IS THE ACTIVE FUND MANAGEMENT INDUSTRY CONCENTRATED ENOUGH?

David Feldman, a Konark Saxena b and Jingrui Xu c,d

Abstract. We introduce a theoretical model of the active fund management industry (AFMI), where performance and size depend on AFMI’s competitiveness (concentration). Under plausible assumptions, as AFMI’s concentration decreases, so do fund managers’ incentives of exerting effort in search for alpha. Consequently, managers produce lower gross alpha, and rational investors, inferring lower expected AFMI performance, allocate a smaller portion of their wealth to active funds. Empirically, we find that a decrease in the U.S. mutual fund industry concentration over our sample period is associated with a decrease in its net alpha and size (relative to stock market capitalization).

JEL Codes: G10, G20, L10
Keywords: Active management, Mutual funds, Effort, Performance, Market concentration, Competition, Herfindahl-Hirschman index (HHI), Industry size, Net alpha, Equilibrium, Returns to scale, Learning

---

a Banking and Finance, UNSW Business School, UNSW Sydney, NSW 2052, Australia. Email: d.feldman@unsw.edu.au.
b Banking and Finance, UNSW Business School, UNSW Sydney, NSW 2052, Australia. Email: k.saxena@unsw.edu.au.
c Wang Yanan Institute for Studies in Economics (WISE), and Department of Finance, School of Economics, Xiamen University, China. Email: jingrui.xu@xmu.edu.cn
1 Introduction

An active area of research in financial economics examines the massive size of the active fund management industry (AFMI) and the high compensation of its managers, despite its unimpressive historical performance. In a pivotal study, Pastor and Stambaugh (2012; PS) present a model in which AFMI’s aggregate ability to outperform a passive benchmark decreases with its size. Intuitively, as more assets under management (AUM) chase opportunities, prices adjust, making opportunities harder to find. Using this model, they argue that the popularity of active management is not puzzling despite its poor track record under industry-level decreasing returns to scale (i.e., decreasing active managers’ ability to outperform passive benchmarks). PS’s findings lead to several interesting questions. For example, how do AFMI competitiveness and incentives of fund managers to exert costly effort in pursuit of fees affect the aggregate performance and size of the industry?2

We argue that fund managers’ effort in searching for mispriced assets depends on the degree of AFMI concentration (or competition3), which in turn influences AFMI performance and size. To formally analyze this relation, we introduce an AFMI equilibrium with a continuum of concentration levels, and allow active fund managers to exert costly effort, in search for mispriced assets, when competing over investment funds. Competing managers optimally choose fee and effort levels, resulting in a Nash equilibrium where they offer similar expected net alphas to investors.4 Mean-variance investors choose optimal portfolios of passive benchmarks and active funds (whose expected net alphas are positive5 to compensate for active funds’ additional risk). In this equilibrium, we study the impact of market concentration on managerial costly effort, net alpha production, and AFMI size, all three endogenously determined.

2 Guercio and Reuter (2014) find evidence consistent with the notion that weaker incentives due to lower competition faced by broker-sold funds leads to their lower performance. Other empirical studies relating mutual fund performance to competition include Wahal and Wang (2011) and Hoberg, Kumar, and Prabhal (2017).
3 We use concentration and competition as opposites.
4 We interpret managers with “lower cost” (for managing the same AUM) as ones with higher skills. In equilibrium, higher manager skills correspond to funds with higher AUM, but the same expected net alpha. See Section 2.3.2.
5 Note that, while AFMI net alphas are positive in our model with risk-averse investors, aggregate net alphas are zero-sum by construction, as they shift wealth between AFMI investors and other (unmodeled) investors. These unmodeled investors could be, for example, individuals with direct equity ownership [see Stambaugh (2014)].
A key quantity that determines the AFMI equilibrium is the “direct benefits” of effort defined as the difference between productivity of managerial effort (e.g., due to opportunities to find mispriced assets) and managerial effort costs (e.g., due to wages and research costs). We find that if higher concentration increases “direct benefits” of effort, then, higher concentration induces higher equilibrium expected net alphas and larger AFMI size. As the level of AFMI concentration decreases (or competition increases), the benefits of fund managers’ effort to search for mispriced assets decreases. As a result, they reduce their “effort” to find mispriced assets. That is, managers reduce their research efforts per dollar of AUM, and thereby are less informed and hold less (informed) active positions. This results in the funds producing lower gross alpha and providing lower net alpha (net of management fees) to investors. Rational investors infer this lower expected performance of active funds and thereby reduce their allocation to AFMI till they are indifferent between putting an extra dollar in AFMI or the passive benchmark.

In summary, a decrease in AFMI concentration reduces the incentives of fund managers to exert effort, resulting in lower AFMI performance and a smaller AFMI size. Our model then has the following empirical predictions: AFMI net alphas, effort, and size increase in AFMI concentration. We evaluate these predictions for the U.S. active equity mutual fund industry. A decrease in the U.S. mutual fund industry concentration (or increase in competition) over our sample should be associated with a decrease in its size and alpha. We find that both mutual fund industry size and net alphas, on average, decrease with measures of AFMI concentration such as the Herfindahl-Hirschman index (HHI). While effort is largely unobservable, we suggest that aggregate managerial effort affects the average AFMI active share and tracking error. This is because any effort to outperform the benchmark must involve taking positions that are different from the benchmark [e.g., Cremers and Petajisto (2009)]. We find that both the average AFMI active share and tracking error increase in concentration, consistent with our model (under the assumption that these measures proxy for effort).

---

6 The technical conditions for such an equilibrium, are first, that higher concentration increases the gap between the marginal benefits and marginal costs of gross alpha production, and, second, that the (further) sensitivity of this gap to increased effort does not reverse this property. See conditions (3) and (5) of Proposition 1 below.

7 While in the real-world, AFMI concentration is likely to be affected by other forces that we do not model (e.g., macroeconomic, regulatory), for convenience and parsimony, we assume exogenous concentration levels in our baseline model. In addition, however, we examine how our main empirical measure of concentration, HHI, can be endogenously determined (see Section 2.4).

8 Cremers and Petajisto (2009) show that fund-level active share predicts fund performance, and that this performance is strongly persistent. Brown and Davies (2017), argue that shirking managers could “jam the signal” in active share by taking uninformed bets to increase their perceived active share, generating a false sense of truly-active management. However, such “signal jamming” behavior is more likely to be an issue if a measure of active
Our analysis is related to several recent papers, though none of them address the question of how AFMI concentration levels impact its performance and size. Our model incorporates features from both Berk and Green (2004; BG) and PS. It is most directly related to PS, who study a framework with decreasing returns to scale, where interactions between active fund managers and investors determine expected net alphas. While they model and compare expected net alpha and industry size in a fully competitive equilibrium with those in a monopolistic equilibrium, that is, a comparison between two extreme “regimes,” we model and study the trade-off across a continuum of market concentration levels. Our model becomes the one in PS if neither AFMI concentration nor managers’ effort affect managers’ search productivity or costs. Even where, in our model, the search productivity for mispriced assets depends on effort, there is a special case of parameter values that leads to a solution where optimal allocated effort is zero. In this case as well, our model results become as those in PS.

We note that the concentration-alpha relation, in our model, is a mechanism distinct from, and additional to, the industry size-alpha relation in PS. Specifically, in our model, the concentration-alpha relation exists also when controlling for industry size. This is the case because concentration affects the optimal level of effort by affecting effort productivity and the cost of effort. And, in turn, a change in the optimal effort by managers leads to a change in expected net alpha. We note that effort and its dependence on concentration is a key feature of our model compared to that of PS. We introduce another feature, fund-level decreasing returns to scale, mainly to facilitate a range of interesting equilibrium concentration levels by allowing for differences in equilibrium fund sizes. It might follow to ask, what if we shut-down our effort channel and introduce (only) fund level decreasing returns to scale to the PS framework? In this case, any effect of fund level decreasing returns to scale on net alpha will be via industry size. Concentration will not influence net alpha (controlling for industry size). Introducing concentration and its influence on effort enables us to model a distinct mechanism that influences net alpha. (For details, please see Section 2.3)

9 Analytically, effort does not affect alpha production in our model, if the third addend of the right side of Equation (7) does not exist, and if we abandon our cost function, Equation (18), in favor of defining funds’ fees to be net of funds’ management costs.

10 For the technical conditions that lead to this case, see Proposition PS and its Corollary. Intuitively, this is the case if, for all concentration levels, costs of efforts producing alphas exceed the benefits of the produced alphas.

11 The effect of such fund heterogeneity via industry size on net alpha is already captured parsimoniously and elegantly by the PS model.
Our model is also related to the BG model. Their equilibrium is compatible with the case in our model where infinitely many small risk-neutral investors compete, and the size of the fund endogenously adjusts to make the gross alpha equal the fee, so that expected net alpha is always zero. Even in this case, there are some significant differences between the models. First, in BG, managers “ability” corresponds to the magnitude of expected excess return (over the passive benchmark) earned on the first dollar actively managed by a fund. However, a manager with more “ability” does not necessarily have larger AUM because they are indifferent between managing a large AUM with a small fee or small AUM with a large fee (as long as their profits stay the same). In our model, competing managers, in equilibrium, optimally charge break-even fees (sufficient to cover costs). Further, managers with more “ability” have larger equilibrium AUM, where “ability” in our model refers to a lower rate at which fund-level returns to scale decrease. (PS’s model identifies the AFMI equilibrium elegantly, without the need to specify fund level size heterogeneity.) Second, in our model, AFMI size depends on AFMI concentration even when AFMI expected net alpha is zero due to competition between infinitely many small risk-neutral investors. This is because higher AFMI concentration incentivizes managers to invest more effort for finding mispriced assets, leading to an increase in the AFMI size.

Recent empirical studies of Pastor, Stambaugh, and Taylor (PST; 2015) find that industry size relates negatively to performance, and of Berk and van Binsbergen (2015) find that most of the growth in the mutual fund industry is due to the growth in the number of funds, not in the median fund size. Interestingly, put together, these empirical findings seem to support our theoretical predictions that performance increases in concentration. Our model, however, provides, in this context, more elaborate predictions. First, we demonstrate that, even when controlling for size, higher concentration levels are associated with increase in alphas. Second, we make predictions of concentration level effects on size as well as second order effects: AFMI expected net alphas and AFMI size are both either concave or convex in market

12 Fund-level decreasing returns scale does not necessarily imply a correlation between fund size and net alpha [Berk and Green (2004)]. In our equilibrium, funds that have lower decreasing returns to scale parameters endogenously have larger sizes till the point where all funds have the same net alpha [In the Berk and Green (2004) equilibrium fund managers with lower decreasing returns to scale parameters are indifferent between larger sizes or higher fees]. While Pastor, Stambaugh, and Taylor (2015) find strong evidence of net alpha decreasing returns with size at the industry-level, they do not find a significant relation at the fund-level. The fund-level evidence is mixed [see, for example, Grinblatt and Titman (1989), Chen, Hong, Huang, and Kubik (2004), Ferreira, Keswani, Miguel, and Ramos (2013a, 2013b), Yan (2008), and Reuter and Zitzewitz (2013)]. In our empirical analysis, consistent with our model, we do not find a significant relation between fund net alpha and fund size.
Section 2 develops the theoretical model, Section 3 presents the empirical methods and results, and Section 4 concludes.

2 Theoretical Framework

Within PS’s world, adopting their notation, we develop a theoretical framework for modeling the effect of AFMI concentration on fund managers’ effort, fund fees, fund performance, AFMI size, and potential direct benefits. We offer below a presentation of the setting and results.

2.1 Setting

For brevity and parsimonious notation, we assume that variables and functions are real, continuous, and at least twice differentiable. Within a one-period market, there are two types of agents: fund managers of $M$ funds, $M > 1$, and $N$ investors, $N \geq 1$. Acting competitively, each manager sets a proportional management fee and chooses an effort level to maximize the fund expected net alpha to attract investments. In this sub-section and sub-section 2.2., we consider the case when infinitely many small mean-variance risk-averse investors allocate their investments to maximize their portfolios’ Sharpe ratios. By infinitely many small investors we mean that $N \to \infty$, and with investors’ finite wealth their choices do not affect fund sizes. We also consider the case of infinitely many small risk-neutral investors.

Our model follows and builds on that of PS. In this partial equilibrium, the passive benchmark portfolio’s returns are exogenously given and are unaffected by interactions between investors and managers. Managers’ outperformance of the passive benchmark portfolio (i.e., gross alphas), may come at the expense of “other investors,” who may be noise traders, liquidity seekers, misinformed, or irrational.

Fund alpha and returns process. Following PS, $r_F$, a vector of $M$ funds’ returns in excess of the riskless rate that investors receive, follows the regression model

$$ r_F = \alpha + \beta r_p + u, $$

where $r_F$ is an $M \times 1$ vector with elements $r_{Fi}$, $i = 1, \ldots, M$. $\alpha$, $\beta$, and $u$ are $M \times 1$ vectors.

Our model predicts that if equilibrium fund expected net alphas are concave in market concentration, then AFMI’s benefits of effort are concave in market concentration. Consequently, equilibrium AFMI size is also concave in market concentration. On the other hand, if equilibrium AFMI size is convex in market concentration, then the AFMI’s benefits of effort are convex in market concentration and, consequently, equilibrium expected fund net alphas are convex in market concentration. Please note that the order of statements in the second (convex) case is different from that in the first case, for reasons explained following Proposition RA3 below.

Please see the detailed discussion in PS, pp. 749.
α is the vector of fund net alphas received by investors, and β is the vector of fund betas. The scalar \( r_p \) is the excess return on the passive benchmark portfolio, with mean \( \mu_p \) and variance \( \sigma_p^2 \): u is the residual vector, with elements that follow

\[
u_i = x + \varepsilon_i, \quad i = 1, \ldots, M,
\]

where \( \varepsilon_i \)'s are mean zero and variance \( \sigma^2 \) idiosyncratic risks, and are uncorrelated with each other, with \( x \), and with \( r_p \). The common factor \( x \) has mean zero, variance of \( \sigma^2_x \), and is uncorrelated with \( r_p \). The values of \( \mu_p, \sigma_p^2, \sigma_x^2 \), and \( \sigma^2_x \) are strictly positive constants that are common knowledge to investors and managers.

The benchmark-adjusted returns on the \( M \) funds that investors receive is

\[ r = \alpha + u. \] (3)

As in PS [see, there, Equations (2) and (3)], the factor structure in Equations (1)-(3), means that the benchmark-adjusted returns of AFMI funds are correlated. An economic rationale for a common component \( x \) in this factor structure is that similar opportunities are likely to be identified by AFMI funds, resulting in correlated benchmark-adjusted returns\(^\text{15}\) (see also PS, pp. 746-747). Technically, this common component \( x \) is necessary to guarantee that investors in AFMI portfolios cannot enjoy expected net alphas without increasing their risk (the variance of their portfolios). (This is the case because they can, plausibly, well diversify the \( \varepsilon_i \)'s in their AFMI returns.) That is, had the common component, \( x \), not existed, risks associated with investing in AFMI funds could be fully diversified away by investing in many of them, while, retaining the benefits of their positive expected net alphas.

Each element in \( \alpha \) has the following structure:

\[ \alpha_i = a - b \frac{S}{W} + A(e_i; H) - f_i, \] (4)

where \( a \) and \( b \) are positive, unknown scalar parameters; \( S \) is the aggregate size of the active management industry and is equal to the sum of all the funds’ sizes (i.e., \( S = \sum_{i=1}^{M} s_i \)); \( W \) is the total wealth managed actively and passively and is equal to \( S \) plus the amount invested in the passive benchmark; \( A(e_i; H) \) is the productivity of manager \( i \)'s proportional effort \( e_i \in [0, \infty) \)

\(^{15}\) For example, Garvey, Kahn, and Savi (2017) decompose fund strategies into a combination of orthogonal and generic insights, and suggest that many funds invest partly in orthogonal insights and partly in generic insights that are common across funds. Investing in multiple fund managers acts to concentrate risk into generic ideas.
to increase gross alpha under AFMI concentration $H$; and $f_i$ is the proportional fee charged by manager $i$.

The expression for net alpha in PS corresponding to Equation (4) does not contain the $A(e_i; H)$ term, which captures the alpha production function due to extra effort under AFMI concentration $H$. This is because PS focus on studying how investor beliefs about the unknown parameters $a$ and $b$ influence AFMI size. We build on their findings and focus on studying how AFMI concentration influences fund managers’ incentives to exert costly effort, thereby influencing AFMI size and alpha. Next, we describe our assumptions about $A(e_i; H)$, $H$, and investor beliefs about $a$ and $b$.

**Productivity of manager effort.** We assume that $A(e_i; H)$, the productivity of effort under $H$, is the same across funds and has the following functional characteristics:

- non-negative,
- increasing and concave in effort,
- increasing in AFMI concentration,
- positive cross-partial derivatives with respect to effort and AFMI concentration.

The assumption that links concentration to gross alpha is that, the more concentrated AFMI is, the relatively more investment (mispriced) opportunities there are, and the more marginally efficient is the use of industry resources. Thus, managers can generate a higher increment in gross alpha for a given effort level $e_i$.

**AFMI concentration.** Our main analysis assumes that $H$ is a known exogenous scalar parameter because it depends mainly on some exogenous factors. For example, industrial organization theory suggests that AFMI concentration not only depends on the number of incumbents, but also on threats of entry, activity-limiting regulation, and the competitiveness of related industries [see, for example, Claessens and Laeven (2003)]. [In Section (2.4) below, we examine endogenous measures of AFMI concentration.] Without loss of generality, we assume that $H \in [0,1)$. If $H = 0$, there are infinitely many small funds in the market, and the market is fully competitive. If $H = 1$, the market is monopolistic. If fund managers are competing (the case we consider), $H$ belongs to $[0,1)$.

---

16 In a more concentrated market, if a fund manager controls most of the industry resources and develops advanced strategies to produce gross alphas, other funds can mimic this fund’s strategy and produce higher gross alphas given a particular effort level, so this assumption is still valid when a dominant fund in the market controls the majority of resources.
Expected alpha and investor’s information about unknown parameters. As noted earlier, the parameters $a$ and $b$ in Equation (4) are positive, unknown scalar parameters. The parameter $a$ represents the expected return on an initial small fraction of wealth invested in active management, net of any proportional costs. The parameter $b$ is the absolute magnitude of the decreasing returns to scale at industry level. As in PS, the first and second conditional moments of $a$ and $b$ are

$$E\left(\begin{bmatrix} a \\ b \end{bmatrix} \mid D\right) \triangleq \begin{bmatrix} a \\ \hat{b} \end{bmatrix},$$

(5)

$$\text{var}\left(\begin{bmatrix} a \\ b \end{bmatrix} \mid D\right) \triangleq \begin{bmatrix} \sigma_a^2 & \sigma_{ab} \\ \sigma_{ab} & \sigma_b^2 \end{bmatrix},$$

(6)

where $D$ denotes investors’ information set.

