Mutual Fund Portfolio Choice in the Presence of Dynamic Flows*

Julien Hugonnier
Department of Mathematical Sciences
Carnegie Mellon University
Pittsburgh PA 15213
jnh@andrew.cmu.edu

Ron Kaniel
McCombs School of Business
University of Texas at Austin
Austin, TX 78712
ron.kaniel@bus.utexas.edu

This revision: April 2002

Preliminary and Incomplete

Abstract

We analyze the implications of a fixed fraction of assets under management fee, the most commonly used fee by mutual funds, on portfolio choice decisions in a continuous time model. In our model, the investor has a log utility function and is allowed to dynamically allocate his capital between an actively managed mutual fund and a money market account. The optimal fund portfolio is shown to be the one that maximizes the market values of the fees received, and is independent of the manager's utility function. The presence of dynamic flows induces "flow hedging" on the part of the fund, even though the investor is log. We predict a positive relationship between a fund's proportional fee rate and a fund's volatility. This is a consequence of higher fee funds holding more extreme equity positions. However, the overall dollar amount of equity held by a fund can be independent of the fee rate, as a higher fee also implies that investors allocate a smaller fraction of their wealth to the fund. While both the fund portfolio and the investor's trading strategy depend on the proportional fee, the equilibrium value functions do not. Finally, we show that our results hold even if in addition to trading the fund and the money market account the investor is allowed to directly trade some of the risky securities, but not all.

JEL Classification Numbers. G11, G12, G13, G23, D81.
Key words. Delegated Portfolio Management, Mutual Funds, Incomplete Markets, Continuous Time Stochastic Game, Backward Stochastic Differential Equations .

*We thank Pierre Collin Dufresne, Mike Gallmeyer, Lane Hughston, Monique Jeanblanc, Hong Liu, Patrice Poncet, Steve Shreve and Stan Zin, seminar participants at Carnegie-Mellon, CIRANO, Université Paris−1, University of Texas at Austin and Washington University at St Louis. Financial support from the National Science Foundation under grant DMS·9819950 (for Hugonnier) is gratefully acknowledged.

1 Introduction

When analyzing the impact of fee structures on the fund's portfolio, most of the previous literature has assumed that the investor is not able to trade dynamically. Instead, the investor turns all his money over to the manager at the initial date. The manager then makes the investment decisions and receives a fee in exchange for his management services. Some examples include Grinblatt and Titman (1989) and Carpenter (2000), who analyze the case where the manager receives convex performance fees. High water mark fees are discussed in Goetzmann, Ingersoll and Ross (1997). In these papers, as well as in many others, the investor is not only prohibited from trading dynamically, but is not even allowed the flexibility of deciding how much funds to allocate initially to the manager. Some exceptions include Cuoco & Kaniel (2001) and Das & Sundaram (1998), who while still restricting the investor from trading dynamically, do allow the investor to determine, at the initial date, what fraction of his or her funds to allocate to the manager and what fraction to hold directly in a riskless asset.

The innovation of our paper is that we allow the investor to move money in and out of the fund dynamically. To this end, we study a dynamic continuous time economy with two agents: a *small investor* and a *fund manager*. The small investor implicitly faces high transaction costs that preclude him from trading directly in the equity market. While he is precluded from holding equity directly, he can invest money in a mutual fund. Thus, the small investor dynamically allocates his money between the mutual fund and the money market account; where in addition we impose the natural restriction that the investor can not short the fund. The fund manager, on the other hand, is allowed to trade dynamically in both the stocks and the money market. For his management services he receives an instantaneous fee that is a fixed proportion of the assets under management; referred to as a fraction of funds fee. We allow the manager to select the fund portfolio strategy and at the same time to trade on his own account.

Our objective is to understand how enabling an investor to dynamically rebalance between a fund and a money market account impacts the dynamic trading strategy of the fund. We focus our analysis on the case of fraction of funds fees. As of 1995, approximately 98% of all mutual funds were using fraction of fund fees without any performance based incentives. It is important to emphasize that we are not taking a stance on whether this is the form of the optimal contract in our setting, but instead rely on its widespread use by mutual funds as the motivation for our analysis.

In order to focus on the impact of the flows to the fund on the manager's choice of the fund portfolio we make the following simplifying assumptions. First, both agents have complete information about the financial markets. Second, agents observe the actions of each other. Third, from the perspective of the fund manager securities markets are complete. Forth, the small investor has a log utility function.

While both agents observe the actions of each other, the fund manager is strategic where as the investor is not. Specifically, when the investor determines his portfolio he takes the fund strategy as given. On the other hand, when the fund manager selects the fund's trading strategy he takes into account what would be the investor's reaction to such a strategy. Thus, the manager can be viewed as a Stackelberg leader, and the investor as a Stackelberg follower, in a stochastic differential game.

When the interest rate and coefficients of the price processes are constant (or more generally deterministic), the investment opportunity set is stationary and the fund's optimal portfolio is constant. In this case we obtain the intuitive result that the equilibrium fund portfolio is that which maximizes the after fees Sharpe ratio of the fund. In equilibrium, the equity component of the fund's portfolio is proportional to the instantaneous fee rate, and the investors portfolio weight in the fund is proportional to 1/(instantaneous fee rate). So that increasing the instantaneous fee rate tilts the fund portfolio towards the risky assets, and at the same time decreases the small investors holdings in the fund. Furthermore, while both the funds portfolio and the investors portfolio depend on the instantaneous fee rate the agents' value functions do not.

When the investment opportunity set is stochastic it is no longer the case that the equilibrium fund portfolio is the one that maximizes the after fees Sharpe ratio of the fund. The presence of dynamic flows induces "flow hedging" on the part of the fund, even though the investor is log. These hedging demands impact the bond to stock mix of the fund portfolio as well as induce distortions within the equity component. Interestingly, the distortions within the equity component do not depend on the fee rate. Even with a stochastic opportunity set, it is still the case that the equilibrium value functions of the two agents are independent of the instantaneous fee rate.

We use our model to investigate a potential contemporaneous relationship between flows and fund returns. For high fee rates: high fund returns will be accompanied by outflows whereas low fund returns will be accompanied by inflows. A reverse relationship holds for low fee rates. In the empirical literature on the relationship between flows and lagged fund returns (see for example Chevalier and Ellison (1997) or Sirri and Tufano (1998)), the potential relationship between measured flows and contemporaneous returns has been ignored. Controlling for this contemporaneous relationship may help shed some light on the empirically observed relationship between flow and lagged fund returns.

In solving our model we rely on tools that were developed for the study of utility maximization problems under incomplete markets as well as on the theory of Backward Stochastic Differential Equations and its links to the study of optimal control problems. For details on the study of utility maximization problems in incomplete/constrained financial markets see for example Karatzas et al. (1991), He & Pearson (1991), Shreve & Xu (1992) and Cvitanič & Karatzas (1992). For a comprehensive survey of backward stochastic equations and some of their applications in finance, see El Karoui, Peng & Quenez (1997) and the research monograph edited by El Karoui & Mazliak (1997). A brief introduction to the theory of such equations as well as a review of the results needed in the paper is given in Appendix B.

The remainder of the paper is organized as follows. The economic setup is described in Section 2. In Section 3 we solve the investor's utility maximization problem and derive the optimal fund portfolio as well as the manager's optimal trading strategy. Section 4 describes the equilibrium trading strategies and discusses the models implications. Section 5 relaxes the assumption that the investor can get exposure to equity only through holding the fund. Section 6 concludes. Appendix A contains all the proofs and Appendix B contains a short review of the necessary results from the theory of Backward Stochastic Differential Equations.

2 The Economy

We consider a continuous-time economy on the finite time span $[0, T]$. There is a single perishable consumption good which serves as the numéraire. The uncertainty is represented by a filtered probability space $(\Omega, \mathbb{F}, \mathcal{F}, P)$ on which is defined an $n-$dimensional standard Brownian motion B. The filtration $\mathbb{F} \equiv \left\{\mathcal{F}(t) : t \in [0, T]\right\}$ is the usual $P-$augmentation of the filtration generated by the Brownian motion and we assume that $\mathcal{F} = \mathcal{F}(T)$. In the sequel, all stochastic processes are assumed to be adapted to the filtration $\mathbb{F}$ and all statements involving random variables and/or stochastic processes are understood to hold either $P-$almost surely or $\lambda \times P-$almost everywhere (here λ denotes Lebesgue measure) depending on the context.

2.1 Securities and Mutual Fund Dynamics

Securities. The financial market consists of $n+1$ long lived securities. The first is The first security S_0 is a locally riskless *bond* satisfying

$$dS_0(t) = r(t)S_0(t)dt, \qquad S_0(0) \equiv 1 \tag{1}$$

for some *interest rate* process r. The remaining n securities, the *stocks*, follow the process

$$dS(t) = \operatorname{diag}\big(S(t)\big)\Big(a(t)dt + \sigma(t)dB(t)\Big), \qquad S(0) \equiv S \in (0, \infty)^n \tag{2}$$

for some $(n \times 1)-$ *drift* process a and $(n \times n)-$ *volatility* process σ.[1]

The conditions imposed on the model coefficients imply that

$$Z_0(t) \equiv \exp\left(-\int_0^t \xi(\tau)^* dB(\tau) - \int_0^t \|\xi(\tau)\|^2 \, d\tau\right)$$

is a strictly positive and uniformly integrable martingale, where

$$\xi \equiv \sigma^{-1}[a - r\mathbf{1}]$$

denotes the *risk premium* process, $\mathbf{1}$ denotes a unit vector of dimension n, and $\|\cdot\|$ denotes the Euclidean norm. As a result, the formula $P^0(A) = \mathbf{E}1_A Z_0(T)$ defines a probability measure which is equivalent to the original probability measure and under which the process

$$B_0(t) \equiv \left(B(t) + \int_0^t \xi(\tau)d\tau\right)$$

is an $n-$dimensional standard Brownian motion by Girsanov theorem. Under this *equivalent martingale measure* P^0 stock dynamics are given by

$$dS(t) = \operatorname{diag}\big(S(t)\big)\Big(r(t)\mathbf{1}\, dt + \sigma(t)dB_0(t)\Big),$$

[1] The model coefficients (r, a and σ) are adapted bounded processes, and the volatility process σ admits a uniformly bounded inverse. Thus, the *risk premium* process $\xi := \sigma^{-1}[a - r\mathbf{1}]$, where $\mathbf{1}$ denotes a unit vector of dimension n, is uniformly bounded.

so that the discounted stock price process S/S_0 is a martingale.

Mutual Fund Dynamics. Trading takes place continuously and there are no market frictions. A trading strategy is an $\mathbb{R}^n-$valued process θ specifying the amount invested in each of the n stocks, so that the remainder is invested in the bond. We denote by Θ the set of admissible trading strategies.[2]

Given a fund trading strategy $\theta \in \Theta$, the value of the fund for one dollar invested is the solution the linear stochastic equation

$$dF(t) = r(t)F(t)dt + \theta(t)^*\sigma(t)\Big(dB(t) + \xi(t)dt\Big) \tag{3}$$

with initial condition $F(0) = 1$. Let now Θ_+ denote the set of admissible trading strategies θ for which the fund value is non negative. For every such trading strategy, the *portfolio weight process* $\pi := \theta/F$ is well defined and (3) becomes

$$dF_\pi(t) = F_\pi(t)\Big(r(t)dt + \pi(t)^*\sigma(t)dB_0(t)\Big) \tag{4}$$

where B_0 is the process defined by (2.1) and we write $F \equiv F_\pi$ to emphasize the dependence of the fund value on the portfolio weight process chosen by the manager. We denote by Θ_f the set of admissible fund portfolio weight processes.

