beta^2 q_4 \eta^2}{\gamma^{m2} \alpha^2 \xi^2 q_2^{d2}} \left( 1 + \frac{q_1^d}{r} \right)^2 + z_1 \right) \quad \text{and} \quad (\text{F.62})$$

$$B^f = \frac{1}{1+r} \left( z_2 - \ln \frac{\rho(1+r)^{1+r}}{r^r} \right). \quad (\text{F.63})$$

Plugging (F.39) into (F.62) yields (6.6). Plugging (F.40) with  $B^d = B^m$  into (F.63), we have

$$B^f = \frac{1}{r} \left( -\ln \frac{\rho(1+r)^{1+r}}{r^r} - \frac{1}{2} \ln(1-q_1^f) - \frac{1}{2} \frac{\gamma^f \beta}{\gamma^m \alpha} \ln \left( \frac{1-q_1^m}{1-q_1^d} \right) \right). \quad (\text{F.64})$$

## F.9.3 Direct investors

Plugging (F.53) into (F.45) and rearranging, we have

$$\mathbb{E}_t[V^d(W_{t+1}^d, S_{t+1})] = -\exp \left( \begin{aligned} & -\frac{1}{2} \left( \frac{\eta^2}{q_2^d} \left( 1 + \frac{q_1^d}{r} \right)^2 + \frac{q_1^d}{\sigma_s^2} \right) S_t^2 \\ & + \frac{1}{2} \ln(1-q_1^d) - \gamma^d (1+r)(W_t^d - c_t^d) - B^d \end{aligned} \right). \quad (\text{F.65})$$

Following the steps from (F.54) to (F.60), we obtain the direct investor's optimal consumption policy as (F.57) with  $m$  replaced with  $d$ , and the equations for  $A^d$  and  $B^d$  as

$$A^d = \frac{1}{1+r} \left( \frac{\eta^2}{q_2^d} \left( 1 + \frac{q_1^d}{r} \right)^2 + \frac{q_1^d}{\sigma_s^2} \right) \quad \text{and} \quad (\text{F.66})$$

$$B^d = \frac{1}{r} \left( -\ln \frac{\rho(1+r)^{1+r}}{r^r} - \frac{\ln(1-q_1^d)}{2} \right). \quad (\text{F.67})$$

### F.9.4 Fund managers who have absconded

The manager who has absconded acts as a direct investor. So his optimization problem is identical to that of Section F.6 except that the risk-aversion coefficient is  $\gamma^m$  rather than  $\gamma^d$ . Likewise, his value function (F.7) is verified as in Section F.9.3, replacing  $\gamma^d$  with  $\gamma^m$ .

### F.10 Solving the system of equations

We can obtain closed-form solutions for  $(\eta, q_1^d, q_2^d, A^d)$  by solving the system of (F.50), (F.23) for  $i = d$ , (F.24) for  $i = d$ , and (F.66). Plugging  $\eta = \sigma^2/\pi$  into (F.52), (F.51), and (F.62) yields (6.1), (6.9), and (6.6), respectively.

## G Robustness

### G.1 Endogenous riskless rate

The model of Section 2 assumes an exogenous, constant riskless rate  $r$ . Although such an assumption is standard in the literature, it is worth investigating what happens to our result if  $r$  is endogenized.

Let us modify the model of Section 2 as follows. Each period, a one-period riskless bond is traded in the market. The bond is in zero net supply, and each unit of the bond delivers one dollar in the next period. Let  $q_t$  denote the (endogenous) period- $t$  bond price, and define the riskless rate by  $r_t \equiv 1/q_t - 1$ . We focus on the case of nonstochastic supply:  $S_t = S > 0$  for all  $t$ . We have not been able to solve the case with stochastic supply (we will discuss the challenge later).

The model is solved in a similar fashion as Section 4. Conjecture and later verify that  $r_t$  (and thus  $q_t$ ) is constant over time in equilibrium, and that the equilibrium stock price is constant,

$$P_t = P \equiv \frac{\hat{\delta} - \eta S}{r}, \quad (\text{G.1})$$

for all  $t$ , where  $\eta$  is a constant that is determined later. These conjectures imply that the stock's excess return  $R_{t+1} \equiv \delta_{t+1} + P_{t+1} - (1+r)P_t$  is, conditional on  $t$ , normally distributed with mean

$$\hat{R} = \eta S \quad (\text{G.2})$$

and variance  $\sigma^2$ . Taking  $\hat{R}$  and  $r$  as given, all the arguments in Sections 4.1–4.4 go through. The stock's market clearing yields, as in (4.8),

$$\hat{R} = \frac{\sigma^2}{\frac{\beta}{\xi\gamma^m} + \frac{1-\alpha}{\gamma^d}} S. \quad (\text{G.3})$$

For our conjectures to be correct, we need (G.2) and (G.3) to be consistent with each other, i.e.,

$$\eta = \frac{\sigma^2}{\frac{\beta}{\xi\gamma^m} + \frac{1-\alpha}{\gamma^d}}. \quad (\text{G.4})$$

The bond's market clearing condition is

$$\underbrace{\alpha \left( \frac{W - X_t}{q} \right)}_{\text{fund investors' demand}} + \underbrace{\alpha \left( \frac{X_t - P_t \theta_t X_t}{q} \right)}_{\text{funds' demand}} + \underbrace{(1 - \alpha) \left( \frac{W - P_t K_t}{q} \right)}_{\text{direct investors' demand}} = 0 \quad (\text{G.5})$$

$$\iff W - P_t (\alpha \theta_t X_t + (1 - \alpha) K_t) = 0. \quad (\text{G.6})$$

Plugging  $\alpha \theta_t X_t + (1 - \alpha) K_t = S$ , (G.1), and (G.4) into (G.6) yields the equilibrium riskless rate:

$$r = \left( \hat{\delta} - \frac{\sigma^2 S}{\frac{\beta}{\xi \gamma^m} + \frac{1 - \alpha}{\gamma^d}} \right) \frac{S}{W}. \quad (\text{G.7})$$

We can choose  $\hat{\delta}$  large enough to ensure that  $r > 0$ .