As we do not focus on $\sigma_{ab}$’s effects on the equilibrium, we assume that $\sigma_{ab} = 0$. In other words, conditional on current information, we assume that how $a$ deviates from $a$ is unrelated to how $\hat{b}$ deviates from $b$. Finally, with $f_i$ being a proportional management fee charged by manager $i$, her fund’s expected net alpha is

$$E(a_i \mid D) = a - \hat{b}\frac{S}{W} + \hat{A}(e_i; H) - f_i.$$  

(7)

Investor’s Problem. Let $\delta_j$ denote the $M \times 1$ vector of weights that investor $j$ places on the $M$ funds, with elements $\delta_{ji}$, $i = 1,\ldots,M$. Thus, investor $j$’s excess return is

---

17 Please see PS, Equations (5) and (6), p.747.

18 We assume that $\sigma_{ab} = 0$, but we note that the value of $\sigma_{ab}$ affects the equilibrium results because it affects portfolio risks. If $\sigma_{ab}$ (in absolute value) is large relative to other risk sources, such as $\sigma_a^2$, $\sigma_b^2$, and $\sigma_e^2$, changes in investors’ wealth allocations to funds, would induce changes in their portfolio risks, affecting in turn their optimal demands. This would make our theoretical results in propositions RA3 and RA4 more complex. We believe that consequences of such an analysis would not be directly material to the issues that we explore here and would obfuscate the analysis. We, thus, assume that precisions of estimates of $a$ and $b$, conditional on current information, are not closely related, making $\sigma_{ab} \rightarrow 0$.

19 Investors observe the passive benchmark and the AFMI funds’ returns. The difference between these returns comes from three components: net alphas, the common risk factor, and idiosyncratic risks. As the distributions of the common risk and idiosyncratic risk are common knowledge, investors know the likelihood function of the net alphas. Given prior beliefs of net alphas, they form posteriors and update their beliefs. In our one-period model, there is no dynamic Bayesian updating, but we suggest that investors reached a fixed-point equilibrium. Further, because investors observe $f_i, H, S$ and $W$, they can also infer $A(e_i, H)$. Here, where equilibrium optimal effort levels of all managers are the same, the estimate $\hat{A}(e_i, H)$ could be subsumed in $\hat{a}$; and in equilibria where managers’ optimal effort levels differ, the estimates $\hat{A}_i(e_i, H)$, could be subsumed in $f_i$. For simplicity and brevity, we depress the notation of $\hat{A}(e_i, H)$ in favor of $A(e_i, H)$ and follow PS formulation.
\[ r_j = \delta_j^T \mathbf{r}_F + (1 - \delta_j^T \mathbf{t}_M) r_p, \] (8)

where \( \mathbf{t}_M \) is an \( M \times 1 \) vector with elements equal to 1. Following PS (see there, p. 750 and Footnote 7), we assume that all funds have beta loadings on the benchmark equal to 1 (i.e., \( \beta \), as defined in Equation (1), fulfills, \( \beta = \mathbf{t}_M \)). With funds’ holding unit beta portfolios, the choice variable \( \delta_j \) represents investors \( j \)'s exposure to the active part in AFMI in excess of her holding of the passive benchmark portfolio. As in PS, this assumption allows parsimonious modelling of the active/passive choice.

Based on Equations (1) and (8), we have

\[ r_j = r_p + \delta_j^T (\mathbf{a} + \mathbf{u}). \] (9)

Further, we have

\[ \mathbb{E}(r_j | D) = \mu_p + \delta_j^T \mathbb{E}(\mathbf{a} | D), \quad \forall j, \] (10)

and,

\[ \text{Var}(r_j | D) = \sigma_p^2 + \left[ \sigma^2_\alpha + \sigma^2_\zeta + \sigma^2_\theta \left( \frac{S}{W} \right) \right]^2 \left( \delta_j^T \mathbf{t}_M \right)^2 + \sigma^2_\delta (\delta_j^T \mathbf{d}_j), \quad \forall j. \] (11)

We focus here on the case of infinitely many (i.e., \( N \to \infty \)) small mean-variance risk-averse investors, none of them can affect fund sizes. Also, investors’ investment in the AFMI dilutes funds’ expected returns due to decreasing returns to scale in funds. In addition, mean-variance risk-averse investors face risk-return tradeoffs in marginal allocations. Investor \( j \)'s objective is to maximize the portfolio’s Sharpe ratio by choosing portfolio weights, \( \delta_j, \ j = 1, \ldots, M \).

\[ \max_{\delta_j} \frac{\mathbb{E}(r_j | D)}{\sqrt{\text{Var}(r_j | D)}}, \] (12)

subject to

\[ \delta_j^T \mathbf{t}_M \leq 1, \] (13)

\[ \delta_{j,i} \geq 0, \quad \forall i. \] (14)

Note that the argument of the objective function in Equation (12) is the ratio of Equations (10) and (11). Condition (13) is a form of wealth constraint, saying that investors

Note that Equations (8) and (9) here are similar to Equations (10) and (11) in PS. Note however, that our functional forms, represented by variables in these equations (such as \( \alpha, \mathbf{r}_F, \delta_j, r_p \)), are different.
cannot borrow from the passive benchmark to invest in the AFMI. Condition (14) says that there is no short sale of funds. Also, as we assume that there are no marginal diversification benefits across funds, we set the idiosyncratic risk of investor j’s portfolio, \( \sigma^2 \delta_j^T \delta_j \) to zero when solving the optimization problem (12). Because the equilibrium is symmetric, we have

\[
\delta_j^T t_M = S/W, \quad \forall j. 
\]  

(15)

**Fund Managers’ Problem.** As noted earlier, \( f_i \) is the proportional fee charged by manager \( i \). The manager sets this fee considering its effect on the fund’s size. The manager also chooses the level of costly proportional effort to exert in order to find mispriced asset and produce additional gross alpha using \( A(e; H) \), which depends on AFMI concentration \( H \). We define manager \( i \)’s average (per-dollar) cost to produce alpha as \( C'(e, s_i; H) \). Therefore, manager \( i \)’s economic profit is

\[
s_i \left( f_i - C'(e, s_i; H) \right),
\]

and for fund \( i \) to survive,

\[
f_i - C'(e, s_i; H) \geq 0.
\]

We assume that average cost functions, \( C'(e, s_i; H) \), contain three independent positive scalar components: \( c_{0,i} \), the average cost for fund \( i \) to operate in the market before receiving investment and before manager \( i \) spends effort; \( c_{1,i} s_i \), the average cost related to fund \( i \)’s size, \( s_i \); and \( c_{2,i}(e; H) \), the average cost of manager \( i \)’s effort under a particular AFMI concentration.\(^{22}\) That is,

\[
C'(e, s_i; H) = c_{0,i} + c_{1,i} s_i + c_{2,i}(e; H).
\]

Equation (18) is also manager \( i \)’s per dollar cost function, which, when multiplied by her fund size, \( s_i \), gives her total cost function. The coefficient \( c_{1,i} \), then, induces a nonlinear (quadratic) increase in manager \( i \)’s total cost function, making it convex in \( s_i \), and representing the extent of decreasing returns to scale in funds gross alpha production. This fund

\[^{21}\text{Here, too, we adopt PS notation. Note that } \sigma^2 \delta_j^T \delta_j = \sigma^2 \delta_j^T I \delta_j \text{ where } I \text{ is an } M \times M \text{ identity matrix, where } \sigma^2 I \text{ stands for the covariance matrix.}\]

\[^{22}\text{To simplify our model, we assume there is no interaction between effort and size in the average cost function because it is unlikely that fund size affects managers’ per dollar effort. We also assume that there is no interaction between concentration and size in the average cost function because it is unlikely that concentration affects managers’ average cost sensitivities to fund sizes. Nevertheless, even if these interacting effects exist, they tend to be small in comparison to effects of other terms in the average cost function.}\]
cost model is consistent with that of BG, who assume decreasing returns to scale at fund level.

Simplifying, we assume that $c_{0,i}$’s, and $c_2(e_i; H)$’s are the same across funds (we, thus, drop the subscript $i$), but that $c_{1,i}$’s are different across funds. (We discuss the effects of similarities and differences in these parameters across funds after Proposition 2.) We assume that the function $c_2(e_i; H)$ has the following characteristics:

- non-negative,
- increasing and convex in effort,

The average cost function implies that as fund $i$’s size, $s_i$, increases, manager $i$’s average cost increases because larger trades are associated with larger price impacts and higher execution costs and because of other factors that create diseconomies of scale in operation. $c_{1,i}$ is the average cost sensitivity to fund $i$’s size. Adding the three cost function components, we get that the average cost function is increasing and convex in effort.

We do not specify whether costs are increasing or decreasing in concentration. Where costs are decreasing in concentration, the advantages of higher concentration is two-fold: more opportunities and lower costs. Where costs do not change as a function of concentration, the advantage of increasing opportunities due to increase in concentration is left unmitigated. Moreover, we show in Proposition 1 below that even increasing costs in concentration, for the plausible parameters set, may not fully mitigate the advantages of the increasing opportunities.

We assume no fixed costs for several reasons. First, fixed costs are lower in comparison to the costs that we model and, we believe, would not affect our analysis.\(^{23}\) Moreover, it is the larger funds, where fixed costs are relatively lower, that determine AFMI concentration. In addition, as we focus on modelling decreasing returns to scale in gross alpha production, positive fixed costs might obfuscate this property.\(^{24}\)

With these assumptions, manager $i$’s problem is

$$\max_{e_i, f_i} s_i \left( f_i - C'(e_i, s_i; H) \right)$$

subject to

$$e_i \geq 0,$$

\(^{23}\) Fixed costs to manage funds, such as registration fees and equipment expenditure, are usually small in comparison to variable costs related to employees’ salaries and managers’ compensation.

\(^{24}\) A non-zero fixed cost and decreasing returns to scale in gross alpha production (i.e., costs component that are increasing and convex in fund size) would induce an average cost function that is U-shape in fund size. Thus, under some cases this may induce instances of increasing returns to scale in gross alpha production.
$f_i \geq 0$.

**Information Structure.** We follow the information structure of PS where relevant, and extend it, in spirit, to the new model structure that we introduce here. Model parameters and functional forms are common knowledge to managers and investors, with the following exceptions. The values of $a$ and $b$, are unknown, but their first two moments specifications are common knowledge. The values of parameters of managers’ cost functions and alpha production functions are private information (manager $i$’s knows her cost and production functions). Sensitivities (assumptions on derivatives) of cost functions and alpha production functions are common knowledge.

### 2.2 Equilibrium

Having articulated the model’s setting and optimization problems we will now identify the AFMI equilibrium.

We commence by discussing why the manager’s optimization problem presented above is equivalent to the problem of maximizing their expected net alpha. The presentation of the latter problem helps to conveniently describe the AFMI equilibrium. Specifically, our risk-averse investors invest only in funds that offer the highest expected net alphas. Fund managers, in turn, compete over expected net alphas to attract investments. Manager $i$’s problem, then, becomes

$$\text{Max}_{e_i, f_i} E(\alpha_i \mid D)$$

subject to

$$f_i - C'(e_i, s_i; H) \geq 0, \quad e_i \geq 0, \quad \text{and} \quad f_i \geq 0.$$ 

See the Appendix, for proof of the managers’ problems equivalence. The proof intuition is as follows. Under competition, funds that offer higher expected net alphas draw (all) investments. Thus, in equilibrium, funds offer similar expected net alphas. The possibility (threat) that other managers increase their fund profits and sizes by improving expected net alphas, induces managers to maximize expected net alphas in order to “survive.” We note that this aspect of the equilibrium is similar to that in PS, but in addition to their result, we show that it holds also in the case of finite number of managers.

To further study the equilibrium, we define the “direct benefits” of effort function of manager $i$ as

---

25 For simplicity and brevity, we omit the condition in Equation (17) from the problem statement as it is implied by the optimization and, thus, is not necessary.
\( B(e_i; H) \doteq A(e_i; H) - c_2(e_i; H) \), \( \forall i \).

\( B(e_i; H) \) captures the direct benefit from effort exerted in active fund management, in terms of increase in gross alpha production minus the effort cost. We note that we should interpret benefits generally, allowing them to be positive or negative. Whether manager \( i \)'s marginal direct benefits of initial effort are positive [i.e., \( B_{e_i}(0; H) > 0, \forall H \)] or not is an important condition affecting the equilibrium. If this condition is not met, our equilibrium becomes the one in PS (see Proposition PS in Section 2.3). Whether the sensitivity of manager \( i \)'s direct benefits, at optimal effort, is positive [i.e., \( \frac{dB(e^*_i; H)}{dH} > 0 \)] or not, is also an important condition affecting the equilibrium.

We note that AFMI’s active search for net alphas might have indirect effects that we do not model here. It might drive security prices toward their true values; it might induce firms to improve governance and performance, and reduce agency costs. It might induce transfer of wealth from less productive firms/investors to more productive ones. As we discussed earlier, here, as in the literature, gross alphas are zero-sum. (See for example PS, pp. 748-750, including Footnote 6, and references therein, and our Footnote 4.) We note that this is the case irrespective of whether any manager’s direct and or indirect benefits are non-zero or zero.

The following proposition describes fund managers’ equilibrium optimal effort levels and fees.

**PROPOSITION 1.** For manager \( i, i = 1, \ldots, M \), if initial effort inputs generate positive direct benefits of effort [i.e., \( B_{e_i}(0; H) > 0, \forall H \)], equilibrium optimal effort-fee combinations \( (e^*_i, f^*_i) \) satisfy the following.

1. Fees are equal to costs.
   \[
   f^*_i - C^i(e^*_i, s^*_i; H) = 0, \forall i.
   \]

2. The impact of marginal efforts on gross alpha is set to be equal to the marginal average costs of effort, thus manager \( i \)'s marginal direct benefits of effort under the optimal effort level are zero).
   \[
   A_e(e^*_i; H) - c_2e(e^*_i; H) = B_{e_i}(e^*_i; H) = 0, \forall i.
   \]

3. Where concentration is higher, equilibrium optimal effort levels are higher (lower) if and only if higher concentration induces a larger (smaller) marginal effort impact on gross alphas than on costs. Or,
\( e^*_i(H) \geq 0(<0) \) iff \( A_{c_e,H}(e^*_i;H) - c_{2e_c,H}(e^*_i;H) \geq 0(<0) \), where \( e^*_i(H) \equiv de^*_i/dH \) (24)

4. Whether higher concentrations induce higher equilibrium optimal fees depends on whether they induce an increase in equilibrium industry sizes and whether they induce an increase in equilibrium optimal effort levels.

5. Where concentrations are higher, equilibrium manager \( i \)'s direct benefits of effort are higher (lower) if and only if higher concentrations induce a larger (smaller) impact on gross alphas than on costs.

6. In equilibrium, managers offer the same market competitive expected net alpha.

7. In equilibrium, managers offer the same market competitive Sharpe ratio.

**Proof of Proposition 1.** See the Mathematical Appendix.

The proof intuition is as follows. While competing for investments, managers maximize fund expected net alphas by choosing optimal effort levels and fees, earning zero economic profits (break-even fees) in equilibrium. The reason for the latter is as follows. If managers increase fees, they would lower fund expected net alphas and lose all investments. If managers decrease fees, they would become insolvent – incurring negative cashflows (costs higher than fees). Deviating from equilibrium effort level would also induce a loss of investments (if decreasing effort) or insolvency (if increasing effort).

If higher concentration induces a higher (lower) marginal effort impact on gross alphas than a marginal effort impact on costs, managers optimally choose higher (lower) effort levels in producing fund net alphas. If higher concentrations induce higher equilibrium optimal effort levels, managers’ costs are driven higher, resulting in higher break-even fees. In addition, higher concentrations have two effects on manager \( i \)'s direct benefits of effort. First, it directly affects the levels of gross alphas production function and of costs, \( A(e^*_i;H) \) and \( c_e(e^*_i;H) \), being a parameter of each of these functions. Second, it changes equilibrium optimal effort levels, consequently changing the levels of gross alphas and costs. In equilibrium, the latter effect is zero because the marginal effort impact on gross alphas is equal to the marginal effort impact on costs and the effect of higher concentration through effort on gross alphas is cancelled out by its effects through effort on costs. Therefore, if higher concentrations induce a higher direct impact on gross alphas than on costs, manager \( i \)'s direct benefits of effort are higher.

Also, as there are no diversification benefits across funds, managers who provide higher expected net alphas dominate, attracting investments. Consequently, their fund costs increase, inducing higher (break-even) fees and lowering expected net alphas. Thus, in equilibrium,
allocation of investments, or fund sizes, set expected net alphas to be equal across funds. If fund managers cannot produce the AFMI highest expected net alpha, even for an infinitesimal fund size, they lose all investments and go out of the market. In addition, as funds have the same expected net alphas, they have the same expected returns. As the source of fund returns’ variance is the same across funds, the fund return variance is the same across funds. Therefore, managers offer the same competitive Sharpe ratio.

The following proposition identifies the relation of different managers’ equilibrium optimal effort levels, fees, direct benefits of effort, and AFMI share.

**PROPOSITION 2.** Under the same \(c_0\) and the same functional form of \(c_2(e; H)\) but different \(c_{1,i}\)’s across funds,

1. Equilibrium effort levels and fees are the same across funds.
2. Pairwise relative fund sizes, \(s_i / s_j, \forall i, j\), are inversely proportional to their corresponding cost coefficients, \(c_{1,i} / c_{1,j}, \forall i, j\) (where \(c_{1,i}\) represents intensity of fund-level decreasing returns to scale in gross alpha production).
3. AFMI fund shares, \(s_i / S\) ’s are \(\frac{s_i}{S} = \left(c_{1,i} \sum_{j=1}^{M} (c_{1,j}^{-1})\right)^{-1}, \forall i\).
4. Equilibrium AFMI’s direct benefits of effort are the same across funds.

*Proof of Proposition 2.* See the Mathematical Appendix.

The second point of Proposition 2 shows that managers’ different costs of producing gross alphas, induce different fund sizes in equilibrium. The third point of Proposition 2 implies that funds’ market shares are deterministic functions of \(c_{1,i}\) and are, thus, unaffected by the AFMI weight in total wealth, \(S/W\). In other words, how investors weight the funds inside the AFMI is unaffected by how investors weight the AFMI as a whole relative to the passive benchmark. This property facilitates later results.

Proposition 2 is driven by the fact that \(c_0\) and the functional form of \(c_2(e; H)\) are the same across funds but \(c_{1,i}\) ’s are different across funds. In contrast, Proposition 1 is valid even without this assumption. In fact, from Proposition 1, we can see that if \(c_{0,i}, c_{1,i}\) and the functional form of \(c_2(e; H)\) are different across funds, managers end up with different levels of effort, different fees, and different fund sizes in equilibrium. If all fund managers have the
same $c_0$, $c_1$, and functional form of $c_2(e_i; H)$, they end up with the same equilibrium levels of effort, fees, and fund sizes.

Finally, as, according to the fourth point of Proposition 2, managers’ direct benefits are the same, we will refer, henceforth, to AFMI’s direct benefits, rather than to manager $i$’s direct benefits. We note that, consequently, previous results related to manager $i$’s direct benefits hold for AFMI in aggregate, that is, for AFMI’s direct benefits.

The next proposition describes the equilibrium.