2.2 Agents

Investor. The first agent is an investor, which we denote by i, and who, while allowed to invest freely in the bond, does not have access to the market for risky assets directly but only through the fund. We assume moreover that even though he is allowed to change his position in the fund continuously through time, he is prohibited from short selling the fund.

The inability of the investor to trade stocks directly should be viewed as a reduced form representing the fact that it is more costly for the investor to trade efficiently stocks than it is for the fund. This could be due, for example, to the fact that given that the investor is small it will be costly for him to hold a portfolio that consists of many stocks. Furthermore, it is implicitly assumed that the opportunity cost for the investor of spending his time in stock trading related activities is high. In the last section of the paper we relax this assumption by allowing the investor to trade directly in a subset of the stocks.

For the portfolio management services provided by the fund, the investor pays an instantaneous fee that is proportional to the amount of money he currently holds in the fund. Specifically, at each time t the investor pays a flow of $\gamma \cdot \phi(t)$ where $\gamma \in (0, 1)$ is a given *fee rate* and $\phi(t)$ represents the dollar amount he currently holds in the fund. The cumulative fee process, which in turn determines the manager's cumulative income process, is thus given by

$$\Phi(t) \equiv \int_0^t \Big(\gamma \cdot \phi(\tau)\Big) d\tau.$$

[2]Our notion of admissibility is standard and the corresponding regularity conditions can be found, e.g., in Karatzas and Shreve (1998).

Given a fund portfolio process $\pi \in \Theta_f$, the investor's trading strategy is a non negative process ϕ that specifies the amount invested in the fund, so that $W_i - \phi$ is invested in the bond. The corresponding wealth process W_i satisfies the dynamic budget constraint

$$\begin{aligned} dW_i(t) &= \Big(W_i(t) - \phi(t)\Big)\frac{dS_0(t)}{S_0(t)} + \phi(t)\frac{dF_\pi(t)}{F_\pi(t)} - \gamma\phi(t)dt \\ &= r(t)W_i(t)dt + \phi(t)\pi(t)^*\sigma(t)\Big(dB(t) + \xi_\pi(t)dt\Big), \end{aligned} \tag{5}$$

where the investor has an initial wealth of $W_i(0)$, and the process

$$\xi_\pi(t) \equiv \xi(t) - \gamma\frac{\sigma(t)^*\pi(t)}{\|\sigma(t)^*\pi(t)\|^2}.$$

is the risk premium associated with an indirect investment in stocks through the fund.

Since the investor's sole resource is his initial wealth of $W_i(0)$, he can avoid bankruptcy only if his wealth remains positive throughout the investment period. We denote by Θ_i^π the set of admissible trading strategies for the investor, such that $\phi \in \Theta_i^\pi$ if $\phi > 0$, $\phi\pi \in \Theta$, and $W_i(t) \geq 0$ for all $t \in [0, T]$.

The investor is assumed to derive utility $\log\big(W_i(T)\big)$ over terminal wealth, taking the fund portfolio process π as given. Thus, he selects a trading strategy ϕ_i^π that maximizes his expected utility

$$\mathbf{E}\log\big(W_i(T)\big)$$

of terminal wealth. The fact that the fund process is driven by n independent Brownian motions and that the investor is not allowed to short the fund imply that the investor effectively faces an incomplete financial market. Thus, the assumption of log utility is made to allow a closed form solution in our possibly non Markovian framework.

Fund Manager. The second agent, which we denote by m, is a fund manager who chooses not only the fund's trading strategy π, but also his own trading strategy θ_m^π. He is endowed with an initial wealth of $W_m(0)$, and in addition receives from the investor the cumulative income process Φ for the portfolio management services that he provides. Since the manager is assumed to be the Stackelberg leader, in determining the fund's portfolio he takes into account that the fund's portfolio π impacts the investor's investment decision and as a results determines the income process he receives. To make the dependence explicit we denote henceforth the income process as Φ^π.

His corresponding wealth process W_m satisfies the dynamic budget constraint

$$\begin{aligned} dW_m(t) &= \Big(W_m(t) - \theta_m(t)^*\mathbf{1}\Big)\frac{dS_0(t)}{S_0(t)} + \sum_{k=1}^{n}\theta_{m,k}(t)\frac{dS_k(t)}{S_k(t)} + d\Phi^\pi(t) \\ &= r(t)W(t)dt + \theta_m(t)^*\sigma(t)\Big(dB(t) + \xi(t)dt\Big) + d\Phi^\pi(t) \end{aligned} \tag{6}$$

with initial condition $W_m(0)$, and he maximizes the expected utility

$$\mathbf{E}U_m\big(W_m(T)\big)$$

of terminal wealth. The function $U_m(\cdot)$ is assumed to be strictly increasing, strictly concave, continuously differentiable and to satisfy the Inada condition at both zero and infinity.

2.3 Equilibrium

An equilibrium is defined by a tuple of dynamic trading strategies $\{\pi,\, \theta_m^\pi,\, \phi_i^\pi\} \in \Theta_f \times \Theta \times \Theta_i^\pi$ such that,

1. the strategy ϕ_i^π is optimal for the investor given the fund trading strategy π : that is,

$$\phi_i^\pi \in \arg\left\{\begin{array}{c} \max\limits_{\phi\in\Theta_i^\pi} \mathbf{E}\Big[\log\Big(W_i(T)\Big)\Big] \\ \text{s.t. } W_i(0) \text{ given} \end{array}\right\};$$

2. the combination of the fund trading strategy π and the manager's trading strategy in his own personal account θ_m^π are optimal for the fund manager given the investor's response to the fund's trading strategy: that is,

$$\{(\pi,\, \theta_m^\pi)\} \in \arg\left\{\begin{array}{c} \max\limits_{(\pi,\theta)\in\Theta_f\times\Theta} \mathbf{E}\Big[U_m\Big(W_m(T))\Big)\Big] \\ \text{s.t. } W_m(0) \text{ given} \\ \qquad\qquad \phi_i^\pi \text{ optimal for investor} \end{array}\right\};$$

The fact that the manager is a Stackelberg leader suggests (see Bagchi (1984)) that the dominating (in the sense of being both individually and collectively stable) equilibrium can be computed through the following sequential procedure.

The first step consists in solving the investor's utility maximization problem given an arbitrary fund portfolio process π. Thus, the investors trading strategy ϕ_i^π can be viewed as his *best response* to the fund's trading strategy π. This optimal strategy induces via (2.2) a cumulative fee process Φ^π, that is payed to the manager.

The second step then consists in solving the manager's utility maximization problem given the cumulative fee process Φ^π in order to obtain his optimal trading strategy θ_m^π as well as his value function

$$V_m^\pi \equiv \max_{\theta\in\Theta} \mathbf{E}\Big[U_m\Big(W_m(T)\Big) \mid W_m(0),\, \Phi^\pi\Big]. \tag{7}$$

Finally, the third step consists in finding the optimal fund portfolio from the manager's perspective: that is

$$\hat{\pi} \in \arg\max_{\pi\in\Theta_f} V_m^\pi. \tag{8}$$

Re-injecting this fund portfolio process $\hat{\pi}$ into the first two steps we obtain, by construction, that the tuple $(\hat{\pi}, \theta_m^{\hat{\pi}}, \phi_i^{\hat{\pi}})$ constitutes a Nash equilibrium and that the corresponding allocations are Pareto optimal. Therefore, this non cooperative equilibrium is dominating.

3 Optimal Trading Strategies

In solving for the optimal trading strategies we follow the sequential procedure described in the previous section. The first step is carried in Section 3.1, and the second and third in Section 3.2.

3.1 The Investor's Problem

In this section we assume that the fund portfolio is a fixed process $\pi \in \Theta_f$. Given the dynamics (4), the net of fees instantaneous Sharpe ratio of the fund is given by

$$\eta_\pi(t) \equiv \frac{(r + \pi(t)^* \sigma(t)\xi(t)) - \gamma - r}{\|\sigma(t)^* \pi(t)\|^2} = \frac{\pi(t)^* \sigma(t)\xi(t) - \gamma}{\|\sigma(t)^* \pi(t)\|^2} \tag{9}$$

Combining the fact that the investor's utility function is log with the fact that he can not short the fund, it is natural to conjecture that at any point in time the investor will hold the fund only if it has a positive net of fees Sharpe ratio, and that in those cases the proportion of his wealth that will be invested in the fund would equal that Sharpe ratio. The following proposition shows that this is indeed the case.

Proposition 1 *Let $\pi \in \Theta_f$ be a given fund portfolio and define a non negative process by setting*

$$\phi_i^\pi(t) \equiv \eta_\pi(t)^+ W_i^\pi(t) = \max\{0, \eta_\pi(t)\} W_i^\pi(t)$$

where W_i^π is the corresponding solution to (5). *Then $\phi_i^\pi \in \Theta_i^\pi$ constitutes the investor best response trading strategy given the fund portfolio process.*

Proof. The derivation of this result relies on the direct verification of the optimality of the trading strategy ϕ_i^π and is carried out in the Appendix. □

3.2 The Manager's Problem

We start (Section 3.2.1) by assuming that the fund portfolio is a fixed process and solve the manager's utility maximization problem (7), given the investor's best response strategy. Then (Section 3.2.2), we complete our sequential optimization procedure by determining the optimal fund portfolio process.

3.2.1 Optimal Trading Strategy

Proposition 1 shows that for a given fund portfolio process, the investor's best response trading strategy generates the cumulative fee process Φ^π defined by

$$\Phi^\pi(t) \equiv \int_0^t \Big(\gamma \cdot \phi_i^\pi(\tau)\Big) d\tau = \int_0^t \Big(\gamma \cdot W_i^\pi(\tau) \eta_\pi(\tau)^+\Big) d\tau$$

where W_i^π is the investor's optimal wealth process.

Because the stock price volatility is invertible, the manager faces a complete financial market. In such a setting the solution to the utility maximization problem (7) is well-known (see Karatzas & Shreve (1999, Chapter 3)). In order to describe it, we introduce the discount factor process $D := 1/S_0$ as well as the *market value process*

$$v_\pi(t) \equiv \mathbf{E}_t^0 \int_t^T \left(\frac{D(\tau)}{D(t)} d\Phi^\pi(\tau)\right) = \mathbf{E}_t \int_t^T \frac{Z_0(\tau)}{Z_0(t)} \left(\frac{D(\tau)}{D(t)} d\Phi^\pi(\tau)\right).$$

of the cumulative fee process generated by the investor's best response strategy.[3] Standard results imply that given Φ^π the manager's problem amounts to maximizing expected utility over the set of positive random variables $W_m(T)$ which satisfy the static budget constraint

$$\mathbf{E}^0\Big(D(T)W_m(T)\Big) = \mathbf{E}\Big(D(T)Z_0(T)W_m(T)\Big) \leq v_\pi(0) + W_m(0). \tag{10}$$

Let now I_m denote the (strictly positive and strictly decreasing) inverse of the manager's marginal utility function U'_m. Since the manager's utility function is strictly concave, the optimal terminal wealth for the manager's utility maximization problem (7)can be written as

$$W_m^\pi(T) \equiv I_m\Big(\lambda^\pi \cdot D(T)Z_0(T)\Big)$$

where λ^π is a strictly positive Lagrange multiplier chosen such that the budget constraint (10) holds as an equality. Let now $\mathcal{X}_m : (0,\infty) \to \mathbb{R}_+$ be the strictly decreasing function defined by

$$\mathcal{X}_m(\lambda^\pi) \equiv \mathbf{E}^0\left[D(T)I_m\Big(\lambda^\pi D(T)Z_0(T)\Big)\right].$$

The following makes use of the Brownian martingale representation theorem (Karatzas & Shreve (1988, Theorem 3.4.2 p.170)) in order to obtain an explicit characterization of the solution to the manager's problem (7).