Do the results presented in Section 5 carry through under endogenous riskless rate? Since  $d\hat{R}/d\alpha > 0$  and  $d\hat{R}/d\beta < 0$ , Result 1 still holds in this model. However, since we have not been able to solve the case with endogenous riskless rate and stochastic stock supply  $S_t$ , we are unable to check the robustness of the other results that rely on the assumption that  $S_t$  is stochastic. The challenge is the following. If  $S_t$  is stochastic, the one-period riskless rate  $r_t$  would also be stochastic, reflecting the agents' asset demands that depend on  $S_t$ . This means that the price conjecture of the form (G.1) would not work out. Intuitively, it would be sensible to conjecture that the equilibrium price is of the form  $P_t = \text{E}_t \left[ \sum_{\nu=t}^{\infty} \prod_{\tau=t}^{\nu} \left( \frac{\delta_{\nu+1}}{1+r_{\tau}} \right) \right] - Z$ , where the first term is the expected discounted future payoffs and  $Z$  is the risk premium. However, the whole analysis is very complicated with such a price function. We leave this problem for future research.

How does the extent of portfolio delegation as measured by  $(\alpha, \beta)$  affect  $r$ ? From (G.7), we have  $dr/d\alpha < 0$  and  $dr/d\beta > 0$ . The intuition is as follows. Given  $r$ , an increase in  $\alpha$  and/or a decrease in  $\beta$  increase  $\hat{R}$  (see Result 1 for the intuition), lowering the stock price  $P = (\hat{\delta} - \hat{R})/r$ . This decreases the stock market capitalization,  $PS$ , and therefore more investor capital flows into the bond market, that is,  $(W - PS)$  increases. Thus, to clear the bond market, that is, to make the market less attractive for investors,  $r$  decreases.

## G.2 Stock supply process

In Section 2, we model the stock supply process as a random walk. While this specification simplifies the analysis, it is more restrictive than needed. All of our results go through qualitatively under more general AR(1) process

$$S_t = a_0 + a_1 S_{t-1} + \epsilon_t, \quad (\text{G.8})$$

where  $a_0 \geq 0$  and  $a_1 \in [0, 1]$  are constants and  $\epsilon_t$  is normal with mean 0 and variance  $\sigma_s^2 > 0$ . Note that (G.8) nests not only the one of Section 2 ( $a_0 = 0$  and  $a_1 = 1$ ) but also, for example, an i.i.d. process ( $a_0 > 0$  and  $a_1 = 0$ ) and a mean-reverting process ( $a_0 = \psi \bar{S}$  and  $a_1 = 1 - \psi$  where  $\bar{S} = \text{E}[S_t]$  and  $\psi \in (0, 1)$ ).

The equilibrium under (G.8) is determined as follows. Conjecture and later verify that the equilibrium stock price is of the form

$$P_t = \frac{\hat{\delta} - \eta(S_t + \frac{a_0}{r})}{r}, \quad (\text{G.9})$$

where  $\eta$  is a constant that is determined later. This conjecture implies that the stock's excess return  $R_{t+1}$  is, conditional on  $t$ , normally distributed with mean

$$\hat{R}_t = \left(1 + \frac{1 - a_1}{r}\right) \eta S_t \quad (\text{G.10})$$

and variance  $\sigma_R^2 \equiv \sigma^2 + \eta^2 \sigma_s^2 / r^2$ . Taking  $\hat{R}_t$  as given, all the argument in Sections 4.1–4.4 carries through. Market clearing yields, as in (4.8),

$$\hat{R}_t = \frac{\sigma_R^2}{\frac{\beta}{\xi \gamma^m} + \frac{1 - \alpha}{\gamma^d}} S_t. \quad (\text{G.11})$$

For (G.10) and (G.11) to be consistent with each other, we need

$$\left(1 + \frac{1 - a_1}{r}\right) \eta = \frac{\sigma_R^2}{\frac{\beta}{\xi \gamma^m} + \frac{1 - \alpha}{\gamma^d}}. \quad (\text{G.12})$$

Solving (G.12) for  $\eta$ , in the stable equilibrium we have

$$\eta = \frac{r^2}{2\sigma_s^2} \left( \left(1 + \frac{1 - a_1}{r}\right) \left(\frac{\beta}{\xi \gamma^m} + \frac{1 - \alpha}{\gamma^d}\right) - \sqrt{\left(1 + \frac{1 - a_1}{r}\right)^2 \left(\frac{\beta}{\xi \gamma^m} + \frac{1 - \alpha}{\gamma^d}\right)^2 - \frac{4\sigma^2 \sigma_s^2}{r^2}} \right). \quad (\text{G.13})$$

Since  $\hat{R}_t$  is proportional to  $S_t$  and the proportionality constant  $\left(1 + \frac{1 - a_1}{r}\right) \eta$  is increasing in  $\alpha$  and decreasing in  $\beta$ , all the results presented in Sections 4 and 5 go through qualitatively. In the main text, we focus on the random walk case because it keeps the analysis very simple, and assuming  $a_0 > 0$  and  $a_1 \neq 1$  does not yield much additional economic insights.