**PROPOSITION RA1. Unique Nash Equilibrium.**

Where $N \to \infty$,

1. there exists a unique Nash equilibrium, $\{e^*, f^*, \delta^*\}$, where
   - $e^*$ is an $M \times 1$ vector with managers’ optimal effort allocations, $e^*_i$,
   - $f^*$ is an $M \times 1$ vector with managers’ optimal fee allocations, $f^*_i$,
   - $\delta^*$ is an $M \times N$ matrix with vectors of investors’ optimal wealth weights allocations to funds, $\delta^*_{ij}$;

2. in this equilibrium, managers produce the same expected net alpha, which drives their economic profits to zero, by charging only break-even fees; and investors allocate the same wealth proportions to each of the funds.

**Proof of Proposition RA1.** See the Mathematical Appendix.

For convenience in describing the equilibrium in the following propositions, we define the equilibrium optimal expected net alphas of an initial marginal investment in the AFMI (where $S = 0$) as $X(e^*_i, H)$. Quantitatively,

$$X(e^*_i, H) \equiv a + A(e^*_i; H) - \left[ c_0 + c_2(e^*_i; H) \right].$$

(25)

For AFMI to exist, we must have positive expected net alphas for initial infinitesimal investments into AFMI:26

$$X(e^*_i, H) > 0.$$  

(26)

If Inequality (26) is violated, investors receive no advantage in diverting funds from the passive index to AFMI and leave their wealth invested in the passive benchmark. Also, to offer meaningful results, we assume that initial marginal allocations of effort generate positive

---

26 The condition in Inequality (26) here is equivalent to the condition that $a > 0$ in PS. See PS, p. 747, for further discussion and insights.
AFMI’s direct benefits of effort, that is,
\[ B_c(i, H) > 0, \quad \forall i, \forall H, \]
such that the optimal effort \( e^*_i \) is positive, finite, and attainable, i.e.,
\[ B_c(e^*_i, H) = 0, \quad e^*_i < K, \quad \forall i, \forall H \]
for some positive constant \( K \). We focus on the case under Proposition 1 in the following analyses.

As in PS, the explicit analytic solutions for \( S/W \) are solutions of a cubic equation and are cumbersome. The following proposition presents the cubic equation and its corollary presents properties of its solution.

**PROPOSITION RA2. Equilibrium Optimal Allocations.**

For \( i = 1, 2, \ldots, M \), we have \( E(\alpha_i \mid D)\big|_{e^*, r^*, s^*} > 0 \); and where \( N \to \infty \), the equilibrium optimal \( S/W \) is either 1 or a real positive solution (smaller than 1), of the following (constrained embedded) first-order condition (a cubic equation) of the investors’ problem [Equations (12)-(14)], after substituting \( \delta^T \iota_M = S/W \),

\[
-\gamma \sigma_b^2 \left( \frac{S}{W} \right)^3 - \left[ \gamma \sigma_a^2 + \gamma \sigma_x^2 + \hat{b} + \left( \sum_{i=1}^M c_{i,i}^{-1} \right)^{-1} \right] \frac{S}{W} + X(e^*_i, H) = 0, \tag{28}
\]

where \( \gamma \triangleq \mu_p / \sigma_p^2 \).

**Proof of Proposition RA2.** See the Mathematical Appendix.

The intuition of Proposition RA2 is as follows. Investors allocate investments to funds based on their risk-return tradeoffs. Investing wealth in AFMI increases portfolio’s risk, so they choose to limit these investments, leaving \( E(\alpha_i \mid D)\big|_{e^*, r^*, s^*} > 0 \). The risk return tradeoff of potentially investing the “last dollar,” the dollar that would drive funds expected net alphas to zero, is “in the variance favor.” That is, the marginal cost of risk, of investing this last dollar, is higher than the marginal benefit of the gained net alpha. This prevents optimizing risk-averse investors from allocating it to AFMI, leaving funds expected net alphas to be positive. The properties of the cubic equation guarantee exactly one real positive root. If the positive root is larger than 1, then \( S/W = 1 \).

**COROLLARY to PROPOSITION RA2.** Where \( N \to \infty \), for large enough \( W \), such that \( S/W = 1 \), we have

1. Higher equilibrium optimal expected net alphas of an initial marginal investment in
the AFMI induce a larger equilibrium AFMI size relative to total wealth.

2. Higher rate of decrease in the aggregate (of fund level and industry level) AFMI’s returns to scale induces a smaller equilibrium AFMI size.

Proof of Corollary to Proposition RA2. See the Mathematical Appendix.

The intuition of this corollary is as follows. Where \( S/W = 1 \), an increase (decrease) in \( X(e^*, H) \) shifts up (down) the cubic function in Proposition RA2, inducing a larger (smaller) \( S/W \) as the maximizer of investors’ objective function. The economic sense is that a higher level of equilibrium optimal expected net alpha of an initial marginal investment, \( X(e^*, H) \), attracts more investments to the AFMI. Also, we can see that \( \hat{b} \) is the expected decreasing returns to scale at the industry level, based on current information, whereas \( \left( \sum_{i=1}^{M} C_{i,i}^{-1} \right)^{-1} W \) may be regarded as the equilibrium decreasing returns to scale factor at the fund level because it is calculated by all the fund average cost sensitivities to size, \( C_{i,i} \). Thus, the factor 

\[
\hat{b} + \left( \sum_{i=1}^{M} C_{i,i}^{-1} \right)^{-1} W
\]

may be regarded as the combined decreasing returns to scale factor. Investors invest less in funds if the effect of decreasing returns to scale is stronger in the AFMI. The next proposition offers comparative statics.

PROPOSITION RA3 AFMI Size and Net Alphas Sensitivity to Concentration.

Where \( N \to \infty \) and \( S/W < 1 \), we have

27

1. Higher concentration induces larger (smaller) equilibrium AFMI size and higher (lower) equilibrium expected net alphas, if and only if higher concentration induces a larger (smaller) impact on gross alphas than on costs.

2. Concave, in concentration, equilibrium direct benefits of effort function indicates concave, in concentration, equilibrium AFMI size. (Convex, in concentration, equilibrium AFMI size indicates convex, in concentration, equilibrium direct benefits of effort function.)

3. Concave, in concentration, equilibrium expected net alphas indicates concave, in concentration, equilibrium direct benefit function. (Convex, in concentration, equilibrium direct benefit function indicates convex, in concentration, equilibrium

---

27 Where \( S/W = 1 \), it is the case that, 1. \( S/W \) is unrelated to industry concentration, 2. higher concentration induces higher (lower) equilibrium expected net alphas if and only if higher concentration induces a larger (smaller) impact on gross alphas than on costs, and 3. equilibrium expected net alphas is concave (convex), in concentration, if and only if the equilibrium direct benefit function is concave (convex), in concentration.
expected net alphas.)

Proof of Proposition RA3. See the Mathematical Appendix.

The intuition of the size related results is as follows. Where \( N \to \infty \), a higher \( H \) affects industry size \( S/W \) through the equilibrium optimal expected net alpha of an initial marginal investment, \( X(e^*_i, H) \). If a higher \( H \) induces a larger (smaller) impact on gross alphas than on costs, then it creates a larger (smaller) \( X(e^*_i, H) \), and, consequently, attracts more (less) investments in the AFMI [if investors have additional wealth to allocate to funds (i.e., \( S/W < 1 \))]. In Proposition 1, we showed that \( \frac{dB(e^*_i; H)}{dH} = A_H(e^*_i; H) - c_2H(e^*_i; H) \); 28 thus, in this case, a higher \( H \) induces a larger \( S/W \) if and only if it induces a higher \( B(e^*_i; H) \).

Examining the second order effects of concentration on size, where \( N \to \infty \) and \( S/W < 1 \), we first note that changes in \( H \) that induce a larger \( S/W \), result in a larger allocation to AFMI funds and, in turn, in a higher investors’ overall portfolio risk. Mean-variance, risk-averse investors facing risk-return tradeoffs respond to increase in marginal portfolio risks, holding other parameters constant, by optimally lowering investment in funds. Thus, how changes in \( H \) affect changes in equilibrium \( S/W \) depend on how changes in \( H \) affect this risk-return tradeoff. The implications for the second-order derivative \( \frac{d^2(S/W)}{dH^2} \) are in the proof of Proposition RA3, which expresses this tradeoff analytically by identifying \( \frac{d^2(S/W)}{dH^2} \) as a sum of two addends. The first addend is negative (positive) if the direct benefits function is concave (convex) in \( H \), and the second one is always negative. This shows that a concave \( B(e^*_i; H) \) in \( H \), implies an \( S/W \) concave in \( H \); and a convex \( S/W \) is in \( H \) implies a convex \( B(e^*_i; H) \) in \( H \). This explains the second point of Proposition RA3.

The intuition of the performance related results is as follows. Where \( N \to \infty \) and \( S/W < 1 \), a higher \( H \) influences \( E(\alpha_i | D) \mid | e^*, t^*, s^* \) in two stages. In the first stage, it changes managers’ ability to produce expected net alphas, which is represented by the first component of \( dE(\alpha_i | D) / dH \mid | e^*, t^*, s^* \). In the second stage, investors react to the changes in fund expected

---

28 We note that this derivative of AFMI direct benefits with respect to \( H \) is the same as its partial derivative with respect to \( H \).
net alphas by adjusting investment levels in funds, consequently affecting $E(\alpha_i \mid D)\big|_{\epsilon_i,r_i,s_i}$
(due to decreasing returns to scale).

The risk-return tradeoff of risk-averse investors, makes their reaction to changes in fund
expected net alphas less intense. That is, they subdue their additional investments to funds
when inferring higher fund expected net alphas and limit their reduction in investments to funds
when observing lower fund expected net alphas. Therefore, whether a higher $H$ increases
$E(\alpha_i \mid D)\big|_{\epsilon_i,r_i,s_i}$ depends on whether it has a larger impact on gross alphas than on costs (i.e.,
the sign of $dE(\alpha_i \mid D) / dH \big|_{\epsilon_i,r_i,s_i}$ depends only on the sign of $A_H(e_i^*;H) - c_{2H}(e_i^*;H)$).

Because in equilibrium, $\frac{dB(e_i^*;H)}{dH} = A_H(e_i^*;H) - c_{2H}(e_i^*;H)$, whether a higher $H$ increases
$E(\alpha_i \mid D)\big|_{\epsilon_i,r_i,s_i}$ depends on whether it increases $B(e_i^*;H)$. Also, as $H$ changes, the change
of marginal $E(\alpha_i \mid D)\big|_{\epsilon_i,r_i,s_i}$ (i.e., $d^2E(\alpha_i \mid D) / dH^2 \big|_{\epsilon_i,r_i,s_i}$) is positively proportional to the
change of marginal $B(e_i^*;H)$, i.e., $d^2B(e_i^*;H) / dH^2$ plus a positive adjustment term that
captures the effects of risk. This adjustment term is positive because, holding all other
parameters constant, if investors’ marginal portfolio risks of investing in funds are higher,
investors optimally invest less in funds. In doing so, they exert a smaller impact on expected
net alphas; thus, a higher $H$ induces a higher marginal $E(\alpha_i \mid D)\big|_{\epsilon_i,r_i,s_i}$. We can see that if
$d^2B(e_i^*;H) / dH^2$ is positive, $d^2E(\alpha_i \mid D) / dH^2 \big|_{\epsilon_i,r_i,s_i}$ must be positive, whereas if
$d^2E(\alpha_i \mid D) / dH^2 \big|_{\epsilon_i,r_i,s_i}$ is negative, $d^2B(e_i^*;H) / dH^2$ must be negative. We also note that
where $S/W = 1$, investors have no additional wealth to allocate to funds, so they exert no
impact on marginal $E(\alpha_i \mid D)\big|_{\epsilon_i,r_i,s_i}$, making the marginal equilibrium optimal expected net
alphas depend only on the effect of $H$ on managers’ ability to produce net alphas.

2.3 Relation to Berk and Green (2004) and Pastor and Stambaugh (2012)

Our model follows the pivotal works of BG and PS in several respects. Central features
of our model include industry-level decreasing returns to scale and risk-averse investors, as in
PS, as well as fund-level decreasing returns to scale, as in BG. We highlight below the main
differences between our model and those of PS and BG. We discuss the special cases where
theirs and our models overlap, and we obtain results similar to theirs.

2.3.1. Relation to Pastor and Stambaugh (2012)

While PS model and compare expected net alpha and AFMI size within, two "extreme regimes," a fully competitive equilibrium \((M \to \infty)\) and a monopolistic equilibrium \((M = 1)\), we model and study trade-offs across a continuum of market concentration levels for any given \(M\), where \(M > 1\). We note that this concentration-alpha relation in our model is a mechanism distinct from, and additional to, the industry size-alpha relation in PS. Specifically, in our model the concentration-alpha relation exists also when controlling for the size of the industry (or the growth of the industry). Analytically, this is the case because concentration affects the optimal level of effort by affecting effort productivity [the third addend of the right side of Equation (7)] and the cost of effort [Equation (18)].

In our model, heterogeneous fund-level decreasing returns to scale are required for making the AFMI concentration non-trivial by allowing funds to have heterogeneous sizes. (PS’s model identifies the AFMI equilibrium elegantly, without the need to specify fund level heterogeneity in fund size.) However, fund-level decreasing returns does not directly influence alpha in our model. It only influences alpha via effort, concentration, and size. If we do not model effort, our model with only fund-level decreasing returns will not generate the concentration-alpha relation. A gist of the argument is that introducing funds heterogeneity would affect industry size and expected net alphas only if, in aggregate, it affects the industry’s alpha production. Such aggregate effects, however, are fully captured within the industry returns to scale structure (as in PS). Introducing fund level heterogeneity that does not have industry aggregate effect would not affect the industry size and expected net alphas. In other words, changes in AFMI concentration due to introducing fund level heterogeneity, without all effort effects, are either captured by industry returns to scale effects or have no effect.

Our model becomes similar to the one in PS in our special case where neither AFMI concentration nor managers’ effort affect managers’ search productivity for mispriced assets. Analytically, effort does not affect alpha production, in our model, if the third addend of the right side of Equation (7) does not exist, and if we abandon our cost function, Equation (18), in favor of defining funds’ fees to be net of funds’ management costs. However, even where, in our model, the search productivity for mispriced assets depends on effort, there is a special case of parameter values that lead to a solution where optimal allocated effort is zero. We report the conditions for this special case in the following proposition and its corollary. Intuitively, this is the case if, for all concentration levels, costs of efforts to improve alpha production,
exceed the benefits of the resulting increase in alphas. That is, market conditions are insufficiently conducive to launching a costly search for favorable investment opportunities. Thus, optimally, no extra effort is exerted and our model results resemble those in PS.

**PROPOSITION PS.** For manager \( i, i = 1,2,\ldots,M \), if initial effort inputs generate non-positive AFMI’s direct benefits of effort (i.e., \( B_{e_i}(0;H) \leq 0, \forall H \)), equilibrium optimal proportional effort levels \( e^*_i \) are zero (i.e., \( e^*_i = 0, \forall i, \)) and the optimal proportional fee \( f^*_i \) equals the average cost of operating funds \( c_0 + c_{1,i} \) (i.e., \( f^*_i = c_{0,i} + c_{1,i} s_i, \forall i \)).

**COROLLARY to PROPOSITION PS.** Under the conditions in Proposition PS, the equilibrium here resembles the one in PS. That is, effort is not exerted, and managers optimally choose not to charge fees above break-even costs.

*Proof of Proposition PS.* See the Mathematical Appendix.

This corollary says that AFMI concentration will not influence AFMI size or AFMI expected net alpha if, for all concentration levels, the optimal effort of fund managers is zero, given their trade-off between effort’s productivity and costs. However, the industry-level decreasing returns to scale mechanism of PS will still function and generate a negative relation between AFMI size and AFMI expected net alpha. While the latter effect represents how managers’ ability to outperform passive benchmarks declines with AFMI size, the former effect represents how manager’s incentives to exert individual effort to outperform passive benchmarks are influenced by AFMI’s concentration (for the same AFMI size and number of managers).

**2.3.2. Relation to Berk and Green (2004)**

As in PS, our baseline model assumes risk-averse investors. This assumption produces positive AFMI expected net alpha. When investors are risk-averse, expected net alphas are positive because investors require compensation to bear the risk of investing in active funds. However, BG do not solve the investor’ optimization problem and fix expected net alpha to be zero by invoking the assumption that non-benchmark risk can be completely diversified away across many funds (see also the discussion in PS, p. 775). This feature of the BG equilibrium is compatible with the case, in our model, where infinitely many small risk-neutral investors compete, and the size of the fund endogenously adjusts to make the gross alpha equal to the fee so that expected net alpha is always zero. In the case of infinitely many risk-neutral investors, our model, and the corresponding model in PS, produce zero AFMI expected net
alpha. The similarities of our model to those of BG and PS, in this case of infinitely many risk-neutral investors, are formally stated in the following proposition.

PROPOSITION RN1. For $N \to \infty$, equilibrium optimal allocations induce AFMI size to be,

$$\delta_j^{*T} \text{max} = \frac{1}{b + \left(\sum_{i=1}^{M} c_{i,j}^{-1}\right)^{-1}} \frac{1}{W}, \forall j,$$

and equilibrium expected net alphas to be,

$$E(\alpha_i | D)|_{\{e^*, r^*, \delta^*\}} = 0, \text{ where } S/W < 1,$$

and

$$E(\alpha_i | D)|_{\{e^*, r^*, \delta^*\}} = X(e_i^*, H) - \left[b + \left(\sum_{i=1}^{M} c_{i,i}^{-1}\right)^{-1} W \right] \geq 0, \text{ where } S/W = 1.$$

Proof of Proposition RN1. See the Mathematical Appendix.

Risk-neutral investors keep investing in AFMI, as long as they earn positive alphas. Thus eventually, they either drive alphas to zero and have $E(\alpha_i | D) = 0$ and $S/W \leq 1$, or run out of funds and have $E(\alpha_i | D) \geq 0$ and $S/W = 1$. Thus, if some of the wealth is passively managed $(S/W < 1)$, then, irrespective of AFMI concentration or AFMI size, equilibrium expected net alpha will be zero [$E(\alpha_i | D) = 0$]. This result parallels the results of BG and the risk-neutral case (with perfect competition) of PS. An additional result, here, is that, even in the risk-neutral case, AFMI size will depend on AFMI concentration through its effect on $X(e_i^*, H)$. The intuition is that, even though the AFMI expected net alphas are driven to zero, higher AFMI concentration incentivizes managers to invest more effort for finding mispriced assets. (Expected net alphas are driven to zero along a path of search for investment opportunities where managers exert more effort.) This increase in optimal effort, increases the AFMI size at which investors are indifferent between investing an additional dollar with AFMI or the passive benchmark.

Another difference between BG and our model is the source of heterogenous “manager ability.” In BG, the source of heterogenous manager ability is the expected excess return (over the passive benchmark) earned on the first dollar actively managed by a fund. In our model, this quantity is the same across funds. The source of heterogenous manager ability in our model is the fund-level decreasing returns to scale parameter $c_{i,j}$, which measures the rate at which the manager’s costs in generating gross alpha increase with size. A more skilled manager in
our model is one who has lower total variable costs of active management for the same AUM and gross alpha. Therefore, in our equilibrium, managers with the lower fund-level decreasing returns to scale parameter \( c_{1,c} \) have more assets under management. While the fund-level decreasing returns to scale parameter in BG influences the fund’s AUM in the same way [large parameter corresponds to a smaller fund, see their Equation (27)], they assume this parameter to be the same across funds.