Proposition 2 *Let $\pi \in \Theta_f$ be a given fund portfolio process, and assume that the function $\mathcal{X}_m$ is real valued.*[4] *Set $\lambda^\pi \equiv \mathcal{X}_m^{-1}(v_\pi(0) + W_m(0))$ in (3.2.1) and define*

$$\begin{aligned} D(t)W_m^\pi(t) &\equiv -v_\pi(t)D(t) + \mathbf{E}_t^0\Big(D(T)W_m^\pi(T)\Big) \\ &\equiv W_m(0) + \int_0^t \Big(h(\tau)^* dB_0(\tau) + D(\tau)d\Phi^\pi(\tau)\Big). \end{aligned}$$

Then the process W_m^π satisfies the manager's dynamic budget constraint (6) with the admissible trading strategy $\theta \equiv \theta_m^\pi := (\sigma^*)^{-1}h/D \in \Theta$ and its terminal value is optimal for the manager's problem given the investor's best response strategy.

Proof. See the Appendix. □

3.2.2 Optimal Fund Portfolio

In order to determine the equilibrium trading strategies, we now need to find a fund portfolio process $\hat{\pi} \in \Theta_f$ that maximizes the manager's value function

$$V_m^\pi = \mathbf{E}\left[U\Big(\mathcal{X}_m^{-1}\big(v_\pi(0) + W_m(0)\big)D(T)Z_0(T)\Big)\right] \equiv V_m\big(v_\pi(0)\big) \tag{11}$$

[3]Lemma A.1 establishes that the process v_π is well defined.

[4]The function $\mathcal{X}_m$ is real valued under various conditions such as a sublinear polynomial growth condition on the manager's utility function (Karatzas (1996, Exercise 2.2.6)) or the asymptotic elasticity restriction of Kramkov & Shachermayer (1999).

where $U := U_m \circ I_m$. By using the increase of the manager's utility function in conjunction with the decrease of the functions I_m and $\mathcal{X}_m$, it is easily seen that the choice of such a process is *independent* from the manager's utility function and amounts to the maximization of the initial market value $v_\pi(0)$ of the cumulative fee process generated by the investor's best response trading strategy. More precisely, the optimal fund portfolio process is the solution to the stochastic control problem with value function

$$\hat{v}(0) \equiv \sup_{\pi \in \Theta_f} v_\pi(0) \equiv \sup_{\pi \in \Theta_f} \mathbf{E}^0 \int_0^T \Big(\gamma \cdot D(t) \phi_i^\pi(t) \Big) dt \tag{12}$$

where the non negative process ϕ_i^π is defined as in (1). Our first result in this section establishes a technical property which will be useful in what follows.

Lemma 1 *Let $\pi \in \Theta_f$ be given and W_i^π denote the corresponding investor's best response wealth process. Then the process M_π defined by*

$$M_\pi(t) \equiv e^{\gamma \int_0^t \eta_\pi(\tau)^+ d\tau} \left(\frac{D(t) W_i^\pi(t)}{W_i(0)} \right)$$

with the process η_π as in (9), is strictly positive and is a square integrable continuous martingale under the risk neutral probability measure P^0.

Proof. See the Appendix. □

Let $\pi \in \Theta_f$ denote an arbitrary fund portfolio process. The integration of the stochastic differential (5) equation verified by the process W_i^π shows that

$$\int_0^t \Big(D(s) d\Phi^\pi(s) \Big) = \int_0^t \Big(D(s) \phi_i^\pi(s) \pi(s)^* \sigma(s) dB_0(s) - d\big(D(s) W_i^\pi(s) \big) \Big).$$

By using (1), it is easily seen that the square integrability of the process M_π implies that the investor's discounted wealth process DW_i^π is square integrable under the risk neutral probability measure and taking expectations on both sides of the above expression we obtain that the optimal fund portfolio problem (12) can be written as

$$\hat{v}(0) \equiv \sup_{\pi \in \Theta_f} \Big(W_i(0) - \mathbf{E}^0 D(T) W_i^\pi(T) \Big). \tag{13}$$

Thanks to the results of the previous sections we can interpret the objective function in (13) as the difference between the investor's capital $W_i(0)$ which also constitutes the minimal amount necessary to finance the optimal terminal wealth $W_i^\pi(T)$ under the dynamic budget constraint (5) and the minimal amount necessary to finance this same terminal wealth in the complete financial market to which the manager has access. The discrepancy between these two amounts generates the initial market value of the cumulative fee process to be received and we are thus prompted to conclude that the optimal fund portfolio is that which "maximizes the incompleteness" of the financial market to which the investor has access.

Although wealthy of interpretations, the above formulation of the optimal fund portfolio problem is not the easiest one to study. By combining the result of Lemma 1 with Girsanov theorem, we have that for an arbitrary admissible fund portfolio process the formula

$P^\pi(A) := \mathbf{E}^0 1_A M_\pi(T)$ defines a probability measure which is equivalent to the risk neutral probability measure and under which the process

$$B_\pi(t) \equiv B_0(t) + \int_0^t \Big(\phi_i^\pi(s)\sigma(s)^*\pi(s)ds\Big)$$

is an $n-$dimensional standard Brownian motion. In conjunction with the non negativity of the investor's initial capital, the definition of this new probability measure shows that (13) can be written as

$$\hat{v}(0) = W_i(0)\left[1 - \inf_{\pi\in\Theta_f} \mathbf{E}^\pi \left(e^{-\gamma\int_0^T \eta_\pi(\tau)^+ d\tau}\right)\right]. \tag{14}$$

The approach we take to solve this problem relies on the interpretation of its objective function as solution to a Backward Stochastic Differential Equation and on the comparison theorems for such equations (see the Appendix for a brief introduction to this theory).
The fund portfolio process being fixed, we start by considering the Backward Stochastic Differential Equation (BSDE in short)

$$\begin{aligned} -dX_\pi(t) &= -\gamma f\big(t,\pi(t)\big)X_\pi(t)dt - Z(t)^* dB_\pi(t) \\ &= f\big(t,\pi(t)\big)\big(\pi(t)^*\sigma(t)Z(t) - \gamma X_\pi(t)\big)dt - Z(t)^* dB_0(t) \end{aligned} \tag{15}$$

with terminal condition 1 where the second equality follows from the definition of the process B_π and we have set $f(t,\pi(t)) := \eta_\pi(t)^+$. In conjunction with the boundedness of the map f (see the proof of Lemma 1 for details), the result of Theorem B.1 in the Appendix shows that the unique adapted and square integrable solution to this *linear* BSDE is given by

$$X_\pi(t) \equiv \mathbf{E}_t^\pi\left[\exp\bigg(-\gamma\int_t^T f(\tau,\pi(\tau))d\tau\bigg)\right]$$

and by the integrand in the representation of the associated $P^\pi-$martingale. By comparing the above expression with (14) it is now easily seen that the value function of our problem can be written as $W_i(0)(1-\hat{X}(0))$ where

$$\hat{X}(t) \equiv \operatorname*{ess\,inf}_{\pi\in\Theta_f} X_\pi(t) \equiv \operatorname*{ess\,inf}_{\pi\in\Theta_f} \mathbf{E}_t^\pi\left[\exp\bigg(-\gamma\int_t^T f(\tau,\pi(\tau))d\tau\bigg)\right] \tag{16}$$

The terminal condition of the BSDE (15) being constant and independent from the fund portfolio process π, the basic comparison Theorem of Peng (1992) (Theorem B.2 in Appendix B) suggests that in order to solve the optimal fund portfolio problem (14), the first step consists in minimizing the *driver*

$$h\big(t,\pi,x,z\big) \equiv f\big(t,\pi\big)\Big(\pi^*\sigma(t)z - \gamma x\Big)$$

of the backward stochastic differential equation over the set of admissible fund portfolio processes. A straightforward (albeit messy) computation shows that the unique minimizer is the process given by

$$\hat{\pi}\big(t,x,z\big) \equiv 1_{\{\|z/x\|\leq\|\xi(t)\|\}}\big[\sigma(t)^*\big]^{-1}\frac{2\gamma\left(\xi(t)-\frac{z}{x}\right)}{\|\xi(t)\|^2 - \left\|\frac{z}{x}\right\|^2}.$$

In view of Corollary B.1 (see Appendix), the next step now consists in proving that the BSDE associated (under the risk neutral probability measure) with the candidate optimal driver

$$\begin{aligned}\hat{h}(t,x,z) &\equiv \min_{\pi\in\mathbb{R}^n} h(t,\pi,x,z) \equiv h(t,\hat{\pi},x,z) \\ &= -1_{\{\|z/x\|\le\|\xi(t)\|\}}\frac{x}{4}\cdot\left\|\xi(t)-\frac{z}{x}\right\|^2\end{aligned}$$

and the terminal condition 1 admits an adapted solution which coincides almost surely with the process $\hat{X}$ of essential infima. As in the traditional dynamic programming approach to Markovian stochastic control problems where the main difficulty is to prove the existence of a classical solution of the associated Hamilton-Jacobi-Bellman equation, it is the proof of the existence of a solution to this *nonlinear* BSDE which constitutes the main difficulty of our approach. In particular, because the driver of the BSDE

$$-dX(t) = \hat{h}(t, X(t), Z(t))dt - Z(t)^* dB_0(t), \qquad X(T) = 1 \tag{17}$$

has quadratic growth in Z, the standard existence result does not apply and it is well-known that even in the forward case such equations do not necessarily admit strong solutions. Nevertheless, the particular structure of the driver and the boundedness of the terminal condition allow us to apply the results of Kobylanski (1997) and Lepeltier & San Martin (1998) on the existence of solution to BSDEs with superlinear and/or quadratic drivers.

Lemma 2 *The backward stochastic differential equation* (17) *admits a unique maximal adapted solution* (X, Z) *such that* X *is uniformly bounded process.*

Proof. See the Appendix. □

By combining the above existence result with the definition of the candidate optimal control $\hat{\pi}(t,x,z)$ and the refined comparison theorem of Lepeltier & San Martin (1998) for BSDEs with superlinear/quadratic drivers, we now obtain a complete solution to the fund portfolio selection problem (13).

Proposition 3 *Let* (X, Z) *be the maximal solution to the* BSDE (17) *and* $\hat{\pi}$ *be the* $n-$*dimensional process defined by*

$$\hat{\pi}(t,x,z) \equiv 1_{\{\|z/x\|\le\|\xi(t)\|\}}\frac{2\gamma}{\|\xi(t)\|^2-\left\|\frac{z}{x}\right\|^2}\left[\sigma(t)^*\right]^{-1}\left(\xi(t)-\frac{z}{x}\right).$$

Then the portfolio process $\hat{\pi}\in\Theta_f$ *attains the infimum in the stochastic control problem* (14).

Proof. See the Appendix. □

One might be tempted to conjecture that given that the investor is log the optimal fund portfolio process should be that which maximizes the fund's net of fees Sharpe ratio. We can use Proposition 3 to demonstrate that in general this does not hold.

A straightforward computation shows that the unique maximizer of the fund's Sharpe ratio is the process defined by

$$\tilde{\pi}(t) \equiv \arg\max_{\pi\in\mathbb{R}^n} f\big(t,\pi\big) = \frac{2\gamma}{\|\xi(t)\|^2}\left[\sigma(t)^*\right]^{-1}\xi(t).$$

Comparing this with the optimal fund portfolio process $\hat{\pi}$ we obtain that the above conjecture does not generally hold.

However, the following corollary shows that the conjecture does hold when the risk premium process ξ is deterministic.

Corollary 1 *Assume that the risk premium process ξ is a deterministic and bounded function. Then the maximal solution to the* BSDE (17) *is given by*

$$X \equiv \exp\left(-\int_t^T \tfrac{1}{4}\|\xi(\tau)\|^2 d\tau\right)$$

and $Z \equiv (0,...,0)^* \in \mathbb{R}^n$. In this case, the optimal fund portfolio process is also deterministic and coincides with the bounded process $\tilde{\pi}$ defined above.