### G.3 Fund investors' choices of funds and funds' competition

This section shows that each fund has a mass  $\alpha/\beta$  of fund investors in equilibrium if each investor can choose a fund to which to allocate capital. This result is not obvious because, a priori, we cannot exclude the possibility of strategic complementarity, that is, each fund investor's expected utility may increase with the mass of investors in the fund. If that is the case, the equilibrium outcome would be that all fund investors allocate capital to a single fund. In the following, we show that it is not the case: each fund investor's expected utility decreases with the mass of fund investors in that fund, and thus the investors will optimally flow in and out of funds until every fund accommodates the same mass of investors. We also show that fund managers are “perfectly competitive” in attracting investors, in the sense that each manager takes the mass of fund investors choosing his fund as given, as he cannot strategically influence the investors' fund choices through his actions.

Suppose that each fund investor chooses a fund at the beginning of each period. Once all investors have chosen funds, each fund determines the fee schedule  $F_t(\cdot)$  and investment profile  $(\theta_t, X_t)$ , as in the model of Section 2. Let  $N$  be the mass of fund investors who choose an arbitrary fund. In this fund, each investor's next-period wealth is given by (2.5), while that of the manager is  $W_t^m = NF_t(\Pi_{t+1})$  if he continues the fund and  $\widetilde{W}_t^m = \xi NY_{t+1}$  if he absconds. The optimal fee schedule is obtained by replacing

$\alpha/\beta$  with  $N$  in Lemma 4.1:

$$F_t(\Pi_{t+1}) = \phi_N \Pi_{t+1} + Q_{N,t}, \quad (\text{G.14})$$

where  $\phi_N \equiv \gamma^f / (N\gamma^m + \gamma^f)$  and

$$Q_{N,t} = (\xi - \phi_N) \text{E}_t[\Pi_{t+1}] + \xi(1+r)X_t - \frac{\gamma^m}{2} N(\xi^2 - \phi_N^2) \text{Var}_t[\Pi_{t+1}]. \quad (\text{G.15})$$

This fund's investment profile given  $\hat{R}_t$  is the one in Lemma 4.4 with  $\phi$  and  $\alpha/\beta$  replaced by  $\phi_N$  and  $N$ , respectively:

$$\theta(\hat{R}_t) = \left( \frac{(1+r)\xi^2}{\xi - \phi_N} \right) \frac{1}{\hat{R}_t} \quad \text{and} \quad X(\hat{R}_t) = \left( \frac{\xi - \phi_N}{(1+r)\gamma^m N \xi^3 \sigma_R^2} \right) \hat{R}_t^2. \quad (\text{G.16})$$

Using (G.16), the certainty equivalent of each fund investor's next-period wealth (given  $\hat{R}_t$ ) is

$$\begin{aligned} U_t^f &= \text{E}_t[W_{t+1}^f] - \frac{\gamma^f}{2} \text{Var}_t[W_{t+1}^f] = (1-\xi)\hat{R}_t\theta_t X_t - \xi(1+r)X_t + (1+r)W - \frac{\gamma^m}{2} N(\phi - \xi^2)\theta_t^2 X_t^2 \sigma_R^2 \\ &= \frac{(\phi_N - \xi^2)\hat{R}_t^2}{2N\gamma^m \xi^2 \sigma_R^2} + (1+r)W. \end{aligned} \quad (\text{G.17})$$

Since  $dU_t^f/dN < 0$ , the investors are worse off as the mass of investors in the fund increases. So, in equilibrium,  $N$  must adjust so that  $U_t^f$  is equalized across all fund investors, i.e.,  $N = \alpha/\beta$  for all funds. The intuition is the standard arbitrage argument: if there is a mass  $N_i$  of investors in fund  $i$  who achieve higher  $U_t^f$  than those in fund  $j$  with mass  $N_j > N_i$ , then investors flee fund  $j$  (i.e.,  $N_j$  decreases) and flock to fund  $i$  (i.e.,  $N_i$  increases) until  $N_i$  and  $N_j$  are equalized.

Likewise, using (G.16), the certainty equivalent of the manager's next-period wealth (given  $\hat{R}_t$ ) is

$$\begin{aligned} U_t^m &= \text{E}_t[W_{t+1}^m] - \frac{\gamma^m}{2} \text{Var}_t[W_{t+1}^m] = N\xi\hat{R}_t\theta_t X_t + N\xi(1+r)X_t - \frac{\gamma^m}{2} N^2 \xi^2 \theta_t^2 X_t^2 \sigma_R^2 \\ &= \left( \frac{1}{2} + \frac{\xi - \phi_N}{\xi^2} \right) \frac{\hat{R}_t^2}{\gamma^m \sigma_R^2}. \end{aligned} \quad (\text{G.18})$$

Since  $dU_t^m/dN > 0$ , the manager is better off as more investors choose his fund. However, in our setting, he is unable to attract investors by strategically influencing their fund choices. The reasons are twofold. First, he cannot increase  $N$  by offering a fee schedule  $F_t(\cdot)$ , as the investors will set  $F_t(\cdot)$  that is best for them anyway. Second, at the time the investors choose funds, each manager cannot credibly commit to choose  $\theta_t$  in the interest of the investors: once the investors have been allocated to funds, it will be optimal for the manager to renege on the promise he made and strategically choose  $\theta_t$  as in Section 4.