Our choice of modeling heterogeneity in \( c_{1,c} \) (as a source of heterogenous manager ability) enables us to obtain heterogeneity in equilibrium fund sizes as well as a positive equilibrium fee charged by managers (see Proposition 3). In contrast, in the competitive equilibrium of PS, the fee is zero. If the fee (net of costs) were, instead, equal to some positive value in PS, then any fund manager would set an infinitesimally lower fee to attract all investment from other funds. However, we model costs explicitly, and in equilibrium fees compensate managers for their costs [fund managers charge (positive) break-even fees], which include a component related to size and a component related to effort (see Proposition 1).

Fund managers in the BG model are indifferent to the fee that they charge as long as two conditions are met (see their Section II.A): a. this fee is less than the hypothetical fee they could charge to maximize their compensation and, b. they can expand their fund by investing in the passive benchmark (i.e., “closet indexing”). They show that, under their assumptions, managers are indifferent between large AUM with a small fee or small AUM with a large fee as long as their profits stay the same. In BG’s framework, fund managers can choose their AUM independently of other fund managers’ skills.

In contrast, we model competition between managers with different returns to scale parameters. In our equilibrium, this competition for finite AUM results in zero profits and break-even fees charged by managers, and relative fund sizes that correspond to the relative rates at which fund-level returns to scale decrease. This implies a tight link between skill of one fund manager \( c_{1,c} \) and the market share and net alpha of all AFMI funds, as we describe in the following proposition.

**PROPOSITION RA4. Relation between Skill, Market Share and Net Alpha.**

Where investors are mean-variance risk-averse, \( N \to \infty \), and where \( S / W < 1 \), or \( S / W = 1 \), an increase (decrease) in \( c_{1,c} \), while \( c_{1,j}, \forall j \neq i \) are unchanged, induces

1. a decrease (increase) in \( s_i / S_i \), and an increase (decrease) in \( s_j / S, \forall j \neq i \),
2. a decrease (increase) in \( E(\alpha_i \mid D)_{[e', t', \delta']} \), and a decrease (increase) in \( E(\alpha_j \mid D)_{[e', t', \delta']} \), \( \forall j \neq i \).

**Proof of Proposition RA4.** See the Mathematical Appendix.

According to Proposition RA4, a change in \( c_{i,i} \) induces both \( s_j / S \) and \( E(\alpha_i \mid D)_{[e', t', \delta']} \) to change in the same direction, while inducing \( s_j / S \), \( \forall j \neq i \), and \( E(\alpha_j \mid D)_{[e', t', \delta']} \), \( \forall j \neq i \), to change in opposite directions. We call the effect of changes in \( c_{i,i} \) on \( s_j / S \), and on \( E(\alpha_i \mid D)_{[e', t', \delta']} \) — internality effect, and we call the effect of changes in \( c_{i,i} \) on \( s_j / S \), \( \forall j \neq i \), and on \( E(\alpha_j \mid D)_{[e', t', \delta']} \) — externality effect.

The intuition of Proposition RA4 is as follows. Based on Proposition 2, we can see that any change in \( c_{i,i} \), keeping \( c_{i,j} \), \( \forall j \neq i \) unchanged, results in a change in \( s_j / S \) in the opposite direction and a change in \( s_j / S \), \( \forall j \neq i \) in the same direction. Also, a higher \( c_{i,i} \), affects \( E(\alpha_i \mid D)_{[e', t', \delta']} \) in two stages. In the first stage, it decreases manager \( i \)'s average cost and, thus, induces higher fund expected net alphas. As manager \( i \) offers a higher fund expected net alpha, investments shift into fund \( i \) from other funds, making all those funds’ fund expected net alphas higher due to decreasing returns to scale at fund level. At the second stage, an increase in fund expected net alphas attracts investments into the AFMI, which in turn drives down fund expected net alphas due to decreasing returns to scale at industry level. Where \( N \to \infty \) and \( S / W < 1 \), as investors’ portfolio risks increase (decrease) when they invest more (less) in AFMI. They subdue AFMI investments increases when observing an increase in fund expected net alphas and limit investment reductions when observing a decrease in fund expected net alphas. Thus, investors’ risk-aversion mitigates the countered effect at the second stage and makes the first stage’s effect dominant. Where \( S / W = 1 \), investors have no additional wealth to allocate to funds, so their investments have no impact on marginal equilibrium optimal expected net alphas, causing the first stage’s effect to dominate. In all these cases, we find that \( E(\alpha_i \mid D)_{[e', t', \delta']} \)'s are driven down by an increase in \( c_{i,i} \), keeping \( c_{i,j} \), \( \forall j \neq i \) unchanged; consequently, we have a positive relation between \( E(\alpha_i \mid D)_{[e', t', \delta']} \) and \( s_j / S \).
(internality effect) and a negative relation between $E(\alpha_j | D)_{t^*,i^*,s^*}$ and $s_j / S$, $\forall j \neq i$ (externality effect).

### 2.4 Endogeneity in Measures of AFMI concentration

Our model allows for an endogenous measure of AFMI concentration. Modeling an endogenous measure of concentration facilitate the use of available and prevalent empirical measures. If we define $H$ to be the Herfindahl-Hirschman index (HHI), which is the sum of market shares squared, then $H$ is endogenous to our model.\(^{29}\) Using funds’ equilibrium market share, as identified in Proposition 2, we can write the equilibrium AFMI concentration $H^*$ as

$$H^* \triangleq \sum_{i=1}^{M} \left( \sum_{j=1}^{M} \left( \frac{S_{ij}}{S} \right) \right)^2 = \sum_{i=1}^{M} \left( c_{1i} \sum_{j=1}^{M} \left( c_{1j}^{-1} \right) \right)^2. \quad (29)$$

We can see that $H^*$ is determined by $c_{1i}$’s. Specifically, depending on the size of $c_{1i}$ relative to that $c_{1,j}$, $\forall j \neq i$, an increase in $c_{1,i}$, holding $c_{1,j}$, $\forall j \neq i$ constant, increases or decreases $H^*$.

When AFMI concentration is defined as $H^*$, propositions RA3 and RA4 imply that the relation between the $c_{1,i}$’s and the equilibrium fund expected net alphas and AFMI size is complex. In the case where there are infinitely many risk-averse investors, an increase in $c_{1,j}$ affects the equilibrium fund expected net alphas in two ways: 1) its direct impact leads to lower equilibrium fund expected net alphas (Proposition RA4), and 2) depending on fund $i$’s size relative to rivals, it increases or decreases $H^*$, which consequently increases (decreases) equilibrium fund expected net alphas if and only if $A_H(e^*_i; H^*) - c_{2H}(e^*_i; H^*) \geq (<) 0$ (Proposition RA3). Similarly, an increase in $c_{1,i}$ affects the equilibrium AFMI size in two ways: 1) its direct impact leads to an (inverse direction) AFMI size change, and 2) it increases or decreases $H^*$, which consequently increases (decreases) the equilibrium AFMI size if and only if $A_{HH}(e^*_i; H^*) - c_{2HH}(e^*_i; H^*) \geq (<) 0$. Thus, in the endogenous AFMI concentration measure case, the relation between the $c_{1,i}$’s and the equilibrium fund expected net alphas and AFMI size

\(^{29}\) In a $M$-firm AFMI, for example, the HHI could have values between the highest concentration, 1, where one of the funds captures practically all the market share, and the lowest concentration, $1/M$, where market shares are evenly divided. That is, in an $M$-firms’ market $HHI \in [\frac{1}{M}, 1)$. 

26
depends on fund i’s size relative to rivals.\textsuperscript{30}

In general, we expect the theoretical concentration level in our framework to be influenced by industry characteristics such as regulation, transaction costs, tax rates, barriers to entry, and funds’ idiosyncratic outcomes, in addition to funds’ cost sensitivity to size (i.e., $c_{i,i}$’s). In these cases, the concentration level can change even when all the cost sensitivities (or fund manager skill) are constant. We do not model the various determinants of concentration levels and simply assume it to be exogenous. As long as real-world concentration is not exactly determined by the $c_{i,i}$’s (or any other exogenous parameter of our model), we are back to the case when concentration is exogenous (that is, has an exogenous component) and where all our predictions, regarding the relation between changes in exogenous AFMI concentration level, the equilibrium fund expected net alphas and AFMI size, remain unaltered.

In the following empirical analysis, we use alternative empirical measures of concentration to evaluate robustness to issues such as endogeneity. We also control for potential endogeneity of fund size, AFMI size and alpha using lagged measures of concentration and the recursive demeaning estimator of PST.

\section{Empirical Predictions and Test Method}

In this section, we describe key empirical predictions that our theoretical model generates, followed by our data and methodology to test these predictions.

\subsection{Empirical Predictions}

Underlying our empirical predictions is the theoretical scenario where an increase in market concentration has a larger effect on the availability of mispriced investment opportunities than on any associated costs of exploiting these opportunities (higher concentration induces a larger marginal effort impact on gross alpha than on costs). For instance, in this scenario, costs associated with an increase in effort may be staff’s increase in compensation (endogenously determined). We assume that these costs are less than the value

\textsuperscript{30} We believe that our cost function, Equation (18), is a concise one that captures essential effects within our model. To assure that all our functional form restrictions of the non-specialized model (exogenous concentration), which we deem basic and simple, hold in the specialized one (endogenous measure of industry concentration); however, we need to impose additional, technical, “second order,” parameter restrictions. For brevity and simplicity, we do not impose these restrictions. We call the parameter values that make the specialized model abide by these restrictions \textit{plausible}. We, later, confirm that the said technical restrictions are not empirically binding. That is, imposing these restrictions would not have changed our empirical results. In other words, the empirically estimated parameters fall within the plausible parameters range.
added to the firm due to the increase in effort. We claim this is a reasonable assumption. [See, for example, Ibert, Kaniel and Van Nieuwerburgh (2017), who find concavity of managerial compensation in firm revenue and weak sensitivity of “pay to performance.”]. According to our definition of “direct benefits of effort”, this assumption means that the direct benefits of effort increase with concentration. Based on this scenario, we predict that a higher concentration level is associated with:

I. Larger AFMI size (Proposition RA3.1)
II. Higher AFMI net alpha (Proposition RA3.1)
III. Higher AFMI effort (Proposition 1.3)

We note that effort per-se is largely unobservable; Even the salaries of managers are difficult to observe. Therefore, it is difficult to test the last prediction directly. Instead, we provide an indirect test that provides suggestive evidence in favor of our model. Specifically, we use aggregate AFMI active share\textsuperscript{32} and tracking error as a proxy for effort. These proxies are likely to be correlated with effort because any attempt to outperform the benchmark must involve taking positions that are different from the benchmark [Cremers and Petajisto (2009)]. One reason why active share can be uncorrelated with effort is that fund managers could “jam the signal” in active share by taking uninformed bets to increase their perceived active share, generating a false sense of truly-active management [Brown and Davies (2017)]. However, such “signal jamming” behavior is more likely to be an issue if a measure of active share is tied with fund manager incentives, which is not common in our sample period [our sample is dominated by years before 2009, when Cremers and Petajisto (2009) published their active share measure.] Also, the moral hazard and signaling problem in Brown and Davies (2017) is associated with fund-level active share. In contrast, our prediction is for aggregate AFMI level, where information asymmetry and its associated “signal jamming” is less likely to be a concern.

Our model also has a second-order prediction that we test:

IV. AFMI net alphas and AFMI size are both either concave or convex in market concentration. (Proposition RA3.2 and RA3.3)

We test these four empirical predictions.

\textsuperscript{31} In our model, predictions I and II follow if and only if direct benefits of effort increase with concentration. Therefore, empirical support for these predictions, also provides additional evidence in support of our assumption that direct benefits of effort increase with concentration. Prediction III relies on an alternative assumption described in Footnote 6.

\textsuperscript{32} Fund-level Active Share is the percentage of fund holdings that is different from the benchmark holdings. A fund that has no holdings in common with the benchmark will have an Active Share of 100%, and a fund that has exactly the same holdings as the benchmark considered will have an Active Share of 0%. We measure industry-level “AFMI Active Share” as the average fund-level Active Share of all active funds in a given quarter.
3.2 Data

We obtain our data from Morningstar Direct. Our sample contains 1,374 actively managed U.S. equity mutual funds from January 1979 to December 2014. The Data Appendix supplements the data description below.

We use keywords in Morningstar to identify U.S. mutual funds (both open-end and closed-end) and exclude index funds, enhanced index funds, funds of funds, and in-house funds of funds. Also, we require funds to be classified as Equity in the Global Broad Category Group, and we further exclude international funds, real estate funds, and sector funds. Next, we use the Fund ID provided by Morningstar to aggregate fund share class-level information to fund-level information (because many mutual funds offer multiple share classes, which represent claims on the same underlying assets but have different fee structures; See PST for a discussion of this issue). Since we use a 5-year rolling window to estimate fund net alphas, we require each of our active equity mutual funds to have at least 10 years’ of monthly return observations. Using these filters, we obtain our sample of 1,374 actively managed U.S. equity mutual funds.

We evaluate a fund’s performance compared to a set of index funds selected as benchmarks. The index funds that we use as benchmarks are also from Morningstar. These index funds include those with the Morningstar Institutional Categories of Small Core, Large Core and S&P 500 Tracking, and the CRSP Fama-French risk-free rate. We require index funds to have no missing observations in our sample period. All the fund returns are net of administrative and management fees and other costs taken out of fund assets.

We obtain quarterly data on fund-level active share and tracking error from Antti Petajisto website (www.petajisto.net/data.html). Petajisto (2013) contains a description of how this data is constructed.

3.3 Variable Definitions

An important goal of our empirical analysis is to analyze how AFMI size and fund net alphas change with AFMI concentration. We also attempt to provide some insights into AFMI effort by using active share and tracking error as a proxy of effort. We now define how we measure these variables.

**AFMI Size:** Our measure of AFMI size \( (\text{SoW}) \), size over wealth) is the sum of funds’ net assets under management divided by the stock market capitalization in the same month.

\[ \text{AFMI Size} = \frac{\text{Sum of fund net assets under management}}{\text{Stock market capitalization}} \]

33 We also omit some rare cases where there is a gap with more than 5 years’ return observations missing.
AFMI Active Share: We measure AFMI active share as the quarterly average active share of active funds (in Petajisto’s database). We exclude index funds and enhanced index funds in the database when identifying active funds.

AFMI Tracking Error: Like AFMI active share, we calculate AFMI tracking error as the average tracking error of active funds in that quarter.

AFMI Concentration: Following the literature, we measure AFMI concentration using three indices [see, for example, Berger and Hannan (1989), Geroski (1990), Berger (1995), Goldberg and Rai (1996), Nickell (1996), Berger, Bonime, Covitz and Hancock (1999), Cremers, Nair, and Peyer (2008), and Giroud and Mueller (2011)]:

1) the HHI:

\[ H_t = \sum_{i} MS_{i,t}^2. \]  

(30)

where \( MS_{i,t} \) is the market share of fund \( i \) at time \( t \), measured by the fund’s assets under management at time \( t \) over the total AFMI assets under management at time \( t \). The HHI is a commonly used measure of concentration [see, for example, Cremers, Nair, and Peyer (2008), and Giroud and Mueller (2011)] and is well grounded in theory [see Tirole (1988), pp. 221-223].

As the value of the HHI is related to the number of funds \( m_t \), for a robustness check, we also use two measures not related to the number of funds to measure market concentration, the NHHI used by Cremers, Nair and Peyer (2008) and the 5-Fund index, another common measure of market concentration:

2) the normalized HHI (NHHI):

\[ NHHI_t = \frac{H_t - \frac{1}{m_t}}{1 - \frac{1}{m_t}}. \]  

(31)

The NHHI induces similarity in possible concentration level distributions. For example, NHHI will be zero for an industry where all firms have equal market shares, regardless of whether it has 3 or 4 firms. In contrast, HHI will be 0.33 for the 3-firm industry with equal market shares, and 0.25 for the 4-firm industry. More generally, these two measures have different possible concentration level intervals: independent of the number (greater than 1) of industry firms, under NHHI all concentration distributions are on, the interval, \([0,1]\). In contrast, under HHI,
an \( M \)-firm industry (\( M > 1 \)), has a concentration distribution on, the interval \( \left[ \frac{1}{M}, 1 \right) \). A 3-firm industry would induce a concentration levels distribution on \([0,33,1)\), and a 4-firm industry would induce a concentration distribution on \([0.25,1)\).

3) the sum of the first five largest funds’ market shares (5-Fund index):

\[
5FI = \sum_{i=1}^{5} MS_{i,t}.
\] (32)

Note that our model provides a relation between these measures of market concentration and fund-level decreasing returns to scale parameters. For instance, if we define \( H \) as the Herfindahl-Hirschman index (HHI), which is the sum of market shares squared, then for an \( M \) firms’ market \( H \in \left[ \frac{1}{M}, 1 \right) \). 34 Using funds’ equilibrium market share, as identified in Proposition 2, we can write the equilibrium market concentration \( H^* \) as in Equation (29). This equation shows a relation between \( H^* \) and the \( c_{1,i} \)'s. Similar relations can be obtained for NHHI and the 5FI. The question of which of these two quantities is exogenous, or whether both are determined together in equilibrium, is a complex one that is beyond the scope of this paper (see Section 2.4 for a discussion of this issue). We simply use these empirical measures as proxies for the true level of AFMI concentration (or competition) and do not enforce restrictions between cost parameters and \( H \).

Net Alpha: Our measure of net alpha (\( \alpha_{i,t} \)) is the difference between a fund’s net return and the net return on the benchmark we assign to the fund. The benchmark against which we judge a fund’s net alpha is a set of (traded) index funds selected using style analysis [Sharpe (1992)]. 35 These index funds are intended to represent the next-best investment opportunity available to investors as a tradable passive index [Berk and van Binsbergen (2015)]. In our model, we assume that a single passive benchmark exists and is common knowledge to investors and managers. In the theoretical analysis, we make this assumption for parsimony. Relaxing it will not alter the key insights from our model. However, in our empirical section, we allow for

34 In an \( M \) AFMI, for example, the Herfindahl-Hirschman index could have values between the highest concentration, 1, where one of the funds captures practically all the market share, and the lowest concentration, \( 1/M \), where market shares are evenly divided.