Proof. See the Appendix. □

Remark 1 Showing that deterministic coefficients imply theat the optimal fund portfolio is obtained by maximizing the investor's net of fees Sharpe ratio can be obtained directly without using BSDEs. Using the definition of the process X_π in (16) in conjunction with the fact that the deterministic process $\tilde{\pi}$ is, by construction, the one that maximizes $f(t,\pi)$ we obtain that the inequality

$$\begin{aligned}\exp\left(-\int_t^T \tfrac{1}{4}\|\xi(\tau)\|^2 d\tau\right) &= \mathbf{E}_t^\pi\left[\exp\left(-\gamma\int_t^T f\big(\tau,\tilde{\pi}(\tau)\big)d\tau\right)\right]\\ &\leq \mathbf{E}_t^\pi\left[\exp\left(-\gamma\int_t^T f\big(\tau,\pi(\tau)\big)d\tau\right)\right] = X_\pi(t)\end{aligned}$$

holds for every admissible fund portfolio process π.

4 Equilibrium

We now gather the results of the previous sections to obtain a complete description of the equilibrium. We start with the general case in which the market coefficients are arbitrary processes and proceed with the special case in which they are deterministic.

Theorem 1 *Let (X,Z) denote the maximal solution to the backward stochastic differential equation* (17) *and define $Y := Z/X$. Then the equilibrium fund portfolio and the investor's equilibrium trading strategy are given by*

$$\hat{\pi}(t) = 1_{\hat{B}}\frac{2\gamma}{\|\xi(t)\|^2-\|Y(t)\|^2}\left[\sigma(t)^*\right]^{-1}(\xi(t)-Y(t)), \tag{18}$$

$$\phi_i^{\hat{\pi}}(t) = \frac{\left(\|\xi(t)\|^2-\|Y(t)\|^2\right)^+}{4\gamma}W_i^{\hat{\pi}}(t) \tag{19}$$

where $\hat{B}$ denotes the non negligible set on which we have $\|Y\| \leq \|\xi\|$. The equilibrium expected utilities of each of the two agents are respectively given by

$$V_i\big(W_i(0)\big) = \mathbf{E}\log\big(W_i(0)S_0(T)\big) + \mathbf{E}\int_0^T \Big(\tfrac{1}{8}\,\|\xi(\tau) - Y(\tau)\|^2\, 1_{\hat{B}}\, d\tau\Big) \tag{20}$$

and by $V_m(W_i(0)(1-X(0)))$ in the notation of (11). *Finally, the manager's equilibrium trading strategy is given by*

$$\theta_m(t) = (\sigma(t)^*)^{-1}h(t)/D(t), \tag{21}$$

where the process h is the integrand in the representation of the martingale defined by (11).

Equation (18) shows that in equilibrium the equity component of the fund is proportional to the fee rate. Thus, our model predicts that *ceteris paribus* funds with high proportional management fees will invest more in equities relative to funds with low proportional management fees. In particular, this implies a positive relationship between a fund's fee rate and its volatility. Equation (19) shows that, as can be expected, the higher the fee rate the lower the fraction of the investor's wealth that is invested in the fund. The fee is charged on assets under management, so that it applies both the equity and the bond components of the fund. Obviously, holding the bond through the fund is more expensive for the investor then holding it directly. To mitigate the investor's reaction to a higher fee rate, the equilibrium fund allocation to equities increases as the fee increases, thereby decreasing the fund's allocation to the bond and enabling the investor to hold a larger fraction of his bond holdings directly versus indirectly. Thus, as the fee rate increases a larger part of the investor's portfolio is invested directly in the bond, both because the fund tilts its portfolio more towards equity and because the investor allocates less to the fund.

The investor's effective equity portfolio weights are given by

$$\frac{\phi_i^{\hat{\pi}}(t)\hat{\pi}(t)}{W_i^{\hat{\pi}}(t)} = 1_{\hat{B}}\frac{1}{2}\big[\sigma(t)^*\big]^{-1}(\xi(t) - Y(t)).$$

The fact that the solution of the BSDE (17) is independent from the fee rate implies that these weights are independent of the fee rate. This implies that the investor's effective overall bond portfolio weight is also independent of the fee rate. Furthermore, multiplying (19) by the fee rate shows that the fraction of the investor's wealth that is paid as management fees is fixed.

Combining the above, yield that both the investor's wealth process and the fee process Φ are independent of the fee rate. This in tern implies that equilibrium value functions of both agents are also independent of the fee rate. In the context of our model, it does not matter whether the investor or the fund manager is the agent setting the management fee. In both cases, the managers ability to select the fund portfolio on the one hand and the investor's ability to adjust the holdings in the fund on the other hand counteract each other so that both agents are equally well off whether one or the other is setting the fee rate.

The first term on the right hand side of (20) is simply the utility that the investor would attain if he was restricted to holding only the bond. The second is the utility gain attributed to being able to trade also the mutual fund. One might have been tempted to conjecture that the fact that the investor is the Stackelberg follower would imply that the investor's equilibrium value function would just be equal to his reservation utility; i.e., the utility he would attain if he could hold only the bond. This argument ignores two important issues. First, in the setting of our model making the investor better off relative to the case where he could invest only in the bond is not inconsistent with making the manager better off. Second, in our model the investor has the ability to move funds in and out of the fund, thereby limiting the managers ability to take advantage of him. This implies that the two agents need to split the surplus, as opposed to the manager being able to expropriate all of the surplus.

If the investor was able to access the financial markets directly without facing any costs, then his expected utility would simply be given by

$$V_i^0 = \mathbf{E} \log\big(W_i(0) S_0(T)\big) + \mathbf{E} \int_0^T \tfrac{1}{2} \left\| \xi(\tau) \right\|^2 d\tau$$

Thus, the *fraction* of his wealth that the investor would be willing to give up to be to directly access the financial markets is

$$p_i = 1 - \exp\Big[\mathbf{E} \int_0^T \Big(\tfrac{1}{8} \left\| \xi(\tau) - Y(\tau) \right\|^2 - \tfrac{1}{2} \left\| \xi(\tau) \right\|^2 \Big) d\tau \Big] \tag{22}$$

In view of the independence of the equilibrium value functions from the fee rate and given that the market value of the fees to be received is $W_i(0)(1 - X(0))$, one may be tempted to conjecture that p_i coincides with $1 - X(0)$. To see that this conjecture is false, first observe that when the market coefficients are constant, or more generally deterministic, Corollary 1 shows that the backward stochastic differential equation (17) degenerates into an ordinary differential equation whose unique solution is given by (1). In particular, the integrand Z is then identically equal to zero and as result we obtain the following:

Corollary 2 *Assume that the market coefficients are deterministic functions. Then, the equilibrium fund portfolio and the investor's equilibrium trading strategy are given by*

$$\hat{\pi}(t) = \frac{2\gamma}{\left\| \xi(t) \right\|^2} \big(\sigma(t)^* \big)^{-1} \xi(t) \tag{23}$$

$$\phi_i^{\hat{\pi}}(t) = \frac{\left\| \xi(t) \right\|^2}{4\gamma} W_i^{\hat{\pi}}(t) \tag{24}$$

and letting X be the deterministic bounded function defined by (1) *we have that the equilibrium expected utilities of each of the two agents are given by*

$$V_i\big(W_i(0)\big) = \log\big(W_i(0) S_0(T)\big) + \int_0^T \Big(\tfrac{1}{8} \left\| \xi(\tau) \right\|^2 d\tau \Big) \tag{25}$$

and by $V_m(W_i(0)(1 - X(0)))$ in the notation of (11)*, where*

$$X(0) = \exp\Big(-\frac{1}{4} \int_0^T \left\| \xi(s) \right\|^2 ds \Big).$$

Assuming constant, or more generally deterministic, market coefficients and setting $Y = 0$ in (22) we obtain

$$\begin{aligned} p_i &= 1 - \exp\Big(- \int_0^T \tfrac{3}{8} \left\|\xi(\tau)\right\|^2 d\tau \Big) \\ &> 1 - \exp\Big(- \int_0^T \tfrac{1}{4} \left\|\xi(\tau)\right\|^2 d\tau \Big) = 1 - X(0), \end{aligned}$$

implying that the presence of delegated portfolio management introduces a dead weight cost on the investor.

4.1 Fund Flows

The empirical literature on the relationship between flow and returns (see for example Chevalier and Ellison (1997) or Sirri and Tufano (1998)) has focused on the relationship between current flows and past returns. Specifically when analyzing this relationship the potential relationship between measured flows and contemporaneous returns has been ignored. We now demonstrate that not controlling for contemporaneous returns potentially biases this relationship.

For fixed price process coefficients, Corollary 2 in conjecture with the fund dynamics (4) and the investors wealth process dynamics (5) implies that

$$\begin{aligned} dF(t) &= F(t)\Big((r + 2\gamma)\, dt + 2\gamma \frac{\xi^*}{\|\xi\|^2}\, dB(t) \Big) \\ dW_i(t) &= W_i(t)\Big((r + \frac{\|\xi\|^2}{4})\, dt + \frac{\xi^*}{2}\, dB(t) \Big) \end{aligned}$$

Using the fact that $\hat{B}(t) \equiv \frac{\xi^*}{\|\xi\|} B(t)$ is a one dimensional Brownian motion we obtain that for $t > s$

$$\begin{aligned} F(t) &= F(s) \exp\left((r + 2\gamma - \frac{1}{2}\left(\frac{2\gamma}{\|\xi\|}\right)^2)(t-s) + \frac{2\gamma}{\|\xi\|}(\hat{B}(t) - \hat{B}(s)) \right) \\ W_i(t) &= W_i(s) \exp\left((r + \frac{\|\xi\|^2}{8})(t-s) + \frac{\|\xi\|}{2}(\hat{B}(t) - \hat{B}(s)) \right). \end{aligned}$$

Using the first equation to get $(\hat{B}(t) - \hat{B}(s))$ as a function of $\frac{F(t)}{F(s)}$ and plugging into the second implies that

$$W_i(t) = W_i(s) \exp\left(A(r, \gamma, \|\xi\|^2)(t-s) \right) \left(\frac{F(t)}{F(s)} \right)^{\frac{\|\xi\|^2}{4\gamma}}$$

where $A(r, \gamma, \|\xi\|^2) \equiv r\Big(1 - \frac{\|\xi\|^2}{4\gamma}\Big) + \frac{\gamma}{2}\Big(1 - 5\frac{\|\xi\|^2}{4\gamma}\Big)$.

We consider two measures for funds flows in the interval $[s,t]$. The first,

$$\begin{aligned} Flow1_{[s,t]}\left(\frac{F(t)}{F(s)}\right) &= \frac{W_i(t) - W_i(s)\frac{F(t)}{F(s)}}{W_i(s)\frac{F(t)}{F(s)}} \\ &= \exp\left(A(r,\gamma,\|\xi\|^2)(t-s)\right)\left(\frac{F(t)}{F(s)}\right)^{\frac{\|\xi\|^2}{4\gamma}-1} - 1 \end{aligned}$$

The second, which is the one typically used in empirical studies,

$$\begin{aligned} Flow2_{[s,t]}\left(\frac{F(t)}{F(s)}\right) &= \frac{W_i(t) - W_i(s)\frac{F(t)}{F(s)}}{W_i(s)} \\ &= \exp\left(A(r,\gamma,\|\xi\|^2)(t-s)\right)\left(\frac{F(t)}{F(s)}\right)^{\frac{\|\xi\|^2}{4\gamma}} - \frac{F(t)}{F(s)} \end{aligned}$$

By analyzing $Flow1_{[s,t]}$ and its derivatives the following proposition is obtained

Proposition 4

1. *As a function of* $\frac{F(t)}{F(s)}$: $Flow1_{[s,t]}$ *is decreasing if* $\frac{\|\xi\|^2}{4\gamma} < 1$ *and increasing* $1 < \frac{\|\xi\|^2}{4\gamma}$. *It is convex if* $\frac{\|\xi\|^2}{4\gamma} < 1$ *or* $2 \leq \frac{\|\xi\|^2}{4\gamma}$, *and concave if* $1 < \frac{\|\xi\|^2}{4\gamma} < 2$. *Its curvature is given by*

$$\frac{Flow1''_{[s,t]}}{Flow1'_{[s,t]}} = \left(\frac{\|\xi\|^2}{4\gamma} - 2\right)\left(\frac{F(t)}{F(s)}\right)^{-1}.$$

2. *For a given level of* $\frac{F(t)}{F(s)}$: $Flow1_{[s,t]}$ *increases in the fee rate iff* $\gamma^2 > \left[\ln\left(\frac{F(t)}{F(s)}\right) - r\right]\frac{\|\xi\|^2}{2}$.