## G.4 Active funds and passive funds

In the model of Section 2, each manager chooses  $\theta_t$  strategically to maximize his own expected utility. Consequently, the fund invests in the riskless asset despite the fact that the investors can buy it on their own. In reality, however, there are a number of passively managed funds, such as index funds, in which the managers do not have discretion over portfolio choice: the investors effectively specify the number of shares of the risky asset to hold, and typically do not allow the managers to buy the riskless asset.



This section shows that our model can accommodate such funds too. We extend the model of Section 2 to have both “active funds,” in which the managers choose  $\theta_t$  strategically as in Section 2, and “passive funds,” in which each investor specifies the number of shares of (only) the stock to buy and the manager passively fulfills it. In both types of funds, the managers can abscond as in Section 2. Each fund investor can choose a fund to which to allocate capital.

In equilibrium, both types of funds coexist in the sense that they all attract positive mass of investors. This result may sound counterintuitive, as it looks like the passive funds dominate the active ones from the investors’ perspective. However, in our setting, the passive funds do not necessarily dominate the active ones for the following two reasons. First, even the passive funds are subject to the same agency problem that the managers can abscond. So, just like in the active funds, the passive funds’ investors also incur agency costs by offering fee contracts to the managers to prevent them from absconding. Second, although the investors in the passive funds are better off than those in the active ones when these funds accommodate the same mass of investors, the investors’ expected utilities change as the allocation of investors across funds change. Specifically, as shown below, as the mass of investors in a fund increases, the investors in that fund are worse off. So, in equilibrium, the investors flow in and out of funds until they become indifferent between any funds.

### G.4.1 Setup

Section 2’s model is modified as follows. A mass  $\omega\beta$  of fund managers run active funds, and a mass  $(1 - \omega)\beta$  of them run passive funds, where  $\omega \in [0, 1]$  is an exogenous parameter. The letters  $a$  and  $p$  are used to represent active funds and passive funds, respectively. Each active fund’s manager can choose  $\theta_t$  strategically as in Section 2. Each passive fund’s manager buys  $X/P_t$  shares of the stock given  $X$  dollars of investor capital and does not buy the riskless asset. That is, from an investor’s perspective, allocating  $X$  dollars of capital to a passive fund is equivalent to asking the manager to buy  $X/P_t$  shares of the stock on her behalf. Each fund investor can choose a fund to which to allocate capital. In equilibrium, a mass  $\lambda\alpha$  of fund investors invest in the active funds and a mass  $(1 - \lambda)\alpha$  of them invest in the passive funds, where  $\lambda \in [0, 1]$  is endogenously determined. In both types of funds, the investors offer fee contracts that depend on the fund profits to the managers who can abscond with a fraction  $\xi$  of the fund assets. It is not important that both types of funds share the same  $\xi$ . The result would be very similar if we assume different  $\xi$  for each type.

### G.4.2 Price conjecture

We conjecture and later verify that the equilibrium stock price is of the form

$$P_t = \ell\hat{\delta} - \frac{\eta}{r}S_t, \tag{G.19}$$

where  $\ell$  and  $\eta$  are constants that are determined later. Given this conjecture, we have

$$\hat{R}_t \equiv E_t[R_{t+1}] = (1 - r\ell)\hat{\delta} + \eta S_t \tag{G.20}$$

and  $\text{Var}_t[R_{t+1}] = \sigma_R^2 \equiv \sigma^2 + \eta^2\sigma_s^2/r^2$ .

### G.4.3 Fee contracts

There is a mass  $\frac{\lambda\alpha}{\omega\beta}$  of fund investors in each active fund, and  $\frac{(1-\lambda)\alpha}{(1-\omega)\beta}$  in each passive fund. Thus, following the same argument of Lemma 4.1, the fee schedule for each active fund is

$$F_t^a(\Pi_{t+1}^a) = \phi^a \Pi_{t+1}^a + Q_t^a, \text{ where } \phi^a \equiv \frac{\omega\beta\gamma^f}{\lambda\alpha\gamma^m + \omega\beta\gamma^f}, \Pi_{t+1}^a \equiv R_{t+1}\theta_t X_t^a, \text{ and}$$

$$Q_t^a = (\xi - \phi^a)E_t[\Pi_{t+1}^a] + \xi(1+r)X_t^a - \frac{\gamma^m \lambda\alpha}{2\omega\beta}(\xi^2 - \phi^{a2})\text{Var}[\Pi_{t+1}^a], \quad (\text{G.21})$$

and the one for each passive fund is

$$F_t^p(\Pi_{t+1}^p) = \phi^p \Pi_{t+1}^p + Q_t^p, \text{ where } \phi^p \equiv \frac{(1-\omega)\beta\gamma^f}{(1-\lambda)\alpha\gamma^m + (1-\omega)\beta\gamma^f}, \Pi_{t+1}^p \equiv R_{t+1}\frac{X_t^p}{P_t}, \text{ and}$$

$$Q_t^p = (\xi - \phi^p)E_t[\Pi_{t+1}^p] + \xi(1+r)X_t^p - \frac{\gamma^m (1-\lambda)\alpha}{2(1-\omega)\beta}(\xi^2 - \phi^{p2})\text{Var}[\Pi_{t+1}^p]. \quad (\text{G.22})$$