35 Notably, as in PST, we do not use the Fama-French factors as our benchmark. PST note that “The Fama-French factors are popular in mutual fund studies because their returns are freely available. Yet the Fama-French factors are not obvious choices because they are long-short portfolios whose returns cannot be costlessly achieved by mutual fund managers or investors.” In addition, Cremers, Petajisto, and Zitzewitz (2012) and Grinblatt and Saxena (2017) argue that the Fama-French model produces biased assessments of alpha. To avoid such problems and remain consistent with our model where investors compare active funds to a traded passive benchmark, we use Sharpe style analysis and identify an appropriate traded benchmark for each mutual fund in our sample.
multiple benchmarks and match each active equity mutual fund to a set of tradable index funds that reasonably replicate passive alternatives available to an average mutual fund investor. Specifically, we assume the following return-generating process:

\[ R_{i,t} = \alpha_{i,t} + b_{i,1} F_{1,t} + b_{i,2} F_{2,t} + \ldots + b_{i,n} F_{n,t}, \]  

where the indices \( i \) and \( t \) represent the fund and time indices, whereas \( n \) indicates the number of tradable index funds in the market. \( R_{i,t} \) is the return net of management fee of a fund, and \( F_{1,t} \) through \( F_{n,t} \) are the returns net of management fees of tradable index funds in different asset classes. We treat the index funds \( F_{1,t} \) through \( F_{n,t} \) as a basis fund set that may be used to replicate the returns on any passive benchmarks used by mutual fund investors.

To perform our analysis, we first need to measure fund net alphas \( (\alpha_{i,t}) \). For each active fund in our sample, we calculate a set of weights on our basis fund set that sum to one and minimize the tracking error between the active fund return and a corresponding passive benchmark portfolio return [Sharpe (1992)]. We note that our empirical design of identifying passive benchmarks, using matching tradable index funds, fits our theoretical structure, which assumes the appropriate passive benchmarks for each fund.

We perform this analysis on a rolling basis, using returns from months \(( t - 60 )\) to \(( t - 1 )\) to avoid look-ahead bias. That is, we identify coefficients \( b_{i,1} \) to \( b_{i,n} \) to minimize the variance of the residual. These coefficients are constrained to be between zero and one (we do not allow short selling), and their sum is constrained to be one. These coefficients identify the portfolio weights in our basis index fund set that provides the estimated minimum “tracking error” passive benchmark of a fund.

To calculate a fund’s net alphas in month \( t \), we subtract the returns on the identified passive portfolio (the style benchmark) for month \( t \) from the active equity fund’s returns in month \( t \) and that of the style benchmark in month \( t \) (see Eq. (33)). This provides us with fund net alphas in each month for each fund.

To ensure the robustness of our results, we also use an alternative method to measure fund net alphas. This method addresses the possibility that traded index funds do not capture unobservable risk factors that drive excess returns. Errors in our set of passive benchmarks or our matching strategy may result in net alphas that measure exposure to such unobservable risk factors instead of fund manager performance. Using the method developed by Connor and Korajczyk (1988), we estimate unobserved common factors in our estimated fund net alphas.
using the principal components of our estimated fund net alphas series. We use these estimated principal components to control for unobserved common factors in fund net alphas. In particular, we regress each fund’s fund net alphas on the first two principal components without a constant term. We refer to the residuals of these regressions as PC-adjusted fund net alphas and use them as the dependent variable in our robustness analysis.

3.4 Methodology

We analyze the impact of AFMI concentration on AFMI size, active share and tracking error (predictions I and III) at an industry level. We analyze the impact of market concentration on AFMI alpha (prediction II) using fund-level data to control for potential effects of fund size on performance, and an associated omitted-variable and finite sample bias (see PST). We study second-order predictions using both industry-level data and fund-level data as prediction IV requires that AFMI net alphas (measured using fund-level data) and AFMI size (measured using industry-level data) are both either concave or convex in market concentration.

3.4.1 Industry-level Analysis

Using monthly data, in analyzing the relation between AFMI size and market concentration, we use vector auto-regression (VAR). The main equation in the VAR system is,

\[ SoW_t = b_0 + b_1 SoW_{t-1} + b_2 H_{t-1} + b_3 H_{t-1}^2 + e_{1, t}, \]

where \( SoW_t \) is AFMI size and \( e_{1, t} \) represents regression residuals. In the VAR system we also have equations where \( H_t \) depends on \( H_{t-1} \), and on \( SoW_{t-1} \), and where \( H_t^2 \) depends on \( H_{t-1}^2 \).

In addition, we use two effort proxies, active share \((AS_t)\) and tracking error \((TE_t)\), which are likely to represent different dimensions of effort. According to Cremers and Petajisto (2009), the active share measure emphasizes stock selection, while tracking error volatility emphasizes bets on systematic risk factors. Therefore, we include the effects of both of these dimensions in testing prediction III, using VAR,

\[
AS_t = b_{01} + b_{11} AS_{t-1} + b_{12} TE_{t-1} + b_{12} H_{t-1} + e_{11, t},
\]

\[
TE_t = b_{02} + b_{21} AS_{t-1} + b_{22} TE_{t-1} + b_{22} H_{t-1} + e_{22, t},
\]

where \( e_{11, t} \) and \( e_{22, t} \) represent regression residuals. In the VAR system we also have the equation where \( H_t \) depends on \( H_{t-1} \), \( AS_{t-1} \), and \( TE_{t-1} \).

3.4.2 Fund-level Analysis

We follow PST’s methodology to control for omitted-variable and finite-sample bias in our alpha analysis. The omitted-variable problem arises from the cross-sectional variation in
performance that is due to differences in skill across funds. PST note that fund fixed effects can control for this heterogeneity as long as fund skill is time-invariant. However, adding fund fixed effects introduces finite-sample bias due to the positive contemporaneous correlation between changes in fund size and unexpected fund returns. Avoiding these biases motivates our use of the PST recursive demeaning (RD) estimator. In particular, we estimate the effects of market share ($\beta_1$), market concentration ($\beta_2$ and $\beta_3$), and AFMI size ($\beta_4$) on fund net alphas using the following panel regression:

$$\alpha_{i,t} = \beta_1 MS_{i,t-1} + \beta_2 H_{i,t-1} + \beta_3 H_{i,t-1}^2 + \beta_4 SoW_{i,t-1} + \epsilon_{i,t}.$$  

(36)

The bar above the variables denotes forward-demeaned variables, defined below:

$$\bar{\alpha}_{i,t} = \alpha_{i,t} - \frac{1}{T_i - t + 1} \sum_{s=t}^{T_i} \alpha_{i,s},$$  

(37)

$$\bar{MS}_{i,t} = MS_{i,t} - \frac{1}{T_i - t + 1} \sum_{s=t}^{T_i} MS_{i,s},$$  

(38)

$$\bar{H}_i = H_i - \frac{1}{T_i - t + 1} \sum_{s=t}^{T_i} H_s,$$  

(39)

$$\bar{H}^2_i = H^2_i - \frac{1}{T_i - t + 1} \sum_{s=t}^{T_i} H^2_s,$$  

(40)

$$\bar{SoW}_i = SoW_i - \frac{1}{T_i - t + 1} \sum_{s=t}^{T_i} SoW_s,$$  

(41)

where $T_i$ is the number of time-series observations of fund $i$. We run robustness checks by replacing $H_i$ (HHI) with the NHHI and with the 5FI.

The RD method in Equation (36) can control for the fund fixed effect. We include market share as a control, not only because the equilibrium market share provides information on a fund’s cost sensitivity to fund size (Proposition 2), but also because empirical studies show a linear relation between changes in market share and fund performance [Spiegel and Zhang (2013)] and use it as a firm-level market power measure [e.g., Berger, Bonime, Covitz and Hancock (1999) and Nickell (1996)]. There may be potential endogeneity (reverse causality) between AFMI shares and fund net alphas because when fund net alphas are higher, corresponding asset values increase and funds attract investments, both leading to a higher market share. This endogeneity issue may bias our results. Following PST, we address this
endogeneity issue using an instrumental variable method. In the first stage, we regress \( MS_{t-1} \) (recursively forward-demeaned market share) on \( MS_{t-1} \) (recursively backward-demeaned market share) without a constant term. In the second stage, we use the fitted value from the first stage to run Equation (36), where

\[
MS_{t,t} = MS_{t,t} - \frac{1}{t-1} \sum_{s=1}^{t-1} MS_{s,t}.
\]

There is no reason to believe that individual fund net alphas, which are fund level variables, are endogenous to industry-level measures such as AFMI concentration ratios (see, for example, Footnote 17 of PST). Thus, to test the concentration-alpha relation, we do not use a backward demeaned instrument. We just use the recursive forward-demeaned market concentration ratios in Equation (36).

## 4 Empirical Results

Table 1 reports the summary statistics. We find that monthly fund net alphas are positive on average, but exhibit a wide variation. We also report summary statistics of the fit of our passive benchmark-matching method using R-squared, which is measured as

\[
R_{sqr,t} = 1 - \frac{Var(\alpha_{i,t})}{Var(R_{sqr,t})},
\]

where \( Var(.) \) denotes variance. On an average, our style-matching model fits well with an average R-squared of 0.86 and a standard deviation of about 0.12. The summary statistics of industry size (total funds’ net assets divided by stock market capitalization) and fund sizes in December 2014 dollars (funds’ net assets divided by stock market capitalization in the same month, multiplied by the stock market capitalization in December 2014) are similar to the sample in PST.

The number of active equity mutual funds in our sample increases over time and the market concentration measures, such as the HHI, NHHI, and 5FI, with fluctuations, tend to decrease over time. Figure 1 shows the HHI value from January 1984 to December 2014. We

---

36 To be a valid instrument of \( MS_{t-1} \), \( MS_{t-1} \) must satisfy the relevance and exclusion conditions. The relevance condition is likely to hold because both \( MS_{t-1} \) and \( MS_{t-1} \) are derived from \( MS_{t-1} \) and are, thus likely to be closely related. The exclusion condition is also likely to hold because the backward-looking information in \( MS_{t-1} \) is unlikely to be helpful in predicting the forward-looking net alpha information in \( \epsilon_{i,t} \), where \( \epsilon_{i,t} \) is the residual in the RD method. We correct the second-stage standard error estimates of \( \beta_t \) by incorporating the estimation errors from the first-stage regression.
can see that before 1990, the HHI value was relatively high, fluctuating from 0.02 to 0.03. After that, it continued decreasing; and in the current years, it has reached 0.006, which is around a quarter of the values before 1990. This figure shows that the concentration of the U.S. active equity mutual fund market decreased substantially. Alternative market concentration measures, such as NHHI and 5FI, show similar trends.

Because our sample differs from PST, we check for any alarming systematic differences by evaluating the returns to scale relation in our sample. In unreported results, we find results consistent with PST’s: fund net alpha is significantly negatively associated with lagged industry size and is negatively (but insignificantly) associated with lagged fund size. Thus, the results suggest decreasing returns to scale at the industry level.

We commence our empirical analysis by evaluating the relation between industry size and market concentration. The results of the main equations of the VARs are shown in Table 2. The first column of each model specification shows how AFMI size is positively associated with HHI. The result of interest in this table is that AFMI size is significantly positively associated with lagged HHI (model specification 1) and is significantly negatively associated with the second order of lagged HHI (model specification 2). If we further include a time trend or year dummies into the model, we find consistent results (model specifications 3 and 4). That is, AFMI size is increasing and concave in market concentration. In Panels B and C, we analyze the sensitivity of our results to alternative measures of market concentration: the NHHI and 5FI. We generally find consistent results. Thus, we conclude that the data supports Prediction I. From our model’s perspective, the positive relation between industry size and market concentration indicates that higher market concentration levels, on average, increase gross alphas more than they increase effort costs.

Next, we evaluate the relation between fund net alphas and market concentration. The results using the RD method are shown in Table 3. Panel A reports the results using fund net alpha as the dependent variable. In the first two columns, we find that the coefficient of the first-order term of lagged HHI is significantly positive, whereas the coefficient of the second-order term is significantly negative. This result is robust to including lagged market share and lagged industry size as controls. This suggests that the effect of concentration is distinct from the effect of decreasing returns to scale at the fund and industry level. To control for the possibility of unaccounted common factors in the estimated net alphas, we also use principal component (PC)-adjusted fund net alphas as the dependent variable (Panel B) and find similar results.
The main result of this table is that fund net alphas, on average, are increasing concave in market concentration. Our theoretical results, then, indicate that for plausible parameter values, higher levels of market concentration induce increases in gross alpha production opportunities that are higher than those in managers’ effort costs.

Table 4 analyzes the relation between AFMI active share, tracking error, and market concentration. We find support for prediction III, in that active share increases with all measures of concentration we consider. The relationship with tracking error is less robust. This could reflect that the relation between concentration and effort is more due to the effort involved in stock picking (as measured by active share), rather than the effort involved in factor timing (as measured by tracking error). Given difficulties in measuring effort discussed earlier, we leave a more comprehensive analysis of the relation between effort and concentration for future research.

4.1 Robustness

In addition to reported tables, we examine the sensitivity of the results in Table 3 by using fund fixed-effect regressions instead of the RD method. Most of the results are consistent, except when regressing the PC-adjusted fund net alpha on market concentration measures; we find that the significance of market concentration measure is reduced. We also analyze whether our results are driven by small funds. We redo our main analyses using observations after restricting our sample to funds with a net asset value above $15 million in any month of our sample period. Again, we find consistent results. To test whether our main results are stable across sub-samples, we redo our analyses in for three sub-periods. We find a significantly positive relation between fund net alphas and lagged HHI in all three sub-periods.

5 Conclusion

We develop a theoretical model to analyze an AFMI equilibrium where we investigate performance, size, and managers’ costly (optimal) effort under a continuum of exogenous market concentration levels. We use Pastor and Stambaugh’s (2012) framework, where gross alpha production is of decreasing returns to scale at the industry level, and we similarly model the decreasing returns to scale effect at the fund level. Higher market concentration levels imply better utilization of industry resources and the existence of more unexplored investment opportunities, making managers’ efforts more productive.

Our model’s comparative statistics characterize the association between fund expected net alphas and a continuum of exogenous market concentration levels, and that between AFMI size and market concentration. In particular, we consider the case of infinitely many mean-
variance risk-averse investors whose portfolio risks increase with investments in funds. The funds’ expected net alphas increase with market concentration if and only if higher concentration induces a larger impact on gross alpha production than on the costs of effort (i.e., higher concentration induces higher AFMI’s direct benefits of effort). Observing an increase in fund expected net alphas, due to higher market concentration, mean-variance risk-averse investors increase their mutual fund holdings but, mitigating the resulting increase in their portfolios risk, reach optimum investment levels at higher expected net alphas. Thus, the equilibrium fund expected net alphas become positively associated with market concentration. In addition, the concavity of fund expected net alphas in market concentration indicates that AFMI’s direct benefits of effort are concave in market concentration. This further induces concavity of AFMI size in market concentration.

We specialize our model to allow for endogenous market concentration levels, which befits empirical market concentration measures and facilitates our empirical studies.

We use Morningstar’s U.S. active equity mutual fund data. First, we find that on average, fund net alphas are negatively associated with fund size and AFMI size, confirming decreasing returns to scale at both fund and industry levels. More importantly, we also find that, on average, both fund net alphas and AFMI size are increasing concave with market concentration.

Our findings have policy implications for the U.S. AFMI. Under the current, empirically identified, tradeoff between changes in managerial productivity and in effort costs due to changes in the AFMI concentration level, increases in concentration levels are likely to increase fund net alphas, AFMI size, and AFMI’s direct benefits of effort. Future research could extend our analysis to international fund markets, pension funds, and hedge funds.

References


Sharpe, W.F., 1992, Asset Allocation: Management Style and Performance Measurement,


Figure 1. HHI Value from January 1984 to December 2014

The HHI value is in decimals. The gray bars represent the recession periods.
Our sample period is from January 1979 to December 2014, and monthly data is used. Panel A reports the summary statistics for fund-level data, and Panel B reports those for industry-level data. Fund Net Return and Fund Net Alpha are in percentages, and both are net of administrative and management fees and other costs taken out of fund assets. HHI, NHHI, and 5FI are Herfindahl-Hirschman index, normalized Herfindahl-Hirschman index, and 5-Fund index, respectively. AFMI Size is the sum of funds’ net assets under management divided by the stock market capitalization in the same month. AFMI active share is measured as the average active share of active funds (in Petajisto’s database) in a quarter. AFMI tracking error is calculated as the average tracking error of active funds in a quarter. The Style-Matching Model R-sqr, AFMI Share, HHI, NHHI, 5FI, AFMI active share, and AFMI tracking error are in decimals. Fund Size is measured in $100 millions and is equal to the fund’s total net assets under management, divided by the stock market capitalization in the same month, and multiplied by the stock market capitalization in December 2014. Number of Funds is in units.

### Panel A: Fund-Level Data

<table>
<thead>
<tr>
<th>Variable</th>
<th>Obs.</th>
<th>Mean</th>
<th>Std.</th>
<th>1st</th>
<th>25th</th>
<th>50th</th>
<th>75th</th>
<th>99th</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Fund Net Return (%)</strong></td>
<td>321,456</td>
<td>0.8736</td>
<td>5.1508</td>
<td>-14.4922</td>
<td>-1.7976</td>
<td>1.2998</td>
<td>3.8907</td>
<td>13.0053</td>
</tr>
<tr>
<td><strong>Fund Net Alpha (%)</strong></td>
<td>246,553</td>
<td>0.0349</td>
<td>1.9499</td>
<td>-5.4465</td>
<td>-0.8570</td>
<td>0.0215</td>
<td>0.9156</td>
<td>5.5982</td>
</tr>
<tr>
<td><strong>Style-Matching Model R-sqr (decimal)</strong></td>
<td>246,557</td>
<td>0.8607</td>
<td>0.1175</td>
<td>0.4223</td>
<td>0.8178</td>
<td>0.8953</td>
<td>0.9408</td>
<td>0.9894</td>
</tr>
<tr>
<td><strong>Fund Size (in 100 Million of 2014 Dec Dollars)</strong></td>
<td>314,083</td>
<td>28.7796</td>
<td>95.3306</td>
<td>0.0399</td>
<td>1.3833</td>
<td>5.5718</td>
<td>20.1835</td>
<td>416.9203</td>
</tr>
<tr>
<td><strong>Fund Market Share (decimal)</strong></td>
<td>314,083</td>
<td>0.0012</td>
<td>0.0041</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0002</td>
<td>0.0007</td>
<td>0.0185</td>
</tr>
</tbody>
</table>

### Panel B: Industry-Level Data

<table>
<thead>
<tr>
<th>Variable</th>
<th>Obs.</th>
<th>Mean</th>
<th>Std.</th>
<th>1st</th>
<th>25th</th>
<th>50th</th>
<th>75th</th>
<th>99th</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>AFMI Size (decimal)</strong></td>
<td>432</td>
<td>0.0982</td>
<td>0.0591</td>
<td>0.0200</td>
<td>0.0389</td>
<td>0.1035</td>
<td>0.1638</td>
<td>0.1801</td>
</tr>
<tr>
<td><strong>Number of Funds (No.)</strong></td>
<td>432</td>
<td>850.2</td>
<td>659.5</td>
<td>86.0</td>
<td>249.0</td>
<td>677.5</td>
<td>1468.5</td>
<td>2126.0</td>
</tr>
<tr>
<td><strong>HHI (decimal)</strong></td>
<td>432</td>
<td>0.0191</td>
<td>0.0230</td>
<td>0.0061</td>
<td>0.0101</td>
<td>0.0157</td>
<td>0.0243</td>
<td>0.0382</td>
</tr>
<tr>
<td><strong>NHHI (decimal)</strong></td>
<td>432</td>
<td>0.0157</td>
<td>0.0139</td>
<td>0.0057</td>
<td>0.0094</td>
<td>0.0141</td>
<td>0.0201</td>
<td>0.0269</td>
</tr>
<tr>
<td><strong>5FI (decimal)</strong></td>
<td>432</td>
<td>0.2166</td>
<td>0.0765</td>
<td>0.1240</td>
<td>0.1640</td>
<td>0.1986</td>
<td>0.2650</td>
<td>0.3438</td>
</tr>
<tr>
<td><strong>AFMI Active Share (decimal)</strong></td>
<td>119</td>
<td>0.8349</td>
<td>0.0416</td>
<td>0.7620</td>
<td>0.7980</td>
<td>0.8440</td>
<td>0.8740</td>
<td>0.8940</td>
</tr>
<tr>
<td><strong>AFMI Tracking Error (decimal)</strong></td>
<td>119</td>
<td>0.0755</td>
<td>0.0260</td>
<td>0.0382</td>
<td>0.0606</td>
<td>0.0707</td>
<td>0.0868</td>
<td>0.1954</td>
</tr>
</tbody>
</table>

**Table 1. Statistical Summary**
Table 2. Industry-level Analysis: AFMI Size and Market Concentration

This table reports the results of the main equations in various VAR models, where AFMI Size is the dependent variable. Sample period is from January 1979 to December 2014, and monthly data is used. AFMI Size is the sum of funds’ net assets under management divided by the stock market capitalization in the same month. HHI, NHHI, and 5FI are Herfindahl-Hirschman index, normalized Herfindahl-Hirschman index, and 5-Fund index, respectively, and HHI^2, NHHI^2, and 5FI^2 are their squared terms. Panel A, B, and C report the results of these three market concentration measures, respectively. Time Trend is set to be one for January 1979 and to increase by one each month. Small-sample adjusted standard errors are used and presented in parentheses. The symbols ***, **, and * represent the 1%, 5%, and 10% significant level in a two-tail t-test, respectively.