The proposition demonstrates that in quantifying the relationship between flows and lagged returns one should control both for the fee rate and the contemporaneous returns. The first part analyzes the contemporaneous flow performance relationship, for different fee rates. The second, holds fixed the level of fund returns and compares measured flows across different fee rates.

Corollary 2 implies that $\frac{\|\xi\|^2}{4\gamma} > 1$ iff the investor is holding a levered position in the fund (i.e. shorts the bond). Given fixed price process coefficients, both the investor and the fund hold fixed portfolios. If the investor is non levered then he holds the bond both directly and indirectly through his holdings of the fund. Thus, the bond constitutes a larger fraction of his portfolio relative to that of the fund. When the investor is levered his portfolio is tilted more towards equities relative to the fund portfolio. The direction of the wedge between the investor's portfolio and the fund portfolio determines whether the contemporaneous relationship between flows and returns will be increasing or decreasing. For a non levered investor: when the return on the fund is higher than the return on the bond, as fund returns go up the investor withdraws money from the fund in order to maintain his fixed portfolio weights. On the other hand, a positive inflow is required to maintain fixed weights following a fund return that is lower than the return on the bond. The opposite occurs when the investor is levered. Furthermore, for both levered and unlevered investor's, the more extreme the fund returns are the more pronounced the rebalancing will be.

The second part of the proposition is a consequence of the fact that a higher fee rate implies that a larger part of the investor's portfolio is held directly in the bond. For example, following a negative fund return the investor needs to inject money into the fund in order to rebalance his portfolio back to the optimal composition. Holding this return level fixed: a higher fee rate implies a higher return on the investor's portfolio (i.e., less negative) which in tern implies less rebalancing on the part of the investor. While the condition in the proposition depends both on the returns and the fee rate, for sufficiently low returns ($\ln\left(\frac{F(t)}{F(s)}\right) < r$) the flow is an increasing function of the fee rate. Given that fee rates have a bounded support, we also obtain that for sufficiently high returns they will be decreasing in the fee rate.

By construction, $Flow2_{[s,t]} = \left(\frac{F(t)}{F(s)}\right) Flow1_{[s,t]}$ so that $Flow2_{[s,t]}$ is positive iff $Flow1_{[s,t]}$ is positive. For positive flows, $Flow2_{[s,t]}$ will give a higher (lower) value than $Flow1_{[s,t]}$ if the gross fund return is above (below) 1. For negative flows $Flow2_{[s,t]}$ will give a lower (higher) value than $Flow1_{[s,t]}$ if the gross fund return is above (below) 1.

Furthermore, by analyzing $Flow2_{[s,t]}$ and its derivatives the following proposition is obtained

Proposition 5

1. *As a function of* $\frac{F(t)}{F(s)}$: $Flow2_{[s,t]}$ *is inverse U-shaped if* $\frac{\|\xi\|^2}{4\gamma} < 1$, *and U-shaped if* $\frac{\|\xi\|^2}{4\gamma} > 1$.

2. *For a given level of* $\frac{F(t)}{F(s)}$: $Flow1_{[s,t]}$ *increases in the fee rate iff* $\gamma^2 > \left[\ln\left(\frac{F(t)}{F(s)}\right) - r\right] \frac{\|\xi\|^2}{2}$.

The second part of the two propositions is the same, but the first is not. A disadvantage of using $Flow2_{[s,t]}$ is that it distorts the measured flow performance relationship. It is no longer the case that the flow performance relationship is monotone, for a given fee rate. This distortion stems from the $-\left(\frac{F(t)}{F(s)}\right)$ term.[5] In order to control for the impact of contemporaneous returns on measured flows it is no longer sufficient just to know the relevant region of the fee rates (i.e., $\frac{\|\xi\|^2}{4\gamma} < 1$ or $\frac{\|\xi\|^2}{4\gamma} > 1$), as is the case with $Flow1_{[s,t]}$. For every fee rate, it is also required to know the fund rate of return at which the derivative of $Flow2_{[s,t]}$ switches sign.

5 Relaxing the Inaccessibility Assumption

We now extend the model of the previous sections in order to allow the investor to have direct access to some of the n stocks traded on the market.

To simplify the exposition, we first introduce some notation. Let $m \leq n$ be the number of stocks that the investor can trade and without loss of generality assume that the prices of these stocks are the first m components of the stock price process defined by (2). Let

[5] Another disadvantage of $Flow2_{[s,t]}$ has been pointed out by Berk and Green (2002).

$\bar{\sigma}$ denote the volatility matrix of these stocks. For an arbitrary fund portfolio process π, define the matrices

$$\Lambda_\pi(t) \equiv \begin{pmatrix} \pi(t)^*\sigma(t) \\ \bar{\sigma}(t) \end{pmatrix} \qquad \Sigma_\pi(t) \equiv \Lambda_\pi(t)\Lambda_\pi(t)^*$$

$$C_\pi(t) \equiv \Big(\Sigma_\pi^{-1}(t)\Lambda_\pi(t)\xi_\pi(t)\Big) \qquad A_\pi(t) \equiv \left(\frac{\Sigma_\pi^{-1}(t)\mathbf{e}}{\mathbf{e}^*\Sigma_\pi^{-1}(t)\mathbf{e}}\right)$$

$$\Omega_0(t) \equiv \bar{\sigma}(t)^*\big(\bar{\sigma}(t)\bar{\sigma}(t)^*\big)^{-1}\bar{\sigma}(t) \qquad \Omega(t) \equiv \Big(\mathbb{I}_n - \Omega_0(t)\Big),$$

where $\mathbf{e} \equiv (1, 0, ..., 0)^*$ and $\mathbb{I}_n$ denotes the $n-$dimensional unit matrix.

5.1 The Investor's Optimal Strategy

Let Θ_f^* denote the set of admissible fund portfolios which are such that the variance covariance matrix Σ_π is invertible. If the variance covariance matrix is not invertible, then the fund is a linear combination of the m stocks available to the investor and is thus strictly dominated, due to the presence of management fees.

Given a fund portfolio process $\pi \in \Theta_f^*$, the investor's wealth process satisfies the dynamic budget constraint

$$\begin{aligned} dW_i(t) &= \Big(r(t)W_i(t) - \gamma\phi(t)\Big)dt + \theta_i(t)^*\Lambda_\pi(t)\Big(dB(t) + \xi(t)dt\Big) \\ &= r(t)W_i(t)dt + \theta_i(t)^*\Lambda_\pi(t)\Big(dB(t) + \xi_\pi(t)dt\Big) \end{aligned} \tag{26}$$

where the first component of the investors trading strategy (θ_i) represents his holding in the fund (i.e., $\phi = \theta_{i,1}$), and is restricted to being non negative. Furthermore, ξ_π denotes the modified risk premium process associated with the risky assets to which the investor has access, that is

$$\xi_\pi(t) \equiv \xi(t) - \gamma\Lambda_\pi(t)^*\Sigma_\pi^{-1}(t)\mathbf{e}.$$

As in the previous sections, the investor's best response problem is to maximize the expected logarithmic utility of his terminal wealth over the set Θ_i^π of trading strategies which satisfy the no short selling constraint on the fund and are such that the corresponding solution to (26) is a non negative process.

The technique we use to deal with the incompleteness of the market to which the investor has access is by now well-known. It was developed by Karatzas et al. (1991,1993), Shreve & Xu (1991), He & Pearson (1991) and Cuoco (1997) among others and consists in constructing the solution to the investor's best response utility maximization problem from that of a *dual* minimization problem which solves for the investor's marginal rates of substitution.

Proposition 6 *Let $\pi \in \Theta_f^*$ denote an arbitrary fund portfolio process. Then the investor's best response trading strategy is given by*

$$\theta_i^\pi(t) = \Big(C_\pi(t) + A_\pi(t)\big(\mathbf{e}^* C_\pi(t)\big)^-\Big) W_i^\pi(t) \tag{27}$$

where W_i^π is the corresponding solution to (26). In particular, the investor's best response investment in the mutual fund is given by $\phi^\pi = \theta_{i,1}^\pi = (\mathbf{e}^* C_\pi)^+ W_i^\pi$.

Proof. The derivation of this proposition is based on results of Cvitanič and Karatzas (1993) on the problem of utility maximization under constraint and is presented in the appendix. □

The intuition for the above result is as follows. First, the investor determines the trading strategy that he would adopt if he was not subject to the no short selling constraint on the fund. This yields the first term in the expression for the best response trading strategy in (27). When the no short selling constraint on the mutual fund binds (that is when the quantity $\mathbf{e}^* C_\pi$ is negative), the investor sets his position in the fund to zero and projects the negative fraction of his wealth that he would have like to invest in the fund on the m available stocks according to the vector A_π. This yields the second term in the expression for the best response trading strategy.

5.2 The Optimal Fund Portfolio

The next two steps that we have to take in order are to first to solve the manager's utility maximization in his account given and then to optimize the value function of this problem to determine the optimal fund portfolio.

As we have shown in Section 3.2.2 the manager's value function depends on the fund portfolio only through the initial market value $v_\pi(0)$ of the management fees to be received, so that the optimal fund portfolio can be obtained by solving the stochastic control problem

$$\hat{v}(0) \equiv \sup_{\pi \in \Theta_f^*} v_\pi(0). \tag{28}$$

In order to facilitate the study of this stochastic control problem, we start by rewriting it in a more convenient form by using the following

Lemma 3 *Let $\pi \in \Theta_f^*$ be a given admissible fund portfolio and denote by W_i^π the corresponding investor's best response wealth process. Then the process*

$$M_\pi(t) \equiv e^{\gamma \int_0^t (\mathbf{e}^* C_\pi(s))^+ ds} \left(\frac{D(t) W_i^\pi(t)}{W_i(0)} \right)$$

is strictly positive and a square integrable continuous martingale under the risk neutral probability measure P^0.