### G.4.4 Optimization problems

Before choosing  $X_t$ , each fund investor chooses a fund. For now, let us take this choice of fund as given, and determine the optimal  $X_t$ . We consider her choice of fund later (in Section G.4.6), after characterizing her expected utilities. Following Lemma 4.2, each active-fund investor's optimal capital allocation is

$$X^a(\theta_t; \hat{R}_t) = \max \left\{ \frac{(1-\xi)\hat{R}_t\theta_t - \xi(1+r)}{\gamma^m \frac{\lambda\alpha}{\omega\beta}(\phi^a - \xi^2)\sigma_R^2\theta_t^2}, 0 \right\}. \quad (\text{G.23})$$

Noting that  $\theta_t = 1/P_t \forall t$  in passive funds, each passive-fund investor's optimal capital allocation is

$$X^p(P_t; \hat{R}_t) = \max \left\{ \frac{(1-\xi)\frac{\hat{R}_t}{P_t} - \xi(1+r)}{\gamma^m \frac{(1-\lambda)\alpha}{(1-\omega)\beta}(\phi^p - \xi^2)\sigma_R^2 \frac{1}{P_t^2}}, 0 \right\}. \quad (\text{G.24})$$

As in Lemma 4.3, given  $X_t^a > 0$ , each active-fund manager chooses

$$\theta(X_t^a; \hat{R}_t) = \frac{\omega\beta\hat{R}_t}{\gamma^m \lambda\alpha\xi\sigma_R^2 X_t^a}. \quad (\text{G.25})$$

From (G.23) and (G.25), each active fund's Nash-equilibrium investment profile given  $\hat{R}_t$  is given by

$$\theta(\hat{R}_t) = \left( \frac{(1+r)\xi^2}{\xi - \phi^a} \right) \frac{1}{\hat{R}_t} \quad \text{and} \quad X^a(\hat{R}_t) = \left( \frac{\omega\beta(\xi - \phi^a)}{(1+r)\gamma^m \lambda\alpha\xi^3\sigma_R^2} \right) \hat{R}_t^2. \quad (\text{G.26})$$

Each direct investor's optimal policy is the same as (3.3).

### G.4.5 Market clearing and belief consistency

The market clearing condition for the stock is

$$\underbrace{\lambda\alpha\theta(\hat{R}_t)X^a(\hat{R}_t)}_{\text{active funds' demand}} + \underbrace{(1-\lambda)\alpha\frac{X^p(\hat{R}_t)}{P_t}}_{\text{passive funds' demand}} + \underbrace{(1-\alpha)K(\hat{R}_t)}_{\text{direct investors' demand}} = S_t \quad (\text{G.27})$$

$$\iff \hat{R}_t = \frac{k_2}{k_1}\ell\hat{\delta} + \frac{1}{k_1}\left(\sigma_R^2 - k_2\frac{\eta}{r}\right)S_t, \quad (\text{G.28})$$

where

$$k_1 \equiv \frac{\omega\beta}{\gamma^m\xi} + \frac{(1-\omega)\beta(1-\xi)}{\gamma^m(\phi^p - \xi^2)} + \frac{1-\alpha}{\gamma^d} \quad \text{and} \quad k_2 \equiv \frac{(1-\omega)\beta\xi(1+r)}{\gamma^m(\phi^p - \xi^2)}. \quad (\text{G.29})$$

For (G.20) and (G.28) to be consistent with each other, we need

$$1 - r\ell = \frac{k_2}{k_1}\ell \iff \ell = \frac{1}{r + \frac{k_2}{k_1}} \quad (\text{G.30})$$

and

$$\eta = \frac{1}{k_1}\left(\sigma_R^2 - k_2\frac{\eta}{r}\right) \iff \eta = \frac{r^2}{2\sigma_s^2} \left( \left(k_1 + \frac{k_2}{r}\right) \pm \sqrt{\left(k_1 + \frac{k_2}{r}\right)^2 - \frac{4\sigma^2\sigma_s^2}{r^2}} \right). \quad (\text{G.31})$$

As discussed in Section 3.3, only the smaller root of (G.31) is relevant in the stable equilibrium. Since  $\hat{R}_t$  and  $P_t$  are affine functions of  $\eta S_t$ , the stock price dynamics are similar to those of the model of Section 2 with the value of  $\eta$  being replaced by (G.31).

### G.4.6 Fund investor's choice of fund

If a fund investor chooses an active fund, the certainty equivalent of her terminal wealth would be

$$U_t^{af} \equiv E_t[W_{t+1}^{af}] - \frac{\gamma^f}{2}\text{Var}_t[W_{t+1}^{af}] = \frac{\left(\frac{\omega\beta\gamma^f}{\lambda\alpha\gamma^m + \omega\beta\gamma^f} - \xi^2\right)\omega\beta}{2\gamma^m\lambda\alpha\xi^2\sigma_R^2}\hat{R}_t^2 + (1+r)W. \quad (\text{G.32})$$

Here, we take  $\hat{R}_t$  and  $P_t$  as given because each investor takes them as given when choosing a fund. If she chooses a passive fund, the certainty equivalent would be

$$\begin{aligned} U_t^{pf} &\equiv E_t[W_{t+1}^{pf}] - \frac{\gamma^f}{2}\text{Var}_t[W_{t+1}^{pf}] \\ &= \frac{(1-\omega)\beta}{2\gamma^m(1-\lambda)\alpha\left(\frac{(1-\omega)\beta\gamma^f}{(1-\lambda)\alpha\gamma^m + (1-\omega)\beta\gamma^f} - \xi^2\right)\sigma_R^2} \left( (1-\xi)\hat{R}_t - \xi(1+r)P_t \right)^2 + (1+r)W. \end{aligned} \quad (\text{G.33})$$