<table>
<thead>
<tr>
<th>Panel A</th>
<th>(1) AFMI Size</th>
<th>(2) AFMI Size</th>
<th>(3) AFMI Size</th>
<th>(4) AFMI Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lagged AFMI Size</td>
<td>1.0033*** (0.0014)</td>
<td>1.0052*** (0.0035)</td>
<td>0.9952*** (0.0063)</td>
<td>0.8772*** (0.0293)</td>
</tr>
<tr>
<td>Lagged HHI</td>
<td>0.0287*** (0.0037)</td>
<td>0.0491* (0.0278)</td>
<td>0.0754*** (0.0310)</td>
<td>0.9341*** (0.1126)</td>
</tr>
<tr>
<td>Lagged HHI^2</td>
<td>-0.1361* (0.0761)</td>
<td>-0.2044** (0.0846)</td>
<td>-2.4601*** (0.3003)</td>
<td></td>
</tr>
<tr>
<td>Time Trend</td>
<td>0.0000* (0.0000)</td>
<td>0.0000 (0.0000)</td>
<td>0.0000 (0.0000)</td>
<td>0.0000 (0.0000)</td>
</tr>
<tr>
<td>Constant</td>
<td>-0.0005** (0.0002)</td>
<td>-0.0009 (0.0008)</td>
<td>-0.0017* (0.0009)</td>
<td>0.0155*** (0.0054)</td>
</tr>
<tr>
<td>Year Dummies</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Observations</td>
<td>431</td>
<td>431</td>
<td>431</td>
<td>431</td>
</tr>
<tr>
<td>R-Sqr</td>
<td>0.999</td>
<td>0.999</td>
<td>0.999</td>
<td>0.999</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel B</th>
<th>(1) AFMI Size</th>
<th>(2) AFMI Size</th>
<th>(3) AFMI Size</th>
<th>(4) AFMI Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lagged AFMI Size</td>
<td>1.0040*** (0.0015)</td>
<td>1.0097*** (0.0029)</td>
<td>0.9999*** (0.0064)</td>
<td>0.8660*** (0.0286)</td>
</tr>
<tr>
<td>Lagged NHHI</td>
<td>0.0497*** (0.0063)</td>
<td>0.1328*** (0.0335)</td>
<td>0.1425*** (0.0341)</td>
<td>1.0085*** (0.1148)</td>
</tr>
<tr>
<td>Lagged NHHI^2</td>
<td>-0.6050*** (0.1493)</td>
<td>-0.6446*** (0.1515)</td>
<td>-4.2975*** (0.4981)</td>
<td></td>
</tr>
<tr>
<td>Time Trend</td>
<td>0.0000 (0.0000)</td>
<td>0.0000 (0.0000)</td>
<td>0.0000 (0.0000)</td>
<td>0.0000 (0.0000)</td>
</tr>
<tr>
<td>Constant</td>
<td>-0.0008*** (0.0002)</td>
<td>-0.0024*** (0.0007)</td>
<td>-0.0025*** (0.0008)</td>
<td>0.0175*** (0.0052)</td>
</tr>
<tr>
<td>Year Dummies</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Observations</td>
<td>431</td>
<td>431</td>
<td>431</td>
<td>431</td>
</tr>
<tr>
<td>R-Sqr</td>
<td>0.999</td>
<td>0.999</td>
<td>0.999</td>
<td>0.999</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel C</th>
<th>(1) AFMI Size</th>
<th>(2) AFMI Size</th>
<th>(3) AFMI Size</th>
<th>(4) AFMI Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lagged AFMI Size</td>
<td>1.0090*** (0.0020)</td>
<td>1.0052*** (0.0032)</td>
<td>0.9961*** (0.0062)</td>
<td>0.8526*** (0.0291)</td>
</tr>
<tr>
<td>Lagged 5FI</td>
<td>0.0199*** (0.0015)</td>
<td>0.0122*** (0.0052)</td>
<td>0.0149*** (0.0055)</td>
<td>0.1353*** (0.0208)</td>
</tr>
<tr>
<td>Lagged 5FI^2</td>
<td>-0.0114*** (0.0043)</td>
<td>-0.0132*** (0.0045)</td>
<td>-0.0977*** (0.0156)</td>
<td></td>
</tr>
<tr>
<td>Time Trend</td>
<td>0.0000 (0.0000)</td>
<td>0.0000 (0.0000)</td>
<td>0.0000 (0.0000)</td>
<td>0.0000 (0.0000)</td>
</tr>
<tr>
<td>Constant</td>
<td>-0.0029*** (0.0005)</td>
<td>-0.0022* (0.0012)</td>
<td>-0.0028** (0.0013)</td>
<td>0.0101 (0.0062)</td>
</tr>
<tr>
<td>Year Dummies</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Observations</td>
<td>431</td>
<td>431</td>
<td>431</td>
<td>431</td>
</tr>
<tr>
<td>R-Sqr</td>
<td>0.999</td>
<td>0.999</td>
<td>0.999</td>
<td>0.999</td>
</tr>
</tbody>
</table>
Table 3. Fund-level Analysis: Fund Net Alpha and Market Concentration

This table reports the results of our RD panel regression model. Results using Fund Net Alpha and the PC-Adjusted Fund Net Alpha (adjusted by the first two principal components of fund net alphas) as the dependent variables are presented. AFMI Size is the sum of funds’ net assets under management divided by the stock market capitalization in the same month. Market Share is equal to a fund’s net assets under management divided by the sum of all funds’ net assets under management in the same month. HHI, NHHI, and 5FI are Herfindahl-Hirschman index, normalized Herfindahl-Hirschman index, and 5-Fund index, respectively, and HHI^2, NHHI^2, and 5FI^2 are their squared terms. Panel A, B, and C report the results of these three market concentration measures, respectively. The unit of coefficients is percentage. Standard errors are clustered by fund and presented in parentheses. The symbols ***, **, and * represent the 1%, 5%, and 10% significant level in a two-tail t-test, respectively.

<table>
<thead>
<tr>
<th>Panel A</th>
<th>Fund Net Alpha</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>PC-Adjusted Fund Net Alpha</th>
<th>(6)</th>
<th>(7)</th>
<th>(8)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lagged HHI</td>
<td></td>
<td>6.5277***</td>
<td>40.0796***</td>
<td>39.1271**</td>
<td>35.3591***</td>
<td>34.8497**</td>
<td>2.5033***</td>
<td>11.8816***</td>
<td>9.4626**</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.0362)</td>
<td>(4.8085)</td>
<td>(17.7560)</td>
<td>(4.6913)</td>
<td>(17.4796)</td>
<td>(0.8485)</td>
<td>(4.3336)</td>
<td>(4.1665)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lagged HHI^2</td>
<td></td>
<td>-1.110.4260***</td>
<td>-1.081.2070*</td>
<td>-1.402.0231***</td>
<td>-1.394.8592**</td>
<td>-310.3817**</td>
<td>-459.8112***</td>
<td>-310.3817**</td>
<td>-459.8112***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(156.2080)</td>
<td>(589.6182)</td>
<td>(184.4615)</td>
<td>(700.8128)</td>
<td>(142.7130)</td>
<td>(161.0837)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lagged Market Share</td>
<td></td>
<td>-12.0701</td>
<td>-15.1453</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(23.6434)</td>
<td>(24.6760)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lagged NHHI</td>
<td></td>
<td>7.8880***</td>
<td>46.4840***</td>
<td>45.7624**</td>
<td>44.3478***</td>
<td>44.2149**</td>
<td>2.9822***</td>
<td>12.4672**</td>
<td>11.4107**</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.2448)</td>
<td>(5.4655)</td>
<td>(20.1084)</td>
<td>(5.4142)</td>
<td>(20.0634)</td>
<td>(0.9912)</td>
<td>(4.9857)</td>
<td>(4.9085)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lagged NHHI^2</td>
<td></td>
<td>-1.468.6397***</td>
<td>-1.452.2241*</td>
<td>-1.884.2991***</td>
<td>-1.917.8119**</td>
<td>-360.9219*</td>
<td>-566.4512***</td>
<td>-360.9219*</td>
<td>-566.4512***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(204.1014)</td>
<td>(766.9316)</td>
<td>(244.4729)</td>
<td>(922.9528)</td>
<td>(187.9899)</td>
<td>(213.4361)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lagged Market Share</td>
<td></td>
<td>-11.3222</td>
<td>-14.6674</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(23.4579)</td>
<td>(24.8045)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lagged AFMI Size</td>
<td></td>
<td>-1.8946***</td>
<td>-1.9440</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.3977)</td>
<td>(1.4847)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td></td>
<td>245.178</td>
<td>245.178</td>
<td>239.537</td>
<td>245.178</td>
<td>239.537</td>
<td>245.179</td>
<td>245.179</td>
<td>245.179</td>
<td></td>
</tr>
<tr>
<td>R-Sqr</td>
<td></td>
<td>0.0002</td>
<td>0.0004</td>
<td>0.0004</td>
<td>0.0006</td>
<td>0.0006</td>
<td>0.0000</td>
<td>0.0001</td>
<td>0.0002</td>
<td></td>
</tr>
<tr>
<td>Adjusted R-Sqr</td>
<td></td>
<td>0.0002</td>
<td>0.0004</td>
<td>0.0004</td>
<td>0.0006</td>
<td>0.0006</td>
<td>0.0000</td>
<td>0.0001</td>
<td>0.0002</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel B</th>
<th>Fund Net Alpha</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>PC-Adjusted Fund Net Alpha</th>
<th>(6)</th>
<th>(7)</th>
<th>(8)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lagged 5FI</td>
<td></td>
<td>7.7880***</td>
<td>46.4840***</td>
<td>45.7624**</td>
<td>44.3478***</td>
<td>44.2149**</td>
<td>2.9822***</td>
<td>12.4672**</td>
<td>11.4107**</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.2448)</td>
<td>(5.4655)</td>
<td>(20.1084)</td>
<td>(5.4142)</td>
<td>(20.0634)</td>
<td>(0.9912)</td>
<td>(4.9857)</td>
<td>(4.9085)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lagged 5FI^2</td>
<td></td>
<td>-1.468.6397***</td>
<td>-1.452.2241*</td>
<td>-1.884.2991***</td>
<td>-1.917.8119**</td>
<td>-360.9219*</td>
<td>-566.4512***</td>
<td>-360.9219*</td>
<td>-566.4512***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(204.1014)</td>
<td>(766.9316)</td>
<td>(244.4729)</td>
<td>(922.9528)</td>
<td>(187.9899)</td>
<td>(213.4361)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lagged Market Share</td>
<td></td>
<td>-11.3222</td>
<td>-14.6674</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(23.4579)</td>
<td>(24.8045)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lagged AFMI Size</td>
<td></td>
<td>-1.6200***</td>
<td>-1.7313</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.3354)</td>
<td>(1.2683)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td></td>
<td>245.178</td>
<td>245.178</td>
<td>239.537</td>
<td>245.178</td>
<td>239.537</td>
<td>245.179</td>
<td>245.179</td>
<td>245.179</td>
<td></td>
</tr>
<tr>
<td>R-Sqr</td>
<td></td>
<td>0.0002</td>
<td>0.0004</td>
<td>0.0004</td>
<td>0.0006</td>
<td>0.0006</td>
<td>0.0000</td>
<td>0.0001</td>
<td>0.0002</td>
<td></td>
</tr>
<tr>
<td>Adjusted R-Sqr</td>
<td></td>
<td>0.0002</td>
<td>0.0004</td>
<td>0.0004</td>
<td>0.0006</td>
<td>0.0006</td>
<td>0.0000</td>
<td>0.0001</td>
<td>0.0002</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel C</th>
<th>Fund Net Alpha</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>PC-Adjusted Fund Net Alpha</th>
<th>(6)</th>
<th>(7)</th>
<th>(8)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lagged 5FI</td>
<td></td>
<td>0.7589***</td>
<td>5.1426***</td>
<td>5.0900</td>
<td>6.9152***</td>
<td>7.0271**</td>
<td>0.2624***</td>
<td>1.3057**</td>
<td>2.1864***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.1499)</td>
<td>(0.8780)</td>
<td>(3.2665)</td>
<td>(0.9451)</td>
<td>(3.5546)</td>
<td>(0.1118)</td>
<td>(0.7587)</td>
<td>(0.8358)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(2.2138)</td>
<td>(8.3851)</td>
<td>(8.4861)</td>
<td>(10.8268)</td>
<td>(2.4875)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lagged Market Share</td>
<td></td>
<td>-9.9791</td>
<td>-14.8748</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(23.0337)</td>
<td>(24.0807)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lagged AFMI Size</td>
<td></td>
<td>-1.8969***</td>
<td>-1.9487</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.3570)</td>
<td>(1.3386)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td></td>
<td>245.178</td>
<td>245.178</td>
<td>239.537</td>
<td>245.178</td>
<td>239.537</td>
<td>245.179</td>
<td>245.179</td>
<td>245.179</td>
<td></td>
</tr>
<tr>
<td>R-Sqr</td>
<td></td>
<td>0.0001</td>
<td>0.0002</td>
<td>0.0002</td>
<td>0.0006</td>
<td>0.0006</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0002</td>
<td></td>
</tr>
<tr>
<td>Adjusted R-Sqr</td>
<td></td>
<td>0.0001</td>
<td>0.0002</td>
<td>0.0002</td>
<td>0.0006</td>
<td>0.0006</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0002</td>
<td></td>
</tr>
</tbody>
</table>
Table 4. Industry-level Analysis: AFMI Active Share, Tracking Error and Market Concentration

This table reports the results of the main equations in various VAR models, where AFMI active share and tracking error are dependent variables. The sample period is from Quarter 1 of 1980 to Quarter 3 and 2009, and quarterly data is used. AFMI active share (AS) is measured as the average active share of active funds (in Petajisto’s database) in a quarter. AFMI tracking error (TE) is calculated as the average tracking error of active funds in a quarter. HHI, NHHI, and 5FI are Herfindahl-Hirschman index, normalized Herfindahl-Hirschman index, and 5-Fund index, respectively. Equation (1), (2), and (3) report the results of these three market concentration measures, respectively. Small-sample adjusted standard errors are used and presented in parentheses. The symbols ***, **, and * represent the 1%, 5%, and 10% significant level in a two-tail t-test, respectively.

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>AFMI AS</td>
<td>AFMI TE</td>
<td>AFMI AS</td>
</tr>
<tr>
<td>Lagged HHI</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.6279***</td>
<td>0.7652*</td>
<td>0.4732**</td>
</tr>
<tr>
<td></td>
<td>(0.1865)</td>
<td>(0.4302)</td>
<td>(0.2365)</td>
</tr>
<tr>
<td>Lagged NHHI</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0.9430***</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.0283)</td>
</tr>
<tr>
<td>Lagged 5FI</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.8979***</td>
<td>-0.0971</td>
<td>0.0406*</td>
</tr>
<tr>
<td></td>
<td>(0.0315)</td>
<td>(0.0728)</td>
<td>(0.0212)</td>
</tr>
<tr>
<td>Lagged AFMI AS</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-0.0299</td>
<td>0.7111***</td>
<td>-0.0111</td>
</tr>
<tr>
<td></td>
<td>(0.0280)</td>
<td>(0.0645)</td>
<td>(0.0279)</td>
</tr>
<tr>
<td>Lagged AFMI TE</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-0.0256</td>
<td>0.0884</td>
<td>0.0406*</td>
</tr>
<tr>
<td></td>
<td>(0.0240)</td>
<td>(0.0554)</td>
<td>(0.0212)</td>
</tr>
<tr>
<td>Constant</td>
<td>0.0753***</td>
<td>0.0884</td>
<td>0.0406*</td>
</tr>
<tr>
<td></td>
<td>(0.0240)</td>
<td>(0.0554)</td>
<td>(0.0212)</td>
</tr>
<tr>
<td>Observations</td>
<td>118</td>
<td>118</td>
<td>118</td>
</tr>
<tr>
<td>R-Sqr</td>
<td>0.969</td>
<td>0.574</td>
<td>0.967</td>
</tr>
</tbody>
</table>
Mathematical Proofs Appendix for

IS THE ACTIVE FUND MANAGEMENT INDUSTRY CONCENTRATED ENOUGH?

David Feldman, Konark Saxena, Jingrui Xu

Latest Revision October 7, 2017

Table of Contents

Proof of Managers' Maximization Problems Equivalence: Profits and Expected Net Alpha ..ii
Proof of Proposition 1 and Proposition 2 ........................................................................ iv
Proof of Proposition RA1 ..............................................................................................vii
Proof of Proposition RA2 and Corollary ..................................................................... viii
Proof of Proposition RA3 .............................................................................................. xi
Proof of Proposition RA4 .............................................................................................. xiii
Proof of Proposition PS and Corollary ........................................................................ xv
Proof of Proposition RN1 .............................................................................................. xvi
Proof of Managers' Maximization Problems Equivalence: Profits and Expected Net Alpha

We will prove that when managers maximize fund expected net alphas they maximize profits, and that they must do so in order to survive.