Proof. **TBA.**□

Combining Lemma 3 with Itô's lemma and Girsanov's theorem we have that for an arbitrary $\pi \in \Theta_f^*$ the formula $P^\pi(A) := E^0 1_A M_\pi(T)$ defines a probability measure which is equivalent to the risk neutral measure and under which

$$B_\pi(t) \equiv B_0(t) + \int_0^t \Lambda_\pi(s)^* \Big(C_\pi(s) + A_\pi(t) \big(\mathbf{e}^* C_\pi(s)\big)^- \Big) ds$$

is an $n-$dimensional standard Brownian motion. In conjunction with the strict non negativity of the investor's initial capital, the definition of this family of equivalent probability measures shows that (28) can be written as

$$\hat{v}(0) = W_i(0) \left[1 - \inf_{\pi \in \Theta_f^*} E^\pi \left(e^{-\gamma \int_0^T (\mathbf{e}^* C_\pi(t))^+ dt} \right) \right]. \tag{29}$$

As in Section 3.2.2, one can convert this minimization problem into one which optimizes over a family of backward stochastic equations. Unfortunately, the associated drift optimization cannot be solved explicitly, unless the price process coefficients are deterministic. To gain some insights on the impact of the ability of the investor to trade directly in a subset of the risky assets on the manager's portfolio choice, we henceforth restrict the analysis to the case of deterministic coefficients. Given deterministic coefficients, the set $\hat{\pi}(t) \in \arg\max_{\pi \in \mathbb{R}^n} (\mathbf{e}^* C_\pi(t))^+$ is deterministic. A similar argument to the one used in Remark 1 yields that

$$\hat{v}(0) = v_{\hat{\pi}}(0) = W_i(0) \left(1 - e^{-\gamma \int_0^T (\mathbf{e}^* C_{\hat{\pi}}(t))^+ dt} \right);$$

implying that any element in the set $\arg\max_{\pi \in \mathbb{R}^n} (\mathbf{e}^* C_\pi(t))^+$ is an optimal fund portfolio process.

Denote by $\Omega(t)^\perp = \mathrm{Ker}(\Omega(t))$ the null space of the matrix $\Omega(t)$ and introduce the set

$$\mathcal{S}(t) \equiv \left[\frac{2\gamma}{\xi(t)^* \Omega(t) \xi(t)} (\sigma(t)^*)^{-1} \left(\Omega(t)^\perp + \xi(t) \right) \right].$$

Since the rows of the matrix $\bar{\sigma}$ are linearly independent, the square matrix $\Omega(t)$ has rank $n-m$ and it therefore follows from standard linear algebra results (see for example in Smith (1984)) that for each t the set $\mathcal{S}(t)$ is a subspace of dimension m. The following proposition provides an explicit characterization of the optimal fund portfolio in the case of constant market coefficients.

Proposition 7 *Assume that the market coefficients are deterministic. Then the set of optimal fund portfolio is the set of non zero $\mathcal{S}-$valued processes.*

Proof. **TBA.**□

The proposition highlights the fact that when the investor has access to a subset m of the risky assets in addition to the fund the manager has m degrees of freedom in choosing the optimal fund portfolio.[6] In particular, the initial market value of the management fees to

[6]The fact that the optimal fund portfolio is not uniquely determined but instead takes values in an $m-$dimensional subspace is reminiscent of the fact that for zero net supply securities, equilibrium considerations can only pin down the risk premium but not the drift and volatility separately.

be received does not depend on the choice of the non zero vector $\hat{\pi} \in \mathcal{S}$, and is given by

$$\hat{v}(0) = W_i(0)\Big(1 - X(0)\Big) \equiv W_i(0)\left(1 - e^{-\frac{1}{4}\int_0^T (\xi(t)^*\Omega(t)\xi(t))dt}\right). \tag{30}$$

Given that the fund has m degrees of freedom in determining its portfolio, it is natural to ask whether there exists a unique optimal portfolio process for the fund that will entail the fund trading only in the bond and the $n-m$ risky assets that the investor cannot trade in. The following results provides an explicit characterization of this unique optimal portfolio process.

Corollary 3 *Assume that the market coefficients are deterministic, then*

$$\hat{\pi}_k(t) \equiv \left\{ \begin{array}{ll} 0 & \text{if } k \le m \\ \frac{2\gamma}{(\xi(t)^*\Omega(t)\xi(t))}((\sigma(t)^*)^{-1}\xi(t))_k & \text{otherwise} \end{array} \right\}, \qquad 1 \le k \le n$$

is the unique optimal fund portfolio with the property that the fund does not invest in any of the stocks that the investor can access directly.

Proof. **TBA.**□

5.3 Equilibrium

Gathering the results of the previous sections we obtain the description of the equilibrium, under the assumption of deterministic market coefficients.

Theorem 2 *Assume that the market coefficients are deterministic functions. Then the investor's equilibrium trading strategy in the mutual fund and the equilibrium fund portfolio are given by*

$$\phi^{\hat{\pi}}(t) = W_i^{\hat{\pi}}(t)\left(\frac{\xi(t)^*\Omega(t)\xi(t)}{4\gamma}\right) \tag{31}$$

$$\hat{\pi}(t) = \frac{2\gamma}{\xi(t)^*\Omega(t)\xi(t)}(\sigma(t)^*)^{-1}\left(L(t) + \xi(t)\right) \tag{32}$$

where $L(t)$ is any vector in the null space of the matrix $\Omega(t)$. The equilibrium expected utility of each of the two agents are respectively given by

$$V_i(W_i(0)) = \log(W_i(0)S_0(T)) + \int_0^T \left(\tfrac{3}{8}\left\|\xi(t)\right\|^2 + \tfrac{1}{8}\xi(t)^*\Omega_0(t)\xi(t)\right) dt \tag{33}$$

and by $V_m(W_i(0)(1-X(0)))$ in the notation of (**??**), *where the non negative constant $X(0)$ is defined in* (30).

Proof. **TBA.**□

The results are consistent with the ones obtained under the assumption that the investor is restricted from trading equity. First, the investor's holding of the fund ($\phi^{\hat{\pi}}$) is inversely related to the fee rate. Second, as can be seen from multiplying $\phi^{\hat{\pi}}$ and $\hat{\pi}(t)$, the investor's effective equity portfolio as well as the fraction of his wealth that is invested in the bond is independent of the fee rate. Third, increasing the fee rate results in the investor holding a larger part of his bond position directly versus indirectly through his holding of the fund. Forth, the fee process Θ is independent of the fee rate, as can be seen from multiplying the fee rate by the investor's holdings in the fund. Fifth, combining the second and the forth above implies that the investor's wealth process is independent of the fee rate, so that his value function should be independent of the fee rate; as it is. Sixth, combining the forth and the fifth implies that the manager's value function should be independent of the fee rate; as it is. Thus, both value functions are again independent from the fee rate, even though the investor can trade some of the risky assets directly.

The utility based price for the investor of having direct access to the markets becomes

$$p_i = 1 - e^{-\frac{3}{8}\int_0^T (\xi(t)^*\Omega(t)\xi(t))dt} \geq 1 - e^{-\frac{1}{4}\int_0^T (\xi(t)^*\Omega(t)\xi(t))dt} = 1 - X(0),$$

where equality holds only in the case that the investor is able to trade in all of the risky securities (i.e., $m = n$) and strict inequality holds in all other cases (i.e., $m < n$). The case in which the investor does not have access to any of the risky securities corresponds to $\Omega(t) = \mathbb{I}_n$. On the other hand when he has full access to equity markets $\Omega(t) = 0$. In that case the fund becomes a dominated redundant asset and consequently the investor does not trade in the fund. Given that the investor does not trade in the fund the fund, of course, does not receive any fees ($X(0) = 1$). Since in this case the investor already has full access to the markets the utility based price becomes 0.

The following points need to be incorporated into the discussion:

As can be expected: the more stocks the investor can trade directly the better off he is.

- Direct access to m stocks increases the investor's welfare by $\frac{1}{8}(\xi^*\Omega_0\xi)T$ (with constant coefficients).
- As $m \uparrow n$ (i.e. full access to risky assets) the matrix Ω_0 becomes the unit matrix and hence the value function of the investor converges to

$$\log(W_i(0)S_0(T)) + \frac{1}{2}\int_0^T \|\xi(s)\|^2 \, ds,$$

  which is the correct direct access value function. At the same time, the matrix Ω becomes the zero matrix and thus the amount of fees extracted by the manager decreases to zero, as it should.
- As $m \downarrow 0$ Ω_0 goes to the zero matrix and we are back to the case considered earlier.

6 Conclusion

In this paper we have analyzed the impact of the ability of investors to dynamically determine their holdings in a fund on a fund's dynamic portfolio choice decision. To the best of our knowledge, this is the first paper to theoretically analyze the impact of fund flows on a fund's portfolio choice decisions in a dynamic setting in which both the fund's portfolio decisions and the flows are determined endogenously.

Our model predicts a positive relationship between a fund's proportional fee rate and a fund's equity exposure, or alternatively a fund's volatility. While funds that have higher proportional fee rates are expected to tilt their portfolio towards stocks, relative to funds with lower fee rates, it is not necessarily true that the dollar amount held in equities by higher fee rate funds will also be higher. In equilibrium, higher fees also imply that investors will allocate less of their wealth to these funds. In the context of our model, the two effects exactly offset each other, so that in equilibrium the amount of equity held by a fund is independent of its fee rate. Furthermore, we show that even when the investors are non strategic and take a fund's portfolio as given their ability to move money in and out of the fund dynamically gives them considerable power. Interestingly, in our model the equilibrium welfare of the agents is not impacted by the fee rate. Thus, the welfare of the different agents is not effected by whether the investors or the fund set the fee rate.

In the current paper, the choice of a log utility function for the investor was made for tractability. Given that we wanted to focus on the fund's portfolio choice decision, a natural simplifying assumption was to make the investor myopic, so that only the fund manager will have hedging demands. As a result, when the investment opportunity set is stationary their are no hedging demands by either the investor or the fund and the fund portfolio is chosen so that to maximize the investors net of fees Sharpe ratio. A non stationary investment opportunity set introduces flow hedging demands on part of the manager. On the other hand, since the investor has a log utility function, and is assumed to be non strategic, his portfolio decisions remain myopic. Allowing for other utility functions for the investor would introduce hedging considerations also for the investor, and would probably make the problem analytically intractable; requiring a non trivial numerical solution.

While we view this paper as an important first step in understanding the dynamic portfolio choice decisions made by fund managers, there are some important extensions that are left for future work. First, while the proportional fee is the most commonly used compensation contract in the mutual fund industry, there are some funds that also use fulcrum performance fees as part of their compensation contract. Furthermore, both hedge funds and pension funds typically have an asymmetric performance fee component; whereas the fund manager shares in the upside but not in the down side. An analysis of the impact of such fees in a setting in which the investor is restricted to making his or her allocation decision only at the initial date is considered in Cuoco and Kaniel (2001). One could attempt to extend that analysis to a dynamic setting similar to the one in the current. Second, in some cases, especially in hedge funds, the manager commits part of her capital to investing in the fund she manages. In such a setting it will no longer be the case that the funds portfolio will be determined so that to maximize the present value of the fees to be received, as such a fund strategy might adversely effect her own portfolio composition. Finally, incorporating asymmetric information and or allowing for multiple mutual funds are both important, but very challenging extensions.

Appendixes

A Proofs

Proof of Proposition 1. Let $\pi \in \Theta_f$ be an arbitrary admissible fund portfolio process, ϕ_i^π be the candidate optimal strategy defined by (1) and W_i^π denote the corresponding wealth process. The fact that ϕ_i^π is an admissible trading strategy for the investor follows from the result of Lemma 1 which also shows that the corresponding wealth process is strictly positive.