As long as  $U_t^{af} > (<)U_t^{pf}$ , investors flock to active (passive) funds and  $\lambda$  increases (decreases). The adjustment of  $\lambda$  continues until every investor is indifferent between all funds, i.e., until  $U_t^{af} = U_t^{pf}$  holds. Such a critical level of  $\lambda$  exists in the interval  $(0, 1)$  because  $\lim_{\lambda \rightarrow 0} U_t^{af} = \infty$  and  $\lim_{\lambda \rightarrow 1} U_t^{af} < \infty$  while  $\lim_{\lambda \rightarrow 0} U_t^{pf} < \infty$  and  $\lim_{\lambda \rightarrow 1} U_t^{pf} = \infty$ .

The intuition for this result is the following. If  $\lambda \rightarrow 0$ , each active fund's manager takes capital from

investors of measure 0. From this manager's perspective, the investment risk is so small that he is willing to bear all of it, that is,  $\phi^a \rightarrow 1$  as  $\lambda \rightarrow 0$ . If  $\phi^a \rightarrow 1$ , we have  $Q_t^a < 0$  (i.e., the investor *receives*  $Q_t^a$  from the manager); thus, the fund is risk-free from the investor's perspective. So she chooses  $X_t^a \rightarrow \infty$ , leading to  $U_t^{af} \rightarrow \infty$ . This induces the investors to choose active funds, so  $\lambda$  increases. This decreases  $\phi^a$ , increasing each investor's risk bearing and decreasing  $U_t^{af}$ . Likewise, if  $\lambda \rightarrow 1$ , the passive funds are risk-free for investors so  $U_t^{pf} \rightarrow \infty$ ; but  $U_t^{pf}$  decreases as  $\lambda$  decreases. At a critical level of  $\lambda \in (0, 1)$ ,  $U_t^{af} = U_t^{pf}$  holds and the investors are indifferent between any funds.

## G.5 Allowing fund managers to invest capital

This section shows that fund managers would not invest their own capital in their funds even if they were allowed to do so.

Suppose that each generation- $t$  manager is born with an endowment of  $W \geq 0$  dollars, and chooses to invest  $Z_t \geq 0$  dollars (per investor) in his fund. That is, given  $X_t$ , he chooses  $\theta_t \in (-\infty, \infty)$  and  $Z_t \geq 0$  to buy  $\theta_t(X_t + Z_t)$  shares of the stock per investor. The fund's profit per investor is  $\Pi_{t+1} \equiv Y_{t+1} - (1+r)X_t$ . Recalling that there is a mass  $\alpha/\beta$  of investors in his fund, the manager's terminal wealth is

$$W_{t+1}^m = \frac{\alpha}{\beta} F_t(\Pi_{t+1}) + (1+r) \left( W - \frac{\alpha}{\beta} Z_t \right), \quad (\text{G.34})$$

where the first term on the RHS is the total fees, and the second term is the proceeds from the riskless asset. His terminal wealth if he absconds would be

$$\widetilde{W}_{t+1}^m = \frac{\alpha}{\beta} \xi Y_{t+1}. \quad (\text{G.35})$$

The fund investor's and the direct investor's budget constraints are the same as those of Section 2. The optimal fee contract is still obtained in the form of Lemma 4.1.

Now, let us consider a manager's problem. Under fee schedule  $F_t(\cdot)$  of Lemma 4.1, his terminal wealth (G.34) is rewritten as

$$W_{t+1}^m = \frac{\alpha}{\beta} \left( \phi \Pi_{t+1} + (\xi - \phi) \text{E}_t[\Pi_{t+1}] + \xi(1+r)X_t - \frac{\gamma^m}{2} \frac{\alpha}{\beta} (\xi^2 - \phi^2) \text{Var}_t[\Pi_{t+1}] - (1+r)Z_t \right) + (1+r)W,$$

which implies that the mean and variance of  $W_{t+1}^m$  are

$$\text{E}_t [W_{t+1}^m] = \frac{\alpha}{\beta} \left( \xi \text{E}_t[\Pi_{t+1}] + \xi(1+r)X_t - \frac{\gamma^m}{2} \frac{\alpha}{\beta} (\xi^2 - \phi^2) \text{Var}_t[\Pi_{t+1}] - (1+r)Z_t \right) + (1+r)W \quad (\text{G.36})$$

and

$$\text{Var}_t [W_{t+1}^m] = \frac{\alpha^2}{\beta^2} \phi^2 \text{Var}_t [\Pi_{t+1}]. \quad (\text{G.37})$$

Since the manager has exponential utility, (G.36) and (G.37) imply that his maximization problem is

$$\max_{\theta_t, Z_t} \text{E}_t [u^m(W_{t+1}^m)] \iff \max_{\theta_t, Z_t} \text{E}_t [W_{t+1}^m] - \frac{\gamma^m}{2} \text{Var}_t [W_{t+1}^m]$$

$$\iff \max_{\theta_t, Z_t} \xi \hat{R}_t \theta_t (X_t + Z_t) - (1+r)(1-\xi)Z_t + \xi(1+r)X_t - \frac{\gamma^m \alpha}{2\beta} \xi^2 \theta_t^2 (X_t + Z_t)^2 \sigma_R^2.$$

The FOC for  $\theta_t$  is

$$\hat{R}_t - \gamma^m \frac{\alpha}{\beta} \xi \theta_t (X_t + Z_t) \sigma_R^2 = 0, \quad (\text{G.38})$$

and the one for  $Z_t$  is

$$\xi \hat{R}_t \theta_t - (1+r)(1-\xi) - \gamma^m \frac{\alpha}{\beta} \xi^2 \theta_t^2 (X_t + Z_t) \sigma_R^2 \leq 0, \text{ with equality iff } Z_t > 0. \quad (\text{G.39})$$

(G.38) and (G.39) imply that the LHS of (G.39) is  $-(1+r)(1-\xi)$ , which is negative. So we have  $Z_t = 0$ . That is, the manager would not invest his own capital in his fund even if he was allowed to do so.