Suppose that a manager chooses profit maximizing optimal effort and fees to set her fund expected net alpha to be $\bar{\alpha}$. We will show that in equilibrium $\bar{\alpha}$ must be the maximum fun expected net alpha that managers can produce (while staying solvent). Substituting $\bar{\alpha}$ into Equation (7) (our “state” equation, that links effort, fees, and fund expected net alphas), yields,

$$f_i^* = a - \hat{b} \frac{S}{W} + A(e_i^*; H) - \bar{\alpha}.$$  \hspace{1cm} (A1)

Denote the profit rate of manager $i$, $pro_i$, as $pro_i \equiv f_i^* - C_i(e_i^*; s_i; H)$. Substituting Equation (A1) into $pro_i$, then, substituting Equation (18) for the cost function, and rearranging, we have

$$\bar{\alpha} = a - \hat{b} \frac{S}{W} + A(e_i^*; H) - pro_i - c_{0,i} - c_{1,i} s_i - c_{2,i}(e_i^*; H).$$ \hspace{1cm} (A2)

If all managers produce the same level of fund expected net alphas (we later demonstrate that this, indeed, is the case in equilibrium), Equation (A2) implies a relation,

$$pro_i + c_{1,i} s_i = pro_j + c_{1,j} s_j, \quad \forall i, j.$$ \hspace{1cm} (A3)

Now, consider manager $i$’s profit function

$$pro_i = s_i[f_i^* - c_{0,i} - c_{1,i} s_i - c_{2,i}(e_i^*; H)], \quad \forall i.$$ \hspace{1cm} (A4)

For a given (non-zero) $pro_i$, the first-order condition for optimal fund size is,

$$s_i^* = \frac{f_i^* - c_{0,i} - c_{2,i}(e_i^*; H)}{2c_{1,i}}.$$ \hspace{1cm} (A5)

If we express the numerator of the last equation in terms of the current $pro_i$, we can rewrite it as

$$s_i^* = \frac{pro_i + s_i}{2c_{1,i}}.$$ \hspace{1cm} (A6)

Equation (A6) relates the optimal fund size and its current size.$^{41}$ It shows that profit

---

$^{41}$Note that Equation (A6) is defined only when profits are non-zero. The size derivative of Equation (A4) is not defined.
rates that are too high (with respect to profit maximizing profit rates) are associated with fund sizes that are too small. Thus, some manager \( j, j \neq i \), might have an incentive to increase fund expected net alpha, reduce profit rates, increase size, and increase profits. Or, vice versa.\(^{42}\) Therefore, although manager \( i \) does not observe other managers’ cost functions, she knows that it is possible that some other manager(s) might have incentive to increase expected net alpha by lowering lower profit rates, attracting investments, increasing fund size, and increasing profits.

We will now analyze a simple game between manager \( i \) and other managers, grouped as an entity “\(-i\)”, that accounts for the possibility of a manager increasing fund expected net alphas. If manager \( i \) improves her fund expected net alpha infinitesimally, then, other managers receive no investments and earn no profits, and manager \( i \)’s profits will change by an infinitesimal amount, say, \( \eta_i \). If, on the other hand, manager \(-i\) (any of the other managers) increases her fund’s expected net alpha infinitesimally, then, manager \( i \) receives no investments and earns no profits, and manager \(-i\)’s profits change by \( \eta_{-i} \). Thus, a necessary condition for all managers to make profits is that all of them produce the same level of fund expected net alphas. Note that \( \eta_i (\eta_{-i}) \) can be positive or negative, depending on whether manager \( i \)’s (\(-i \)’s) fund size is below or above optimal level. Suppose that manager \( i \) believes that manager \(-i\)’s strategy is to improve her fund expected net alpha, \( \bar{\alpha} \), with probability \( p \) and maintain \( \bar{\alpha} \) with probability \( 1 - p \). (Again, this is because manager \( i \) knows that it is possible for some other managers to improve their fund expected net alpha to attract investments in order to improve their funds’ profits, and that this probability, \( p \), is nontrivial.) Suppose that manager \( i \)’s strategy is to improve her fund expected net alpha with probability \( \theta \) and maintain \( \bar{\alpha} \) with probability \( 1 - \theta \).

We will show that in this game, manager \( i \) optimally chooses \( \theta = 1 \), till reaching the highest level of fund expected net alphas (beyond which she becomes insolvent). As manager \( i \) is a generic manager, this implies that all managers do that. We will also show that once managers reach the points of producing the highest level of fund expected net alphas, they are in (a Nash) equilibrium.

The payoffs of such a game are illustrated in the following table, with the row (column) representing manager \( i \)’s (\(-i \)’s) action, and with manager \( i \)’s (\(-i \)’s) payoffs in the first

\(^{42}\) That is, profit rates that are too low are associated with fund sizes that are too large. Thus, some manager \( j, j \neq i \), might have an incentive to decrease expected net alpha, increase profit rates, decrease size, and increase profits.
(second) figures in the brackets.

<table>
<thead>
<tr>
<th>Maintain $\bar{\alpha}$</th>
<th>Improve Infinitesimally</th>
</tr>
</thead>
<tbody>
<tr>
<td>$1 - p$</td>
<td>$p$</td>
</tr>
<tr>
<td>$1 - \theta$</td>
<td>$(\text{pro}<em>i s_i, \text{pro}</em>{-i} s_{-i})$</td>
</tr>
<tr>
<td>$\theta$</td>
<td>$(\text{pro}_i s_i + \eta_i, 0)$</td>
</tr>
</tbody>
</table>

Then, the expected payoffs of manager $i$ is

$$\pi_i = (1 - \theta)(1 - p) \text{pro}_i s_i + \theta(1 - p)(\text{pro}_i s_i + \eta_i) + p(\text{pro}_i s_i + \eta_i). \quad (A7)$$

The first-order derivative is with respect to $\theta$, is

$$\frac{d\pi_i}{d\theta} = \eta_i + p \times \text{pro}_i s_i. \quad (A8)$$

Equation (A8) shows that $\eta_i \to 0$, implies that $d\pi_i/d\theta > 0$. Thus, manager $i$’s optimal choice to maximize $\pi_i$ is $\theta = 1$. That is, increasing fund expected net alphas increases profits.

As managers keep increasing fund expected net alphas, they reach the highest level of fund expected net alpha where $\bar{\alpha}$ becomes the maximum fund expected net alpha. At this point, managers’ profit rates are zero (otherwise managers could use profits to increase fund expected net alphas). Moreover, further increases of fund expected net alphas (by increasing effort levels or decreasing fees) make managers insolvent. Thus, at this point, where $\bar{\alpha}$ is the optimal fund expected net alpha, $\eta_i$ and $\eta_{-i}$ are negative. Managers are, then, in a Nash equilibrium (Maintain $\bar{\alpha}$, Maintain $\bar{\alpha}$).

Because at any fund expected net alpha level, which is below the maximizing level, managers attract no investments, thus not in equilibrium as they have incentives to increase fund expected net alphas, and because further increasing fund expected net alpha above the maximizing level, drives managers to insolvency, this Nash equilibrium is unique.

Thus, managers’ problems of maximizing profits is equivalent to their maximizing fund expected net alphas.

Q.E.D.

**Proof of Proposition 1 and Proposition 2**

To maximize $E(\alpha_i \mid D)$, manager $i$ chooses the breakeven management fee (higher fee would decrease alpha and lower fee would induce insolvency), that is

$$f_i^* - C'(e_i^*, s_i; H) = 0. \quad (A9)$$
This proves Point 1 of Proposition 1.

We defined the direct benefits of effort in Equation (21). If their partial derivative with respect to effort, at zero effort is positive, i.e., \( A_{e_i} (0; H) - c_{2e_i} (0; H) > 0 \), then, it pays to exert effort and the optimal effort level is strictly positive, i.e., \( e_i^* > 0 \). The first-order condition, with respect to effort, to maximize \( E(\alpha_i \mid D) \) becomes,

\[
A_{e_i} (e_i^*; H) - c_{2e_i} (e_i^*; H) = B_{e_i} (e_i^*; H) = 0.
\] (A10)

The related second-order condition, \( A_{e_i e_i} (e_i^*; H) - c_{2e_i e_i} (e_i^*; H) = B_{e_i e_i} (e_i^*; H) < 0 \), is satisfied by assumptions. (This is because we assume that productivity effort decreases in scale, i.e., \( A_{e_i e_i} (e_i; H) < 0, \forall e_i \), and that the costs of effort increase in scale, i.e., \( c_{2e_i e_i} (e_i; H) > 0, \forall e_i \). Thus \( e_i^* \) is a maximum. (We assume that functional forms of effort productivities and effort costs induce a finite \( e_i^* \).)

This proves Point 2 of Proposition 1.

Next, as both \( f_i^* \) and \( e_i^* \) are functions of \( H \), we can write,

\[ e_i^* = e_j^* (H), \] (A11)

and,

\[ f_i^* = f_j^* (H). \] (A12)

Complete differentiation, of the left-hand side, of (A10) gives

\[
e_i^* (H) = \frac{A_{e_i e_i} (e_i^*; H) - c_{2e_i e_i} (e_i^*; H)}{A_{e_i} (e_i^*; H) - c_{2e_i} (e_i^*; H)}. \] (A13)

Thus, if the numerator of Equation (A13), \( A_{e_i e_i} (e_i^*; H) - c_{2e_i e_i} (e_i^*; H) \geq 0(<0) \), then \( e_i^* (H) \geq 0(<0) \). (We showed above that the denominator of Equation (A13) is negative.)

This proves Point 3 of Proposition 1.

The optimal manager effort \( e_i^* \) is determined only by the functions \( c_{2} (e_i; H) \), and \( A(e_i; H) \), which are the same across funds. Thus, we have \( e_i^* = e_j^* \) and \( B(e_i^*; H) = B(e_j^*; H), \forall i, j \). Because, in equilibrium, managers produce the (same) level of fund expected net alphas (Point 6 of Proposition 1), and as we proved above, and (as we just showed) exert the same optimal effort levels (i.e., \( e_i^* = e_j^*, \forall i, j \)), from Equation (A1) we have that
\( f_i^* = f_j^*, \forall i, j. \)

These prove Points 1 and 4 of Proposition 2.

In addition, by Equation (A9), we further have \( C^i(e_i^*, s_i; H) = C^j(e_j^*, s_j; H), \forall i, j. \)

Recall that \( C^i(e_i^*, s_i; H) = c_0 + c_{1,i} s_i + c_2(e_i^*; H) \). As \( c_0, e_i^*, \) and \( C^i(e_i^*, s_i; H) \) are the same across funds, we have the following relationship between different funds’ sizes and costs,

\[
c_{1,i} s_i = c_{1,j} s_j, \forall i, j,
\]

or, \( s_i / s_j = c_{1,i} / c_{1,j}, \forall i, j. \)

This proves Point 2 of Proposition 2.

Summing \( s_i / s_j \) with respect to \( i, i=1,2,...,M \), we have \( \sum_{i=1}^{M} s_i / s_j = S / s_j = \sum_{i=1}^{M} c_{1,i} / c_{1,j} \), where we use Equation (A14) to write the second equality. Inversing the second equality, and exchanging the subscripts \( j \) and \( i \), gives

\[
\frac{s_i}{S} = \left( c_{1,i} \sum_{j=1}^{M} (c_{1,j}^{-1}) \right)^{-1}, \forall i.
\]

This proves Point 3 of Proposition 2.

Using the breakeven fee condition and Equation (A15), we can write

\[
f_i^* = C^i(e_i^*, s_i; H) = c_0 + c_{1,i} s_i + c_2(e_i^*; H)
= c_0 + c_{1,i} s_i \frac{S}{s_j} W + c_2(e_i^*; H)\]
\[
= c_0 + c_{1,i} \frac{S}{s_j} W \left( c_{1,i} \sum_{j=1}^{M} (c_{1,j}^{-1}) \right)^{-1} + c_2(e_i^*; H).
\]

Differentiating \( f_i^* \), as in the last equation, with respect to \( H \), yields

\[
f_i^*(H) = c_{1,i} \frac{S}{s_j} W \left( c_{1,i} \sum_{j=1}^{M} (c_{1,j}^{-1}) \right)^{-1} d(S/W) + c_2(e_i^*; H)e_i^*(H) + c_2(e_i^*; H).
\]

Thus, whether higher concentrations induce higher equilibrium optimal fees depends on whether they induce an increase in equilibrium industry sizes \( (d(S/W)/dH) \) and whether they induce an increase in equilibrium optimal effort levels \( (e_i^*(H)) \).

This proves Point 4 of Proposition 1.

Differentiation of \( B(e_i^*; H) \) with respect to \( H \), and use of Equation (A10) give
Thus, if \( A_H(e_i^*; H) - c_{2H}(e_i^*; H) \geq 0(<0) \), then \( \frac{dB(e_i^*; H)}{dH} \geq 0(<0) \).

This proves Point 5 of Proposition 1.

Taking expectations of both sides of Equation (1), for fund \( i \), yields

\[
E(r_{F,i}|D) = E(\alpha_i|D) + \mu_p.
\]

(A19)

Because \( E(\alpha_i|D) \)'s are the same across funds in equilibrium, \( E(r_{F,i}|D) \)'s are the same across funds in equilibrium. Further, we have

\[
\text{Var}(r_{F,i}|D) = \sigma_p^2 + \sigma_u^2 + \sigma_e^2 \left( \frac{S}{W} \right)^2 + \sigma_s^2 + \sigma_v^2,
\]

(A20)

Implying that fund returns’ variances, \( \text{Var}(r_{F,i}|D) \)'s, are the same across funds. Combining (A19) and (A20) shows that all managers offer the same market competitive Sharpe ratio.

This proves Point 7 of Proposition 1.

Q.E.D.

Proof of Proposition RA1

\( \{e^*, f^*, \delta^*\} \) is a Nash equilibrium because,

1. Given other managers’ optimal choices, a manager has incentives to not deviate from \( e^* \) and/or \( f^* \). If a manager deviates from \( e^* \) or \( f^* \) they either decrease their fund expected net alpha losing all investment or becoming insolvent. They also cannot deviate from both in offsetting ways and gain. This is because, effort increases do not sufficiently improve performance to justify costs, and fee increase, and effort reductions will cause loss of too much performance that cannot be returned to investors by fee reductions. This is because, a manager’s optimal effort and fee together determine her fund expected net alpha. If she deviates from the equilibrium and produces a higher fund expected net alpha, she incurs a loss, and if she deviates and produces a lower fund expected net alpha, she receives no investments. We proved these results in the previous proof.

2. Given managers’ and other investors’ optimal choices, an investor has no incentive to deviate from \( \delta_{j}^* \). This is because, where there are infinitely many small mean-variance
risk-averse investors, each investor’s choice does not affect fund sizes and, thus, the AFMI size, changing allocations across funds does not improve an investor’s portfolio’s Sharpe ratio, whereas changing allocations between the AFMI and the passive benchmark decreases the portfolio’s Sharpe ratio.

\( \{e^*, f^*, \delta^*\} \) is unique because,

1. \( e^* \) is unique because, for each fund, \( e_i^* \) is a unique solution to \( B_i(e_i, H) = 0 \);
2. \( f^* \) is unique because for each fund \( f_i^* - C'(e_i^*, s_i; H) = 0 \), where \( C'(e_i^*, s_i; H) \) is a deterministic function of \( e_i^* \), and \( e_i^* \) is unique;
3. \( \delta^* \) is unique because allocations to funds maximize investor portfolios’ Sharpe ratios, driving fund expected net alphas to the same values. Deviating, thus, cannot help and to the extent that large deviation would affect fund sizes, they will decrease Sharpe ratios. Moreover, the uniqueness of \( e^* \) and \( f^* \) rules out existence of additional equilibrium allocations.

Q.E.D.

**Proof of Proposition RA2 and Corollary**

**Infinitely Many Small Mean-Variance Risk-Averse Investors**

In this case, investors maximize their portfolio Sharpe ratios subject to the constraints described by Equations (13) and (14) in the paper (wealth constrains, and no fund short selling constraints). Investors are small, so none affects fund sizes (i.e., \( s_i, i = 1, \ldots, M \)) and industry size \( S/W \).

We note that

\[
E(r_j | D) = \mu_p + \delta_j^T E(a | D) = \mu_p + \delta_j^T \left[ a - \hat{b} \frac{S}{W} + A(e_i^*; H) - f_i^* \right]_M, \ \forall i, j. \quad \text{(A21)}
\]

The first equality is Equation (10). The second equality holds because equilibrium fund expected net alphas are the same, as we show in Point 6 of Proposition 1. Also, as we assume that there are no marginal diversification benefits across funds, implying that the term \( \sigma_e^2 \delta_j^T \delta_j \) is zero, and get

\[
\text{Var}(r_j | D) = \sigma_p^2 + \left[ \sigma_a^2 + \sigma_s^2 + \sigma_h^2 \left( \frac{S}{W} \right)^2 \right] \left( \delta_j^T t_M \right)^2, \ \forall j. \quad \text{(A22)}
\]
In this section, we analyze the behavior of investors in a market where they can choose between a risky asset and two risk-free assets. Investors maximize
\[
\frac{E(r_j | D)}{\sqrt{\text{Var}(r_j | D)}}
\]
by choosing \( \delta_j \). The first-order condition (i.e., the derivative of \( \frac{E(r_j | D)}{\sqrt{\text{Var}(r_j | D)}} \) with respect to \( \delta_j \), set to be 0), when substituting the constraints
\[
f_j^* - C^i(e_j^*, s_i; H) = 0, \quad \forall i,
\]
into it, is
\[
-\left( \frac{\mu_i}{\sigma_p^2} \right) \left[ \sigma_a^2 + \sigma_b^2 \left( \frac{S}{W} \right)^2 + \sigma_x^2 \right] \delta_j^* \mathbf{1}_M + a - \hat{b} \frac{S}{W} + A(e_j^*; H) - c_0 - c_1 s_i - c_2 (e_j^*; H) = 0.
\]
\[
= -\left( \frac{\mu_i}{\sigma_p^2} \right) \left[ \sigma_a^2 + \sigma_b^2 \left( \frac{S}{W} \right)^2 + \sigma_x^2 \right] \delta_j^* \mathbf{1}_M - \hat{b} \left( \sum_{i=1}^{M} c_i^{-1} \right) \frac{S}{W} + X(e_j^*; H) = 0.
\]
\[
(A23)
\]
The first equality of Equation (A23) holds because
\[
c_{i,s} = \left( c_{i,s} \frac{s}{S} \right) W \frac{S}{W},
\]
and we obtain the second equality by using the definition of \( X(e_j^*; H) \) and Equation (A15).

Substituting \( \gamma \overset{\Delta}{=} \frac{\mu_i}{\sigma_p^2} \) and \( S/W = \delta_j^* \mathbf{1}_M \) (symmetric equilibrium) into Equation (A23), we have
\[
-\gamma \sigma_a^2 \left( \frac{S}{W} \right)^3 - \left[ \gamma \sigma_a^2 + \gamma \sigma_x^2 + \hat{b} + \left( \sum_{i=1}^{M} c_i^{-1} \right) \frac{S}{W} + X(e_j^*; H) \right] = 0.
\]
\[
(A25)
\]
If the constraint \( \delta_j^* \mathbf{1}_M \leq 1 \) is not binding (i.e., \( S/W < 1 \)), the equilibrium optimal \( S/W \) is a real positive solution of this cubic equation. This is because, the condition \( X(e_j^*; H) > 0, \quad \forall H \) (positivity of the lowest order polynomial coefficient) and the negativity of the two higher order polynomial coefficients \( -\gamma \sigma_a^2 \) and \( -\left[ \gamma \sigma_a^2 + \gamma \sigma_x^2 + \hat{b} + \left( \sum_{i=1}^{M} c_i^{-1} \right) \frac{S}{W} \right] \), (i.e., \( -\gamma \sigma_a^2 < 0 \) and \( -\left[ \gamma \sigma_a^2 + \gamma \sigma_x^2 + \hat{b} + \left( \sum_{i=1}^{M} c_i^{-1} \right) \frac{S}{W} \right] < 0 \) ) guarantee the existence of exactly one positive real solution for \( S/W \) (and two imaginary ones). Also, as each investor cannot affect the value of \( S/W \), Equation (A23) shows that the solution for \( \delta_j^* \mathbf{1}_M = S/W \)
is unique given the parameter values and the market $S/W$.