In order to verify that ϕ_i^π identifies the investor's best response trading strategy, let ϕ be an arbitrary admissible trading strategy for the investor given the fund portfolio process and denote by W_i the corresponding wealth process. Using Itô's product rule in conjunction with (5) and (1) we have

$$d\left(\frac{W_i(t)}{W_i^\pi(t)}\right) = \Big(\phi(t) - \max\{0, \eta_\pi(t)\} W_i(t)\Big) \frac{\pi(t)^* \sigma(t)}{W_i^\pi(t)} dB(t) - \frac{\|\sigma(t)^* \pi(t)\|^2}{W_i^\pi(t)} \Big(\max\{0, -\eta_\pi(t)\} \phi(t)\Big) dt$$

where η_π is the process defined by (9). Observing that both ϕ and W_i/W_i^π are non negative processes and applying Fatou's lemma we conclude that the latter process is a supermartingale under the objective measure. Using this in conjunction with the concavity and increase of the logarithm we obtain that

$$\begin{aligned}\mathbf{E}\left[\log(W_i(T))\right] - \mathbf{E}\left[\log(W_i^\pi(T))\right] &= \mathbf{E}\left[\log\left(\frac{W_i(T)}{W_i^\pi(T)}\right)\right] \\ &\leq \log\left(\mathbf{E}\left[\frac{W_i(T)}{W_i^\pi(T)}\right]\right) \leq \log\left(\frac{W_i(0)}{W_i^\pi(0)}\right) = 0\end{aligned}$$

holds and it now only remains to observe that the investor's admissible trading strategy $\phi \in \Theta_i^\pi$ was arbitrary in order to conclude our proof. □

Lemma A.1 *For every admissible fund portfolio process $\pi \in \Theta_f$, the cumulative fee process Φ^π defined by* (3.2.1) *is integrable under the probability measure P^0.*

Proof. The process W_i^π being non negative we deduce from (3.2.1) that Φ^π is increasing and since its initial value is equal to zero, it is also non negative. Let f be the progressively measurable mapping defined by

$$f\big(t, \pi(t)\big) \equiv \eta_\pi(t)^+ \equiv \left(\frac{\pi(t)^* \sigma(t) \xi(t) - \gamma}{\|\sigma(t)^* \pi(t)\|^2}\right)^+. \tag{A.1}$$

By definition, we have $f(t, \pi(t)) \leq \sup_{\pi \in \mathbb{R}^n} f(t, \pi)$ for every admissible fund portfolio process. A simple computation shows that the supremum is uniquely attained by the progressively measurable process

$$\tilde{\pi}(t) \equiv \arg\max_{\pi\in\mathbb{R}^n} f\big(t,\pi\big) = 2\gamma \cdot \big[\sigma(t)^*\big]^{-1} \frac{\xi(t)}{\|\xi(t)\|^2} \tag{A.2}$$

and that the corresponding value is given by $(4\gamma)^{-1}\|\xi(t)\|^2$. Let now $\ell \in \mathbb{R}_+$ denote a uniform bound on the norm of the risk premium process. The increase of the cumulative fee process Φ^π and the boundedness of the discount factor imply that we have

$$\begin{aligned}
\mathbf{E}^0 \sup_{t\in[0,T]} \big|\Phi^\pi(t)\big| &\equiv \mathbf{E}^0\big[\Phi^\pi(T)] \le c_1 \cdot \mathbf{E}^0 \int_0^T \Big(D(t)d\Phi^\pi(t)\Big) \\
&= c_1 \cdot \mathbf{E} \int_0^T \Big(\gamma D(t) Z_0(t) W_i^\pi(t) f\big(t,\pi(t)\big)\Big) dt \\
&\le c_2 \cdot \mathbf{E} \int_0^T \Big(D(t) Z_0(t) W_i^\pi(t)\, \|\xi(t)\|^2\Big) dt \\
&\le c_2 \cdot \mathbf{E} \int_0^T \Big(\ell D(t) Z_0(t) W_i^\pi(t)\Big) dt \le c_3 \cdot W_i(0)
\end{aligned}$$

for some non negative constants $\{c_i\}_{i=1}^3$ where the second inequality follows from Dellacherie's integration formula (see Brémaud (1981, Theorem A.19)) and the last one obtains upon observing that the investor's discounted wealth process is a non negative supermartingale under the risk neutral probability measure. □

Proof of Proposition 2. Under the conditions of the statement, the function $\mathcal{X}_m$ is strictly decreasing and satisfies $\mathcal{X}_m(0+) = \infty$ as well $\mathcal{X}_m(\infty) = 0$ so that it admits a strictly decreasing inverse function. Now let λ^π be as in the statement. The random variable $W_m^\pi(T)$ being feasible by construction all we have to prove is that it is optimal for the manager's problem. By concavity of the manager's utility function we have that

$$\begin{aligned}
\mathbf{E}\Big[U_m\Big(W(T)\Big) - U_m\Big(W_m^\pi(T)\Big)\Big] &\le \mathbf{E}\Big[U_m'\Big(W_m^\pi(T)\Big)\Big(W(T) - W_m^\pi(T)\Big)\Big] \\
&= \lambda^\pi \cdot \Big(\mathbf{E}^0\big[D(T)W(T)\big] - v_\pi(0) - W_m(0)\Big) \le 0
\end{aligned} \tag{A.3}$$

holds for every feasible random variable where the second inequality follows from the budget constraint (10) and the fact that λ^π is strictly positive. The above expression shows that the random variable of (3.2.1) is optimal for the manager's problem and it only remains to observe that uniqueness of the solution follows from the strict concavity of U_m to complete the proof. □

Proof of Lemma 1. Applying Itô's lemma to the investor's best response wealth process and using (1), we obtain that the non negative process M_π is the stochastic exponential of the local martingale

$$N_\pi(t) \equiv \int_0^t \Big(f\big(s,\pi(s)\big)\pi(s)^*\sigma(s)dB_0(s)\Big), \tag{A.4}$$

hence a local martingale itself. Coming back to the definition of the progressively measurable mapping f in (A.1), it is easily seen that we have

$$(dt)^{-1}d\langle N_\pi\rangle(t) \equiv \left\|f\big(s,\pi(s)\big)\sigma(s)^*\pi(s)\right\|^2 \leq \|\xi(t)\|^2 \leq \ell \tag{A.5}$$

for some $\ell \in \mathbb{R}_+$ where the last inequality follows from the boundedness of the risk premium process. The above expression now implies that the local martingale M_π satisfies the Novikov condition and our proof is complete. □

Proof of Lemma 2. Let $\ell \in \mathbb{R}_+$ denote a uniform bound on the Euclidean norm of the risk premium process. Coming back to the definition of the driver in (17) it is easily seen that the mapping $\hat{h}$ is almost surely continuous in (t,x,z). On the other hand, an easy computation shows that

$$\left|\hat{h}\big(t,x,z\big)\right| \leq 2|x|\,\|\xi(t)\|^2 \leq 2\ell|x| \tag{A.6}$$

holds almost surely for all (t,x,z) and it now only remains to apply the result of Lepeltier & San Martin (see Theorem B.3 in Appendix B) to obtain the existence of maximal solution with the desired properties. □

Proof of Proposition 3. The solution to (17) being continuous, the process $\hat{\pi}$ is locally bounded. Applying Itô's lemma, we have that the solution $F \equiv F_{\hat{\pi}}$ to the stochastic differential equation (4) is given by

$$D(t)F(t) = \exp\left(-\frac{1}{2}\int_0^t \|\sigma(s)^*\hat{\pi}(s)\|^2\,ds + \int_0^t \hat{\pi}(s)^*\sigma(s)dB_0(s)\right). \tag{A.7}$$

The process DF is thus a non negative local martingale under the risk neutral probability measure and it now follows from the non negativity of the discount factor that $\hat{\pi}$ is an admissible fund portfolio process. Using (17) in conjunction with the uniqueness of the solution to (15) and the maximality of X in the set of solutions to (17) we obtain that

$$\hat{X}(t) \leq X_{\hat{\pi}}(t) \leq X(t) \tag{A.8}$$

holds almost everywhere. On the other hand, the definition of the map $\hat{h}$ and Theorem B.4 show that we have $X \leq X_\pi$ for every admissible fund portfolio process and taking essential infimum over $\pi \in \Theta_f$ on both sides we conclude that equality holds in (A.8). Using this fact in conjunction with (12) and (13) it is now easily seen that

$$\begin{aligned} v_{\hat{\pi}}(0) &= W_i(0) - \mathbf{E}^0 D(T)W_i^{\hat{\pi}}(T) = W_i(0)\big(1 - X_{\hat{\pi}}(0)\big) = W_i(0)\big(1 - \hat{X}(0)\big) \\ &\geq W_i(0)\big(1 - X_\pi(0)\big) = W_i(0) - \mathbf{E}^0 D(T)W_i^\pi(T) = v_\pi(0) \end{aligned} \tag{A.9}$$

holds for every admissible fund portfolio process. This implies that $\hat{\pi}$ constitutes the optimal fund portfolio process and our proof is complete. □

Proof of Corollary 1. Assume that the risk premium satisfies the conditions of the statement and consider the candidate solution $(X,0)$. The deterministic function $X : [0,T] \to (0,1]$

satisfies the prescribed terminal condition. On the other hand, differentiating (1) and comparing the result with (17) we obtain

$$\begin{aligned} -dX(t) &= -\tfrac{1}{4}X(t)\,\|\xi(t)\|^2\,dt \\ &= \Big(\hat{h}\big(t, X(t), Z(t)\big)dt - Z(t)^* dB_0(t)\Big)\Big|_{Z(t)=0} \end{aligned} \tag{A.10}$$

and it follows that X belongs to the set of solutions to (17). The definition of the mapping f and (16) then show that X coincides with the unique solution $X_{\tilde{\pi}}$ to (15) it now only remains to observe that by Theorem B.4 the maximal solution to (17) is dominated by $X_{\tilde{\pi}}$ to complete the proof. □

B Backward Stochastic Differential Equations

In this appendix, we introduce the notion of any adapted solution to a backward stochastic differential equation (BSDE) driven by a $n-$dimensional Brownian motion and briefly review the different results that we shall be using throughout the paper. In order to simplify the presentation, and since results for the vector valued case will not be used in the text, we restrict ourselves to the real valued case. The research monograph edited by El Karoui and Mazliak (1997) provides a comprehensive survey of most mathematical aspects of BSDEs. Examples of applications of these results in mathematical finance can be found in Duffie and Epstein (1992), El Karoui et al. (1997,2000) , Chen and Epstein (2001) and Collin Dufresne and Hugonnier (2000) among others.

B.1 An Existence Result

Throughout this appendix, we work on a probability space $(\Omega, \mathcal{F}, P)$ similar to that of Section 2. On this stochastic basis, a solution to the BSDE is a pair of progressively measurable processes with value in $\mathbb{R} \times \mathbb{R}^n$ such that

$$-dX(t) = f\big(t, X(t), Z(t)\big)dt - Z(t)^* dB(t), \qquad X(T) = \xi. \tag{B.1}$$

The progressively measurable mapping f is referred to as the *driver* and the measurable random variable ξ is called the *terminal condition*. Before giving sufficient conditions for the existence and uniqueness of the solution to (B.1) let us first fix some notation. For every $p \in [1, \infty)$ we denote by $\mathbb{H}_d^p$ the set of progressively measurable, $\mathbb{R}^d-$valued processes h with

$$\|h\|_p \equiv \left(\mathbf{E}\int_0^T \|h(t)\|^p\,dt\right)^{1/p} < \infty \tag{B.2}$$

and let $\mathbb{H}_d^\infty$ denote the set of progressively measurable, $\mathbb{R}^d-$valued processes whose Euclidean norm is essentially bounded. Using these notations, we now define the notion of *standard parameters* for the BSDE (B.1).

Definition B.1 *The data* (ξ, f) *are said to be standard parameters if* (i) *the progressively measurable process* $f(\cdot, 0, 0)$ *belongs to the space* $\mathbb{H}_1^2$*,* (ii) *we have*

$$\big|f(t, x, z) - f(t, \bar{x}, \bar{z})\big| \leq K\Big(\,\big|x - \bar{x}\big| + \big|z - \bar{z}\big|\,\Big) \tag{B.3}$$

almost surely for all t*,* (x, z)*,* $(\bar{x}, \bar{z})$ *and some strictly positive constant* K *and* (iii) *the terminal condition* ξ *is a square integrable random variable.*

Theorem B.1 *Given a pair* (ξ, f) *of standard parameters, there exists a unique solution* (X, Z) *to the* BSDE (B.1) *in the product space* $\mathbb{M} := \mathbb{H}_1^2 \times \mathbb{H}_n^2$.