The result holds because a manager can control the fund's stock holding,  $\theta_t(X_t + Z_t)$ , by choosing two variables:  $\theta_t$  and  $Z_t$ . While choosing  $Z_t > 0$  entails an opportunity cost (i.e., he could invest it in the riskless asset and earn  $1+r$ ), choosing  $\theta_t > 0$  is costless. Thus, it is optimal for him to set  $Z_t = 0$  and use only  $\theta_t$  to control the stock holding.

## G.6 Fee contract space

In the model of Section 2, we assume that the fee  $F_t$  depends only on the fund profit,  $\Pi_{t+1}$ . This assumption is stronger than necessary: all the results (including the fee schedule) would be identical to those of Section 4 even if we allow for a larger contract space that includes not only  $\Pi_{t+1}$  but also the stock return  $R_{t+1}$  and the fund portfolio's total proceeds  $Y_{t+1}$ .

To see this, let us allow the fee to depend on  $R_{t+1}$ ,  $\Pi_{t+1}$ , and  $Y_{t+1}$ , and solve for the optimal fee contract. The maximization problem is the same as the one presented in Appendix A. Since  $R_{t+1}$  is the only random variable in  $R_{t+1}$ ,  $\Pi_{t+1}$ , and  $Y_{t+1}$ , the Lagrangian is (we omit time scripts)

$$\mathcal{L} = \int u^f (\Pi - F(R, \Pi, Y) + (1+r)W) h(R) dR - \psi \left( \mathbb{E} \left[ u^m (\widetilde{W}^m) \right] - \int u^m \left( \frac{\alpha}{\beta} F(R, \Pi, Y) \right) h(R) dR \right),$$

where  $h(\cdot)$  is the probability density function (conditional on  $t$ ) of  $R_{t+1}$ , and  $\psi$  is the Lagrange multiplier. The FOC for  $F_t(R_{t+1}, \Pi_{t+1}, Y_{t+1})$  is

$$\frac{u^{m'} \left( \frac{\alpha}{\beta} F(R, \Pi, Y) \right)}{u^{f'} (\Pi - F(R, \Pi, Y) + (1+r)W)} = \frac{\beta}{\alpha} \frac{1}{\psi}. \quad (\text{G.40})$$

Log-differentiating (G.40) with respect to  $R_{t+1}$ ,  $\Pi_{t+1}$ , and  $Y_{t+1}$ , we have

$$-\gamma^m \frac{\alpha}{\beta} (F_R + F_\Pi \theta X + F_Y \theta X) + \gamma^f (\theta X - F_R - F_\Pi \theta X - F_Y \theta X) = 0, \quad (\text{G.41})$$

$$-\gamma^m \frac{\alpha}{\beta} (F_\Pi + F_Y) + \gamma^f (1 - F_\Pi - F_Y) = 0, \quad (\text{G.42})$$

and

$$-\gamma^m \frac{\alpha}{\beta} F_Y + \gamma^f (1 - F_Y) = 0, \quad (\text{G.43})$$

respectively, where  $F_j$  is the partial derivative of  $F(R, \Pi, Y)$  with respect to  $j \in \{R, \Pi, Y\}$ . From (G.41),

(G.42), and (G.43), we have  $F_R = 0$ ,  $F_\Pi = 0$ , and  $F_Y = \phi$ , which imply that the optimal fee schedule is

$$F_t(R_{t+1}, \Pi_{t+1}, Y_{t+1}) = \phi Y_{t+1} + Q_t^Y \quad (\text{G.44})$$

for some  $Q_t^Y$  that is known as of period  $t$ . The value of  $Q_t^Y$  satisfies the manager's incentive compatibility constraint (2.7) with equality:

$$\begin{aligned} \mathbb{E}_t [u^m(W_{t+1}^m)] &= \mathbb{E}_t [u^m(\widetilde{W}_{t+1}^m)] \\ \iff \mathbb{E}_t [W_{t+1}^m] - \frac{\gamma^m}{2} \text{Var}_t [W_{t+1}^m] &= \mathbb{E}_t [\widetilde{W}_{t+1}^m] - \frac{\gamma^m}{2} \text{Var}_t [\widetilde{W}_{t+1}^m] \\ \iff \frac{\alpha}{\beta} (\phi \mathbb{E}_t [Y_{t+1}] + Q_t^Y) - \frac{\gamma^m}{2} \frac{\alpha^2}{\beta^2} \phi^2 \text{Var}_t [Y_{t+1}] &= \frac{\alpha}{\beta} \xi \mathbb{E}_t [Y_{t+1}] - \frac{\gamma^m}{2} \frac{\alpha^2}{\beta^2} \xi^2 \text{Var}_t [Y_{t+1}] \\ \iff Q_t^Y &= (\xi - \phi) \mathbb{E}_t [Y_{t+1}] - \frac{\gamma^m}{2} \frac{\alpha}{\beta} (\xi^2 - \phi^2) \text{Var}_t [Y_{t+1}], \end{aligned} \quad (\text{G.45})$$