If the constraint $\delta^T s \leq 1$ is binding, (i.e., $S/W = 1$), there is, obviously unique solution, where investors maximize their portfolios Sharpe ratios by allocating all their wealth to AFMI (no passive index holdings).

In addition, we have

$$E(\alpha_i | D)\bigg|_{[e^*, t^*, s^*]} = a - \hat{b} \left( \frac{S}{W} \right) + \hat{A} \left( e^*_i; H \right) - f^*_i$$

$$= a - \hat{b} \left( \frac{S}{W} \right) + A \left( e^*_i; H \right) - c_0 - c_{1,i} s_i - c_2 \left( e^*_i; H \right)$$

$$= - \left[ \hat{b} + \left( \sum_{i=1}^{M} \gamma c_{1,i} \right)^{-1} W \right] \left( \frac{S}{W} \right) + X \left( e^*_i; H \right).$$

(A26)

The first equality of Equality (A26) follows from the equilibrium breakeven management fee condition, and the second equality of Equation (A26) follows from Equation (A24). Substituting Equation (A26) into Equation (A25), and rearranging, we have

$$E(\alpha_i | D)\bigg|_{[e^*, t^*, s^*]} = \frac{\gamma \left( \sigma_a^2 + \sigma_s^2 \right) X \left( e^*_i; H \right)}{\gamma \left( \sigma_a^2 + \sigma_s^2 \right) + \hat{b} + \left( \sum_{i=1}^{M} \gamma c_{1,i} \right)^{-1} W}.$$  

(A27)

where

$$\sigma_a^2 \triangleq \text{Var}(\alpha_j | D) = \sigma_b^2 \left( \frac{S}{W} \right)^2 + \sigma_a^2.$$  

(A28)

Because all the components of Equation (A27) are positive, $E(\alpha_i | D)\bigg|_{[e^*, t^*, s^*]}$ is positive. The intuition is as follows. From investors’ portfolio variance formulas, Equation (A22), we can easily see that portfolios with allocations to AFMI have higher variance than those holding the passive benchmark. If $E(\alpha_i | D)\bigg|_{[e^*, t^*, s^*]} = 0$, because of a sufficiently large amount of investment in funds, investors can always improve their portfolio Sharpe ratios (in particular, reduce their portfolios risk) by shifting wealth from AFMI to the passive benchmark.

This proves Proposition RA2.

Where the equilibrium optimal AFMI size is less than one (i.e., $S/W < 1$), differentiation of Equation (A25) with respect to $X \left( e^*_i; H \right)$, gives

$$\frac{d \left( \frac{S}{W} \right)}{d X} = \frac{1}{\gamma \left( 3 \sigma_b^2 \left( \frac{S}{W} \right)^2 + \sigma_a^2 + \sigma_s^2 \right) + \hat{b} + \left( \sum_{i=1}^{M} \gamma c_{1,i} \right)^{-1} W}.$$  

(A29)
As all the components of Equation (A29) are positive, we have that
\[
\frac{d(S/W)}{dX} > 0.
\] (A30)

This proves Point 1 of the corollary of Proposition RA2.

Differentiation of Equation (A25) with respect to \( \hat{b} + \left( \sum_{i=1}^{M} c_{i,1}^{-1} \right)^{-1} W \) gives
\[
\frac{d(S/W)}{d \left[ \hat{b} + \left( \sum_{i=1}^{M} c_{i,1}^{-1} \right)^{-1} W \right]} = -\frac{(S/W)}{\gamma} \left[ 3\sigma_b^2 \left( \frac{S}{W} \right)^2 + \sigma_a^2 + \sigma_x^2 \right] + \hat{b} + \left( \sum_{i=1}^{M} c_{i,1}^{-1} \right)^{-1} W
\]
\[
= -\frac{S}{W} \frac{d(S/W)}{dX} < 0.
\] (A31)

The last inequality holds because of Equation (A30) and the positivity of \( S/W \).

This proves Point 2 of the corollary of Proposition RA2.

**Q.E.D.**

**Proof of Proposition RA3**

By the chain rule,
\[
\frac{d(S/W)}{dH} = \frac{d(S/W)}{dx} \frac{dx(e_i^*,H)}{dH} = \frac{d(S/W)}{dx} \frac{dB(e_i^*,H)}{dH}.
\] (A32)

The second equality of Equation (A32) follows from the definitions of \( x(e_i^*,H) \) and \( B(e_i^*,H) \) in equations (25) and (21) respectively. Recalling that \( \frac{d(S/W)}{dx} > 0 \), we see that the sign of \( d(S/W)/dH \) is determined by the sign of \( dB(e_i^*,H)/dH \).

Also, differentiating Equation (A32) again with respect to \( H \), we have
\[
\frac{d^2(S/W)}{dH^2} = \left( \frac{d(S/W)}{dx} \right)^2 \frac{d^2B(e_i^*,H)}{dH^2} + \left[ \frac{dB(e_i^*,H)}{dH} \right] \frac{d^2(S/W)}{dx^2} + \left[ \frac{dB(e_i^*,H)}{dH} \right]^2 \left[ \frac{d(S/W)}{dx} \right]^2 \gamma [6\sigma_b^2 \frac{S}{W}].
\] (A33)

The first equality of Equation (A33) holds because of Equation (A36) and the 2nd derivative chain rule. To show that the second equality holds, we first, differentiate Equation (A29) again,
to get,

\[ \frac{d^2 (S/W)}{dX^2} = - \left[ \frac{d (S/W)}{dX} \right]^3 \gamma \left[ 6\sigma^2 \frac{S}{W} \right], \]

(A34)

and then substitute the result.

We can write the second order derivative of \( B(e^*_i, H) \) with respect to \( H \), as

\[ \frac{d^2 B(e^*_i; H)}{dH^2} = \frac{d}{dH} \left[ \frac{dB(e^*_i; H)}{dH} \right] = \frac{d}{dH} \left[ A_{ii}(e^*_i; H) - c_{2ii}(e^*_i; H) \right], \]

(A35)

where the second equality of Equation (A35) holds because

\[ \frac{dB(e^*_i; H)}{dH} = \left[ A_{ii}(e^*_i; H) - c_{2ii}(e^*_i; H) \right] e^*_i(H) + A_{ii}(e^*_i; H) - c_{2ii}(e^*_i; H) \]

(A36)

where the last equality of Equation (A36) holds because of the effort optimality condition.

Noting that \( \frac{d(S/W)}{dX} > 0 \) from Equation (A30), \( \gamma \left[ 6\sigma^2 \frac{S}{W} \right] > 0 \), and

\[ \left[ A_{ii}(e^*_i; H) - c_{2ii}(e^*_i; H) \right]^2 \geq 0, \] Equation (A33) implies that \( \frac{d^2 B(e^*_i; H)}{dH^2} < 0 \) implies \( \frac{d^2 (S/W)}{dH^2} < 0 \), whereas \( \frac{d^2 (S/W)}{dH^2} > 0 \) implies \( \frac{d^2 B(e^*_i; H)}{dH^2} > 0 \).

Differentiating Equation (A26) with respect to \( H \), and using Equations (A32) and (A36), we have

\[ \frac{dE(\alpha_i | D)}{dH} \bigg|_{e^*, \delta^*} = \left\{ 1 - \left[ \hat{b} + \left( \sum_{i=1}^{M_i} c_{i,i}^{-1} \right)^{-1} \frac{W}{dS/W} \right] \frac{d(S/W)}{dX} \right\} \frac{dB(e^*_i; H)}{dH}. \]

(A37)

Substituting Equation (A29) into the left hand side of Equation (A38), yields,

\[ 1 - \left[ \hat{b} + \left( \sum_{i=1}^{M_i} c_{i,i}^{-1} \right)^{-1} \frac{W}{dS/W} \right] \frac{d(S/W)}{dX} = \frac{\gamma \left[ 3\sigma^2_a + \sigma^2_s \right]}{\gamma \left[ 3\sigma^2_b \left( \frac{S}{W} \right)^2 + \sigma^2_a + \sigma^2_x \right] + \hat{b} + \left( \sum_{i=1}^{M_i} c_{i,i}^{-1} \right)^{-1} \frac{W}{dS/W}} > 0. \]

(A38)

The last inequality of Equation (A38) holds because all the parameters and variables values in the equation are positive. Then, from Equation (A38), Equation (A37) implies that the sign of
\[ dE(\alpha_i \mid D) / dH \mid_{e_i, r_i, A} \] is determined by the sign of \( dB(e_i^*, H) / dH \).

Also, differentiating Equation (A37) again, with respect to \( H \), from Equations (A36), (A35), and (A34), we have

\[
\frac{d^2 E(\alpha_i \mid D)}{dH^2} \bigg|_{e_i, r_i, \theta} = \frac{d^2 B(e_i^*; H)}{dH^2} \left[ 1 - \left( \hat{b} + \left( \sum_{i=1}^{M} c_{i,j} \right)^{-1} W \right) \frac{d(S / W)}{dX} \right] \left( \frac{d(S / W)}{dX} \right)^2 \gamma \left[ 6\sigma^2_b \frac{S}{W} \right].
\]

As \( 1 - \left( \hat{b} + \left( \sum_{i=1}^{M} c_{i,j} \right)^{-1} W \right) \frac{d(S / W)}{dX} > 0 \) by Equation (A38), \( d(S / W) / dH > 0 \) by Equation (A30), \( \left[ A_{ij}(e_i^*; H) - c_{2i}(e_i^*; H) \right] \geq 0 \), \( \hat{b} + \left( \sum_{i=1}^{M} c_{i,j} \right)^{-1} W > 0 \), and \( \gamma \left[ 6\sigma^2_b \frac{S}{W} \right] > 0 \), we have that \( d^2 E(\alpha_i \mid D) / dH^2 \bigg|_{e_i, r_i, \theta} < 0 \) implies \( d^2 B(e_i^*; H) / dH^2 < 0 \), whereas \( d^2 B(e_i^*; H) / dH^2 > 0 \) implies \( d^2 E(\alpha_i \mid D) / dH^2 \bigg|_{e_i, r_i, \theta} > 0 \).

Putting the results in this section together, we have

\[
\frac{d(S / W)}{dH} \geq 0(< 0) \Leftrightarrow \left( \frac{dB(e_i^*; H)}{dH} \right) \geq 0(< 0) \Leftrightarrow \left( \frac{dE(\alpha_i \mid D)}{dH} \right) \geq 0(< 0),
\]

\[
\frac{d^2 (S / W)}{dH^2} > 0 \Rightarrow \left( \frac{dB(e_i^*; H)}{dH^2} \right) > 0 \Rightarrow \left( \frac{d^2 E(\alpha_i \mid D)}{dH^2} \right) > 0,
\]

and

\[
\frac{d^2 E(\alpha_i \mid D)}{dH^2} \bigg|_{e_i, r_i, \theta} < 0 \Rightarrow \left( \frac{dB(e_i^*; H)}{dH^2} \right) < 0 \Rightarrow \frac{d(S / W)}{dH^2} > 0.
\]

This proves Proposition RA3.

\[ Q.E.D. \]

**Proof of Proposition RA4**

Differentiating Equation (A25) with respect to \( c_{i,j} \) gives
\[
\frac{d(S/W)}{dc_{i,t}} = -\frac{(S/W)W}{\left\{ \gamma \left[ 3\sigma_y^2 (S/W)^2 + \sigma_a^2 + \sigma_x^2 \right] + \hat{b} + \left( \sum_{l=1}^{M} c_{l,t}^{-1} \right) W \right\} \left( \sum_{l=1}^{M} c_{l,t}^{-1} \right)^2 c_{i,t}^2} < 0.
\]

(A43)

where the last inequality holds because all the parameter and variable values in Equation (A43) are positive, and we have a negative sign in the numerator.

Differentiating Equation (A26) with respect to \( c_{i,t} \), we have

\[
\frac{dE(\alpha_i | D)}{dc_{i,t}} \bigg|_{[e', r', \delta']} = \left\{ \gamma \left[ 3\sigma_y^2 (S/W)^2 + \sigma_a^2 + \sigma_x^2 \right] + \hat{b} + \left( \sum_{l=1}^{M} c_{l,t}^{-1} \right) W \right\} \left( \sum_{l=1}^{M} c_{l,t}^{-1} \right)^2 c_{i,t}^2 < 0,
\]

(A44)

where the second equality of Equation (A44) holds after substituting Equation (A43). The last inequality of Equation (A44) holds because all the parameters and variables in Equation (A44) are positive and we have a negative sign in the numerator.

Similarly, differentiating Equation (A26) with respect to \( c_{i,j} \), yields

\[
\frac{dE(\alpha_i | D)}{dc_{i,j}} \bigg|_{[e', r', \delta']} = \left\{ \gamma \left[ 3\sigma_y^2 (S/W)^2 + \sigma_a^2 + \sigma_x^2 \right] + \hat{b} + \left( \sum_{l=1}^{M} c_{l,t}^{-1} \right) W \right\} \left( \sum_{l=1}^{M} c_{l,t}^{-1} \right)^2 c_{i,j}^2 < 0.
\]

(A45)

For the case where \( S/W = 1 \), from Equation (A26), we have

\[
E(\alpha_i | D) \big|_{[e', r', \delta']} = \left[ \hat{b} + \left( \sum_{l=1}^{M} c_{l,t}^{-1} \right) W \right] + X(e_i^*; H). \quad \text{Then,}
\]

\[
\frac{dE(\alpha_i | D)}{dc_{i,t}} \bigg|_{[e', r', \delta']} = \frac{-W}{\left( \sum_{l=1}^{M} c_{l,t}^{-1} \right)^2 c_{i,t}^2} < 0,
\]

(A46)

\[
\frac{dE(\alpha_i | D)}{dc_{i,j}} \bigg|_{[e', r', \delta']} = \frac{-W}{\left( \sum_{l=1}^{M} c_{l,t}^{-1} \right)^2 c_{i,j}^2} < 0,
\]

(A47)
where the last inequalities of Equations (A46) and (A47) hold because all the parameter and variable values are positive and we have a negative sign in the numerators.

This proves Point 2 of Proposition RA4.

As $s_i / S$ decreases in $c_{i,i}$, whereas $s_j / S, \forall j \neq i$ increases in $c_{i,j}$, from Equations (A15), (A44), (A45), (A46), and (A47), we find that $E(\alpha_i | D)_{|s^*,s^*}^*$ and $s_i / S$, are increasing/decreasing in the same direction due to changes in $c_{i,i}$, and that $E(\alpha_j | D)_{|s^*,s^*}^*$ and $s_j / S, \forall j \neq i$, are increasing/decreasing inversely due to changes in $c_{i,i}$, whether $S/W < 1$ or $S/W = 1$.

This proves Point 1 of Proposition RA4.

This proves Proposition RA4.

Q.E.D.

Proof of Proposition PS and Corollary

As we know from previous propositions, in equilibrium, managers are earning zero economic profit. Thus, we substitute $f_i = C_i(e_i, s_i; H)$ into $E(\alpha_i | D)$, and then perform first-order differentiation with respect to $e_i$. We get $A_{e_i}(e_i^*; H) - c_{2e_i}(e_i^*; H) = 0$. The second-order condition $A_{e_i,e_i}(e_i; H) - c_{2e_i,e_i}(e_i; H) < 0$, holds for all effort levels by our assumptions on the functional forms.

By construction, effort levels cannot be negative, i.e., $e_i^* \geq 0$. If $A_{e_i}(0; H) - c_{2e_i}(0; H) < 0$, managers (unable to exert negative effort levels) spend no effort, i.e., $e_i^* = 0$. In this case, managers charge an optimal proportional fee (break even fee) equal to $f_i^* = c_0 + c_{1i}s_i$.

This proves Proposition PS.

Then, our equilibrium is similar to that in Pastor and Stambaugh (2012), where managerial effort is not modeled [our $A(e_i; H)$ and $c_2(e_i; H)$ are both zero] and managers do not charge fees above opportunity costs.

This proves the corollary to Proposition PS.

Q.E.D.
Proof of Proposition RN1

Infinitely Many Small Risk-Neutral Investors

Risk neutral investors keep allocating wealth to AFMI as long as fund expected net alphas are positive, maximizing their portfolio expected net returns.

In the case that investors have sufficient wealth, fund expected net alphas are driven down to zero. In equilibrium, if investors have additional wealth available to allocate to AFMI but funds have exhausted their abilities to produce positive fund expected net alphas, technically stating, \( S/W < 1 \), and
\[
E(\alpha_i \mid D)\big|_{\{e^*, r^*, \delta^*\}} = a - \hat{b} \left( \frac{S}{W} \right) + A(e^*_i; H) - f^*_i = 0, \quad \forall i. \quad (A48)
\]
Substituting, into Equation (A48) the equilibrium conditions, \( f^*_i - C^i(e^*_i, s^*_i; H) = 0 \), \( A_i(e^*_i; H) - c_{2i}(e^*_i; H) = 0 \), and \( \frac{S_i}{S} = \left( c_{1i} \sum_{j=1}^{M} \left( c_{i,j}^{-1} \right) \right)^{-1} \), \( \forall i \), [Equations (A9), (A10), (A15)], and rearranging, we have
\[
\frac{S_i}{W} = \frac{a + A(e^*_i; H) - c_0 - c_2(e^*_i; H)}{\hat{b} + \left( c_{1i} \sum_{j=1}^{M} c_{i,j}^{-1} \right)^{-1}} \frac{X(e^*_i, H)}{W}. \quad (A49)
\]
In the case that investors have do not have sufficient wealth to allocate to AFMI, despite having positive fund expected net alphas, then \( S/W = 1 \). Now, substituting the equilibrium conditions, \( S/W = 1 \), \( f^*_i - C^i(e^*_i, s^*_i; H) = 0 \), \( A_i(e^*_i; H) - c_{2i}(e^*_i; H) = B_i(e^*_i; H) = 0 \), and \( \frac{S_i}{S} = \left( c_{1i} \sum_{j=1}^{M} c_{i,j}^{-1} \right)^{-1} \), \( \forall i \) into Equation (A26), we have
\[
E(\alpha_i \mid D)\big|_{\{e^*, r^*, \delta^*\}} = X(e^*_i, H) - \left[ \hat{b} + \left( \sum_{j=1}^{M} c_{i,j}^{-1} \right)^{-1} W \right]. \quad (A50)
\]
This proves Proposition RN1.

Q.E.D.