Proof. This result was originally proved by Pardoux and Peng (1990). Various extensions and refinements of this existence result can be found in the research monograph edited by El Karoui and Mazliak (1997). □

B.2 The Linear Case

In this section we specialize the above existence result to the case where the driver is linear in (x,z). Let (γ,φ,β) denote a triple of progressively measurable bounded processes with values in the product space $\mathbb{R}\times\mathbb{R}\times\mathbb{R}^n$ and consider the BSDE (B.1) with driver given by

$$f\big(t,x,z\big)\equiv\gamma(t)-\varphi(t)x-\beta(t)^*z. \tag{B.4}$$

Such equations are referred to as *Linear* backward stochastic differential equations (LBSDEs) and admit an explicit solution as we now demonstrate.

Proposition B.1 *If* $(0,\xi)$ *are standard parameters, then the unique solution to the* LBSDE *with terminal condition* ξ *and driver* f *is explicitly given by*

$$X(t)=\frac{1}{H(t)}\mathbf{E}_t\left[H(T)\xi+\int_t^T\gamma(s)H(s)ds\right] \tag{B.5}$$

where the adjoint process H *is defined as the unique solution to the linear forward stochastic differential equation*

$$dH(t)=-H(t)\Big[\varphi(t)dt+\beta(t)^*dB(t)\Big],\qquad H(0)=1. \tag{B.6}$$

In particular, if the random variable ξ *and the process* γ *are non negative, then* X *is a non negative process and if moreover* $X(0)=0$ *then we have* $\xi\equiv 0$ *almost surely as well* $X(t)=\gamma(t)\equiv 0$ *almost everywhere.*

Proof. See the proof of Proposition 2.4 in Chapter 1 of the research monograph edited by El Karoui and Mazliak (1997). □

As an immediate consequence of the above proposition we obtain the celebrated comparison theorem which was first stated by Peng (1992). This result holds only in the one dimensional case and plays a role similar to that of the maximum principle in the theory of partial differential equations.

Theorem B.2 *Let* (ξ_k,f_k) *be standard parameters, assume that* $(\xi_1,f_1)\le(\xi_2,f_2)$ *holds almost everywhere and denote by* (X_k,Z_k) *the corresponding solutions. Then the inequality* $X_1\le X_2$ *holds almost everywhere.*

Proof. See the proof of Theorem 2.5 in Chapter 1 of the research monograph edited by El Karoui and Mazliak (1997). □

B.3 Optimization of a Family of BSDEs

Let us now suppose that we are given a set $\mathcal{A}$ of progressively measurable processes α and a family $\{(\xi_\alpha,f_\alpha)\}$ of standard parameters. Since it gives very simple conditions for the solution of a backward equation to dominate that of another, the comparison Theorem B.2 makes BSDEs a practical tool for the study of non Markovian stochastic control problems.

Corollary B.1 *If there exists a process* $\hat{\alpha} \in \mathcal{A}$ *such that* $(\xi_\alpha, f_\alpha) \le (\xi_{\hat{\alpha}}, f_{\hat{\alpha}})$ *holds almost everywhere for all* $\alpha \in \mathcal{A}$ *and if the parameters associated with the process* $\hat{\alpha}$ *are standard, then we have* $X_\alpha \le X_{\hat{\alpha}}$ *almost everywhere for all* $\alpha \in \mathcal{A}$.

The difficulty in applying the above result to solve a particular stochastic control problem is that even if the set $\mathcal{A}$ is well behaved enough for the data (ξ_α, f_α) to be standard parameters, the *optimal control* might not possess this property and the corresponding BSDE might not even admit an adapted solution. In particular, the above result does not apply to the stochastic control problem studied in Section 3.2.2 since the driver under consideration fails to satisfy the requirements of Definition B.1.
The next theorem extends the results of the previous subsections by allowing for a driver with superlinear growth in x and quadratic growth in z. Note that because we lessen the conditions imposed on the driver, we are forced to strengthen our integrability assumption on the terminal condition and that uniqueness of the solution is lost. In the following statement, we say that the pair (X, Z) of adapted processes is a *maximal solution* to the BSDE if its trajectory dominates that of any other solution.

Theorem B.3 *Assume that* ξ *is a bounded random variable, that* f *is continuous in* (t, x, z) *and that there exist non negative constants* (C, D, E) *such that*

$$\left|f\big(t, \omega, x, z\big)\right| \le C + D|x| + E\left\|z\right\|^2 \tag{B.7}$$

holds almost surely for all $(t, x, z) \in [0, T] \times \mathbb{R} \times \mathbb{R}^n$. *Then the* BSDE *with parameters* (ξ, f) *admits a maximal solution in the product space* $\mathbb{H}_1^\infty \times \mathbb{H}_n^2$.

Proof. For the case of a driver bounded in x and with quadratic growth in z, this result was proved by Kobylanski (1997) who also finds sufficient conditions for uniqueness of the solution. The version presented here is from Lepeltier and San Martin (1998) to which the reader is referred. □

Our last result in this appendix is also taken from the paper by Lepeltier and San Martin who obtain it as a by product of the method they use to prove Theorem B.3. It establishes a comparison result for the solution to the BSDE and was used in the resolution the stochastic control problem associated with the determination of the equilibrium fund portfolio process in Section 3.2.2.

Theorem B.4 *Let* h *denote a progressively measurable mapping,* θ *be a bounded measurable random variable and assume that* (ξ, f) *satisfy the conditions of Theorem* B.3. *If* $(\xi, f) \le (\theta, h)$ *holds almost surely then any bounded solution to the backward equation associated with the pair* (θ, h) *dominates the maximal solution to the backward equation associated with the pair* (ξ, f).

Proof. See the proof of Corollary 2 in Lepeltier and San Martin (1998). □

Example B.1 In some special cases, the BSDEs covered by Theorem B.3 admit explicit solutions. For example, the equation with driver given by $\beta\left\|z\right\|^2$ for some constant $\beta \neq 0$ and bounded terminal condition ξ admits a unique solution whose trajectory is given by

$$X(t) = \tfrac{1}{2\beta} \cdot \log \mathbf{E}_t\left[\exp\big(2\beta\xi\big)\right] \tag{B.8}$$

For another example, let θ denote an arbitrary bounded, predictable process with values in $\mathbb{R}^n$, define an equivalent probability measure by setting

$$P^{\theta}(A) \equiv \mathbf{E}\left[1_A \exp\left(-\int_0^T \tfrac{1}{8}\,\|\theta(t)\|^2\,dt + \int_0^T \tfrac{1}{2}\theta(t)^* dB(t)\right)\right] \tag{B.9}$$

and consider the backward equation with constant terminal condition $\xi = 1$ and driver given by $-(x/4)\,\|-z/x + \theta(t)\|^2$. Assuming that its maximal solution is strictly positive and applying Itô's lemma to the square root of this maximal solution, we obtain a candidate solution in the form

$$\big(X(t)\big)^{1/2} = \mathbf{E}_t^{\theta}\left[\exp\left(-\int_t^T \tfrac{1}{8}\,\|\theta(s)\|^2\,ds\right)\right] \tag{B.10}$$

where $\mathbf{E}_t^{\theta}$ denotes the conditional expectation operator under the probability measure of (B.9). It is now straightforward to check that the process X is strictly positive and constitutes the unique bounded solution to the BSDE under consideration. Other examples of explicitly solvable quadratic equations may be found in El Karoui and Rouge (2000).

References

[1] A. Bagchi. *Stackelberg Differential Games in Economic Models*, volume 64 of *Lecture Notes in Control and Information Sciences*. Springer Verlag, New York, 1984.

[2] J.B. Berk and R.C. Green. Mutual fund flows and performance in rational markets. *working paper, UC Berkeley*, 2002.

[3] P. Brémaud. *Point processes and queues: martingale dynamics*. Series in Statistics. Springer Verlag, New-York, 1981.

[4] Z. Chen and L. Epstein. Ambiguity, risk and asset returns in continous time. Working Paper, University of Rochester, May 2001.

[5] J. Chevalier and G. Ellison. Risk taking by mutual funds as a response to incentives. *Journal of Political Economy*, 105:1167–1200, 1997.

[6] P. Collin-Dufresne and J. Hugonnier. Event risk, contingent claims and the temporal resolution of uncertainty. June 2000. Working paper GSIA & Department of Mathematics, Carnegie Mellon University.

[7] D. Cuoco. Optimal consumption and equilibrium prices with portfolio constraints and stochastic income. *Journal of Economic Theory*, 72:33–73, 1997.

[8] D. Cuoco and R. Kaniel. General equilibrium implications of fund managers' compensation fees. *Working paper, Wharton school of business and University of Texas at Austin*, 1999.

[9] J. Cvitanič and I. Karatzas. Convex duality in convex portfolio optimization. *Annals of applied probability*, 2:767–818, 1992.

[10] S. Das and R. Sundaram. On the regulation of mutual fund fee structures. *working paper, NYU*, 1998.

[11] D. Duffie and L. Epstein. Stochastic differential utility. *Econometrica*, 60:353–94, 1992. With an Appendix by C. Skiadas.

[12] H. He and N. Pearson. Consumption and portfolio policies with incomplete markets and short-sale constraints: the infinite dimensional case. *Journal of Economic Theory*, 54:259–304, 1991.

[13] I. Karatzas. *Lectures on the mathematics of finance*, volume 8 of *CRM monograph series*. AMS, New York, 1997.

[14] I. Karatzas, J. P. Lehoczky, S. E. Shreve, and G. L. Xu. Martingale and duality methods for utility maximization in incomplete markets. *SIAM journal on control & optimization*, 29:702–730, 1991.

[15] I. Karatzas and S. E. Shreve. *Brownian motion and stochastic calculus*, volume 113 of *Graduate texts in Mathematics*. Springer Verlag, New York, 1991.

[16] I. Karatzas and S. E. Shreve. *Methods of mathematical finance*, volume 39 of *Applications of mathematics*. Springer Verlag, New York, 1998.

[17] N. El Karoui and L. Mazliak (editors). *Backward stochastic differential equations*, volume 364 of *Pitman research notes in Mathematics*. Addison Wesley Longman, 1997.

[18] N. El Karoui, M-C. Quenez, and S. Peng. Backward stochastic differential equations in finance. *Mathematical Finance*, 7:1–71, 1997.

[19] N. El Karoui and R. Rouge. Pricing via utility maximization and entropy. *Mathematical Finance*, 2000. To appear.

[20] M. Kobylanski. Résultats d'existence et d'unicité pour des équations différentielles stochastiques rétrogrades avec des générateurs à croissance quadratique. *Comptes Rendus de l'Académie des Sciences*, 324:81–86, 1997.

[21] D. Kramkov and W. Schachermayer. The asymptotic elasticity of utility functions and optimal investment in incomplete markets. *Annals of applied probability*, 1999.

[22] J-P. Lepeltier and J. San Martin. Existence for bsde with superlinear-quadratic coefficient. *Stochastics & Stochastics reports*, 63:227–40, 1998.

[23] A. Lynch and D. Musto. Understanding fee structures in the asset management business. *working paper, NYU*, 1997.

[24] E. Pardoux and S. Peng. Adapted solution of a backward stochastic differential equation. *Systems and control letters*, 14:55–61, 1990.

[25] S. Peng. A generalized dynamic programming principle and hamilton-jacobi-bellman equations. *Stochastics*, 38:119–34, 1992.

[26] S. E. Shreve and G. L. Xu. A duality method for optimal consumption and investment under short-sales prohibition. Part I: General markets & Part II: Constant market coefficients. *Annals of Applied Probability*, 2:87–112 and 314–328, 1992.

[27] E.R. Sirri and P. Tufano. Costly search and mutual fund flows. *Journal of Finance*, 53:1589–1622, 1998.