which is different from  $Q_t$  obtained in (4.2). However, plugging (G.45) into (G.44), we end up with the same fee schedule as Lemma 4.1:

$$\begin{aligned} &F_t(R_{t+1}, \Pi_{t+1}, Y_{t+1}) \\ &= \phi Y_{t+1} + \left( (\xi - \phi) \mathbb{E}_t [Y_{t+1}] - \frac{\gamma^m}{2} \frac{\alpha}{\beta} (\xi^2 - \phi^2) \text{Var}_t [Y_{t+1}] \right) \\ &= \phi (\Pi_{t+1} + (1+r)X_t) + \left( (\xi - \phi) \mathbb{E}_t [\Pi_{t+1} + (1+r)X_t] - \frac{\gamma^m}{2} \frac{\alpha}{\beta} (\xi^2 - \phi^2) \text{Var}_t [\Pi_{t+1}] \right) \\ &= \phi \Pi_{t+1} + \left( (\xi - \phi) \mathbb{E}_t [\Pi_{t+1}] + \xi(1+r)X_t - \frac{\gamma^m}{2} \frac{\alpha}{\beta} (\xi^2 - \phi^2) \text{Var}_t [\Pi_{t+1}] \right) \\ &= \phi \Pi_{t+1} + Q_t. \end{aligned} \quad (\text{G.46})$$

Following similar steps, it is easy to show that the fee contracts depending on any one or two of  $(R_{t+1}, \Pi_{t+1}, Y_{t+1})$  would also result in (G.46). For example, if we allow  $F_t$  to depend on only  $R_{t+1}$ , then the optimal proportional fee rate would be  $\phi \theta_t X_t$  and the fixed fee is  $Q_t$ , so we have  $F_t(R_{t+1}) = (\phi \theta_t X_t) R_{t+1} + Q_t = \phi \Pi_{t+1} + Q_t$ , which is identical to (G.46).

## G.7 Allowing managers to modify $\theta_t$ when absconding

In the model of Section 6, an implicit assumption is that each manager who absconds in period  $t$  is not allowed to change the risk exposure of that period,  $\theta_t$ . This section shows that the results would not change even if we allow him to modify  $\theta_t$  when absconding.

Let us assume that each manager can abscond with a fraction  $\xi$  of the fund portfolio and can change that portfolio's risk exposure from  $\theta_t$  to  $\theta_t + \Delta_t$ , where  $\Delta_t \in (-\infty, \infty)$  is his choice variable. That is, the next-period wealth of an absconding manager, i.e., the counterpart of (F.3), is

$$\widetilde{W}_{t+1}^m = R_{t+1}(\theta_t + \Delta_t) \frac{\alpha}{\beta} \xi X_t + (1+r) \frac{\alpha}{\beta} \xi X_t + (1+r)(W_t^m - c_t^m). \quad (\text{G.47})$$

Then, following the steps of Appendix F, we have

$$\begin{aligned}
\mathbb{E}_t [V^m(W_{t+1}^m, S_{t+1})] &= \mathbb{E}_t [\tilde{V}^m(\tilde{W}_{t+1}^m, S_{t+1})] \\
&= -\exp \left( \begin{array}{c} -\gamma^m \frac{\alpha}{\beta} \xi \left( \left( \hat{R}_t + \frac{\eta}{r} q_1^d S_t \right) (\theta_t + \Delta_t) X_t - \frac{\gamma^m}{2} \frac{\alpha}{\beta} \xi q_2^d (\theta_t + \Delta_t)^2 X_t^2 \right) \\ -\gamma^m \frac{\alpha}{\beta} \xi (1+r) X_t + \frac{1}{2} \ln(1 - q_1^d) - \frac{1}{2} \frac{q_1^d}{\sigma_s^2} S_t^2 - \gamma^m (1+r) (W_t^m - c_t^m) - B^d \end{array} \right). \quad (\text{G.48})
\end{aligned}$$

On the one hand, the manager chooses  $\theta_t$  to maximize  $\mathbb{E}_t [V^m(W_{t+1}^m, S_{t+1})]$ , which is (G.48). On the other hand, if he absconds, he chooses  $\Delta_t$  to maximize  $\mathbb{E}_t [\tilde{V}^m(\tilde{W}_{t+1}^m, S_{t+1})]$ , which is also (G.48) due to the binding IC. But since  $\theta_t$  and  $\Delta_t$  appear in (G.48) only in the form of their sum,  $\theta_t + \Delta_t$ , the FOCs for  $\theta_t$  and  $\Delta_t$  are the same. Thus, the optimal level of the sum  $\theta_t + \Delta_t$  is determined but its composition is indeterminate. So  $\Delta_t = 0$ , which we assume in Section 6, is an equilibrium outcome, though not a unique one. It is easy to show that if the manager incurs a (very small) adjustment cost for choosing  $\Delta_t > 0$ , the unique outcome would be that he chooses  $\Delta_t = 0$  and the same  $\theta_t$  as in Section 6. The economic intuition is the following. Since the IC is binding, each manager chooses  $\theta_t$  as if he is maximizing the outside option that he would obtain if absconding. Thus, he has no reason to change such a  $\theta_t$  when he actually absconds.